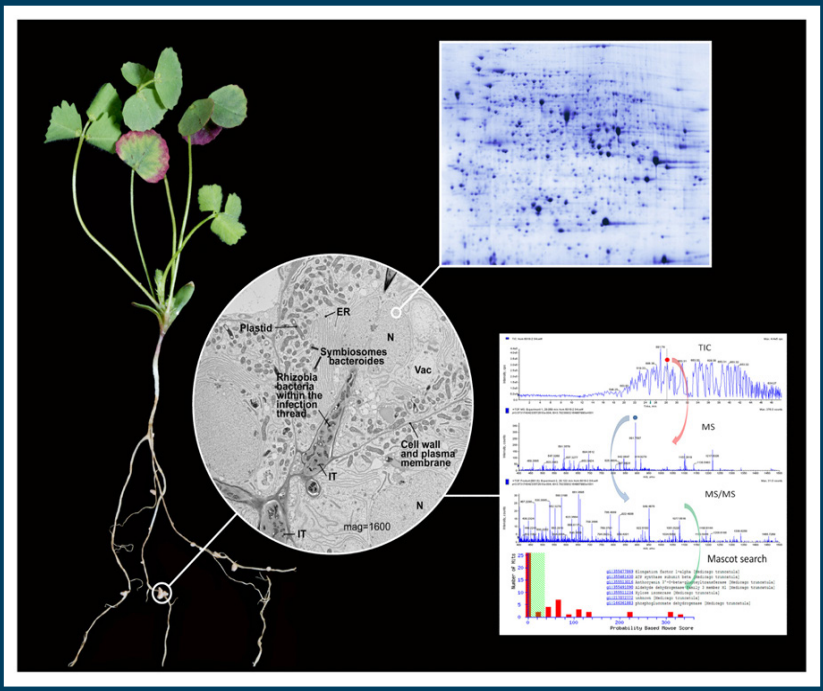
frontiers RESEARCH TOPICS



SUB-CELLULAR PROTEOMICS

Topic Editors
Nicolas L. Taylor and A. Harvey Millar

frontiers in PLANT SCIENCE

FRONTIERS COPYRIGHT STATEMENT

© Copyright 2007-2014
Frontiers Media SA.
All rights reserved.

All content included on this site, such as text, graphics, logos, button icons, images, video/audio clips, downloads, data compilations and software, is the property of or is licensed to Frontiers Media SA ("Frontiers") or its licensees and/or subcontractors. The copyright in the text of individual articles is the property of their respective authors, subject to a license granted to Frontiers.

The compilation of articles constituting this e-book, wherever published, as well as the compilation of all other content on this site, is the exclusive property of Frontiers. For the conditions for downloading and copying of e-books from Frontiers' website, please see the Terms for Website Use. If purchasing Frontiers e-books from other websites or sources, the conditions of the website concerned apply.

Images and graphics not forming part of user-contributed materials may not be downloaded or copied without permission.

Individual articles may be downloaded and reproduced in accordance with the principles of the CC-BY licence subject to any copyright or other notices. They may not be re-sold as an e-book.

As author or other contributor you grant a CC-BY licence to others to reproduce your articles, including any graphics and third-party materials supplied by you, in accordance with the Conditions for Website Use and subject to any copyright notices which you include in connection with your articles and materials.

All copyright, and all rights therein, are protected by national and international copyright laws.

The above represents a summary only. For the full conditions see the Conditions for Authors and the Conditions for Website Use.

ISSN 1664-8714
ISBN 978-2-88919-302-8
DOI 10.3389/978-2-88919-302-8

ABOUT FRONTIERS

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

FRONTIERS JOURNAL SERIES

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing.
All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

DEDICATION TO QUALITY

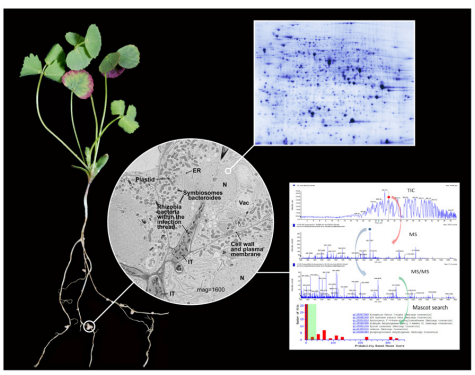
Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.
By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

WHAT ARE FRONTIERS RESEARCH TOPICS?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area!
Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: researchtopics@frontiersin.org

SUB-CELLULAR PROTEOMICS

Topic Editors:
Nicolas L. Taylor, The University of Western Australia, Australia
A. Harvey Millar, The University of Western Australia, Australia



Sub-cellular proteomics of Medicago truncatula
Image taken from: Azpeitia E, Weinstein N, Benítez M, Mendoza L and Alvarez-Buylla ER (2013) Finding missing interactions of the Arabidopsis thaliana root stem cell niche gene regulatory network. *Front. Plant Sci.* 4:110. doi: 10.3389/fpls.2013.00110

Whilst significant advances have been made in whole organismal proteomics approaches, many researchers still rely on combinations of tissue selection and subcellular prefractionation methods to reduce the complexity of protein extracts from plants prior to proteomic analysis. Often this will allow identification of many lower abundance proteins of the target proteome and it may involve the selection of specific organs, cell types or the isolation of specific subcellular components.

These subcellular proteomes provide insight into functions following various treatments and also contribute to the wider understanding of the entire organismal proteome by cataloguing a series of sub-proteome contents. The aim of this Research Topic is to bring together knowledge of sub cellular components in different plant species to provide a basis for accelerated research. It aims to provide a mini-review for each proposed section that summarizes the current understanding of a particular proteome, with the anticipation that every 5 - 10 years we can update these definitive publications.

Table of Contents

05 ***Subcellular Proteomics—Where Cell Biology Meets Protein Chemistry***
A. Harvey Millar and Nicolas L. Taylor

07 ***Plant Cell Wall Proteomics: The Leadership of Arabidopsis Thaliana***
Cécile Albenne, Hervé Canut and Elisabeth Jamet

24 ***Cell Wall Proteomics of Crops***
Setsuko Komatsu and Yuki Yanagawa

32 ***Progress Toward the Tomato Fruit Cell Wall Proteome***
Eliel Ruiz-May and Jocelyn K. C. Rose

39 ***The Arabidopsis Cytosolic Proteome: The Metabolic Heart of the Cell***
Jun Ito, Harriet T. Parsons and Joshua L. Heazlewood

47 ***The Arabidopsis Cytosolic Ribosomal Proteome: From Form to Function***
Adam J. Carroll

61 ***Proteomic Dissection of the Arabidopsis Golgi and Trans-Golgi Network***
Harriet T. Parsons, Georgia Drakakaki and Joshua L. Heazlewood

68 ***Recent Advances in the Composition and Heterogeneity of the Arabidopsis Mitochondrial Proteome***
Chun Pong Lee, Nicolas L. Taylor and A. Harvey Millar

76 ***3D Gel Map of Arabidopsis Complex I***
Katrin Peters, Katharina Belt and Hans-Peter Braun

85 ***The Rice Mitochondria Proteome and its Response During Development and to the Environment***
Shaobai Huang, Rachel N. Shingaki-Wells, Nicolas L. Taylor and A. Harvey Millar

91 ***The Mitochondrial Complexome of Medicago Truncatula***
Leonard Muriithi Kiirika, Christof Behrens, Hans-Peter Braun and Frank Colditz

98 ***Biochemistry, Proteomics, and Phosphoproteomics of Plant Mitochondria From Non-Photosynthetic Cells***
Jesper F. Havelund, Jay J. Thelen and Ian M. Møller

108 ***Comparative Analyses of Nuclear Proteome: Extending Its Function***
Kanika Narula, Asis Datta, Niranjan Chakraborty and Subhra Chakraborty

122 ***Recent Advances in Maize Nuclear Proteomic Studies Reveal Histone Modifications***
Paula Casati

127 ***The Seed Nuclear Proteome***
Ombretta Repetto, Hélène Rogniaux, Colette Larré, Richard Thompson and Karine Gallardo

134 ***Arabidopsis Peroxisome Proteomics***
John D. Bussell, Christof Behrens, Wiebke Ecke and Holger Eubel

147 ***Advancements in the Analysis of the Arabidopsis Plasma Membrane Proteome***
Koste A. Yadeta, J. Mitch Elmore and Gitta Coaker

157 ***Developmental Distribution of the Plasma Membrane-Enriched Proteome in the Maize Primary Root***
Zhe Zhang, Priyamvada Voothuluru, Mineo Yamaguchi, Robert E. Sharp and Scott C. Peck

169 ***Detergent-Resistant Plasma Membrane Proteome to Elucidate Microdomain Functions in Plant Cells***
Daisuke Takahashi, Yukio Kawamura and Matsuo Uemura

176 ***Defining the Core Proteome of the Chloroplast Envelope Membranes***
Stefan Simm, Dimitrios G. Papasotiriou, Mohamed Ibrahim, Matthias S. Leisegang, Bernd Müller, Tobias Schorge, Michael Karas, Oliver Mirus, Maik S. Sommer and Enrico Schleiff

194 ***The Hydrogen Peroxide-Sensitive Proteome of the Chloroplast in Vitro and in Vivo***
Meenakumari Muthuramalingam, Andrea Matros, Renate Scheibe, Hans-Peter Mock and Karl-Josef Dietz

208 ***Functional Proteomics of Barley and Barley Chloroplasts – Strategies, Methods and Perspectives***
Jørgen Petersen, Adelina Rogowska-Wrzesinska and Ole N. Jensen

218 ***Dissecting Plasmodesmata Molecular Composition by Mass Spectrometry-Based Proteomics***
Magali S. Salmon and Emmanuelle M. F. Bayer

226 ***When Proteomics Reveals Unsuspected Roles: The Plastoglobule Example***
Houda Nacir and Claire Bréhélin

234 ***Fluorescent Protein Tagging as a Tool to Define the Subcellular Distribution of Proteins in Plants***
Sandra K. Tanz, Ian Castleden, Ian D. Small and A. Harvey Millar

243 ***Sub-Cellular Proteomics of Medicago Truncatula***
Jeonghoon Lee, Zhentian Lei, Bonnie S. Watson and Lloyd W. Sumner

249 ***Plant Secretome Proteomics***
Erik Alexandersson, Ashfaq Ali, Svante Resjö and Erik Andreasson

frontiers in
PLANT SCIENCE

EDITORIAL
published: 26 February 2014
doi: 10.3389/fpls.2014.00055

Subcellular proteomics—where cell biology meets protein chemistry

A. Harvey Millar and Nicolas L. Taylor*

ARC Centre of Excellence in Plant Energy Biology, Centre for Comparative Analysis of Biomolecular Networks, The University of Western Australia, Perth, WA, Australia
**Correspondence: nicolas.taylor@uwa.edu.au*

Edited by:
Joshua L. Heazlewood, Lawrence Berkeley National Laboratory, USA

Keywords: sub-cellular proteomics, mass spectrometry, organelles, model plants, crop plants

The development of compartments in eukaryotic cells and the distribution of nuclear-encoded proteins underlies the expansion of plant genomes, the proliferation of multigene families and the specialization of cellular functions. The exploration of the proteome of the cell in terms of the collection of its subcompartments is therefore both a practical approach and also a function led necessity that recognizes that proper interpretation of proteomic data requires information about compartmentation of protein machinery.

Subcellular proteomics decreases the complexity of proteome discovery. With the typical compartment representing 500–4000 proteins, its analysis by gel based and MS based systems approach the resolution of the analytical techniques. In contrast, whole cell proteomes of 12,000–40,000 proteins extend well beyond the ability of proteomic tools to resolve them, leaving whole cell proteome studies being "tip of the iceberg" activities. Current shotgun studies can identify ∼500–3000 proteins with 2–20 h of MS time, making organelle proteomes and their quantitative comparisons within the reach of many research laboratories that either perform their own MS or use MS services.

Subcellular proteomics stands on the shoulders of decades of biochemical research that has developed methods for isolation of subcellular compartments. Extensive laboratory work involving the tinkering with density, size, and charge separation techniques has enabled incremental limitation of contamination in isolation methods from a range of subcellular structures. However, in depth MS studies over the last decade have also revealed that typical 90–95% enrichment still leaves much room for contaminants in preparations (Eubel et al., 2007; Huang et al., 2009; Ito et al., 2014). Studies from relatively abundant, or easily isolated homogenous compartments dominate the literature. In this class of structures are plastids, mitochondria, peroxisomes, and nuclei. Currently over 8000 proteins have been experimentally identified in these organelles in Arabidopsis (Tanz et al., 2013). Many fewer studies have attempted to untangle the intracellular membrane systems of ER, golgi, and PM. Separate techniques for these are complex, lack high levels of enrichment, and the protein populations of these structures are often transitory and differ between tissue types. Currently over 6000 proteins have been experimentally identified in these membranes in Arabidopsis (Tanz et al., 2013). All these structures bathe in the cytosol of the cell that itself contains a large and complex proteome. Isolation of pure cytosol without breaking organelles is extremely challenging and thus cytosolic proteomes are best defined through subtractive analysis of soluble proteomes against enriched organelle datasets. Quantitative comparisons of fractions collected during the subcompartment enrichment process, or across gradient separation of organelles, are key tools to differentiate the low level protein component from the small fraction of a contaminating protein from another location in the cell. Bringing together subcellular proteomics studies in aggregation databases has been very revealing to confirm location of proteins for which there are multiple conflicting claims in the literature (Tan et al., 2012; Tanz et al., 2013).

Analysis of multiple subcellular proteomes from the same tissues has begun to show the way in which multigene families have dispersed particular protein classes across subcellular boundaries to maintain translational, metabolic, signaling, and degradative machinery through the cell. Subcellular proteomes and targeted metabolic engineering are also showing how steps in metabolic pathways have been, and can be, redistributed in plants (compared to animals) to enable unique chemistries and accumulation of end-products in plants.

This special research topic aimed to bring together knowledge across sub cellular components and plant species to provide a basis for accelerated research in plant subcellular protein research. We have brought together a wide array of 26 publications including original research articles, reviews, and mini-reviews. They are focused on the model plants Arabidopsis (Parsons et al., 2012; Albenne et al., 2013; Bussell et al., 2013; Carroll, 2013; Lee et al., 2013a; Peters et al., 2013; Simm et al., 2013; Yadeta et al., 2013; Ito et al., 2014), rice (Huang et al., 2013; Komatsu and Yanagawa, 2013) and medicago (Kiirika et al., 2013; Lee et al., 2013b; Simm et al., 2013) as well as crop plants wheat, barley, maize, and tomato (Casati, 2012; Komatsu and Yanagawa, 2013; Petersen et al., 2013; Ruiz-May and Rose, 2013; Zhang et al., 2013). They include studies of the easily isolated subcellular proteomes of the chloroplast, mitochondria, peroxisome, and nuclei (Casati, 2012; Repetto et al., 2012; Bussell et al., 2013; Havelund et al., 2013; Huang et al., 2013; Lee et al., 2013a; Narula et al., 2013; Peters et al., 2013; Petersen et al., 2013; Simm et al., 2013), as well as less easily isolated golgi, plasma membrane, cytosolic ribosome, and cell wall proteomes(Parsons et al., 2012; Carroll, 2013; Takahashi et al., 2013; Yadeta et al., 2013; Zhang et al., 2013). Articles have also begun to investigate sub-organellar proteomes including the subcompartments of chloroplast (Simm et al.,

2013) and mitochondria (Peters et al., 2013), plasma membrane microdomains (Takahashi et al., 2013), and cell wall plasmodesmata (Salmon and Bayer, 2012). In addition to cataloguing these proteomes, researchers are beginning to investigate the post-translational modifications present on proteins in these locations (Havelund et al., 2013).

ACKNOWLEDGMENTS

This work is supported by the Australian Research Council Centre of Excellence in Plant Energy Biology. A. Harvey Millar and Nicolas L. Taylor are supported by the Australian Research Council as Future Fellows.

REFERENCES

Albenne, C., Canut, H., and Jamet, E. (2013). Plant cell wall proteomics: the leadership of *Arabidopsis thaliana*. *Front. Plant Sci.* 4:111. doi: 10.3389/fpls.2013.00111

Bussell, J. D., Behrens, C., Ecke, W., and Eubel, H. (2013). Arabidopsis peroxisome proteomics. *Front. Plant Sci.* 4:101. doi: 10.3389/fpls.2013.00101

Carroll, A. J. (2013). The Arabidopsis cytosolic ribosomal proteome: from form to function. *Front. Plant Sci.* 4:32. doi: 10.3389/fpls.2013.00032

Casati, P. (2012). Recent advances in maize nuclear proteomic studies reveal histone modifications. *Front. Plant Sci.* 3:278. doi: 10.3389/fpls.2012.00278

Eubel, H., Lee, C. P., Kuo, J., Meyer, E. H., Taylor, N. L., and Millar, A. H. (2007). Free-flow electrophoresis for purification of plant mitochondria by surface charge. *Plant J.* 52, 583–594. doi: 10.1111/j.1365-313X.2007.03253.x

Havelund, J. F., Thelen, J. J., and Moller, I. M. (2013). Biochemistry, proteomics, and phosphoproteomics of plant mitochondria from non-photosynthetic cells. *Front. Plant Sci.* 4:51. doi: 10.3389/fpls.2013.00051

Huang, S., Shingaki-Wells, R. N., Taylor, N. L., and Millar, A. H. (2013). The rice mitochondria proteome and its response during development and to the environment. *Front. Plant Sci.* 4:16. doi: 10.3389/fpls.2013.00016

Huang, S., Taylor, N. L., Narsai, R., Eubel, H., Whelan, J., and Millar, A. H. (2009). Experimental analysis of the rice mitochondrial proteome, its biogenesis, and heterogeneity. *Plant Physiol.* 149, 719–734. doi: 10.1104/pp.108.131300

Ito, J., Parsons, H. T., and Heazlewood, J. L. (2014). The Arabidopsis cytosolic proteome: the metabolic heart of the cell. *Front. Plant Sci.* 5:21. doi: 10.3389/fpls.2014.00021

Kiirika, L. M., Behrens, C., Braun, H. P., and Colditz, F. (2013). The Mitochondrial complexome of *Medicago truncatula*. *Front. Plant Sci.* 4:84. doi: 10.3389/fpls.2013.00084

Komatsu, S., and Yanagawa, Y. (2013). Cell wall proteomics of crops. *Front. Plant Sci.* 4:17. doi: 10.3389/fpls.2013.00017

Lee, C. P., Taylor, N. L., and Millar, A. H. (2013a). Recent advances in the composition and heterogeneity of the Arabidopsis mitochondrial proteome. *Front. Plant Sci.* 4:4. doi: 10.3389/fpls.2013.00004

Lee, J., Lei, Z., Watson, B. S., and Sumner, L. W. (2013b). Sub-cellular proteomics of *Medicago truncatula*. *Front. Plant Sci.* 4:112. doi: 10.3389/fpls.2013.00112

Narula, K., Datta, A., Chakraborty, N., and Chakraborty, S. (2013). Comparative analyses of nuclear proteome: extending its function. *Front. Plant Sci.* 4:100. doi: 10.3389/fpls.2013.00100

Parsons, H. T., Christiansen, K., Knierim, B., Carroll, A., Ito, J., Batth, T. S., et al. (2012). Isolation and proteomic characterization of the Arabidopsis Golgi defines functional and novel components involved in plant cell wall biosynthesis. *Plant Physiol.* 159, 12–26. doi: 10.1104/pp.111.193151

Peters, K., Belt, K., and Braun, H. P. (2013). 3D Gel Map of Arabidopsis Complex I. *Front. Plant Sci.* 4l:153. doi: 10.3389/fpls.2013.00153

Petersen, J., Rogowska-Wrzesinska, A., and Jensen, O. N. (2013). Functional proteomics of barley and barley chloroplasts - strategies, methods and perspectives. *Front. Plant Sci.* 4:52. doi: 10.3389/fpls.2013.00052

Repetto, O., Rogniaux, H., Larre, C., Thompson, R., and Gallardo, K. (2012). The seed nuclear proteome. *Front. Plant Sci.* 3:289. doi: 10.3389/fpls.2012.00289

Ruiz-May, E., and Rose, J. K. (2013). Progress toward the tomato fruit cell wall proteome. *Front. Plant Sci.* 4:159. doi: 10.3389/fpls.2013.00159

Salmon, M. S., and Bayer, E. M. (2012). Dissecting plasmodesmata molecular composition by mass spectrometry-based proteomics. *Front. Plant Sci.* 3:307. doi: 10.3389/fpls.2012.00307

Simm, S., Papasotiriou, D. G., Ibrahim, M., Leisegang, M. S., Muller, B., Schorge, T., et al. (2013). Defining the core proteome of the chloroplast envelope membranes. *Front. Plant Sci.* 4:11. doi: 10.3389/fpls.2013.00011

Takahashi, D., Kawamura, Y., and Uemura, M. (2013). Detergent-resistant plasma membrane proteome to elucidate microdomain functions in plant cells. *Front. Plant Sci.* 4:27. doi: 10.3389/fpls.2013.00027

Tan, Y. F., Millar, A. H., and Taylor, N. L. (2012). Components of mitochondrial oxidative phosphorylation vary in abundance following exposure to cold and chemical stresses. *J. Proteome Res.* 11, 3860–3879. doi: 10.1021/pr3003535

Tanz, S. K., Castleden, I., Hooper, C. M., Vacher, M., Small, I., and Millar, H. A. (2013). SUBA3: a database for integrating experimentation and prediction to define the SUBcellular location of proteins in Arabidopsis. *Nucleic Acids Res.* 41, D1185–D1191. doi: 10.1093/nar/gks1151

Yadeta, K. A., Elmore, J. M., and Coaker, G. (2013). Advancements in the analysis of the Arabidopsis plasma membrane proteome. *Front. Plant Sci.* 4:86. doi: 10.3389/fpls.2013.00086

Zhang, Z., Voothuluru, P., Yamaguchi, M., Sharp, R. E., and Peck, S. C. (2013). Developmental distribution of the plasma membrane-enriched proteome in the maize primary root growth zone. *Front. Plant Sci.* 4:33. doi: 10.3389/fpls.2013.00033

Received: 04 February 2014; accepted: 04 February 2014; published online: 26 February 2014.

Citation: Millar AH and Taylor NL (2014) Subcellular proteomics—where cell biology meets protein chemistry. Front. Plant Sci. ***5****:55. doi: 10.3389/fpls.2014.00055*

This article was submitted to Plant Proteomics, a section of the journal Frontiers in Plant Science.

Copyright © 2014 Millar and Taylor. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in
PLANT SCIENCE

REVIEW ARTICLE
published: 01 May 2013
doi: 10.3389/fpls.2013.00111

Plant cell wall proteomics: the leadership of *Arabidopsis thaliana*

Cécile Albenne*[1,2], *Hervé Canut*[1,2] and *Elisabeth Jamet*[1,2]

[1] *Laboratoire de Recherche en Sciences Végétales, Université de Toulouse, UPS, UMR 5546, Castanet-Tolosan, France*
[2] *CNRS, UMR 5546, Castanet-Tolosan, France*

Edited by:
Nicolas L. Taylor, The University of Western Australia, Australia

Reviewed by:
Zhiyong Wang, Carnegie Institution, USA
Subhra Chakraborty, National Institute of Plant Genome Research, India

****Correspondence:***
Elisabeth Jamet, Laboratoire de Recherche en Sciences Végétales, UMR 5546 UPS/CNRS, 24 chemin de Borde Rouge – Auzeville, BP 42617, F-31326 Castanet-Tolosan, France.
e-mail: jamet@lrsv.ups-tlse.fr

Plant cell wall proteins (CWPs) progressively emerged as crucial components of cell walls although present in minor amounts. Cell wall polysaccharides such as pectins, hemicelluloses, and cellulose represent more than 90% of primary cell wall mass, whereas hemicelluloses, cellulose, and lignins are the main components of lignified secondary walls. All these polymers provide mechanical properties to cell walls, participate in cell shape and prevent water loss in aerial organs. However, cell walls need to be modified and customized during plant development and in response to environmental cues, thus contributing to plant adaptation. CWPs play essential roles in all these physiological processes and particularly in the dynamics of cell walls, which requires organization and rearrangements of polysaccharides as well as cell-to-cell communication. In the last 10 years, plant cell wall proteomics has greatly contributed to a wider knowledge of CWPs. This update will deal with (i) a survey of plant cell wall proteomics studies with a focus on *Arabidopsis thaliana*; (ii) the main protein families identified and the still missing peptides; (iii) the persistent issue of the non-canonical CWPs; (iv) the present challenges to overcome technological bottlenecks; and (v) the perspectives beyond cell wall proteomics to understand CWP functions.

Keywords: *Arabidopsis thaliana*, cell wall, mass spectrometry, peptidomics, proteomics

INTRODUCTION

Plant primary cell walls are mainly composed of polysaccharide networks such as cellulose microfibrills, hemicelluloses wrapping and interlacing cellulose microfibrills and pectins (Carpita and Gibeaut, 1993). After the end of cell growth, secondary walls which contain additional compounds such as lignins, wax or cutin, are synthesized. Cell wall proteins (CWPs) play critical roles in plant cell walls during development and adaptation to environmental cues (Fry, 2004; Passardi et al., 2004). For this reason, extensive studies leading to their identification and characterization have been undertaken. Cell wall proteomics started about 10 years ago when the first plant genome sequences became available. Nowadays, there are about 40 papers covering this field (**Figure 1**), half of them concerning *Arabidopsis thaliana* whose genome was available in 2000 (Arabidopsis Genome Initiative, 2000). The availability of new genome sequences such as those of *Oryza sativa* (International Rice Genome Sequencing Project., 2005), *Populus trichocarpa* (Tuskan et al., 2006) and *Solanum lycopersicum* (Tomato Genome Consortium, 2012) enlarged the range of plant proteomics studies.

For plant cell wall proteomics studies, organs or cell suspension cultures have been used as starting materials containing cells surrounded by primary and/or secondary walls. Various experimental approaches were undertaken to characterize cell wall proteomes. Five specific features of CWPs need to be emphasized to understand them. (i) CWPs represent only 5–10% of the cell wall mass (Cassab and Varner, 1988). They are embedded in a complex matrix of carbohydrate polymers, aromatic compounds, wax or cutin depending on the type of cell walls. (ii) CWPs may interact with cell wall components by non-covalent linkages (Carpin et al., 2001; Spadoni et al., 2006). They can also be covalently linked, thus forming insoluble networks, like structural proteins networks of Proline-Rich Proteins (PRPs) or extensins (Brisson et al., 1994; Brady et al., 1996). (iii) Contrary to other sub-cellular compartments, plant cell walls constitute an open space connecting the cells in a tissue. It is located between the cell plasma membrane and the cuticle in aerial organs or the suberin layer in roots conferring to the plant surface waterproof qualities and protection against biotic and abiotic stresses (Thomas et al., 2007; Javelle et al., 2010). (iv) Most CWPs are basic proteins (Jamet et al., 2008). (v) Most CWPs undergo post-translational modifications (PTMs), like hydroxylation of proline (Pro) residues converting them to hydroxyproline (Hyp), *N*-glycosylation, *O*-glycosylation or addition of a glycosylphosphatidylinositol (GPI)-anchor (Kieliszewski and Lamport, 1994; Spiro, 2002; Faye et al., 2005).

For each step of the cell wall proteomics flowchart, the specificities of CWPs must be taken into account: plant fractionation, protein extraction, protein separation, protein identification by mass spectrometry (MS), and bioinformatics. Indeed, CWPs can be tightly trapped into the extracellular matrix and escape the extraction procedure. They may not be resolved at the step of separation by two-dimensional electrophoresis (2D-E) because they are mainly basic glycoproteins (Jamet et al., 2008). Finally, the databases used for protein identification using MS data

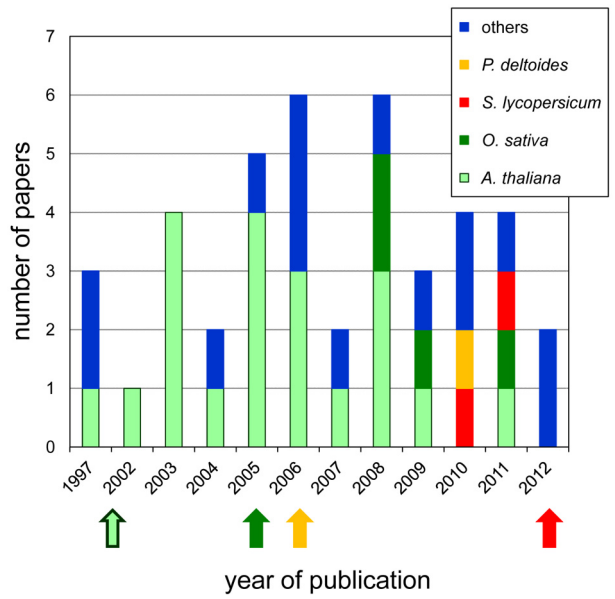


FIGURE 1 | Occurrence of cell wall proteomics papers since 1997. The number of papers published each year is represented with colored bars for each plant: *A. thaliana* (green), *O. sativa* (pink), *P. deltoides* (yellow), *S. lycopersicum* (red), and others (blue). The arrows show the year of the genomic sequence release for each of them. Note that the genomic sequence of *P. trichocarpa* was used for protein identification in *P. deltoides*.

contain no information about PTMs such as glycosylation, thus preventing the identification of some of them.

In this review, we will give a survey of plant cell wall proteomics studies with a focus on *A. thaliana* because this plant provides the best documented cell wall proteomes. The main protein families identified and the persistent issue of the non-canonical CWPs will be discussed. Finally, we will provide perspectives in the field of plant cell wall proteomics, going beyond the present data with systems biology approaches and peptidomics to decipher the roles of proteins and peptides in cell walls.

MATERIALS AND METHODS

All the *A. thaliana* proteins reported in this review have been analyzed with different bioinformatics software to predict their sub-cellular localization and their functional domains using *ProtAnnDB* (http://www.polebio.lrsv.ups-tlse.fr/ProtAnnDB/index.php) as previously described (San Clemente et al., 2009). Briefly, the following programs have been used for prediction of sub-cellular localization: TargetP (http://www.cbs.dtu.dk/services/TargetP/), SignalP (http://www.cbs.dtu.dk/services/SignalP/), Predotar (http://urgi.versailles.inra.fr/predotar/predotar.html), Aramemnon (http://aramemnon.botanik.uni-koeln.de/) and TMHMM (http://www.cbs.dtu.dk/services/TMHMM-2.0/). The programs used for prediction of functional domains were PROSITE (http://prosite.expasy.org/), Pfam (http://pfam.sanger.ac.uk/), and InterPro (http://www.ebi.ac.uk/interpro/).

A SURVEY OF PLANT CELL WALL PROTEOMICS

EXTRACELLULAR PROTEOMES

In this review, we will focus on different type of extracellular proteomes, commonly named as call wall proteomes. For example, secretome, in which all of the secreted proteins of a cell suspension culture, roots or seedling are collected in liquid culture media. Another type of extracellular proteome encompasses apoplastic proteomes in which proteins from the cell wall can be eluted by vacuum infiltration with various solutions. Extraction of proteins from purified cell walls with various solutions is the third category of cell wall proteome that have been used to elute loosely bound CWPs. In addition, sub-proteomes such as *N*-glycoproteomes and a GPI-anchored proteome have been analyzed. All the extracellular proteomes have been obtained with different plants like *A. thaliana* (see **Table 1**), *Cicer arietinum* (Bhushan et al., 2006, 2011), *Glycine max* (Komatsu et al., 2010), *Helianthus annuus* (Pinedo et al., 2012), *O. sativa* (Chen et al., 2008a; Jung et al., 2008; Cho et al., 2009; Zhou et al., 2011), *Medicago sativa* (Watson et al., 2004; Verdonk et al., 2012), *Nicotiana benthamiana* (Goulet et al., 2010), *Nicotiana tabacum* (Robertson et al., 1997; Dani et al., 2005; Morel et al., 2006; Delannoy et al., 2008; Millar et al., 2009), *Populus deltoides* (Pechanova et al., 2010), *S. lycopersicum* (Robertson et al., 1997; Yeats et al., 2010; Catalá et al., 2011), *Solanum tuberosum* (Fernández et al., 2012; Lim et al., 2012) and *Zea mays* (Zhu et al., 2006, 2007). Besides, several xylem sap proteomes have been analyzed and were found to be very close to cell wall proteomes (Kehr et al., 2005; Alvarez et al., 2006; Dafoe and Constabel, 2009; Ligat et al., 2011). With 20 published papers (**Table 1**) and 500 proteins with predicted signal peptide identified, the most studied plant is *A. thaliana*. Its genome was the first one to be sequenced, thus allowing a precise identification of the proteins. Altogether, between one fourth and one third of the expected cell wall proteome of *A. thaliana* has been identified (Jamet et al., 2006). The second most studied plant is *O. sativa* with 270 proteins with predicted signal peptide identified. When no genome information is available, protein identification relies on the availability of expressed sequenced tags (ESTs) or cDNAs (Lim et al., 2012). Alternatively, proteins are identified by sequence homology. In this case, it is not possible to obtain precise identification of proteins and to distinguish between members of multigene families like in *C. arietinum* or *H. annuus* (Bhushan et al., 2006; Pinedo et al., 2012).

STRATEGIES IN PLANT CELL WALL PROTEOMICS

Many different strategies have been used to identify extracellular proteins of plants. A synopsis of the different experimental procedures is presented in **Figure 2** and **Table A1** in five general steps. Steps 1 and 2 lead to protein extraction. Step 3 consists in protein separation. Steps 4 and 5 lead to protein identification. The first step distinguishes: (i) studies of secretomes in which only proteins spontaneously released in culture media are analyzed (Charmont et al., 2005; Oh et al., 2005; Basu et al., 2006; Tran and Plaxton, 2008; Cheng et al., 2009); (ii) the release of proteins by non-destructive methods in which the integrity of the cell plasma membranes is preserved either by vacuum infiltration of tissues (Haslam et al., 2003; Boudart et al., 2005; Casasoli et al., 2008) or by washing of cells cultured in liquid medium (Robertson et al.,

Table 1 | Cell wall proteomes of *A. thaliana*.

Organ	References	Type of proteome	Total number of identified proteins (% without predicted signal peptide)[c]	Number of proteins with predicted signal peptide[c]
Culture medium of roots	Basu et al., 2006	Secretome	52 (31)	36
Stems	Minic et al., 2007	*N*-glycoproteome	102 (12)	90
Leaves	Haslam et al., 2003	Apoplastic proteome	10 (10)	9
Leaves	Boudart et al., 2005	Apoplastic proteome	93 (7)	87
Etiolated hypocotyls (5 days)	Feiz et al., 2006	Cell wall proteome	96 (13)	84
Etiolated hypocotyls (5 and 11 days)	Irshad et al., 2008	Cell wall proteome	173 (21)	137
Etiolated hypocotyls (11 days)	Zhang et al., 2011	Cell wall glycoproteome	129 (2)	127
Seedlings (10 days) (oligogalacturonides)	Casasoli et al., 2008[a]	Apoplastic proteome	20 (30)	14
Cell suspension cultures (3 days)	Bayer et al., 2006	Cell wall proteome	792 (87)	106
Cell suspension cultures (5 days)	Kwon et al., 2005	Cell wall proteome[b]	39 (13)	34
Cell suspension cultures (5 days)	Chivasa et al., 2002	Cell wall proteome	72 (64)	26
Cell suspension cultures (5 days)	Robertson et al., 1997	Cell wall proteome[b]	33 (46)	18
Cell suspension cultures (7 days)	Borderies et al., 2003	Cell wall proteome[b]	95 (48)	50
Culture medium of calli (21 days)	Borner et al., 2003	GPI-anchored proteome	30 (0)	30
Culture medium of cell suspension cultures (4 days)	Oh et al., 2005	Secretome	45 (47)	24
Cell suspension cultures (7 days) (fungal elicitors)	Ndimba et al., 2003[a]	Cell wall proteome	6 (0)	6
Culture medium of cell suspension cultures (7 days) (salicylic acid treatment)	Cheng et al., 2009	Secretome	74 (47)	39
Culture medium of cell suspension cultures (7 days) (phosphate deficiency)	Tran and Plaxton, 2008[a]	Secretome	37 (43)	21
Protoplasts (1 h)	Kwon et al., 2005	Cell wall proteome[b]	71 (65)	25
Protoplasts (3 h)	Kwon et al., 2005	Cell wall proteome[b]	66 (65)	23
Liquid cultured etiolated seedlings	Schultz et al., 2004	AGP proteome	12 (0)	12
Culture medium of etiolated seedlings	Charmont et al., 2005	Secretome	49 (10)	44

[a] *Protein identification has been only performed on protein spots showing variation between control and treated samples.*

[b] *These proteomes have been obtained by washing of cells or protoplasts with various salt solutions (see* ***Table A1****).*

[c] *All the bioinformatic predictions of sub-cellular localization have been done as described in Materials and Methods to allow the comparison between the A. thaliana cell wall proteomes.*

1997; Borderies et al., 2003; Kwon et al., 2005); and (iii) the release of proteins by destructive methods starting with a grinding of the tissues, thus mixing intracellular and extracellular compartments. In this case, either cell walls were purified prior to protein extraction (Chivasa et al., 2002; Ndimba et al., 2003; Feiz et al., 2006; Minic et al., 2007; Irshad et al., 2008; Zhang et al., 2011) or the tissues were ground prior to isolation of *N*-glycosylated proteins by lectin affinity chromatography (Minic et al., 2007). In the case of the GPI-anchored proteome, the first step consisted in the preparation of a membrane fraction followed by the cleavage of GPI-anchors by phosphatidylinositol-specific phospholipase C (Pi-PLC) (Borner et al., 2003).

The second step (**Figure 2**) is very diverse using different solutions to extract proteins. These solutions can be acidic or basic (Feiz et al., 2006; Casasoli et al., 2008). Their main components are: salts (NaCl, $CaCl_2$, $MgCl_2$, KCl, or LiCl) or osmotic agents (mannitol) (Borderies et al., 2003; Boudart et al., 2005; Kwon et al., 2005; Feiz et al., 2006); chelating agents (EDTA or CDTA) (Robertson et al., 1997; Boudart et al., 2005); detergents (SDS, Triton or CHAPS) (Chivasa et al., 2002; Borner et al., 2003); phenol (Bayer et al., 2006); and/or chaotropic agents (urea and thiourea) (Chivasa et al., 2002). The β-glucosyl Yariv reagent has been used to isolate arabinogalactan proteins (AGPs) (Schultz et al., 2004). In some cases, several salt solutions have been used successively (Chivasa et al., 2002; Borderies et al., 2003; Boudart et al., 2005; Feiz et al., 2006; Irshad et al., 2008; Zhang et al., 2011). As mentioned above, in the case of the GPI-anchored proteome, a Pi-PLC has been used (Borner et al., 2003). The methods used to extract CWPs have been previously described in detail (Feiz et al., 2006; Jamet et al., 2008).

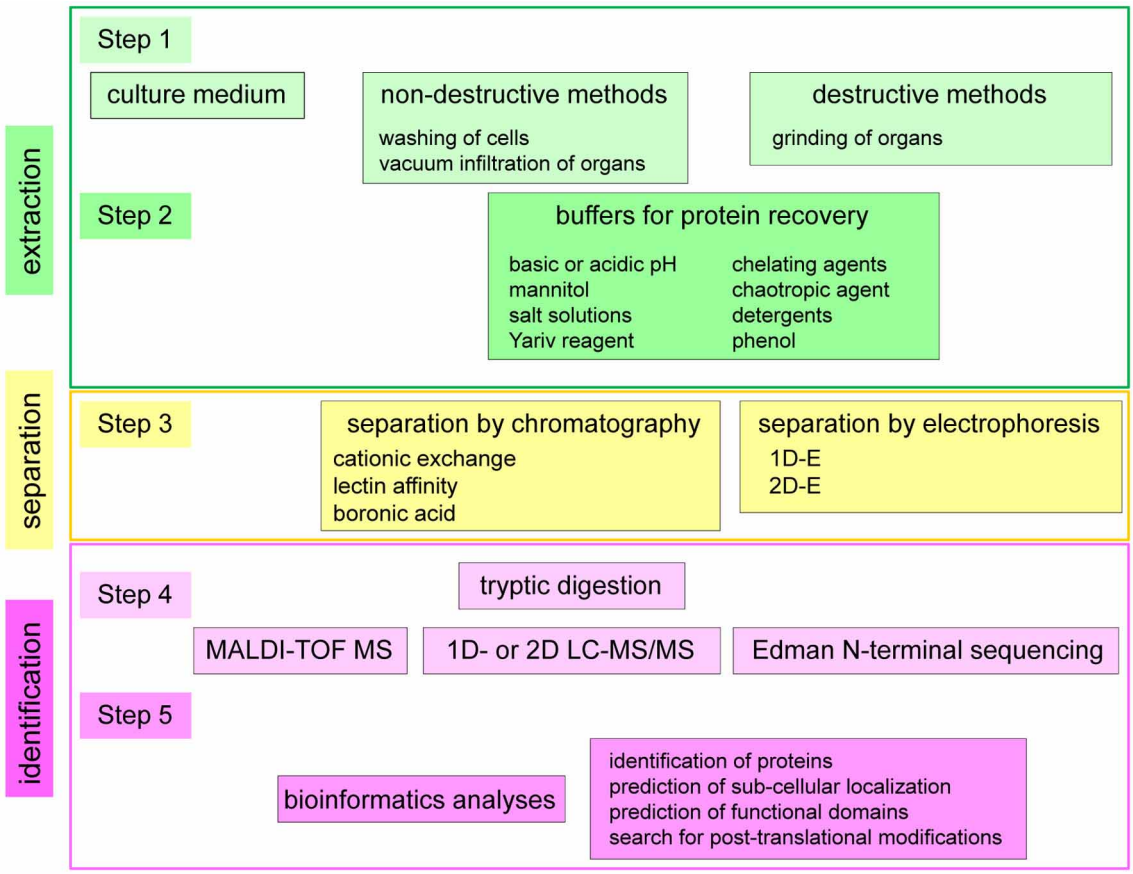


FIGURE 2 | Synopsis of the different strategies used for the study of cell wall proteomes and secretomes. Five main steps have been identified in the published strategies and different combinations of the five steps have been used. Steps 1 and 2 lead to protein extraction. Step 3 consists in protein separation. Steps 4 and 5 lead to protein identification by combining MS or Edman N-terminal sequencing and bioinformatics. In some cases, one step is performed twice or even several times, e.g., steps 2 and 3, by modifying the composition of the buffer used for protein recovery or doing two subsequent steps of protein separation. In other cases, one step can be skipped, like step 2 for the analysis of secreted proteins present in culture medium, or step 3 when proteins are directly analyzed by MS.

The third possible step is protein separation (**Figure 2**). It can be done by chromatography (cationic exchange, lectin affinity, boronic acid), and/or by 1D- or 2D-E. Cationic exchange chromatography has been performed under physico-chemical conditions similar to those found in cell walls, that is an acidic medium at a pH around 4.5 at which basic proteins are positively charged (Boudart et al., 2005; Irshad et al., 2008). Affinity chromatography on Concanavalin A (ConA) has been artfully used to isolate *N*-glycoproteins from a total extract of proteins (Minic et al., 2007). As expected, most of the identified proteins were predicted to be addressed to the secretion pathway where *N*-glycosylation occurs. Other lectins have been used to separate proteins extracted from cell walls: *Artocarpus integrifolia* Lectin (AIL) specific for α-Gal residues, PeaNut Agglutinin (PNA) specific for β-Gal residues, and wheat germ agglutinin (WGA) specific for GlcNAc residues (Zhang et al., 2011). With regard to separation of proteins by electrophoresis, 2D-E has shown limitations due to the fact that CWPs are mainly basic glycoproteins (Jamet et al., 2008). Considering the number of identified proteins, the most efficient cell wall proteomics analyses have been performed using two steps of protein separation (Boudart et al., 2005; Minic et al., 2007; Irshad et al., 2008; Zhang et al., 2011).

Two additional steps are necessary to achieve protein identification (steps 4 and 5, **Figure 2**). The fourth step consists in proteolytic digestion of proteins and MS analyses of peptides, using Matrix-Assisted Laser Desorption Ionization-Time Of Flight (MALDI-TOF) MS (Boudart et al., 2005; Kwon et al., 2005; Irshad et al., 2008), liquid chromatography (LC)-MS/MS (Minic et al., 2007; Casasoli et al., 2008; Zhang et al., 2011) or 2D-LC-MS/MS (Basu et al., 2006; Bayer et al., 2006; Cheng et al., 2009). Trypsin is the most widely used protease. In a few cases, Edman N-terminal sequencing has been performed (Robertson et al., 1997; Schultz et al., 2004). When proteins are heavily glycosylated, like *O*-glycoproteins, it is necessary to deglycosylate them with hydrogen fluoride (HF) to get access to their polypeptide skeleton (Schultz et al., 2004). The fifth step consists in bioinformatics analyses to identify proteins, predict their

sub-cellular localization, their functional domains and eventually get information about their PTMs (San Clemente et al., 2009).

Different complementary strategies are now available to study plant cell wall proteomes. It is possible to design the most relevant flowchart for a new cell wall proteomics study and to perform it in an efficient way. The main limitation remains the availability of genomic sequences for many plants of agronomic interest.

PLANT CELL WALL GLYCOPROTEOMES

During their secretion, proteins undergo glycosylation which is one of the most common and complex PTM known to control many physiological processes (Faye et al., 2005). Glycosylation is of two main types, namely *N*- and *O*-glycosylation, depending on the nature of the amino acid bearing them. Unlike yeast and mammalian glycoproteins which are extensively studied, plant glycoproteins are still poorly characterized. Hyp-rich glycoproteins (HRGPs) undoubtedly constitute the most documented plant cell wall *O*-glycoprotein superfamily (Kieliszewski, 2001; Tan et al., 2012; Velasquez et al., 2012). A few *N*-glycoproteins have been studied in detail, e.g., a peroxidase (Lige et al., 2001), an α-mannosidase (Kimura et al., 1999) or a polygalacturonase inhibiting protein (PGIP) (Lim et al., 2009) for which glycosylation has been shown to contribute to activity. Beyond the study on targeted glycoproteins, the concept of glycoproteomics is now emerging in plants. New analytical pipelines are available (Song et al., 2011; Ruiz-May et al., 2012). They aim at detection, enrichment and MS analysis of large sets of glycoproteins.

A few systematic surveys have been carried out so far on plants. Minic et al. were the first ones to use a ConA lectin chromatography step to capture *N*-glycoproteins from a protein extract of *A. thaliana* (Minic et al., 2007). A similar approach has been developed to characterize the *N*-glycoproteome of *S. lycopersicum* (Catalá et al., 2011). Finally, Zhang et al. enlarged the coverage of the *A. thaliana* cell wall glycoproteome using multi-dimensional lectin chromatography and boronic acid chromatography (Zhang et al., 2011). The obtained sub-proteomes mostly corresponded to *N*-glycoproteomes with only few *O*-glycoproteins detected. Plant glycoproteomics is only at its premise and is undoubtedly a very promising approach toward an integrated study of both sugar and proteins moieties to gain new insight into the structure and function of glycoproteins.

MAIN CWP FAMILIES

GROUPING OF CWPs: PRINCIPLES AND DRAWBACKS

A major challenge is to interpret cell wall proteomics data, in other words, to get a biological message from a list of proteins. In a first effort, it is tempting to group proteins to get an overview of the extracellular proteome, to highlight specific proteins or protein families in the physiological context of interest, or to identify house-keeping proteins. This is a difficult exercise because most of the identified proteins have no experimentally defined function. Another difficulty is that two series of proteins can be distinguished in all cell wall proteomes: those having a signal peptide predicted with at least two bioinformatics programs, and those having no predicted signal peptide or having a motif that addresses them to an intracellular compartment. Only those having a *bona fide* predicted signal peptide are named CWPs in this review. This point will be discussed below.

From the fifth step of the proteomics flowchart (**Figure 2**), bioinformatics analyses lead to group proteins in families either by sequence comparison to proteins present in databases and already annotated, or by search for functional domains as defined in domain repertoires like PROSITE, Pfam, or InterProScan. Two types of classification have then been proposed. Both of them have drawbacks and suffer from ambiguity. The first type is based on the physiological processes in which the proteins are assumed to be involved, like growth and development, stress or defense against pathogens. The drawback is that it can be difficult to sort the proteins. For example, glycoside hydrolases (GHs) could be involved in both plant development and defense (Kasprzewska, 2003). In the same way, proteases could be involved in protein turnover or in signaling by releasing biologically active peptides (Berger and Altmann, 2000; Hunt et al., 2010; Leasure and He, 2012). The second type of classification is based on predicted functional domains and possible partners or targets (i.e., polysaccharides, lipids or proteins) of CWPs in cell walls. The drawbacks are the followings: (i) all the proteins do not have a predicted biochemical activity; (ii) it is difficult to assign proteins with several functional domains to a class as exemplified below. Of course, all these classifications need to evolve to take into account new experimental results demonstrating protein functions.

AN EXAMPLE OF FUNCTIONAL CLASSIFICATION OF CWPs

In this review, we present a functional classification taking into account predicted functional domains as well as possible partners or targets in cell walls. *A. thaliana* CWPs are taken as an example. Nine functional classes listed in **Table 2** have been proposed (Jamet et al., 2008). The most populated functional class is that of proteins acting on carbohydrates. It represents about one fourth of the proteomes (25.6%) and it includes GHs, carbohydrate esterases (CEs), polysaccharide lyases (PLs) and expansins. The importance of such proteins is not surprising since polysaccharides constitute the largest fraction of cell walls and are constantly submitted to remodeling during plant development or in response to environmental cues (Fry, 2004; Cosgrove, 2005). The second most predominant class of CWPs is that of oxido-reductases (14.6%), like peroxidases, multicopper oxidases, blue copper binding proteins, and berberine bridge enzymes. Again, the importance of this class was expected because many oxidation reactions occur in the extracellular matrix to modify polymer networks involving carbohydrates, aromatic compounds, or structural proteins (Passardi et al., 2004). However, the biochemical functions of proteins homologous to berberine bridge enzymes and of blue copper binding proteins in cell walls are not known (Nersissian and Shipp, 2002). Then, numerous proteases have been found in cell wall proteomes (11.2%). Until recently, the roles of such enzymes have probably been under-estimated in cell walls. They could be involved in protein turnover, protein maturation or release of biologically active peptides (Van Der Hoorn, 2008). Nothing is known about CWP turnover. The maturation of enzymes having N- or C-terminal pro-peptides or N-terminal inhibitory domains has

Table 2 | A classification of *A. thaliana* CWPs in 9 functional classes according to their predicted functional domains and their possible partners in cell walls.

Functional class	Protein families	Possible partners in cell walls	% *A. thaliana* proteomes
Proteins acting on carbohydrates			25.7
	Glycoside hydrolases (GHs)	Carbohydrates	19.4
	Carbohydrate esterases (CEs)	Pectins	2.4
	Polysaccharide lyases (PLs)	Pectins	0.6
	Expansins	Cellulose/hemicellulose	2.4
Oxido-reductases			14.6
	Peroxidases	Carbohydrates/Structural proteins/Lignins	6.2
	Multicopper oxidases		2.0
	Blue copper binding proteins		2.2
	Berberine bridge enzymes		2.2
Proteases			11.2
	Asp proteases	Proteins	2.8
	Cys proteases	Proteins	1.6
	Ser proteases	Proteins	3.6
	Ser carboxypeptidases	Proteins	2.2
Proteins with interaction domains			11.0
	Lectins	Carbohydrates	2.4
	Leucine-rich-repeat domains (LRRs)	Proteins	2.6
	Enzyme inhibitors	Enzymatic proteins	4.0
Proteins possibly involved in signaling			6.6
	Arabinogalactan proteins (AGPs)		2.4
	Fasciclin-AGPs (FLAs)		1.2
	Receptors		2.4
Structural proteins			1.6
	Glycine-rich proteins (GRPs)		0.2
	Leucine-rich-repeat extensins (LRXs)	Proteins	0.8
	Proline-rich proteins (PRPs)	Proteins	0.6
Proteins related to lipid metabolism			5.8
	Lipase GDSL	Lipids	2.4
	Lipid transfer proteins (LTPs)	Lipids	1.8
Miscellaneous proteins			11.0
	Purple acid phosphatases (PAPs)		1.2
	Phosphate-induced (phi) proteins		1.0
	Germin		1.6
Unknown function			12.5
	Domains of unknown function (DUFs)		5.0
	No clue		4.4

The annotation of proteins is based on the presence of functional domains as defined in the PROSITE, Pfam, and InterPro bioinformatics programs. Only the major protein families are mentioned in each functional class.

been demonstrated only in a few cases, such as type I pectin methylesterases (PMEs) (Wolf et al., 2009) and some GHs (Lee et al., 2003; Minic et al., 2004; Albenne et al., 2009). It was also shown that the AtSBT1.7 Ser protease plays a role in mucilage release from *A. thaliana* seed coat (Rautengarten et al., 2008). It is assumed that this protease contributes to the degradation or the maturation of cell wall modifying enzymes. The class of CWPs possibly involved in signaling (6.6%) is a difficult one. It comprises proteins like AGPs and proteins with transmembrane domains which are predicted to be plasma membrane receptors having extracellular domains. The roles of AGPs are not completely understood. Besides their roles in signaling (Seifert and Blaukopf, 2010), AGPs could also contribute to cell wall mechanical properties (Seifert and Roberts, 2007). The identification of receptors mostly relies on peptides located in the extracellular domain, but they do not really belong to cell wall proteomes. Finally a protease like stomatal density and distribution 1 (SDD1) could also be included in the class of signaling proteins because it is assumed to generate a extracellular signal to control the stomatal pattern (Von Groll et al., 2002). The next class of CWPs is that of proteins predicted to be related to lipid metabolism which is unexpectedly populated (5.8%). Such proteins are assumed to be involved in cuticle formation. However, the cuticle does not represent a major part of the organs analyzed. Other roles might be possible for these proteins. For example, a lipid transfer protein (LTP) has been assumed to be involved in cell wall extension by interacting with the cellulose/xyloglucan network of tobacco cell walls (Nieuwland et al., 2005). The class of structural proteins (1.6%) only groups a few proteins like Glycine-Rich Proteins (GRPs), PRPs, and Leucine-Rich Repeat Extensins (LRXs). The problem of CWP classification appears again with LRXs which could also be involved in signaling or be classified among proteins having predicted interaction domains (Baumberger et al., 2001; Leiber et al., 2010). No extensin has been identified in the published cell wall proteomes probably because they are covalently cross-linked (Wilson and Fry, 1986). The strategies presently used for cell wall proteomics fail to efficiently isolate such proteins. The class of CWPs with interaction domains (11.0%) presently groups proteins having predicted carbohydrate binding domains, Leucine-Rich-Repeat (LRR) domains assumed to be involved in protein–protein interactions and enzyme inhibitors. This class can be split according to these three categories of CWPs (Catalá et al., 2011). A better knowledge of the function of proteins interacting with polysaccharides will also contribute to a more precise classification. The group of miscellaneous proteins (11.0%) is the Achilles'heel of the classification since it comprises all the proteins which cannot be put elsewhere. A few protein families emerge from this group like purple acid phosphatases (Wang et al., 2011), phosphate-induced proteins (Farrar et al., 2003) and germins (Membré et al., 2000). Finally, about one eighth of the cell wall proteomes correspond to proteins of yet unknown function with no predicted functional domain or a predicted domain of unknown function (DUF). This is a puzzling class of proteins which will probably reveal new functions in cell walls. It is expected to disappear when these proteins are characterized.

WallProtDB, A CELL WALL PROTEOMICS DATABASE

The classification of proteins described above has been used to build up the *WallProtDB* database (http://www.polebio.scsv.ups-tlse.fr/WallProtDB/) (Pont-Lezica et al., 2010; Ligat et al., 2011). The 20 published *A. thaliana* cell wall proteomes listed in **Table 1** (described in more details in **Table A1**) have been subjected to the same bioinformatics software pipeline (*ProtAnnDB*) in order to compare them more accurately. The number of CWPs identified in these proteomes is very variable, ranging from 6 to 137. The less populated proteomes are a leaf apoplast proteome (Haslam et al., 2003), the AGP proteome (Schultz et al., 2004) and those focused on proteins which level of accumulation changes in response to a treatment (Ndimba et al., 2003; Casasoli et al., 2008; Tran and Plaxton, 2008). On the contrary, the most populated proteomes are those relying on efficient CWP extraction (Boudart et al., 2005) or separation (Minic et al., 2007; Irshad et al., 2008; Zhang et al., 2011), or on the most sensitive MS techniques (Bayer et al., 2006). In addition to the *A. thaliana* cell wall proteomes (500 CWPs), the cell wall proteomes of *O. sativa* and a *B. oleracea* xylem sap proteome have been included in *WallProtDB*, thus representing about 1000 CWPs.

THE CASE OF NON-CANONICAL CWPs

Apart from the proteins having predicted signal peptides, all the cell wall proteomes contain proteins which are not predicted to be secreted and proteins predicted to be endoplasmic reticulum resident proteins. The proportion of these non-canonical CWPs varies from none in the case of the AGP and GPI-anchored proteomes (Borner et al., 2003; Schultz et al., 2004) to 87% (Bayer et al., 2006), the average being 30% (**Table 1**). The cell wall proteomes containing the lowest proportion of non-canonical CWPs have been obtained with the following strategies: secretome analysis (Charmont et al., 2005), extraction of apoplastic fluids with salt solutions (Boudart et al., 2005), affinity chromatography on lectins to isolate glycoproteins (Minic et al., 2007; Catalá et al., 2011; Zhang et al., 2011) and cell wall purification with an adapted protocol followed by extraction of proteins with salt solutions (Feiz et al., 2006). Apart from the limitations of bioinformatics programs to predict sub-cellular localization (Imai and Nakai, 2010), the difficulties mentioned above to preserve membrane integrity or to purify cell walls have to be taken into account to understand these contrasting results. Moreover, the facts that the percentage of non-canonical CWPs varies between experiments and that these proteins are not always the same, indicate that most of them are probably intracellular contaminants. However, it cannot be excluded that some of them are present in cell walls. This point has been recently reviewed (Rose and Lee, 2010).

Several authors used prediction of sub-cellular localization with SecretomeP which performs *ab initio* predictions of non-classical, i.e., not signal peptide-triggered, protein secretion for mammalian proteins (http://www.cbs.dtu.dk/services/SecretomeP/) (Bendtsen et al., 2004). However, this software is not well-adapted to plant proteins since it has been designed for mammalian proteins. Moreover, only a small proportion of the non-canonical proteins identified in cell wall proteomics studies

gave a score above threshold (Jamet et al., 2008; Pechanova et al., 2010; Bhushan et al., 2011; Fernández et al., 2012).

Unfortunately, experimental data are too scarce to confirm the localization of all the non-canonical CWPs. In animal cells, several alternative mechanisms of protein secretion have been proposed and partly demonstrated (Nickel and Seedorf, 2008). Unconventional secretory proteins seem to share several common features like (i) no leader sequence, (ii) absence of PTMs specific for ER or Golgi apparatus, and (iii) secretion not affected by brefeldin A which blocks the classical ER/Golgi-dependent secretion pathway. A jacalin-related lectin has been first identified in extracellular fluids of sunflower seedlings and then demonstrated to be extracellular by immunolocalization (Pinedo et al., 2012). This is the only case where the three criteria defined for animal unconventional secretory proteins described above were met, leading to the assumption of the release of exosomes to the extracellular matrix (Regente et al., 2012). In addition, a few moonlighting proteins were described like a rice α-amylase (GH13 family) which was shown to be present in both cell walls and plastids (Chen et al., 2004).

There is an urgent need for systematic localization of plant proteins by (i) biochemical strategies like isotope tagging successfully used for the membrane organelle proteome of *A. thaliana* (Dunkley et al., 2006), (ii) immunolocalization to get a reliable protein atlas as done in the Human Protein Atlas project (Lundberg and Uhlén, 2010), (iii) green fluorescent protein (GFP) tagging (Heazlewood et al., 2005) or (iv) a yeast secretion trap assay (Lee and Rose, 2012). An interesting tool is SUBA3 (SUBcellular localization database for Arabidopsis proteins) which collects bioinformatics and experimental data of sub-cellular localization (http://suba.plantenergy.uwa.edu.au/) (Heazlewood et al., 2007). The precise identification of secreted proteins devoid of predicted signal peptide will allow the demonstration of the existence of alternative secretion pathways in plant cells and the design of bioinformatics software able to predict non-classical secretion of plant proteins.

PRESENT CHALLENGES: OVERCOMING TECHNOLOGICAL BOTTLENECKS

Plant cell wall proteomics has been a very active research area during the last 10 years, and is rapidly expanding with the availability of new genome sequences. However, knowledge on plant CWPs will gain new insight thanks to new methodological and technological developments aiming at the identification of low-abundant proteins, the characterization of protein–protein complexes and the description of PTMs.

TOWARD COMPLETE CELL WALL PROTEOMES

Proteomics studies aim at providing a global description of proteins present in a biological extract. However, the complexity of protein samples renders difficult their exhaustive analysis since (i) a few highly-abundant proteins can mask low-abundant proteins, and (ii) the dynamic range of proteins can be very broad, i.e., up to 12 orders of magnitude (Corthals et al., 2000). To overcome these limitations, new separation techniques have been developed, namely depletion and equalization methods. These methods have first proven their efficiency for mammalian and microbial systems and are now emerging for plants. Plant depletion methods described so far mostly concern the depletion of storage proteins or ribulose-1,5-biphosphate carboxylase/oxygenase (RUBISCO). A fast and simple fractionation technique to precipitate legume seed storage proteins has been developed, allowing the detection of 541 low-abundant proteins of *G. max* seeds after a 2D-E separation (Krishnan et al., 2009). A similar approach has been carried out to precipitate RUBISCO from soybean leaf soluble protein extract, permitting the detection of 230 new protein spots (Krishnan and Natarajan, 2009). Another RUBISCO depletion method based on immunocapture has been successfully performed to detect low-abundant proteins differentially regulated during *A. thaliana* defense (Widjaja et al., 2009). Even if storage proteins are not CWPs, they can be found as contaminants in specific cell wall proteomes, like seed cell wall proteomes (Merah et al., unpublished data). These depletion methods should then be useful to remove such major contaminants and improve the identification of low-abundant CWPs. Even more interesting and relevant for cell wall proteomics studies, is the new equalization technology based on the use of combinatorial hexapeptide ligand libraries (CPLLs) to reduce the dynamic range of protein concentrations (Fröhlich and Lindermayr, 2011). CPLLs consist in 64 millions different peptides fixed to a single bead commercially available. Specific binding of proteins depends on the physico-chemical properties of each protein. Highly-abundant proteins quickly saturate their ligands whereas all low-abundant ones are bound, resulting after elution, in a narrower dynamic range of all the proteins initially present. First successful results have been obtained for proteomics of spinach leaves (Fasoli et al., 2011), leaf extracts of *A. thaliana*, and phloem exudates of pumpkin (Fröhlich et al., 2012). Protein extraction is a critical step since native conformation of proteins is required for interaction with CPLLs. Notwithstanding the different limitations of this technique (Fröhlich and Lindermayr, 2011), it could be applied to the study of plant cell wall proteomes providing sufficient amounts of proteins are obtained. This approach would undoubtedly permit to identify new low-abundant proteins.

PROTEIN–PROTEIN INTERACTIONS IN CELL WALLS

Many protein/protein interactions are expected in plant cell walls. Indeed, enzymes and their inhibitors like proteases and protease inhibitors, PMEs and PME inhibitors (PMEIs), or proteins with LRR domains have been detected in cell wall proteomes (Jamet et al., 2008). However, the present knowledge on plant CWPs suffers from the lack of data on protein–protein interaction mapping since most of the protein extraction methods used did not preserve supramolecular assemblies. One of the future challenges in plant cell wall proteomics will consist in developing extraction and capture methods to analyze CWP complexes. Concerning the purification of protein complexes, tandem affinity purification (TAP), in combination with MS, has become the method of choice to explore *in vivo* protein interactions (Xu et al., 2010). This method is based on the expression of a target protein fused to a double affinity tag. The first successful study of nuclear and cytoplasmic plant protein complexes using the TAP

method has been carried out in a transient expression system of *N. benthamiana* (Rohila et al., 2004). The method has been optimized for use in plants. Since this first report, only a limited number of plant protein complexes through TAP have been reported from *A. thaliana* and *O. sativa* (Rohila et al., 2009; Andrès et al., 2011). It would be of special interest to carry out this method to analyze CWPs with predicted protein interaction domains, thus permitting to identify their partners. Further optimization will be necessary for (i) CWP extraction, possibly associated with protein cross-linking treatment, and (ii) protein complex capture which will require the design of a new TAP tag to preserve the level of accumulation of CWPs as well as their localization, stability, and function. Alternatively, the analysis of intact CWP assemblies could be conducted by applying low energy MS methods preserving non-covalent interactions developed in the frame of the analysis of mammalian or microbial protein complexes (Stengel et al., 2012). Such approaches should provide a more detailed description of plant CWP complex architecture.

NEW MS TOOLS TO IMPROVE CELL WALL PROTEOME DESCRIPTION

Overcoming the future challenges in plant cell wall proteomics including analysis of low-abundance proteins, PTMs, protein–protein interactions, and quantitative proteomics will be facilitated by significant advances in MS technologies (Thelen and Miernyk, 2012). MS instrumentation evolves very quickly and impressive improvement in sensitivity, mass accuracy and fragmentation has been achieved in recent years. Instruments like Fourier Transformed-Ion Cyclotron Resonance (FT-ICR) are capable of mass accuracy of less than 2 ppm and have a high resolution (above 10^6). Sensitivity of new generation MS instruments reaches the femtomole or the attomole range. New fragmentation methods such as electron capture dissociation (ECD) and electron transfer dissociation (ETD) (Bond and Kohler, 2007) are also very promising. They will provide new insight into the structure of CWPs, as recently achieved for the AGP31, an *A. thaliana* cell wall *O*-glycoprotein (Hijazi et al., 2012). Finally, progresses in bioinformatics will be very helpful to characterize cell wall glycoproteins. Several computer programs like GlycoMod (Cooper et al., 2001), GlysodeIQ™(Joshi et al., 2004), GlycoMiner (Ozohanics et al., 2008), or Peptoonist (Goldberg et al., 2007) have been developed, but most of them do not consider plant glycan specificities. The *ProTerNyc* software has been developed in this purpose and efficiently used to predict *N*-glycan motifs on cell wall glycoproteins (Albenne et al., 2009; Zhang et al., 2011). However, additional bioinformatics tools should be developed to improve automatic data interpretation.

BEYOND CELL WALL PROTEOMICS

IDENTIFICATION OF CANDIDATE PROTEINS AND SEARCH FOR FUNCTION

In addition to the basic work of protein identification resulting in lists of proteins, cell wall proteomics has become a new tool to identify candidate proteins involved in developmental processes or in response to environmental cues. Some examples involve quantitative analyses. Up to now, label-free techniques have been favored like quantification of stained spots on polyacrylamide gels (Ndimba et al., 2003; Oh et al., 2005; Tran and Plaxton, 2008), spectral counting (Irshad et al., 2008) or calculation of area under the curve (AUC) (Cheng et al., 2009). Only one study has been performed with the difference in gel electrophoresis (DIGE) technique which requires the labeling of proteins with fluorescent dyes prior to electrophoresis (Casasoli et al., 2008). Quantifications performed on stained spots are difficult to interprete since some proteins are present in different spots for different reasons such as presence of PTMs, degradation or maturation of proteins. In addition, contrary to staining with fluorescent molecules like Sypro® Ruby, staining with Coomassie blue or silver nitrate has a narrow dynamic range, i.e., about two orders of magnitude (Moritz and Meyer, 2003). Only a few of these proteomics studies has given rise to functional or structural studies of proteins. A protein containing a GDSL motif lipase/hydrolase (GLIP1) has been identified as one of the salicylic acid (SA) responsive proteins secreted by *A. thaliana* cell suspension cultures (Oh et al., 2005). The increase in protein level was calculated to be three-fold after comparison of proteins extracted from control and from SA-treated cells separated by 2D-E and stained with silver nitrate. Two *glip1* T-DNA insertion mutants have been found to be more resistant to the *Alternaria brassicicola* necrotrophic fungus. It has also been shown that the recombinant GLIP1 protein has a lipase activity and an antimicrobial activity able to disrupt the integrity of fungal spores. The AGP AGP31 has been identified as a major protein in the cell wall proteome of etiolated hypocotyls of *A. thaliana* (Irshad et al., 2008). AGP31 is a multi-domain proteins having a N-terminal AGP, a central Pro-rich and a C-terminal Cys-rich domains. The combination of several MS technologies has allowed the first description of the Pro hydroxylation and *O*-glycosylation patterns of its Pro-rich domain (Hijazi et al., 2012). Finally, the *N. tabacum* NtSCP1 serine carboxypeptidase III identified in leaf intercellular fluids has been later shown to be involved in cell elongation (Delannoy et al., 2008; Bienert et al., 2012). The protease activity of NtSCP1 has been demonstrated *in vitro* and its cell wall localization has been confirmed by expression of the protein as a GFP fusion protein *in vivo*. Over-expression of *NtSCP1* has led to reduce flower length due to decrease in cell size and to etiolated seedlings with short hypocotyls.

In addition to proteomics data, it would be interesting to consider other data to identify proteins of interest such as transcriptomics or gene regulatory networks. Such data are available online (e.g., http://bar.utoronto.ca/welcome.htm, https://www.genevestigator.com/gv/, http://aranet.mpimp-golm.mpg.de/, http://atted.jp/). The feeding of new portals like MASCP Gator which aims at unifying the *A. thaliana* proteomics resources in a single interface for the research community is also essential (Joshi et al., 2011). This systems biology approach would allow a better understanding of gene regulation, from gene transcription to protein synthesis and even PTMs, allowing the description of protein active forms. Indeed, several studies have shown that proteomics and transcriptomics data are complementary and do not give exactly the same picture of a physiological situation depending on the level of regulation of gene expression (Jamet et al., 2009; Minic et al., 2009).

FROM PROTEOMICS TO PEPTIDOMICS

During the last decade, it has become evident that secreted peptides function as signaling molecules in cell-to-cell communication in plants. Recently, they have been recognized as hormones that coordinate and specify cellular functions in complex developmental processes (Shinohara and Matsubayashi, 2010; Murphy et al., 2012). The secreted signaling peptides identified so far can be categorized in two groups: (i) the small post-translationally modified peptides are less than 20 amino acids and undergo extensive proteolytic processing from a longer precursor and PTMs such as Tyr sulfation, Pro hydroxylation and arabinosylation on Hyp; (ii) the Cys-rich peptides are larger (<160 amino acids), cationic at the extracellular pH and have multiple intramolecular disulfide bonds. All of them have a predicted signal peptide. Peptidomics is becoming a stimulating field especially because of the description of the active forms of the signaling peptides (Shinohara and Matsubayashi, 2010; Murphy et al., 2012). Indeed, the PTMs are essential for their biological activity. For instance, STOMAGEN, a Cys-rich peptide that positively regulates stomatal density in *A. thaliana*, is active at nanomolar (10 nM) concentrations when forming three disulphide bonds. When Cys residues were replaced by Ser residues, STOMAGEN was unable to increase stomatal density even at very high (10 μM) concentrations (Ohki et al., 2011). Tyr sulfation and arabinosylation were also required for the full activity of small post-translationally modified peptides (Ohyama et al., 2009; Matsuzaki et al., 2010). Like cell wall proteomics, peptidomics has the potential to reveal new secreted signaling peptides as well as new functions of plant cell walls.

Today, the description of cell wall peptidomes, or secreted peptidomes, defined as a set of peptides present in cell walls at a specified physiological state, is lacking. Two main reasons can explain such a gap. First, the signaling peptides are believed to be present in very low quantity in plant tissues, they are active at nanomolar concentrations and their transcripts have been found to be transiently expressed (Ito et al., 2006; Chen et al., 2008b). Most of the well-characterized signaling peptides have been identified by genetics and *in silico* approaches (Murphy et al., 2012). In order to fully characterize the structure of their mature forms, they have been produced by cells or plants over-expressing the corresponding genes to obtain sufficient amounts of peptides amenable to LC-MS-based structure analysis (Amano et al., 2007; Ohyama et al., 2008, 2009; Sugano et al., 2010) or to *in situ* MALDI-TOF-MS analysis (Kondo et al., 2006). The latter study used *A. thaliana* plants over-expressing *CLV3*, a gene encoding a 96 amino acid propeptide containing a signal peptide. *CLV3* has been shown to be involved in the control of the size of the shoot apical meristem. The identified mature peptide contained 12 amino acids from Arg^{70} to His^{81} in CLV3 in which two of three Pro residues were modified to Hyp. Nevertheless, a number of studies that employed LC purification and Edman sequencing or MS identification have been developed and applied successfully to the analysis of native peptide sources (Pearce et al., 1991; Matsubayashi and Sakagami, 1996; Ito et al., 2006; Chen et al., 2008b). When these studies used high amounts of plants as starting material, they have allowed the identification of mature signaling peptides active at a concentration of 10^{-11}M in the case of TDIF (Tracheary element Differentiation Inhibitory Factor) (Ito et al., 2006). The corresponding cDNA of *Zinnia elegans* encodes a protein of 132 amino acids, but only 12, from His^{120} to Asn^{131}, match the TDIF sequence with two Hyp residues (Hyp^{123} and Hyp^{126}).

Second, peptide-encoding genes are frequently overlooked during the annotation of genomes. Indeed, gene prediction programs hardly distinguish between short, often intronless peptide-encoding genes and random small open reading frames (ORFs). To minimize incorrect gene predictions, it is common that small ORFs are rejected (Olsen et al., 2002). To overcome such a deficiency, a bioinformatics approach has been undertaken to identify candidate peptide-encoding genes in the *A. thaliana* genome (Lease and Walker, 2006). It has led to an unannotated secreted peptide database containing 33,809 ORFs. The identified peptides have been characterized by the presence of a predicted N-terminal signal peptide and by the absence of transmembrane domains and ER retention sequences. Since the expression of some ORFs has been detected by RT-PCR, it is suggested that the number and diversity of plant peptides is broader than currently assumed (Lease and Walker, 2006). The secreted peptide database will permit the necessary retrieval of information required for the identification of *A. thaliana* signaling peptides. Together with the progress of MS sensitivity, cell wall peptidomics is now a reachable objective.

CONCLUSION: UP AND COMING OF CELL WALL PROTEOMICS

Within the last 10 years, cell wall proteomics studies have received full credit among the OMICS strategies. They have allowed not only the precise identification of proteins in particular physiological conditions, but also their quantification and the characterization of their PTMs. Proteomics could also provide information about the dynamics of CWPs by kinetics analysis to follow the *de novo* synthesis of proteins or their degradation during plant development or in response to environmental cues. All the knowledge presently available on cell wall proteomics contributes to a better understanding of CWP structures and functions in cell walls. However, it is not yet possible to distinguish proteomes of primary and secondary walls notably because it is difficult to separate the cells surrounded by either of them. Micro-dissection of tissues should help solving this problem providing enough material can be obtained, but the extraction of proteins from the intricate macromolecular networks of secondary walls remains a great challenge. Next development will take advantage of cutting-edge MS technologies for a better coverage of cell wall proteomes, a more precise description of protein forms and protein complexes and for an insight into cell wall peptidomics.

ACKNOWLEDGMENTS

The authors are thankful to Université Paul Sabatier (Toulouse, France) and CNRS for supporting their research work. The authors also wish to thank Hélène San Clemente for her contribution to the bioinformatics work. This work has been done at LRSV, part of the "Laboratoire d'Excellence" (LABEX) entitled TULIP (ANR -10-LABX-41). The authors wish to apologize for the papers not quoted in this review, due to lack of space.

REFERENCES

Albenne, C., Canut, H., Boudart, G., Zhang, Y., San Clemente, H., Pont-Lezica, R., et al. (2009). Plant cell wall proteomics: mass spectrometry data, a trove for research on protein structure/function relationships. *Mol. Plant.* 2, 977–989.

Alvarez, S., Goodger, J. Q., Marsh, E. L., Chen, S., Asirvatham, V. S., and Schachtman, D. P. (2006). Characterization of the maize xylem sap proteome. *J. Proteome Res.* 5, 963–972.

Amano, Y., Tsubouchi, H., Shinohara, H., Ogawa, M., and Matsubayashi, Y. (2007). Tyrosine-sulfated glycopeptide involved in cellular proliferation and expansion in Arabidopsis. *Proc. Natl. Acad. Sci. U.S.A.* 104, 18333–18338.

Andrès, C., Agne, B., and Kessler, F. (2011). Preparation of multiprotein complexes from Arabidopsis chloroplasts using tandem affinity purification. *Methods Mol. Biol.* 775, 31–49.

Arabidopsis Genome Initiative. (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408, 796–815.

Basu, U., Francis, J. L., Whittal, R. W., Stephens, J. L., Wang, Y., Zaiane, O. R., et al. (2006). Extracellular proteomes of *Arabidopsis thaliana* and *Brassica napus* roots: analysis and comparison by MUdPIT and LC-MS/MS. *Plant Soil* 286, 357–376.

Baumberger, N., Ringli, C., and Keller, B. (2001). The chimeric leucine-rich repeat/extensin cell wall protein LRX1 is required for root hair morphogenesis in *Arabidopsis thaliana*. *Genes Dev.* 15, 1128–1139.

Bayer, E. M., Bottrill, A. R., Walshaw, J., Vigouroux, M., Naldrett, M. J., Thomas, C. L., et al. (2006). Arabidopsis cell wall proteome defined using multidimensional protein identification technology. *Proteomics* 6, 301–311.

Bendtsen, J. D., Jensen, L. J., Blom, N., Von Heijne, G., and Brunak, S. (2004). Feature-based prediction of non-classical and leaderless protein secretion. *Protein Eng. Des. Sel.* 17, 349–356.

Berger, D., and Altmann, T. (2000). A subtilisin-like serine protease involved in the regulation of stomatal density and distribution in *Arabidopsis thaliana*. *Genes Dev.* 14, 1119–1131.

Bhushan, D., Jaiswal, D. K., Ray, D., Basu, D., Datta, A., Chakraborty, S., et al. (2011). Dehydration-responsive reversible and irreversible changes in the extracellular matrix: comparative proteomics of chickpea genotypes with contrasting tolerance. *J. Proteome Res.* 10, 2027–2046.

Bhushan, D., Pandey, A., Chattopadhyay, A., Choudhary, M. K., Chakraborty, S., Datta, A., et al. (2006). Extracellular matrix proteome of chickpea (*Cicer arietinum* L.) illustrates pathway abundance, novel protein functions and evolutionary perspect. *J. Proteome Res.* 5, 1711–1720.

Bienert, M. D., Delannoy, M., Navarre, C., and Boutry, M. (2012). NtSCP1 from tobacco is an extracellular serine carboxypeptidase III that has an impact on cell elongation. *Plant Physiol.* 158, 1220–1229.

Bond, M. R., and Kohler, J. J. (2007). Chemical methods for glycoprotein discovery. *Curr. Opin. Chem. Biol.* 11, 52–58.

Borderies, G., Jamet, E., Lafitte, C., Rossignol, M., Jauneau, A., Boudart, G., et al. (2003). Proteomics of loosely bound cell wall proteins of *Arabidopsis thaliana* cell suspension cultures: a critical analysis. *Electrophoresis* 24, 3421–3432.

Borner, G. H., Lilley, K. S., Stevens, T. J., and Dupree, P. (2003). Identification of glycosylphosphatidylinositol-anchored proteins in Arabidopsis. A proteomic and genomic analysis. *Plant Physiol.* 132, 568–577.

Boudart, G., Jamet, E., Rossignol, M., Lafitte, C., Borderies, G., Jauneau, A., et al. (2005). Cell wall proteins in apoplastic fluids of *Arabidopsis thaliana* rosettes: identification by mass spectrometry and bioinformatics. *Proteomics* 5, 212–221.

Brady, J. D., Sadler, I. H., and Fry, S. C. (1996). Di-isodityrosine, a novel tetrametric derivative of tyrosine in plant cell wall proteins: a new potential cross-link. *Biochem. J.* 315, 323–327.

Brisson, L. F., Tenhaken, R., and Lamb, C. (1994). Function of oxidative cross linking of cell wall structural proteins in plant disease resistance. *Plant Cell* 6, 1703–1712.

Carpin, S., Crevecoeur, M., De Meyer, M., Simon, P., Greppin, H., and Penel, C. (2001). Identification of a Ca(2+)-pectate binding site on an apoplastic peroxidase. *Plant Cell* 13, 511–520.

Carpita, N. C., and Gibeaut, D. M. (1993). Structural models of primary cell walls in flowering plants, consistency of molecular structure with the physical properties of the walls during growth. *Plant J.* 3, 1–30.

Casasoli, M., Spadoni, S., Lilley, K. S., Cervone, F., De Lorenzo, G., and Mattei, B. (2008). Identification by 2-D DIGE of apoplastic proteins regulated by oligogalacturonides in *Arabidopsis thaliana*. *Proteomics* 8, 1042–1054.

Cassab, G. I., and Varner, J. E. (1988). Cell wall proteins. *Annu. Rev. Plant Physiol. Plant Biol.* 39, 321–353.

Catalá, C., Howe, K. J., Hucko, S., Rose, J. K., and Thannhauser, T. W. (2011). Towards characterization of the glycoproteome of tomato (*Solanum lycopersicum*) fruit using Concanavalin A lectin affinity chromatography and LC-MALDI-MS/MS analysis. *Proteomics* 11, 1530–1544.

Charmont, S., Jamet, E., Pont-Lezica, R., and Canut, H. (2005). Proteomic analysis of secreted proteins from *Arabidopsis thaliana* seedlings: improved recovery following removal of phenolic compounds. *Phytochemistry* 66, 453–461.

Chen, M. H., Huang, L. F., Li, H. M., Chen, Y. R., and Yu, S. M. (2004). Signal peptide-dependent targeting of a rice alpha-amylase and cargo proteins to plastids and extracellular compartments of plant cells. *Plant Physiol. Biochem.* 135, 1367–1377.

Chen, X. Y., Kim, S. T., Cho, W. K., Rim, Y., Kim, S., Kim, S. W., et al. (2008a). Proteomics of weakly bound cell wall proteins in rice calli. *J. Plant Physiol.* 166, 665–685.

Chen, Y. C., Siems, W. F., Pearce, G., and Ryan, C. A. (2008b). Six peptide wound signals derived from a single precursor protein in *Ipomoea batatas* leaves activate the expression of the defense gene sporamin. *J. Biol. Chem.* 283, 11469–11476.

Cheng, F.-Y., Blackburn, K., Lin, Y.-M., Goshe, M. B., and Wiliamson, J. D. (2009). Absolute protein quantification by LC/MS for global analysis of salicylic acid-induced plant protein secretion responses. *J. Proteome Res.* 8, 82–93.

Chivasa, S., Ndimba, B. K., Simon, W. J., Robertson, D., Yu, X.-L., Knox, J. P., et al. (2002). Proteomic analysis of the *Arabidopsis thaliana* cell wall. *Electrophoresis* 23, 1754–1765.

Cho, W. K., Chen, X. Y., Chu, H., Rim, Y., Kim, S., Kim, S. T., et al. (2009). The proteomic analysis of the secretome of rice calli. *Physiol. Plant.* 135, 331–341.

Cooper, C. A., Gasteiger, E., and Packer, N. H. (2001). GlycoMod-a software tool for determining glycosylation compositions from mass spectrometric data. *Proteomics* 1, 340–349.

Corthals, G. L., Wasinger, V. C., Hochstrasser, D. F., and Sanchez, J. C. (2000). The dynamic range of protein expression: a challenge for proteomic research. *Electrophoresis* 21, 1104–1115.

Cosgrove, D. J. (2005). Growth of the plant cell wall. *Nat. Rev. Mol. Cell. Biol.* 6, 850–861.

Dafoe, N. J., and Constabel, P. (2009). Proteomic analysis of hybrid poplar xylem sap. *Phytochemistry* 70, 856–863.

Dani, V., Simon, W. J., Duranti, M., and Croy, R. R. (2005). Changes in the tobacco leaf apoplast proteome in response to salt stress. *Proteomics* 5, 737–745.

Delannoy, M., Alves, G., Vertommen, D., Ma, J., Boutry, M., and Navarre, C. (2008). Identification of peptidases in *Nicotiana tabacum* leaf intercellular fluid. *Proteomics* 8, 2285–2298.

Dunkley, T. P., Hester, S., Shadforth, I. P., Runions, J., Weimar, T., Hanton, S. L., et al. (2006). Mapping the Arabidopsis organelle proteome. *Proc. Natl. Acad. Sci. U.S.A.* 103, 6518–6523.

Farrar, K., Evans, I. M., Topping, J. F., Souter, M. A., Nielsen, J. E., and Lindsey, K. (2003). EXORDIUM-a gene expressed in proliferating cells and with a role in meristem function, identified by promoter trapping in Arabidopsis. *Plant J.* 33, 61–73.

Fasoli, E., D'amato, A., Kravchuk, A. V., Boschetti, E., Bachi, A., and Righetti, P. G. (2011). Popeye strikes again: the deep proteome of spinach leaves. *J. Proteomics* 74, 127–136.

Faye, L., Boulaflous, A., Benchabane, M., Gomord, V., and Michaud, D. (2005). Protein modifications in the plant secretory pathway: current status and practical implications in molecular pharming. *Vaccine* 23, 1770–1778.

Feiz, L., Irshad, M., Pont-Lezica, R. F., Canut, H., and Jamet, E. (2006). Evaluation of cell wall preparations for proteomics: a new procedure for purifying cell walls from Arabidopsis hypocotyls. *Plant Methods* 2, 10.

Fernández, M. B., Pagano, M. R., Daleo, G. R., and Guevara, M. G. (2012). Hydrophobic proteins secreted into the apoplast may contribute to resistance against *Phytophthora infestans* in potato. *Plant Physiol. Biochem.* 60, 59–66.

Fröhlich, A., Gaupels, F., Sarioglu, H., Holzmeister, C., Spannagl, M., Durner, J., et al. (2012). Looking deep inside: detection of low-abundance proteins in leaf extracts of Arabidopsis and phloem exudates of pumpkin. *Plant Physiol.* 159, 902–914.

Fröhlich, A., and Lindermayr, C. (2011). Deep insights into the plant proteome by pretreatment with combinatorial hexapeptide ligand libraries. *J. Proteomics* 74, 1182–1189.

Fry, S. C. (2004). Primary cell wall metabolism: tracking the careers of wall polymers in living plant cells. *New Phytol.* 161, 641–675.

Goldberg, D., Bern, M., Parry, S., Sutton-Smith, M., Panico, M., Morris, H. R., et al. (2007). Automated *N*-glycopeptide identification using a combination of single- and tandem-MS. *J. Proteome Res.* 6, 3995–4005.

Goulet, C., Goulet, C., Goulet, M.-C., and Michaud, D. (2010). 2-DE proteome maps for the leaf apoplast of *Nicotiana benthamiana*. *Proteomics* 10, 2536–2544.

Haslam, R. P., Downie, A. L., Raventon, M., Gallardo, K., Job, D., Pallett, K. E., et al. (2003). The assessment of enriched apoplastic extracts using proteomic approaches. *Ann. Appl. Biol.* 143, 81–91.

Heazlewood, J., Verboom, R., Tonti-Filippini, J., Small, I., and Millar, A. (2007). SUBA: the Arabidopsis subcellular database. *Nucleic Acids Res.* 35, D213–D218.

Heazlewood, J. L., Tonti-Filippini, J., Verboom, R. E., and Millar, A. H. (2005). Combining experimental and predicted datasets for determination of the subcellular location of proteins in Arabidopsis. *Plant Physiol.* 139, 598–609.

Hijazi, M., Durand, J., Pichereaux, C., Pont, F., Jamet, E., and Albenne, C. (2012). Characterization of the Arabinogalactan Protein 31 (AGP31) of *Arabidopsis thaliana*: new advances on the Hyp-*O*-glycosylation of the Pro-rich domain. *J. Biol. Chem.* 287, 9623–9632.

Hunt, L., Bailey, K. J., and Gray, J. E. (2010). The signalling peptide EPFL9 is a positive regulator of stomatal development. *New Phytol.* 183, 609–614.

Imai, K., and Nakai, K. (2010). Prediction of subcellular locations of proteins: where to proceed? *Proteomics* 10, 3970–3983.

International Rice Genome Sequencing Project. (2005). The map-based sequence of the rice genome. *Nature* 436, 793–800.

Irshad, M., Canut, H., Borderies, G., Pont-Lezica, R., and Jamet, E. (2008). A new picture of cell wall protein dynamics in elongating cells of *Arabidopsis thaliana*: confirmed actors and newcomers. *BMC Plant Biol.* 8:94. doi: 10.1186/1471-2229-8-94

Ito, Y., Nakanomyo, I., Motose, H., Iwamoto, K., Sawa, S., Dohmae, N., et al. (2006). Dodeca-CLE peptides as suppressors of plant stem differentiation. *Science* 313, 8842–8845.

Jamet, E., Albenne, C., Boudart, G., Irshad, M., Canut, H., and Pont-Lezica, R. (2008). Recent advances in plant cell wall proteomics. *Proteomics* 8, 893–908.

Jamet, E., Canut, H., Boudart, G., and Pont-Lezica, R. F. (2006). Cell wall proteins: a new insight through proteomics. *Trends Plant Sci.* 11, 33–39.

Jamet, E., Roujol, D., San Clemente, H., Irshad, M., Soubigou-Taconnat, L., Renou, J.-P., et al. (2009). Cell wall biogenesis of *Arabidopsis thaliana* elongating cells: transcriptomics complements proteomics. *BMC Genomics* 10:505. doi: 10.1186/1471-2164-10-505

Javelle, M., Vernoud, V., Rogowsky, P. M., and Ingram, G. C. (2010). Epidermis: the formation and functions of a fundamental plant tissue. *New Phytol.* 189, 17–39.

Joshi, H. J., Harrison, M. J., Schulz, B. L., Cooper, C. A., Packer, N. H., and Karlsson, N. G. (2004). Development of a mass fingerprinting tool for automated interpretation of oligosaccharide fragmentation data. *Proteomics* 4, 1650–1664.

Joshi, H. J., Hirsch-Hoffmann, M., Baerenfaller, K., Gruissem, W., Baginsky, S., Schmidt, R., et al. (2011). MASCP Gator: an aggregation portal for the visualization of Arabidopsis proteomics data. *Plant Physiol.* 155, 259–270.

Jung, Y.-H., Jeong, S.-H., Kim, S., Singh, R., Lee, J.-E., Cho, Y.-S., et al. (2008). Systematic secretome analyses of rice leaf and seed callus suspension-cultured cells: workflow development and establishment of high-density two-dimensional gel reference maps. *J. Proteome Res.* 7, 5187–5210.

Kasprzewska, A. (2003). Plant chitinases-regulation and function. *Cell Mol. Biol. Lett.* 8, 809–824.

Kehr, J., Buhtz, A., and Giavalisco, P. (2005). Analysis of xylem sap proteins from *Brassica napus*. *BMC Plant Biol.* 5:11. doi: 10.1186/1471-2229-5-11

Kieliszewski, M. J. (2001). The latest hype on Hyp-O-glycosylation codes. *Phytochemistry* 57, 319–323.

Kieliszewski, M. J., and Lamport, D. T. A. (1994). Extensin: repetitive motifs, functional sites, post-translational codes, and phylogeny. *Plant J.* 5, 157–172.

Kimura, Y., Hess, D., and Sturm, A. (1999). The *N*-glycans of jack bean a-mannosidase. *Eur. J. Biochem.* 264, 168–175.

Komatsu, S., Kobayashi, Y., Nishizawa, N., Nanjo, Y., and Furukawa, K. (2010). Comparative proteomics analysis of differentially expressed proteins in soybean cell wall during flooding stress. *Amino Acid* 39, 1435–1449.

Kondo, T., Sawa, S., Kinoshita, A., Mizuno, S., Kakimoto, T., Fukuda, H., et al. (2006). A plant peptide encoded by CLV3 identified by *in situ* MALDI-TOF MS analysis. *Science* 313, 845–848.

Krishnan, H. B., and Natarajan, S. S. (2009). A rapid method for depletion of Rubisco from soybean (*Glycine max*) leaf for proteomic analysis of lower abundance proteins. *Phytochemistry* 70, 1958–1964.

Krishnan, H. B., Oehrle, N. W., and Natarajan, S. S. (2009). A rapid and simple procedure for the depletion of abundant storage proteins from legume seeds to advance proteome analysis: a case study using *Glycine max*. *Proteomics* 9, 3174–3188.

Kwon, H.-K., Yokoyama, R., and Nishitani, K. (2005). A proteomic approach to apoplastic proteins involved in cell wall regeneration in protoplasts of Arabidopsis suspension-cultured cells. *Plant Cell Physiol.* 46, 843–857.

Lease, K. A., and Walker, J. C. (2006). The Arabidopsis unannotated secreted peptide database, a resource for plant peptidomics. *Plant Physiol.* 142, 831–838.

Leasure, C. D., and He, Z. H. (2012). CLE and RGF family peptide hormone signaling in plant development. *Mol. Plant*. 5, 1173–1175.

Lee, R. C., Hrmova, M., Burton, R. A., Lahnstein, J., and Fincher, G. B. (2003). Bifunctional family 3 glycoside hydrolases from barley with α -L-arabinofuranosidase and β-D-xylosidase activity. Characterization, primary structures, and COOH-terminal processing. *J. Biol. Chem.* 278, 5377–5387.

Lee, S.-J., and Rose, J. K. C. (2012). A yeast secretion trap assay for identification of secreted proteins from eukaryotic phytopathogens and their plant hosts. *Methods Mol. Biol.* 835, 519–530.

Leiber, R. M., John, F., Verhertbruggen, Y., Diet, A., Knox, J. P., and Ringli, C. (2010). The TOR pathway modulates the structure of cell walls in Arabidopsis. *Plant Cell* 22, 1898–1908.

Ligat, L., Lauber, E., Albenne, C., San Clemente, H., Valot, B., Zivy, M., et al. (2011). Analysis of the xylem sap proteome of *Brassica oleracea* reveals a high content in secreted proteins. *Proteomics* 11, 1798–1813.

Lige, B., Shengwu, M., and Van Huystee, R. B. (2001). The effects of the site-directed removal of N-glycosylation from cationic peanut peroxidase on its function. *Arch. Biochem. Biophys.* 386, 17–24.

Lim, J.-M., Aoki, K., Angel, P., Garrison, D., King, D., Tiemeyer, M., et al. (2009). Mapping glycans onto specific *N*-linked glycosylation sites of *Pyrus communis* PGIP redefines the interface for EPG-PGIP interactions. *J. Proteome Res.* 8, 673–680.

Lim, S., Chisholm, K., Coffin, R. H., Peters, R. D., Al-Mughrabi, K. I., Wang-Pruski, G., et al. (2012). Protein profiling in potato (*Solanum tuberosum* L.) leaf tissues by differential centrifugation. *J. Proteome Res.* 11, 2594–2601.

Lundberg, E., and Uhlén, M. (2010). Creation of an antibody-based subcellular protein atlas. *Proteomics* 10, 3984–3996.

Matsubayashi, Y., and Sakagami, Y. (1996). Phytosulfokine, sulfated peptides that induce the proliferation of single mesophyll cells of *Asparagus officinalis* L. *Proc. Natl. Acad. Sci. U.S.A.* 93, 7623–7627.

Matsuzaki, Y., Ogawa-Ohnishi, M., Mori, A., and Matsubayashi, Y. (2010). Secreted peptide signals required for maintenance of root stem cell niche in Arabidopsis. *Science* 329, 1065–1067.

Membré, N., Bernier, F., Staiger, D., and Berna, A. (2000). *Arabidopsis thaliana* germin-like proteins: common and specific features point to a variety of functions. *Planta* 211, 345–354.

Millar, D. J., Whitelegge, J. P., Bindschedler, L. V., Rayon, C., Boudet, A. M., Rossignol, M., et al. (2009). The cell wall and secretory proteome of a tobacco cell line synthesising secondary wall. *Proteomics* 9, 2355–2372.

Minic, Z., Jamet, E., Negroni, L., Der Garabedian, P. A., Zivy, M., and Jouanin, L. (2007). A sub-proteome of *Arabidopsis thaliana* trapped on Concanavalin A is enriched in cell wall glycoside hydrolases. *J. Exp. Bot.* 58, 2503–2512.

Minic, Z., Jamet, E., San Clemente, H., Pelletier, S., Renou, J.-P., Rihouey, C., et al. (2009). Transcriptomic analysis of Arabidopsis developing stems: a close-up on cell wall

genes. *BMC Plant Biol.* 9:6. doi: 10.1186/1471-2229-9-6

Minic, Z., Rihouey, C., Do, C. T., Lerouge, P., and Jouanin, L. (2004). Purification and characterization of enzymes exhibiting β-D-xylosidase activities in stem tissues of Arabidopsis. *Plant Physiol.* 135, 867–878.

Morel, J., Claverol, S., Mongrand, S., Furt, F., Fromentin, J., Bessoule, J. J., et al. (2006). Proteomics of plant detergent-resistant membranes. *Mol. Cell. Proteomics* 5, 1393–1411.

Moritz, B., and Meyer, H. E. (2003). Approaches for the quantification of protein concentration ratios. *Proteomics* 3, 2208–2220.

Murphy, E., Smith, S., and De Smet, I. (2012). Small signaling peptides in Arabidopsis development: how cells communicate over a short distance. *Plant Cell* 24, 3198–3217.

Ndimba, B. K., Chivasa, S., Hamilton, J. M., Simon, W. J., and Slabas, A. R. (2003). Proteomic analysis of changes in the extracellular matrix of Arabidopsis cell suspension cultures induced by fungal elicitors. *Proteomics* 3, 1047–1059.

Nersissian, A. M., and Shipp, E. L. (2002). Blue copper-binding domains. *Adv. Protein Chem.* 60, 271–340.

Nickel, W., and Seedorf, M. (2008). Unconventional mechanisms of protein transport to the cell surface of eukaryotic cells. *Annu. Rev. Cell. Dev. Biol.* 24, 287–308.

Nieuwland, J., Feron, R., Huisman, B. A., Fasolino, A., Hilbers, C. W., Derksen, J., et al. (2005). Lipid transfer proteins enhance cell wall extension in tobacco. *Plant Cell* 17, 2009–2019.

Oh, I. S., Park, A. R., Bae, M. S., Kwon, S. J., Kim, Y. S., Lee, J. E., et al. (2005). Secretome analysis reveals an Arabidopsis lipase involved in defense against *Alternaria brassicicola*. *Plant Cell* 17, 2832–2847.

Ohki, S., Takeuchi, M., and Mori, M. (2011). The NMR structure of stomagen reveals the basis of stomatal density regulation by plant peptide hormones. *Nat. Commun.* 2:512. doi: 10.1038/ncomms1520

Ohyama, K., Ogawa, M., and Matsubayashi, Y. (2008). Identification of a biologically active, small, secreted peptide in Arabidopsis by *in silico* gene screening, followed by LC-MS-based structure analysis. *Plant J.* 55, 152–160.

Ohyama, K., Shinohara, H., Ogawa-Ohnishi, M., and Matsubayashi, Y. (2009). A glycopeptide regulating stem cell fate in *Arabidopsis thaliana*. *Nat. Chem. Biol.* 5, 578–580.

Olsen, A. N., Mundy, J., and Skriver, K. (2002). Peptomics, identification of novel cationic Arabidopsis peptides with conserved sequence motifs. *In Silico Biol.* 2, 441–451.

Ozohanics, O., Krenyacz, J., Ludányi, K., Pollreisz, F., Vékey, K., and Drahos, L. (2008). GlycoMiner: a new software tool to elucidate glycopeptide composition. *Rapid Commun. Mass Spectrom.* 22, 3245–3254.

Passardi, F., Penel, C., and Dunand, C. (2004). Performing the paradoxical: how plant peroxidases modify the cell wall. *Trends Plant Sci.* 9, 534–540.

Pearce, G., Strydom, D., Johnson, S., and Ryan, C. A. (1991). A polypeptide from tomato leaves induces wound-inducible proteinase inhibitor proteins. *Science* 253, 895–898.

Pechanova, O., Hsu, C.-Y., Adams, J. P., Pechan, T., Vandervelde, L., Drnevich, J., et al. (2010). Apoplast proteome reveals that extracellular matrix contributes to multistress response in poplar. *BMC Genomics* 11:674. doi: 10.1186/1471-2164-11-674

Pinedo, M., Regente, M., Elizalde, M., Quiroga, I. Y., Pagnussat, L. A., Jorrin-Novo, J., et al. (2012). Extracellular sunflower proteins: evidence on non-classical secretion of a jacalin-related lectin. *Protein Pept. Lett.* 19, 270–276.

Pont-Lezica, R., Minic, Z., San Clemente, H., Roujol, D., and Jamet, E. (2010). "Plant cell wall functional genomics: novelties from proteomics," in *Advances in Genetics Research*, ed M. A. Osborne (Hauppauge, NY: Nova Science Publishers, Inc.).

Rautengarten, C., Usadel, B., Neumetzler, L., Hartmann, J., Büssis, D., and Altmann, T. (2008). A subtilisin-like serine protease essential for mucilage release from Arabidopsis seed coats. *Plant J.* 54, 466–480.

Regente, M., Pinedo, M., Elizalde, M., and De La Canal, L. (2012). Apoplastic exosome-like vesicles: a new way of protein secretion in plants. *Plant Signal. Behav.* 7, 1–3.

Robertson, D., Mitchell, G. P., Gilroy, J. S., Gerrish, C., Bolwell, G. P., and Slabas, A. R. (1997). Differential extraction and protein sequencing reveals major differences in patterns of primary cell wall proteins from plants. *J. Biol. Chem.* 272, 15841–15848.

Rohila, J. S., Chen, M. H., Cerny, R., and Fromm, M. (2004). Improved tandem affinity purification tag and methods for isolation of protein heterocomplexes from plants. *Plant J.* 38, 172–181.

Rohila, J. S., Chen, M. H., Chen, S., Chen, J., Cerny, R. L., Dardick, C., et al. (2009). Protein-protein interactions of tandem affinity purified protein kinases from rice. *PLoS ONE* 4:e6685. doi: 10.1371/journal.pone.0006685

Rose, J. K. C., and Lee, S.-J. (2010). Straying off the highway: trafficking of secreted plant proteins and complexity in the plant cell wall proteome. *Plant Physiol. Biochem.* 153, 433–436.

Ruiz-May, E., Thannhauser, T. W., Zhang, S., and Rose, J. K. C. (2012). Analytical technologies for identification and characterization of the plant *N*-glycoproteome. *Front. Plant Sci.* 3:150. doi: 10.3389/fpls.2012.00150

San Clemente, H., Pont-Lezica, R., and Jamet, E. (2009). Bioinformatics as a tool for assessing the quality of sub-cellular proteomic strategies and inferring functions of proteins: plant cell wall proteomics as a test case. *Bioinform. Biol. Insights* 3, 15–28.

Schultz, C. J., Ferguson, K. L., Lahnstein, J., and Bacic, A. (2004). Post-translational modifications of arabinogalactan-peptides of *Arabidopsis thaliana*. Endoplasmic reticulum and glycosylphosphatidylinositol-anchor signal cleavage sites and hydroxylation of proline. *J. Biol. Chem.* 279, 45503–45511.

Seifert, G. J., and Blaukopf, C. (2010). Irritable walls: the plant extracellular matrix and signaling. *Plant Physiol.* 153, 467–478.

Seifert, G. J., and Roberts, K. (2007). The biology of arabinogalactan proteins. *Annu. Rev. Plant Biol.* 58, 137–161.

Shinohara, H., and Matsubayashi, Y. (2010). Arabinosylated glycopeptide hormones: new insights into CLAVATA3 structure. *Curr. Opin. Plant Biol.* 13, 515–519.

Song, W., Henquet, M. G., Mentink, R. A., Van Dijk, A. J., Cordewener, J. H., Bosch, D., et al. (2011). *N*-glycoproteomics in plants: perspectives and challenges. *J. Proteomics* 74, 1463–1474.

Spadoni, S., Zabotina, O., Di Matteo, A., Mikkelsen, J. D., Cervone, F., De Lorenzo, G., et al. (2006). Polygalacturonase-inhibiting protein interacts with pectin through a binding site formed by four clustered residues of arginine and lysine. *Plant Physiol.* 141, 557–564.

Spiro, R. G. (2002). Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology* 12, 43R–56R.

Stengel, F., Aebersold, R., and Robinson, C. V. (2012). Joining forces: integrating proteomics and cross-linking with the mass spectrometry of intact complexes. *Mol. Cell. Proteomics* 11:R111.014027. doi: 10.1074/mcp.R111.014027

Sugano, S. S., Shimada, T., Imai, Y., Okawa, K., Tamai, A., Mori, M., et al. (2010). Stomagen positively regulates stomatal density in Arabidopsis. *Nature* 463, 241–246.

Tan, L., Showalter, A. M., Egelund, J., Hernandez-Sanchez, A., Doblin, M. S., and Bacic, A. (2012). Arabinogalactan-proteins and the research challenges for these enigmatic plant cell surface proteoglycans. *Front. Plant Sci.* 3:140. doi: 10.3389/fpls.2012.00140

Thelen, J. J., and Miernyk, J. A. (2012). The proteomic future: where mass spectrometry should be taking us. *Biochem. J.* 444, 169–181.

Thomas, R., Fang, X., Ranathunge, K., Anderson, T. R., Peterson, C. A., and Bernards, M. A. (2007). Soybean root suberin: anatomical distribution, chemical composition, and relationship to partial resistance to *Phytophthora sojae*. *Plant Physiol.* 144, 299–311.

Tomato Genome Consortium. (2012). The tomato genome sequence provides insights into fleshy fruit evolution. *Nature* 485, 635–641.

Tran, H. T., and Plaxton, W. C. (2008). Proteomic analysis of alterations in the secretome of *Arabidopsis thaliana* suspension cells subjected to nutritional deficiency. *Proteomics* 8, 4317–4326.

Tuskan, G. A., Difazio, S., Jansson, S., Bohlmann, J., Grigoriev, I., Hellsten, U., et al. (2006). The genome of black cottonwood, *Populus trichocarpa* (Torr. and Gray). *Science* 313, 1593–1604.

Van Der Hoorn, R. A. (2008). Plant proteases: from phenotypes to molecular mechanisms. *Annu. Rev. Plant Biol.* 59, 191–223.

Velasquez, M., Salter, J. S., Dorosz, J. G., Petersen, B. L., and Estevez, J. M. (2012). Recent advances on the posttranslational modifications of EXTs and their roles in plant cell walls. *Front. Plant Sci.* 3:93. doi: 10.3389/fpls.2012.00093

Verdonk, J., Hatfield, R., and Sullivan, M. (2012). Proteomics analysis of cell walls of two developmental

stages of alfalfa stems. *Front. Plant Sci.* 3:279. doi: 10.3389/fpls.2012.00279

Von Groll, U., Berger, D., and Altmann, T. (2002). The subtilisin-like serine protease SDD1 mediates cell-to-cell signaling during Arabidopsis stomatal development. *Plant Cell* 14, 1527–1539.

Wang, L., Li, Z., Qian, W., Guo, W., Gao, X., Huang, L., et al. (2011). The Arabidopsis purple acid phosphatase AtPAP10 is predominantly associated with the root surface and plays an important role in plant tolerance to phosphate limitation. *Plant Physiol.* 157, 1283–1299.

Watson, B. S., Lei, Z., Dixon, R. A., and Sumner, L. W. (2004). Proteomics of *Medicago sativa* cell walls. *Phytochemistry* 65, 1709–1720.

Widjaja, I., Naumann, K., Roth, U., Wolf, N., Mackey, D., Dangl, J. L., et al. (2009). Combining subproteome enrichment and Rubisco depletion enables identification of low abundance proteins differentially regulated during plant defense. *Proteomics* 9, 138–147.

Wilson, L. G., and Fry, J. C. (1986). Extensin, a major cell wall glycoprotein. *Plant Cell Environ.* 9, 239–260.

Wolf, S., Rausch, T., and Greiner, S. (2009). The N-terminal pro region mediates retention of unprocessed type-I PME in the Golgi apparatus. *Plant J.* 58, 361–375.

Xu, X., Song, Y. R., Li, Y., Chang, J., Zhang, H., and An, L. (2010). The tandem affinity purification method: an efficient system for protein complex purification and protein interaction identification. *Protein Expr. Purif.* 72, 149–156.

Yeats, T. H., Howe, K. J., Matas, A. J., Buda, G., Thannhauser, T. W., and Rose, J. K. (2010). Mining the surface proteome of tomato (*Solanum lycopersicum*) fruit for proteins associated with cuticle biogenesis. *J. Exp. Bot.* 61, 3759–3771.

Zhang, Y., Giboulot, A., Zivy, M., Valot, B., Jamet, E., and Albenne, C. (2011). Combining various strategies to increase the coverage of the plant cell wall glycoproteome. *Phytochemistry* 72, 1109–1123.

Zhou, L., Bokhari, S. A., Dong, C.-H., and Liu, J.-Y. (2011). Comparative proteomics analysis of the root apoplasts of rice seedlings in response to hydrogen peroxide. *PLoS ONE* 6:e16723. doi: 10.1371/journal.pone.0016723

Zhu, J., Alvarez, S., Marsh, E. L., Lenoble, M. E., Cho, I.-J., Sivaguru, M., et al. (2007). Cell wall proteome in the maize primary root elongation zone. II. Region-specific changes in water soluble and lightly ionically bound proteins under water deficit. *Plant Physiol.* 145, 1533–1548.

Zhu, J., Chen, S., Alvarez, S., Asirvatham, V. S., Schachtman, D. P., Wu, Y., et al. (2006). Cell wall proteome in the maize primary root elongation zone. I. Extraction and identification of water-soluble and lightly ionically bound proteins. *Plant Physiol.* 140, 311–325.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 19 December 2012; paper pending published: 05 March 2013; accepted: 10 April 2013; published online: 01 May 2013.

Citation: Albenne C, Canut H and Jamet E (2013) Plant cell wall proteomics: the leadership of Arabidopsis thaliana. Front. Plant Sci. ***4****:111. doi: 10.3389/fpls.2013.00111*

This article was submitted to Frontiers in Plant Proteomics, a specialty of Frontiers in Plant Science.

Copyright © 2013 Albenne, Canut and Jamet. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.

APPENDIX

Table A1 | Cell wall proteomes and secretomes of *A. thaliana*: detailed information.

Organ	References	Destructive or non-destructive method	Type of proteome	Buffer for protein extraction	Protein separation	Protein identification	Total number of identified proteins	Number of proteins with predicted target signal and no predicted intracellular retention signal
Step (see **Figure 1**)		Step 1		Step 2	Step 3	Step 4		
Culture medium of roots	Basu et al. (2006)		Secretome			2D-LC-MS/MS	52	36
Stems	Minic et al. (2007)	Destructive (grinding)	N-glycoproteome (ConA affinity chromatography)	Salt solution (200 mM CaCl2)	2D-E	LC-MS/MS	102	90
Leaves	Haslam et al. (2003)	Non-destructive (vacuum infiltration)	Apoplastic proteome	low ionic strength buffer	2D-E	MALDI-TOF MS	10	9
Leaves	Boudart et al. (2005)	Non-destructive (vacuum infiltration)	Apoplastic proteome	0.3 M mannitol or salt solutions (1 M NaCl, 200 mM CaCl2, 2 M LiCl, 50 mM CDTA), acidic pH	1D-E or 2D-E	MALDI-TOF MS	93	87
Etiolated hypocotyls (5 days)	Feiz et al. (2006)	Destructive (grinding in low ionic strength acidic buffer), cell wall purification	Cell wall proteome	Salt solutions (200 mM CaCl2, 2 M LiCl), acidic pH	1D-E	MALDI-TOF MS	96	84
Etiolated hypocotyls (5 and 11 days)	Irshad et al. (2008)	Destructive (grinding in low ionic strength acidic buffer), cell wall purification	Cell wall proteome	Salt solutions (200 mM CaCl2, 2 M LiCl), acidic pH	Cationic exchange chromatography followed by 1D-E	MALDI-TOF MS	173	137
Etiolated hypocotyls (11 days)	Zhang et al. (2011)	Destructive (grinding in low ionic strength acidic buffer), cell wall purification	Cell wall glycoproteome (lectin affinity chromatography, boronic acid chromatography)	Salt solutions (200 mM CaCl2, 2 M LiCl), acidic pH	1D-E	MALDI-TOF MS and LC-MS/MS	129	127

(Continued)

Table A1 | Continued

Organ	References	Destructive or non-destructive method	Type of proteome	Buffer for protein extraction	Protein separation	Protein identification	Total number of identified proteins	Number of proteins with predicted target signal and no predicted intracellular retention signal
Seedlings (10 days) (oligogalacturonides)	Casasoli et al. (2008)	Non-destructive	Apoplastic proteome	Salt solution (150 mM MgCl2, 50 mM EDTA), basic pH	2D-E	MALDI-TOF MS and LC-MS/MS	20	14
Cell suspension cultures (3 days)	Bayer et al. (2006)	Destructive (grinding in salt solution)	Cell wall proteome	Phenol buffer		2D-LC-MS/MS	792	106
Cell suspension cultures (5 days)	Kwon et al. (2005)	Non-destructive (washing of cells)	Cell wall proteome	Salt solution (1 M KCl)	2D-E	MALDI-TOF MS	39	34
Cell suspension cultures (5 days)	Chivasa et al. (2002)	Destructive (grinding)	Cell wall proteome	Salt solution, chaotropic agent (urea and thiourea) and detergent (CHAPS)	2D-E	MALDI-TOF MS	72	26
Cell suspension cultures (5 days)	Robertson et al. (1997)	Non-destructive (washing of cells)	Cell wall proteome	Salt solutions (0.2 M CaCl2, 50 mM CDTA, 1 M NaCl, 0.2 M borate)	1D-E	Edman N-terminal sequencing	33	18
Cell suspension cultures (7 days)	Borderies et al. (2003)	Non-destructive (washing of cells)	Cell wall proteome	Glycerol solution (50%) or salt solutions (successively 0.15 M NaCl, 1 M NaCl, 50 mM EDTA, 2 M LiCl in 50% glycerol)	1D-E or 2D-E	MALDI-TOF MS	95	50
Culture medium of calli (21 days	Borner et al. (2003)	Destructive (purification of a membrane fraction with detergent)	GPI-anchored proteome	Pi-PLC treatment, detergent (Triton X-114)	1D-E or 2D-E	LC-MS/MS	30	30
Culture medium of cell suspension cultures (4 days)	Oh et al. (2005)		Secretome		2D-E	MALDI-TOF MS	45	24

(Continued)

Table A1 | Continued

Organ	References	Destructive or non-destructive method	Type of proteome	Buffer for protein extraction	Protein separation	Protein identification	Total number of identified proteins	Number of proteins with predicted target signal and no predicted intracellular retention signal
Cell suspension cultures (7 days) (fungal elicitors)	Ndimba et al. (2003)	Destructive (grinding)	Cell wall proteome	Salt solution (200 mM CaCl2)	2D-E	MALDI-TOF MS	6	6
Culture medium of cell suspension cultures (7 days) (salicylic acid treatment)	Cheng et al. (2009)		Secretome			2D-LC-MS/MS	74	39
Culture medium of cell suspension cultures (7 days) (phosphate deficiency)	Tran and Plaxton (2008)		Secretome		2D-E	MALDI-TOF MS	37	21
Protoplasts (1 h)	Kwon et al. (2005)	Non-destructive (washing of cells)	Cell wall proteome	Salt solution (1 M KCl)	2D-E	MALDI-TOF MS	71	25
Protoplasts (3 h)	Kwon et al. (2005)	Non-destructive (washing of cells)	Cell wall proteome	Salt solution (1 M KCl)	2D-E	MALDI-TOF MS	66	23
Liquid cultured etiolated seedlings	Schultz et al. (2004)	Destructive (grinding)	Arabinogalactan proteins (AGPs) proteome	Detergent-containing buffer (Triton X-100), Yariv precipitation, HF deglycosylation	HPLC	Edman N-terminal sequencing	12	12
Culture medium of etiolated seedlings	Charmont et al. (2005)		Secretome		2D-E	MALDI-TOF MS	49	44

frontiers in
PLANT SCIENCE

REVIEW ARTICLE
published: 07 February 2013
doi: 10.3389/fpls.2013.00017

Cell wall proteomics of crops

Setsuko Komatsu[1]* and Yuki Yanagawa[2]

[1] *National Institute of Crop Science, National Agriculture and Food Research Organization, Tsukuba, Japan*
[2] *Plant Science Center, RIKEN Yokohama Institute, Yokohama, Japan*

Edited by:
Nicolas L. Taylor, The University of Western Australia, Australia

Reviewed by:
Niranjan Chakraborty, National Institute of Plant Genome Research, India
Toshiaki Mitsui, Niigata University, Japan
Phil A. Jackson, Instituto de Tecnologia Química e Biológica, Portugal

****Correspondence:***
Setsuko Komatsu, National Institute of Crop Science, National Agriculture and Food Research Organization, 2-1-18 Kannondai, Tsukuba 305-8518, Japan.
e-mail: skomatsu@affrc.go.jp

Cell wall proteins play key roles in cell structure and metabolism, cell enlargement, signal transduction, responses to environmental stress, and many other physiological events. Agricultural crops are often used for investigating stress tolerance because cultivars with differing degrees of tolerance are available. Abiotic and biotic stress factors markedly influence the geographical distribution and yields of many crop species. Crop cell wall proteomics is of particular importance for improving crop productivity, particularly under unfavorable environmental conditions. To better understand the mechanisms underlying stress response in crops, cell wall proteomic analyses are being increasingly utilized. In this review, the methods of purification and purity assays of cell wall protein fractions from crops are described, and the results of protein identification using gel-based and gel-free proteomic techniques are presented. Furthermore, protein composition of the cell walls of rice, wheat, maize, and soybean are compared, and the role of cell wall proteins in crops under flooding and drought stress is discussed. This review will be useful for clarifying the role of the cell wall of crops in response to environmental stresses.

Keywords: crop, proteomics, cell wall, drought stress, flooding stress

INTRODUCTION

The cell wall is an important sub-cellular organelle for the modulation of some stress signals and an important target of stress response-related changes in protein abundance. In addition to sensing environmental stresses, plant cell walls are dynamic structures that are essential for cell division, enlargement, and differentiation (Roberts, 2001; Huckelhoven, 2007). Although cell wall proteins account for only 10% of the extracellular matrix mass, they comprise several hundred different molecules with diverse cellular functions, including cell structure maintenance and responses to abiotic and biotic stresses (Somerville et al., 2004). Furthermore, it has been demonstrated in yeast cells that transient damage to the cell wall induces cell wall-related genes as part of a homeostatic response to maintain cell integrity (Garcia et al., 2004). However, it remains to be elucidated if plant cells utilize a similar mechanism to that found in yeast. Thus, crop cell wall proteomics could provide further information of the underlying mechanisms of plant responses to environmental stresses, which may prove useful for the bioengineering of more tolerant crops.

The cell wall is also thought to an important material as a biomass material, because it consists of layers primarily composed of cellulose and hemicelluloses, which are polymers of β-1,4 linked glucose and mixtures of sugars such as xylose, mannose, and galactose, respectively, being sugar sources of ethanol fermentation. Bioethanol is produced by ethanol fermentation from sugars such as glucose, fructose, and galactose. Maize and sugarcane are being used to produce bioethanol due to their high productivity and high levels of carbohydrates providing simple sugars; however, these plants are also important crops to us as food supplies. To avoid the use of food crops, the identification of alternative source species for bioethanol production is desirable. The plant cell wall consists of cellulose, hemicellulose and pectins (Demura and Ye, 2010; Pauly and Keegstra, 2010) which, upon acid or enzymatic hydrolysis can yield monosaccharides for bioethanol production. The non-edible portions of plants, such as rice straw and sugarcane pomace, and other crops including cell wall components are potential sources of sugar needed for bioethanol production, although low productivity of bioethanol is a problem to be solved. To overcome this, these plants need improvements of cell wall components to provide high quality and a large quantity of sugar sources producing high productivity of bioethanol. Thus, cell wall proteomic studies are expected to provide useful information for understanding the mechanisms controlling quality and quantity of plant cell wall components. However, such a research is very little.

Various cell wall proteins have been characterized in *Arabidopsis* (Bayer et al., 2006; Minic et al., 2007; for review Jamet et al., 2006, 2008a,b; Zhang et al., 2011), *Medicago* (Watson et al., 2004; Soares et al., 2007), chickpea (Bhushan et al., 2006), maize (Zhu et al., 2006), rice (Jung et al., 2008; Chen et al., 2009; Cho et al., 2009), and potato (Lim et al., 2012). In addition, many types of stress-associated cell wall proteins have been identified in crops, including flooding stress-induced proteins in soybean (Komatsu et al., 2010) and wheat (Kong et al., 2009), drought stress-induced proteins in rice (Pandey et al., 2010), maize (Zhu et al., 2007), and chickpea (Bhushan et al., 2007), hydrogen peroxide-induced proteins in rice (Zhou et al., 2011), and/or pathogen-induced proteins in maize or tomato (Chivasa et al., 2005; Dahal et al., 2010). Also, cell wall proteins have been studied in wounded

Abbreviations: MudPIT, multidimensional protein identification technology.

Medicago (Soares et al., 2009). Although many proteomic studies of primary cell wall have been conducted in *Arabidopsis* (Chivasa et al., 2002; Boudart et al., 2005; Jamet et al., 2006, 2008a), there have correspondingly fewer proteomic studies devoted to systematically mapping the proteins of the secondary cell wall (Millar et al., 2009). The utility of plant secondary cell wall biomass for industrial and biofuel purposes depends upon improving cellulose amount, availability, and extractability. The possibility of engineering such biomass requires much more knowledge of the genes and proteins involved in the synthesis, modification and assembly of cellulose, lignin and xylans (Millar et al., 2009).

Research on the plant cell wall has primarily focused on carbohydrate components due to their structural role and commercial value, whereas study of the complex mechanisms of stress responses mediated by cell wall proteins has remained secondary (Bhushan et al., 2007). In this review, the current methods of purification and purity test of crop cell wall proteins are presented, and the results of protein identification using gel-based and gel-free proteomic techniques are described. Furthermore, the role of cell wall proteomics of rice, wheat, maize and soybean under flooding and drought stresses is discussed.

CELL WALL PURIFICATION AND PURITY TEST

Cell wall proteins can be classified into three categories according to their interaction with other cell wall components (Jamet et al., 2006). The first is a soluble protein group, which has little or no interaction with cell wall components and thus moves freely in the extracellular space. Such proteins can be found in the culture media of cell suspensions and seedlings or can be extracted with low ionic strength buffers. The second is a group of weakly bound cell wall proteins that bind the extracellular matrix by Van der Waals forces, hydrogen bonds, and hydrophobic or ionic interactions. These proteins can be extracted from cell walls using salts. The third is a group of strongly bound cell wall proteins, and there is no efficient procedure to release these proteins from the extracellular matrix up to now. Within the past few years, there have been rapid advances in cell wall research (Jamet et al., 2008a).

The purification of plant cell walls is hampered by a number of technical difficulties such as contamination from other organelles. Thus, characterization of the cell wall proteome remains challenging and requires a combination of various treatment and analytical approaches (Watson et al., 2004). For example, mass spectrometry (MS) analyses have identified many proteins not previously believed to be extracellular, while multidimensional peptide analysis has facilitated the identification and characterization of over 250 *Arabidopsis* cell wall proteins, including new subsets of proteins (Bayer et al., 2006; Rossignol et al., 2006). With this approach, the presence of numerous extracellular proteases in the matrix was also confirmed (Bayer et al., 2006; Rossignol et al., 2006). For *Arabidopsis*, Jamet et al. (2008b) described a protocol for purifying soluble and weakly bound cell wall proteins with only low levels of contamination by intracellular proteins, thus providing a more acurate description of protein functions in the apoplast (Jamet et al., 2008b). However, extraction methods developed for *Arabidopsis* may not be useful for other species. Structural differences in the cell walls between species could have consequences to cell susceptibility to rupture by infiltration techniques, whereas compositional differences in the matrix, such as the content of homogalacturonic acid, might require the use of different extraction buffers for more effective protein release from the matrix. This implies that for the first proteomic studies of a plant species, the level of contamination of the cell wall extract must be monitored carefully and the extraction protocol adjusted accordingly to maximize its content of cell wall proteins. Here, the methods of purification and purity assay of cell wall proteome fraction from crops are explained.

RICE

Rice production in Asia has doubled since 1961 due to the breeding of new rice cultivars using intensive cultivation systems. In addition to being an important agricultural crop, rice is a useful model plant for biological research because it has a smaller genome than those of other cereals (Devos and Gale, 2000). The International Rice Genome Sequencing Project (2005) produced a map-based, high-quality sequence covering 95% of the 389-Mb rice genome. The annotation of the rice genome has progressed rapidly (The Rice Annotation Project, 2007), and a high proportion of the predicted genes are supported by full-length cDNAs (Rice Full-Length cDNA Consortium, 2003). Since completing the rice genome sequence, the challenge for the rice research community has been to identify the function and regulation of rice genes and proteins. Proteomic studies, including those directed toward the cell wall proteome, are expected to improve our understanding of these processes.

Chen et al. (2009) presented a proteomic analysis of weakly bound rice cell wall proteins extracted from rice calli with mannitol/$CaCl_2$, followed by back extraction with water-saturated phenol. The isolated cell wall proteins were evaluated for contamination with cytosolic proteins by measuring glucose-6-phosphate dehydrogenase activity, which indicated the presence of only low levels of intracellular proteins and a significant enrichment of cell wall proteins. Using multidimensional protein identification technology (MudPIT), a total of 292 proteins were identified, and bioinformatic analysis showed that 72.6% of these proteins possessed a signal peptide, indicating a total of 198 (67.8%) proteins were determined to be cell wall proteins in rice. Functional classification divided the extracellular proteins into different groups, including glycosyl hydrolases (23%), antioxidant proteins (12%), cell wall structure-related proteins (6%), metabolic pathways (9%), protein modification (4%), defense (4%), and protease inhibitors (3%). Furthermore, comparative analysis of the identified rice cell wall proteins with those of *Arabidopsis* identified 25 novel cell wall proteins that were specific to rice (Chen et al., 2009).

High-purity secreted rice proteins were obtained from the medium of a suspension of callus culture (Cho et al., 2009). To check for contamination from other cellular compartments, Western blot analyses were performed using antibodies specific for the cytoskeletal actin protein and Hsp70, and showed that the purified rice proteins were secretory. Using MudPIT, a total of 555 rice proteins were identified, and bioinformatics analysis indicated that 154 proteins (27.7%) were considered to be secreted proteins because they possessed a signal peptide. Functional classification divided the majority of identified proteins into

stress response proteins (27%), metabolic proteins (26%) and factors involved in protein modification (24%). Comparative analysis demonstrated that one third of the secreted rice proteins overlapped with those of *Arabidopsis*. Furthermore, 25 novel rice-specific secreted proteins were found (Cho et al., 2009).

Using MudPIT, the two proteomic studies described above identified more than 150 secreted or cell wall proteins, which included 25 novel rice-specific secreted or cell wall proteins. Out of rice specific secreted and/or cell wall proteins, LYM1/LYM2 and gibberellin-regulated proteins were the same between secreted and/or cell wall proteins. Two proteins, LYM1 and LYM2, which contain LysM (lysin motif) domains, a widely distributed protein motif which binds to peptidoglycans and chitin, most likely recognize the N-acetylglucosamine moiety (Cho et al., 2009). In addition, two gibberellin-regulated proteins belonging to the GASA family were increased by gibberellin and might have functions in plant development as shown in expression studies during silique development and seed germination (Chen et al., 2009). These results indicate that MudPIT is a suitable technique for cell wall proteomics.

MAIZE

Maize is an important food crop as well as a classic genetic model plant, and has been used to explore the genetic mechanism of heterosis because it exhibits high levels of phenotypic (Flint-Garcia et al., 2005), sequence (Schnable et al., 2009), transcriptional (Ma et al., 2006; Stupar and Springer, 2006), and translational variation (Hoecker et al., 2008).

Zhu et al. (2006) performed the cell wall proteomic analysis of maize. In the paper, they presented the effectiveness of a vacuum infiltration–centrifugation technique for extracting water-soluble and loosely ionically bound cell wall proteins from the root elongation zone of maize. The purity of the extracted cell wall proteins was evaluated by comparison with total soluble proteins extracted from homogenized tissue. In addition, MS analysis of proteins separated on two-dimensional gel electrophoresis (2-DE) indicated that 84% of the cell wall proteins differed from the total soluble proteins. It means that only about 16% of the proteins identified overlapped with those from the total soluble protein gel. Several lines of evidence indicated that the vacuum infiltration-centrifugation technique effectively enriched for cell wall proteins. Approximately 40% of the loosely ionically bound cell wall proteins had traditional signal peptides and 33% were predicted to be non-classical secretory proteins, whereas only 3 and 11%, respectively, of the total soluble proteins were classified into these categories. Many of the identified cell wall proteins were previously shown to be involved in cell wall metabolism and cell elongation. Primary cell walls can be divided into two types: type I cell wall were found in all flower plants expect the grass family and type II cell wall were in the grass family. The identified cell wall proteins of maize were type II cell walls, were not detected in proteomic studies that focused on type I cell walls. These newly identified proteins included endo-1,3;1,4-β-D-glucanase and α-L-arabinofuranosidase, which act on the major polysaccharides of type II cell walls (Zhu et al., 2006). Furthermore, Malate dehydrogenase and β-glucosidase, which were found in apoplastic fluid in barley and oat primary leaves, were identified in maize. Enolase, which was detected in the cell walls of *Candida albicans*, *Arabidopsis*, and *Medicago sativa,* was also identified in maize (Zhu et al., 2006). With this approach, several novel cell wall proteins of maize were found. The same purification technique was also used for the identification of cell wall proteins involved in drought stress (Zhu et al., 2007).

CHICKPEA

Legumes are valuable agricultural and commercial crops that serve as important nutrient sources for the human diet and animal feed. Legume plants typically form symbiotic relationships with both nitrogen-fixing bacteria and arbuscular mycorrhizal fungi. Many secondary metabolites in legumes have been implicated in defense against pathogen and are of particular interest as novel pharmaceuticals (Haridas et al., 2001). Legume such as soybean is one group of the most important agricultural crops in many countries. As the entire genome sequences of numerous legumes have been completed (Sato et al., 2008; Schmutz et al., 2010; Young et al., 2011; Katayose et al., 2012), proteomic approaches to identify and quantify protein molecules have been facilitated.

To better understand the function of the extracellular proteins in chickpea, Bhushan et al. (2006) examined its' extracellular proteome. In their study, chickpea seedling powder was added to homogenizing buffer (Averyhart-Fullard et al., 1988), and the cell wall fraction was recovered by differential centrifugation. Microscopy showed that the extracellular matrix fraction was free of plasma membranes and other ultrastructural cytoplasmic organelles. This finding was confirmed by performing catalase and ATPase activity assays to assess contamination by cytosol components and plasma membrane, respectively. Using 2-DE and MALDI-TOF-MS together with LC-MS/MS, 131 spots were analyzed, of which a total of 121 proteins were detected. Functions were assigned to 69 extracellular matrix proteins, whereas 43 proteins belonged to unidentified function categories, and 9 had no significant match. The proteins were classified into six different functional classes, which included metabolism (19%), cell signaling (13%), cellular transport (10%), development (9%), stress response (6%), and unidentified function.

Bhushan et al. (2006) also reported evidence for the presence of several unexpected proteins in the cell wall fraction with known biochemical activities that had never been associated with the extracellular matrix. Although many researchers have reported non-canonical proteins in cell wall fractions, this is not sufficient to unambiguously verify that the extracellular matrix fraction in this study is without any contamination as suggested by the presence of RuBisCO subunits (Wang et al., 2004). The presence of RuBisCO is most likely inevitable because it is the most abundant plant protein and has been shown in several earlier reports (Wang et al., 2004; Watson et al., 2004; Bhushan et al., 2006). This purification technique was further used for the identification of cell wall proteins of chickpea involved in drought stress (Bhushan et al., 2007).

CELL WALL PROTEOMICS UNDER STRESSES

Climate change poses tremendous global challenges for agriculture, as plant growth and productivity are strongly

influenced by environmental stresses. As the global climate changes, high temperature, flooding, and drought are becoming the most important environmental factors influencing the yield and quality of crops. Research conducted to date on changes in the cell wall proteome in response to environmental stress is summarized here (**Table 1**). Despite significant efforts, quite a large number of cell wall proteins remain unidentified up to date. To investigate the function of cell wall proteins under abiotic and biotic stresses, the remaining unidentified proteins should be identified.

RICE CELL WALL PROTEOMICS UNDER DROUGHT STRESS AND IN RESPONSE TO HYDROGEN PEROXIDE

For a better understanding of the underlying molecular mechanism of the dehydration stress response in rice seedlings, Pandey et al. (2010) used a cell wall proteomic technique. In this study, 4-week-old rice seedlings were subjected to progressive dehydration by withdrawing water, and proteomic analyses of the drought-induced changes in the extracellular matrix proteins were performed using 2-DE. Dehydration-responsive temporal changes in protein levels indicated that 192 proteins had individual intensities that differed by more than 2.5-fold from baseline levels at one or more time points during the dehydration stress. The proteomic analysis led to the identification of nearly 100 differentially regulated proteins that were predicted to be involved in a variety of cellular functions including carbohydrate metabolism, cell defense and rescue, cell wall modification, cell signaling, protein folding, and stabilization, and all of which were suggested to play key roles in the dehydration tolerance cascade (Pandey et al., 2010).

The plant apoplast is a major site for signal transduction and plays a major role in defensive responses to environmental

Table 1 | Summary of published cell wall proteomics analyses in response to environmental stress.

Plant	Stress	Proteomics techniques	IP[a]	Major findings	References
Rice	Dehydration	2-DE LC-TOF MS	94	Dehydration-responsive proteins mainly involved in a variety of functions, including carbohydrate metabolism, cell defense and rescue, cell wall modification, cell signaling and molecular chaperones	Pandey et al., 2010
Rice	Hydrogen peroxide	2-DE MALDI TOF/TOF MS	54	These proteins identified are involved in a variety of processes, including redox homeostasis, cell wall modification, signal transduction, cell defense and carbohydrate metabolism, indicating a complex regulative network in the apoplast of seedling roots under H_2O_2 stress	Zhou et al., 2011
Maize	Water deficit	2-DE Q-TOF MS	152	Water stress-responsive proteins were identified and categorized, into five groups; apoplastic ROS level increases in the apical region of the elongation zone of water stressed maize roots and hence provides novel information about the complex mechanisms regulating root growth under water stress	Zhu et al., 2007
Maize	Pathogen	2-DE MALDI-TOF MS	12	Secretion of a new class of putative enzyme inhibitor, the apparent recruitment of classical cytosolic proteins into the cell wall, and the change in phosphorylation status of extracellular matrix proteins	Chivasa et al., 2005
Chickpea	Dehydration	2-DE LC-TOF MS	134	More than hundred extracellular matrix proteins with a variety of cellular functions which are cell wall modification, signal transduction, metabolism, and cell defense and rescue, play crucial roles in dehydration stress sensing and tolerance mechanism	Bhushan et al., 2007
Soybean	Flooding	2-DE MALDI-TOF MS and nanoLC MS/MS	16	Two lipoxygenases, germin-like protein precursors, glycoprotein precursors, super oxide dismutase were decreased, and then lignification was suppressed under flooding stress	Komatsu et al., 2010
Wheat	Flooding	2-DE nanoLC MS/MS Gel-free nanoLC MS/MS	18 15	Four proteins were upregulated and 14 proteins were down-regulated by flooding stress. Both gel-based proteomic system and LC-MS/MS-based proteomic system under flooding stress, most were disease/defense proteins	Kong et al., 2009

[a] *IP, Number of identified proteins.*

stresses. Zhou et al. (2011) reported their comparative proteomic analysis of the root apoplasts of rice seedlings in response to hydrogen peroxide. Two week-old rice seedlings were treated with low concentrations of hydrogen peroxide and a modified vacuum infiltration method was used to extract apoplastic proteins from the roots. A 2-DE analysis revealed 58 differentially expressed protein spots under low hydrogen peroxide conditions. Of these, 54 protein spots were identified by MS as matches to 35 different proteins including known and novel hydrogen peroxide-responsive proteins. Almost all of these identities (98%) were indeed apoplast proteins confirmed either by previous experiments or through publicly available prediction programs. These identified proteins were involved in a variety of processes, including redox homeostasis, cell wall modification, signal transduction, cell defense, and carbohydrate metabolism, indicating a complex regulative network in the apoplast of seedling roots under hydrogen peroxide stress (Zhou et al., 2011). They indicated that the identified proteins might work cooperatively to establish a complex network of apoplast response to exogenous hydrogen peroxide in the rice seedling root, and depict the strategies of the root apoplast to oxidative challenge.

MAIZE CELL WALL PROTEOMICS UNDER DROUGHT STRESS AND IN RESPONSE TO PATHOGENS

To gain a comprehensive understanding of how cell wall protein composition changes in association with the differential growth responses to water deficit in different regions of the elongation zone, Zhu et al. (2007) used a proteomic approach for water soluble and loosely ionically bound cell wall proteins. As different regions, 20 mm of each root was sectioned into four regions (distances are from the root cap junction): R1, 0–3 mm plus the root cap; R2, 3–7 mm; R3, 7–12 mm; R4, 12–20 mm. The analyses demonstrated region-specific changes in the protein profiles of control and water-stressed roots. In total, 152 water stress-responsive proteins were identified and categorized into five functional groups: metabolism of cell wall reactive oxygen species, defense and detoxification, hydrolases, carbohydrate metabolism and other/unknown functions. In particular, the changes in protein abundance related to reactive oxygen species metabolism predicted an increase in apoplastic reactive oxygen species production in the apical region of the elongation zone of water-stressed roots. This was verified by quantification of hydrogen peroxide content in extracted apoplastic fluid and by *in situ* imaging of apoplastic reactive oxygen species levels. This response could contribute directly to the enhancement of wall loosening in this region (Zhu et al., 2007). This large-scale proteomic analysis provided novel insights into the complexity of mechanisms that regulated root growth under water deficit conditions and highlighted the spatial differences in cell wall protein composition in the root elongation zone.

Chivasa et al. (2002) have reported their proteomic analysis of elicited maize suspension cultures. Using phosphotyrosine antibodies raised against synthetic phosphotyrosine peptides, they identified a number of phosphotyrosine protein spots secreted into the growth medium of cell cultures. Elicitor treatment of cell cultures-induced a rapid change in the phosphorylation status of extracellular peroxidases, the apparent disappearance of a putative extracellular β-*N*-acetylglucosamonidase, and accumulation of a putative secreted xylanase inhibitor protein. The onset of the defense response was accompanied by accumulation of glyceraldehyde-3-phosphate dehydrogenase and a fragment of a putative heat shock protein. Several distinct spots of both proteins, which preferentially accumulated in cell wall protein fractions, were identified. The novel observations of the secretion of a new class of putative enzyme inhibitor, the apparent recruitment of classical cytosolic proteins into the cell wall, and the change in phosphorylation status of extracellular matrix proteins, suggested that the extracellular matrix plays a complex role in defense (Chivasa et al., 2002). Taken together, it is suggested that cell wall proteins may have an important role in stress signal transduction.

CHICKPEA CELL WALL PROTEOMICS UNDER DROUGHT STRESS

Comprehensive extracellular matrix proteins analysis of chickpea under dehydration stress has been performed using comparative proteomics. This approach led to the identification of 134 differentially accumulated proteins, which include both predicted and novel dehydration-responsive proteins. This comparative proteomic study demonstrates that more than 100 extracellular matrix proteins with a variety of cellular functions, such as cell wall modification, signal transduction, metabolism, and cell defense and rescue, play crucial roles in dehydration stress sensing and tolerance mechanisms in chickpea (Bhushan et al., 2007).

SOYBEAN CELL WALL PROTEOMICS UNDER FLOODING STRESS

To investigate the cell wall function of soybean under flooding stress, cell wall proteomics was performed using cell wall proteins purified from 2-day-old soybean plants treated with flooding for 2 days. The purity of the cell wall protein extract was confirmed by measuring the activity of glucose-6-phosphate dehydrogenase, a cytoplasmic marker enzyme. Using 2-DE and MS, it was found that 16 out of 204 cell wall proteins responded to flooding stress. Of these, two lipoxygenases, four germin-like protein precursors, three stem glycoprotein precursors and one Cu-Zn superoxide dismutase were found to be present at lower levels than those found in the control plants. These changed proteins indicated that the roots and hypocotyls of soybean under flooding stress suppressed lignification through decreasing these cell wall proteins by down-regulation of reactive oxygen species and inhibition of jasmonate biosynthesis (Komatsu et al., 2010).

WHEAT CELL WALL PROTEOMICS UNDER FLOODING STRESS

A procedure for extracting and purifying cell wall proteins was adopted for wheat seedling roots, and the purity of the cell wall protein extract was confirmed by measuring the activity of glucose-6-phosphate dehydrogenase. To identify flooding-stress responsive proteins in the wheat cell wall, gel-based and LC-MS/MS-based proteomic techniques were applied. A total of 18 and 15 proteins were shown to accumulate in response to flooding by the former and latter proteomic techniques, respectively. Among the proteins detected at lower levels in

response to flooding, most were related to the glycolysis pathway and cell wall structure and modification. In contrast, the cell wall proteins of highest abundance after flooding treatment belonged to the category of defense and disease-response proteins. Among the identified proteins, only methionine synthase, β-1,3-glucanases and β-glucosidase were consistently detected by both techniques. The decrease of these three proteins suggested that wheat seedlings respond to flooding stress by restricting cell growth to avoid energy consumption; by coordinating methionine assimilation and cell wall hydrolysis, cell wall proteins played critical roles in flooding responsiveness (Kong et al., 2009).

DIFFERENCES IN STRESS RESPONSES AMONG CROPS

DROUGHT STRESS

Water deficit and dehydration are the most crucial environmental factors that limit the productivity and geographic distribution of many crop species. Evidence suggests that dehydration-responsive changes in protein levels may lead to cellular adaptation against water-deficit conditions. Plants exposed to dehydration mostly rely on the protection of cellular integrity to prevent mechanical damage by changing the composition of the cell wall (Vicre et al., 2004; Moore et al., 2006).

Using cell wall proteomic techniques, the mechanisms of drought response in the cereals rice (Pandey et al., 2010) and maize (Zhu et al., 2007), and in the legume chickpea (Bhushan et al., 2007) were analyzed. In the case of chickpea and rice, cell wall proteins were purified from shoots, whereas root tissue was used for the analysis of maize. Although many proteins unique to each crop were identified, proteins associated with cell defense and rescue, ascorbate peroxidases and glyoxalase, respectively, and one protein involved in carbohydrate metabolism, fructose-bisphosphate aldolase, were common in all three crops. It has been proposed that functionally important proteins are more evolutionarily conserved than less vital proteins.

Pandey et al. (2010) speculated that the preponderance of differential protein networks among crop species may be attributable to the evolutionary species-specific dynamics of the cell wall proteomics because the protein expression profile is a reflection of the cellular environment and ecological niche of the corresponding organism. A closer look at their results indicated that chickpea and rice share many common proteins in shoot. Notably, methyltransferases and the chaperone class of proteins were absent in maize exposed to drought stress, suggesting the type of sampled tissue can influence the proteomics. Interestingly, rice and maize showed the highest degree of commonality, possibly because both species belong to the Poaceae family. Together, these data provide evidence for molecular diversity among different plant species as opposed to commonality with respect to protein profiles at the level of the organism and/or tissue. The higher percentage of crop-specific dehydration-responsive proteins identified in this study signifies that examining the cell wall proteomics of different crops, particularly the major lineages of higher plants, at different tissue levels is needed to better understand the critical role of the cell wall in drought stress tolerance (Pandey et al., 2010).

FLOODING STRESS

Flooding caused by heavy or continuous rainfall in an area with poorly drained soil can be a devastating environmental stressor, as many types of crops cannot tolerate flooding. Soybean and other crops are particularly susceptible to stress arising from flooding (Komatsu et al., 2012). Previous studies showed that genes encoding transcription factors (Liu et al., 2005), signal transduction components (Baxter-Burrell et al., 2002), non-symbiotic hemoglobin (Dordas et al., 2004), ethylene signaling components (Reggiani, 2006), as well as genes involved in nitrogen metabolism (Mattana et al., 1994) and cell wall loosening (Saab and Sachs, 1996), were up-regulated under low-oxygen condition. Since flooding produces hypoxia in plant tissues, these genes may be regulated by flooding.

The primary cell wall plays in regulating extension growth, cell adhesion, and cell morphology (Cosgrove, 2001). Pereira et al. (2011) reported that the primary cell wall is essentially a hydrated matrix and sensitive to global changes in the water content of the plant by altering its biophysical properties. The partial adaptation to flooding stress might be required a compensatory modification of the extracellular matrix. The mechanisms underlying the flooding response in soybean (Komatsu et al., 2010) and wheat (Kong et al., 2009) have been analyzed using cell wall proteomic techniques. Although soybean is a dicot and wheat is a monocot, the roots of both crops were affected by flooding stress. In fact, the physiological and morphological response to flooding stress was to suppress root growth (Kong et al., 2009; Komatsu et al., 2012). In soybean and wheat, the levels of methionine synthase were commonly reduced under flooding stress, indicating that both crops restrict cell growth to avoid energy consumption by coordinating methionine assimilation and cell wall hydrolysis, which is consistent with the observed growth suppression of wheat seedlings by flooding stress. However, many kinds of proteins differentially accumulated in soybean and wheat under flooding stress. In roots and hypocotyls of soybean, lignification was suppressed by the down regulation of reactive oxygen species and jasmonate biosynthesis. On the other hands, in the case of wheat, the reduction of cell wall hydrolytic enzymes in response to flooding stress might allow not only cell wall polysaccharides to be used as carbohydrate sources, but also restricts cell elongation, thus better preparing wheat seedlings for long periods of submergence.

FUTURE PERSPECTIVE

Plant cell walls are involved in both maintenance of the cell structure and resistance against environmental stresses, because the location of the cell wall places it in direct contact with environmental factors. Indeed, flooding stress leads to morphological suppression of roots, and in severe cases, the root tissue is destroyed, as was demonstrated in soybean (Kong et al., 2009; Komatsu et al., 2012). When plant pathogens and insects attack plants, the cell wall represents the first line of host defense and resistant to damage. Thus, examination of the structural and functional mechanisms underlying stress response in the cell wall is expected to aid in the development of stress-tolerant crops. Based on the data of proteomic analysis of cell wall, the future

of abiotic stress research should focus on targeting multiple gene regulation, because the change on a single gene expression may not have enough results on the field in question, since genetic regulation is complex. It may be possible to generate stress-tolerant crops using mutants with loss/gain of function in cell wall stress signal transduction mechanisms. Furthermore, the identified factors may aid in the genetic screening for crop tolerance against other stresses. It could be that proteomics is a useful tool to identify targets/markers proteins possibly involved in stress response of crops and to compare proteins identified in rice, wheat, maize, and soybean upon flooding and drought stresses.

REFERENCES

Averyhart-Fullard, V., Datta, K., and Marcus, A. (1988). A hydroxyproline-rich protein in the soybean cell wall. *Proc. Natl. Acad. Sci. U.S.A.* 85, 1082–1085.

Baxter-Burrell, A., Yang, Z., Springer, P. S., and Bailey-Serres, J. (2002). RopGAP4-dependent Rop GTPase rheostat control of *Arabidopsis* oxygen deprivation tolerance. *Science* 296, 2026–2028.

Bayer, E. M., Bottrill, A. R., Walshaw, J., Vigouroux, M., Naldrett, M. J., Thomas, C. L., et al. (2006). *Arabidopsis* cell wall proteome defined using multidimensional protein identification technology. *Proteomics* 6, 301–311.

Bhushan, D., Pandey, A., Chattopadhyay, A., Choudhary, M. K., Chakraborty, S., Datta, A., et al. (2006). Extracellular matrix proteome of chickpea (*Cicer arietinum* L.) illustrates pathway abundance, novel protein functions and evolutionary perspect. *J. Proteome Res.* 5, 1711–1720.

Bhushan, D., Pandey, A., Choudhary, M. K., Datta, A., Chakraborty, S., and Chakraborty, N. (2007). Comparative proteomics analysis of differentially expressed proteins in chickpea extracellular matrix during dehydration stress. *Mol. Cell. Proteomics* 6, 1868–1884.

Boudart, G., Jamet, E., Rossignol, M., Lafitte, C., Borderies, G., Jauneau, A., et al. (2005). Cell wall proteins in apoplastic fluids of *Arabidopsis thaliana* rosettes: identification by mass spectrometry and bioinformatics. *Proteomics* 5, 212–221.

Chen, X. Y., Kim, S. T., Cho, W. K., Rim, Y., Kim, S., Kim, S. W., et al. (2009). Proteomics of weakly bound cell wall proteins in rice calli. *J. Plant Physiol.* 166, 675–685.

Chivasa, S., Ndimba, B. K., Simon, W. J., Robertson, D., Yu, X. L., Knox, J. P., et al. (2002). Proteomic analysis of the *Arabidopsis thaliana* cell wall. *Electrophoresis* 23, 1754–1765.

Chivasa, S., Simon, W. J., Yu, X.-L., Yalpani, N., and Slabas, A. R. (2005). Pathogen elicitor-induced changes in the maize extracellular matrix proteome. *Proteomics* 5, 4894–4904.

Cho, W. K., Chen, X. Y., Chu, H., Rim, Y., Kim, S., Kim, S. T., et al. (2009). Proteomic analysis of the secretome of rice calli. *Physiol. Plant.* 135, 331–341.

Cosgrove, D. J. (2001). Wall structure and wall loosening. A look backwards and forwards. *Plant Physiol.* 125, 131–134.

Dahal, D., Pich, A., Braun, H. P., and Wydra, K. (2010). Analysis of cell wall proteins regulated in stem of susceptible and resistant tomato species after inoculation with *Ralstonia solanacearum*: a proteomic approach. *Plant Mol. Biol.* 73, 643–658.

Demura, T., and Ye, Z.-H. (2010). Regulation of plant biomass production. *Curr. Opin. Plant Biol.* 13, 299–304.

Devos, M. K., and Gale, D. M. (2000). Genome relationships: the grass model in current research. *Plant Cell* 12, 637–643.

Dordas, C., Hasinoff, B. B., Rivoal, J., and Hill, R. D. (2004). Class-1 hemoglobins, nitrate and NO levels in anoxic maize cell-suspension cultures. *Planta* 219, 66–72.

Flint-Garcia, S. A., Thuillet, A. C., Yu, J., and Pressoir, G. (2005). Maize association population: a high-resolution platform for quantitative trait locus dissection. *Plant J.* 44, 1054–1064.

Garcia, R., Bermejo, C., Grau, C., Perez, R., Rodriguez-Pena, J. M., Francois, J., et al. (2004). The global transcriptional response to transient cell wall damage in *Saccharomyces cerevisiae* and its regulation by the cell integrity signaling pathway. *J. Biol. Chem.* 279, 15183–15195.

Haridas, V., Higuchi, M., Jaytilake, G. S., Baily, D., Mujoo, K., Blake, M. E., et al. (2001). Avicins: triterpenoid saponins from *Acacia victoriae* (Bentham) induce apoptosis by mitochondrial perturbation. *Proc. Natl. Acad. Sci. U.S.A.* 98, 5821–5826.

Hoecker, N., Lamkemeyer, T., Sarholz, B., Paschold, A., Fladerer, C., Madlung, J., et al. (2008). Analysis of nonadditive protein accumulation in young primary roots of a maize (Zea mays L.) F1-hybrid compared to its parental inbred lines. *Proteomics* 8, 3882–3894.

Huckelhoven, R. (2007). Cell wall-associated mechanisms of disease resistance and susceptibility. *Annu. Rev. Phytopathol.* 45, 101–127.

Jamet, E., Albenne, C., Boudart, G., Irshad, M., Canut, H., and Pont-Lezica, R. (2008a). Recent advances in plant cell wall proteomics. *Proteomics* 8, 893–908.

Jamet, E., Boudart, G., Borderies, G., Charmont, S., Lafitte, C., Rossignol, M., et al. (2008b). Isolation of plant cell wall proteins. *Methods Mol. Biol.* 425, 187–201.

Jamet, E., Canut, H., Boudart, G., and Pont-Lezica, R. F. (2006). Cell wall proteins: a new insight through proteomics. *Trends Plant Sci.* 11, 33–39.

Jung, Y. H., Jeong, S. H., Kim, S. H., Singh, R., Lee, J. E., Cho, Y. S., et al. (2008). Systematic secretome analyses of rice leaf and seed callus suspension-cultured cells: workflow development and establishment of high-density two-dimensional gel reference maps. *J. Proteome Res.* 7, 5187–5210.

Katayose, Y., Kanamori, H., Shimomura, M., Ohyanagi, H., Ikawa, H., Minami, H., et al. (2012). DaizuBase, an integrated soybean genome database including BAC-based physical maps. *Breed. Sci.* 61, 661–664.

Komatsu, S., Hiraga, S., and Yanagawa, Y. (2012). Proteomics techniques for the development of flood tolerant crops. *J. Proteome Res.* 11, 68–78.

Komatsu, S., Kobayashi, Y., Nishizawa, K., Nanjo, Y., and Furukawa, K. (2010). Comparative proteomics analysis of differentially expressed proteins in soybean cell wall during flooding stress. *Amino Acids* 39, 1435–1449.

Kong, F. J., Oyanagi, A., and Komatsu, S. (2009). Cell wall proteome of wheat roots under flooding stress using gel-based and LC MS/MS-based proteomics approaches. *Biochim. Biophys. Acta* 1804, 124–136.

Lim, S., Chisholm, K., Coffin, R. H., Peters, R. D., Al-Mughrabi, K. I., Wang-Pruski, G., et al. (2012). Protein profiling in potato (*Solanum tuberosum* L.) leaf tissues by differential centrifugation. *J. Proteome Res.* 11, 2594–2601.

Liu, F., Vantoai, T., Moy, L., Bock, G., Linford, L. D., and Quackenbush, J. (2005). Global transcription profiling reveals novel insights into hypoxic response in *Arabidopsis*. *Plant Physiol.* 137, 1115–1129.

Ma, J., Morrow, D. J., Fernandes, J., and Walbot, V. (2006). Comparative profiling of the sense and antisense transcriptome of maize lines. *Genome Biol.* 7:R22. doi: 10.1186/gb-2006-7-3-r22

Mattana, M., Coraggio, I., Bertani, A., and Reggiani, R. (1994). Expression of the enzyme of nitrate reduction during the anaerobic germination of rice. *Plant Physiol.* 106, 1605–1608.

Millar, D. J., Whitelegge, J. P., Bindschedler, L. V., Rayon, C., Boudet, A. M., Rossignol, M., et al. (2009). The cell wall and secretory proteome of a tobacco cell line synthesizing secondary wall. *Proteomics* 9, 2355–2372.

Minic, Z., Jamet, E., Négroni, L., Arsene der Garabedian, P., Zivy, M., and Jouanin, L. (2007). A sub-proteome of *Arabidopsis thaliana* mature stems trapped on Concanavalin A is enriched in cell wall glycoside hydrolases. *J. Exp. Bot.* 58, 2503–2512.

Moore, J. P., Nguema-Ona, E., Chevalier, L., Lindsey, G. G., Brandt, W. F., Lerouge, P., et al. (2006). Response of the leaf cell wall to desiccation in the resurrection plant *Myrothamnus flabellifolius*. *Plant Physiol.* 141, 651–662.

Pandey, A., Rajamani, U., Verma, J., Subba, P., Chakraborty, N., Datta, A., et al. (2010). Identification of extracellular matrix proteins of rice (*Oryza sativa* L.) involved in dehydration-responsive network: a proteomic approach. *J. Proteome Res.* 9, 3443–3464.

Pauly, M., and Keegstra, K. (2010). Plant cell wall polymers as precursors for biofuels. *Curr. Opin. Plant Biol.* 13, 305–312.

Pereira, C. S., Ribeiro, J. M., Vatulescu, A. D., Findlay, K., MacDougall, A. J., and Jackson, P. A. (2011). Extensin network formation in Vitis vinifera callus cells is an essential and

causal event in rapid and H(2)O(2)-induced reduction in primary cell wall hydration. *BMC Plant Biol.* 11:106. doi: 10.1186/1471-2229-11-106

Reggiani, R. (2006). A role for ethylene in low-oxygen signaling in rice roots. *Amino Acids* 30, 299–301.

Rice Full-Length cDNA Consortium. (2003). Collection mapping, and annotation of over 28000 cDNA clones from japonica rice. *Science* 301, 376–379.

Roberts, K. (2001). How the cell wall acquired a cellular context. *Plant Physiol.* 125, 127–130.

Rossignol, M., Peltier, J.-B., Mock, H.-P., Matros, A., Maldonado, A. M., and Jorrin, J. V. (2006). Plant proteome analysis: a 2004–2006 update. *Proteomics* 6, 5529–5548.

Saab, I. N., and Sachs, M. M. (1996). A flooding-induced xyloglucan endotransglycosylase homolog in maize is responsive to ethylene and associated with aerenchyma. *Plant Physiol.* 112, 385–391.

Sato, S., Nakamura, Y., Kaneko, T., Asamizu, E., Kato, T., Nakao, M., et al. (2008). Genome structure of the legume, *Lotus japonicus*. *DNA Res.* 15, 227–239.

Schmutz, J., Cannon, S. B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., et al. (2010). Genome sequence of the palaeopolyploid soybean. *Nature* 463, 178–183.

Schnable, P. S., Ware, D., Fulton, R. S., Stein, J. C., Wei, F., Pasternak, S., et al. (2009). The B73 maize genome: complexity, diversity, and dynamics. *Science* 326, 1112–1115.

Soares, N. C., Francisco, R., Ricardo, C. P., and Jackson, P. A. (2007). Proteomics of ionically bound and soluble extracellular proteins in *Medicago truncatula* leaves. *Proteomics* 7, 2070–2082.

Soares, N. C., Francisco, R., Vielba, J. M., Ricardo, C. P., and Jackson, P. A. (2009). Associating wound-related changes in the apoplast proteome of Medicago with early steps in the ROS signal-transduction pathway. *J. Proteome Res.* 8, 2298–2309.

Somerville, C., Bauer, S., Brininstool, G., Facette, M., Hamann, T., Milne, J., et al. (2004). Toward a systems approach to understanding plant cell walls. *Science* 306, 2206–2211.

Stupar, R. M., and Springer, N. M. (2006). Cis-transcripioanl variation in maize inbred lines B73 and Mo17 leads to additive expression patterns in the F1 hybrid. *Genetics* 173, 2199–2210.

The International Rice Genome Sequencing Project. (2005). The map-based sequence of the rice genome. *Nature* 436, 793–800.

The Rice Annotation Project. (2007). Crated genome annotation of *Oryza sativa* ssp. Japonica and comparative genome analysis with *Arabidopsis thaliana*. *Genome Res.* 17, 175–183.

Vicre, M., Lerouxel, O., Farrant, J., Lerouge, P., and Driouich, A. (2004). Composition and desiccation-induced alterations of the cell wall in the resurrection plant *Craterostigma wilmsii*. *Physiol. Plant.* 120, 229–239.

Wang, S. B., Hu, Q., Sommerfeld, M., and Chen, F. (2004). Cell wall proteomics of the green alga *Haematococcus pluvialis* (Chlorophyceae). *Proteomics* 4, 692–708.

Watson, B. S., Lei, Z., Dixon, R. A., and Sumner, L. W. (2004). Proteomics of *Medicago sativa* cell walls. *Phytochemistry* 65, 1709–1720.

Young, N. D., Debellé, F., Oldroyd, G. E., Geurts, R., Cannon, S. B., Udvardi, M. K., et al. (2011). The *Medicago* genome provides insight into the evolution of rhizobial symbioses. *Nature* 480, 520–524.

Zhang, Y., Giboulot, A., Zivy, M., Valot, B., Jamet, E., and Albenne, C. (2011). Combining various strategies to increase the coverage of the plant cell wall glycoproteome. *Phytochemistry* 72, 1109–1123.

Zhou, L., Bokhari, S. A., Dong, C. J., and Liu, J. Y. (2011). Comparative proteomics analysis of the root apoplasts of rice seedlings in response to hydrogen peroxide. *PLoS ONE* 6:e16723. doi: 10.1371/journal.pone.0016723

Zhu, J., Alvarez, S., Marsh, E. L., Lenoble, M. E., Cho, I. J., Sivaguru, M., et al. (2007). Cell wall proteome in the maize primary root elongation zone. II. Region-specific changes in water soluble and lightly ionically bound proteins under water deficit. *Plant Physiol.* 145, 1533–1548.

Zhu, J., Chen, S., Alvarez, S., Asirvatham, V. S., Schachtman, D. P., Wu, Y., et al. (2006). Cell wall proteome in the maize primary root elongation zone. I. Extraction and identification of water-soluble and lightly ionically bound proteins. *Plant Physiol.* 140, 311–325.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 13 November 2012; paper pending published: 03 December 2012; accepted: 23 January 2013; published online: 07 February 2013.

Citation: Komatsu S and Yanagawa Y (2013) Cell wall proteomics of crops. Front. Plant Sci. ***4****:17. doi: 10.3389/fpls.2013.00017*

This article was submitted to Frontiers in Plant Proteomics, a specialty of Frontiers in Plant Science.

Copyright © 2013 Komatsu and Yanagawa. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.

frontiers in
PLANT SCIENCE

REVIEW ARTICLE
published: 29 May 2013
doi: 10.3389/fpls.2013.00159

Progress toward the tomato fruit cell wall proteome

Eliel Ruiz-May and Jocelyn K. C. Rose*

Department of Plant Biology, Cornell University, Ithaca, NY, USA

Edited by:
Harvey Millar, The University of Western Australia, Australia

Reviewed by:
Ning Li, The Hong Kong University of Science and Technology, China
Jose M. M. Palma, Estacion Experimental del Zaidin, Consejo Superior de Investigaciones Cientificas, Spain

****Correspondence:***
Jocelyn K. C. Rose, Department of Plant Biology, Cornell University, 412 Mann Library Building, Ithaca, NY 14853, USA
e-mail: jr286@cornell.edu

The plant cell wall (CW) compartment, or apoplast, is host to a highly dynamic proteome, comprising large numbers of both enzymatic and structural proteins. This reflects its importance as the interface between adjacent cells and the external environment, the presence of numerous extracellular metabolic and signaling pathways, and the complex nature of wall structural assembly and remodeling during cell growth and differentiation. Tomato fruit ontogeny, with its distinct phases of rapid growth and ripening, provides a valuable experimental model system for CW proteomic studies, in that it involves substantial wall assembly, remodeling, and coordinated disassembly. Moreover, diverse populations of secreted proteins must be deployed to resist microbial infection and protect against abiotic stresses. Tomato fruits also provide substantial amounts of biological material, which is a significant advantage for many types of biochemical analyses, and facilitates the detection of lower abundance proteins. In this review, we describe a variety of orthogonal techniques that have been applied to identify CW localized proteins from tomato fruit, including approaches that: target the proteome of the CW and the overlying cuticle; functional "secretome" screens; lectin affinity chromatography; and computational analyses to predict proteins that enter the secretory pathway. Each has its merits and limitations, but collectively they are providing important insights into CW proteome composition and dynamics, as well as some potentially controversial issues, such as the prevalence of non-canonical protein secretion.

Keywords: tomato, fruit, cell wall, proteomics, secretion

INTRODUCTION

Tomato (*Solanum lycopersicum*), one of the world's most important horticultural crops, is recognized as the pre-eminent experimental model to study fleshy fruit development, physiology, and pathology (Klee and Giovannoni, 2011; Tomato Genome Consortium, 2012; Seymour et al., 2013). Particular attention has been paid to attributes of tomato fruit that are associated with consumer desirability, such as color, flavor, aroma, and texture, and considerable progress has been made in understanding the molecular processes that underlie these traits. In some cases these are directly and clearly associated with specific biochemical or regulatory pathways that are now well understood, such as the synthesis of pigments, or the gaseous hormone ethylene. However, in other instances they relate broadly to subcellular features whose role in fruit development is highly complex and less well defined. An example of the latter is the intimate, but still poorly understood relationship between fruit texture and the metabolism of the cell wall (CW). This is reflected in an extensive literature, spanning many decades, describing the putative roles of numerous, functionally diverse families of CW resident (i.e., apoplastic, or secreted) proteins and their cognate genes in ripening-related textural changes (Brummell, 2006; Prasanna et al., 2007; Vicente et al., 2007; Payasi et al., 2009; Palma et al., 2011). Progress to date in experimentally addressing the enzymatic basis of CW modification and the relationship with softening has typically been piecemeal, targeting specific activities or individual genes and related proteins. Perhaps the best studied example of such is the pectin degrading enzyme polygalacturonase (PG), whose expression was suppressed in the first commercially available genetically engineered whole food, the Flavr Savr tomato (Smith et al., 1988; Kramer and Redenbaugh, 1994). Since then, other genes encoding CW modifying proteins have been targeted in transgenic tomato fruits in an effort to prevent over-softening and textural deterioration, although mostly without success (Brummell, 2006; Vicente et al., 2007). Such experiments have provided many insights into CW metabolism, but an important outcome has also been the realization that CW disassembly is the consequence of the synergistic actions of many proteins, and that a significant understanding of the dynamics of wall architecture and the mechanistic basis of softening will require a more complete compendium of the CW proteome. This has spurred efforts to study the CW proteome, or "secretome," of tomato fruit during ripening, as well as during cuticle and CW biosynthesis (Faurobert et al., 2007; Yeats et al., 2010, 2012a,b; Catalá et al., 2011; Palma et al., 2011).

As with any species, genomic and transcriptomic studies of tomato provide an invaluable, if not essential platform for equivalent proteomics analysis and the recent publication of the tomato genome sequence includes more than 700 gene models annotated with CW-related functions (Tomato Genome Consortium, 2012). Transcript analysis of tomato fruit development and

Abbreviations: CW, cell wall; GH, glycosyl hydrolase; GT, glycosyltransferase; IG, immature green; MG, mature green; PG, polygalacturonase; PMEI, pectin methylesterase inhibitor; PR, pathogenesis related; PTM, post-translational modifications; *rin*, *ripening inhibitor*; RR, red ripe; SP, signal peptide.

ripening showed the differential expression of more than 50 genes associated with CW modification (Tomato Genome Consortium, 2012), while a more detailed analysis of the cell/tissue type-specific transcriptomes of tomato fruit at the maximal stage of growth revealed 253 glycosyltransferases (GTs) and 293 glycosyl hydrolases (GHs) related to CW biogenesis and disassembly, respectively (Matas et al., 2011). Such information, together with a wealth of analytical tools and bioinformatic resources, has laid the foundation for cataloging the tomato fruit CW proteome in a more comprehensive fashion. This will include as assessment of alternate splicing of transcripts to reveal the multiple proteins variants that can result from single genes, as well as the broad range of post-translational modifications (PTMs).

THE CHALLENGES OF IDENTIFYING FRUIT CW PROTEINS

The pioneering study of the plant CW proteome involved an analysis of cell suspension cultures derived from several plant species, including tomato, that were washed sequentially with buffers of different ionic strengths in order to isolate proteins that were designated as soluble, weakly or strongly bound to the CW (Robertson et al., 1997). Since then, different approaches have been used to characterize the CW proteomes of different tissues from a broad range of plants (Blee et al., 2001; Chivasa et al., 2002; Borderies et al., 2003; Watson et al., 2004; Bayer et al., 2006; Jamet et al., 2006, 2008; Zhu et al., 2006, 2007; Negri et al., 2008; Albenne et al., 2009; Chen et al., 2009; Cho et al., 2009; Millar et al., 2009; Zhang et al., 2010). However, as previously described (Lee et al., 2004; Rose and Lee, 2010), characterization of the plant CW proteome is technically challenging compared to that of other subcellular fractions. Firstly, the "cell wall," in the context discussed here, is not bound by a distinct membrane that can facilitate isolation, but rather corresponds to the apoplastic continuum and associated extracellular matrix that extends throughout the plant. Attempts to isolate a comprehensive set of wall localized proteins must therefore contend with the conflicting needs to isolate proteins that can be extremely tightly linked to the wall matrix, and thus require harsh treatments to liberate them, as well as proteins that are mobile in the apoplast and are readily lost upon tissue disruption. Any cellular breakage will also immediately result in severe contamination of the protein fraction with cytoplasmic proteins and possibly also those from intracellular organelles (Lee et al., 2004; Feiz et al., 2006). Isolation strategies can therefore be designated as "disruptive," where researchers must address the challenges of identifying proteins resulting from the inevitable contamination; or "non-disruptive," where techniques such as vacuum infiltration typically yield a very small fraction of the total CW proteome, and only those protein species that are weakly affiliated with the CW. We also note that even non-disruptive approaches will almost certainly cause some cellular lysis.

In one sense, fleshy fruits provide a "worst case scenario," as not only are their walls often particularly rich in highly charged anionic pectin polymers that form gels and confound protein isolation, but also the degree of charge and physicochemical properties of the wall matrix change during ripening. This means that the ease of extraction of a particular protein may change substantially at different developmental stages. Thus, a protein may be present and detected in a soluble fraction derived from one developmental stage, but "vanish" from a similar extract from a subsequent developmental stage as it is more tightly bound to the CW and thus less easily extracted. The reverse may also be true and we have observed that this can be a chronic problem for some proteins. In addition, most CW proteins are glycosylated, sometimes to a large degree, which also affects ease of isolation and downstream analysis. This makes accurate quantitative analysis extremely difficult, if not impossible in some cases. We contend that CW proteome studies should be considered as analyses of the "extractome," rather than the true proteome.

It is important to bear in mind that there is no one perfect approach and that an extensive catalog of the CW proteome requires multiple orthogonal strategies, including techniques that enrich for wall proteins and bioinformatic analyses (Lee et al., 2004; Feiz et al., 2006; Rose and Lee, 2010; Ruiz-May et al., 2012b). Assessment of protein localization *in silico*, based on the predicted presence or absence of subcellular targeting sequences, can provide a valuable tool for biologists. Indeed, current algorithms are generally highly effective; however, they are not perfect predictors (Rose and Lee, 2010) and for a high confidence determination of true wall localization, confirmatory analysis, such as fluorescence protein fusion localization or immunolocalization is essential. It is notable that, to our knowledge, no CW proteome profiling study to date has followed up the identification of a potentially "non-classically secreted CW protein" with such a confirmatory analysis. Until this is done such reports should be treated with great caution.

INSIGHTS INTO THE TOMATO FRUIT PROTEOME

To date, there have been few systematic studies of the tomato fruit CW proteome compared to those that have targeted *Arabidopsis thaliana* (Nikolovski et al., 2012; Parsons et al., 2012; Zielinska et al., 2012). However, several reports have focused on specific aspects of tomato fruit biology, as summarized below.

Perhaps the most important question in arena of tomato fruit CW proteomics is the relationship between CW resident proteins and the complex textural changes that occur during ripening, which are loosely referred to as "softening" (Vicente et al., 2007; Seymour et al., 2013). One obvious approach is to identify the suites of wall localized proteins that are expressed during ripening, while another is take advantage of the diversity of texture associated phenotypes that are collectively exhibited by different cultivars, and to correlate those differences with patterns of CW protein expression. To this end, Konozy et al. (2013) used a proteomic approach to qualitatively compare the CW proteomes of fruits from three tomato cultivars with distinctly different fruit textural traits. Both non-disruptive and disruptive approaches were used to isolate soluble apoplastic proteins and those that were weakly bound to the CW, respectively. The former used vacuum infiltration-centrifugation of tomato pericarp samples, while the disruptive assay involved pericarp tissue homogenization and consecutive washing of the CW enriched pellet in order to reduce contamination with cytosolic proteins, followed by elution of a CW protein fraction with a buffer containing a moderate salt concentration. A total of 75 proteins were identified, many with a predicted or known CW localization, although no major

differences were observed between cultivars. Further experiments will be needed to determine whether any of these CW proteins is responsible for the textural characteristics associated with each cultivar. However, this study represents one of the first efforts to profile the CW proteome of tomato fruit pericarp using sequential extraction approaches that have previously been applied to other plant organs and complex tissues (Watson et al., 2004; Zhu et al., 2006).

In addition to those involved in CW metabolism, substantial numbers of apoplastic proteins and peptides function in plant defense against microbial pathogens, including many of the classical pathogenesis-related (PR) proteins (van Loon et al., 2006; Ferreira et al., 2007; Lagaert et al., 2009; Benko-Iseppon et al., 2010; De-la-Peña and Vivanco, 2010). The susceptibility of ripening fruit to infection can also be influenced by endogenous CW disassembly (Cantu et al., 2008) and so characterization of extracellular proteins in the microenvironment of the infection site may provide insights into the complex factors that affect the nature and timing of the interaction between fruits and pathogens. Shah et al. (2012) used a non-disruptive shotgun proteomics approach to isolate and identify extracellular proteins associated with the infection of tomato fruit by the necrotrophic fungus *Botrytis cinerea*. A total of 558 tomato proteins were identified from the mature green (MG) and red ripe (RR) stage of wild type fruits and those from the non-ripening *ripening inhibitor* (*rin*) mutant, all of which had been inoculated with the fungus. These included proteins belonging to many of the classical PR families, as well as members of diverse protein families such as proteases, protease inhibitors, and peroxidases, revealing a complicated CW proteome cocktail. Interestingly, substantially fewer defense-related proteins were identified in the MG fruit compared with the non-ripening *rin* or RR wild type fruits, although the authors point out that this may reflect differences in tissue homogenization, and thus presumably extractability, in addition to possible difference in the host response to the pathogen. Regardless, this study provided a useful simultaneous qualitative snapshot of the fruit and pathogen CW-related proteomes.

Another important factor that provides a critical barrier against microbial pathogens, as well as protection against pests and abiotic stresses such as desiccation and UV radiation, is the plant cuticle, a specialized lipid rich plant CW that covers the aerial epidermis of land plants (Isaacson et al., 2009; Reina-Pinto and Yephremov, 2009; Yeats et al., 2012b). Tomato fruit are an excellent system in which to identify the proteins involved in cuticle formation and restructuring since, like many fruits, their cuticle is typically much thicker than that of vegetative organs. Yeats et al. (2010) took advantage of this in a study aimed at identifying proteins associated with tomato fruit cuticle biosynthesis. The authors employed a non-disruptive protocol by briefly dipping fruits in an organic solvent, followed by several protein fractionation strategies and two mass spectrometry techniques [LC-ESI-MS/MS (liquid chromatography–electrospray ionization tandem mass spectrometry) and MALDI-TOF/TOF (matrix-assisted laser desorption ionization–time-of-flight tandem mass spectrometer)]. A total of 202 proteins were identified, of which approximately 40% had a predicted N-terminal signal peptide (SP) suggesting targeting to the CW, although the tomato genome sequence was not available at the time of the analysis and so missing N-terminal sequences would likely have resulted in an underestimation. A number of lipid metabolism-related proteins were identified, one of which was a GDSL (Gly-Asp-Ser-Leu)-motif lipase/hydrolase that was recently shown to be cutin synthase, the enzyme that catalyzes the polymerization of cutin monomers at the polysaccharide CW–cuticle interface (Yeats et al., 2012b). This exemplifies the value of CW proteomics for targeting specific biological questions.

Beyond identifying CW protein sequences, a crucial level of information that is slowly starting to emerge relates to PTMs, and while the presence of glycoproteins and phosphoproteins in the CW is now well established (Chivasa et al., 2005; Jamet et al., 2006; Kaida et al., 2010; Melo-Braga et al., 2012; Ruiz-May et al., 2012a) this represents a relatively unexplored area of plant wall biology. Almost nothing is known about the functional significance of such decorations in CW proteins, but suppressing the expression of key enzymes associated with modification of the *N*-glycans (α-mannosidase and β-*N*-acetyl hexosaminidase) in either tomato or pepper (*Capsicum annuum*) has been reported to have profound effects on the ripening (Meli et al., 2010; Ghosh et al., 2011). This therefore represents a potentially exciting area of future research, which will be aided by new analytical pipelines for identifying PTMs and structurally characterizing complex *N*-glycan modifications (Ruiz-May et al., 2012b). An example of such an approach, and the first reported study of the tomato fruit *N*-glycoproteome, was described by Catalá et al. (2011), who used the *N*-glycan binding lectin Concanavalin A, coupled with two-dimensional liquid chromatography, to identify 133 CW proteins from RR stage tomato fruit pericarp. Of these, 89% had a predicted N-terminal secretory SP, suggesting that such as lectin affinity approach both allows a substantial enrichment in CW proteins and provides an opportunity to characterize the sites of *N*-glycosylation and structures of the associated *N*-glycans.

A COMPARISON OF TOMATO FRUIT CW PROTEIN STUDIES

Robust bioinformatic tools, such as SignalP 4.1 (Petersen et al., 2011), have been developed to predict the presence of an N-terminal SP and such computational approaches provide a reasonably reliable, although certainly not perfect, indicator of targeting to the secretory pathway. A portion of these secretory proteins eventually traffic to the apoplast, while other subsets localize within various compartment of the endomembrane system, or even other organelles (Rose and Lee, 2010). While a few CW proteins may travel through a non-canonical secretion pathway proteins (Cheng et al., 2009; Rose and Lee, 2010; Zhang et al., 2011), there is little evidence to suggest that these are anything other than rare exceptions. Thus, SP prediction represents a useful "first pass" means to predict the secretome and to estimate the degree of contamination of CW protein extracts with intracellular proteins. We examined the data derived from several of the tomato fruit proteome studies described above using SignalP 4.1, and determined that a relatively high proportion of predicted non-secreted proteins was identified in most cases, including well-known cytosolic proteins, which underscores the

extent of the problem of contamination (**Figure 1**). The lectin affinity chromatography (Catalá et al., 2011) yielded the highest percentage of predicted secreted proteins (75%) and thus, by inference, the lowest degree of contamination (**Figure 1**). We note that our analysis used a newer version of the software than was used in the original studies and so the values differ slightly from those that were originally reported. The lowest percentages of predicted CW proteins were evident in studies of the RR stage fruit, where the deterioration in tomato fruit integrity or pathogen mediated tissue damage likely resulted in cell rupture and consequent contamination. Analysis of the functional categories of the tomato fruit CW proteins identified in the various studies suggests that most are associated with CW modifications and they can be assigned to a range of GH families (**Figure 2**). The pectin esterase, pectin methylesterase inhibitor (PMEI) and peroxidase families are particularly well represented, as are members of the GH17 and GH19 families. These correspond to endo-1,3-β-glucanases and chitinases, respectively, which are well-known families of defense-related proteins and accordingly they were particularly prevalent in RR stage tomato fruit challenged with *B. cinerea* (**Figure 2A**; Shah et al., 2012). These defenses-related proteins did not show the same level of representation in tomato cultivars without infection (**Figure 2B**), suggesting that fungal infection triggered the expression of defense-related proteins. Interesting, during the infection with *B. cinerea*, β-galactosidase (GH35 family) was absent in the MG stage but was well represented in RR fruit (**Figure 2A**; Shah et al., 2012). This is in agreement with previous studies where β-galactosidase expression has been observed in the later stages of tomato fruit ripening (Smith et al., 2002; Sozzi et al., 1998). Furthermore, high β-galactosidase activity has been associated with fruit ripening, during which a substantial amount of galactose is lost from the wall (Seymour et al., 1990; Sozzi et al., 1998). In contrast, PMEI was one of the most abundant functional categories in the MG stage but was present at very low levels in both infected (**Figure 2A**) and non-infected (**Figure 2B**) RR stage tomatoes from different cultivars. Moreover, the representation of the PMEI was low in RR fruit even after lectin enrichment of these putative glycoproteins (**Figure 2C**). PMEI proteins were more frequently identified in fruits of the non-ripening *rin* mutant compared with wild type (**Figure 2A**), which further supports the notion that expression declines during ripening. PMEI regulates the expression of pectin methylesterase (PME) enzymes, which catalyze the de-esterification of pectins in the CW during ripening (Prasanna et al., 2007; Reca et al., 2012). This can make pectin polymers more susceptible to hydrolytic depolymerization by PG (Brummell and Harpster, 2001; Wakabayashi et al., 2003; Prasanna et al., 2007). Our analysis suggests that PMEI proteins may suppress the activity of PMEs in MG fruit during *B. cinerea* infection, thereby limiting pectin depolymerization, which in turn may strengthen the CW and deter microbial invasion. Indeed, the infection of MG tomato fruit by *B. cinerea* was limited even though the proportions of defense-related proteins were lower than in RR fruit (Shah et al., 2012).

It is noteworthy that when looking at the various studies as a whole, many known fruit CW proteins were not detected, and those that were identified were generally abundantly expressed.

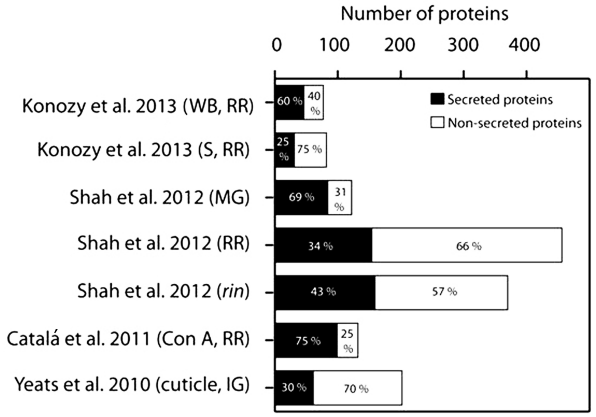


FIGURE 1 | Analysis of the proportions of proteins identified as entering the secretory pathway, as reported in the cited publications, from immature green (IG), mature green (MG), and red ripe (RR) stage wild type fruit and those of the *rin* mutant, using SignalP 4.1 Server (www.cbs.dtu.dk/services/SignalP). S, soluble; WB, wall-bound.

This suggests that additional enrichment and fractionation steps, together with higher sensitivity MS platforms will be necessary to provide more holistic coverage of the fruit CW proteome.

SUMMARY

Of the various plant subcellular proteomes that have been studied, arguably the most technically challenging is that of the CW, for the reasons described above. Moreover, the CW proteomes of fleshy fruit, such as tomato, represent extreme examples of such challenges, due largely to the composition and properties of the extracellular matrix that often limits effective and representative protein extraction. This raises the question of whether a comprehensive and quantitatively significant assessment of the whole fruit CW proteome is achievable. Remarkably, there are still no reports of large scale proteomic profiling initiatives of the tomato fruit CW spanning the various stage of fruit development and ripening in a single study and this likely reflects the major technical hurdles. The absence of equivalent data sets therefore currently limits biologically informative comparisons between studies (e.g., **Figure 2**). However, "proteomics" comes in many shapes and flavors and recent reports have shown that a great deal of useful information can be learnt from well-established and emerging analytical approaches (Nikolovski et al., 2012; Parsons et al., 2012; Zielinska et al., 2012). A promising emerging area is in the field of PTMs and, in particular, studies of the CW glycoproteome and phosphoproteome will doubtless shed new light on many aspects of fruit biology. This is further suggested by a recent study describing changes in *N*-glycosylation, phosphorylation, and Lys-acetylation during grape berry infection by the pathogen *Lobesia botrana* (Melo-Braga et al., 2012). A similar analysis of such changes during tomato fruit development and responses to environmental changes will doubtless give equivalent information and take analysis of this dynamic subcellular proteome to the next level.

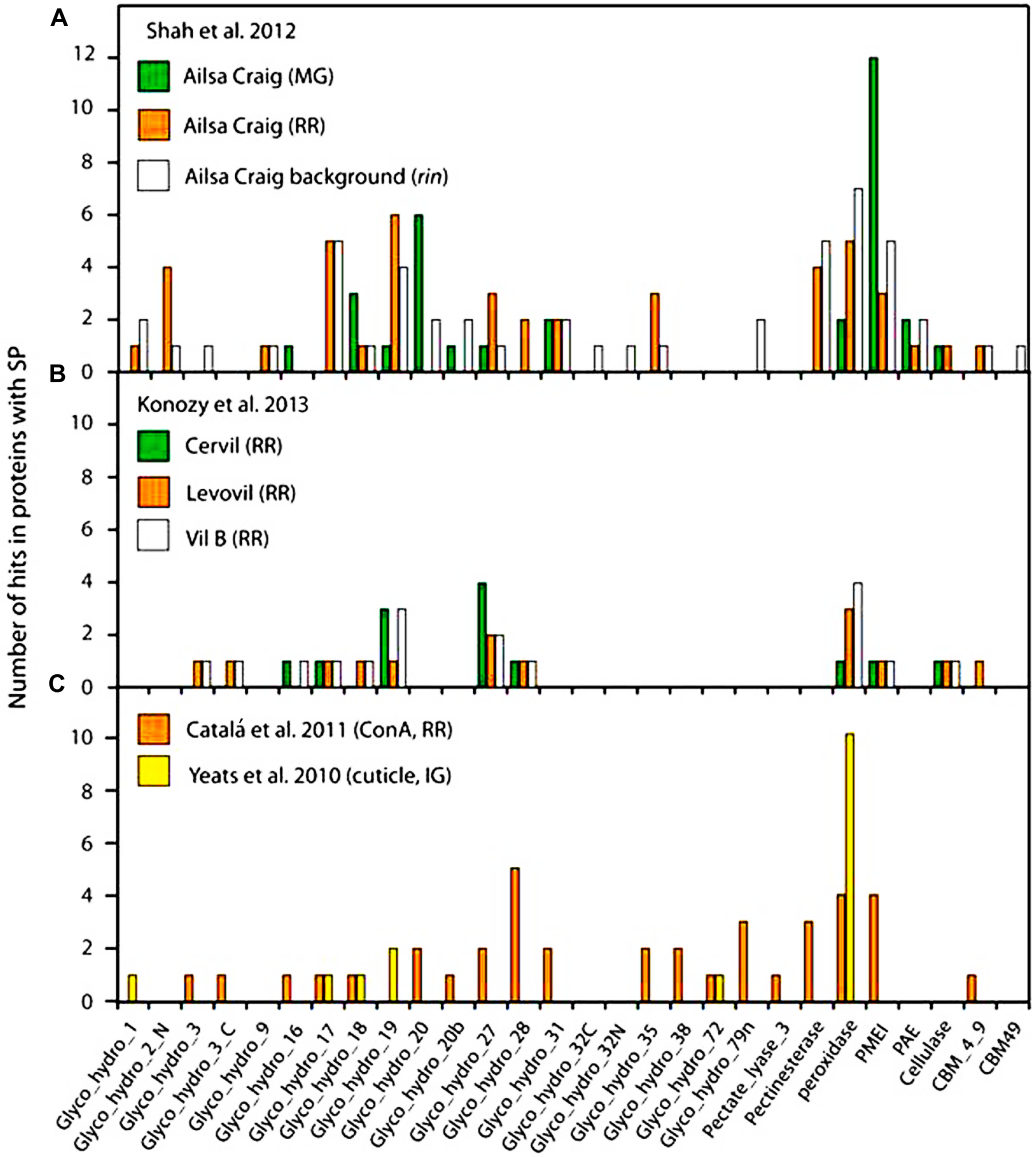


FIGURE 2 | Classification of tomato fruit proteins with a SP described in the cited publications, generated using the Pfam database (http://pfam.sanger.ac.uk). (A) Classification of proteins identified in MG stage, RR stage, and *rin* mutant and of tomato fruit infected with *B. cinerea*. **(B)** Classification of proteins identified in RR stage in different cultivars of tomato. **(C)** Classification of proteins in enriched proteome studies targeting the cuticle [immature green (IG) stage] or after using lectin affinity chromatography (RR stage).

ACKNOWLEDGMENTS

Funding to Jocelyn K. C. Rose for research in this area is provided by the NSF Plant Genome Research Program (DBI-0606595 and EAGER award IOS1313887), by the Agriculture and Food Research Initiative competitive grant #2011-04197 of the USDA National Institute of Food and Agriculture, and the New York State Office of Science, Technology and Academic Research (NYSTAR).

REFERENCES

Albenne, C., Canut, H., Boudart, G., Zhang, Y., San Clemente, H., Pont-Lezica, R., et al. (2009). Plant cell wall proteomics: mass spectrometry data, a trove for research on protein structure/function relationships. *Mol. Plant* 2, 977–989.

Bayer, E. M., Bottrill, A. R., Walshaw, J., Vigouroux, M., Naldrett, M. J., Thomas, C. L., et al. (2006). *Arabidopsis* cell wall proteome defined using multidimensional protein identification technology. *Proteomics* 6, 301–311.

Benko-Iseppon, A. M., Galdino, S. L., Calsa, T., Kido, E. A., Tossi, A., Belarmino, L. C., et al. (2010). Overview on plant antimicrobial peptides. *Curr. Protein Pept. Sci.* 11, 181–188.

Blee, K. A., Wheatley, E. R., Bonham, V. A., Mitchell, G. P., Robertson, D., Slabas, A. R., et al. (2001). Proteomic analysis reveals a novel set of cell wall proteins in a transformed tobacco cell culture that synthesises secondary walls as determined by biochemical and morphological parameters. *Planta* 212, 404–415.

Borderies, G., Jamet, E., Lafitte, C., Rossignol, M., Jauneau, A., Boudart, G., et al. (2003). Proteomics of loosely bound cell wall proteins of *Arabidopsis thaliana* cell suspension cultures: a critical analysis. *Electrophoresis* 24, 3421–3432.

Brummell, D. A. (2006). Cell wall disassembly in ripening fruit. *Funct. Plant Biol.* 33, 103–119.

Brummell, D. A., and Harpster, M. H. (2001). Cell wall metabolism in fruit softening and quality and its manipulation in transgenic plants. *Plant Mol. Biol.* 47, 311–340.

Cantu, D., Vicente, A. R., Greve, L. C., Dewey, F. M., Bennett, A. B., Labavitch, J. M., et al. (2008). The intersection between cell wall disassembly, ripening, and fruit susceptibility to *Botrytis cinerea. Proc. Natl. Acad. Sci. U.S.A.* 105, 859–864.

Catalá, C., Howe, K. J., Hucko, S., Rose, J. K., and Thannhauser, T. W. (2011). Towards characterization of the glycoproteome of tomato (*Solanum lycopersicum*) fruit using concanavalin A lectin affinity chromatography and LC-MALDI-MS/MS analysis. *Proteomics* 11, 1530–1544.

Chen, X. Y., Kim, S. T., Cho, W. K., Rim, Y., Kim, S., Kim, S. W., et al. (2009). Proteomics of weakly bound cell wall proteins in rice calli. *J. Plant Physiol.* 166, 675–685.

Cheng, F. Y., Zamski, E., Guo, W. W., Pharr, D. M., and Williamson, J. D. (2009). Salicylic acid stimulates secretion of the normally symplastic enzyme mannitol dehydrogenase: a possible defense against mannitol-secreting fungal pathogens. *Planta* 230, 1093–1103.

Chivasa, S., Ndimba, B. K., Simon, W. J., Robertson, D., Yu, X. L., Knox, J. P., et al. (2002). Proteomic analysis of the *Arabidopsis thaliana* cell wall. *Electrophoresis* 23, 1754–1765.

Chivasa, S., Simon, W. J., Yu, X. L., Yalpani, N., and Slabas, A. R. (2005). Pathogen elicitor-induced changes in the maize extracellular matrix proteome. *Proteomics* 5, 4894–4904.

Cho, W. K., Chen, X. Y., Chu, H., Rim, Y., Kim, S., Kim, S. T., et al. (2009). Proteomic analysis of the secretome of rice calli. *Physiol. Plant.* 135, 331–341.

De-la-Peña, C., and Vivanco, J. M. (2010). Root–microbe interactions: the importance of protein secretion. *Curr. Proteomics* 7, 265–274.

Faurobert, M., Mihr, C., Bertin, N., Pawlowski, T., Negroni, L., Sommerer, N., et al. (2007). Major proteome variations associated with cherry tomato pericarp development and ripening. *Plant Physiol.* 143, 1327–1346.

Feiz, L., Irshad, M., Pont-Lezica, R. F., Canut, H., and Jamet, E. (2006). Evaluation of cell wall preparations for proteomics: a new procedure for purifying cell walls from *Arabidopsis hypocotyls. Plant Methods* 2, 10.

Ferreira, R. B., Monteiro, S., Freitas, R., Santos, C. N., Chen, Z., Batista, L. M., et al. (2007). The role of plant defence proteins in fungal pathogenesis. *Mol. Plant Pathol.* 8, 677–700.

Ghosh, S., Meli, V. S., Kumar, A., Thakur, A., Chakraborty, N., Chakraborty, S., et al. (2011). The *N*-glycan processing enzymes α-mannosidase and β-D-*N*-acetylhexosaminidase are involved in ripening-associated softening in the non-climacteric fruits of capsicum. *J. Exp. Bot.* 62, 571–582.

Isaacson, T., Kosma, D. K., Matas, A. J., Buda, G. J., He, Y., Yu, B., et al. (2009). Cutin deficiency in the tomato fruit cuticle consistently affects resistance to microbial infection and biomechanical properties, but not transpirational water loss. *Plant J.* 60, 363–377.

Jamet, E., Albenne, C., Boudart, G., Irshad, M., Canut, H., and Pont-Lezica, R. (2008). Recent advances in plant cell wall proteomics. *Proteomics* 8, 893–908.

Jamet, E., Canut, H., Boudart, G., and Pont-Lezica, R. F. (2006). Cell wall proteins: a new insight through proteomics. *Trends Plant Sci.* 11, 33–39.

Kaida, R., Serada, S., Norioka, N., Norioka, S., Neumetzler, L., Pauly, M., et al. (2010). Potential role for purple acid phosphatase in the dephosphorylation of wall proteins in tobacco cells. *Plant Physiol.* 153, 603–610.

Klee, H. J., and Giovannoni, J. J. (2011). Genetics and control of tomato fruit ripening and quality attributes. *Annu. Rev. Genet.* 45, 41–59.

Konozy, E. H., Rogniaux, H., Causse, M., and Faurobert, M. (2013). Proteomic analysis of tomato (*Solanum lycopersicum*) secretome. *J. Plant Res.* 126, 251–266.

Kramer, M. G., and Redenbaugh, K. (1994). Commercialization of a tomato with an antisense polygalacturonase gene: the FLAVR SAVR tomato story. *Euphytica* 79, 293–297.

Lagaert, S., Belien, T., and Volckaert, G. (2009). Plant cell walls: protecting the barrier from degradation by microbial enzymes. S*emin. Cell Dev. Biol.* 20, 1064–1073.

Lee, S.-J., Saravanan, R. S., Damasceno, C. M. B., Yamane, H., Kim, B.-D., and Rose, J. K. C. (2004). Digging deeper into the plant cell wall proteome. *Plant Physiol. Biochem.* 42, 979–988.

Matas, A. J., Yeats, T. H., Buda, G. J., Zheng, Y., Chatterjee, S., Tohge, T., et al. (2011). Tissue and cell type specific transcriptome profiling of expanding tomato fruit provides insights into metabolic and regulatory specialization and cuticle formation. *Plant Cell* 23, 3893–3910.

Meli, V. S., Ghosh, S., Prabha, T. N., Chakraborty, N., Chakraborty, S., and Datta, A. (2010). Enhancement of fruit shelf life by suppressing *N*-glycan processing enzymes. *Proc. Natl. Acad. Sci. U.S.A.* 107, 2413–2418.

Melo-Braga, M. N., Verano-Braga, T., Leon, I. R., Antonacci, D., Nogueira, F. C., Thelen, J. J., et al. (2012). Modulation of protein phosphorylation, *N*-glycosylation and Lys-acetylation in grape (*Vitis vinifera*) mesocarp and exocarp owing to *Lobesia botrana* infection. *Mol. Cell. Proteomics* 11, 945–956.

Millar, D. J., Whitelegge, J. P., Bindschedler, L. V., Rayon, C., Boudet, A. M., Rossignol, M., et al. (2009). The cell wall and secretory proteome of a tobacco cell line synthesising secondary wall. *Proteomics* 9, 2355–2372.

Negri, A. S., Prinsi, B., Scienza, A., Morgutti, S., Cocucci, M., and Espen, L. (2008). Analysis of grape berry cell wall proteome: a comparative evaluation of extraction methods. *J. Plant Physiol.* 165, 1379–1389.

Nikolovski, N., Rubtsov, D., Segura, M. P., Miles, G. P., Stevens, T. J., Dunkley, T. P., et al. (2012). Putative glycosyltransferases and other plant Golgi apparatus proteins are revealed by LOPIT proteomics. *Plant Physiol.* 160, 1037–1051.

Palma, J. M., Corpas, F. J., and Del Rio, L. A. (2011). Proteomics as an approach to the understanding of the molecular physiology of fruit development and ripening. *J. Proteomics* 74, 1230–1243.

Parsons, H. T., Christiansen, K., Knierim, B., Carroll, A., Ito, J., Batth, T. S., et al. (2012). Isolation and proteomic characterization of the *Arabidopsis* Golgi defines functional and novel components involved in plant cell wall biosynthesis. *Plant Physiol.* 159, 12–26.

Payasi, A., Mishra, N. N., Chaves, A. L. S., and Singh, R. (2009). Biochemistry of fruit softening: an overview. *Physiol. Mol. Biol. Plants* 15, 103–113.

Petersen, T. N., Brunak, S., Von Heijne, G., and Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* 8, 785–786.

Prasanna, V., Prabha, T. N., and Tharanathan, R. N. (2007). Fruit ripening phenomena – an overview. *Crit. Rev. Food Sci. Nutr.* 47, 1–19.

Reca, I. B., Lionetti, V., Camardella, L., D'Avino, R., Giardina, T., Cervone, F., et al. (2012). A functional pectin methylesterase inhibitor protein (SolyPMEI) is expressed during tomato fruit ripening and interacts with PME-1. *Plant Mol. Biol.* 79, 429–442.

Reina-Pinto, J. J., and Yephremov, A. (2009). Surface lipids and plant defenses. *Plant Physiol. Biochem.* 47, 540–549.

Robertson, D., Mitchell, G. P., Gilroy, J. S., Gerrish, C., Bolwell, G. P., and Slabas, A. R. (1997). Differential extraction and protein sequencing reveals major differences in patterns of primary cell wall proteins from plants. *J. Biol. Chem.* 272, 15841–15848.

Rose, J. K. C., and Lee, S. J. (2010). Straying off the highway: trafficking of secreted plant proteins and complexity in the plant cell wall proteome. *Plant Physiol.* 153, 433–436.

Ruiz-May, E., Kim, S. J., Brandizzi, F., and Rose, J. K. C. (2012a). The secreted plant *N*-glycoproteome and associated secretory pathways. *Front. Plant Sci.* 3:117. doi: 10.3389/fpls.2012.00117

Ruiz-May, E., Thannhauser, T. W., Zhang, S., and Rose, J. K. C. (2012b). Analytical technologies for identification and characterization of the plant *N*-glycoproteome. *Front. Plant Sci.* 3:150. doi: 10.3389/fpls.2012.00150

Seymour, G. B., Chapman, N., Chew, B. L., and Rose, J. K. C. (2013). Regulation of ripening and opportunities for control in tomato and other fruits. *Plant Biotechnol.* 11, 269–278.

Seymour, G. B., Colquhoun, I. J., Dupont, M. S., Parsley, K. R., and Sevendran, R. R. (1990). Composition and structural features of cell wall polysaccharides from tomato fruits. *Phytochemistry* 29, 725–731.

Shah, P., Powell, A. L., Orlando, R., Bergmann, C., and Gutierrez-Sanchez, G. (2012). Proteomic analysis of ripening tomato fruit infected by *Botrytis cinerea. J. Proteome Res.* 11, 2178–2192.

Smith, C. J. S., Watson, C. F., Ray, J., Bird, C. R., Morris, P. C., Schuch, W., et al. (1988). Antisense RNA inhibition of polygalacturonase gene-expression in transgenic tomatoes. *Nature* 334, 724–726.

Smith, D. L., Abbott, J. A., and Gross, K. C. (2002). Down-regulation of

tomato beta-galactosidase 4 results in decreased fruit softening. *Plant Physiol.* 129, 1755–1762.

Sozzi, G. O., Camperi, S. A., Cascone, O., and Fraschina, A. A. (1998). Galactosidases in tomato fruit ontogeny: decreased galactosidase activities in antisense ACC synthase fruit during ripening and reversal with exogenous ethylene. *Aust. J. Plant Physiol.* 25, 237–244.

Tomato Genome Consortium. (2012). The tomato genome sequence provides insights into fleshy fruit evolution. *Nature* 485, 635–641.

van Loon, L. C., Rep, M., and Pietersen, C. M. (2006). Significance of inducible defense-related proteins in infected plants. *Annu. Rev. Phytopathol.* 44, 135–162.

Vicente, A. R., Saladié, M., Rose, J. K. C., and Labavitch, J. M. (2007). The linkage between cell wall metabolism and fruit softening: looking to the future. *J. Sci. Food Agric.* 87, 1435–1448.

Wakabayashi, K., Hoson, T., and Huber, D. J. (2003). Methyl de-esterification as a major factor regulating the extent of pectin depolymerization during fruit ripening: a comparison of the action of avocado (*Persea americana*) and tomato (*Lycopersicon esculentum*) polygalacturonases. *J. Plant Physiol.* 160, 667–673.

Watson, B. S., Lei, Z. T., Dixon, R. A., and Sumner, L. W. (2004). Proteomics of *Medicago sativa* cell walls. *Phytochemistry* 65, 1709–1720.

Yeats, T. H., Buda, G. J., Wang, Z., Chehanovsky, N., Moyle, L. C., Jetter, R., et al. (2012a). The fruit cuticles of wild tomato species exhibit architectural and chemical diversity, providing a new model for studying the evolution of cuticle function. *Plant J.* 69, 655–666.

Yeats, T. H., Martin, L. B., Viart, H. M., Isaacson, T., He, Y., Zhao, L., et al. (2012b). The identification of cutin synthase: formation of the plant polyester cutin. *Nat. Chem. Biol.* 8, 609–611.

Yeats, T. H., Howe, K. J., Matas, A. J., Buda, G. J., Thannhauser, T. W., and Rose, J. K. C. (2010). Mining the surface proteome of tomato (*Solanum lycopersicum*) fruit for proteins associated with cuticle biogenesis. *J. Exp. Bot.* 61, 3759–3771.

Zhang, H., Zhang, L., Gao, B., Fan, H., Jin, J., Botella, M. A., et al. (2011). Golgi apparatus-localized synaptotagmin 2 is required for unconventional secretion in *Arabidopsis*. *PLoS ONE* 6:e26477. doi: 10.1371/journal.pone.0026477

Zhang, Y., Giboulot, A., Zivy, M., Valot, B., Jamet, E., and Albenne, C. (2010). Combining various strategies to increase the coverage of the plant cell wall glycoproteome. *Phytochemistry* 72, 1109–1123.

Zhu, J., Alvarez, S., Marsh, E. L., Lenoble, M. E., Cho, I. J., Sivaguru, M., et al. (2007). Cell wall proteome in the maize primary root elongation zone. II. Region-specific changes in water soluble and lightly ionically bound proteins under water deficit. *Plant Physiol.* 145, 1533–1548.

Zhu, J., Chen, S., Alvarez, S., Asirvatham, V. S., Schachtman, D. P., Wu, Y., et al. (2006). Cell wall proteome in the maize primary root elongation zone. I. Extraction and identification of water-soluble and lightly ionically bound proteins. *Plant Physiol.* 140, 311–325.

Zielinska, D. F., Gnad, F., Schropp, K., Wisniewski, J. R., and Mann, M. (2012). Mapping *N*-glycosylation sites across seven evolutionarily distant species reveals a divergent substrate proteome despite a common core machinery. *Mol. Cell* 46, 542–548.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 08 January 2013; accepted: 09 May 2013; published online: 29 May 2013.

Citation: Ruiz-May E and Rose JKC (2013) Progress toward the tomato fruit cell wall proteome. Front. Plant Sci. ***4****:159. doi: 10.3389/fpls.2013.00159*

This article was submitted to Frontiers in Plant Proteomics, a specialty of Frontiers in Plant Science.

Copyright © 2013 Ruiz-May and Rose. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.

frontiers in
PLANT SCIENCE

MINI REVIEW ARTICLE
published: 05 February 2014
doi: 10.3389/fpls.2014.00021

The *Arabidopsis* cytosolic proteome: the metabolic heart of the cell

Jun Ito[1,2], Harriet T. Parsons[1,2,3] and Joshua L. Heazlewood[1,2]*

[1] Joint BioEnergy Institute, Emeryville, CA, USA
[2] Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA
[3] Department of Plant and Environmental Sciences, University of Copenhagen, Copenhagen, Denmark

Edited by:
Nicolas L. Taylor, The University of Western Australia, Australia

Reviewed by:
Ján A. Miernyk, University of Missouri, USA
Stefanie Wienkoop, University of Vienna, Austria

***Correspondence:**
Joshua L. Heazlewood, Joint BioEnergy Institute, Lawrence Berkeley National Laboratory, One Cyclotron Road MS 978-4466, Berkeley, CA 94720, USA
e-mail: jlheazlewood@lbl.gov

The plant cytosol is the major intracellular fluid that acts as the medium for inter-organellar crosstalk and where a plethora of important biological reactions take place. These include its involvement in protein synthesis and degradation, stress response signaling, carbon metabolism, biosynthesis of secondary metabolites, and accumulation of enzymes for defense and detoxification. This central role is highlighted by estimates indicating that the majority of eukaryotic proteins are cytosolic. *Arabidopsis thaliana* has been the subject of numerous proteomic studies on its different subcellular compartments. However, a detailed study of enriched cytosolic fractions from *Arabidopsis* cell culture has been performed only recently, with over 1,000 proteins reproducibly identified by mass spectrometry. The number of proteins allocated to the cytosol nearly doubles to 1,802 if a series of targeted proteomic characterizations of complexes is included. Despite this, few groups are currently applying advanced proteomic approaches to this important metabolic space. This review will highlight the current state of the *Arabidopsis* cytosolic proteome since its initial characterization a few years ago.

Keywords: cytosol, ribosome, proteasome, localization, *Arabidopsis*

INTRODUCTION

The cytosol is the liquid portion of a cell that contains principle cellular constituents comprising membrane-bound organelles. The cytosol itself lacks membrane compartmentalization. Within its highly concentrated aqueous setting of dissolved ionic solutes, small molecule metabolites and macromolecules, which include nucleic acids and proteins, a wide range of biochemical reactions are known to occur. These include an involvement in glycolysis (Plaxton, 1996), the oxidative branch of the pentose phosphate pathway (Schnarrenberger et al., 1995), protein biosynthesis and degradation (Bailey-Serres et al., 2009; Vierstra, 2009), signal transduction (Lecourieux et al., 2006; Klimecka and Muszynska, 2007), primary and secondary metabolite biosynthesis and transportation (Lundmark et al., 2006; Lunn, 2007; Martinoia et al., 2007; Weber and Fischer, 2007; Krueger et al., 2009), stress response signaling (Yamada and Nishimura, 2008; Cazale et al., 2009; Sugio et al., 2009), and the accumulation of enzymes for defense and detoxification (Laule et al., 2003; Dixon et al., 2009; Sappl et al., 2009). Furthermore, nuclear-encoded organellar proteins are synthesized in the cytosol prior to their import into organelles by targeting peptides (Jarvis, 2008; Prassinos et al., 2008; Huang et al., 2009). Although the cytosol has a multitude of prominent biochemical processes in the eukaryotic cell (**Figure 1**), only two proteome surveys have been carried to date on the plant cytosol. The first study identified 69 abundant proteins in cytosolic samples of soybean root nodules (Oehrle et al., 2008) while the second study identified 1,071 proteins from a large-scale mass spectrometry (MS) analysis of cytosol-enriched fractions from *Arabidopsis thaliana* cell suspension cultures (Ito et al., 2010). Many of the identified proteins were from well-known cytosolic processes (**Figure 1**); although a significant portion of the functionally unclassifiable proteins likely undertake novel roles in the cytosol (Ito et al., 2010). In this review, we will discuss further developments that have occurred from these initial proteomic analyses of the *Arabidopsis* cytosol.

THE *Arabidopsis* CYTOSOLIC 80S RIBOSOME

The cytosolic ribosome is a major component of the *Arabidopsis* cytosol and has been targeted by a number of studies for analysis by proteomics. A significant proportion of the proteins identified in the cytosolic proteome of *Arabidopsis* are involved in the core biological process of protein biosynthesis and degradation (Book et al., 2010; Ito et al., 2010; Hummel et al., 2012). The ribosome was well-represented amongst these proteins, with 92 previously identified ribosomal protein subunits from 61 of the 80 gene families (Ito et al., 2010). *Arabidopsis* ribosomal proteins have highly conserved sequences that belong to small gene families of two to six members, most of which are expressed (Carroll et al., 2008). A total of 79 of the 80 ribosomal protein families were characterized in purified ribosome preparations from *Arabidopsis* leaves (Giavalisco et al., 2005) and cell suspension cultures (Chang et al., 2005; Carroll et al., 2008). This included the identification of post-translational modifications (PTMs) such as initiator methionine removal, N-terminal acetylation, N-terminal methylation, lysine *N*-methylation, and phosphorylation. These studies represent basic proteomic surveys of the ribosome; more recent analyses have undertaken quantitative approaches to characterize this important protein complex of the cytosol.

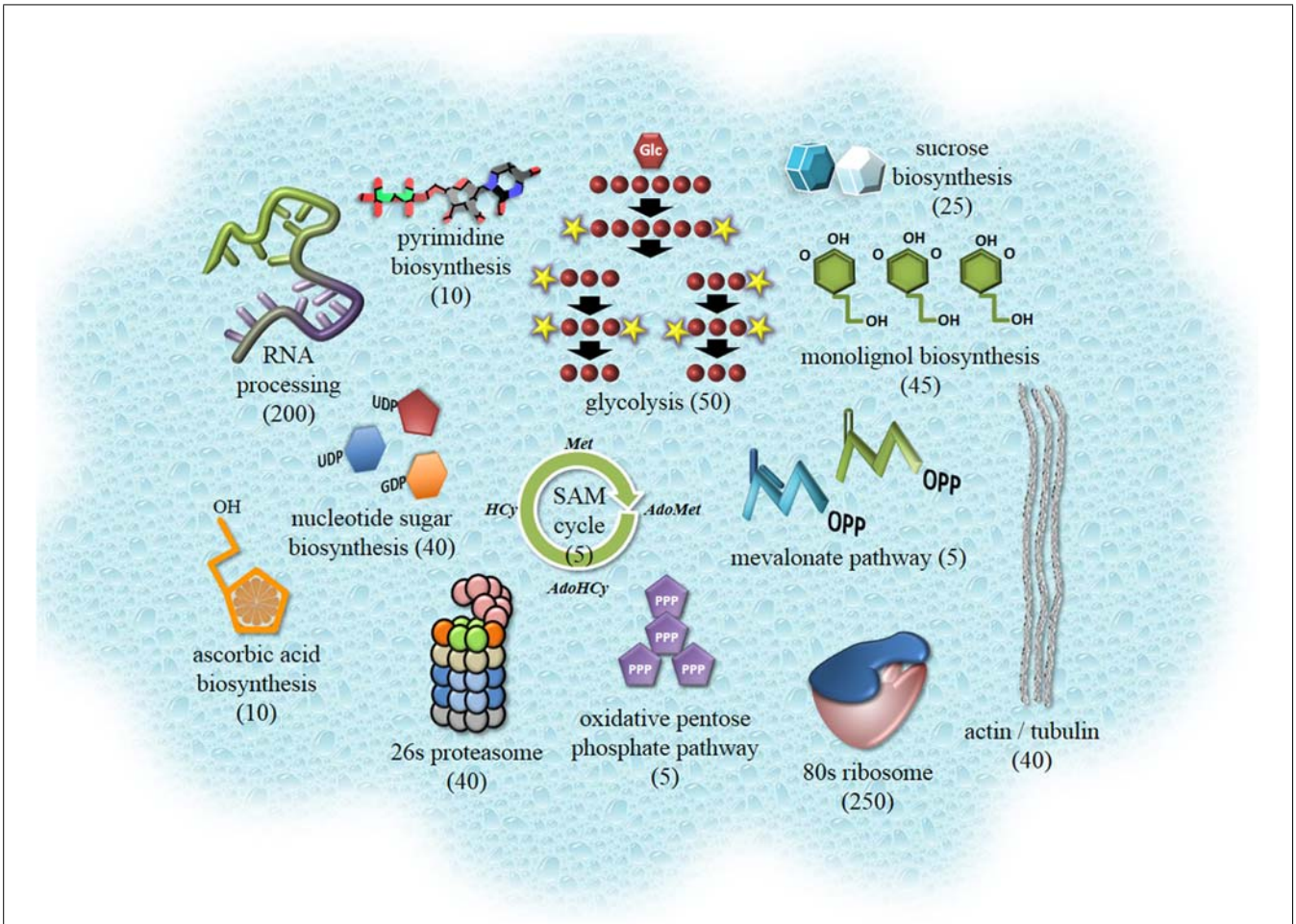


FIGURE 1 | Cartoon highlighting prominent metabolic processes, proteins, and protein complexes of the plant cytosol. Components were selected based on the proteomic data outlined in supplementary material of Ito et al. (2010). The number of proteins contributing to each of these components is shown in brackets. This information was obtained from Plant Metabolic Network resource (Zhang et al., 2005) and from MapMan pathway dataset for *Arabidopsis* (Thimm et al., 2004). Collectively, over 2,600 proteins can be assigned to the *Arabidopsis* cytosol when considering proteomics studies, localization with fluorescent proteins and functional curation.

Two quantitative proteomic studies have attempted to measure changes in the *Arabidopsis* ribosomal proteome under defined growing conditions. The first quantitative study investigated differential phosphorylation of purified ribosomal proteins from *Arabidopsis* leaves at day and night cycles as a possible mechanism to regulate diurnal protein synthesis (Turkina et al., 2011). Phosphorylation was detected by liquid chromatography (LC)–MS/MS on eight serine residues of six ribosomal proteins: S2-3, S6-1, S6-2, P0-2, P1, and L29-1. Relative quantification of phosphopeptides by differential stable isotope labeling and LC–MS/MS showed significant increases in day to night phosphorylation ratios of ribosomal proteins S6 at Ser-231 (2.2-fold), S6-1 and S6-2 variants at Ser-240 (4.2- and 1.8-fold, respectively), and L29-1 at Ser-58 (1.6-fold). This indicated that differential phosphorylation of these ribosomal proteins are likely mechanisms in modulating diurnal translation in plants (Turkina et al., 2011). The second study performed a label-free absolute quantitative analysis by LC–MS^E of immune-purified ribosomal protein paralogs from transgenic *Arabidopsis* leaves in response to sucrose feeding – a treatment known to have a profound effect on plant physiology and gene regulation (Hummel et al., 2012). The extensive families of ribosomal protein paralogs, the ambiguity of their incorporation into ribosomes and the potential alterations to ribosome composition in response to environmental and developmental cues were all factors in carrying out this study. Indeed, out of 204 ribosomal proteins identified by LC–MS/MS, 13 paralogs including S8A, S3aA, L12C, L19A-C, L30B, L8C, L28A, S12A, S12C, L22B, and S7C, as well as the ribosomal scaffold protein RACK1A, showed significant changes in their abundances up to 2.7-fold by LC–MS^E in response to sucrose treatments (Hummel et al., 2012). While L28A, L19A, and RACK1 have been shown to be important in normal plant growth and development (Tzafrir et al., 2004; Chen et al., 2006; Yao et al., 2008), the majority display limited phenotypic traits in their mutant plants. Concurrently, multiple ribosomal protein paralogs were shown to be incorporated into ribosomes in both sucrose fed and unfed

plants. It was surmised from these results that the *Arabidopsis* cytosolic ribosomes undergo variable alteration to their protein paralog compositions in reaction to changing external conditions (Hummel et al., 2012).

THE *Arabidopsis* CYTOSOLIC 26S PROTEASOME

The 26S proteasome is a complex of approximately 2.5 MDa which is responsible for the proteolytic degradation of most ubiquitylated proteins. Ubiquitylated protein degradation regulates processes such as the cell cycle, organ morphogenesis, circadian rhythms, and environmental response (Vierstra, 2009). The proteasome consists of a 28-subunit core protease (CP), which houses the active sites for protein and peptide hydrolysis, and a regulatory particle (RP) of at least 18 subunits which regulates substrate recognition, unfolding, and access to the CP. The architecture is highly conserved amongst eukaryotes but recent affinity purification of the 26S complex from *Arabidopsis* has revealed that although the plant 26S proteasome is analogous to that of the human and yeast (Kim et al., 2011), important differences exist.

In *Arabidopsis*, as in other plant groups, almost all subunits in both the CP and RP are encoded by duplicate genes of at least 90% homology, of which few appear to be pseudogenes (Book et al., 2010). Complexes containing all subunit duplicants have been purified from whole plants and characterized by MS (Yang et al., 2004; Book et al., 2010). It is not known yet whether duplicants are inserted into the 26S proteasome randomly or specifically. If these subunit "duplicants" are functionally specific, this raises the possibility of localized regulation of specific protein groups by populations of 26S proteasomes containing specific subunit duplicants/variants. In mutant backgrounds for the RPT2a/b subunit (Lee et al., 2011), complementation studies revealed functional redundancy between duplicants. However, double rpt2a/rpt2b knockout mutants exhibited a more severe phenotype that either single mutant, suggesting redundancy is only partial. RPN2a has uniquely been shown to be unregulated in response to increased sucrose concentrations, implicating a RPN2a-complex in the degradation hexokinase signaling pathway proteins (Sun et al., 2012). Likewise, single RPN5a/b mutants are phenotypically different and double mutants are lethal (Book et al., 2009; Serino and Pick, 2013). Together, these pieces of evidence point toward neofunctionalization of gene duplicants, supporting the idea of multiple populations of complexes within a whole plant.

Most of what is known about the plant 26S proteasome comes from yeast studies and has been reviewed previously (Finley, 2009; Vierstra, 2009). However, a recent study of RPN10 in *Arabidopsis* shows that important functional differences exist, at least in recognition of ubiquitylated substrates (Lin et al., 2011). Further unique properties of the *Arabidopsis* 26S proteasome include a much greater degree of ubiquitylation of subunits than has been observed in yeast (Peng et al., 2003; Book et al., 2010). Subunits became ubiquitylated when still assembled as a complex, implying that this modification performed a function beyond tagging subunits for degradation after complex disassembly. Accessory proteins help assemble the complex and recognize and recruit ubiquitylated substrates. A number of proteins homologous to yeast accessory proteins co-purified with the *Arabidopsis* 26S proteasome, as well as some novel putative accessory proteins not found in yeast (Book et al., 2010). An interesting question for future studies is whether certain accessory proteins associate with particular subunit variants/duplicants.

An important aim in understanding plant 26S proteasome function is to understand the relationship between subunit composition, and specific protein degradation in response to changes in internal and external environments. Given the high identity of many of these subunits, this will involve a significant challenge for characterization by MS. Nonetheless, together with the recent analysis of the ubiquitylated proteome in *Arabidopsis* (Kim et al., 2013), such work will undoubtedly expand our understanding of signaling and process regulation related to this important cytosolic protein complex.

POST-TRANSLATIONAL MODIFICATIONS

The ability to routinely identify and quantify PTMs represents a grand challenge in the field of proteomics (Heazlewood, 2011). However, few proteomic studies have targeted a subcellular compartment to specifically characterize PTMs (de la Fuente van Bentem et al., 2006; Ito et al., 2009). To the best of our knowledge, no such survey has ever been conducted on highly purified cytosolic fractions from *Arabidopsis*. Aside from the detailed analyses of the purified cytosolic complexes 80S ribosome and 26S proteasome outlined above, PTMs identified on cytosolic localized proteins are largely the result of large-scale PTM-targeted studies. In *Arabidopsis*, this has included phosphorylation (Heazlewood et al., 2008), *N*-linked glycosylation (Zielinska et al., 2012), ubiquitination (Kim et al., 2013), methionine oxidation (Marondedze et al., 2013), *S*-nitrosylation (Fares et al., 2011), and acetylation (Finkemeier et al., 2011). With few exceptions, these studies comprise collections of identified sites and do not generally explore the functional implication of a PTM. However, a number of more detailed investigations have identified the importance of PTMs on proteins localized to the cytosol. Entry into the cytosolic oxidative pentose phosphate pathway (OPPP) is catalyzed by glucose-6-phosphate dehydrogenase (G6PD) which is encoded by AT3G27300 and AT5G40760 in *Arabidopsis*. Large-scale phosphoproteomic studies have identified phosphorylation sites on both cytosolic isoforms. Recently it was demonstrated that the phosphorylation of AT5G40760 at Thr-467 increased G6PD activity fourfold (Dal Santo et al., 2012). Glycolysis represents a key metabolic pathway in the plant cytosol. The sixth step in this pathway is catalyzed by glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and represents the beginning of a net gain in ATP and NADH. In *Arabidopsis*, the step is encoded by a small gene family, a member of which has been identified as lysine acetylated (AT1G13440) in *Arabidopsis*. It was also demonstrated that the acetylation of Lys-130 inhibited the activity of this enzyme *in vitro* and consequently this PTM may represent a regulatory mechanism for this step in the pathway (Finkemeier et al., 2011). GAPDH encoded by AT1G13440 also contains *N*-glycosylation and numerous phosphorylation sites according to a number of targeted PTM studies (Heazlewood et al., 2008; Zielinska et al., 2012). The functional roles, if any, of the many thousands of

PTMs on cytosolic localized proteins will likely take many years to accurately characterize. Recently many of these sites were incorporated into the MASCP Gator, the *Arabidopsis* proteomics aggregation portal (Mann et al., 2013). It is envisaged that the inclusion of this information into such a utility will enable the community to better leverage these data for future functional analyses.

UTILIZATION OF THE *Arabidopsis* CYTOSOLIC PROTEOME

Establishing the subcellular location of a protein is an important factor in determining its function (Chou and Cai, 2003). MS analysis of purified organelles or cellular compartments and chimeric fluorescent fusion proteins are two common experimental methods used to define subcellular localizations of *Arabidopsis* proteins (Heazlewood et al., 2007; Tanz et al., 2013). Over 2,200 proteins contain information indicating a cytosolic localization in *Arabidopsis* (**Table 1**), which comprises nearly 25% of all experimentally localized proteins in the SUBcellular Arabidopsis database (SUBA). A large proportion of these cytosolic proteins have been identified in multiple subcellular compartments, especially in the case of proteomic approaches. It is therefore ideal, though often not the case, that protein localization is confirmed using complementary methods (Millar et al., 2009).

Several recent reports have used data from the *Arabidopsis* cytosolic proteome to confirm functional interpretations supporting a localization in the cytosol. Overall, they exemplify the practicality of this subcellular proteome for verifying the cytosolic localizations of different proteins. Glyoxylate reductase (GLYR) is a central enzyme in the γ-aminobutyrate (GABA) metabolic pathway, where it catalyzes the detoxification of glyoxylate and succinic semialdehdye (Ching et al., 2012). The two plant isoforms GLYR1 and GLYR2 were believed to localize to the cytosol or peroxisomes, and plastid, respectively. Conflicting reports of *Arabidopsis* GLYR1 (At3g25530) localizing in the cytosol (Simpson et al., 2008) or the peroxisome (Reumann et al., 2009) had implications for defining its exact metabolic roles and the compartmentation of the GABA and photorespiratory pathways. This was resolved by visualizing N-terminal green fluorescent protein (GFP)-tagged GLYR1 in *Arabidopsis* suspension-cultured cells, leaves and seedlings and tobacco BY-2 suspension-cultured cells, where it was observed to exclusively localize in the cytosol (Ching et al., 2012). Its identification by MS as a major protein in the cytosolic proteome of *Arabidopsis* cell suspensions was cited as further evidence of this finding (Ito et al., 2010; Ching et al., 2012).

The *Arabidopsis* translation elongation factor eEF-1Bβ1 (EF1Bβ, At1g30230) is involved in plant cell wall biosynthesis and it is essential for normal plant development (Hossain et al., 2012). *Arabidopsis* plants with T-DNA insertions in their EF1Bβ gene display a dwarf phenotype, with alterations to their vascular morphology and inflorescence stem structures and 38 and 20% reductions in total lignin and crystalline cellulose content, respectively. By transforming *Arabidopsis* plants with a 35S promoter-controlled EF1Bβ fused with yellow fluorescent protein (EF1Bβ-YFP), the subcellular locations of EF1Bβ were visualized in the plasma membrane and cytosol (Hossain et al., 2012). These observations agreed with MS analyses of the *Arabidopsis* plasma membrane (Mitra et al., 2009) and cytosol proteomes (Ito et al., 2010), with EF1Bβ identified in both subcellular compartments.

An evolutionary and structural analysis of a human disrupted in schizophrenia 1 (DISC1) protein conducted orthology searches of non-vertebrate reference organisms such as *Dictyostelium*, *Trichoplax*, *Monosiga*, and *Arabidopsis* (Sanchez-Pulido and Ponting, 2011). This study found that while most DISC1 orthologs lacked any experimental evidence of their functions, the *Arabidopsis* DISC1 ortholog (At5g25070) is ubiquitously expressed in various tissues and developmental stages and is a constituent of the *Arabidopsis* cytosolic proteome (Ito et al., 2010). This was strikingly similar to human DISC1, which is expressed in a wide range of tissues and also cytosol-localized (Sanchez-Pulido and Ponting, 2011).

Table 1 | A survey of cytosolic proteins experimentally localized in *Arabidopsis* from the SUBA database as of November 2013 (Tanz et al., 2013).

	MS/MS	FP	Total	Overlap
All locations	7891	2647	9319	1219
Cytosol	1808	580	2262	126

MS/MS indicates proteins identified through subcellular proteomics studies; FP are proteins localized using a fluorescent protein tag. The overlap between FP and MS/MS for cytosolic proteins is significantly worse than all proteins localized in the SUBA database. Possibly reflecting poor attention to this subcellular space and its processes by the research community.

EXPANDING THE *Arabidopsis* CYTOSOLIC PROTEOME

A computational analysis of the *Arabidopsis* proteome estimated that the cytosolic proteome may contain around 5,400 ± 650 proteins (Ito et al., 2010). This indicates that the current experimental set of 2,262 proteins likely represents only about 40% of the cytosolic proteome (**Table 1**). A dissection of fluorescent protein-based localization studies of *Arabidopsis* proteins (**Table 1**) reveals that many members were also identified in the *Arabidopsis* cytosolic proteome (recent examples include Ching et al., 2012; Christ et al., 2012; Hossain et al., 2012; Li et al., 2012; Lu et al., 2012; McLoughlin et al., 2012; Witz et al., 2012). However, there are many examples of FP-tagged proteins that have been localized to the cytosol and not identified by proteomic surveys (some recent studies include Gaber et al., 2012; Hernandez et al., 2012; Kwon et al., 2012; Lu et al., 2012; McLoughlin et al., 2012; Rautengarten et al., 2012; Vadassery et al., 2012; Witz et al., 2012). The inclusion of complementary subcellular datasets such as those available from the gene ontology database AmiGO (Carbon et al., 2009) and UniProtKB (Magrane and UniProt Consortium, 2011) can also be used to capture some of these missing cytosolic proteins. Nearly 2000 *Arabidopsis* proteins are designated as cytosolic by AmiGO, while about 1,300 *Arabidopsis* proteins are allocated to the cytosol by the UniProt Protein Knowledgebase. Incorporating these data with the proteomic and fluorescent protein information, the total number of *Arabidopsis* proteins with some cytosolic designation is 2604 distinct members or about 50% of the computationally

derived proteome. It should be noted that the "experimental" figure of ca. 2,600 does not account for false positives resulting from proteins with multiple subcellular designations. Over 1,400 of these proteins also have non-cytosolic assignments by either MS or fluorescent protein localizations according to SUBA (Tanz et al., 2013).

While proteomics has identified a considerable proportion of the computationally derived cytosolic proteome (around 30%), the shortfall can be readily explained and include: many proteins are not abundant and thus not easily detected by MS, many proteins could be expressed in tissue(s) other than cell suspension cultures or only under certain conditions (i.e., at a specific stage of plant development or in response to stress) and most significantly only one out of the nearly 120 proteomic analyses of various subcellular compartments from *Arabidopsis* has been performed on its cytosolic fraction (Heazlewood et al., 2007; Ito et al., 2010). In contrast, studies in *Arabidopsis* in the areas of respiration and photosynthesis have benefited tremendously from the characterization of their proteomes across different organs and tissues, developmental stages, and growth conditions (Lee et al., 2008; van Wijk and Baginsky, 2011). In order to better understand its dynamics, future analyses of the *Arabidopsis* cytosolic proteome will also need to reach this level of diversity.

A critical factor in performing in-depth proteomic analysis of the cytosol from plants will be to obtain relatively pure cytosolic fractions from this material. Isolating the cytosolic fraction from *Arabidopsis* cell suspensions relies on enzymatic generation of protoplasts and their disruption by gentle pressure to maintain organelle integrity, followed by organelle removal by differential centrifugation (Ito et al., 2010). Unlike uniform heterotrophic cell suspensions, cytosol purification from plants requires extra steps including the removal of chloroplasts. A study of protein localization between cytosol and chloroplasts of *Arabidopsis* seedlings developed a method for isolating the cytosolic fraction from protoplasts of seedlings (Estavillo et al., 2011). The addition of density centrifugation was necessary to remove broken protoplasts and intact chloroplasts, respectively, from the seedling cytosolic fraction (Estavillo et al., 2011, 2014). Employing immunoblotting or MS-based quantitation against subcellular markers to assess organelle contamination during the extraction process (Ito et al., 2010), this method could be further refined to generate high-purity cytosolic fractions from many types of *Arabidopsis* plant material for proteomic analysis.

Sub-fractionation of the cytosol is an effective way to reduce its protein complexity and to improve MS/MS identification of low abundant cytosolic proteins. Unlike mitochondria and plastids, the cytosol lacks defined membrane-bound compartments that can be further sub fractionated (Eubel et al., 2007; Ferro et al., 2010). However, isolating soluble protein complexes from the *Arabidopsis* cytosol has been shown to be relatively straight forward. As outlined above, both the 80S ribosome and the 26S proteasome have been isolated and extensively characterized by MS (Yang et al., 2004; Chang et al., 2005; Giavalisco et al., 2005; Carroll et al., 2008; Book et al., 2010; Turkina et al., 2011; Hummel et al., 2012). Beyond these examples, sub-fractionation of other cytosolic protein groups will likely rely on affinity purification techniques tailored to the physiochemical properties of target proteins to simplify complex mixtures and enrich for low abundant proteins. In non-plant systems approaches have included immobilized heparin chromatography to fractionate cytosolic proteins from human breast cancer MCF-7 cells (Shefcheck et al., 2003). Approximately 300 low-abundant cytosolic proteins were detected by two-dimensional gel electrophoresis (2-DE) of heparin fractions, and they were not present on 2-DE separations of total cytosolic protein mixtures (Shefcheck et al., 2003). Finally, LC–MS/MS analysis of tandem biomimetic affinity pre-fractionation of rat liver cytosol proteins identified 665 unique rat proteins, which was significantly more than the 371 proteins in the unfractionated cytosol (Tan et al., 2009).

PERSPECTIVES

There is tremendous scope to extend our current knowledge of the multitude of reactions that take place in the plant cytosol. Few studies have employed quantitative proteomic approaches to study cytosolic components revealing a lack of attention to this important compartment. Similarly, the characterization and analysis of PTMs of cytosolic proteins will be a significant challenge in the future. Recent reports of cytosolic localizations of *Arabidopsis* proteins by fluorescent protein tagging showed that while a number of them were identified in the cytosolic proteome, many others were not. Future comparative analysis of cytosolic proteomes of different plant tissues grown under various environmental conditions is essential to better understand its dynamics and to unravel its complexity. Isolating pure cytosolic fractions and their subfractions from diverse sources of plant material for LC–MS/MS analysis will be key factors to achieve this aim.

AUTHOR CONTRIBUTIONS

The manuscript was devised by Jun Ito and written by Jun Ito, Harriet T. Parsons, and Joshua L. Heazlewood. Figure and Table were constructed by Joshua L. Heazlewood.

ACKNOWLEDGMENTS

This work was part of the DOE Joint BioEnergy Institute (http://www.jbei.org) supported by the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, through contract DE-AC02-05CH11231 between Lawrence Berkeley National Laboratory and the U.S. Department of Energy. Harriet T. Parsons was supported by a Marie Curie Fellowship.

REFERENCES

Bailey-Serres, J., Sorenson, R., and Juntawong, P. (2009). Getting the message across: cytoplasmic ribonucleoprotein complexes. *Trends Plant Sci.* 14, 443–453. doi: 10.1016/j.tplants.2009.05.004

Book, A. J., Gladman, N. P., Lee, S. S., Scalf, M., Smith, L. M., and Vierstra, R. D. (2010). Affinity purification of the *Arabidopsis* 26 S proteasome reveals a diverse array of plant proteolytic complexes. *J. Biol. Chem.* 285, 25554–25569. doi: 10.1074/jbc.M110.136622

Book, A. J., Smalle, J., Lee, K.-H., Yang, P., Walker, J. M., Casper, S., et al. (2009). The RPN5 subunit of the 26s proteasome is essential for gametogenesis, sporophyte development, and complex assembly in *Arabidopsis. Plant Cell* 21, 460–478. doi: 10.1105/tpc.108.064444

Carbon, S., Ireland, A., Mungall, C. J., Shu, S., Marshall, B., Lewis, S., et al. (2009). AmiGO: online access to ontology and annotation data. *Bioinformatics* 25, 288–289. doi: 10.1093/bioinformatics/btn615

Carroll, A. J., Heazlewood, J. L., Ito, J., and Millar, A. H. (2008). Analysis of the *Arabidopsis* cytosolic ribosome proteome provides detailed insights into its components and their post-translational modification. *Mol. Cell. Proteomics* 7, 347–369. doi: 10.1074/mcp.M700052-MCP200

Cazale, A. C., Clement, M., Chiarenza, S., Roncato, M. A., Pochon, N., Creff, A., et al. (2009). Altered expression of cytosolic/nuclear HSC70-1 molecular chaperone affects development and abiotic stress tolerance in *Arabidopsis thaliana*. *J. Exp. Bot.* 60, 2653–2664. doi: 10.1093/jxb/erp109

Chang, I. F., Szick-Miranda, K., Pan, S., and Bailey-Serres, J. (2005). Proteomic characterization of evolutionarily conserved and variable proteins of *Arabidopsis* cytosolic ribosomes. *Plant Physiol.* 137, 848–862. doi: 10.1104/pp.104.053637

Chen, J. G., Ullah, H., Temple, B., Liang, J., Guo, J., Alonso, J. M., et al. (2006). RACK1 mediates multiple hormone responsiveness and developmental processes in *Arabidopsis*. *J. Exp. Bot.* 57, 2697–2708. doi: 10.1093/jxb/erl035

Ching, S. L., Gidda, S. K., Rochon, A., Van Cauwenberghe, O. R., Shelp, B. J., and Mullen, R. T. (2012). Glyoxylate reductase isoform 1 is localized in the cytosol and not peroxisomes in plant cells. *J. Integr. Plant Biol.* 54, 152–168. doi: 10.1111/j.1744-7909.2012.01103.x

Chou, K. C., and Cai, Y. D. (2003). Prediction and classification of protein subcellular location-sequence-order effect and pseudo amino acid composition. *J. Cell. Biochem.* 90, 1250–1260. doi: 10.1002/jcb.10719

Christ, B., Schelbert, S., Aubry, S., Sussenbacher, I., Muller, T., Krautler, B., et al. (2012). MES16, a member of the methylesterase protein family, specifically demethylates fluorescent chlorophyll catabolites during chlorophyll breakdown in *Arabidopsis*. *Plant Physiol.* 158, 628–641. doi: 10.1104/pp.111.188870

Dal Santo, S., Stampfl, H., Krasensky, J., Kempa, S., Gibon, Y., Petutschnig, E., et al. (2012). Stress-induced GSK3 regulates the redox stress response by phosphorylating glucose-6-phosphate dehydrogenase in *Arabidopsis*. *Plant Cell* 24, 3380–3392. doi: 10.1105/tpc.112.101279

de la Fuente van Bentem, S., Anrather, D., Roitinger, E., Djamei, A., Hufnagl, T., Barta, A., et al. (2006). Phosphoproteomics reveals extensive in vivo phosphorylation of *Arabidopsis* proteins involved in RNA metabolism. *Nucleic Acids Res.* 34, 3267–3278. doi: 10.1093/nar/gkl429

Dixon, D. P., Hawkins, T., Hussey, P. J., and Edwards, R. (2009). Enzyme activities and subcellular localization of members of the *Arabidopsis* glutathione transferase superfamily. *J. Exp. Bot.* 60, 1207–1218. doi: 10.1093/jxb/ern365

Estavillo, G. M., Crisp, P. A., Pornsiriwong, W., Wirtz, M., Collinge, D., Carrie, C., et al. (2011). Evidence for a SAL1-PAP chloroplast retrograde pathway that functions in drought and high light signaling in *Arabidopsis*. *Plant Cell* 23, 3992–4012. doi: 10.1105/tpc.111.091033

Estavillo, G. M., Verhertbruggen, Y., Scheller, H. V., Pogson, B. J., Heazlewood, J. L., and Ito, J. (2014). Isolation of the plant cytosolic fraction for proteomic analysis. *Methods Mol. Biol.* 1072, 453–467. doi: 10.1007/978-1-62703-631-3_31

Eubel, H., Heazlewood, J. L., and Millar, A. H. (2007). "Isolation and subfractionation of plant mitochondria for proteomic analysis," in *Plant Proteomics: Methods and Protocols*, eds H. Thiellement, M. Zivy, C. Damerval, and V. Méchin. (Totowa, NJ: Humana Press Inc.), 49–62

Fares, A., Rossignol, M., and Peltier, J. B. (2011). Proteomics investigation of endogenous *S*-nitrosylation in *Arabidopsis*. *Biochem. Biophys. Res. Commun.* 416, 331–336. doi: 10.1016/j.bbrc.2011.11.036

Ferro, M., Brugière, S., Salvi, D., Seigneurin-Berny, D., Court, M., Moyet, L., et al. (2010). AT_CHLORO, a comprehensive chloroplast proteome database with subplastidial localization and curated information on envelope proteins. *Mol. Cell. Proteomics* 9, 1063–1084. doi: 10.1074/mcp.M900325-MCP200

Finkemeier, I., Laxa, M., Miguet, L., Howden, A. J. M., and Sweetlove, L. J. (2011). Proteins of diverse function and subcellular location are lysine acetylated in *Arabidopsis*. *Plant Physiol.* 155, 1779–1790. doi: 10.1104/pp.110.171595

Finley, D. (2009). Recognition and processing of ubiquitin–protein conjugates by the proteasome. *Annu. Rev. Biochem.* 78, 477–513. doi: 10.1146/annurev.biochem.78.081507.101607

Gaber, A., Ogata, T., Maruta, T., Yoshimura, K., Tamoi, M., and Shigeoka, S. (2012). The involvement of *Arabidopsis* glutathione peroxidase 8 in the suppression of oxidative damage in the nucleus and cytosol. *Plant Cell Physiol.* 53, 1596–1606. doi: 10.1093/pcp/pcs100

Giavalisco, P., Wilson, D., Kreitler, T., Lehrach, H., Klose, J., Gobom, J., et al. (2005). High heterogeneity within the ribosomal proteins of the *Arabidopsis thaliana* 80S ribosome. *Plant Mol. Biol.* 57, 577–591. doi: 10.1007/s11103-005-0699-3

Heazlewood, J. L. (2011). The green proteome: challenges in plant proteomics. *Front. Plant Sci.* 2:6. doi: 10.3389/fpls.2011.00006

Heazlewood, J. L., Durek, P., Hummel, J., Selbig, J., Weckwerth, W., Walther, D., et al. (2008). PhosPhAt: a database of phosphorylation sites in *Arabidopsis thaliana* and a plant-specific phosphorylation site predictor. *Nucleic Acids Res.* 36, D1015–D1021

Heazlewood, J. L., Verboom, R. E., Tonti-Filippini, J., Small, I., and Millar, A. H. (2007). SUBA: the Arabidopsis Subcellular Database. *Nucleic Acids Res.* 35, D213–D218. doi: 10.1093/nar/gkl863

Hernandez, L., Whitehead, L., He, Z., Gazda, V., Gilday, A., Kozhevnikova, E., et al. (2012). A cytosolic acyltransferase contributes to triacylglycerol synthesis in sucrose-rescued *Arabidopsis* seed oil catabolism mutants. *Plant Physiol.* 160, 215–225. doi: 10.1104/pp.112.201541

Hossain, Z., Amyot, L., Mcgarvey, B., Gruber, M., Jung, J., and Hannoufa, A. (2012). The translation elongation factor eEF-1Bβ1 is involved in cell wall biosynthesis and plant development in *Arabidopsis thaliana*. *PLoS ONE* 7:e30425. doi: 10.1371/journal.pone.0030425

Huang, S., Taylor, N. L., Whelan, J., and Millar, A. H. (2009). Refining the definition of plant mitochondrial presequences through analysis of sorting signals, N-terminal modifications, and cleavage motifs. *Plant Physiol.* 150, 1272–1285. doi: 10.1104/pp.109.137885

Hummel, M., Cordewener, J. H., De Groot, J. C., Smeekens, S., America, A. H., and Hanson, J. (2012). Dynamic protein composition of *Arabidopsis thaliana* cytosolic ribosomes in response to sucrose feeding as revealed by label free MSE proteomics. *Proteomics* 12, 1024–1038. doi: 10.1002/pmic.201100413

Ito, J., Batth, T. S., Petzold, C. J., Redding-Johanson, A. M., Mukhopadhyay, A., Verboom, R., et al. (2010). Analysis of the *Arabidopsis* cytosolic proteome highlights subcellular partitioning of central plant metabolism. *J. Proteome Res.* 10, 1571–1582. doi: 10.1021/pr1009433

Ito, J., Taylor, N. L., Castleden, I., Weckwerth, W., Millar, A. H., and Heazlewood, J. L. (2009). A survey of the *Arabidopsis thaliana* mitochondrial phosphoproteome. *Proteomics* 9, 4229–4240. doi: 10.1002/pmic.200900064

Jarvis, P. (2008). Targeting of nucleus-encoded proteins to chloroplasts in plants. *New Phytol.* 179, 257–285. doi: 10.1111/j.1469-8137.2008.02452.x

Kim, D. Y., Scalf, M., Smith, L. M., and Vierstra, R. D. (2013). Advanced proteomic analyses yield a deep catalog of ubiquitylation targets in *Arabidopsis*. *Plant Cell* 25, 1523–1540. doi: 10.1105/tpc.112.108613

Kim, H. M., Yu, Y., and Cheng, Y. (2011). Structure characterization of the 26S proteasome. *Biochim. Biophys. Acta* 1809, 67–79. doi: 10.1016/j.bbagrm.2010.08.008

Klimecka, M., and Muszynska, G. (2007). Structure and functions of plant calcium-dependent protein kinases. *Acta Biochim. Pol.* 54, 219–233

Krueger, S., Niehl, A., Lopez Martin, M. C., Steinhauser, D., Donath, A., Hildebrandt, T., et al. (2009). Analysis of cytosolic and plastidic serine acetyltransferase mutants and subcellular metabolite distributions suggests interplay of the cellular compartments for cysteine biosynthesis in *Arabidopsis*. *Plant Cell Environ.* 32, 349–367. doi: 10.1111/j.1365-3040.2008.01928.x

Kwon, Y., Yu, S. I., Lee, H., Yim, J. H., Zhu, J. K., and Lee, B. H. (2012). *Arabidopsis* serine decarboxylase mutants implicate the roles of ethanolamine in plant growth and development. *Int. J. Mol. Sci.* 13, 3176–3188. doi: 10.3390/ijms13033176

Laule, O., Furholz, A., Chang, H. S., Zhu, T., Wang, X., Heifetz, P. B., et al. (2003). Crosstalk between cytosolic and plastidial pathways of isoprenoid biosynthesis in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* 100, 6866–6871. doi: 10.1073/pnas.1031755100

Lecourieux, D., Ranjeva, R., and Pugin, A. (2006). Calcium in plant defence-signalling pathways. *New Phytol.* 171, 249–269. doi: 10.1111/j.1469-8137.2006.01777.x

Lee, C. P., Eubel, H., O'Toole, N., and Millar, A. H. (2008). Heterogeneity of the mitochondrial proteome for photosynthetic and non-photosynthetic *Arabidopsis* metabolism. *Mol. Cell. Proteomics* 7, 1297–1316. doi: 10.1074/mcp.M700535-MCP200

Lee, K. H., Minami, A., Marshall, R. S., Book, A. J., Farmer, L. M., Walker, J. M., et al. (2011). The RPT2 subunit of the 26S proteasome directs complex assembly, histone dynamics, and gametophyte and sporophyte development in *Arabidopsis*. *Plant Cell* 23, 4298–4317. doi: 10.1105/tpc.111.089482

Li, H. M., Chen, H., Yang, Z. N., and Gong, J. M. (2012). Cdi gene is required for pollen germination and tube growth in *Arabidopsis*. *FEBS Lett.* 586, 1027–1031. doi: 10.1016/j.febslet.2012.02.046

Lin, Y. L., Sung, S. C., Tsai, H. L., Yu, T. T., Radjacommare, R., Usharani, R., et al. (2011). The defective proteasome but not substrate recognition function is responsible for the null phenotypes of the *Arabidopsis* proteasome subunit RPN10. *Plant Cell* 23, 2754–2773. doi: 10.1105/tpc.111.086702

Lu, W., Tang, X., Huo, Y., Xu, R., Qi, S., Huang, J., et al. (2012). Identification and characterization of fructose 1,6-bisphosphate aldolase genes in *Arabidopsis* reveal a gene family with diverse responses to abiotic stresses. *Gene* 503, 65–74. doi: 10.1016/j.gene.2012.04.042

Lundmark, M., Cavaco, A. M., Trevanion, S., and Hurry, V. (2006). Carbon partitioning and export in transgenic *Arabidopsis thaliana* with altered capacity for sucrose synthesis grown at low temperature: a role for metabolite transporters. *Plant Cell Environ.* 29, 1703–1714. doi: 10.1111/j.1365-3040.2006.01543.x

Lunn, J. E. (2007). Compartmentation in plant metabolism. *J. Exp. Bot.* 58, 35–47. doi: 10.1093/jxb/erl134

Magrane, M., and UniProt Consortium. (2011). UniProt Knowledgebase: a hub of integrated protein data. *Database* 2011, bar009. doi: 10.1093/database/bar009

Mann, G. W., Calley, P. C., Joshi, H. J., and Heazlewood, J. L. (2013). MASCP gator: an overview of the *Arabidopsis* proteomic aggregation portal. *Front. Plant Sci.* 4:411. doi: 10.3389/fpls.2013.00411

Marondedze, C., Turek, I., Parrott, B., Thomas, L., Jankovic, B., Lilley, K. S., et al. (2013). Structural and functional characteristics of cGMP-dependent methionine oxidation in *Arabidopsis thaliana* proteins. *Cell Commun. Signal.* 11, 1. doi: 10.1186/1478-811X-11-1

Martinoia, E., Maeshima, M., and Neuhaus, H. E. (2007). Vacuolar transporters and their essential role in plant metabolism. *J. Exp. Bot.* 58, 83–102. doi: 10.1093/jxb/erl183

McLoughlin, F., Galvan-Ampudia, C. S., Julkowska, M. M., Caarls, L., Van Der Does, D., Lauriere, C., et al. (2012). The Snf1-related protein kinases SnRK2.4 and SnRK2.10 are involved in maintenance of root system architecture during salt stress. *Plant J.* 72, 436–449. doi: 10.1111/j.1365-313X.2012.05089.x

Millar, A. H., Carrie, C., Pogson, B., and Whelan, J. (2009). Exploring the function-location nexus: using multiple lines of evidence in defining the subcellular location of plant proteins. *Plant Cell* 21, 1625–1631. doi: 10.1105/tpc.109.066019

Mitra, S. K., Walters, B. T., Clouse, S. D., and Goshe, M. B. (2009). An efficient organic solvent based extraction method for the proteomic analysis of *Arabidopsis* plasma membranes. *J. Proteome Res.* 8, 2752–2767. doi: 10.1021/pr801044y

Oehrle, N. W., Sarma, A. D., Waters, J. K., and Emerich, D. W. (2008). Proteomic analysis of soybean nodule cytosol. *Phytochemistry* 69, 2426–2438. doi: 10.1016/j.phytochem.2008.07.004

Peng, J., Schwartz, D., Elias, J. E., Thoreen, C. C., Cheng, D., Marsischky, G., et al. (2003). A proteomics approach to understanding protein ubiquitination. *Nat. Biotechnol.* 21, 921–926. doi: 10.1038/nbt849

Plaxton, W. C. (1996). The organization and regulation of plant glycolysis. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 47, 185–214. doi: 10.1146/annurev.arplant.47.1.185

Prassinos, C., Haralampidis, K., Milioni, D., Samakovli, D., Krambis, K., and Hatzopoulos, P. (2008). Complexity of Hsp90 in organelle targeting. *Plant Mol. Biol.* 67, 323–334. doi: 10.1007/s11103-008-9322-8

Rautengarten, C., Ebert, B., Ouellet, M., Nafisi, M., Baidoo, E. E., Benke, P., et al. (2012). *Arabidopsis* deficient in cutin ferulate encodes a transferase required for feruloylation of omega-hydroxy fatty acids in cutin polyester. *Plant Physiol.* 158, 654–665. doi: 10.1104/pp.111.187187

Reumann, S., Quan, S., Aung, K., Yang, P., Manandhar-Shrestha, K., Holbrook, D., et al. (2009). In-depth proteome analysis of *Arabidopsis* leaf peroxisomes combined with in vivo subcellular targeting verification indicates novel metabolic and regulatory functions of peroxisomes. *Plant Physiol.* 150, 125–143. doi: 10.1104/pp.109.137703

Sanchez-Pulido, L., and Ponting, C. P. (2011). Structure and evolutionary history of DISC1. *Hum. Mol. Genet.* 20, R175–R181. doi: 10.1093/hmg/ddr374

Sappl, P. G., Carroll, A. J., Clifton, R., Lister, R., Whelan, J., Millar, A. H., et al. (2009). The *Arabidopsis* glutathione transferase gene family displays complex stress regulation and co-silencing multiple genes results in altered metabolic sensitivity to oxidative stress. *Plant J.* 58, 53–68. doi: 10.1111/j.1365-313X.2008.03761.x

Schnarrenberger, C., Flechner, A., and Martin, W. (1995). Enzymatic evidence for a complete oxidative pentose phosphate pathway in chloroplasts and an incomplete pathway in the cytosol of spinach leaves. *Plant Physiol.* 108, 609–614. doi: 10.1104/pp.108.2.609

Serino, G., and Pick, E. (2013). Duplication and familial promiscuity within the proteasome lid and COP9 signalosome kin complexes. *Plant Sci.* 203, 89–97. doi: 10.1016/j.plantsci.2012.12.018

Shefcheck, K., Yao, X., and Fenselau, C. (2003). Fractionation of cytosolic proteins on an immobilized heparin column. *Anal. Chem.* 75, 1691–1698. doi: 10.1021/ac026153h

Simpson, J. P., Di Leo, R., Dhanoa, P. K., Allan, W. L., Makhmoudova, A., Clark, S. M., et al. (2008). Identification and characterization of a plastid-localized *Arabidopsis* glyoxylate reductase isoform: comparison with a cytosolic isoform and implications for cellular redox homeostasis and aldehyde detoxification. *J. Exp. Bot.* 59, 2545–2554. doi: 10.1093/jxb/ern123

Sugio, A., Dreos, R., Aparicio, F., and Maule, A. J. (2009). The cytosolic protein response as a subcomponent of the wider heat shock response in *Arabidopsis*. *Plant Cell* 21, 642–654. doi: 10.1105/tpc.108.062596

Sun, H. H., Sako, K., Suzuki, Y., Maekawa, S., Yasuda, S., Chiba, Y., et al. (2012). Sugar-inducible RPT2a, a subunit of 26S proteasome, participates in sugar response in *Arabidopsis*. *Plant Biotechnol.* 29, 279–284. doi: 10.5511/plantbiotechnology.12.0409a

Tan, Q., Dong, D., Ye, L., Huo, C., Huang, F., and Li, R. (2009). Pre-fractionation of rat liver cytosol proteins prior to mass spectrometry-based proteomic analysis using tandem biomimetic affinity chromatography. *J. Mol. Recognit.* 23, 93–100. doi: 10.1002/jmr.995

Tanz, S. K., Castleden, I., Hooper, C. M., Vacher, M., Small, I., and Millar, H. A. (2013). SUBA3: a database for integrating experimentation and prediction to define the SUBcellular location of proteins in *Arabidopsis*. *Nucleic Acids Res.* 41, D1185–D1191. doi: 10.1093/nar/gks1151

Thimm, O., Blasing, O., Gibon, Y., Nagel, A., Meyer, S., Kruger, P., et al. (2004). MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *Plant J.* 37, 914–939. doi: 10.1111/j.1365-313X.2004.02016.x

Turkina, M. V., Klang Arstrand, H., and Vener, A. V. (2011). Differential phosphorylation of ribosomal proteins in *Arabidopsis thaliana* plants during day and night. *PLoS ONE* 6:e29307. doi: 10.1371/journal.pone.0029307

Tzafrir, I., Pena-Muralla, R., Dickerman, A., Berg, M., Rogers, R., Hutchens, S., et al. (2004). Identification of genes required for embryo development in *Arabidopsis*. *Plant Physiol.* 135, 1206–1220. doi: 10.1104/pp.104.045179

Vadassery, J., Reichelt, M., Hause, B., Gershenzon, J., Boland, W., and Mithofer, A. (2012). CML42-mediated calcium signaling coordinates responses to *Spodoptera* herbivory and abiotic stresses in *Arabidopsis*. *Plant Physiol.* 159, 1159–1175. doi: 10.1104/pp.112.198150

van Wijk, K. J., and Baginsky, S. (2011). Plastid proteomics in higher plants: current state and future goals. *Plant Physiol.* 155, 1578–1588. doi: 10.1104/pp.111.172932

Vierstra, R. D. (2009). The ubiquitin-26S proteasome system at the nexus of plant biology. *Nat. Rev. Mol. Cell Biol.* 10, 385–397. doi: 10.1038/nrm2688

Weber, A. P., and Fischer, K. (2007). Making the connections – the crucial role of metabolite transporters at the interface between chloroplast and cytosol. *FEBS Lett.* 581, 2215–2222. doi: 10.1016/j.febslet.2007.02.010

Witz, S., Jung, B., Furst, S., and Mohlmann, T. (2012). De novo pyrimidine nucleotide synthesis mainly occurs outside of plastids, but a previously undiscovered nucleobase importer provides substrates for the essential salvage pathway in *Arabidopsis*. *Plant Cell* 24, 1549–1559. doi: 10.1105/tpc.112.096743

Yamada, K., and Nishimura, M. (2008). Cytosolic heat shock protein 90 regulates heat shock transcription factor in *Arabidopsis thaliana*. *Plant Signal. Behav.* 3, 660–662. doi: 10.4161/psb.3.9.5775

Yang, P., Fu, H., Walker, J., Papa, C. M., Smalle, J., Ju, Y. M., et al. (2004). Purification of the *Arabidopsis* 26 S proteasome: biochemical and molecular analyses revealed the presence of multiple isoforms. *J. Biol. Chem.* 279, 6401–6413. doi: 10.1074/jbc.M311977200

Yao, Y., Ling, Q., Wang, H., and Huang, H. (2008). Ribosomal proteins promote leaf adaxial identity. *Development* 135, 1325–1334. doi: 10.1242/dev.017913

Zhang, P., Foerster, H., Tissier, C. P., Mueller, L., Paley, S., Karp, P. D., et al. (2005). MetaCyc and AraCyc. Metabolic pathway databases for plant research. *Plant Physiol.* 138, 27–37. doi: 10.1104/pp.105.060376

Zielinska, D. F., Gnad, F., Schropp, K., Wisniewski, J. R., and Mann, M. (2012). Mapping *N*-glycosylation sites across seven evolutionarily distant species reveals a divergent substrate proteome despite a common core machinery. *Mol. Cell* 46, 542–548. doi: 10.1016/j.molcel.2012.04.031

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 16 November 2013; accepted: 19 January 2014; published online: 05 February 2014.

Citation: Ito J, Parsons HT and Heazlewood JL (2014) The Arabidopsis cytosolic proteome: the metabolic heart of the cell. Front. Plant Sci. ***5****:21. doi: 10.3389/fpls.2014.00021*

This article was submitted to Plant Proteomics, a section of the journal Frontiers in Plant Science.

Copyright © 2014 Ito, Parsons and Heazlewood. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in
PLANT SCIENCE

REVIEW ARTICLE
published: 01 March 2013
doi: 10.3389/fpls.2013.00032

The *Arabidopsis* cytosolic ribosomal proteome: from form to function

Adam J. Carroll*

Australian Research Council Centre of Excellence in Plant Energy Biology, Australian National University, Canberra, ACT, Australia

Edited by:
Harvey Millar, The University of Western Australia, Australia

Reviewed by:
Peta Bonham-Smith, University of Saskatchewan, Canada
Kenichi Yamaguchi, Nagasaki University, Japan

Correspondence:
Adam J. Carroll, Australian Research Council Centre of Excellence in Plant Energy Biology, Australian National University, ACT 0200, Canberra, Australia.
e-mail: adam.carroll@anu.edu.au

The cytosolic ribosomal proteome of *Arabidopsis thaliana* has been studied intensively by a range of proteomics approaches and is now one of the most well characterized eukaryotic ribosomal proteomes. Plant cytosolic ribosomes are distinguished from other eukaryotic ribosomes by unique proteins, unique post-translational modifications and an abundance of ribosomal proteins for which multiple divergent paralogs are expressed and incorporated. Study of the *A. thaliana* ribosome has now progressed well beyond a simple cataloging of protein parts and is focused strongly on elucidating the functions of specific ribosomal proteins, their paralogous isoforms and covalent modifications. This review summarises current knowledge concerning the *Arabidopsis* cytosolic ribosomal proteome and highlights potentially fruitful areas of future research in this fast moving and important area.

Keywords: *Arabidopsis*, sub-cellular proteomics, proteomics, ribosomes, cytosolic ribosomes, plants, translation, 80S ribosomes

RIBOSOMES – A FUNDAMENTALLY IMPORTANT TARGET FOR BASIC AND APPLIED SCIENCE

Ribosomes – the ribonucleoprotein complexes responsible for catalyzing translation – the mRNA-guided synthesis of proteins from aminoacyl-tRNA, GTP and ATP substrates – have fascinated biologists since their Nobel Prize-winning discovery by George E. Palade in 1955 (Zorca and Zorca, 2011). Understanding how ribosomes are made by cells (ribosome biogenesis and structure), how they work (ribosome molecular mechanics) and how they are controlled through transcriptional, translational, and post-translational mechanisms is of fundamental importance for several reasons. The most obvious reason relates to the fundamental role of ribosomes in the generation of proteomes. Like DNA replication and transcription, translation is a basic requirement for life and an integral component of the Central Dogma of molecular biology.

Understanding the molecular mechanics of different ribosomes will increase our capacity to: (1) design bioactive agents to alter their function (Kannan et al., 2012) and (2) rationally engineer them to modify their performance (Piekna-Przybylska et al., 2008; Santoro et al., 2009) or even provide them with completely new functions – e.g., the residue-specific incorporation of unnatural amino acids into designer polypeptides with novel research and industrial applications (Bain et al., 1992; Benner, 1994; Taira et al., 2005; Neumann et al., 2010; Neumann, 2012). Custom-engineered ribosomes have even been used to create synthetic Boolean information processing networks that control gene expression according to rationally designed logic (Rackham and Chin, 2005, 2006). Despite these numerous examples illustrating the power of ribosome engineering in non-plant species, there are, to the author's knowledge, no published examples of applied ribosome engineering in plants. It seems inevitable that powerful applications of plant ribosome engineering will emerge in time.

Another reason for the fundamental importance of ribosome research relates to chemical and energy resource usage. In rapidly dividing yeast cells, up to at least 60% of transcriptional activity is devoted to ribosome biogenesis alone, consuming vast amounts of nitrogen (N), phosphorous (P), and energy while translation itself represents a further major demand on N and energy reserves (Warner, 1999; Piques et al., 2009). Understanding the mechanisms controlling ribosome biogenesis and translation in plants could therefore have profound implications for the management, engineering, and utilization of the enormous chemical energy fluxes in natural and agricultural ecosystems.

EUKARYOTIC RIBOSOMES – MORE COMPLEX MACHINES TO BUILD MORE COMPLEX ORGANISMS

The gross structure of ribosomes is essentially the same between prokaryotes and eukaryotes in that they are both comprised of ribosomal RNAs (rRNAs) and proteins (r-proteins) in large and small subunits. However, the ribosomes of eukaryotes exhibit greater structural complexity, reflecting the greater complexity of molecular mechanics observed in eukaryotic translation (Kapp and Lorsch, 2004). In eukaryotes, nuclear-encoded proteins (i.e., the vast majority of proteins) are synthesized on 80S cytosolic ribosomes which are distinguished from the 70S prokaryotic-type ribosomes of bacteria, mitochondria, and plastids by their larger size and higher number of proteins (~80 versus ~54). Each 80S eukaryotic ribosome is comprised of a large 60S subunit (50S in prokaryotic ribosomes) containing three rRNA molecules (5S, 5.8S, and a 23S-like rRNA ranging between 25S and 28S in plants) and up to 47 different r-proteins and a small 40S subunit (30S in prokaryotic ribosomes) containing a single 18S rRNA and up to 33 different r-proteins (Wilson and Doudna Cate, 2012).

The precise reasons for all the specific differences between prokaryotic and eukaryotic translation machineries and processes remain largely unknown. However, it seems likely that the higher

complexity of eukaryotic translation evolved in response to the following needs which, intuitively, seem likely to be characteristic of eukaryotic organisms: (1) to translate with a greater priority toward fidelity and control over speed of ribosome biogenesis; (2) to efficiently and accurately translate mRNAs having (and encoding proteins having) a wider range of primary and secondary structures; (3) to have greater control over the relative rates of translation of specific mRNAs; (4) to have a greater capacity for spatiotemporal ribosome heterogeneity in order to tailor the translation process for different subcellular locations, cell types, and developmental stages (Giavalisco et al., 2005; Komili et al., 2007; Sugihara et al., 2010; Xue and Barna, 2012).

Among eukaryotes, the cytosolic ribosomes of yeast (*Saccharomyces cerevisiae*), rat (*Rattus norvegicus*), human (*Homo sapiens*), and *Arabidopsis* (*Arabidopsis thaliana*) have been the most extensively characterized. Primarily through the extensive r-protein sequencing and gene cloning efforts of Wool et al. (1995), rat liver ribosomes were the first eukaryotic ribosomes for which a presumed complete list of r-proteins became available and has since served as a useful model for r-protein nomenclature in yeast (Mager et al., 1997) and plants (Barakat et al., 2001) although some inconsistencies do still exist between the r-protein nomenclatures of yeast and other eukaryotes. Efforts to characterize the cytosolic ribosomes of other eukaryote lineages have since revealed that all 79 of the r-protein families present in mammalian ribosomes are also represented in the ribosomes of yeast and plants (Wilson and Doudna Cate, 2012) although an additional plant-specific r-protein family known as acidic stalk protein P3 has been identified in the ribosomes of plants (Szick et al., 1998; Barakat et al., 2001; Chang et al., 2005; Carroll et al., 2008). This deep conservation of the protein composition of eukaryotic ribosomes suggests that the archetypal eukaryotic ribosome evolved very early in eukaryote evolution and that all of the r-protein families are important for ribosome function. That said, considerable primary sequence divergence has occurred between r-protein orthologs of different eukaryote lineages (Wool et al., 1995) and between r-protein paralogs within individual species that have emerged through gene duplication events during eukaryote evolution (Barakat et al., 2001). Hence, a major current focus of ribosome-related research is to elucidate the adaptive and physiological significance of these divergences and the ribosome heterogeneity that they enable (Wool et al., 1995; Wilson and Doudna Cate, 2012; Xue and Barna, 2012).

PLANT RIBOSOMES – A CHALLENGING TARGET FOR PROTEOMICS

Plants offer unique technical challenges to researchers of ribosomal proteomes. Firstly, in addition to the cytosolic and mitochondrial ribosomes found in mammals and fungi, plants contain a third type of ribosome in the plastid thus introducing more potential for cross-contamination of ribosome preparations and more potential for ambiguity with respect to the localization of r-proteins when they are detected in multiple cellular fractions. Protocols for the isolation of cytosolic ribosomes from plants must therefore incorporate special measures to avoid contamination from organelle ribosomes.

Another challenge associated with the study of plant ribosomes is that the possible degree of heterogeneity is particularly high (Giavalisco et al., 2005). While mammalian r-proteins are usually represented by only a single expressed gene (Sugihara et al., 2010) and yeast r-proteins are each represented by only one or two, often encoding identical proteins (McIntosh and Warner, 2007), the situation is far more complex in higher plants. Indeed, high heterogeneity appears to be particularly characteristic of higher plants with much less paralog heterogeneity being observed in 80S ribosomes of the green alga, *Chlamydomonas reinhardtii* (Manuell et al., 2005). A survey of the *Arabidopsis* genome (Barakat et al., 2001) revealed that none of the 80 different r-protein families were encoded by a single-copy gene. Rather, most were found to be encoded by three or four transcribed genes. These paralogs could theoretically combine to form more than 10^{34} different ribosomes, not including different post-translational modifications (PTMs; Hummel et al., 2012). This striking potential for heterogeneity is likely attributable to the sessile nature of plants and their greater need to be adaptable under changing environments than animals, who have more capacity to avoid environmental fluctuations.

A major ongoing challenge has been to determine not only which r-protein families but precisely which of the 251 r-protein genes encoded by the *A. thaliana* genome (Barakat et al., 2001; Chang et al., 2005) are transcribed and translated into proteins that are incorporated into ribosomes. As will be discussed shortly, a collection of proteomic studies (see **Table 1**) have confirmed the presence of all but one of the 81 predicted r-protein families and, in the case of many families, the presence of multiple distinct paralogous family members in *Arabidopsis* ribosomes (Chang et al., 2005; Giavalisco et al., 2005; Carroll et al., 2008; Piques et al., 2009; Turkina et al., 2011; Hummel et al., 2012).

Determining precisely which members of r-protein families are incorporated into ribosomes depends on the ability to confidently discriminate between those protein isoforms. In theory, proteotypic peptides – peptides specific to a single gene-product – may be generated from trypsin digestion of most r-proteins and confident detection of these may be used as evidence for the presence of their corresponding specific gene products. However, *in silico* analysis has revealed that 10 r-protein families (S18, S29, S30, L11, L21, L23, L36a, L38, L40, and L41) exhibit no sequence divergence within them while others (S15a, S16, S2, S20, S4, L11, L35a, L39, and L9) include some members predicted to generate proteotypic peptides and others that would not (Carroll et al., 2008). Hence, proteomics alone will be of limited use for assessing the heterogeneity of the ribosomal proteome in these perfectly homologous r-proteins. This is non-trivial given intriguing observations that, in yeast, independent deletion of paralogous genes encoding sequence-identical proteins cause readily distinguishable phenotypes, suggesting that these paralogous genes are functionally non-equivalent despite the fact that the proteins they encode are predicted to have identical amino acid sequences (Komili et al., 2007; McIntosh and Warner, 2007). Promoter analysis using reporter gene constructs expressed under the promoters of different r-protein paralogs, as exemplified in the L16 family (Williams and Sussex, 1995), will continue to be valuable in determining the physiological significance of these different paralogs.

Table 1 | Proteomic studies of *Arabidopsis thaliana* ribosomes – an overview.

Study	Tissue	Methods	Detected proteins/information provided
Giavalisco et al. (2005)	Leaves	Differential centrifugation and sucrose gradients, 2D-E, and MALDI-TOF PMF	Representatives of 60 r-protein families including 87 distinct r-protein identifications based mainly on intact peptide masses
Chang et al. (2005)	Cell suspension	Detergent extraction, ultracentrifugation and sucrose gradient followed by 2D-E, MALDI-TOF PMF, and some LC-MS/MS. Barium hydroxide treatment of phospho-proteins	Representatives of 74 r-protein families including 77 family member-specific claims based mainly on intact peptide masses
Carroll et al. (2008)	Cell suspension	Differential centrifugation and double sucrose cushion purification followed by 1D-SDS-PAGE and LC-MS/MS. Titanium dioxide phosphopeptide enrichment. Use of pepsin and chymotrypsin in addition to trypsin to increase coverage of low MW r-proteins. Bioinformatic analysis to quantify gene family member-specific MS/MS evidence	Representatives of 80 r-protein families including 87 family member-specific identifications based on detailed MS/MS analysis. 30 residue-specific post-translational modification sites including Initiator methionine removal, N-terminal acetylation, serine phosphorylation, lysine mono and tri-methylation, and N-terminal proline dimethylation
Piques et al. (2009)	Leaves sampled at different times of the diurnal cycle	Polysome fractions were isolated by detergent extraction and sucrose gradient fractionation as described in Kawaguchi et al. (2003). Proteins were then acetone-precipitated, trypsin digested, and analyses directly by nano-LC-MS/MS with data-dependent MS/MS acquisition. Absolute protein concentrations were estimated using the emPAI scoring method of Ishihama et al. (2005)	Estimates of absolute (mol%) concentrations of r-proteins in different polysome fractions at different times of day and night. MS/MS evidence provided to support the detection of representatives from 71 r-protein families. MS/MS evidence for proteotypic peptides from 92 specific r-proteins collectively across the various polysome samples
Turkina et al. (2011)	Leaves sampled at day and night	Detergent extraction followed by ultracentrifugation through sucrose cushion and then nano-LC-MS/MS for qualitative analysis. Also, quantitative phosphoproteomic analysis using differential isotopic labeling of tryptic peptides by methyl esterification with IMAC enrichment of phosphopeptides prior to LC-MS/MS. Provide counts of family member-specific peptides	Representatives from 72 r-protein families including family member-specific identifications of 71 r-proteins based on MS/MS. New phosphorylation sites on S6, S2 and L29. Diurnal changes in phosphorylation levels of S6 and L29
Hummel et al. (2012)	Leaves ± sucrose feeding	Immunopurification of ribosomes from a transgenic *A. thaliana* line expressing a His FLAG Tagged form of RPL18B (Zanetti et al., 2005; Mustroph et al., 2009). Relative quantitation with simultaneous identification by label-free LC-MS^E analysis	Representatives of 74 r-protein families including 166 family member-specific identifications. Changes in the levels of RPS3aA, RPS5A, RPL8B, and RACK1 in response to sucrose feeding were particularly evident

Brief details about each of the major proteomic studies of A. thaliana ribosomes are provided. PMF, Peptide Mass Fingerprinting; 2D-E, Two-dimensional gel electrophoresis; IMAC, Immobilized Metal Affinity Chromatography; LC-MS/MS, Liquid Chromatography – Tandem Mass Spectrometry; MALDI-TOF, Matrix Assisted Laser Desorption Ionization Time-of-Flight Mass Spectrometry.

DEFINING THE *ARABIDOPSIS* CYTOSOLIC RIBOSOMAL PROTEOME – PROGRESS TO DATE

Several proteomic studies of *Arabidopsis* cytosolic ribosomes have been reported in the literature – each employing its own unique combination of methods for purification, gel separation, mass-spectrometric detection and data analysis (summarized in **Table 1**). In the earliest of these reports, Giavalisco et al. (2005) combined differential centrifugation and sucrose density gradient purification of *Arabidopsis* leaf ribosomes with 2D-gel electrophoresis and Matrix Assisted Laser Desorption Ionization Time-of-Flight Mass Spectrometry (MALDI-TOF-MS)-based Peptide Mass Fingerprinting (PMF) to identify protein spots corresponding to 87 distinct r-protein gene products representing 60 of the 80 r-protein families. The authors highlighted low molecular weights, high pIs and low numbers of tryptic cleavage sites as possible reasons for the non-detection of the other 20 predicted families of r-proteins in their study. A key finding of the study was that at least 21 of the 60 detected r-protein families were represented by two or more distinct r-proteins (distinct AGIs) and that >45% of the distinct r-proteins detected were represented

by 2–13 separate spots. Indeed, this confirmed earlier predictions by Barakat et al. (2001) of high ribosomal heterogeneity in plants due to the frequent expression of multiple divergent paralogous r-protein genes. However, it also suggested that plant r-proteins exist in a much wider variety of modification states than those of other organisms (Giavalisco et al., 2005).

Shortly after Giavalisco et al. (2005) published their study of *Arabidopsis* leaf ribosomes, Chang et al. (2005) published an independent proteomic survey of 80S ribosomes isolated from heterotrophic *Arabidopsis* cell suspensions. This study combined detergent-based tissue lysis with differential centrifugation, sucrose gradient purification, 1D- and 2D-gel electrophoresis with MALDI-TOF-MS PMF and in some cases Liquid Chromatography – Electrospray Ionization – Quadrupole – Time-of-Flight – Tandem Mass Spectrometry (LC-ESI-Q-TOF-MS/MS). Protein assignments based on ∼850 peptide identifications (mostly based on MALDI-TOF-MS with 172 based on MS/MS) provided evidence for the ribosomal incorporation of 14 previously undetected r-protein families, bringing the total number of detected families to 74 and leaving only six undetected. This study also presented new evidence to support the identification of particular r-protein family members by reporting the masses of ions assigned to tryptic peptides predicted to belong to only one specific gene product. On this basis, Chang et al. (2005) provided paralog-specific evidence for 77 r-proteins with the following 25 r-protein families being represented in the cytosolic ribosomal proteome by more than one structurally distinct member: S3a, S6, S7, S10, S12, S14, S15, S15a, S16, S19, S23, S24, Sa, P0, P2, L4, L7, L7a, L8, L10, L10a, L18a, L26, L27, L31.

Even after considerable efforts by Giavalisco et al. (2005) and Chang et al. (2005) to define the *Arabidopsis* cytosolic ribosomal proteome, clear opportunities to gain further insight remained. *In silico* analyses suggested that many r-protein families for which gene-specific peptides were predicted to exist had still not been resolved, suggesting that a higher-coverage proteomic analysis based on peptide MS/MS rather than PMF might yield further evidence with which to resolve particular paralogs. Moreover, while limited tissue-type sampling was one possible reason for the non-detection of some r-proteins, *in silico* analyses also suggested that the six small and basic r-protein families remaining undetected might have been missed by the previous studies because their tryptic peptides were very small and that their detection may have been aided through the use of complementary proteases yielding larger fragments (Carroll et al., 2008).

Prompted by the above observations, Carroll et al. (2008) undertook a systematic analysis of highly pure 80S ribosomes isolated from *Arabidopsis* cell suspensions by combining an optimized ribosome isolation procedure with 1D gel electrophoresis, LC-ESI-Q-TOF-MS/MS analysis of excised gel bands (using three different proteases on low MW bands to capture larger peptides) and a custom data analysis pipeline to provide deep proteome coverage and high-confidence paralog-specific identifications. This analysis, based on 1446 high-quality MS/MS spectra matching to 795 peptide sequences, provided high-confidence evidence for the presence of 79 of the 80 predicted r-protein families in the ribosomes of *Arabidopsis*, including five previously undetected r-protein families: S29, S30, L29, L36a, and L39.

To date, the only predicted r-protein family yet to be detected in *Arabidopsis* ribosomes is the extremely basic (predicted pI of 13.4) and small (3.5 kDa) L41. The four paralogous L41 genes in the *A. thaliana* genome (At2g40205, At3g08520, At3g11120, and At3g56020) encode identical proteins with the amino acid sequence MRAKWKKKRMRRLKRKRRKMRQRSK. The strong conservation between eukaryotes of genes encoding this putative r-protein suggests that it is most likely a component of *Arabidopsis* ribosomes. X-ray crystallography has shown that the yeast ortholog of *Arabidopsis* L41 forms a bridge between the 40S and 60S subunits (Wilson and Doudna Cate, 2012), deep in the ribosome. For this reason, its non-detection so far in *Arabidopsis* ribosomes seems more likely to be due to technical limitations of the LC and MS detection approaches used rather than its absence in samples. Given that trypsin is not expected to yield useful peptides from L41, its detection in ribosomes is likely to require either targeted top-down LC-MS methods (Odintsova et al., 2003) or x-ray crystallography. Top-down LC-MS analyses are likely to require special chromatographic conditions as, with a predicted pI of 13.4, L41 is likely to be highly charged and therefore unlikely to be retained under typical reverse-phase pH conditions used in non-targeted top-down proteomics. Perhaps synthetic L41 peptides will prove useful as positive controls for method development and validation.

Carroll et al. (2008) provided strong MS/MS evidence to support the identification of 87 specific r-protein paralogs in total, including 32 not previously reported by Chang et al. (2005). These paralog-specific identifications confirmed previous reports of heterogeneity within S10, S12, S14, S15, S19, S24, S3a, S6, S7, Sa, P0, P1, P2, L10, L10a, L18a, L26, L27, L4, L7, L7a, and L8 (Chang et al., 2005) and provided strong evidence for previously unreported heterogeneity within a further 19 families, namely: S11, S2, S21, S25, S27a, S3, P1, L13a, L17, L18, L22, L23a, L28, L32, L35, L36, L37a, L5, and L6. In the case of six families – namely S15a, S16, S23, L19, and L31 – the paralog-specific detection of only a single family member contrasted with reports of Chang et al. (2005) of heterogeneity within those families. While the fact that Carroll et al. (2008) used much higher-stringency filters for their paralog-specific identifications should be considered when comparing these datasets, it is possible that these discrepancies were due at least in part to the tendencies of MALDI-TOF-MS and LC-ESI-Q-TOF-MS/MS to preferentially ionize different peptides (Stapels and Barofsky, 2004). Hence, each platform may have preferentially detected paralog-specific peptides from different r-proteins. Other possible contributors to differences in detected r-protein profiles are, of course, differences in tissue types and growth conditions. Given these considerations, the limited range of analytical techniques employed to date and the fact that new r-protein PTMs have only recently been detected (Turkina et al., 2011), it seems likely that the true extent of ribosome heterogeneity is greater than indicated by any individual study or, indeed, all the studies collectively. For the reader's convenience, Table S1 in Supplementary Material aligns and summarizes the r-protein identifications and post-translational modification detections reported to date across all of the major proteomic analyses of *A. thaliana* ribosomes (listed in **Table 1**).

TYPE II S15a PROTEINS: COMPONENTS OR CONTAMINANTS OF THE *ARABIDOPSIS* CYTOSOLIC RIBOSOME?

The *Arabidopsis* genome encodes for two evolutionarily distinct classes of S15a r-protein, commonly denoted type I and type II (Chang et al., 2005). There is strong evidence that the type II forms obtained functional mitochondrial targeting sequences to become part of the mitochondrial ribosome during the evolution of higher plants (Adams et al., 2002). Indeed, Carroll et al. (2008) detected paralog-specific peptides for both type II S15a proteins (S15aB and S15aE) in *Arabidopsis* mitochondrial ribosome preparations.

The detection of type II S15a sequences in their crude ribosomal pellet led Chang et al. (2005) to hypothesize that type II S15a proteins might be part of the cytosolic ribosome. However, an alternative explanation for this observation lies in the use of four membrane-solubilizing detergents (1% each of Triton X-100, Brij 35, Tween-40, and NP-40) – which would have dissolved mitochondrial membranes (Gurtubay et al., 1980) and released mitochondrial ribosomes and other mitochondrial proteins prior to pelleting of ribosomes by ultracentrifugation – in the ribosome extraction buffers of Chang et al. (2005). Hence, although other mitochondrial ribosomal proteins were not detected, the crude ribosome pellet of Chang et al. (2005) in which the type II proteins were detected (they were not reported in the sucrose gradient purified ribosomes) most probably contained at least a considerable portion of the mitochondrial ribosome population of their experimental cells-albeit at inherently low molar % levels reflecting their low cellular abundance relative to cytosolic ribosomes (Piques et al., 2009). A mitochondrial origin of the type II S15a proteins cannot, therefore, be ruled out on the basis of that analysis. In contrast, Carroll et al. (2008), who used a detergent-free ribosome extraction buffer containing 0.45 M mannitol as osmoticum to prevent osmotic bursting of organelles and subjected their tissue homogenates to $1500 \times g \times 5$ min, $16{,}000 \times g \times 15$ min, and $30{,}000 \times g \times 30$ min centrifugation steps to remove nuclei/chloroplasts, mitochondria and large aggregates of poorly defined insoluble materials prior to ultracentrifugation, did not observe a single peptide mapping to type II S15a proteins in their 80S ribosome preparations despite finding strong MS/MS evidence for proteotypic peptides from type I S15a proteins.

Methods are available to resolve 80S and 70S chloroplast ribosomes (Yamaguchi, 2011). However, given that the 70S and 80S ribosomes of *C. reinhardtii* sedimented closely on sucrose gradients (Yamaguchi et al., 2003) and that mitochondrial ribosomes from higher plants have been observed to sediment anywhere between 70S (Vasconcelos and Bogorad, 1971; Pinel et al., 1986) and 78S (Leaver and Harmey, 1973, 1976; Pring, 1974), the above observations highlight the importance of early fractionation steps, orthogonal to sucrose gradient purification, in obtaining pure cytosolic ribosomes required for confident discrimination of cytosolic and organellar ribosomal proteomes. In this author's view, this technical point is worth highlighting given the potential functional and evolutionary significance of parallel-targeting of r-proteins to multiple ribosomes in eukaryotic cells and the fact that just a few simple protocol modifications could greatly enhance the utility of future studies in addressing this important possibility.

"NON-RIBOSOMAL" RIBOSOME-ASSOCIATED PROTEINS

WHAT IS A NON-RIBOSOMAL PROTEIN?

Each of the major efforts to qualitatively define the *Arabidopsis* ribosomal proteome has reported the detection of "non-ribosomal" proteins in purified ribosomes (Chang et al., 2005; Giavalisco et al., 2005; Carroll et al., 2008; Hummel et al., 2012). However, the reporting of "non-ribosomal" proteins in purified ribosomes begs the question "how do we define ribosomal proteins?" As many new proteins not orthologous to the original set of 79 proteins originally labeled as core r-proteins by Wool et al. (1995) continue to be confidently detected in purified ribosome samples (Hummel et al., 2012), the classic view of ribosomes as a well-defined proteomic entity with a consistent stoichiometry is rapidly giving way to an increasingly fuzzy model of the ribosomal proteome in which a well-defined set core r-proteins (some of which may not always be associated with ribosomes) serve as a docking station for a poorly defined set of ribosome-associated regulatory proteins for which the natures and functions of their ribosome interactions are unclear (Gilbert, 2011; Xue and Barna, 2012). Due to the large number of ribosome-associated proteins that have now been reported and the fact that further experiments will be required to determine which associations represent *bona-fide* interactions as opposed to non-specific binding, an exhaustive list will not be provided here. Rather, the following sections highlight and discuss some examples for which *bona-fide* functions are either well established or worthy of further investigation based on independent information (which will be explained below).

RACK1 AND eIF6

Of all the ribosome-associated proteins detected so far in *A. thaliana* ribosomes, orthologs of the mammalian Receptor of Activated C Kinase (RACK1) are the most consistently detected. The RACK1A protein encoded by At1g18080 has been reported in all major proteomic surveys of *A. thaliana* ribosomes to date (Chang et al., 2005; Giavalisco et al., 2005; Carroll et al., 2008; Hummel et al., 2012). A second RACK1 ortholog (RACK1B) encoded by At1g48630 has also been detected in three independent studies (Chang et al., 2005; Carroll et al., 2008; Hummel et al., 2012) but RACK1C (At3g18130), the third of the three known RACK1 genes in the *A. thaliana* genome has not yet been detected in *A. thaliana* ribosomes. The close association of RACK1 with mammalian and yeast ribosomes has been known for some time and its role as a key regulatory component of the eukaryotic translation machinery is now well appreciated (Jakob et al., 2004). While RACK1 does not appear to be essential for translation in yeast, its absence decreases the efficiency of translation and steady state levels of numerous proteins (Shor et al., 2003). RACK1 is believed to play a key role in 80S ribosome assembly by directing the phosphorylation (by activated C Kinase) and release of eukaryotic Translation Initiation Factor 6 (eIF6) from the 60S subunit, thus allowing assembly of the 80S ribosome (Ceci et al., 2003; Guo et al., 2011).

In *A. thaliana*, a collection of studies have identified RACK1 as a key integrator and mediator of hormonal control over translation (Chen et al., 2006; Guo and Chen, 2008; Guo et al., 2009a,b, 2011). In particular, evidence suggests that abscisic acid (ABA) may down-regulate translation generally by inhibiting the transcriptional expression of RACK1 and eIF6 mRNAs,

although the mechanism by which ABA controls the levels of these mRNAs is currently unclear (Guo et al., 2011). Interestingly, the amount of RACK1A (but not RACK1B) associated with ribosomes/polysomes increased significantly in response to sucrose feeding in *A. thaliana* (Hummel et al., 2012) – a response that is known to involve ABA signaling pathways (Laby et al., 2000). The *Arabidopsis* genome encodes two homologs of eIF6 – At3g55620 (eIF6A) and At2g39820 (eIF6B). While both proteins have been demonstrated to interact physically with RACK1 (Guo et al., 2011), only the eIF6A protein has been detected in the ribosomes of *A. thaliana* leaf and suspension cells (Carroll et al., 2008; Hummel et al., 2012). This is consistent with mRNA expression patterns which indicate that while eIF6A mRNA is expressed ubiquitously, eIF6B mRNA is mainly expressed in flower buds, stamens, and pollen (Guo et al., 2011). It remains to be seen whether eIF6B is present in the ribosomes of these tissues.

20S PROTEASOME

The 20S proteasome forms part of the 26S proteasome complex responsible for the proteolysis of many proteins (particularly those carrying poly ubiquitin tails) in eukaryotic cells (Yang et al., 2004). Subunits of the 20S proteasome were detected in polysomal bands on sucrose density gradients by Chang et al. (2005) and also in crude ribosomal pellet by Giavalisco et al. (2005) but not in the highly purified ribosome samples of Carroll et al. (2008). Because of the high abundance and similar sedimentation coefficient of the proteasome complex (when associated with other complexes), there has been some uncertainty whether the association of the proteasome with ribosomes was due to a *bona-fide in vivo* interaction or simply a non-specific interaction between abundant complexes or simple co-sedimentation (Chang et al., 2005). Indeed, the fact that proteasome subunits were not reported in epitope-tag purified *A. thaliana* ribosomes in the same manner as RACK1 (Chang, 2006) suggests that if a *bona-fide* interaction between the proteasome complex and *A. thaliana* ribosomes exists, it is more labile than the interaction between ribosomes and RACK1. That said, given that the proteasome is thought to play a role in degrading defective ribosomal products (proteins that result from errors in translation or folding) representing some 30% of newly synthesized proteins, it would seem efficient to have proteasome complexes localized at the point of protein synthesis to prevent the escape of potentially toxic defective proteins into the cytoplasm. Another possible explanation may lie in the major role played by the proteasome in ribosome biogenesis (Stavreva et al., 2006).

FERRITIN

Only four ribosome-associated proteins were detected in the ribosome preparations of Carroll et al. (2008). Three of these – namely RACK1A, RACK1B, and eIF6 – have been discussed above. The fourth ribosome-associated protein detected was identified as FERRITIN 3 (FER3; At3g56090). The fact that so few ribosome-associated proteins were detected in these ribosomes and the striking absence of obvious abundant non-specific binding proteins suggests that FER3 was indeed tightly associated with these ribosomes. The FER3 protein has also been detected in small polysome fractions isolated from the leaves of *A. thaliana* plants in the dark (Piques et al., 2009). This is intriguing given that FER3 is a nuclear-encoded chloroplast-targeted protein that, according to the SUBA database (Heazlewood et al., 2007), has been repeatedly detected in chloroplast preparations by mass spectrometry. In humans, ferritin has been demonstrated to regulate folate metabolism by controlling the translation of cytosolic serine hydroxymethyltransferase (cSHMT) via binding to ferritin-responsive internal ribosome entry site (IRES) in the 5′UTR of the cSHMT mRNA (Woeller et al., 2007). The H ferritin involved was also shown to interact physically with the mRNA-binding protein CUGBP1 which is known to interact with the α and β subunits of eukaryotic translation initiation factor 2 (eIF2; Woeller et al., 2007). Together, the above observations suggest that the existence of similar mechanisms involving FER3 in *A. thaliana* should be investigated. The link with chloroplasts is particularly intriguing since it is possible to imagine a mechanism whereby FER3 mediates coordination between the translational activity of cytosolic ribosomes and the function of chloroplasts in response to iron-based signals.

POST-TRANSLATIONAL MODIFICATIONS AND THE NEED FOR "TOP-DOWN" APPROACHES

Eukaryotic ribosomes are well-known to be rich in many kinds of PTMs. The diversity and conservation of PTMs of r-proteins observed across different eukaryote lineages has been reviewed elsewhere (Carroll et al., 2008) and will not be covered again here. Instead, the discussion of PTMs in this review will focus on providing an updated overview of current knowledge concerning PTMs of *A. thaliana* r-proteins. The different types of PTMs detected in *A. thaliana* cytosolic ribosomes include initiator methionine removal, N-terminal acetylation, serine phosphorylation, lysine mono-, and tri-methylation, and N-terminal proline dimethylation (Chang et al., 2005; Carroll et al., 2008; Turkina et al., 2011). Specific PTM reports are listed in **Table 2**. Some particularly important issues concerning *A. thaliana* r-protein PTMs are discussed below.

That phosphorylation sites exist on S6 and the acidic stalk P proteins has been well established for some time, primarily from work in *Zea mays* (Szick-Miranda and Bailey-Serres, 2001; Williams et al., 2003). The conservation of these modifications in *A. thaliana* ribosomes has been confirmed more recently (Chang et al., 2005; Carroll et al., 2008; Turkina et al., 2011). However, new phosphorylation sites continue to emerge as new tissues are analyzed and new methods of analysis are employed. For example, Carroll et al. (2008) recently reported a previously undiscovered phosphorylation site on L13(At3g49010). It should be noted that this r-protein is not homologous to the human r-protein L13a which has been shown to act as an mRNA-binding translational suppressor upon being released from human ribosomes by phosphorylation following treatment of cells with interferon-γ (Mazumder et al., 2003). Interestingly, phosphorylated L13 was not detected in a recent quantitative phosphoproteomic analysis of *A. thaliana* leaf cytosolic ribosome despite the detection of previously undetected phosphorylation sites at Ser_{231} of S6 and Ser_{58} of L29A (Turkina et al., 2011). Phosphorylation of the human ortholog of L29 has also been detected (Molina et al., 2007; Wang et al., 2008). Together, these observations highlight the likely plasticity of *A. thaliana* L13 and L29A phosphorylation

Table 2 | Post-translational modifications reported to date in *A. thaliana* ribosomal proteins.

Family	Loci	Post-translational modifications	Reference
L10a	At1g08360 (L10aA)	$^{90}K_{m3}$	a
L10a	At2g27530 (L10aB)	$^{90}K_{m3}$	a
L10a	At1g08360 (L10aA), At2g27530 (L10aB), At5g22440 (L10aC)	−Met, $^{N\text{-term}}S_{Ac}$	a
L12	At2g37190 (L12A), At3g53430 (L12B), At5g60670 (L12C)	−Met, $^{N\text{-term}}P_{m2}$, $^{3}K_{m3}$	a
L12	At2g37190 (L12A), At3g53430 (L12B), At5g60670 (L12C)	−Met, $^{N\text{-term}}P_{Ac}$	b
L13	At3g49010 (L13B)	$^{137}S_{phospho}$	a
L15	At4g16720 (L15A)	$^{N\text{-term}}G_{Ac}$	b
L18	At3g05590 (L18B), At5g27850 (L18C)	−Met	a
L21	At1g09590 (L21A), At1g09690 (L21C), At1g57660 (L21E), At1g57860 (L21G)	−Met	a
L28	At2g19730 (L28A), At4g29410 (L28C)	$^{N\text{-term}}A_{Ac}$	a,b
L28	At4g29410 (L28C)	$^{N\text{-term}}A_{Ac}$	a
L29	At3g06700 (L29A)	$^{58}S_{phospho}$	c
L32	At4g18100 (L32A), At5g46430 (L32B)	−Met	a
L36	At5g02450 (L36C)	−Met	a
L36a	At3g23390 (L36aA), At4g14320 (L36aB)	$^{55}K_{m1}$	a
P0	At3g09200 (P0B)	$^{305}S_{phospho}$	c
P0	At3g11250 (P0C)	$^{305}S_{phospho}$	a
P1	At1g01100 (P1A), At4g00810 (P1B), At5g47700 (P1C)	$^{102/103}S_{phospho}$	a,c
P2	At2g27720 (P2A), At2g27710 (P2B), At3g44590 (P2D)	$^{105}S_{phospho}$	a
P3	At4g25890 (P3A), At5g57290 (P3B)	$^{107}S_{phospho}$	a
S2	At2g41840 (S2C)	$^{273}S_{phospho}$	c
S3	At2g31610 (S3A), At5g35530 (S3C)	−Met, $^{N\text{-term}}A_{Ac}$	a
S5	At2g37270 (S5A)	−Met, $^{N\text{-term}}A_{Ac}$	a,b
S5	At3g11940 (S5B)	−Met, $^{N\text{-term}}A_{Ac}$	a
S6	At4g31700 (S6A)	$^{240}S_{phospho}$	a,b,c
S6	At5g10360 (S6B)	$^{240}S_{phospho}$	a,c
S6	At4g31700 (S6A), At5g10360 (S6B)	Mono-, di-, tri-, and tetra-phospho at unknown sites in C terminal region	b
S6	At4g31700 (S6A), At5g10360 (S6B)	$^{231}S_{phospho}$	c
S7	At1g48830 (S7A)	$^{N\text{-term}}M_{Ac}$	a
S15	At1g04270 (S15A), At5g09510 (S15D)	−Met, $^{N\text{-term}}A_{Ac}$	a,b
S16	At2g09990 (S16A), At5g18380 (S16C)	−Met, $^{N\text{-term}}A_{Ac}$	b
S18	At1g22780 (S18A), At1g34030 (S18B), At4g09800 (S18C)	−Met, $^{N\text{-term}}S_{Ac}$	a,b
S20	At3g45030 (S20A), At5g62300 (S20C)	−Met, $^{N\text{-term}}A_{Ac}$	a
S20	At3g47370 (S20B)	−Met, $^{N\text{-term}}A_{Ac}$	a,b
S21	At3g53890 (S21B)	$^{N\text{-term}}M_{Ac}$	a,b
S27	At2g45710 (S27A), At3g61110 (S27B)	−Met	a
Sa	At1g72370 (SaA)	−Met, $^{N\text{-term}}A_{Ac}$	a

Reports of post-translational modifications of A. thaliana r-proteins are listed in order of r-protein family. Each row corresponds to an individual report of a particular modified peptide. The paralog(s) to which each peptide may be mapped theoretically is indicated in column "Loci" with the listing of several paralogs indicating ambiguity with respect to which paralog carried the indicated modification and the listing of a single paralog indicating that the detected modified peptide was specific to a single paralog. In the column "Post-translational modifications," "–Met" indicates removal of the initiator methionine while other modification(s) reported to be associated with the peptide are indicated in the format $^{\text{residue position}}X_{\text{modification type}}$ *where X is the one-letter code of the modified amino acid residue. All residue positions are given with the initiator methionine as position 1 regardless of whether this methionine is removed. Abbreviations used include: N-term, N terminus; Ac, Acetyl; phospho, phosphorylation. Methylation modification types are abbreviated as mY where Y, the number of methyl groups added. In the references column, (a) Carroll et al. (2008), (b) Chang et al. (2005), and (c) Turkina et al. (2011).*

and suggest that potential roles of L13 and L29 phosphorylation in translational control need targeted research. The extent to which these and other modifications are conserved across different plant species remains to be seen. However, given the divergence of PTMs seen between the major eukaryote lineages (Carroll et al., 2008), the possibility that the gain or loss of specific r-protein PTM sites by different plant species during plant evolution could have played a role in ecological specialization of plants is too tantalizing not be explored.

Another important point relates to the potential role of post-translational modification in ribosome heterogeneity. Giavalisco

et al. (2005) suggested that the observation that >45% of specific r-proteins were detected in 2–13 spots was indicative of each being present in various modification states. In contrast, none of the 30 modification sites identified by Carroll et al. (2008) were also detected in an unmodified form. One possible explanation for the greater diversity of modification states observed by Giavalisco et al. (2005) may lay in the higher diversity of cell types expected in leaves compared to relatively homogeneous and undifferentiated cell cultures analyzed by Chang et al. (2005) and Carroll et al. (2008). Another possibility is that the extra spots observed by Giavalisco et al. (2005) were multiply phosphorylated forms that generate multiply phosphorylated peptides that are notoriously difficult to detect directly by mass spectrometry (Choi et al., 2008). Proteolytic degradation, either via natural *in vivo* mechanisms or during analysis, may have also contributed detection of r-proteins across multiple spots (Finnie and Svensson, 2002; Vohradsky et al., 2008).

The application of "top-down" proteomics techniques involving the direct analysis of intact proteins by LC/MS without protease digestion will be invaluable in resolving the issues discussed above (Zhang and Ge, 2011; Zhou et al., 2011). Top-down approaches complement bottom-up approaches by revealing modification states – such as multiple modifications at distal sites or proteolytic protein truncation – that are masked by protease digestion. Similarly, modifications that are difficult to detect in peptide form (e.g., multiply phosphorylated peptides) may be more amenable to detection in the form of modified whole proteins. Top-down approaches have been used extensively to study the composition and PTMs of mammalian (Louie et al., 1996; Odintsova et al., 2003; Yu et al., 2005) and yeast (Arnold et al., 1999; Lee et al., 2002) ribosomes. However, top-down analyses of plant ribosomes have still to be carried out.

A recent quantitative phosphoproteomic analysis has confirmed that phosphorylation not only contributes to cytosolic ribosome heterogeneity in *Arabidopsis*, but the relative abundance of different phosphorylated forms of S6 and L29 change during the diurnal cycle (Turkina et al., 2011). The levels of mRNA transcripts encoding the phosphorylated acidic stalk P proteins P1, P2A, P2B, and P3 have been shown to be highly variable across different organs and tissues in *Zea mays* (Szick-Miranda and Bailey-Serres, 2001). Importantly, while levels of P1, P2A, and P2B (but not P3) proteins in ribosomal extracts were also shown to be variable across different organs and tissues, these levels were poorly correlated with the observed variation in mRNA levels, clearly demonstrating the importance of proteome level studies. This study also demonstrated that the phosphorylation levels of the P1, P2A, and P3 proteins of root tip ribosomes decreased under anoxic conditions.

THE ENZYMES THAT MODIFY *ARABIDOPSIS* R-PROTEINS ARE LARGELY UNKNOWN

The identification of enzymes responsible for the post-translation modification of ribosomal proteins has progressed much more slowly in *Arabidopsis* than in other eukaryotic systems. While the kinase responsible for *Arabidopsis* S6 phosphorylation has been known for some time (Mizoguchi et al., 1995; Mahfouz et al., 2006), little is known about the enzymes responsible for other modifications of *A. thaliana* r-proteins. In contrast, a variety of *N*-methyltransferases and acetyltransferases responsible for the modification of r-proteins have been identified in yeast and humans (Arnold et al., 1999; Bachand and Silver, 2004; Porras-Yakushi et al., 2005, 2008; Ren et al., 2010; Webb et al., 2010a,b, 2011; Forte et al., 2011). Proteomic analysis of ribosomes isolated from *A. thaliana* mutants perturbed in orthologous or homologous candidate genes encoding potential ribosome-modifying enzymes will almost certainly be a fruitful line of research.

THE TRANSITION FROM QUALITATIVE TO QUANTITATIVE RIBOSOMAL PROTEOMICS

With an abundance of qualitative proteomics data suggesting the extreme heterogeneity of cytosolic ribosome populations from whole plant tissues, ribosomal proteomics is now heavily focused on understanding the spatiotemporal distribution and physiological function of this heterogeneity. Ribosome heterogeneity could and probably does occur at many different spatiotemporal scales – from slow developmental changes or constitutive differences in the ribosome populations of distinct organs to rapid changes in minor subcellular populations of ribosomes within single cells. Hence, a major task that will be important for reverse-engineering the physiological function of ribosome heterogeneity will be the use of quantitative proteomic approaches to correlate variations in the relative abundances and modification states of different r-protein paralogs across tissues, cell types, subcellular-fractions, developmental stages, and environmental and genetic perturbations with ribosome properties and processes upstream and downstream of ribosomes. With this goal in mind, Hummel et al. (2012) and colleagues recently demonstrated, through a highly impressive large-scale label-free MS^E quantitative proteomic approach, that the paralog composition (particularly in RPS3aA, RPS5A, RPL8B, and RACK1) of *A. thaliana* leaf cytosolic ribosomes responded significantly to sucrose feeding – a treatment that elicits dramatic changes in gene expression. Similar experiments involving different treatments seem likely to reveal even broader ribosome dynamics involving a wider range of r-proteins.

FROM FORM TO FUNCTION: GENETIC STUDIES OF R-PROTEIN FUNCTION IN *ARABIDOPSIS*

While the large-scale use of qualitative and quantitative proteomics approaches to study the composition and dynamics of ribosomes will be essential for elucidating their role in plant physiology, unraveling the precise functions of specific r-proteins will also be greatly assisted by functional genetic studies. A considerable number of genetic studies involving the characterization of *A. thaliana* r-protein mutants have already emerged (see **Table 3** for a list of studies, mutants, and phenotypes). Together, these studies highlight, perhaps unsurprisingly, the important role of ribosomes and translation in many aspects of plant development (Byrne, 2009). While the leaf abaxialisation phenotypes of the various r-protein/ASYMMETRIC LEAVES double mutants appear to be somewhat qualitatively independent of which particular r-protein gene is disrupted, the relative severity of different phenotypic sub-elements does seem to depend to some degree on the identity

Table 3 | Published *A. thaliana* r-protein mutants and their phenotypes.

r-Protein (AGI)	Type of mutant	Mutant name	Phenotype	Reference
L5A(At3G25520)	*rpl5A/as2* double EMS mutant	*ae6-1 as2-101*	Abaxialised leaves. Increased number of lotus- and needle-like leaves	Yao et al. (2008)
L5A(At3G25520)	*rpl5A/as1* double EMS mutant	*ae6-1 as1-101*	Abaxialised leaves. Increased number of lotus- and needle-like leaves	Yao et al. (2008)
L5A(At3G25520)	*rpl5A* single EMS mutant	*ae6-1*	Normal wildtype phenotype	Yao et al. (2008)
L5A(At3g25520)	*rpl5A/as1* double EMS mutant	*pgy3*	Dramatic ectopic lamina outgrowths on the adaxial side of the leaf	Pinon et al. (2008)
L5B(At5G39740)	*rpl5B* single EMS mutant	*rpl5b*	Pale green leaves	Yao et al. (2008)
L5B(At5G39740)	*rpl5Bas2-101* double EMS mutant	*rpl5b as2-101*	Abaxialised leaves. Almost all leaves needle-like	Yao et al. (2008)
L9C(At1g33140)	*rpl9C/as1* double EMS mutant	*pgy2*	Dramatic ectopic lamina outgrowths on the adaxial side of the leaf	Pinon et al. (2008)
L10A(At1G14320)	*rpl10A* homozygous knockout		Lethal	Ferreyra et al. (2010b)
L10A(At1G14320)	heterozygous T-DNA insertion causing reduced mRNA levels		Conditional translational deficiency under UV-B stress	Ferreyra et al. (2010b)
L10A(At1g14320)	*rpl10a/acl5* double EMS mutant	*acl5-1*	Semi-dominant mutation in *L10A* rescues the severe dwarf phenotype resulting from mutation of acl5 which encodes thermospermine synthase	Imai et al. (2008)
L10B(At1G26910)	heterozygous T-DNA insertion causing reduced mRNA levels		Abnormal growth including reduced size, narrow, and pointed first leaves, 77% reduction in seedling leaf size, smaller but similar numbers of leaves until flowering time at which point the mutant continued producing leaves and started showing increased rosette branching. Shorter primary roots and reduced silique length were also observed	Ferreyra et al. (2010b)
L10aB(At2g27530)	*rpl10aB/as1* double EMS mutant	*piggyback1 (pgy1)*	Dramatic ectopic lamina outgrowths on the adaxial side of the leaf	Pinon et al. (2008)
L23aA(At2G39460)	RNAi knockdown		Growth retardation, irregular root and leaf morphology, abnormal phyllotaxy, and vasculature and loss of apical dominance	Degenhardt and Bonham-Smith (2008)
L23aB(At3G55280)	RNAi knockdown		No visible phenotype	Degenhardt and Bonham-Smith (2008)
L24B(At3G53020)	*rpl24B/as2-101* double EMS mutant	*stv1 as2-101*	Abaxialised leaves. Almost all leaves needle-like	Yao et al. (2008)
L24B(At3G53020)	EMS mutant	*short valve (stv1)*	Pale green leaves	Yao et al. (2008)
L24B(At3g53020)	EMS mutant	*short valve (stv1)*	Basal region of ovary shortened. Gynophore elongated	Nishimura et al. (2004, 2005)
L27aC(At1g70600)	*rpl27aC/as1* double EMS mutant	*pgy6/rpl27ac-1d*	Altered shoot development, including leaf patterning, inflorescence and floral meristem function, and seed set. A temporal delay in initiation and outgrowth of cotyledon primordia leads to development of an enlarged globular embryo prior to apical domain patterning	Szakonyi and Byrne (2011)

(Continued)

Table 3 | Continued

r-Protein (AGI)	Type of mutant	Mutant name	Phenotype	Reference
L27aC(At1g70600)	*rpl27aC* knockdown mutant (T-DNA insertion in promoter) – homozygous	*rpl27ac-2*	Pointed and serrated leaves	Szakonyi and Byrne (2011)
L27aC(At1g70600)	*rpl27aC* knockdown mutant (T-DNA insertion in promoter) – heterozygous	*rpl27ac-2/+*	No visible shoot phenotype	Szakonyi and Byrne (2011)
L27aC(At1g70600)	*rpl27aC* knockout mutant (T-DNA insertion in 5′-UTR) – homozygous	*rpl27ac-3*	Pointed and serrated leaves (less so than in *rpl27ac-2*)	Szakonyi and Byrne (2011)
L27aC(At1g70600)	*rpl27aC* knockout mutant (T-DNA insertion in 5′-UTR) – heterozygous	*rpl27ac-3/+*	No visible shoot phenotype	Szakonyi and Byrne (2011)
L28A(At2G19730)	*rpl28A/as2* double EMS mutant	*ae5-1 as2-101*	Abaxialised leaves. Increased number of lotus- and needle-like leaves	Yao et al. (2008)
L28A(At2G19730)	*rpl28A/as1* double EMS mutant	*ae5-1 as1-101*	Abaxialised leaves. Increased number of lotus- and needle-like leaves	Yao et al. (2008)
L28A(At2G19730)	*rpl28A* single EMS mutant	*ae5-1*	Pale green leaves. First few leaves slightly longer than wildtype	Yao et al. (2008)
S5A(At2g37270)	T-DNA insertion knockout	*Arabidopsis Minute-like 1 (aml1)*	Most cell-division processes delayed or disturbed in the heterozygous mutant. Development is completely arrested at an early embryonic stage in the homozygous mutant	Weijers et al. (2001)
S6B(At5g10360)	Antisense knockdown		Reduced apical dominance and irregular positioning of leaves and flowers	Morimoto et al. (2002)
S10B(At5g41520)	*rps10B-1/max2-1* double EMS mutant		Recessive mutation in *S10B* suppresses the excessive branching phenotype of *max2-1*	Stirnberg et al. (2012)
S13A(At3g60770)	Transposon-mediated knockout	*Pointed First Leaf 2 (pfl2)*	Aberrant leaf and trichome morphology, retarded root growth, and late flowering. Reproductive growth otherwise not altered. Reduced number of palisade cells. No phenotypic changes observed when crossed with a S18 mutant, *pfl1*, having a similar phenotype	Ito et al. (2000)
S15aE(At4g29430)	T-DNA insertion knockdown	*rps15aE-mut1*	Greater mean rosette radii and leaf areas and longer roots	Szick-Miranda et al. (2010)
S18A(At1g22780)	T-DNA insertion knockout	*Pointed First Leaf 1 (pfl1)*	Pointed first leaves, reduced fresh weight, and growth retardation	Van Lijsebettens et al. (1994)
S27A(At2g45710)	T-DNA insertion knockout	*rps27A*	Conditional growth inhibition under genotoxic stress (growth on methylmethanesulfonate-containing medium). Impairment in mRNA degradation after UV irradiation	Revenkova et al. (1999)

EMS, ethylmethanesulfonate; UV, Ultra-violet.

of the disrupted r-protein. These observations support the suggestion that different r-proteins do contribute differently to leaf development (Horiguchi et al., 2011).

In addition to the plethora of developmental defects observed in many r-protein mutants, other interesting phenotypes associated with r-protein mutants include the conditional translational deficiency phenotype of L10A mutants exposed the UV-B stress (Ferreyra et al., 2010a,b) and the conditional growth inhibition of S27A mutants grown on genotoxic methyl methane sulfonate-containing medium (Revenkova et al., 1999). Also interesting is a defect in mRNA degradation seen in S27A mutants exposed to UV light (Revenkova et al., 1999).

FUTURE STRATEGIES TO REVERSE-ENGINEER THE PHYSIOLOGICAL ROLE OF RIBOSOME HETEROGENEITY

One of the fundamental goals of ribosome research is to understand the role of ribosome heterogeneity in translational specialization and control. Indeed, the fact that different mRNA profiles are associated with polysomes isolated from different cell types (Mustroph et al., 2009) or under different environmental conditions (Branco-Price et al., 2005, 2008; Piques et al., 2009; Liu et al., 2012), combined with the fact that ribosome heterogeneity is also under environmental and developmental control (Szick-Miranda and Bailey-Serres, 2001; Branco-Price et al., 2005, 2008; Turkina et al., 2011; Hummel et al., 2012) suggests that there may well be a link. The existence of so many paralogs of each r-protein family in the genome of *Arabidopsis* means that more than 10^{34} theoretical r-protein combinations could potentially be formed *in-vivo* (Hummel et al., 2012). Such incredible capacity for ribosome heterogeneity makes the notion of a ribosome "code" – whereby different ribosomes are optimized for or dedicated to the translation of specific mRNAs (Komili et al., 2007) – particularly alluring. However, proving the existence or otherwise of a ribosome code will be extremely challenging.

It will probably never be possible to resolve and characterize every single one of the 10^{34} potential ribosomes. However, we may be able to significantly deepen our understanding of the role of changes or differences in ribosome composition in translational specificity by fractionating ribosome and polysome populations, analyzing the fractions by translatomic (Mustroph et al., 2009) and quantitative proteomic (Hummel et al., 2012) approaches and then mining the resulting data for correlations between ribosome composition and translational behavior. The phosphorylation of S6 has already been correlated with differential mRNA recruitment to ribosomes (Scharf and Nover, 1982; Turck et al., 2004). However, global integrated proteomic and translatomic analyses across a much wider range of ribosome types will be essential if we hope to properly decipher the ribosome code and resolve causations from correlations with a high degree of confidence.

The correlative approach described above depends on building up sufficient covariance between ribosome composition and translatome profiles. This variation could be obtained in a wide variety of ways. The quantitative proteomic analysis of polysomes from different cell types for which translatome data are already available (Mustroph et al., 2009) may be a fruitful place to start. However, other possibilities might include the separation of free cytoplasmic polysomes and polysomes bound to various subcellular membrane structures such as the endoplasmic reticulum, mitochondria and chloroplast surfaces (Suissa and Schatz, 1982; Kaltimbacher et al., 2006; Fu et al., 2012). Alternatively, complex polysome populations might be fractionated directly by non-denaturing preparative separation techniques such as free-flow electrophoresis which separates protein complexes and even organelles on the basis of surface charge (Wagner, 1989).

More targeted approaches might include the affinity purification of ribosomes translating "bait" mRNAs containing aptamers enabling their selective immunopurification (along with the ribosomes translating them). Quantitative proteomic comparisons of these ribosomes with those pulled down using control mRNAs might help reveal regulatory elements within test mRNAs that promote their recruitment by polysomes while also revealing the types of ribosomes they attract.

Another targeted approach might be to genetically perturb the expression of particular r-proteins and then monitor and cross-correlate changes in ribosome composition with changes in the translatome. However, given the high potential for pleiotropic effects when disrupting translation machinery, perhaps an appropriate approach would be to employ inducible, possibly cell type-specific, silencing, or over expression of r-proteins so that time-course profiling of the ribosomal proteome and translatome can be used to distinguish primary (early) and secondary (later) effects of specific r-protein perturbations.

Another potentially powerful approach to understand how changes in ribosome composition are related to changes in translation may be to apply next-generation ribosome footprinting whereby the exact locations of ribosomes on transcripts is determined by deep sequencing the regions of transcripts that are protected by ribosomes (Ingolia et al., 2009; Lee et al., 2012). Combining this technique with polysome fractionation, genetic-, and environmental-perturbation and quantitative proteomics may reveal, on a genome-wide scale and with single-base-pair resolution, how ribosome composition is related to mRNA occupancy and, through motif analysis, the affinity of ribosomes for particular mRNA sequence elements.

Clearly, as far as the structures and functions of ribosomes are concerned, there are still many more questions than answers. Despite being discovered so long ago, ribosomes remain one of the most interesting and crucially important targets for basic and applied biological research. However, being the complex natural nano-machines that they are, they do not give up their secrets easily and will no doubt remain the focus of many research careers well into the foreseeable future.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at http://www.frontiersin.org/Plant_Proteomics/10.3389/fpls.2013.00032/abstract

REFERENCES

Adams, K. L., Daley, D. O., Whelan, J., and Palmer, J. D. (2002). Genes for two mitochondrial ribosomal proteins in flowering plants are derived from their chloroplast or cytosolic counterparts. *Plant Cell* 14, 931–943.

Arnold, R. J., Polevoda, B., Reilly, J. P., and Sherman, F. (1999). The action of N-terminal acetyltransferases on yeast ribosomal proteins. *J. Biol. Chem.* 274, 37035–37040.

Bachand, F., and Silver, P. A. (2004). PRMT3 is a ribosomal protein methyltransferase that affects the cellular levels of ribosomal subunits. *EMBO J.* 23, 2641–2650.

Bain, J. D., Switzer, C., Chamberlin, R., and Bennert, S. A. (1992). Ribosome-mediated incorporation of a non-standard amino acid into a peptide through expansion of the genetic code. *Nature* 356, 537–539.

Barakat, A., Szick-Miranda, K., Chang, I.-F., Guyot, R., Blanc, G., Cooke, R., et al. (2001). The organization of cytoplasmic ribosomal protein genes in the *Arabidopsis* genome. *Plant Physiol.* 127, 398–415.

Benner, S. A. (1994). Expanding the genetic lexicon: incorporating non-standard amino acids into proteins by ribosome-based synthesis. *Trends Biotechnol.* 12, 158–163.

Branco-Price, C., Kaiser, K. A., Jang, C. J., Larive, C. K., and Bailey-Serres, J. (2008). Selective mRNA translation coordinates energetic and metabolic adjustments to cellular oxygen deprivation and reoxygenation in *Arabidopsis thaliana*. *Plant J.* 56, 743–755.

Branco-Price, C., Kawaguchi, R., Ferreira, R. B., and Bailey-Serres, J. (2005). Genome-wide analysis of transcript abundance and translation in *Arabidopsis* seedlings subjected to oxygen deprivation. *Ann. Bot.* 96, 647–660.

Byrne, M. E. (2009). A role for the ribosome in development. *Trends Plant Sci.* 14, 512–519.

Carroll, A. J., Heazlewood, J. L., Ito, J., and Millar, A. H. (2008). Analysis of the *Arabidopsis* cytosolic ribosome proteome provides detailed insights into its components and their post-translational modification. *Mol. Cell. Proteomics* 7, 347–369.

Ceci, M., Gaviraghi, C., Gorrini, C., Sala, L. A., Offenhauser, N., Carlo Marchisio, P., et al. (2003). Release of eIF6 (p27BBP) from the 60S subunit allows 80S ribosome assembly. *Nature* 426, 579–584.

Chang, I.-F. (2006). Mass spectrometry-based proteomic analysis of the epitope-tag affinity purified protein complexes in eukaryotes. *Proteomics* 6, 6158–6166.

Chang, I.-F., Szick-Miranda, K., Pan, S., and Bailey-Serres, J. (2005). Proteomic characterization of evolutionarily conserved and variable proteins of *Arabidopsis* cytosolic ribosomes. *Plant Physiol.* 137, 848–862.

Chen, J.-G., Ullah, H., Temple, B., Liang, J., Guo, J., Alonso, J. M., et al. (2006). RACK1 mediates multiple hormone responsiveness and developmental processes in *Arabidopsis*. *J. Exp. Bot.* 57, 2697–2708.

Choi, H., Lee, H.-S., and Park, Z.-Y. (2008). Detection of multiphosphorylated peptides in LC-MS/MS analysis under low pH conditions. *Anal. Chem.* 80, 3007–3015.

Degenhardt, R. F., and Bonham-Smith, P. C. (2008). *Arabidopsis* ribosomal proteins RPL23aA and RPL23aB are differentially targeted to the nucleolus and are disparately required for normal development. *Plant Physiol.* 147, 128–142.

Ferreyra, M. L. F., Biarc, J., Burlingame, A. L., and Casati, P. (2010a). *Arabidopsis* L10 ribosomal proteins in UV-B responses. *Plant Signal. Behav.* 5, 1222–1225.

Ferreyra, M. L. F., Pezza, A., Biarc, J., Burlingame, A. L., and Casati, P. (2010b). Plant L10 ribosomal proteins have different roles during development and translation under ultraviolet-B stress. *Plant Physiol.* 153, 1878–1894.

Finnie, C., and Svensson, B. (2002). Proteolysis during the isoelectric focusing step of two-dimensional gel electrophoresis may be a common problem. *Anal. Biochem.* 311, 182–186.

Forte, G. M. A., Pool, M. R., and Stirling, C. J. (2011). N-terminal acetylation inhibits protein targeting to the endoplasmic reticulum. *PLoS Biol.* 9:e1001073. doi:10.1371/journal.pbio.1001073

Fu, S., Fan, J., Blanco, J., Gimenez-Cassina, A., Danial, N. N., Watkins, S. M., et al. (2012). Polysome profiling in liver identifies dynamic regulation of endoplasmic reticulum translatome by obesity and fasting. *PLoS Genet.* 8:e1002902. doi:10.1371/journal.pgen.1002902

Giavalisco, P., Wilson, D., Kreitler, T., Lehrach, H., Klose, J., Gobom, J., et al. (2005). High heterogeneity within the ribosomal proteins of the *Arabidopsis thaliana* 80S ribosome. *Plant Mol. Biol.* 57, 577–591.

Gilbert, W. V. (2011). Functional specialization of ribosomes? *Trends Biochem. Sci.* 36, 127–132.

Guo, J., and Chen, J.-G. (2008). RACK1 genes regulate plant development with unequal genetic redundancy in *Arabidopsis*. *BMC Plant Biol.* 8:108. doi:10.1186/1471-2229-8-108

Guo, J., Wang, J., Xi, L., Huang, W.-D., Liang, J., and Chen, J.-G. (2009a). RACK1 is a negative regulator of ABA responses in *Arabidopsis*. *J. Exp. Bot.* 60, 3819–3833.

Guo, J., Wang, S., Wang, J., Huang, W.-D., Liang, J., and Chen, J.-G. (2009b). Dissection of the relationship between RACK1 and heterotrimeric G-proteins in *Arabidopsis*. *Plant Cell Physiol.* 50, 1681–1694.

Guo, J., Wang, S., Valerius, O., Hall, H., Zeng, Q., Li, J.-F., et al. (2011). Involvement of *Arabidopsis* RACK1 in protein translation and its regulation by abscisic acid. *Plant Physiol.* 155, 370–383.

Gurtubay, J. I. G., Goñi, F. M., Gómez-Fernández, J. C., Otamendi, J. J., and MacArulla, J. M. (1980). Triton X-100 solubilization of mitochondrial inner and outer membranes. *J. Bioenerg. Biomembr.* 12, 47–70.

Heazlewood, J. L., Verboom, R. E., Tonti-Filippini, J., Small, I., and Millar, A. H. (2007). SUBA: the *Arabidopsis* subcellular database. *Nucleic Acids Res.* 35, D213–D218.

Horiguchi, G., Molla-Morales, A., Perez-Perez, J. M., Kojima, K., Robles, P., Ponce, M. R., et al. (2011). Differential contributions of ribosomal protein genes to *Arabidopsis thaliana* leaf development. *Plant J.* 65, 724–736.

Hummel, M., Cordewener, J. H. G., De Groot, J. C. M., Smeekens, S., America, A. H. P., and Hanson, J. (2012). Dynamic protein composition of *Arabidopsis thaliana* cytosolic ribosomes in response to sucrose feeding as revealed by label free MSE proteomics. *Proteomics* 12, 1024–1038.

Imai, A., Komura, M., Kawano, E., Kuwashiro, Y., and Takahashi, T. (2008). A semi-dominant mutation in the ribosomal protein L10 gene suppresses the dwarf phenotype of the acl5 mutant in *Arabidopsis thaliana*. *Plant J.* 56, 881–890.

Ingolia, N. T., Ghaemmaghami, S., Newman, J. R. S., and Weissman, J. S. (2009). Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *Science* 324, 218–223.

Ishihama, Y., Oda, Y., Tabata, T., Sato, T., Nagasu, T., Rappsilber, J., et al. (2005). Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Mol. Cell. Proteomics* 4, 1265–1272.

Ito, T., Kim, G. T., and Shinozaki, K. (2000). Disruption of an *Arabidopsis* cytoplasmic ribosomal protein S13-homologous gene by transposon-mediated mutagenesis causes aberrant growth and development. *Plant J.* 22, 257–264.

Jakob, N., Jayati, S., Joachim, F., and Poul, N. (2004). Regulation of eukaryotic translation by the RACK1 protein: a platform for signalling molecules on the ribosome. *EMBO Rep.* 5, 1137–1141.

Kaltimbacher, V., Bonnet, C., Lecoeuvre, G., Forster, V., Sahel, J.-A., and Corral-Debrinski, M. (2006). mRNA localization to the mitochondrial surface allows the efficient translocation inside the organelle of a nuclear recoded ATP6 protein. *RNA* 12, 1408–1417.

Kannan, K., Vázquez-Laslop, N., and Mankin, A. S. (2012). Selective protein synthesis by ribosomes with a drug-obstructed exit tunnel. *Cell* 151, 508–520.

Kapp, L. D., and Lorsch, J. R. (2004). The molecular mechanics of eukaryotic translation. *Annu. Rev. Biochem.* 73, 657–704.

Kawaguchi, R., Williams, A. J., Bray, E. A., and Bailey-Serres, J. (2003). Water-deficit-induced translational control in *Nicotiana tabacum*. *Plant Cell Environ.* 26, 221–229.

Komili, S., Farny, N. G., Roth, F. P., and Silver, P. A. (2007). Functional specificity among ribosomal proteins regulates gene expression. *Cell* 131, 557–571.

Laby, R. J., Kincaid, M. S., Kim, D., and Gibson, S. I. (2000). The *Arabidopsis* sugar-insensitive mutants sis4 and sis5 are defective in abscisic acid synthesis and response. *Plant J.* 23, 587–596.

Leaver, C. J., and Harmey, M. A. (1973). Plant mitochondrial nucleic acids. *Biochem. Soc. Symp.* 175–193.

Leaver, C. J., and Harmey, M. A. (1976). Higher-plant mitochondrial ribosomes contain a 5S ribosomal ribonucleic acid component. *Biochem. J.* 157, 275–277.

Lee, S., Liu, B., Lee, S., Huang, S.-X., Shen, B., and Qian, S.-B. (2012). Global mapping of translation initiation sites in mammalian cells at single-nucleotide resolution. *Proc. Natl. Acad. Sci. U.S.A.* 109, E2424–E2432.

Lee, S. W., Berger, S. J., Martinovic, S., Pasa-Tolic, L., Anderson, G. A., Shen, Y., et al. (2002). Direct mass spectrometric analysis of intact proteins of the yeast large ribosomal subunit using capillary LC/FTICR. *Proc. Natl. Acad. Sci. U.S.A.* 99, 5942–5947.

Liu, M.-J., Wu, S.-H., Chen, H.-M., and Wu, S.-H. (2012). Widespread translational control contributes to the regulation of *Arabidopsis* photomorphogenesis. *Mol. Syst. Biol.* 8, 566.

Louie, D. F., Resing, K. A., Lewis, T. S., and Ahn, N. G. (1996). Mass spectrometric analysis of 40 S ribosomal proteins from Rat-1 fibroblasts. *J. Biol. Chem.* 271, 28189–28198.

Mager, W. H., Planta, R. J., Ballesta, J.-P. G., Lee, J. C., Mizuta, K., Suzuki, K., et al. (1997). A new nomenclature for the cytoplasmic ribosomal proteins of *Saccharomyces cerevisiae*. *Nucleic Acids Res.* 25, 4872–4875.

Mahfouz, M. M., Kim, S., Delauney, A. J., and Verma, D. P. S. (2006). *Arabidopsis* TARGET OF RAPAMYCIN interacts with RAPTOR, which regulates the activity of S6 kinase in response to osmotic stress signals. *Plant Cell* 18, 477–490.

Manuell, A. L., Yamaguchi, K., Haynes, P. A., Milligan, R. A., and Mayfield, S. P. (2005). Composition and structure of the 80S ribosome from

the green alga *Chlamydomonas reinhardtii*: 80S ribosomes are conserved in plants and animals. *J. Mol. Biol.* 351, 266–279.

Mazumder, B., Sampath, P., Seshadri, V., Maitra, R. K., Dicorleto, P. E., and Fox, P. L. (2003). Regulated release of L13a from the 60S ribosomal subunit as a mechanism of transcript-specific translational control. *Cell* 115, 187–198.

McIntosh, K. B., and Warner, J. R. (2007). Yeast ribosomes: variety is the spice of life. *Cell* 131, 450–451.

Mizoguchi, T., Hayashida, N., Yamaguchi-Shinozaki, K., Kamada, H., and Shinozaki, K. (1995). Two genes that encode ribosomal-protein S6 kinase homologs are induced by cold or salinity stress in *Arabidopsis thaliana. FEBS Lett.* 358, 199–204.

Molina, H., Horn, D. M., Tang, N., Mathivanan, S., and Pandey, A. (2007). Global proteomic profiling of phosphopeptides using electron transfer dissociation tandem mass spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* 104, 2199–2204.

Morimoto, T., Suzuki, Y., and Yamaguchi, I. (2002). Effects of partial suppression of ribosomal protein S6 on organ formation in *Arabidopsis thaliana. Biosci. Biotechnol. Biochem.* 66, 2437–2443.

Mustroph, A., Zanetti, M. E., Jang, C. J., Holtan, H. E., Repetti, P. P., Galbraith, D. W., et al. (2009). Profiling translatomes of discrete cell populations resolves altered cellular priorities during hypoxia in *Arabidopsis. Proc. Natl. Acad. Sci. U.S.A.* 106, 18843–18848.

Neumann, H. (2012). Rewiring translation – genetic code expansion and its applications. *FEBS Lett.* 586, 2057–2064.

Neumann, H., Wang, K., Davis, L., Garcia-Alai, M., and Chin, J. W. (2010). Encoding multiple unnatural amino acids via evolution of a quadruplet-decoding ribosome. *Nature* 464, 441–444.

Nishimura, T., Wada, T., and Okada, K. (2004). A key factor of translation reinitiation, ribosomal protein L24, is involved in gynoecium development in *Arabidopsis. Biochem. Soc. Trans.* 32, 611–613.

Nishimura, T., Wada, T., Yamamoto, K. T., and Okada, K. (2005). The *Arabidopsis* STV1 protein, responsible for translation reinitiation, is required for auxin-mediated gynoecium patterning. *Plant Cell* 17, 2940–2953.

Odintsova, T. I., Muller, E. C., Ivanov, A. V., Egorov, T. A., Bienert, R., Vladimirov, S. N., et al. (2003). Characterization and analysis of post-translational modifications of the human large cytoplasmic ribosomal subunit proteins by mass spectrometry and Edman sequencing. *J. Protein Chem.* 22, 249–258.

Piekna-Przybylska, D., Przybylski, P., Baudin-Baillieu, A., Rousset, J.-P., and Fournier, M. J. (2008). Ribosome performance is enhanced by a rich cluster of pseudouridines in the A-site finger region of the large subunit. *J. Biol. Chem.* 283, 26026–26036.

Pinel, C., Douce, R., and Mache, R. (1986). A study of mitochondrial ribosomes from the higher plant Solanum tuberosum L. *Mol. Biol. Rep.* 11, 93–97.

Pinon, V., Etchells, J. P., Rossignol, P., Collier, S. A., Arroyo, J. M., Martienssen, R. A., et al. (2008). Three PIGGYBACK genes that specifically influence leaf patterning encode ribosomal proteins. *Development* 135, 1315–1324.

Piques, M., Schulze, W. X., Hohne, M., Usadel, B., Gibon, Y., Rohwer, J., et al. (2009). Ribosome and transcript copy numbers, polysome occupancy and enzyme dynamics in *Arabidopsis. Mol. Syst. Biol.* 5, 314.

Porras-Yakushi, T. R., Whitelegge, J. P., and Clarke, S. (2008). A novel SET domain methyltransferase in yeast: Rkm2-dependent trimethylation of ribosomal protein L12ab at the N terminus. VOLUME 281 (2006) PAGES 35835-35845. *J. Biol. Chem.* 283, 15512.

Porras-Yakushi, T. R., Whitelegge, J. P., Miranda, T. B., and Clarke, S. (2005). A novel SET domain methyltransferase modifies ribosomal protein Rpl23ab in yeast. *J. Biol. Chem.* 280, 34590–34598.

Pring, D. R. (1974). Maize mitochondria: purification and characterization of ribosomes and ribosomal ribonucleic acid. *Plant Physiol.* 53, 677–683.

Rackham, O., and Chin, J. W. (2005). A network of orthogonal ribosome x mRNA pairs. *Nat. Chem. Biol.* 1, 159–166.

Rackham, O., and Chin, J. W. (2006). Synthesizing cellular networks from evolved ribosome-mRNA pairs. *Biochem. Soc. Trans.* 34, 328–329.

Ren, J., Wang, Y., Liang, Y., Zhang, Y., Bao, S., and Xu, Z. (2010). Methylation of ribosomal protein S10 by protein-arginine methyltransferase 5 regulates ribosome biogenesis. *J. Biol. Chem.* 285, 12695–12705.

Revenkova, E., Masson, J., Koncz, C., Afsar, K., Jakovleva, L., and Paszkowski, J. (1999). Involvement of *Arabidopsis thaliana* ribosomal protein S27 in mRNA degradation triggered by genotoxic stress. *EMBO J.* 18, 490–499.

Santoro, R., Lienemann, P., and Fussenegger, M. (2009). Epigenetic engineering of ribosomal RNA genes enhances protein production. *PLoS ONE* 4:e6653. doi:10.1371/journal.pone.0006653

Scharf, K. D., and Nover, L. (1982). Heat-shock-induced alterations of ribosomal protein phosphorylation in plant cell cultures. *Cell* 30, 427–437.

Shor, B., Calaycay, J., Rushbrook, J., and McLeod, M. (2003). Cpc2/RACK1 is a ribosome-associated protein that promotes efficient translation in *Schizosaccharomyces pombe. J. Biol. Chem.* 278, 49119–49128.

Stapels, M. D., and Barofsky, D. F. (2004). Complementary use of MALDI and ESI for the HPLC-MS/MS analysis of DNA-binding proteins. *Anal. Chem.* 76, 5423–5430.

Stavreva, D. A., Kawasaki, M., Dundr, M., Koberna, K., Müller, W. G., Tsujimura-Takahashi, T., et al. (2006). Potential roles for ubiquitin and the proteasome during ribosome biogenesis. *Mol. Cell. Biol.* 26, 5131–5145.

Stirnberg, P., Liu, J. P., Ward, S., Kendall, S. L., and Leyser, O. (2012). Mutation of the cytosolic ribosomal protein-encoding RPS10B gene affects shoot meristematic function in *Arabidopsis. BMC Plant Biol.* 12:160. doi:10.1186/1471-2229-12-160

Sugihara, Y., Honda, H., Iida, T., Morinaga, T., Hino, S., Okajima, T., et al. (2010). Proteomic analysis of rodent ribosomes revealed heterogeneity including ribosomal proteins L10-like, L22-like 1, and L39-like. *J. Proteome Res.* 9, 1351–1366.

Suissa, M., and Schatz, G. (1982). Import of proteins into mitochondria. Translatable mRNAs for imported mitochondrial proteins are present in free as well as mitochondria-bound cytoplasmic polysomes. *J. Biol. Chem.* 257, 13048–13055.

Szakonyi, D., and Byrne, M. E. (2011). Ribosomal protein L27a is required for growth and patterning in *Arabidopsis thaliana. Plant J.* 65, 269–281.

Szick, K., Springer, M., and Bailey-Serres, J. (1998). Evolutionary analyses of the 12-kDa acidic ribosomal P-proteins reveal a distinct protein of higher plant ribosomes. *Proc. Natl. Acad. Sci. U.S.A.* 95, 2378–2383.

Szick-Miranda, K., and Bailey-Serres, J. (2001). Regulated heterogeneity in 12-kDa P-protein phosphorylation and composition of ribosomes in Maize (Zea mays L.). *J. Biol. Chem.* 276, 10921–10928.

Szick-Miranda, K., Zanial, A., Zanial, A., Abidayo, S., and Slater, K. C. (2010). Analysis of RPS15aE, an isoform of a plant-specific evolutionarily distinct ribosomal protein in *Arabidopsis thaliana*, reveals its potential role as a growth regulator. *Plant Mol. Biol. Rep.* 28, 239–252.

Taira, H., Fukushima, M., Hohsaka, T., and Sisido, M. (2005). Four-base codon-mediated incorporation of nonnatural amino acids into proteins in a eukaryotic cell-free translation system. *J. Biosci. Bioeng.* 99, 473–476.

Turck, F., Zilbermann, F., Kozma, S. C., Thomas, G., and Nagy, F. (2004). Phytohormones participate in an S6 kinase signal transduction pathway in *Arabidopsis. Plant Physiol.* 134, 1527–1535.

Turkina, M. V., Klang Årstrand, H., and Vener, A. V. (2011). Differential phosphorylation of ribosomal proteins in *Arabidopsis thaliana* plants during day and night. *PLoS ONE* 6:e29307. doi:10.1371/journal.pone.0029307

Van Lijsebettens, M., Vanderhaeghen, R., De Block, M., Bauw, G., Villarroel, R., and Van Montagu, M. (1994). An S18 ribosomal protein gene copy at the *Arabidopsis* PFL locus affects plant development by its specific expression in meristems. *EMBO J.* 13, 3378.

Vasconcelos, A. C., and Bogorad, L. (1971). Proteins of cytoplasmic, chloroplast, and mitochondrial ribosomes of some plants. *Biochim. Biophys. Acta* 228, 492–502.

Vohradsky, J., Branny, P., Li, X.-M., and Thompson, C. J. (2008). Effect of protein degradation on spot Mr distribution in 2-D gels – a case study of proteolysis during development of Streptomyces coelicolor cultures. *Proteomics* 8, 2371–2375.

Wagner, H. (1989). Free-flow electrophoresis. *Nature* 341, 669–670.

Wang, B., Malik, R., Nigg, E. A., and KöRner, R. (2008). Evaluation of the low-specificity protease elastase for large-scale phosphoproteome analysis. *Anal. Chem.* 80, 9526–9533.

Warner, J. R. (1999). The economics of ribosome biosynthesis in yeast. *Trends Biochem. Sci.* 24, 437–440.

Webb, K. J., Al-Hadid, Q., Zurita-Lopez, C. I., Young, B. D., Lipson, R. S., and

Clarke, S. G. (2011). The ribosomal L1 protuberance in yeast is methylated on a lysine residue catalyzed by a seven-β-strand methyltransferase. *J. Biol. Chem.* 286, 18405–18413.

Webb, K. J., Lipson, R. S., Al-Hadid, Q., Whitelegge, J. P., and Clarke, S. G. (2010a). Identification of protein N-terminal methyltransferases in yeast and humans. *Biochemistry* 49, 5225–5235.

Webb, K. J., Zurita-Lopez, C. I., Al-Hadid, Q., Laganowsky, A., Young, B. D., Lipson, R. S., et al. (2010b). A novel 3-methylhistidine modification of yeast ribosomal protein Rpl3 is dependent upon the YIL110W methyltransferase. *J. Biol. Chem.* 285, 37598–37606.

Weijers, D., Franke-Van Dijk, M., Vencken, R. J., Quint, A., Hooykaas, P., and Offringa, R. (2001). An *Arabidopsis* minute-like phenotype caused by a semi-dominant mutation in a RIBOSOMAL PROTEIN S5 gene. *Development* 128, 4289–4299.

Williams, A. J., Werner-Fraczek, J., Chang, I.-F., and Bailey-Serres, J. (2003). Regulated phosphorylation of 40S ribosomal protein S6 in root tips of maize. *Plant Physiol.* 132, 2086–2097.

Williams, M. E., and Sussex, I. M. (1995). Developmental regulation of ribosomal protein L16 genes in *Arabidopsis thaliana*. *Plant J.* 8, 65–76.

Wilson, D. N., and Doudna Cate, J. H. (2012). The structure and function of the eukaryotic ribosome. *Cold Spring Harb. Perspect. Biol.* 4, pii: a011536.

Woeller, C. F., Fox, J. T., Perry, C., and Stover, P. J. (2007). A ferritin-responsive internal ribosome entry site regulates folate metabolism. *J. Biol. Chem.* 282, 29927–29935.

Wool, I. G., Chan, Y.-L., and Glück, A. (1995). Structure and evolution of mammalian ribosomal proteins. *Biochem. Cell Biol.* 73, 933–947.

Xue, S., and Barna, M. (2012). Specialized ribosomes: a new frontier in gene regulation and organismal biology. *Nat. Rev. Mol. Cell Biol.* 13, 355–369.

Yamaguchi, K. (2011). "Preparation and proteomic analysis of chloroplast ribosomes," in *Chloroplast Research in Arabidopsis*, ed. R. P. Jarvis (New York: Humana Press), 241–264.

Yamaguchi, K., Beligni, M. V., Prieto, S., Haynes, P. A., Mcdonald, W. H., Yates, J. R., et al. (2003). Proteomic characterization of the *Chlamydomonas reinhardtii* chloroplast ribosome. *J. Biol. Chem.* 278, 33774–33785.

Yang, P., Fu, H., Walker, J., Papa, C. M., Smalle, J., Ju, Y.-M., et al. (2004). Purification of the *Arabidopsis* 26 S proteasome: biochemical and molecular analyses revealed the presence of multiple isoforms. *J. Biol. Chem.* 279, 6401–6413.

Yao, Y., Ling, Q., Wang, H., and Huang, H. (2008). Ribosomal proteins promote leaf adaxial identity. *Development* 135, 1325–1334.

Yu, Y., Ji, H., Doudna, J. A., and Leary, J. A. (2005). Mass spectrometric analysis of the human 40S ribosomal subunit: native and HCV IRES-bound complexes. *Protein Sci.* 14, 1438–1446.

Zanetti, M. E., Chang, I.-F., Gong, F., Galbraith, D. W., and Bailey-Serres, J. (2005). Immunopurification of polyribosomal complexes of *Arabidopsis* for global analysis of gene expression. *Plant Physiol.* 138, 624–635.

Zhang, H., and Ge, Y. (2011). Comprehensive analysis of protein modifications by top-down mass spectrometry. *Circ. Cardiovasc. Genet.* 4, 711.

Zhou, H., Ning, Z. E., Starr, A., Abu-Farha, M., and Figeys, D. (2011). Advancements in top-down proteomics. *Anal. Chem.* 84, 720–734.

Zorca, S. M., and Zorca, C. E. (2011). The legacy of a founding father of modern cell biology: George Emil Palade (1912-2008). *Yale J. Biol. Med.* 84, 113–116.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 01 November 2012; accepted: 10 February 2013; published online: 01 March 2013.

Citation: Carroll AJ (2013) The Arabidopsis cytosolic ribosomal proteome: from form to function. Front. Plant Sci. ***4****:32. doi: 10.3389/fpls.2013.00032*

This article was submitted to Frontiers in Plant Proteomics, a specialty of Frontiers in Plant Science.

Copyright © 2013 Carroll. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.

frontiers in
PLANT SCIENCE

MINI REVIEW ARTICLE
published: 03 January 2013
doi: 10.3389/fpls.2012.00298

Proteomic dissection of the *Arabidopsis* Golgi and *trans*-Golgi network

Harriet T. Parsons[1]*, Georgia Drakakaki[2] and Joshua L. Heazlewood[3,4]

[1] Department of Plant and Environmental Sciences, University of Copenhagen, Copenhagen, Denmark
[2] Department of Plant Sciences, University of California at Davis, Davis, CA, USA
[3] Joint BioEnergy Institute, Lawrence Berkeley National Laboratory, Berkeley, CA, USA
[4] Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA

Edited by:
Nicolas L. Taylor, The University of Western Australia, Australia

Reviewed by:
Holger Eubel, Leibniz Universität Hannover, Germany
Karine Gallardo, National Institute for Agronomic Research, France

****Correspondence:***
Harriet T. Parsons, Department of Plant and Environmental Sciences, University of Copenhagen, Thorvaldsensvej 40, 1871 Frederiksberg C, Copenhagen, Denmark.
e-mail: htpa@life.ku.dk

The plant Golgi apparatus and *trans*-Golgi network are major endomembrane trafficking hubs within the plant cell and are involved in a diverse and vital series of functions to maintain plant growth and development. Recently, a series of disparate technical approaches have been used to isolate and characterize components of these complex organelles by mass spectrometry in the model plant *Arabidopsis thaliana*. Collectively, these studies have increased the number of Golgi and vesicular localized proteins identified by mass spectrometry to nearly 500 proteins. We have sought to provide a brief overview of these technical approaches and bring the datasets together to examine how they can reveal insights into the secretory pathway.

Keywords: Golgi, *trans*-Golgi network, proteomics, LOPIT, free-flow electrophoresis, *Arabidopsis*, SYP61

BACKGROUND

At its simplest level, subcellular proteomics attempts to identify all proteins in a particular compartment. However, even with such a basic definition in mind, the Golgi proteome presents conceptual difficulties; functional proteins in the Golgi may also be functional elsewhere (Ondzighi et al., 2008), whilst endoplasmic reticulum (ER)–Golgi connections (Boevink et al., 1998) makes absolute divisions between the proteomes of these compartments somewhat futile. A number of proteins are known to form functional associations on the cytoplasmic face of cisternae but are part of the cytosol (Ito et al., 2011), so the very definition of the Golgi proteomes is problematic. Furthermore, in such an architecturally heterogeneous organelle, simply identifying all the proteins present in the Golgi is not that helpful unless we can classify them according to sub-Golgi location, post-Golgi compartments, cargo, resident, or dual-localized proteins. The plant Golgi poses a challenge in terms of isolation, not least because of its fragmented morphology. In mammalian cells Golgi stacks tend to be less numerous per cell with fewer, longer cisternae which are less tightly associated with the ER and could be relatively easily isolated (Morre and Mollenhauer, 2009). Excepting highly conserved pathways such as protein *N*-linked glycan processing, few similarities exist between plant and mammalian Golgi. Thus assuming Golgi-residency between the two systems based on homology alone is not possible. Earlier work on Golgi from rat liver was therefore of limited help either in terms of providing an isolation strategy or a comprehensive bank of marker proteins (Taylor et al., 1997). The plant Golgi is much less structurally defined during and after cell homogenization than, for example, plastids or mitochondria. Consequently, quality control of and improvements to isolation strategies have been tricky and therefore purity limited when using sucrose density centrifugation strategies (Morre and Mollenhauer, 1964). In short, it is easy to understand why progress in Golgi proteomics has trailed behind other subcellular compartments in plants. In light of the shortcomings of sucrose density centrifugation for plant Golgi purification, two more technical but very different approaches have been successfully applied, namely localization of organelle proteins by isotope tagging (LOPIT) and free-flow electrophoresis (FFE). The LOPIT approach does not distinguish between Golgi and the *trans*-Golgi network (TGN) localized proteins but identifies resident proteins (Dunkley et al., 2004, 2006; Nikolovski et al., 2012), whilst the FFE approach identified proteins in fractions of purified Golgi, that were estimated to be enriched in medial Golgi cisternae (Parsons et al., 2012a). Immunoisolation of compartments has recently been used to great effect in separating components of the TGN, enabling comparative proteomics at the sub-Golgi level (Drakakaki et al., 2012). Characterization of Golgi-enriched fractions has been attempted in various plant systems (Tanaka et al., 2004; Asakura et al., 2006; Mast et al., 2010), major, large-scale proteomic characterizations have exclusively occurred in the model plant *Arabidopsis thaliana*.

AN OVERVIEW OF THE *Arabidopsis* GOLGI–TGN PROTEOMES

Initial attempts to characterize the *Arabidopsis* Golgi by mass spectrometry were undertaken nearly a decade ago with the aim of distinguishing between ER- and Golgi-resident proteins (Dunkley et al., 2004). The LOPIT approach involves quantitative mass spectrometry of proteins labeled with isotope tags. A cell homogenate separated along a linear gradient is fractionated and pairwise comparisons of fractions allow abundance ratios of isotope masses to be calculated for each protein. Proteins physically located in the same compartment will have similar ratios and so cluster together

during partial least squares discriminant analysis (**Figure 1**). Using LOPIT, 89 proteins were initially localized to the Golgi (Dunkley et al., 2006) but the requirement that proteins carry all four tags limited the number of proteins for which a statistically credible localization could be assigned. Recent reanalysis and analysis of existing and new datasets, incorporating values for "missing" tags assigned using partial least squares regression models and training sets based on fully tagged proteins, enabled the collective localization of 204 proteins to the Golgi/TGN (Dunkley et al., 2006; Nikolovski et al., 2012).

Although a major motivation for the development of LOPIT was the difficulty in separating the Golgi, particularly from ER contaminants, a recent study has managed to isolate Golgi vesicles with an estimated 80% purity based on protein composition. This was achieved using a combination of sucrose density centrifugation and FFE (Parsons et al., 2012a). The power of FFE for organelle isolation was demonstrated in plants several years ago when applied to the separation of mitochondria and peroxisomes, two organelles which are typically hard to separate using density centrifugation alone (Eubel et al., 2008). As separation by FFE is dependent on surface charge, the Golgi, which carries a more negative surface charge than ER vesicles and most other contaminants, is amenable to separation using this technique, which resulted in 371 proteins being localized to the Golgi (**Figure 1**).

A dissection of the complexity of the Golgi proteome was recently attempted using immunoisolation of specific TGN trafficking populations. Affinity purified TGN compartments from plants expressing a syntaxin from plants (SYP61)-CFP construct were enriched for the TGN by sucrose density centrifugation then exposed to anti-FP antibodies coupled to agarose beads and analyzed by mass spectrometry (Drakakaki et al., 2012). Although widely used in mammalian systems, application of this approach in plants was precedential. The technique was able to identify 145 proteins from affinity purified samples of SYP61 vesicles, providing the foundation of a TGN proteome in plants.

THE SIZE OF THE PLANT GOLGI PROTEOME

In total, 452 proteins have been characterized by mass spectrometry to the Golgi apparatus and 145 to the TGN from the model plant *Arabidopsis*. An ever-present question in subcellular

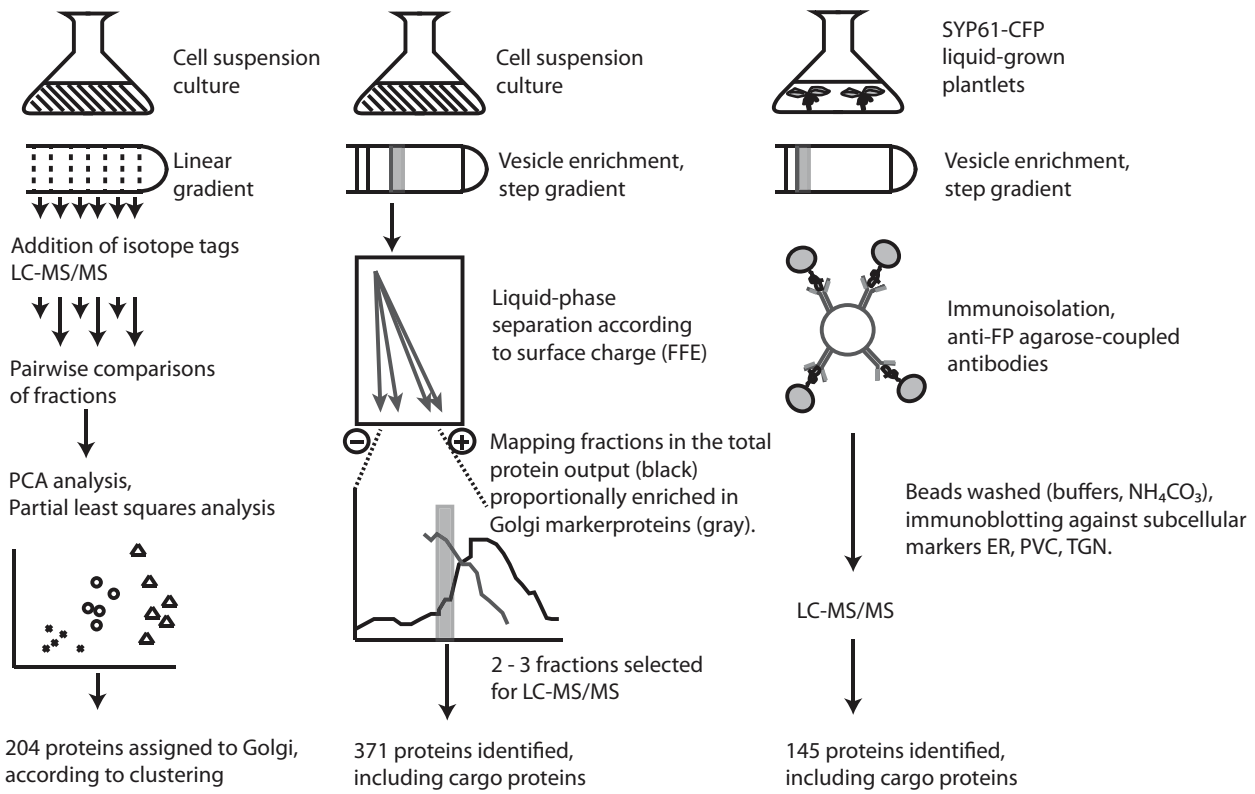


FIGURE 1 | Overview of the three different techniques employed in proteomic characterization of the *Arabidopsis* Golgi and TGN. **(A)** Clustered proteins in LOPIT studies were assigned to the Golgi according to co-clustering with known and predicted Golgi marker proteins (for details, see Dunkley et al., 2004, 2006; Nikolovski et al., 2012). **(B)** FFE purified fractions were estimated at ca. 80% purity according to the proportion of previously localized Golgi proteins and contaminants present in each fraction; based on experimental data in SUBA (Heazlewood et al., 2007; for details, see Parsons et al., 2012a,c). **(C)** Isolation of SYP61 vesicles by affinity purification. Successful removal of contaminants during immunoisolation was assayed by the presence of the ER/*cis*-Golgi marker, BiP, and the prevacuolar compartment marker SYP21 (for details, see Drakakaki et al., 2012; Parsons et al., 2012c).

proteomics concerns the total number of proteins present in an organelle. Given the residential/transitory definitions raised above, this is an especially difficult question to answer in the case of the Golgi and TGN, since proteins with ambiguous localization profile cannot be clearly assigned to a particular sub-compartment. Therefore dual-localized but Golgi-functional proteins or those at the *cis*-Golgi extremity will potentially be excluded from many analyses. Given the extensive subcellular localization data in the model plant *Arabidopsis* and the collection of subcellular prediction algorithms that are outlined in the SUBA database (Heazlewood et al., 2007), it is possible to make an estimation of the size of an organelle proteome based on an experimentally determined collection (Ito et al., 2011). Collectively, 491 proteins (excluding the defined cargo proteins) have been localized to the Golgi/TGN proteomes (Dunkley et al., 2006; Drakakaki et al., 2012; Nikolovski et al., 2012; Parsons et al., 2012a) and 145 proteins to the Golgi/TGN by fluorescent marker studies (Heazlewood et al., 2007). In total 575 unique proteins have been experimentally localized to the Golgi/TGN. Of the 22 subcellular prediction algorithms that have been applied to the entire *Arabidopsis* proteome, 14 provide a "Golgi" prediction output (**Table 1**).

Employing the relational capabilities of the SUBA database, it is possible to compute a size estimate of the Golgi/TGN proteome based on each algorithms performance. The overall performance of each prediction program can vary considerably with regard to the total predicted "Golgi" proteins in *Arabidopsis* (contrast AdaBoost, 66 Golgi and PProwler, 8885 Golgi) and positive prediction rate of the experimental proteome (contrast AdaBoost $<1\%$ and PProwler $>50\%$). However, after calculating false positive and false negative rates for each program, the final predicted Golgi proteomes are remarkably similar. Based on this analysis, the *Arabidopsis* Golgi/TGN proteome is estimated to be 2239 ± 465, employing the average of the predicted proteomes of these 14 subcellular prediction programs.

USING THE PROTEOME: WHAT ARE THE ROLES OF UNCHARACTERIZED PROTEIN FAMILIES?

A number of large gene families have been identified by both the FFE and LOPIT studies (Nikolovski et al., 2012; Parsons et al., 2012a). The quantitative mass spectrometry performed when applying LOPIT (Nikolovski et al., 2012) and spectral counts from FFE isolates (Parsons et al., 2012a), combined with localization data (Heazlewood et al., 2007), provide an important starting guide as to which members of these large families are major components and should be initially investigated in future studies.

The cyclophilin-like peptidyl-prolyl *cis–trans* isomerase family is consistently represented in the Golgi proteomes. These are known to catalyze conversion of *cis* to *trans* conformation of peptide bonds preceding prolyl residues in newly synthesized peptides (Chou and Gasser, 1997). In plants, they are classically associated with the thylakoid lumen where they are thought to help protein folding and assembly of photosystem complexes although their exact role is not clear (Ingelsson et al., 2009). The cyclophilins found by both FFE and LOPIT approaches (Nikolovski et al., 2012; Parsons et al., 2012a) localize either exclusively to the Golgi or are dually localized to the Golgi and plasma membrane (Dunkley et al., 2006; Benschop et al., 2007; Marmagne et al., 2007; Parsons et al., 2012a), implying a secretory-specific function, although no cyclophilins were found during immunoisolation of the TGN (Drakakaki et al., 2012).

The prenylated RAB acceptor B2 (PRA1.B2, AT2G40380) is found in both Golgi proteomes (FFE and LOPIT) but not the TGN, implying involvement with cisternal-specific interactions and vesicle docking. Examining proteins present uniquely in the TGN, besides those involved in trafficking such as the RAB GTPases, soluble *N*-ethylmaleimide-sensitive factor attachment protein receptors (SNARE; Blatt et al., 1999; Surpin and Raikhel, 2004), transport protein particle (TRAPP) components (Barrowman et al., 2010) or present as cargo, e.g., specific cellulose synthase A (CESA) subunits (Paredez et al., 2006), one endomembrane protein/transmembrane 9 protein (EMP/TMN9) and two *S*-adenosyl-L-methionine-dependent methyltransferases appear to stand out. Most EMP/TMN9 proteins are found in the Golgi cisternae: 11 members from a total of 12 were identified in FFE-purified samples (Parsons et al., 2012a) and 10 during LOPIT studies. EMP/TMN9 proteins interact with COPI and COPII proteins and membrane proteins destined for post-Golgi locations but are only recently studied in plants (Gao et al., 2012). The presence of two EMP/TMN9 proteins in both the Golgi and TGN implies *trans*-Golgi localization. With only one EMP/TMN9 identified uniquely in the TGN, members of the family may fulfill niche roles in trafficking depending on their location along the Golgi stack and are likely interesting subjects for future study. Apart from QUA2 (Mouille et al., 2007), a pectin methyltransferase in the *S*-adenosyl-L-methionine-dependent methyltransferases superfamily, no clear function has been assigned to any other members of this family of proteins in plants. The *S*-adenosyl-L-methionine-dependent methyltransferases which include QUA2 are prevalent in the Golgi and Golgi/TGN proteomes. A total of 20 were identified by LOPIT, 15 by FFE, and 3 in SYP61, resulting in 22 distinct proteins from this family (Drakakaki et al., 2012; Nikolovski et al., 2012; Parsons et al., 2012a). One member, AT5G64030, has been found in the plasma membrane proteome (Mitra et al., 2009; Zhang and Peck, 2011), so could conceivably function there. Assuming that all family members perform some kind of polysaccharide methylation, proteomic comparisons could be used to reveal late-acting enzymes in cell wall biosynthesis such as these examples.

Many functionally important Golgi proteins may actually be the sole members of their protein family. Of the 111 proteins not assigned to a functional protein category in the FFE proteome, 30 were also identified by LOPIT studies and many different protein families were represented. Amongst datasets such as these, dataset overlaps can provide a means to shortlist potentially important proteins about which little information is available.

Interestingly, although the proteomes comprised by the LOPIT studies and Parsons et al. (2012a) were both derived from similar starting tissues, a number of proteins are found in Parsons et al. (2012a) but not LOPIT studies and vice versa. Parsons et al. (2012a) identified more proteins overall and results included cargo proteins, unlike in LOPIT studies. Nevertheless after eliminating those annotated by Parsons et al. (2012a) as either transient or involved in protein synthesis, 81 proteins identified by LOPIT are not found in Parsons et al. (2012a) and 205 are in

Table 1 | Estimated size of the *Arabidopsis* Golgi/TGN proteomes utilizing data in the SUBA database and the current integrated proteome (575) employing the abilities of subcellular prediction algorithms.

Predictor	Predicted Golgi *Arabidopsis*	Expt. any location (575)	Expt. in Golgi (575)	Expt. non-Golgi	FPR Golgi prediction	Est. correct predictions	FNR Golgi prediction	Predicted Golgi	Non-predictable expt. Golgi	Reference
AdaBoost	66	13	2	11	0.85	10	1.00	2919	2909	Niu et al. (2008)
BaCelLo	5178	1101	104	997	0.91	489	0.82	2704	2215	Pierleoni et al. (2006)
EpiLoc	290	121	30	91	0.75	72	0.95	1378	1306	Brady and Shatkay (2008)
iPSORT	5688	1439	177	1262	0.88	700	0.69	2273	1573	Bannai et al. (2002)
MultiLoc2	916	299	114	185	0.62	349	0.80	1762	1412	Blum et al. (2009)
Plant-mPLoc	1074	330	100	230	0.70	325	0.83	1871	1546	Chou and Shen (2010)
Predotar	5851	1438	186	1252	0.87	757	0.68	2340	1583	Small et al. (2004)
PredSL	8780	2082	259	1823	0.88	1092	0.55	2425	1333	Unpublished
PProwler	8885	2121	296	1825	0.86	1240	0.49	2409	1169	Boden and Hawkins (2005)
SLPFA	7738	1733	142	1591	0.92	634	0.75	2567	1933	Tamura and Akutsu (2007)
SLP-Local	7406	1599	167	1432	0.90	773	0.71	2663	1890	Matsuda et al. (2005)
TargetP	6492	1574	200	1374	0.87	825	0.65	2372	1547	Emanuelsson et al. (2000)
WoLF PSORT	112	45	11	34	0.76	27	0.98	1431	1404	Horton et al. (2007)
YLoc	1235	318	105	213	0.67	408	0.82	2233	1825	Briesemeister et al. (2010)

Predicted Golgi Arabidopsis: All proteins predicted to be Golgi in Arabidopsis (TAIR10).
Expt. any location (575): Predicted Golgi and experimentally determined to be in any location by MS or FP (SUBA) or Golgi (575).
Expt. in Golgi (575): Predicted Golgi and experimentally determined to be in Golgi (575).
Expt. non-Golgi: Predicted Golgi but experimentally found to be non-Golgi [(Expt. any location) − (Expt. non-Golgi)].
FPR Golgi prediction: False positive rate for Golgi prediction. [(Expt. non-Golgi)/(Expt. in Golgi)].
Est. correct predictions: Estimation of correct predictions from total Golgi predictions in Arabidopsis [(Expt. any location) − (Expt. any location) × (FPR Golgi prediction)].
FNR Golgi prediction: False negative rate for Golgi prediction [1 − (Expt. in Golgi)/575].
Predicted Golgi: The predicted size of the proteome based on validated performance for each predictor program [(Expt. any location) × (1 − FPR)/(1-FNR)].
Non-predictable expt. Golgi: The size of the unpredictable Golgi proteome [(Predicted Golgi) − (Est. correct predictions)].

Parsons et al. (2012a) but not LOPIT. No clear pattern, e.g., protein abundance, exists between the proteins observed in either study; most probably differences arise from variations in methodologies, highlighting the value of multi-facetted approaches to proteomic characterization of the Golgi.

WHAT IS MISSING FROM THE EXPERIMENTAL GOLGI PROTEOME?

Specific questions concerning what has not been identified so far are obviously difficult to answer but they can be addressed in part by examining what sorts of protein have been localized by fluorescent tagging but not identified by subcellular proteomic techniques. Fluorescent localization of proteins is generally motivated by interest in a specific protein and so is more likely to represent low-abundant polypeptides. It therefore provides an initial guide to the completeness of subcellular proteomic approaches.

Notably absent from proteomic surveys, but localized to the Golgi stack by fluorescent tagging are the Golgins and GRIP domain proteins (Latijnhouwers et al., 2007). Several glycosyltransferases such as cellulose synthase-like D5 (CSLD5; Bernal et al., 2007), rhamnogalacturonan II xylosyltransferase (RGXT) 1 and 2 (Egelund et al., 2006), irregular xylem 9 (IRX9; Pena et al., 2007), reversibly glycosylated polypeptide (RGP)1–4 (Drakakaki et al., 2006; Rautengarten et al., 2011), galacturonic acid transferase like (GATL) members from the GT8 family and a number of small GTPases are also either absent or poorly represented. Common methodological steps between these technically very different proteomes may in part explain these absences. Both the FFE and LOPIT approaches (Nikolovski et al., 2012; Parsons et al., 2012a) used cell suspension cultures whilst the immunoisolation approach (Drakakaki et al., 2012) used 14-day-old liquid grown plantlets as the starting tissue, meaning that all proteomes were based on primary cell wall-rich tissue. This may explain the absence of CSLD5 and IRX9, which are both implicated in secondary cell wall biosynthesis and localized to the Golgi stack (Bernal et al., 2007; Lee et al., 2007). RGXT1 and 2 may have been also have been missed because of tissue-specific or low expression (Egelund et al., 2006). Members of the GATL clade, although localized to the Golgi stack (Kong et al., 2011), are absence from all Golgi proteomes, which could point toward some specific spatial or temporal function of these glycosyltransferases. Golgins are Golgi matrix proteins with coiled coil domains that typically locate to the *cis*- and *trans*-extremities of the Golgi stack and cisternal peripheries. They are involved in regulation of stack architecture and tethering events during trafficking (Osterrieder, 2012). Their location to *cis*- or *trans*-extremities of the Golgi stack may have precluded detection (Nikolovski et al., 2012; Parsons et al., 2012a). Peripheral golgins and those with GRIP domains which localize to the TGN, have no predicted transmembrane domain and appear to be recruited from the cytosol by interactions with small GTPases. Their absence from either the Golgi or the SYP61 proteome (Drakakaki et al., 2012) may be due to carbonate washes used to remove cytosolic contaminants and/or centrifugation steps. Electron micrographs taken during FFE isolation procedure (Parsons et al., 2012a) show loss of vesicles from cisternal edges in with progressive centrifugation steps. Two of four data sets used in the LOPIT approach (Nikolovski et al., 2012) had been subjected to carbonate washes resulting in reduced peripheral proteins. This may explain why no RGPs have been detected, as these are peripheral membrane associated proteins (Delgado et al., 1998).

Several RAB GTPases have been localized by fluorescent protein assay to the Golgi stack (Batoko et al., 2000; Feraru et al., 2012). LOPIT approaches have identified two RAB GTPases localized to the Golgi, five were found by FFE purification (Parsons et al., 2012a) and 19 by immunoisolation (Drakakaki et al., 2012). RAB GTPases are involved in cargo-vesicle docking (Woollard and Moore, 2008) and are not Golgi-residents. This likely explains why fewer were present in the LOPIT Golgi proteome (Nikolovski et al., 2012). Step gradients employed prior to FFE purifications (Parsons et al., 2012a) were designed for maximal cisternal enrichment at the cost of small vesicles, so as to minimize ER contamination prior to FFE. This exemplifies the role of methodology in these technically diverse proteomes and shows how removal of contaminants may risk removal of Golgi-associated proteins.

Judging from these inconsistencies between the subcellular proteomics data and fluorescent protein localizations, it is clear that Golgi proteomics must be applied to other tissue types if the proteome is to be "completed." This presents an even greater technical challenge as young, softer tissues are more easily homogenized to maintain Golgi stack integrity (Morre and Mollenhauer, 2009). However, useful information may be gleaned from less pure preparations using tougher, challenging tissue types, or preparations which are less pure but contains Golgi-associated and Golgi matrix proteins, as there is now a sufficiently broad base of proteins from which to compile ever more extensive markers and training sets.

SUB-GOLGI PROTEOMICS AND THE GOLGI IN AN ENDOMEMBRANE CONTEXT

Comparative analyses such as those discussed above can now be formulated since a post-Golgi compartment has been characterized. The potential for distinguishing resident and cargo Golgi components can also be applied. Almost 30% of proteins identified in the TGN proteome comprise non-Golgi proteins as determined by the LOPIT approach (Drakakaki et al., 2012; Nikolovski et al., 2012). It is conceivable that with a few more post-Golgi compartments characterized, many of the endomembrane proteins currently assigned to multiple locations (Heazlewood et al., 2007) could be reassigned and more light shed on the various protein cycling routes through the secretory pathway. This could be reasonably achieved in a number of ways. For the smaller compartments such as endosomal compartments, the immunoisolation approach (Drakakaki et al., 2012) would hold the most promise as a number of syntaxin proteins known to associate with this compartment have been identified (Sanderfoot and Raikhel, 1999). Such an approach may not be appropriate for isolating individual cisternae from the main stack as trafficked proteins destined for later cisternae and TGN may also be detected by antibodies, whilst stack architecture could prove too complex for such an approach. Several fractions containing a high proportion of known Golgi proteins were not included in the FFE proteome owing to slightly higher level of contaminants. The number of fractions in which over 25% of proteins had been localized to the Golgi by LOPIT

studies suggest partial electrophoretic separation of cisternae may have been occurring during the isolation process (Parsons et al., 2012a,b). A collection of sub-Golgi markers have been characterized (Saint-Jore-Dupas et al., 2006), so if proteins from FFE fractions could be accurately quantified profiles of co-migrating proteins could be created to enable sub-Golgi differentiation.

CONCLUDING REMARKS

Although one of the most technically challenging organelles to isolate, a diversity of technologies have led to two Golgi proteomes and one proteome of TGN vesicles, resulting in nearly 500 proteins now localized to the Golgi and/or TGN by mass spectrometry. As the hub of protein trafficking, its proteome is best understood within the context of other proteomes; comparisons between these compartments bring a new level of understanding to protein distribution through the endomembrane system and show the potential for expansion through proteomic analysis of other post-Golgi compartments. It is estimated here that only about 20% of Golgi proteins have been identified thus far by mass spectrometry. So far all studies have been carried out in rapidly dividing, developing tissue (either cell suspension culture or liquid-grown plantlets). Exploration of other tissue types is needed to increase the coverage of the Golgi proteome. Efforts must also be concentrated in getting the proteomes of *cis*-, *medial*-, and *trans*-Golgi subcompartments and specific vesicle populations. This will incur further technical challenges but will help identify more lowly expressed proteins and provide invaluable insight into plant Golgi functions.

ACKNOWLEDGMENTS

This work conducted by the Joint BioEnergy Institute was supported by the Office of Science, Office of Biological and Environmental Research, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. The work conducted by Drakakaki et al. (2012) was funded by a grant from the DOE (DEFG03-02ER15295) and UC Davis startup funds.

REFERENCES

Asakura, T., Hirose, S., Katamine, H., Sato, M., Hujiwara, M., Shimamoto, K., et al. (2006). Rice Golgi proteome: analysis of GFP-syp31 labeled *cis* Golgi membrane. *Plant Cell Physiol.* 47, S26–S26.

Bannai, H., Tamada, Y., Maruyama, O., Nakai, K., and Miyano, S. (2002). Extensive feature detection of N-terminal protein sorting signals. *Bioinformatics* 18, 298–305.

Barrowman, J., Bhandari, D., Reinisch, K., and Ferro-Novick, S. (2010). TRAPP complexes in membrane traffic: convergence through a common Rab. *Nat. Rev. Mol. Cell Biol.* 11, 759–763.

Batoko, H., Zheng, H.-Q., Hawes, C., and Moore, I. (2000). A Rab1 GTPase is required for transport between the endoplasmic reticulum and Golgi apparatus and for normal Golgi movement in plants. *Plant Cell* 12, 2201–2217.

Benschop, J. J., Mohammed, S., O'Flaherty, M., Heck, A. J. R., Slijper, M., and Menke, F. L. H. (2007). Quantitative phosphoproteomics of early elicitor signaling in *Arabidopsis*. *Mol. Cell. Proteomics* 6, 1198–1214.

Bernal, A. J., Jensen, J. K., Harholt, J., Sørensen, S., Moller, I., Blaukopf, C., et al. (2007). Disruption of ATCSLD5 results in reduced growth, reduced xylan and homogalacturonan synthase activity and altered xylan occurrence in *Arabidopsis*. *Plant J.* 52, 791–802.

Blatt, M. R., Leyman, B., and Geelen, D. (1999). Tansley review no. 108 – molecular events of vesicle trafficking and control by SNARE proteins in plants. *New Phytol.* 144, 389–418.

Blum, T., Briesemeister, S., and Kohlbacher, O. (2009). MultiLoc2: integrating phylogeny and gene ontology terms improves subcellular protein localization prediction. *BMC Bioinformatics* 10:274. doi: 10.1186/1471-2105-10-274

Boden, M., and Hawkins, J. (2005). Prediction of subcellular localization using sequence-biased recurrent networks. *Bioinformatics* 21, 2279–2286.

Boevink, P., Oparka, K., Cruz, S. S., Martin, B., Betteridge, A., and Hawes, C. (1998). Stacks on tracks: the plant Golgi apparatus traffics on an actin/ER network. *Plant J.* 15, 441–447.

Brady, S., and Shatkay, H. (2008). EpiLoc: a (working) text-based system for predicting protein subcellular location. *Pac. Symp. Biocomput.* 13:604–615.

Briesemeister, S., Rahnenfuhrer, J., and Kohlbacher, O. (2010). YLoc–an interpretable web server for predicting subcellular localization. *Nucleic Acids Res.* 38, W497–W502.

Chou, I. T., and Gasser, C. S. (1997). Characterization of the cyclophilin gene family of *Arabidopsis thaliana* and phylogenetic analysis of known cyclophilin proteins. *Plant Mol. Biol.* 35, 873–892.

Chou, K. C., and Shen, H. B. (2010). Plant-mPLoc: a top-down strategy to augment the power for predicting plant protein subcellular localization. *PLoS ONE* 5:e11335. doi: 10.1371/journal.pone.0011335

Delgado, I. J., Wang, Z. H., De Rocher, A., Keegstra, K., and Raikhel, N. V. (1998). Cloning and characterization of AtRGP1 – a reversibly autoglycosylated *Arabidopsis* protein implicated in cell wall biosynthesis. *Plant Physiol.* 116, 1339–1349.

Drakakaki, G., Van De Ven, W., Pan, S. Q., Miao, Y. S., Wang, J. Q., Keinath, N. F., et al. (2012). Isolation and proteomic analysis of the SYP61 compartment reveal its role in exocytic trafficking in *Arabidopsis*. *Cell Res.* 22, 413–424.

Drakakaki, G., Zabotina, O., Delgado, I., Robert, S., Keegstra, K., and Raikhel, N. (2006). *Arabidopsis* reversibly glycosylated polypeptides 1 and 2 are essential for pollen development. *Plant Physiol.* 142, 1480–1492.

Dunkley, T. P., Watson, R., Griffin, J. L., Dupree, P., and Lilley, K. S. (2004). Localization of organelle proteins by isotope tagging (LOPIT). *Mol. Cell. Proteomics* 3, 1128–1134.

Dunkley, T. P. J., Hester, S., Shadforth, I. P., Runions, J., Weimar, T., Hanton, S. L., et al. (2006). Mapping the *Arabidopsis* organelle proteome. *Proc. Natl. Acad. Sci. U.S.A.* 103, 6518–6523.

Egelund, J., Petersen, B. L., Motawia, M. S., Damager, I., Faik, A., Olsen, C. E., et al. (2006). *Arabidopsis thaliana* RGXT1 and RGXT2 encode Golgi-localized (1,3)-alpha-D-xylosyltransferases involved in the synthesis of pectic rhamnogalacturonan-II. *Plant Cell* 18, 2593–2607.

Emanuelsson, O., Nielsen, H., Brunak, S., and Von Heijne, G. (2000). Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J. Mol. Biol.* 300, 1005–1016.

Eubel, H., Meyer, E. H., Taylor, N. L., Bussell, J. D., O'Toole, N., Heazlewood, J. L., et al. (2008). Novel proteins, putative membrane transporters, and an integrated metabolic network are revealed by quantitative proteomic analysis of *Arabidopsis* cell culture peroxisomes. *Plant Physiol.* 148, 1809–1829.

Feraru, E., Feraru, M. I., Asaoka, R., Paciorek, T., De Rycke, R., Tanaka, H., et al. (2012). BEX5/RabA1b regulates *trans*-Golgi network-to-plasma membrane protein trafficking in *Arabidopsis*. *Plant Cell* 24, 3074–3086.

Gao, C., Yu, C. K. Y., Qu, S., San, M. W. Y., Li, K. Y., Lo, S. W., et al. (2012). The Golgi-localized *Arabidopsis* endomembrane protein12 contains both endoplasmic reticulum export and Golgi retention signals at its C terminus. *Plant Cell* 24, 2086–2104.

Heazlewood, J. L., Verboom, R. E., Tonti-Filippini, J., Small, I., and Millar, A. H. (2007). SUBA: the *Arabidopsis* subcellular database. *Nucleic Acids Res.* 35, D213–D218.

Horton, P., Park, K. J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C. J., et al. (2007). WoLF PSORT: protein localization predictor. *Nucleic Acids Res.* 35, W585–W587.

Ingelsson, B., Shapiguzov, A., Kieselbach, T., and Vener, A. V. (2009). Peptidylprolyl isomerase activity in chloroplast thylakoid lumen is a dispensable function of immunophilins in *Arabidopsis thaliana*. *Plant Cell Physiol.* 50, 1801–1814.

Ito, J., Batth, T. S., Petzold, C. J., Redding-Johanson, A. M., Mukhopadhyay, A., Verboom, R., et al. (2011). Analysis of the *Arabidopsis* cytosolic proteome highlights subcellular partitioning of central plant metabolism. *J. Proteome Res.* 10, 1571–1582.

Kong, Y., Zhou, G., Yin, Y., Xu, Y., Pattathil, S., and Hahn, M. G. (2011). Molecular analysis of a family of *Arabidopsis* genes related to galacturonosyltransferases. *Plant Physiol.* 155, 1791–1805.

Latijnhouwers, M., Gillespie, T., Boevink, P., Kriechbaumer, V., Hawes, C., and Carvalho, C. M. (2007). Localization and domain characterization of *Arabidopsis* golgin candidates. *J. Exp. Bot.* 58, 4373–4386.

Lee, C. H., O'Neill, M. A., Tsumuraya, Y., Darvill, A. G., and Ye, Z. H. (2007). The irregular xylem9 mutant is deficient in xylan xylosyltransferase activity. *Plant Cell Physiol.* 48, 1624–1634.

Marmagne, A., Ferro, M., Meinnel, T., Bruley, C., Kuhn, L., Garin, J., et al. (2007). A high content in lipid-modified peripheral proteins and integral receptor kinases features in the *Arabidopsis* plasma membrane proteome. *Mol. Cell. Proteomics* 6, 1980–1996.

Mast, S., Peng, L. F., Jordan, T. W., Flint, H., Phillips, L., Donaldson, L., et al. (2010). Proteomic analysis of membrane preparations from developing *Pinus radiata* compression wood. *Tree Physiol.* 30, 1456–1468.

Matsuda, S., Vert, J. P., Saigo, H., Ueda, N., Toh, H., and Akutsu, T. (2005). A novel representation of protein sequences for prediction of subcellular location using support vector machines. *Protein Sci.* 14, 2804–2813.

Mitra, S. K., Walters, B. T., Clouse, S. D., and Goshe, M. B. (2009). An efficient organic solvent based extraction method for the proteomic analysis of *Arabidopsis* plasma membranes. *J. Proteome Res.* 8, 2752–2767.

Morre, D. J., and Mollenhauer, H. H. (1964). Isolation of Golgi apparatus from plant cells. *J. Cell Biol.* 23, 295–305.

Morre, D. J., and Mollenhauer, H. H. (2009). *The Golgi Apparatus: The First 100 Years.* New York: Springer.

Mouille, G., Ralet, M. C., Cavelier, C., Eland, C., Effroy, D., Hematy, K., et al. (2007). Homogalacturonan synthesis in *Arabidopsis thaliana* requires a Golgi-localized protein with a putative methyltransferase domain. *Plant J.* 50, 605–614.

Nikolovski, N., Rubtsov, D., Segura, M. P., Miles, G. P., Stevens, T. J., Dunkley, T. P. J., et al. (2012). Putative glycosyltransferases and other plant Golgi apparatus proteins are revealed by LOPIT proteomics. *Plant Physiol.* 160, 1037–1051.

Niu, B., Jin, Y. H., Feng, K. Y., Lu, W. C., Cai, Y. D., and Li, G. Z. (2008). Using AdaBoost for the prediction of subcellular location of prokaryotic and eukaryotic proteins. *Mol. Divers.* 12, 41–45.

Ondzighi, C. A., Christopher, D. A., Cho, E. J., Chang, S. C., and Staehelin, L. A. (2008). *Arabidopsis* protein disulfide isomerase-5 inhibits cysteine proteases during trafficking to vacuoles before programmed cell death of the endothelium in developing seeds. *Plant Cell* 20, 2205–2220.

Osterrieder, A. (2012). Tales of tethers and tentacles: golgins in plants. *J. Microsc.* 247, 68–77.

Paredez, A. R., Somerville, C. R., and Ehrhardt, D. W. (2006). Visualization of cellulose synthase demonstrates functional association with microtubules. *Science* 312, 1491–1495.

Parsons, H. T., Christiansen, K., Knierim, B., Carroll, A., Ito, J., Batth, T. S., et al. (2012a). Isolation and proteomic characterization of the *Arabidopsis* Golgi defines functional and novel components involved in plant cell wall biosynthesis. *Plant Physiol.* 159, 12–26.

Parsons, H. T., González Fernández-Niño, S. M., and Heazlewood, J. L. (2012b). "Separation of the plant Golgi apparatus and endoplasmic reticulum by free-flow electrophoresis," in *Plant Proteomics: Methods and Protocols*, 2 Edn, eds J. V. Jorrín-Novo, S. Komatsu, W. Weckwerth, and S. Weinkoop (New York: Springer) (in press).

Parsons, H. T., Ito, J., Park, E., Carroll, A. W., Joshi, H. J., Petzold, C. J., et al. (2012c). "The current state of the Golgi proteomes," in *Proteomic Applications in Biology*, eds J. L. Heazlewood and C. J. Petzold (Rijeka: InTech), 167–188.

Pena, M. J., Zhong, R. Q., Zhou, G. K., Richardson, E. A., O'Neill, M. A., Darvill, A. G., et al. (2007). *Arabidopsis* irregular xylem8 and irregular xylem9: implications for the complexity of glucuronoxylan biosynthesis. *Plant Cell* 19, 549–563.

Pierleoni, A., Martelli, P. L., Fariselli, P., and Casadio, R. (2006). BaCelLo: a balanced subcellular localization predictor. *Bioinformatics* 22, e408–e416.

Rautengarten, C., Ebert, B., Herter, T., Petzold, C. J., Ishii, T., Mukhopadhyay, A., et al. (2011). The interconversion of UDP-arabinopyranose and UDP-arabinofuranose is indispensable for plant development in *Arabidopsis*. *Plant Cell* 23, 1373–1390.

Saint-Jore-Dupas, C., Nebenfuhr, A., Boulaflous, A., Follet-Gueye, M. L., Plasson, C., Hawes, C., et al. (2006). Plant N-glycan processing enzymes employ different targeting mechanisms for their spatial arrangement along the secretory pathway. *Plant Cell* 18, 3182–3200.

Sanderfoot, A. A., and Raikhel, N. V. (1999). The specificity of vesicle trafficking: coat proteins and SNAREs. *Plant Cell* 11, 629–641.

Small, I., Peeters, N., Legeai, F., and Lurin, C. (2004). Predotar: a tool for rapidly screening proteomes for N-terminal targeting sequences. *Proteomics* 4, 1581–1590.

Surpin, M., and Raikhel, N. (2004). Traffic jams affect plant development and signal transduction. *Nat. Rev. Mol. Cell Biol.* 5, 100–109.

Tamura, T., and Akutsu, T. (2007). Subcellular location prediction of proteins using support vector machines with alignment of block sequences utilizing amino acid composition. *BMC Bioinformatics* 8:466. doi: 10.1186/1471-2105-8-466

Tanaka, N., Fujita, M., Handa, H., Murayama, S., Uemura, M., Kawamura, Y., et al. (2004). Proteomics of the rice cell: systematic identification of the protein populations in subcellular compartments. *Mol. Genet. Genomics* 271, 566–576.

Taylor, R. S., Jones, S. M., Dahl, R. H., Nordeen, M. H., and Howell, K. E. (1997). Characterization of the Golgi complex cleared of proteins in transit and examination of calcium uptake activities. *Mol. Biol. Cell* 8, 1911–1931.

Woollard, A. A., and Moore, I. (2008). The functions of Rab GTPases in plant membrane traffic. *Curr. Opin. Plant Biol.* 11, 610–619.

Zhang, Z. J., and Peck, S. C. (2011). Simplified enrichment of plasma membrane proteins for proteomic analyses in *Arabidopsis thaliana*. *Proteomics* 11, 1780–1788.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 30 October 2012; paper pending published: 20 November 2012; accepted: 12 December 2012; published online: 03 January 2013.

Citation: Parsons HT, Drakakaki G and Heazlewood JL (2013) Proteomic dissection of the Arabidopsis Golgi and trans-Golgi network. Front. Plant Sci. ***3****:298. doi: 10.3389/fpls.2012.00298*

This article was submitted to Frontiers in Plant Proteomics, a specialty of Frontiers in Plant Science.

Copyright © 2013 Parsons, Drakakaki and Heazlewood. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.

frontiers in
PLANT SCIENCE

MINI REVIEW ARTICLE
published: 25 January 2013
doi: 10.3389/fpls.2013.00004

Recent advances in the composition and heterogeneity of the *Arabidopsis* mitochondrial proteome

Chun Pong Lee[1]*, Nicolas L. Taylor[2,3] and A. Harvey Millar[2,3]

[1] *Department of Plant Sciences, University of Oxford, Oxford, UK*
[2] *ARC Centre of Excellence in Plant Energy Biology, The University of Western Australia, Crawley, WA, Australia*
[3] *Centre for Comparative Analysis of Biomolecular Networks, The University of Western Australia, Crawley, WA, Australia*

Edited by:
Joshua L. Heazlewood, Lawrence Berkeley National Laboratory, USA

Reviewed by:
Sebastien Carpentier, KU Leuven, Belgium
Christian Lindermayr, Helmholtz Zentrum München – German Research Center for Environmental Health, Germany

****Correspondence:***
Chun Pong Lee, Department of Plant Sciences, University of Oxford, South Parks Road, Oxford OX1 3RB, UK.
e-mail: chun.lee@plants.ox.ac.uk

Mitochondria are important organelles for providing the ATP and carbon skeletons required to sustain cell growth. While these organelles also participate in other key metabolic functions across species, they have a specialized role in plants of optimizing photosynthesis through participating in photorespiration. It is therefore critical to map the protein composition of mitochondria in plants to gain a better understanding of their regulation and define the uniqueness of their metabolic networks. To date, <30% of the predicted number of mitochondrial proteins has been verified experimentally by proteomics and/or GFP localization studies. In this mini-review, we will provide an overview of the advances in mitochondrial proteomics in the model plant *Arabidopsis thaliana* over the past 5 years. The ultimate goal of mapping the mitochondrial proteome in *Arabidopsis* is to discover novel mitochondrial components that are critical during development in plants as well as genes involved in developmental abnormalities, such as those implicated in mitochondrial-linked cytoplasmic male sterility.

Keywords: *Arabidopsis thaliana*, mitochondria, proteomics, heterogeneity, protein complex, post-translational modifications, functional proteomics

INTRODUCTION

Mitochondria are semi-autonomous, double membrane bound organelles with unique morphologies and highly specialized functions. While these organelles are well-recognized for energy metabolism via coupling the oxidation of organic acids with oxidative phosphorylation (OXPHOS), they also have diverse functional roles such as metabolism of amino acids and biosynthesis of cofactors and vitamins. Mitochondria in plants are set apart from their mammalian counterparts by their mediation of photosynthesis through providing alternative electron sinks for photosynthetic products and their participating in photorespiration (Padmasree et al., 2002). In order to fully understand the functional roles of mitochondria in photosynthetic cells, it is essential to establish their total protein make-up (proteome) and their post-translational modifications (PTMs), as well as to generate a protein atlas that collects information about mitochondrial protein expression patterns during stress and in different cells, tissues, and organs.

Arabidopsis thaliana became the first model system for plants after its genome was fully sequenced and made publicly available in 2000 (The Arabidopsis Genome Initiative, 2000). In the last decade, tremendous progress has been made, by both experimental and bioinformatics approaches, to define the mitochondrial proteome in this model plant species. Like its yeast and mammalian counterparts, most of the mitochondrial proteins in *Arabidopsis* are encoded by the nuclear genome. Based on the analyses of the N-terminal targeting peptide sequences in *Arabidopsis*, there are about 2500 predicted nuclear-encoded mitochondrial proteins (representing 7–10% of all encoded proteins) with broad functional roles (Heazlewood et al., 2007; Cui et al., 2011). In comparison, the mitochondrial genome encodes for only 57 gene products (Unseld et al., 1997). The first extensive experimental studies of the mitochondrial proteome in *Arabidopsis* identified ∼100–150 proteins (Kruft et al., 2001; Millar et al., 2001; Werhahn and Braun, 2002; Millar and Heazlewood, 2003). Improvement of organelle purification procedure, availability of different protein mapping strategies, enhanced sensitivity of peptide detection by mass spectrometry (MS), and improved genomic resources and peptide identification software have driven a significant increase in the number of mitochondrial proteins identified across different model species – from 843 in *Arabidopsis* (Table S1A in Supplementary Material) and 851 in yeast (Reinders et al., 2006), to 1404 in mouse (Forner et al., 2009).

Given the number of proteins identified so far in *Arabidopsis* mitochondrion, it is clear that our understanding of its composition and functions in plants is far from complete. In this mini-review, we will provide an update on the status of *Arabidopsis* mitochondrial proteomics research based on published data in the past 5 years (2007–2012). We would refer readers to previous review articles for more comprehensive overviews on the progress of plant mitochondrial proteomics in the preceding years (Millar et al., 2005, 2011; Ito et al., 2007; Dudkina et al., 2010).

HOW FAR ARE WE FROM COMPILING THE COMPLETE SET OF *ARABIDOPSIS* MITOCHONDRIAL PROTEINS?

A recent in-depth analysis of the proteome in Percoll-purified mitochondria has identified a non-redundant set of 572 proteins in *Arabidopsis* cell culture (Taylor et al., 2011). With a combined

proteomics, localization experiment and literature confirmation approach, a set of 38 mitochondrial proteins have been found in or associated with the mitochondrial outer membrane (Duncan et al., 2011). More recently, a total of 66 novel integral membrane proteins have been identified in mitochondria using a MS-based quantitative enrichment approach (Tan et al., 2012). A new set of components with unknown functions have also been identified in a number of recent studies, including the analysis of the mitochondrial fraction from: (i) separated protein complexes (Klodmann et al., 2010, 2011; Klodmann and Braun, 2011; Schertl et al., 2012); (ii) enriched phospho-proteome (Ito et al., 2009); (iii) different tissue types (Lee et al., 2012); (iv) various time points of a diurnal cycle (Lee et al., 2010); and (v) cells subjected to biotic stress (Livaja et al., 2008).

While various large-scale proteomics studies over the last 5 years have led to the identification of a non-redundant set of 843 putative mitochondrial proteins (Table S1A in Supplementary Material), it remains difficult to discriminate true mitochondrial proteins from contaminants, particularly for low abundant proteins, in a sample. It has been estimated that about 11% of the total spot intensity on a 2-D map of mitochondria from *Arabidopsis* cell culture are proteins originated from other compartments (Taylor et al., 2011). By querying previous evidence from literature and/or consensus subcellular localization prediction score from publicly available databases [SUBA3 (Heazlewood et al., 2007); ARAMEMNON7.0 (Schwacke et al., 2003)], we define a set of 504 proteins which can be assigned to be mitochondrial-localized in *Arabidopsis* with high confidence (Table S1B in Supplementary Material). This approach is biased toward proteins that are highly abundant and does not explicitly imply that the remaining proteins are in fact contaminants from other compartments. Some of these proteins lack a predictable targeting presequence, may be dual-targeted to multiple compartments and/or are present in relatively low amount, thus their localization should be confirmed in the future through multiple independent proteomic analyses and/or by fluorescent protein localization.

According to the SUBA database, a number of GFP tagging studies have revealed the mitochondrial localization of 222 proteins that cannot be identified through proteomic approaches (Table S1C in Supplementary Material), most of which are low abundance proteins involved in the processing and maintenance of the mitochondrial genome. Together with the proteomics set, 726 proteins can be confidently assigned as mitochondrial, <30% of the presumed number of predicted proteins. To further expand the current *Arabidopsis* mitochondrial protein compendium, it is essential to overcome the challenge of identifying low abundance proteins. To achieve this a number of approaches could be employed including protein enrichment tools, such as proteominer (Fröhlich et al., 2012) or protein fractionation approaches including strong cation exchange (SCX) or off-gel electrophoresis (OGE) prior to RP-LC-MS (Chenau et al., 2008; Ito et al., 2011). Together with biological fractionation approaches such as investigation of pre-fractionated sub-mitochondrial compartments or enrichment by metal or co-factor binding approaches and advances in LC-MS techniques and equipment, it is likely that an increasing number of low abundance proteins will be revealed.

FUNCTIONS OF THE MITOCHONDRIAL PROTEOME IN *ARABIDOPSIS*

MITOCHONDRIAL PROTEIN FUNCTIONS AND ABUNDANCE

Of the confirmed set of mitochondrial proteins (Table S1B in Supplementary Material), ∼22% are components of pyruvate metabolism/TCA cycle and OXPHOS, while a similar number (∼20%) are identified as subunits of machinery for mitochondrial gene expression and maintenance (**Figure 1A**). In the yeast mitochondrial proteome, a similar proportion (∼15%) of identified proteins are involved in energy metabolism (Schmidt et al., 2010). When comparing the abundance of proteins in these functional categories using the recently published LC-MS/MS data (Taylor et al., 2011), energy metabolism comprises over 50% of the total protein abundance in mitochondria, whereas <2% is associated with processing mitochondrial DNA/RNA (**Figure 1B**). The observed abundance of proteins in energy metabolism is consistent with the main role of mitochondria in the cell and bulk of the chemical reactions performed in the organelle; in contrast, the low abundance of proteins for mitochondrial DNA/RNA processing can probably be attributed to their relatively less stable nature so that they can respond rapidly to external stimuli or to changes in energy cost (Schwanhausser et al., 2011), the transient need for their functions during the life of cells and presumably the high specific activity of their functions. At the whole cellular level, components in this functional category have recently been shown to have a high turnover rate in *Arabidopsis* (Li et al., 2012). Mitochondrial proteins involving nucleic acid processing appear to perform highly specialized functions and do not seem to have overlapping specificity. Only ∼12% of the proteins in the yeast mitochondrial proteome are dedicated to genome maintenance and processing (Schmidt et al., 2010). The proportion is higher in *Arabidopsis* due to the presence of multiple plant-specific pentatricopeptide repeat (PPR) proteins and/or its larger genome size which may require more proteins to maintain and process. Each PPR protein recognizes and acts on a single site in a specific transcript sequence (Delannoy et al., 2007).

Several of the unknown proteins identified by our earlier study (Heazlewood et al., 2004) have since been re-assigned as plant-specific components of OXPHOS (Klodmann et al., 2011). The most nebulous subset of the known proteome is the more than 18% of the identified proteins that remain without any functional class. However, while this subset are great in number they contribute to <2% of mitochondrial protein abundance. Interestingly, these include a number of plant-specific proteins. It is therefore clear that many more studies are required to elucidate the functions of this subset of proteins which can potentially lead to the discovery of novel plant-specific mitochondrial metabolic pathways/functions.

PROTEIN COMPLEXES AND INTERACTOME

Multiple proteins/isoforms often assembled into large complexes which serve vital metabolic and regulatory roles. While earlier reports have extensively analyzed the structure and function of individual enzyme complexes of interest, such as glycine decarboxylase complex (Douce et al., 2001), it is uncertain whether other mitochondrial proteins could also organize into macromolecular structures. Using 2-D blue-native/SDS-PAGE, Klodmann et al. (2011) found 35 different protein complexes in mitochondria

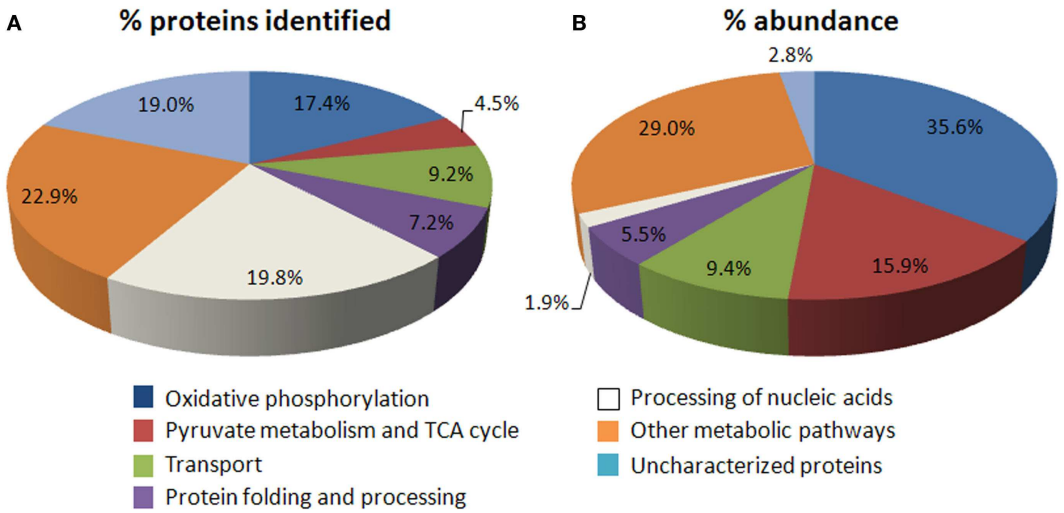


FIGURE 1 | Overview of the number and abundance of mitochondrial proteins across functional categories. (A) Pie chart showing of functional categories of the confirmed set of 726 mitochondrial proteins (see Table S1B,C in Supplementary Material). A comparison with the more complete yeast mitochondrial proteome shows that similar proportion of proteins involving energy metabolism as well as proteins with unknown functions has been found (Schmidt et al., 2010). In addition, more proteins are involved in mitochondrial genome maintenance (white) in plants (~20%) than in yeast (~12%), due to the presence of numerous plant-specific pentatricopeptide repeat (PPR) proteins and a larger genome size. **(B)** Distribution of the abundance of proteins that can be identified by gel-free MS (Taylor et al., 2011) across seven functional categories.

from *Arabidopsis* cell culture. OXPHOS complexes are amongst the largest and the most abundant protein complexes in mitochondria. Mitochondrial complex assemblies are also dominated by components in the TCA cycle, amino acid metabolism, PPR proteins, and pre-protein import apparatus. While the preliminary compositions of these proteins complexes have been proposed based on the number of subunits identified and their migration on the first and second dimension, they must be verified through independent biochemical analysis.

A number of mitochondrial proteins of diverse function have been identified to interact with metal ions (Tan et al., 2010) and/or have binding affinity with ATP (Ito et al., 2006) in *Arabidopsis*. In contrast, studies on the more transient direct interactions (functional and physical) between multiple mitochondrial proteins in plants are lacking. Such detailed studies in the future will lead to the construction of plant mitochondrial interactome, to sit alongside side the complexome, and help to define unique metabolic regulations in plants that differentiate them from yeast and mammals.

POST-TRANSLATIONAL MODIFICATIONS

The complexity of *Arabidopsis* mitochondrial proteome is further implicated by the dynamic regulation of PTMs which can control activity, stability, and structural characteristics of proteins. Proteins with PTMs often appear as multiple spots with different pI and/or molecular mass on a 2-D gel, and the region of a peptide with modified residues can be detected as an altered m/z ion species by MS. Recent large-scale proteomic studies have reported a number of PTMs in *Arabidopsis* mitochondrial proteome (**Table 1**), including oxidation (Tan et al., 2010; Solheim et al., 2012), phosphorylation (Ito et al., 2009; Taylor et al., 2011), S-nitrosylation (Palmieri et al., 2010), N-terminal acetylation (Huang et al., 2009), and lysine acetylation (Finkemeier et al., 2011). However, there appears to be no evidence for specific preference of PTMs to particular functional categories of identified proteins (**Table 1**), suggesting that PTMs have a wide variety of functional targets in the mitochondrion.

The total number of identified proteins with PTMs is very likely a gross underestimation due to a number of technical challenges, such as the loss of PTMs during mitochondrial purification procedures and the relatively low abundance of the modified peptides compared to their unmodified counterparts. Also, it is not clear how many proteins, including those listed in **Table 1**, are functionally modified through enzyme-catalyzed mitochondrial processes *in vivo*. For example, degradation products observed on a 2-D gel often perceive as artificial post-purification events. These concerns can be at least partially overcome by enrichment of modified peptides/proteins and/or repeat analysis of multiple replicates to ensure that similar changes can be observed in all samples. Alternatively, the incorporation of radioactive tracers into proteins *in vivo* (cells) or *in vitro* (isolated mitochondria) can be used to identify proteins with reversible PTMs. For instance, 18 phosphoproteins have recently been identified by [γ^{32}P]-ATP labeling and affinity enrichment of isolated mitochondria (Ito et al., 2009).

CHANGES IN THE MITOCHONDRIAL PROTEOME IN DIFFERENT TISSUES AND IN RESPONSE TO OXIDATIVE STRESS

The mitochondrial proteome is not static, but has many components that are dynamically regulated in order to meet energy and metabolic needs required by the cell in response to developmental and/or environmental changes. There are many different cell/tissue/organ types which have functions that are unique to plants. Thus, mitochondrial composition, metabolism, and stress

Table 1 | The set of mitochondrial proteins in *Arabidopsis* with known post-translational modifications.

AGI accession	Protein ID	PTM	Reference
OXIDATIVE PHOSPHORYLATION			
AT1G15120	Complex III QCR6-1, Hinge protein	N-acetylation	Huang et al. (2009)
AT1G22840	Cytochrome	Lys acetylation	Finkemeier et al. (2011)
		N-acetylation	Huang et al. (2009)
AT1G47420	SDH5 succinate dehydrogenase subunit 5	Phosphorylation	Ito et al. (2009)
AT1G51980	MPP alpha-1 mitochondrial processing peptidase alpha subunit	Oxidation	Solheim et al. (2012)
AT3G12260	NADH dehydrogenase B14 subunit	N-acetylation	Huang et al. (2009)
AT3G62810	Complex 1 family protein/LVR family protein	N-acetylation	Huang et al. (2009)
AT4G21105	COX X4	Oxidation	Solheim et al. (2012)
AT5G08670	ATP synthase beta subunit	Oxidation	Solheim et al. (2012)
AT5G52840	NADH dehydrogenase B13 subunit	Phosphorylation	Taylor et al. (2011)
AT5G66760	SDH1-1 succinate dehydrogenase flavoprotein subunit	Phosphorylation	Ito et al. (2009)
ATMG01190	ATP synthase alpha-1 subunit	S-Nitrosylation	Palmieri et al. (2010)
		Oxidation	Solheim et al. (2012)
PYRUVATE METABOLISM AND TCA CYCLE			
AT1G24180	E1 alpha-2 (pyruvate dehydrogenase)	Phosphorylation	Ito et al. (2009)
AT1G48030	mtLPD-1 (mtlipoamide dehydrogenase-1)	S-Nitrosylation	Palmieri et al. (2010)
		Oxidation	Solheim et al. (2012)
AT1G53240	Malate dehydrogenase-1	Oxidation	Solheim et al. (2012)
At1G59900	E1 alpha-1 (pyruvate dehydrogenase)	Phosphorylation	Ito et al. (2009)
AT2G05710	Aconitate hydratase-2	Phosphorylation	Taylor et al., 2011; ref therein
		Oxidation	Tan et al. (2010)
AT2G20420	Succinyl-CoA ligase (GDP-forming) beta-chain	Oxidation	Solheim et al. (2012)
AT3G13930	E3-1 (dihydrolipoamide dehydrogenase)	Phosphorylation	Taylor et al., 2011; ref therein
AT3G15020	Malate dehydrogenase-2	Oxidation	Solheim et al. (2012)
AT3G17240	mtLPD-2 (mtlipoamide dehydrogenase-2)	S-Nitrosylation	Palmieri et al. (2010)
AT4G26970	Aconitate hydratase-1	Oxidation	Tan et al. (2010)
AT5G03290	Isocitrate dehydrogenase-5	Oxidation	Solheim et al. (2012)
TRANSPORT			
AT2G29530	Translocase inner membrane subunit 10, TIM10	N-acetylation	Huang et al. (2009)
AT3G08580	AAC1 (ADP/ATP carrier 1)	Oxidation	Solheim et al. (2012)
AT3G46560	Translocase inner membrane subunit 9, TIM9	N-acetylation	Huang et al. (2009)
AT5G13490	AAC2 (ADP/ATP carrier 2)	Oxidation	Solheim et al. (2012)
AT5G14040	mt phosphate transporter	Oxidation	Solheim et al. (2012)
AT5G50810	Translocase inner membrane subunit 8, TIM8	N-acetylation	Huang et al. (2009)
NUCLEIC ACID PROCESSING AND PROTEIN FOLDING AND STABILITY			
AT1G74230	GR-RBP5 (glycine-rich RNA-binding protein 5)	Phosphorylation	Ito et al. (2009)
AT3G13160	PPR8-2	Oxidation	Solheim et al. (2012)
AT3G23990	HSP60-3B	Oxidation	Solheim et al. (2012)
AT4G02930	Elongation factor Tu	Oxidation	Solheim et al. (2012)
AT4G26780	Co-chaperone grpE	Phosphorylation	Ito et al. (2009)
AT4G37910	Heat shock protein HSP70-1	Phosphorylation	Taylor et al., 2011; ref therein
AT5G26860	LON1 (LON protease 1)	Phosphorylation	Ito et al. (2009)
AT5G40770	Prohibitin-3	N-acetylation	Huang et al. (2009)
AT5G61030	GR-RBP3 (glycine-rich RNA-binding protein 3)	Phosphorylation	Ito et al. (2009)
PHOTORESPIRATION			
AT1G11860	GDT1 aminomethyltransferase	S-Nitrosylation	Palmieri et al. (2010)
		Phosphorylation	Taylor et al. (2011)
AT1G32470	GDH Glycine decarboxylase H subunit	S-Nitrosylation	Palmieri et al. (2010)
		Oxidation	Solheim et al. (2012)
AT2G35370	GDH Glycine decarboxylase H subunit	S-Nitrosylation	Palmieri et al. (2010)
		Oxidation	Solheim et al. (2012)

(Continued)

Table 1 | Continued

AGI accession	Protein ID	PTM	Reference
AT4G33010	Glycine decarboxylase P-protein 1	S-Nitrosylation	Palmieri et al. (2010)
		Oxidation	Solheim et al. (2012)
AT4G37930	SHM1 (serine transhydroxymethyltransferase 1)	S-Nitrosylation	Palmieri et al. (2010)
		Oxidation	Solheim et al. (2012)
AT5G26780	SHM2 (serine hydroxymethyltransferase 2)	Oxidation	Solheim et al. (2012)
OTHER METABOLISM			
AT3G61440	Cyanoalanine synthase	Phosphorylation	Taylor et al., 2011; ref therein
AT3G22200	4-Aminobutyrate aminotransferase	Oxidation	Solheim et al. (2012)
AT4G15940	Fumarylacetoacetate hydrolase family protein	N-acetylation	Huang et al. (2009)
AT5G07440	GDH2 (glutamate dehydrogenase-2)	Phosphorylation	Ito et al. (2009)
		N-acetylation	Huang et al. (2009)
		Oxidation	Solheim et al. (2012)
AT5G14780	FDH Formate dehydrogenase	Phosphorylation	Ito et al. (2009)
		Oxidation	Solheim et al. (2012)
AT5G18170	GDH1 (glutamate dehydrogenase-1)	Phosphorylation	Ito et al. (2009)
		N-acetylation	Huang et al. (2009)
AT5G50370	Adenylate kinase family	Phosphorylation	Taylor et al. (2011)
		N-acetylation	Huang et al. (2009)
AT5G63400	ADK1 Adenylate kinase 1	N-acetylation	Huang et al. (2009)
PROTEINS WITH UNKNOWN FUNCTIONS			
AT2G39795	mt glycoprotein family protein	Phosphorylation	Ito et al. (2009)
AT3G18240	Unknown protein	Phosphorylation	Ito et al. (2009)
AT3G55605	Mitochondrial glycoprotein family protein	Phosphorylation	Ito et al. (2009)
AT4G21460	Unknown protein	Phosphorylation	Ito et al. (2009)
AT4G27585	Stomatin-like protein	Phosphorylation	Ito et al. (2009)
AT4G23885	Unknown protein	N-acetylation	Huang et al. (2009)

PTM, post-translational modification(s).

response in these cells/tissues/organs from *Arabidopsis* will be different from what has been observed in yeast and animals.

Analysis of the mitochondria proteome from photosynthetic shoots, non-photosynthetic cell culture, and roots identified major differences in the abundance of enzymes of the TCA cycle and photorespiration (Lee et al., 2008, 2011). Quantitative comparison of the mitochondrial proteome across 10 different time points covering 24-h of the life of *Arabidopsis* shoots also uncovers day (photosynthetic)- and night (non-photosynthetic)-enhanced proteins in central carbon metabolism (Lee et al., 2010). In these studies, the abundances of OXPHOS complexes in purified mitochondria generally remain unaltered but their respiratory capacity differs depending on the choice and/or availability of substrates (Lee et al., 2008, 2011). However, on a whole tissue basis differences in mitochondrial electron transport chain complex ratios between tissues has been reported (Peters et al., 2012). Lee et al. (2012) have reported changes in the isolated *Arabidopsis* mitochondrial proteome beyond differences in the cellular photosynthetic capacity. Changes in the abundance of a wide variety of mitochondrial proteins can be observed from cells/tissues from various vegetative and reproductive phases of development. Differences in protein accumulation and metabolic specializations of these mitochondria generally coincide with the main physiological role of each corresponding tissue type, such as glycine cleavage via photorespiration in shoot and maintenance of mitochondrial redox environment in flowers. In mouse, it has been reported that just over half of all proteins identified by gel-free MS approach can be found in all the investigated organs (Pagliarini et al., 2008). However, the number of mitochondrial proteins that are highly tissue-specific (i.e., totally absent in at least one tissue) in *Arabidopsis* remains to be defined. Such analysis will assist in identifying mitochondrial components that causes plant-specific developmental phenotypes, e.g., cytoplasm male sterility.

Using a gel-free quantitative MS approach, Tan et al. (2012) recently identify a number of integral membrane proteins in mitochondria that are altered in abundance in response to cold and/or various chemical stresses. These proteins include the components of the alternative NADH dehydrogenases, alternative oxidase, and uncoupling proteins, but also several stress-sensitive subunits within the OXPHOS complexes. Together with a similar study by Sweetlove et al. (2002), it is concluded that the reduction in respiration in response to chemical-induced oxidative stress is a consequence of coordinated changes in the mitochondrial proteome, particularly OXPHOS complex subunits and stress-related components.

APPLICATION OF PROTEOMICS TO ANALYZE MITOCHONDRIAL PROTEIN FUNCTIONS

Over the last decade, advances in the understanding of mitochondrial composition and protein complex assembly have led to the

identification of many genes associated with genetic diseases in humans (Calvo and Mootha, 2010). In contrast, plant proteomics still needs to discover novel mitochondrial components that are associated with known developmental defects in plants. Nevertheless, by combining proteomics and reverse-genetics strategies, a number of recent studies have highlighted the unique role of a mitochondrial component of interest in *Arabidopsis* that had not been unraveled by other biochemical and molecular techniques. Metabolite analyses of malate dehydrogenase (MDH) antisense and knockout lines in tomato and *Arabidopsis* respectively show an elevated foliar ascorbate level (Nunes-Nesi et al., 2005; Tomaz et al., 2010). Such accumulation coincides with the reduction of Complex I-associated galacton-1,4-lactone dehydrogenase (GLDH) abundance in the mitochondrial proteome of a MDH double mutant (*mmdh1mmdh2*; Tomaz et al., 2010), indicating that there might be a complex metabolic regulation/interaction between OXPHOS, TCA cycle, and cellular ascorbate biosynthesis. A mutation in mitochondrial Lon protease leads to a retarded growth phenotype (Rigas et al., 2009), which can be explained by an altered abundance of enzymes in the TCA cycle and OXPHOS, a decrease in the abundance of breakdown products and a small increase in the number of proteins with oxidized peptides, but not by heightened oxidative stress (Solheim et al., 2012). In contrast, knockout of the protease AtFtsH4 does not significantly affect *Arabidopsis* growth under long day conditions, but changes rosette development under short-day conditions (Gibala et al., 2009). The phenotypes correlate with elevated levels of oxidative stress, increased abundance of Hsp70 and prohibitins, and decreased abundance of ATP synthase subunits.

CONCLUSION AND PERSPECTIVES

The availability of the full genome sequence of *Arabidopsis* for more than a decade, advances in various proteomic technologies, as well as their wider adoption, have provided an opportunity to understand the protein make-up of mitochondria and their underlying metabolism in this plant more than in any other. Significant progress in extracting information on PTMs and protein abundances has also improved our insight into the dynamic regulation of the mitochondrial proteome in a cellular/organismal context. However, further work is needed to characterize mitochondrial proteins according to their sub-organellar localization. In-depth identification of components in the intermembrane space has not been reported since the improvements in MS analysis in recent years. Recent discoveries of a pyruvate transporter (Herzig et al., 2012) and a calcium uniporter (Baughman et al., 2011) in mouse mitochondria have been conducted through an integrated proteomics, bioinformatics, and genetics strategy. Thus, identification of low abundance proteins should allow us to complete the catalog of mitochondrial proteins in *Arabidopsis*, which will provide us several candidates for identifying plant-specific transporters or metabolic pathways by a similar approach.

ACKNOWLEDGMENTS

This work was funded through a grant to the ARC Centre of Excellence in Plant Energy Biology (CE0561495; A. Harvey Millar). A. Harvey Millar is funded as an ARC Australian Future Fellow (FT110100242) and Chun Pong Lee is a receipt of an EMBO Fellowship (ALTF1140-2011).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at http://www.frontiersin.org/Plant_Proteomics/10.3389/fpls.2013.00004/abstract

Table S1 | Mitochondrial proteins identified by mass spectrometry or fluorescent protein localization studies. (A) All proteins identified from isolated mitochondria from *Arabidopsis* using proteomics in the last 5 years. **(B)** A set of proteins which has a high probably being located in the mitochondrion. For the inclusion of a protein in the list, it should meet the following criteria: (i) A protein is automatically considered mitochondrial if at least two studies have identified it in isolated mitochondrial fraction. However, a protein is considered to be non-mitochondrial if it is identified in equal number of or more non-mitochondrial proteomics studies than the mitochondrial ones. (ii) If the location of a protein is verified independently by fluorescent protein localization analysis, then (ii) is ignored and it is included in the list. (iii) If a protein is identified by one study, the localization based on SUBAcon score and/or ARAMEMNON localization consensus score is also considered. **(C)** Mitochondrial protein confirmed through fluorescent protein localization studies (according to SUBA) only and not by proteomics.

REFERENCES

Baughman, J. M., Perocchi, F., Girgis, H. S., Plovanich, M., Belcher-Timme, C. A., Sancak, Y., et al. (2011). Integrative genomics identifies MCU as an essential component of the mitochondrial calcium uniporter. *Nature* 476, 341–345.

Calvo, S. E., and Mootha, V. K. (2010). The mitochondrial proteome and human disease. *Annu. Rev. Genomics Hum. Genet.* 11, 25–44.

Chenau, J., Michelland, S., Sidibe, J., and Seve, M. (2008). Peptides OFFGEL electrophoresis: a suitable pre-analytical step for complex eukaryotic samples fractionation compatible with quantitative iTRAQ labeling. *Proteome Sci.* 6, 9.

Cui, J., Liu, J., Li, Y., and Shi, T. (2011). Integrative identification of Arabidopsis mitochondrial proteome and its function exploitation through protein interaction network. *PLoS ONE* 6:e16022. doi:10.1371/journal.pone.0016022

Delannoy, E., Stanley, W. A., Bond, C. S., and Small, I. D. (2007). Pentatricopeptide repeat (PPR) proteins as sequence-specificity factors in post-transcriptional processes in organelles. *Biochem. Soc. Trans.* 35, 1643–1647.

Douce, R., Bourguignon, J., Neuburger, M., and Rébeillé, F. (2001). The glycine decarboxylase system: a fascinating complex. *Trends Plant Sci.* 6, 167–176.

Dudkina, N. V., Kouril, R., Peters, K., Braun, H.-P., and Boekema, E. J. (2010). Structure and function of mitochondrial supercomplexes. *Biochim. Biophys. Acta* 1797, 664–670.

Duncan, O., Taylor, N. L., Carrie, C., Eubel, H., Kubiszewski-Jakubiak, S., Zhang, B., et al. (2011). Multiple lines of evidence localize signaling, morphology, and lipid biosynthesis machinery to the mitochondrial outer membrane of Arabidopsis. *Plant Physiol.* 157, 1093–1113.

Finkemeier, I., Laxa, M., Miguet, L., Howden, A. J. M., and Sweetlove, L. J. (2011). Proteins of diverse function and subcellular location are lysine acetylated in Arabidopsis. *Plant Physiol.* 155, 1779–1790.

Forner, F., Kumar, C., Luber, C. A., Fromme, T., Klingenspor, M., and Mann, M. (2009). Proteome differences between brown and white fat mitochondria reveal specialized metabolic functions. *Cell Metab.* 10, 324–335.

Fröhlich, A., Gaupels, F., Sarioglu, H., Holzmeister, C., Spannagl, M., Durner, J., et al. (2012). Looking deep inside: detection of low-abundance proteins in leaf extracts of Arabidopsis and phloem exudates of pumpkin. *Plant Physiol.* 159, 902–914.

Gibala, M., Kicia, M., Sakamoto, W., Gola, E. M., Kubrakiewicz, J., Smakowska, E., et al. (2009). The lack of mitochondrial AtFtsH4 protease alters Arabidopsis leaf morphology at the late stage of rosette development under short-day photoperiod. *Plant J.* 59, 685–699.

Heazlewood, J. L., Tonti-Filippini, J. S., Gout, A. M., Day, D. A., Whelan, J., and Millar, A. H. (2004). Experimental analysis of the Arabidopsis mitochondrial proteome highlights signaling and regulatory components, provides assessment of targeting prediction programs, and indicates plant-specific mitochondrial proteins. *Plant Cell* 16, 241–256.

Heazlewood, J. L., Verboom, R. E., Tonti-Filippini, J., Small, I., and Millar, A. H. (2007). SUBA: the Arabidopsis subcellular database. *Nucleic Acid Res.* 35, D213–D218.

Herzig, S., Raemy, E., Montessuit, S., Veuthey, J.-L., Zamboni, N., Westermann, B., et al. (2012). Identification and functional expression of the mitochondrial pyruvate carrier. *Science* 337, 93–96.

Huang, S., Taylor, N. L., Whelan, J., and Millar, A. H. (2009). Refining the definition of plant mitochondrial presequences through analysis of sorting signals, N-terminal modifications, and cleavage motifs. *Plant Physiol.* 150, 1272–1285.

Ito, J., Batth, T. S., Petzold, C. J., Redding-Johanson, A. M., Mukhopadhyay, A., Verboom, R., et al. (2011). Analysis of the Arabidopsis cytosolic proteome highlights subcellular partitioning of central plant metabolism. *J. Proteome Res.* 10, 1571–1582.

Ito, J., Heazlewood, J. L., and Millar, A. H. (2006). Analysis of the soluble atp-binding proteome of plant mitochondria identifies new proteins and nucleotide triphosphate interactions within the matrix. *J. Proteome Res.* 5, 3459–3469.

Ito, J., Heazlewood, J. L., and Millar, A. H. (2007). The plant mitochondrial proteome and the challenge of defining the posttranslational modifications responsible for signalling and stress effects on respiratory functions. *Physiol. Plant.* 129, 207–224.

Ito, J., Taylor, N. L., Castleden, I., Weckwerth, W., Millar, A. H., and Heazlewood, J. L. (2009). A survey of the *Arabidopsis thaliana* mitochondrial phosphoproteome. *Proteomics* 9, 4229–4240.

Klodmann, J., and Braun, H.-P. (2011). Proteomic approach to characterize mitochondrial complex I from plants. *Phytochemistry* 72, 1071–1080.

Klodmann, J., Senkler, M., Rode, C., and Braun, H.-P. (2011). Defining the protein complex proteome of plant mitochondria. *Plant Physiol.* 157, 587–598.

Klodmann, J., Sunderhaus, S., Nimtz, M., Jänsch, L., and Braun, H.-P. (2010). Internal architecture of mitochondrial complex I from *Arabidopsis thaliana*. *Plant Cell* 22, 797–810.

Kruft, V., Eubel, H., Jaensch, L., Werhahn, W., and Braun, H.-P. (2001). Proteomic approach to identify novel mitochondrial proteins in Arabidopsis. *Plant Physiol.* 127, 1694–1710.

Lee, C. P., Eubel, H., and Millar, A. H. (2010). Diurnal changes in mitochondrial function reveal daily optimization of light and dark respiratory metabolism in Arabidopsis. *Mol. Cell. Proteomics* 9, 2125–2139.

Lee, C. P., Eubel, H., O'Toole, N., and Millar, A. H. (2008). Heterogeneity of the mitochondrial proteome for photosynthetic and non-photosynthetic Arabidopsis metabolism. *Mol. Cell. Proteomics* 7, 1297–1316.

Lee, C. P., Eubel, H., O'Toole, N., and Millar, A. H. (2011). Combining proteomics of root and shoot mitochondria and transcript analysis to define constitutive and variable components in plant mitochondria. *Phytochemistry* 72, 1092–1108.

Lee, C. P., Eubel, H., Solheim, C., and Millar, A. H. (2012). Mitochondrial proteome heterogeneity between tissues from the vegetative and reproductive stages of *Arabidopsis thaliana* development. *J. Proteome Res.* 11, 3326–3343.

Li, L., Nelson, C. J., Solheim, C., Whelan, J., and Millar, A. H. (2012). Determining degradation and synthesis rates of Arabidopsis proteins using the kinetics of progressive 15N labeling of two-dimensional gel-separated protein spots. *Mol. Cell Proteomics* 11, M111.010025.

Livaja, M., Palmieri, M. C., von Rad, U., and Durner, J. (2008). The effect of the bacterial effector protein harpin on transcriptional profile and mitochondrial proteins of *Arabidopsis thaliana*. *J. Proteomics* 71, 148–159.

Millar, A. H., and Heazlewood, J. L. (2003). Genomic and proteomic analysis of mitochondrial carrier proteins in Arabidopsis. *Plant Physiol.* 131, 443–453.

Millar, A. H., Heazlewood, J. L., Kristensen, B. K., Braun, H.-P., and Moller, I. M. (2005). The plant mitochondrial proteome. *Trends Plant Sci.* 10, 36–43.

Millar, A. H., Sweetlove, L. J., Giege, P., and Leaver, C. J. (2001). Analysis of the Arabidopsis mitochondrial proteome. *Plant Physiol.* 127, 1711–1727.

Millar, A. H., Whelan, J., Soole, K. L., and Day, D. A. (2011). Organization and regulation of mitochondrial respiration in plants. *Annu. Rev. Plant Biol.* 62, 79–104.

Nunes-Nesi, A., Carrari, F., Lytovchenko, A., Smith, A. M., Loureiro, M. E., Ratcliffe, R. G., et al. (2005). Enhanced photosynthetic performance and growth as a consequence of decreasing mitochondrial malate dehydrogenase activity in transgenic tomato plants. *Plant Physiol.* 137, 611–622.

Padmasree, K., Padmavathi, L., and Raghavendra, A. S. (2002). Essentiality of mitochondrial oxidative metabolism for photosynthesis: optimization of carbon assimilation and protection against photoinhibition. *Crit. Rev. Biochem. Mol. Biol.* 37, 71–119.

Pagliarini, D. J., Calvo, S. E., Chang, B., Sheth, S. A., Vafai, S. B., Ong, S. E., et al. (2008). A mitochondrial protein compendium elucidates complex I disease biology. *Cell* 134, 112–123.

Palmieri, M. C., Lindermayr, C., Bauwe, H., Steinhauser, C., and Durner, J. (2010). Regulation of plant glycine decarboxylase by S-nitrosylation and glutathionylation. *Plant Physiol.* 152, 1514–1528.

Peters, K., Nießen, M., Peterhänsel, C., Späth, B., Hölzle, A., Binder, S., et al. (2012). Complex I-complex II ratio strongly differs in various organs of *Arabidopsis thaliana*. *Plant Mol. Biol.* 79, 273–284.

Reinders, J., Zahedi, R. P., Pfanner, N., Meisinger, C., and Sickmann, A. (2006). Toward the complete yeast mitochondrial proteome: multidimensional separation techniques for mitochondrial proteomics. *J. Proteome Res.* 5, 1543–1554.

Rigas, S., Daras, G., Laxa, M., Marathias, N., Fasseas, C., Sweetlove, L. J., et al. (2009). Role of Lon1 protease in post-germinative growth and maintenance of mitochondrial function in *Arabidopsis thaliana*. *New Phytol.* 181, 588–600.

Schertl, P., Sunderhaus, S., Klodmann, J., Grozeff, G. E. G., Bartoli, C. G., and Braun, H.-P. (2012). l-Galactono-1,4-lactone dehydrogenase (GLDH) forms part of three subcomplexes of mitochondrial complex I in *Arabidopsis thaliana*. *J. Biol. Chem.* 287, 14412–14419.

Schmidt, O., Pfanner, N., and Meisinger, C. (2010). Mitochondrial protein import: from proteomics to functional mechanisms. *Nat. Rev. Mol. Cell Biol.* 11, 655–667.

Schwacke, R., Schneider, A., van der Graaff, E., Fischer, K., Catoni, E., Desimone, M., et al. (2003). ARAMEMNON, a novel database for Arabidopsis integral membrane proteins. *Plant Physiol.* 131, 16–26.

Schwanhausser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., et al. (2011). Global quantification of mammalian gene expression control. *Nature* 473, 337–342.

Solheim, C., Li, L., Hatzopoulos, P., and Millar, A. H. (2012). Loss of Lon1 in Arabidopsis changes the mitochondrial proteome leading to altered metabolite profiles and growth retardation without an accumulation of oxidative damage. *Plant Physiol.* 160, 1187–1203.

Sweetlove, L. J., Heazlewood, J. L., Herald, V., Holtzapffel, R., Day, D. A., Leaver, C. J., et al. (2002). The impact of oxidative stress on Arabidopsis mitochondria. *Plant J.* 32, 891–904.

Tan, Y.-F., Millar, A. H., and Taylor, N. L. (2012). Components of mitochondrial oxidative phosphorylation vary in abundance following exposure to cold and chemical stresses. *J. Proteome Res.* 11, 3860–3879.

Tan, Y.-F., O'Toole, N., Taylor, N. L., and Millar, A. H. (2010). Divalent metal ions in plant mitochondria and their role in interactions with proteins and oxidative stress-induced damage to respiratory function. *Plant Physiol.* 152, 747–761.

Taylor, N. L., Heazlewood, J. L., and Millar, A. H. (2011). The *Arabidopsis thaliana* 2-D gel mitochondrial proteome: Refining the value of reference maps for assessing protein abundance, contaminants and post-translational modifications. *Proteomics* 11, 1720–1733.

The Arabidopsis Genome Initiative. (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408, 796–815.

Tomaz, T., Bagard, M., Pracharoenwattana, I., Lindén, P., Lee, C. P., Carroll, A. J., et al. (2010). Mitochondrial malate dehydrogenase lowers leaf respiration and alters photorespiration and plant growth in Arabidopsis. *Plant Physiol.* 154, 1143–1157.

Unseld, M., Marienfeld, J. R., Brandt, P., and Brennicke, A. (1997). The mitochondrial genome of *Arabidopsis thaliana* contains 57 genes in 366,924 nucleotides. *Nat. Genet.* 15, 57–61.

Werhahn, W., and Braun, H.-P. (2002). Biochemical dissection of the mitochondrial proteome from *Arabidopsis thaliana* by three-dimensional gel electrophoresis. *Electrophoresis* 23, 640–646.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 18 November 2012; accepted: 03 January 2013; published online: 25 January 2013.

Citation: Lee CP, Taylor NL and Millar AH (2013) Recent advances in the composition and heterogeneity of the Arabidopsis mitochondrial proteome. Front. Plant Sci. ***4****:4. doi: 10.3389/fpls.2013.00004*

This article was submitted to Frontiers in Plant Proteomics, a specialty of Frontiers in Plant Science.

Copyright © 2013 Lee, Taylor and Millar. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.

frontiers in
PLANT SCIENCE

ORIGINAL RESEARCH ARTICLE
published: 04 June 2013
doi: 10.3389/fpls.2013.00153

3D gel map of *Arabidopsis* complex I

Katrin Peters, Katharina Belt and Hans-Peter Braun*

Institute for Plant Genetics, Faculty of Natural Sciences, Leibniz Universität Hannover, Hannover, Germany

Edited by:
Harvey Millar, The University of Western Australia, Australia

Reviewed by:
Etienne H. Meyer, Max Planck Society, Germany
Pierre Cardol, Université de Liège, Belgium

****Correspondence:***
Hans-Peter Braun, Institute for Plant Genetics, Faculty of Natural Sciences, Leibniz Universität Hannover, Herrenhäuser Straße 2, 30419 Hannover, Germany
e-mail: braun@genetik.uni-hannover.de

Complex I has a unique structure in plants and includes extra subunits. Here, we present a novel study to define its protein constituents. Mitochondria were isolated from *Arabidopsis thaliana* cell cultures, leaves, and roots. Subunits of complex I were resolved by 3D blue-native (BN)/SDS/SDS-PAGE and identified by mass spectrometry. Overall, 55 distinct proteins were found, seven of which occur in pairs of isoforms. We present evidence that *Arabidopsis* complex I consists of 49 distinct types of subunits, 40 of which represent homologs of bovine complex I. The nine other subunits represent special proteins absent in the animal linage of eukaryotes, most prominently a group of subunits related to bacterial gamma-type carbonic anhydrases. A GelMap http://www.gelmap.de/arabidopsis-3d-complex-i/ is presented for promoting future complex I research in *Arabidopsis thaliana*.

Keywords: mitochondria, OXPHOS system, respiratory chain, NADH dehydrogenase, blue-native, BN/SDS/SDS-PAGE, *Arabidopsis thaliana*

INTRODUCTION

The NADH dehydrogenase complex (complex I) of the Oxidative Phosphorylation (OXPHOS) system is present in the cytoplasmic membrane of aerobic bacteria and the inner mitochondrial membrane of eukaryotes. It is composed of two elongated arms: the "membrane arm," and the so-called "peripheral arm" which protrudes into the cytoplasm of the bacterial cell or the matrix of mitochondria (reviewed in Friedrich and Böttcher, 2004; Brandt, 2006; Vogel et al., 2007; Remacle et al., 2008; Zickermann et al., 2008, 2009; Lazarou et al., 2009). The two arms form an L-like structure as originally revealed by electron microscopy (Hofhaus et al., 1991). Very recently, the structure of the entire bacterial enzyme complex has been resolved by X-ray crystallography (Baradaran et al., 2013). Complex I represents a NADH:ubiquinone oxidoreductase. Electron transfer entirely takes place within the peripheral arm and involves an electron transfer chain composed of seven FeS clusters (Hinchliffe and Sazanov, 2005). Quinone reduction takes place at the interface between the two arms and was proposed to induce an electrostatical chain reaction throughout the membrane arm which drives proton translocation across the bacterial or mitochondrial membrane (Baradaran et al., 2013).

Complex I is by far the largest complex of the OXPHOS system. In its simplest form, the bacterial complex consists of 14 subunits (seven subunits per arm) and has a molecular mass of about 500 kDa. However, in eukaryotes, complex I is much larger and consists of more than 40 subunits. Bovine complex I, which extensively was investigated with respect to its subunit composition, consists of 44 subunits, 16 of which are localized in the peripheral and 28 in the membrane arm (Carroll et al., 2006; Balsa et al., 2012). Complex I composition is remarkably conserved in different eukaryotic lineages (Cardol, 2011). However, some lineage-specific complex I subunits occur (Cardol, 2011).

Additional subunits were especially described for plants. Using electron microscopy, complex I of plants was shown to have a very unique shape (Dudkina et al., 2005; Sunderhaus et al., 2006; Peters et al., 2008; Bultema et al., 2009). It has an extra spherical domain which is attached to the membrane arm at a central position and, like the peripheral arm, protrudes into the mitochondrial matrix. It was shown to include extra subunits which resemble gamma-type carbonic anhydrases (Perales et al., 2005; Sunderhaus et al., 2006). In *Arabidopsis*, three carbonic anhydrase subunits form part of complex I (termed CA1, CA2, and CA3) and additionally two more derived "carbonic anhydrase-like" proteins (CAL1 and CAL2). Proteomic studies were initiated to systematically characterize complex I subunits in plants (Heazlewood et al., 2003; Cardol et al., 2004; Sunderhaus et al., 2006; Meyer et al., 2008; Klodmann et al., 2010, 2011; Klodmann and Braun, 2011; Li et al., 2013). These projects led to the identification of several proteins homologous to subunits of bovine complex I and some additional subunits specifically occurring in plants. However, resulting protein sets slightly differ between the presented studies (reviewed in Meyer, 2012).

Here, we present a new study to thoroughly characterize complex I subunits in the model plant *Arabidopsis thaliana*. Our study is based on a 3D gel-electrophoretic approach introduced by Meyer et al. (2008). Using mass spectrometry (MS), 55 complex I proteins were identified, seven of which occur in pairs of isoforms. We present evidence that complex I of *Arabidopsis* includes at least 49 types of proteins, 40 of which represent homologs of bovine complex I and 9 of which are special to plants. A 3D GelMap is presented at http://www.gelmap.de/Arabidopsis-3D-complex-I to facilitate future complex I research in *Arabidopsis*.

MATERIALS AND METHODS

PLANT MATERIAL

A cell culture of *Arabidopsis thaliana* (Col-0) was established as described by May and Leaver (1993). Callus was maintained

as suspension culture according to Sunderhaus et al. (2006). Leaves were harvested from 3 weeks old *Arabidopsis thaliana* (Col-0) plants grown in soil at long day conditions (16 h light, 8 h dark) at 22 °C during the day and 20 °C at night. *Arabidopsis* roots were cultured in liquid medium as described by Lee et al. (2011). For this approach, 50–100 seeds of *Arabidopsis thaliana* Col-0 were surface-sterilized in 70% ethanol for 5 min followed by 5 min incubation in 5% bleach/0.1% Tween 20. Seeds were then washed five times in sterilized water. Length of the single washing steps was increased from 10 s to finally 5 min. All incubation steps took place in a rotary shaker. After the final washing step an appropriate volume of 0.15% agarose solution was added to the seeds. The seeds immediately were carefully dispensed on a stainless steel wire mesh platform which is part of the hydroponic culture system adapted from Schlesier et al. (2003). Conditions for hydroponic culture were according to the protocol of Schlesier et al. (2003). *Arabidopsis* plants were grown under 16/8 h light/dark period with light intensity 100–125 μmol m^{-2} s^{-1} at 22 °C. Liquid medium was replaced with freshly made liquid medium after 2 weeks. After 4 weeks the roots were harvested, pre-washed in root culture medium [0.38% (w/v) Gamborg's B5 salt with vitamins, 3% sucrose, pH 5.8] and transferred into Erlenmeyer flasks containing 50 ml root culture medium. The root culture was kept at 22 °C in the dark under constant agitation at 100 rpm (Lee et al., 2011). It was maintained by transferring small amounts of roots into a new culture flask containing freshly prepared sterilized root culture medium every 3 weeks.

ISOLATION OF MITOCHONDRIA

Mitochondria from cell culture were isolated as described by Werhahn et al. (2001). Isolation of mitochondria from green leaves and roots was performed according to the protocol of Keech et al. (2005).

3D BN/SDS/SDS-PAGE

One-dimensional blue-native PAGE (1D BN-PAGE) was performed according to Wittig et al. (2006). Mitochondrial membranes were solubilized by digitonin at a concentration of 5 g/g mitochondrial protein (Eubel et al., 2003). The two further gel dimensions represented a 2D SDS/SDS-PAGE as originally suggested by Rais et al. (2004). Combining 1D BN-PAGE and 2D SDS/SDS-PAGE was carried out according to Meyer et al. (2008). For this approach, bands corresponding to complex I were excised from the blue-native (BN) gel. Three bands of complex I were used to build a stack on top of a SDS gel (10% polyacrylamide). Electrophoresis was carried out in the presence of 6 M urea. After end of the electrophoretic run, the lane was cut out from the second gel dimension and incubated in acidic solution (Meyer et al., 2008). The gel strip then was horizontally transferred on top of a third dimension SDS gel (16% polyacrylamide) and gel electrophoresis was carried out in the absence of urea.

GEL STAINING PROCEDURES

Polyacrylamide gels were stained with Coomassie Brilliant Blue G250 according to the protocol of Neuhoff et al. (1988, 1990).

PROTEIN IDENTIFICATION BY MASS SPECTROMETRY

Tryptic digestion of proteins and identification of proteins by MS were performed as described by Klodmann et al. (2010). Procedures were based on peptide separation using the EASY-nLC System (Proxeon; Thermo Scientific, Bremen, Germany) and coupled MS analyses using the MicrOTOF-Q II mass spectrometer (Bruker Bremen, Germany). MS data evaluation was carried out using ProteinScape2.1 software (Bruker, Bremen, Germany), the Mascot search engine (Matrix Science, London, UK), and (1) the *Arabidopsis* protein database[1] as well as (2) an updated version of the complex I database used by Klodmann et al. (2010). The latter database is also based on the TAIR protein database (release 10) and includes additionally proteins known to co-migrate with complex I on Blue-native gels (like prohibitins). The following Mascot search parameters were used: enzyme, trypsin/P (up to one missed cleavage allowed); global modification, carbamidomethylation (C), variable modifications, acetyl (N), oxidation (M); precursor ion mass tolerance, 15 ppm; fragment ion mass tolerance, 0.05 Da; peptide charge, 1+, 2+, and 3+; instrument type, electrospray ionization quadrupole time of flight. Minimum ion score was 15, minimum peptide length was four amino acids, significance threshold was set to 0.05 and protein and peptide assessments were carried out if the Mascot Score was greater than 30 for proteins and 20 for peptides.

IMAGE PROCESSING AND DATABASE GENERATION USING GELMAP

Coomassie-blue stained 3D BN/SDS/SDS gels of complex I were scanned using the Image Scanner III (GE Healthcare). Spot coordinates were generated using Microsoft Office Paint. The gel image and a file containing all relevant MS data including the spot coordinates were exported into the GelMap software package available at www.gelmap.de following the instructions given on the website and in Senkler and Braun (2012).

RESULTS AND DISCUSSION

SEPARATION OF COMPLEX I SUBUNITS BY 3D GEL ELECTROPHORESIS

To further investigate the subunit composition of *Arabidopsis* complex I, isolated mitochondria from leaves, roots, and cell cultures were analyzed by 3D BN/SDS/SDS-PAGE according to Meyer et al. (2008) (Figure S1 in Supplementary Material). In the first gel dimension intact mitochondrial protein complexes are resolved by BN-PAGE. Bands representing mitochondrial complex I are cut out from the gel and staples of up to three bands are transferred onto the 2D SDS/SDS-PAGE system as published by Rais et al. (2004). The latter electrophoresis system combines the advantages of high resolution SDS-PAGE with differential resolution of hydrophilic versus hydrophobic proteins. The first SDS gel dimension contains 10% polyacrylamide (PAA) plus 6 M urea while the second SDS gel dimension contains no urea and has a PAA concentration of 16%. On the resulting SDS/SDS gels proteins are dispersed around a diagonal line. This variation in electrophoretic mobility is presumably

[1] www.Arabidopsis.org; release TAIR 10

caused by an altered interaction between SDS and proteins in the presence or absence of urea (Rais et al., 2004). Furthermore, highly hydrophobic proteins show a differential electrophoretic mobility in gels with varying PAA concentrations. In low PAA gels, hydrophobic proteins run slightly faster than hydrophilic ones and in high PPA gels the other way round. On the 2D gel system suggested by Rais et al. (2004) hydrophobic proteins run above the diagonal line. Since complex I likewise includes highly hydrophobic and hydrophilic subunits this gel system nicely allows to investigate its composition (Rais et al., 2004; Meyer et al., 2008; Angerer et al., 2011; Dröse et al., 2011). Upon optimization of protocols, 3D BN/SDS/SDS-PAGE of complex I from *Arabidopsis* cell culture, leaves, and roots allowed to visualize 52 protein spots per fraction based on Coomassie-staining (**Figure 1**; Figure S2 in Supplementary Material). Variation in subunit composition between the three *Arabidopsis* tissues was not observed.

ANALYSIS OF COMPLEX I SUBUNITS

All 52 protein spots of complex I from cell culture and selected subunits of complex I from leaves and roots were analyzed by ESI MS/MS (**Figure 2**; **Table 1**; Figure S2 and Tables in Supplementary Material). Overall, 55 distinct proteins were identified. Analyses of two spots in the low-molecular-mass range did not allow identifying any proteins (spots 51 and 52 on **Figure 2**). Due to spot overlappings, some proteins were detected in more than one spot. The main locations of all proteins (here: highest Mascot score) as well as their secondary locations on the gel are given in **Table 1**. Overall, 7 out of the 55 subunits of *Arabidopsis* complex I occur in pairs of isoforms. This reduces the number of distinct types of subunits detected in our complex I fraction to 48. The subunit ND4L was not detected by MS in our or any previous investigation on *Arabidopsis* complex I which is most likely due to its extreme hydrophobicity (gravy score + 0.976). Systematic analysis of the subunit composition of complex I in the model organism *Yarrowia lipolytica* also did not led to the identification of this subunit (Abdrakhmanova et al., 2004). ND4L belongs to the "core" set of subunits present in all complex I particles. Its gene is localized on the mitochondrial genome in *Arabidopsis*, transcribed and edited (Giegé and Brennicke, 1999). We speculate that ND4L is represented by spots 51 or 52 in the 7 kDa range of our 3D gel, both of which could not be identified (**Figure 2**; Figure S3 in Supplementary Material). ND4L has a calculated mass of 10.9 kDa but is very hydrophobic and therefore should run at ∼7 kDa upon SDS-PAGE. We conclude that *Arabidopsis* complex I consists of at least 49 subunits, 48 of which were detected by our analyses, seven of which occur in pairs of isoforms.

For a limited number of subunits, MS analysis also was carried out for the *Arabidopsis* leaves and roots fractions (**Table 1**; Table S1 in Supplementary Material). Identifications confirm the results obtained for the *Arabidopsis* cell culture. However, in some cases the main locations of corresponding subunits slightly vary between the fractions. It cannot be excluded that these differences are caused by minor gel to gel variations which in some cases made it difficult to precisely assign spots between different fractions. Possible variations in complex I subunit composition between different *Arabidopsis* fractions should be further addressed by future studies.

Based on previous topological investigations for *Arabidopsis* and other model organisms (Carroll, 2003; Hunte et al., 2010; Klodmann et al., 2010; Angerer et al., 2011; Cardol, 2011; Dröse et al., 2011), all 49 subunits can be assigned to the membrane or the peripheral arm of complex I. The peripheral arm consists

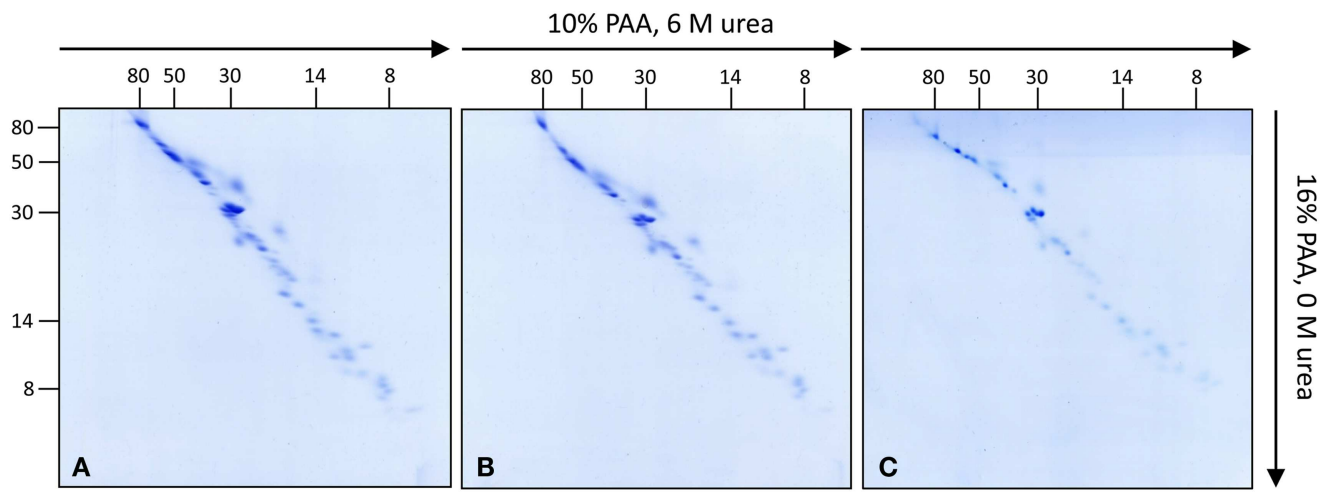


FIGURE 1 | Investigation of complex I subunits from different tissues of *Arabidopsis thaliana* by 3D BN/SDS/SDS-PAGE. Total mitochondrial protein from cell culture, leaves, and roots (1200 μg each) was resolved by BN-PAGE in a first dimension. Complex I was cut out from the BN gel and used for second gel dimensions [SDS-PAGE within a 10% polyacrylamide (PAA) gel in the presence of 6 M urea]. Lanes from the second dimension gels were again cut out and transferred horizontally onto third gel dimensions (SDS-PAGE within a 16% PAA gel in the absence of urea). Gels were stained with Coomassie colloidal. **(A)** Complex I of cell cultures, **(B)** of leaves, **(C)** of roots. Molecular masses (in kilodaltons) are given to the left and on the top of the gels.

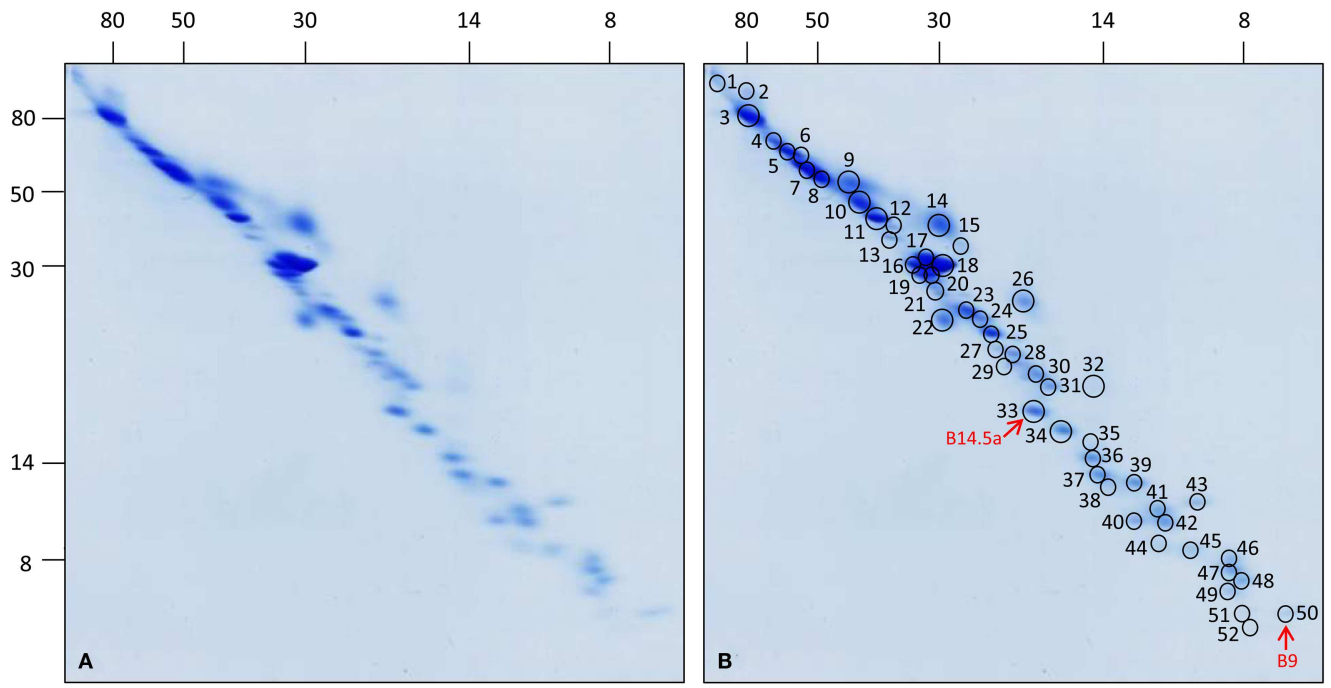


FIGURE 2 | 3D map of complex I from *Arabidopsis thaliana* cell culture. Total mitochondrial protein (1200 μg each) was resolved by 3D BN/SDS/SDS-PAGE. **(A)** Coomassie-stained gel, **(B)** same gel as in **(A)** indicating protein spots which have been analyzed by mass spectrometry. Numbers correspond to those given in **Table 1**. Red arrows indicate the newly identified subunits B14.5a and B9. Molecular masses (in kilodaltons) are given to the left and on the top of the gels.

of 15 subunits, the membrane arm of 34 subunits (**Table 1**). Five subunits of the membrane arm form part of the so-called carbonic anhydrase (CA/CAL) domain, which is absent in mitochondria of opisthokonts (animals and fungi; Gawryluk and Gray, 2010; Cardol, 2011). Of the 49 subunits, 40 represent homologs of subunits present in bovine complex I (**Table 1**). Two of these proteins (subunits B14.5a and B9) were identified for the first time in *Arabidopsis* but previously predicted to form part of complex I by genome analyses (Cardol, 2011). The high number of homologs in bovine and *Arabidopsis* complex I underlines the remarkable conservation of this protein complex in Eukaryotes (Cardol, 2011). Bovine complex I consists of 44 subunits (Carroll et al., 2006; Balsa et al., 2012), only four of which were not found in *Arabidopsis* (10 kDa, 42 kDa, SDAP, and B17 subunits; Meyer, 2012). On the contrary, *Arabidopsis* complex I includes nine subunits absent in the bovine complex (for summary, see Figure S4 in Supplementary Material).

Of the nine extra subunits in plants, five represent members of the CA/CAL family. Since deletion of single CA or CAL genes does not cause complete loss of intact complex I (Perales et al., 2005; Sunderhaus et al., 2006; Meyer et al., 2011; Wang et al., 2012a) it cannot be excluded that they present isoforms which alternatively are present in complex I particles. However, deletion of the ca2 gene leads to highly reduced levels of complex I (Perales et al., 2005) indicating that CA2 cannot easily be replaced by CA1 or CA3. Sequence identity between CA1, CA2, and CA3 is in the range of 75%. In contrast, sequences of the CAL1 and CAL2 subunits of *Arabidopsis* are very similar (90% sequence identity), possibly indicating that these proteins represent isoforms. Indeed, deletion of the cal1 or cal2 gene in *Arabidopsis* does not visibly affect *Arabidopsis* development but the double mutant is not viable (Wang et al., 2012a). Considering the size of the CA/CAL domain upon single particle EM of *Arabidopsis* complex I it was concluded that it consists of at least three copies of CA/CAL proteins (Sunderhaus et al., 2006). Further experiments have to be carried out in order to clarify the number of CA/CAL subunits per individual complex I particles.

The plant-specific GLDH subunit binds to three complex I assembly intermediates of 420, 480, and 850 kDa (Schertl et al., 2012) but so far was not detected in preparations of intact complex I. Our data point to the possibility that GLDH also binds to the intact complex. However, it cannot be excluded that the 1000 kDa complex I band excised from the BN gel also included small amounts of the band representing the 850 kDa subcomplex. Three further plant-specific subunits were detected on our 3D gels: P1, P2, and a protein encoded by At1g18320. The P1 and P2 proteins were consistently detected in complex I fractions from plants (Meyer, 2012). Both form part of the membrane arm (Sunderhaus et al., 2006). The At1g18320 protein was previously found to co-migrate with complex I on a BN/SDS gel (Klodmann et al., 2011). However, its status representing an integral complex I subunit in *Arabidopsis* should be further investigated.

Table 1 | Complex I subunits in *Arabidopsis thaliana*.

Subunit[1]		Accession[2]	Spot[3]		Organ[4]
Plant subunit	**Bovine homolog**		**Main spot**	**Further spots**	
Membrane arm					
15 kDa-1	15 kDa	At3g62790	41		c
15 kDa-2	15 kDa	At2g47690	41		c
AGGG	AGGG	At1g76200	47		c
ASHI	ASHI	At5g47570	44		c
B9	B9	At2g46540	50		c
B12-1	B12	At1g14450	45		c, l
B12-2	B12	At2g02510	45		c, l
B14	B14	At3g12260	36		c
B14.5b	B14.5b	At4g20150	47		c, l
B14.7	B14.7	At2g42210	33		c, l
B15	B15	At2g31490	46		c, l
B16.6-1	B16.6	At1g04630	34		c, l
B16.6-2	B16.6	At2g33220	34		c
B18	B18	At2g02050	38	37	c, l, r
B22	B22	At4g34700	35		c
ESSS-1	ESSS	At2g42310	42		c
ESSS-2	ESSS	At3g57785	42		c
KFYI	KFYI	At4g00585	41	32	c
MNLL	MNLL	At4g16450	40		c, l, r
MWFE	MWFE	At3g08610	49		c, l
ND1	ND1	AtMg00516/AtMg01120/AtMg01275[5]	26		c, l
ND2	ND2	AtMg00285/AtMg01320[5]	14		c, l, r
ND3	ND3	AtMg00990	43		c
ND4	ND4	AtMg00580	14		c
ND4L	ND4L	AtMg00650	–		–
ND5	ND5	AtMg00060/AtMg00513/AtMg00665[5]	9	2	c, l
ND6	ND6	AtMg00270	15		c
PDSW-2	PDSW	At1g49140	36		c, l, r
PDSW-1	PDSW	At3g18410	36		c, l, r
PGIV-1	PGIV	At3g06310	37		c
PGIV-2	PGIV	At5g18800	38	37	c
GLDH	–	At3g47930	4		c
P1	–	At1g67350	39		c, l
P2	–	At2g27730	37		c, l, r
At1g18320	–	At1g18320	29		c
Carbonic anhydrase domain (membrane arm)					
CA1	–	At1g19580	16	17, 18, 19, 20	c, l
CA2	–	At1g47260	16	13, 17, 18, 19, 20	c, l
CA3	–	At5g66510	21	19, 20	c, l
CAL1	–	At5g63510	23		c, l
CAL2	–	At3g48680	23	22, 24	c, l, r
Peripheral arm					
13 kDa	13 kDa	At3g03070	44		c
18 kDa	18 kDa	At5g67590	33		c, l
24 kDa	24 kDa	At4g02580	21		c, l, r
39 kDa	39 kDa	At2g20360	11	12	c, l, r
51 kDa	51 kDa	At5g08530	8	5, 6, 7, 9	c, l
75 kDa	75 kDa	At5g37510	3	1, 2, 4	c, l
B8	B8	At5g47890	42	41	c, l
B13	B13	At5g52840	28	27, 29	c, l, r
B14.5a	B14.5a	At5g08060	33		c
B17.2	B17.2	At3g03100	31		c
ND7	49 kDa	AtMg00510	10		c, l
ND9	30 kDa	AtMg00070	25		c, l, r
PSST	PSST	At5g11770	30		c, l, r
SGDH	SGDH	At1g67785	48		c, l
TYKY-1	TYKY	At1g79010	22		c, l
TYKY-2	TYKY	At1g16700	22		c, l

(Continued)

Table 1 | Continued

[1] Subunits of complex I from Arabidopsis were named according to their homologs in bovine complex I (40 homologous subunits). Exceptions: Arabidopsis homologs to the 30 and 49 kDa subunits of bovine complex I are designated ND7 and ND9 because the corresponding proteins are encoded by the mitochondrial genome in plants. Seven subunits occur in pairs of isoforms in Arabidopsis. The names of these proteins were extended by "−1" and "−2." Arabidopsis complex I includes nine additional subunits absent in bovine complex I. These proteins are named in accordance to the literature: CA1, CA2, CA3, CAL1, CAL2, GLDH (L-galactone 1-4 lactone dehydrogenase), P1, P2, and At1g18320.

[2] Accession numbers as given by TAIR http://www.arabidopsis.org.

*[3] Spot number in accordance with **Figure 2**.*

[4] Organ/culture in which the subunit was identified; c, cell culture; l, leaf; r, root.

[5] Two to three accession numbers are given for the ND1, ND2, and ND5 proteins because they are encoded by a corresponding number of gene fragments on the mitochondrial genome in Arabidopsis. Transcripts encoding the complete proteins are generated by trans-splicing (Knoop et al., 1991; Knoop and Brennicke, 1993; Lippok et al., 1996).

Table 2 | Candidates of additional complex I subunits in *Arabidopsis thaliana*.

Accession	Evidence	Remark
At5g14105	Klodmann et al. (2010), Klodmann and Braun (2011)	Subunit P3[1]
At1g68680	Meyer et al. (2008)	
At3g10110	Klodmann et al. (2011)	Similar to TIM22
At1g72170	Klodmann et al. (2011)	
At2g28430	Klodmann et al. (2011)	
At1g72750	Wang et al. (2012b)	Similar to TIM23

[1] At5g14105 was suggested to be named P3 in Meyer (2012).

FURTHER COMPLEX I SUBUNITS IN PLANTS?

In previous investigations based on BN-PAGE six additional complex I proteins were identified in *Arabidopsis* (summarized in Meyer, 2012): At5g14105 (Klodmann et al., 2010; Klodmann and Braun, 2011), At1g68680 (Meyer et al., 2008), At1g72170, At3g10110 and At2g28430 (Klodmann et al., 2011), and At1g72750 (Wang et al., 2012b) (**Table 2**). However, detection of these proteins is not consistent. It currently cannot be excluded that these proteins co-migrate with complex I on blue-native gels but form part of separate complexes. Interestingly, some of these proteins are known components of the pre-protein translocase of the inner mitochondrial membrane, the TIM complex (At1g72750 and At3g10110; the latter protein represents an isoform of At1g18320 which was identified in the course of our current study; **Table 1**). It recently has been suggested that complex I and the TIM complex are physically linked in plant mitochondria (Murcha et al., 2012).

3D REFERENCE MAP OF COMPLEX I

To facilitate identifying complex I subunits upon 3D BN/SDS/SDS-PAGE, a GelMap was generated for the MS dataset of the gel presented in **Figure 2**. GelMap is a software tool for the building and presentation of proteome reference maps (www.gelmap.de; Senkler and Braun, 2012). In contrast to alternative software packages, it allows assignment of multiple proteins per protein spot and at the same time functional annotation of all proteins. By clicking onto protein spots, widespread information is offered. Several GelMaps on *Arabidopsis* mitochondria are presented at the GelMap homepage, including a map on SDS-induced complex I subcomplexes[2].

For the 3D GelMap of *Arabidopsis* complex I, the 55 identified proteins are grouped into functional categories according to their localization within the peripheral arm, the membrane arm, or the carbonic anhydrase domain attached to the membrane arm (**Figure 3**; http://www.gelmap.de/arabidopsis-3d-complex-i/). Furthermore, the six candidates for additional complex I subunits are given in another category. The proteins of the latter category are linked to an "extra" spot below the gel. By clicking onto any protein spot on the map, all included proteins are displayed. Proteins are sorted according to their MASCOT scores. Upon clicking onto an individual protein, a tooltip opens which includes additional information. Extensive further information on each protein is offered by links to several external databases. The new GelMap is intended to be a helpful tool for future complex I research in *Arabidopsis*.

ACKNOWLEDGMENTS

Katrin Peters was supported by the "Wege in die Forschung II" program offered by Leibniz University Hannover. We acknowledge support by Deutsche Forschungsgemeinschaft (DFG) and Open Access Publishing Fund of Leibniz Universität Hannover.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://www.frontiersin.org/Plant_Proteomics/10.3389/fpls.2013.00153/abstract

Figure S1 | Principle of 3D BN/SDS/SDS-PAGE.

Figure S2 | Replicates of 3D BN/SDS/SDS gels for complex I from cell cultures of *Arabidopsis*.

Figure S3 | Regions on 3D BN/SDS/SDS gels showing the smallest complex I subunits. Gels were Coomassie stained (left, middle) or silver stained (right). The three smallest proteins (corresponding to spots 50, 51, and 52 on **Figure 2**) only become clearly visible upon silver staining. Spot 50 represents the B9 subunit. Spot 51 might represent subunit ND4L. Spot 52 could not be identified.

Figure S4 | Species specific complex I subunits in *B. taurus* and *A. thaliana*.

[2] http://www.gelmap.de/177

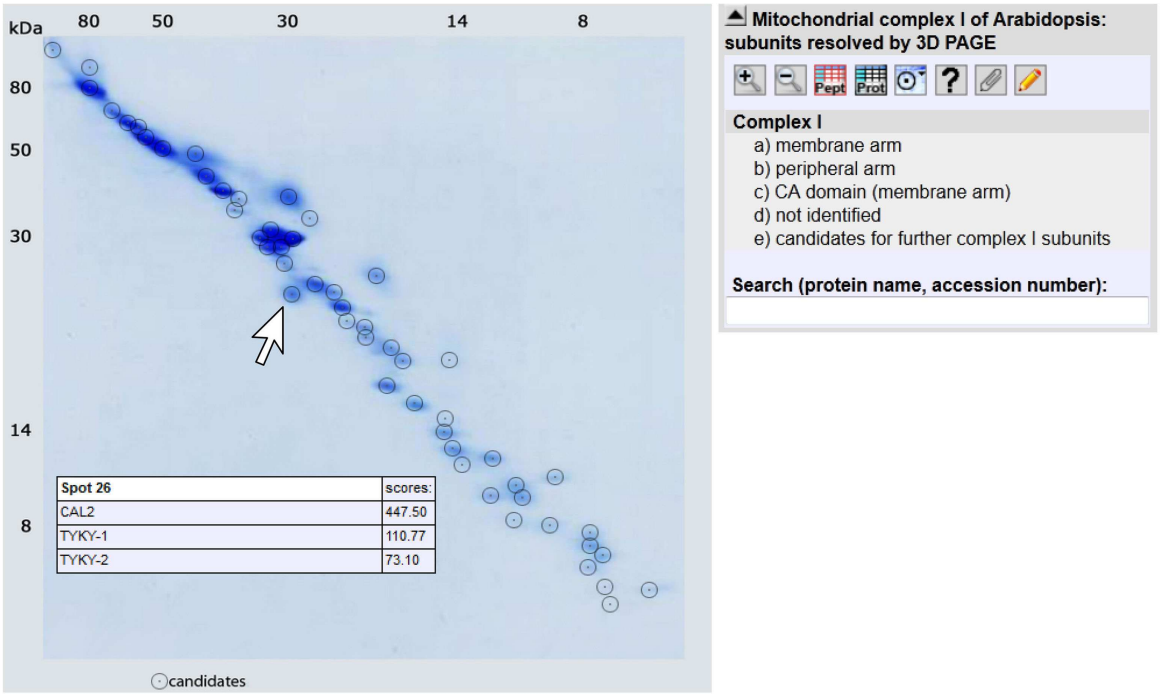


FIGURE 3 | GelMap of complex I as resolved by 3D BN/SDS/SDS-PAGE (http://www.gelmap.de/arabidopsis-3d-complex-i/). Upon hovering with the cursor over a spot, a tooltip including information on all included proteins is opened. In the example given on the figure, the indicated spot includes the CAL2 protein and two isoforms of the TYKY subunit. Upon clicking into the spot the protein names are converted into stable links which can be used to obtain further information. Protein information also can be obtained by clicking into the menu given to the right or by entering protein names or accessions into the search field below the menu.

Figure S5 | Identity of complex I subunits of *Arabidopsis* upon analysis by 3D BN/SDS/SDS PAGE.

Table S1 | Protein table of the GelMap (http://www.gelmap.de/arabidopsis-3d-complex-i/).

Table S2 | Protein table of complex I subunits in leaves of *Arabidopsis thaliana*.

Table S3 | Protein table of complex I subunits in roots of *Arabidopsis thaliana*.

REFERENCES

Abdrakhmanova, A., Zickermann, V., Bostina, M., Radermacher, M., Schägger, H., Kerscher, S., et al. (2004). Subunit composition of mitochondrial complex I from the yeast *Yarrowia lipolytica*. *Biochim. Biophys. Acta* 1658, 148–156. doi:10.1016/j.bbabio.2004.04.019

Angerer, H., Zwicker, K., Wumaier, Z., Sokolova, L., and Heide, H. (2011). A scaffold of accessory subunits links the peripheral arm and the distal proton-pumping module of mitochondrial complex I. *Biochem. J.* 437, 279–288. doi:10.1042/BJ20110359

Balsa, E., Marco, R., Perales-Clemente, E., Szklarczyk, R., Calvo, E., Landazuri, M. O., et al. (2012). NDUFA4 is a subunit of complex IV of the mammalian electron transport chain. *Cell Metab.* 16, 378–386. doi:10.1016/j.cmet.2012.07.015

Baradaran, R., Berrisford, J. M., Minhas, G. S., and Sazanov, L. A. (2013). Crystal structure of the entire respiratory complex I. *Nature* 494, 443–448. doi:10.1038/nature11871

Brandt, U. (2006). Energy converting NADH:quinone oxidoreductase (complex I). *Annu. Rev. Biochem.* 75, 69–92. doi:10.1146/annurev.biochem.75.103004.142539

Bultema, J., Braun, H.-P., Boekema, E., and Kouril, R. (2009). Megacomplex organization of the oxidative phosphorylation system by structural analysis of respiratory supercomplexes from potato. *Biochim. Biophys. Acta* 1787, 60–67. doi:10.1016/j.bbabio.2008.10.010

Cardol, P. (2011). Mitochondrial NADH:ubiquinone oxidoreductase (complex I) in eukaryotes: a highly conserved subunit composition highlighted by mining of protein databases. *Biochim. Biophys. Acta* 1807, 1390–1397. doi:10.1016/j.bbabio.2011.06.015

Cardol, P., Vanrobaeys, F., Devreese, B., van Beeumen, J., Matagne, R. F., and Remacle, C. (2004). Higher plant-like subunit composition of mitochondrial complex I from *Chlamydomonas reinhardtii*: 31 conserved components among eukaryotes. *Biochim. Biophys. Acta* 1658, 212–224. doi:10.1016/j.bbabio.2004.06.001

Carroll, J. (2003). Analysis of the subunit composition of complex I from bovine heart mitochondria. *Mol. Cell. Proteomics* 2, 117–126.

Carroll, J., Fearnley, I. M., Skehel, J. M., Shannon, R. J., Hirst, J., and Walker, J. E. (2006). Bovine complex I is a complex of 45 different subunits. *J. Biol. Chem.* 281, 32724–32727. doi:10.1074/jbc.M607135200

Dröse, S., Krack, S., Sokolova, L., Zwicker, K., Barth, H. D., Morgner, N., et al. (2011). Functional dissection of the proton pumping modules of mitochondrial complex I. *PLoS Biol.* 9:e1001128. doi:10.1371/journal.pbio.1001128

Dudkina, N. V., Eubel, H., Keegstra, W., Boekema, E. J., and Braun, H.-P. (2005). Structure of a mitochondrial supercomplex formed by respiratory-chain complexes I and III. *Proc. Natl. Acad. Sci. U.S.A.* 102, 3225–3229. doi:10.1073/pnas.0408870102

Eubel, H., Jänsch, L., and Braun, H.-P. (2003). New insights into the respiratory chain of plant mitochondria. Supercomplexes and a unique composition of complex II. *Plant Physiol.* 133, 274–286. doi:10.1104/pp.103.024620

Friedrich, T., and Böttcher, B. (2004). The gross structure of the respiratory complex I: a Lego System. *Biochim. Biophys. Acta* 1608, 1–9. doi:10.1016/j.bbabio.2003.10.002

Gawryluk, R. M., and Gray, M. W. (2010). Evidence for an early evolutionary emergence of gamma-type carbonic anhydrase as components of mitochondrial

respiratory complex I. *BMC Evol. Biol.* 10:176. doi:10.1186/1471-2148-10-176

Giegé, P., and Brennicke, A. (1999). RNA editing in *Arabidopsis* mitochondria effects 441 C to U changes in ORFs. *Proc. Natl. Acad. Sci. U.S.A.* 96, 15324–15329. doi:10.1073/pnas.96.26.15324

Heazlewood, J. L., Howell, K. A., and Millar, A. H. (2003). Mitochondrial complex I from *Arabidopsis* and rice: orthologs of mammalian and fungal components coupled with plant-specific subunits. *Biochim. Biophys. Acta* 1604, 159–169. doi:10.1016/S0005-2728(03)00045-8

Hinchliffe, P., and Sazanov, L. A. (2005). Organization of iron-sulfur clusters in respiratory complex I. *Science* 309, 771–774. doi:10.1126/science.1113988

Hofhaus, G., Weiss, H., and Leonard, K. (1991). Electron microscopic analysis of the peripheral and membrane parts of mitochondrial NADH dehydrogenase (complex I). *J. Mol. Biol.* 221, 1027–1043. doi:10.1016/0022-2836(91)80190-6

Hunte, C., Zickermann, V., and Brandt, U. (2010). Functional modules and structural basis of conformational coupling in mitochondrial complex I. *Science* 329, 448–451. doi:10.1126/science.1191046

Keech, O., Dizengremel, P., and Gardeström, P. (2005). Preparation of leaf mitochondria from *Arabidopsis thaliana*. *Physiol. Plant* 124, 403–409. doi:10.1111/j.1399-3054.2005.00521.x

Klodmann, J., and Braun, H.-P. (2011). Proteomic approach to characterize mitochondrial complex I from plants. *Phytochemistry* 72, 1071–1080. doi:10.1016/j.phytochem.2010.11.012

Klodmann, J., Senkler, M., Rode, C., and Braun, H.-P. (2011). Defining the protein complex proteome of plant mitochondria. *Plant Physiol.* 157, 587–598. doi:10.1104/pp.111.182352

Klodmann, J., Sunderhaus, S., Nimtz, M., Jansch, L., and Braun, H.-P. (2010). Internal architecture of mitochondrial complex I from *Arabidopsis thaliana*. *Plant Cell* 22, 797–810. doi:10.1105/tpc.109.073726

Knoop, V., and Brennicke, A. (1993). "Group II introns in plant mitochondria – trans-splicing, RNA editing, evolution and promiscuity," in *Plant Mitochondria*, eds A. Brennicke and U. Kück (Weinheim: VCH Verlagsgesellschaft), 221–232.

Knoop, V., Schuster, W., Wissinger, B., and Brennicke, A. (1991). Trans splicing integrates an exon of 22 nucleotides into the nad5 mRNA in higher plant mitochondria. *EMBO J.* 10, 3483–3493.

Lazarou, M., Thorburn, D. R., Ryan, M. T., and McKenzie, M. (2009). Assembly of mitochondrial complex I and defects in disease. *Biochim. Biophys. Acta* 1793, 78–88. doi:10.1016/j.bbamcr.2008.04.015

Lee, C. P., Eubel, H., O'Toole, N., and Millar, A. H. (2011). Combining proteomics of root and shoot mitochondria and transcript analysis to define constitutive and variable components in plant mitochondria. *Phytochemistry* 72, 1092–1108. doi:10.1016/j.phytochem.2010.12.004

Li, L., Nelson, C. J., Carrie, C., Gawryluk, R. M., Solheim, C., Gray, M. W., et al. (2013). Subcomplexes of ancestral respiratory complex I subunits rapidly turn over *in vivo* as productive assembly intermediates in *Arabidopsis*. *J. Biol. Chem.* 288, 5707–5717. doi:10.1074/jbc.M112.432070

Lippok, B., Brennicke, A., and Unseld, M. (1996). The rps4-gene is encoded upstream of the nad2-gene in *Arabidopsis* mitochondria. *Biol. Chem. Hoppe-Seyler* 377, 251–257. doi:10.1515/bchm3.1996.377.4.251

May, M. J., and Leaver, C. J. (1993). Oxidative stimulation of glutathione synthesis in *Arabidopsis thaliana* suspension cultures. *Plant Physiol.* 103, 621–627.

Meyer, E. H. (2012). Proteomic investigations of complex I composition: how to define a subunit? *Front. Plant Sci.* 3:106. doi:10.3389/fpls.2012.00106

Meyer, E. H., Solheim, C., Tanz, S. K., Bonnard, G., and Millar, A. H. (2011). Insights into the composition and assembly of the membrane arm of plant complex I through analysis of subcomplexes in *Arabidopsis* mutant lines. *J. Biol. Chem.* 286, 26081–26092. doi:10.1074/jbc.M110.209601

Meyer, E. H., Taylor, N. L., and Millar, A. H. (2008). Resolving and identifying protein components of plant mitochondrial respiratory complexes using three dimensions of gel electrophoresis. *J. Proteome Res.* 7, 786–794. doi:10.1021/pr700595p

Murcha, M. W., Wang, Y., and Whelan, J. (2012). A molecular link between mitochondrial preprotein transporters and respiratory chain complexes. *Plant Signal. Behav.* 7, 1594–1597. doi:10.4161/psb.22250

Neuhoff, V., Stamm, R., and Eibl, H. (1988). Clear background and highly sensitive protein staining with coomassie blue dyes in polyacrylamide gels: a systematic analysis. *Electrophoresis* 11, 427–448.

Neuhoff, V., Stamm, R., Pardowitz, I., Arold, N., Ehrhardt, W., and Taube, D. (1990). Essential problems in quantification of proteins following colloidal staining with Coomassie brilliant blue dyes in polyacrylamide gels, and their solution. *Electrophoresis* 11, 101–117. doi:10.1002/elps.1150110202

Perales, M., Eubel, H., Heinemeyer, J., Colaneri, A., Zabaleta, E., and Braun, H.-P. (2005). Disruption of a nuclear gene encoding a mitochondrial gamma carbonic anhydrase reduces complex I and supercomplex I+III2 levels and alters mitochondrial physiology in *Arabidopsis*. *J. Mol. Biol.* 350, 263–277. doi:10.1016/j.jmb.2005.04.062

Peters, K., Dudkina, N. V., Jänsch, L., Braun, H.-P., and Boekema, E. J. (2008). A structural investigation of complex I and I+III2 supercomplex from *Zea mays* at 11-13 A resolution: assignment of the carbonic anhydrase domain and evidence for structural heterogeneity within complex I. *Biochim. Biophys. Acta* 1777, 84–93. doi:10.1016/j.bbabio.2007.10.012

Rais, I., Karas, M., and Schägger, H. (2004). Two-dimensional electrophoresis for the isolation of integral membrane proteins and mass spectrometric identification. *Proteomics* 4, 2567–2571. doi:10.1002/pmic.200400829

Remacle, C., Barbieri, M. R., Cardol, P., and Hamel, P. P. (2008). Eukaryotic complex I: functional diversity and experimental systems to unravel the assembly process. *Mol. Genet. Genomics* 280, 93–110. doi:10.1007/s00438-008-0350-5

Schertl, P., Sunderhaus, S., Klodmann, J., Grozeff, G. E. G., Bartoli, C. G., and Braun, H.-P. (2012). L-galactono-1,4-lactone dehydrogenase (GLDH) forms part of three subcomplexes of mitochondrial complex I in *Arabidopsis thaliana*. *J. Biol. Chem.* 287, 14412–14419. doi:10.1074/jbc.M111.305144

Schlesier, B., Bréton, F., and Mock, H.-P. (2003). A hydroponic culture system for growing *Arabidopsis thaliana* plantlets under sterile conditions. *Plant Mol. Biol. Rep.* 21, 449–456. doi:10.1007/BF02772594

Senkler, M., and Braun, H.-P. (2012). Functional annotation of 2D protein maps: the GelMap portal. *Front. Plant Sci.* 3:87. doi:10.3389/fpls.2012.00087

Sunderhaus, S., Dudkina, N., Jansch, L., Klodmann, J., Heinemeyer, J., Perales, M., et al. (2006). Carbonic anhydrase subunits form a matrix-exposed domain attached to the membrane arm of mitochondrial complex I in plants. *J. Biol. Chem.* 281, 6482–6488. doi:10.1074/jbc.M511542200

Vogel, R. O., Smeitink, J. A. M., and Nijtmans, L. G. J. (2007). Human mitochondrial complex I assembly: a dynamic and versatile process. *Biochim. Biophys. Acta* 1767, 1215–1227. doi:10.1016/j.bbabio.2007.07.008

Wang, Q., Fristedt, R., Yu, X., Chen, Z., Liu, H., Lee, Y., et al. (2012a). The gamma-carbonic anhydrase subcomplex of mitochondrial complex I is essential for development and important for photomorphogenesis of *Arabidopsis*. *Plant Physiol.* 160, 1373–1383. doi:10.1104/pp.112.204339

Wang, Y., Carrie, C., Giraud, E., Elhafez, D., Narsai, R., Duncan, O., et al. (2012b). Dual location of the mitochondrial preprotein transporters B14.7 and Tim23-2 in complex I and the TIM17:23 complex in *Arabidopsis* links mitochondrial activity and biogenesis. *Plant Cell* 24, 2675–2695. doi:10.1105/tpc.112.098731

Werhahn, W., Niemeyer, A., Jansch, L., Kruft, V., Schmitz, U. K., and Braun, H. (2001). Purification and characterization of the preprotein translocase of the outer mitochondrial membrane from *Arabidopsis*. Identification of multiple forms of TOM20. *Plant Physiol.* 125, 943–954. doi:10.1104/pp.125.2.943

Wittig, I., Braun, H.-P., and Schägger, H. (2006). Blue native PAGE. *Nat. Protoc.* 1, 418–428. doi:10.1038/nprot.2006.62

Zickermann, V., Drose, S., Tocilescu, M. A., Zwicker, K., Kerscher, S., and Brandt, U. (2008). Challenges in elucidating structure and

mechanism of proton pumping NADH:ubiquinone oxidoreductase (complex I). *J. Bioenerg. Biomembr.* 40, 475–483. doi:10.1007/s10863-008-9171-9

Zickermann, V., Kerscher, S., Zwicker, K., Tocilescu, M. A., Radermacher, M., and Brandt, U. (2009). Architecture of complex I and its implications for electron transfer and proton pumping. *Biochim. Biophys. Acta* 1787, 574–583. doi:10.1016/j.bbabio.2009.01.012

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 22 March 2013; accepted: 04 May 2013; published online: 04 June 2013.

Citation: Peters K, Belt K and Braun H-P (2013) 3D gel map of Arabidopsis complex I. Front. Plant Sci. ***4****:153. doi: 10.3389/fpls.2013.00153*

This article was submitted to Frontiers in Plant Proteomics, a specialty of Frontiers in Plant Science.

Copyright © 2013 Peters, Belt and Braun. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.

frontiers in
PLANT SCIENCE

MINI REVIEW ARTICLE
published: 07 February 2013
doi: 10.3389/fpls.2013.00016

The rice mitochondria proteome and its response during development and to the environment

Shaobai Huang*, Rachel N. Shingaki-Wells, Nicolas L. Taylor and A. Harvey Millar

Australian Research Council Centre of Excellence in Plant Energy Biology and Centre for Comparative Analysis of Biomolecular Networks, The University of Western Australia, Crawley, WA, Australia

Edited by:
Scott C. Peck, University of Missouri, USA

Reviewed by:
Brian Mooney, University of Missouri, USA
Ian M. Møller, Aarhus University, Denmark

****Correspondence:***
Shaobai Huang, Australian Research Council Centre of Excellence in Plant Energy Biology and Centre for Comparative Analysis of Biomolecular Networks, The University of Western Australia, Bayliss Building M316, 35 Stirling Highway, Crawley, WA 6009, Australia.
e-mail: shaobai.huang@uwa.edu.au

Rice (*Oryza sativa* L.) is both a major crop species and the key model grass for molecular and physiological research. Mitochondria are important in rice, as in all crops, as the main source of ATP for cell maintenance and growth. However, the practical significance of understanding the function of mitochondria in rice is increased by the widespread farming practice of using hybrids to boost rice production. This relies on cytoplasmic male sterile (CMS) lines with abortive pollen caused by dysfunctional mitochondria. We provide an overview of what is known about the mitochondrial proteome of rice seedlings. To date, more than 320 proteins have been identified in purified rice mitochondria using mass spectrometry. The insights from this work include a broad understanding of the major subunits of mitochondrial respiratory complexes and TCA cycle enzymes, carbon and nitrogen metabolism enzymes as well as details of the supporting machinery for biogenesis and the subset of stress-responsive mitochondrial proteins. Many proteins with unknown functions have also been found in rice mitochondria. Proteomic analysis has also revealed the features of rice mitochondrial protein presequences required for mitochondrial targeting, as well as cleavage site features for processing of precursors after import. Changes in the abundance of rice mitochondrial proteins in response to different stresses, especially anoxia and light, are summarized. Future research on quantitative analysis of the rice mitochondrial proteomes at the spatial and developmental level, its response to environmental stresses and recent advances in understanding of the basis of rice CMS systems are highlighted.

Keywords: rice, mitochondrial proteome, mitochondrial proteins, development, stress response

INTRODUCTION

Rice is the one of the key model plants for research and also the major food crop in developing countries. Dramatic increases in rice production have occurred in the past few decades through large scale hybrid rice cultivation using cytoplasmic male sterile (CMS) lines with abortive pollen caused by dysfunctional mitochondria (Eckardt, 2006; Wang et al., 2006). This tremendous advance highlights the significance of understanding rice mitochondrial function. The role and nature of rice mitochondria also takes on special significance due to its early growth in hypoxic or even anoxic environments (Perata and Voesenek, 2007) as well as its need for rapid mitochondrial biogenesis and function during re-oxygenation (Millar et al., 2004; Howell et al., 2007). Mitochondria contain many hundreds of different proteins that initiate or co-ordinate the biochemical processes essential for its function. It is estimated that while only ~300 proteins are components of the respiratory apparatus, up to 2000 proteins are housed in plant mitochondria with the majority encoded in the nucleus and transported into mitochondria as cytosolic precursor proteins by the mitochondrial protein import machinery (Millar et al., 2005; Cui et al., 2011). Because software-based subcellular targeting prediction offers low fidelity in actual subcellular localization (Heazlewood et al., 2004, 2005), direct experimental analysis of mitochondrial proteomes, including that of rice, is required to obtain precise information on which proteins are mitochondrially located. More broad advances in rice proteomics have been well summarized recently (Agrawal and Rakwal, 2011). In this review, recent research on rice mitochondrial purification, contaminant removal, and rice mitochondrial proteomic analysis are discussed. The rice mitochondrial protein composition, functional classifications, and features of mitochondrial protein presequences are summarized. We also discuss the effects of the environment, in particular anoxia and light, on rice mitochondrial proteome composition and how the proteome differs in CMS lines. Finally, we propose future directions for research on the rice mitochondrial proteome.

PURIFICATION AND PROTEOMIC ANALYSIS OF RICE MITOCHONDRIA

Removal of contaminants, like chloroplasts, from purified rice mitochondria is critical for downstream protein separation and identification of mitochondrial proteins. Classically, differential and gradient centrifugation methods based on size and density have been applied to plant mitochondrial proteomic analysis (Kruft et al., 2001; Millar et al., 2001; Bardel et al., 2002). Using these approaches, mitochondria have been purified on Percoll density gradients from dark-grown rice seedlings (Heazlewood et al., 2003) and from green rice seedlings (Kristensen et al., 2004).

The purified mitochondria from dark-grown seedlings were then separated using 2-D IEF/SDS-PAGE, blue native (BN)-PAGE and 122 proteins were identified using LC-MS/MS (Heazlewood et al., 2003). In another similar study that used a sucrose gradient for mitochondrial purification, a set of 112 non-redundant rice mitochondrial proteins were identified after 2-D IEF/SDS-PAGE separation[1]. (Komatsu, 2005). Comparison of these two studies revealed less than 20% overlap in the two datasets of highly abundant proteins, highlighting the importance of optimized methods for mitochondria purification prior to proteomic analysis.

Free-flow electrophoresis in zone electrophoresis mode (ZE-FFE) can be used to separate organelles based on differential surface charge and this has allowed the comprehensive analysis of *Arabidopsis* organellar proteomes including the exclusion of contaminating proteins through quantitative analysis (Eubel et al., 2007, 2008). The combination of traditional differential and gradient centrifugation with this new FFE separation technique has allowed isolation of highly purified rice mitochondria for proteomic analysis (Huang et al., 2009a). Quantitative analysis using differential in gel electrophoresis (DIGE) and spectral counting have allowed the identification of contaminant proteins removed by FFE purification (Huang et al., 2009a). The purity of isolated mitochondria was >95% based on calculating the number of peptides from contaminant proteins compared to peptides from mitochondrial proteins in these preparations (Huang et al., 2009a). In total, 322 proteins from FFE purified rice mitochondria were identified through the direct analysis of trypsin-digested peptides by LC-MS/MS and gel-based analysis (Huang et al., 2009a). The annotations of rice mitochondrial protein spots on 2-D IEF/SDS/PAGE gel are available online[2] using the gel-map tool (Klodmann et al., 2011; Senkler and Braun, 2012). Seventy-eight proteins identified previously as components of the rice mitochondrial proteome (Heazlewood et al., 2003) were also in this study. Half of the unconfirmed proteins from Heazlewood et al. (2003) were proteins now predicted to be retrotransposon sequences with unknown function.

THE PROTEIN COMPOSITION OF RICE MITOCHONDRIA

A refined dataset of 322 proteins allowed us to assess the functional distribution of the rice mitochondrial proteome as shown in **Figure 1**. There are 99 proteins identified as either components of the five oxidative phosphorylation/respiratory complexes or TCA cycle enzymes, representing 31% of the total set (**Figure 1**). The genes encoding electron transport chain (ETC) proteins are highly expressed across all tissues, which is consistent with the fundamental role of mitochondria in energy production throughout the plant. Interestingly, a series of genes encoding TCA cycle components are highly expressed in anthers, suggesting a high energy requirement for metabolism in this tissue (Huang et al., 2009a). There were 64 proteins identified (20% of total set) that are

[1]http://gene64.dna.affrc.go.jp/RPD

[2]www.gelmap.de/oryza

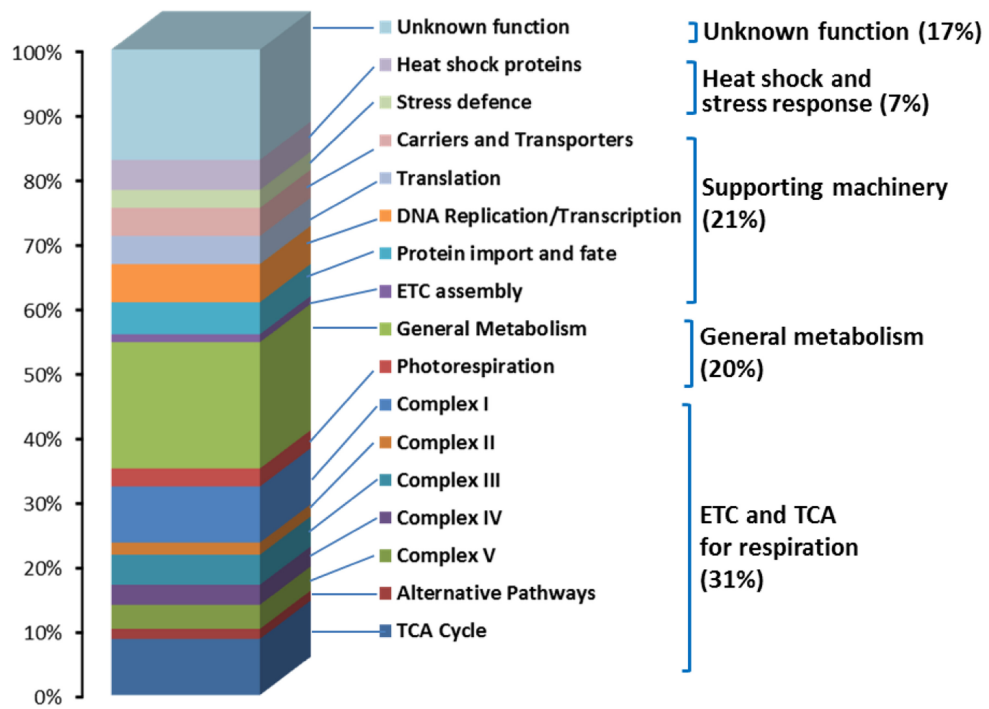


FIGURE 1 | Functional distribution of 322 identified rice mitochondrial proteins. Rice mitochondria were purified from 10-day-old dark-grown seedlings using differential and gradient centrifugation combined with surface charged ZE-FFE mode (Huang et al., 2009a). Rice mitochondria proteins were separated using gel-based and non-gel based methods and digested with trypsin before identification using mass spectrometry (Huang et al., 2009a). Rice mitochondria protein data were extracted from Supplemental Table S3 of Huang et al. (2009a).

thought to be involved in central carbon and nitrogen metabolism (**Figure 1**), such as the inter-conversion of amino acids, photorespiratory glycine oxidation, synthesis of lipids, vitamins, as well as export of organic acids. Within this group, the identification of a 4-methyl-5-thiazole monophosphate biosynthesis protein (Os01g11880) provided new insight into the involvement of rice mitochondria in the process of thiamine biosynthesis. Furthermore, the highly selective expression of genes for components of photorespiratory glycine oxidation in leaf tissues is consistent with the role of mitochondria in photorespiration during photosynthesis in green tissues (Huang et al., 2009a). Proteins involved in supporting machinery such as those for DNA replication, transcription and translation, protein import and fate, ETC assembly as well as carriers and transporters accounted for 21% of the total number of proteins identified. Thirty-three proteins were listed to be involved in DNA replication, transcription, and translation, and 19 proteins were assigned the protein import and fate category (**Figure 1**). Genes encoding mitochondrial enzymes involved in DNA replication, transcription, and translation, as well as protein import and fate are highly expressed in early germinated rice seeds as well as in suspension culture cells, consistent with their role in mitochondrial biogenesis (Huang et al., 2009a). Fifteen heat shock proteins and 9 putative stress response proteins were also identified (**Figure 1**). A total of 55 proteins (17%) were identified for which no known function has been reported (**Figure 1**).

From the 313 nuclear-encoded rice mitochondrial proteins identified, ~65% were predicted to be located in mitochondria by four different subcellular localization prediction software packages (Huang et al., 2009a). The low fidelity of the prediction software is due in part to the use of a limited number of targeting signals in training sets for these software packages (Heazlewood et al., 2004), which again highlights the importance of building experimental evidence for the mitochondrial location of proteins. The number of identified proteins involved in the ETC and TCA cycle in monocotyledon rice mitochondria is similar to number of identified in *Arabidopsis* mitochondrial datasets (Heazlewood et al., 2004) and the corresponding proteins are also largely conserved (Huang et al., 2011). Proteins involved in supporting machinery and stress response were also conserved between the rice and *Arabidopsis* datasets (Huang et al., 2011). The conservation of the proteomes between these diverse species highlights the fundamental role of mitochondria in energy production and metabolism in plants.

THE RICE MITOCHONDRIAL PROTEIN PRESEQUENCE AND ITS CLEAVAGE

N-terminal presequences carry the targeting signals required to import nuclear-encoded mitochondrial proteins and these are cleaved off following the import process to generate mature proteins (Zhang and Glaser, 2002). Analysis of the peptides derived from the digestion of mature rice mitochondrial proteins allowed us to experimentally identify cleavage sites and thus determine 52 rice mitochondrial presequences (Huang et al., 2009b). The average length of these presequences is 45 amino acids. The average pI of the first 10 amino acids was 11.8 with a hydrophobicity index of −1.4. Nearly 90% of the presequences were predicted to form α-helices in this region (Huang et al., 2009b). These features are very similar to those observed for *Arabidopsis* mitochondrial proteins (Huang et al., 2009b).

Amongst the rice mitochondrial presequences three groups of cleavage sites were found: -2 Arg (class I), -3 Arg (class II); and one without any conserved Arg (class III; **Figure 2**). The majority of presequences were -3 Arg (58%) with a smaller contingent of -2 Arg (13%), and a surprisingly large percentage without any conserved arginine (29%; **Figure 2**). In the dominating -3 Arg group, the occurrence of Tyr/Phe/Leu at the -1 position was evenly distributed (**Figure 2**), which differs from the similar *Arabidopsis* -3 Arg group which predominantly features Phe at the -1 position (Huang et al., 2009b). In yeast, an intermediate cleaving peptidase (Icp55, P40051) removes one residue from the presequence after cleavage by the mitochondrial processing peptidase (MPP) when it contains an Arg residue at the -3 position (Vögtle et al., 2009). It is likely that in the mitochondria of rice and *Arabidopsis*, the observed -3 Arg proteins are a consequence of an Icp55-like cleavage, after MPP cutting by an uncharacterized protease (**Figure 2**). Yeast Icp55 (P40051) does have a rice ortholog (Os12g37640; $E = 2 \times e^{-91}$). We have not found Os12g37640 in the rice mitochondrial protein data set (Huang et al., 2009a), but it is predicted to be located in mitochondria by Target P and Mitoprot II. Future functional analysis of Icp55-like peptidase in plant mitochondria is needed to understand its role in stabilizing mitochondrial proteins following MPP cleavage.

CHANGES IN THE RICE MITOCHONDRIA PROTEOME DURING ENVIRONMENTAL STRESS AND PLANT DEVELOPMENT

Most rice proteomic analyses in response to environmental stresses have been conducted at the whole tissue level in leaves or leaf sheaths. Mitochondrial proteins contribute only ~2–5% of total cellular protein and this makes them difficult to quantify in whole tissue protein extracts. For example, there no mitochondrial proteins with significant changes in abundance were detected in rice leaves or sheaths under drought (Salekdeh et al., 2002; Ali

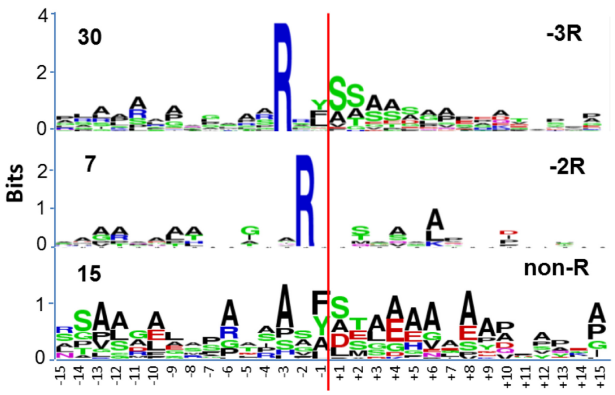


FIGURE 2 | Sequence logo analysis of 52 rice mitochondria protein precursors at the region of the cleavage site. The red line indicates the cleavage sites as supported by proteomic evidence. The cleavage sites are grouped into three classes (class II, -3 Arg; Class I -2 Arg; non-conserved Arg, class III). The numbers on the left side represent the total sequences identified in different groups for sequence logo analysis. Data adapted from Huang et al. (2009b).

and Komatsu, 2006) or infected with the fungi *Rhizoctoni solani* (Lee et al., 2006). A few cases, two highly abundant mitochondrial proteins, both glycine dehydrogenase subunits (Os06g40940; Os01g51410), were significantly changed after treatments of heat (Lee et al., 2007), cold (Lee et al., 2009; Yan et al., 2006), and salt (Kim et al., 2005). Isolated single observations of mitochondrial proteins changing in abundance from whole tissue extracts include pyruvate dehydrogenase (Os02g50620), NADH-ubiquinone oxidoreductase 75 kDa protein (Os03g50540), aconitase hydratase (Os08g09200), dihydrolipoyl dehydrogenase (Os01g22520) after treatment with heat (Lee et al., 2007), cold (Lee et al., 2009), and salt (Kim et al., 2005; Chitteti and Peng, 2007). It is clear that to obtain a more detailed picture of the mitochondrial proteome in response to different environmental stresses, purified mitochondria would be required. Using isolated rice mitochondria for protein oxidation analysis, it was found that a number of proteins are oxidized in the matrix *in vivo* and a group of proteins are particularly susceptible to mild oxidation *in vitro* (Kristensen et al., 2004).

The early growth habitat of rice is often hypoxic or even anoxic (Perata and Voesenek, 2007), meaning that the role and nature of rice mitochondria is especially interesting given their central role in respiration. An early study showed that anoxic rice shoots had the ability to synthesize the same range of mitochondrial proteins as aerobic shoots as long as ATP was supplied, which could be provided *in vivo* by glycolytic reactions even in the absence of oxygen (Couée et al., 1992). Analysis of the soluble rice mitochondrial proteome using 2-D IEF/SDS-PAGE gel separation showed no significant difference between samples derived from anoxic and reoxygenated coleoptiles (Millar et al., 2004). However, BN-SDS-PAGE gels of mitochondrial membrane-associated complexes showed a very low abundance of assembled b/c_1 complex and cytochrome *c* oxidase in anoxic samples and a dramatic increase in the abundance of these complexes after 1 day of air adaptation (Millar et al., 2004). These results suggested that anoxic rice does have the capacity to develop its respiratory machinery but with a discrete and reversible blockage of full mitochondrial biogenesis at Complex III (Millar et al., 2004). Howell et al. (2007) showed that anoxic conditions reduced the efficiency of the general import pathway but not the carrier import pathway in rice mitochondria. Rice mitochondria from anoxic seed embryos 48 h after germination had much lower abundance of TCA cycle enzymes and cytochrome-containing complexes of the respiration chain (Howell et al., 2007). In a whole-cell proteomic analysis, malate dehydrogenase and two ATP synthase subunits were lower in abundance in 6-day-old anoxic coleoptiles compared to similar sized 4-day-old aerated coleoptiles (Shingaki-Wells et al., 2011). The lower abundance of enzymes involved in the TCA cycle or ETC agrees with the previous observation that there is a reduced respiratory capacity in mitochondria isolated from anoxic coleoptiles when compared to aerated or re-oxygenated samples (Millar et al., 2004).

To date the analysis of rice mitochondrial integral membrane proteins has identified seven membrane carrier proteins, one of which was only routinely found in mitochondrial samples from anoxic tissue (Taylor et al., 2010). Further quantitative analysis of the relative abundance of this basic amino acid carrier (BAC; Os10g42299) by QqQ SRM mass spectrometry revealed that Os10g42299 was threefold more abundant in anoxic than in aerated samples (Taylor et al., 2010). Along with the observed anoxic induction of mitochondrial arginase and the accumulation of Arg and Orn, this mitochondrial BAC is likely to play a role in Arg metabolism during O_2 deprivation (Taylor et al., 2010). Such mitochondrial responses may contribute to the exceptional anoxia tolerance of rice seedlings.

Decreasing the rate of photorespiration has become a key target in the further improvement of rice production (Hibberd et al., 2008). Mitochondria are specifically involved in photorespiration via the oxidation of glycine and the export of serine (Walker and Oliver, 1986). The light responsiveness of mitochondrial functions and the induction of photorespiration that occurs when etiolated rice seedlings are exposed to light was recently investigated using a proteomic and metabolomic approach (Huang et al., 2013). Specific steps in mitochondrial TCA cycle metabolism were decreased under high light which correlates with lower respiration rate (Huang et al., 2013). Light treatment reduced the abundance of mitochondrial enzymes in branched chain amino acid metabolism, correlating with a decrease of the abundance of a range of amino acids after a 24 h light treatment of etiolated shoots (Huang et al., 2013). These results have parallels in the diurnal changes observed in mitochondrial function in *Arabidopsis* shoots (Lee et al., 2010). Significant accumulation of glycine decarboxylase (GDC) P, T subunits and serine hydroxymethyltransferase were observed upon light treatment in rice (Huang et al., 2013), which is similar to what has been observed in pea (Walker and Oliver, 1986; Turner et al., 1993) and *Arabidopsis* (Lee et al., 2010). However, the abundance of the GDC H subunit protein in rice was unchanged by light, and the abundance of GDC L subunit protein was halved under high light. The differential change in the stoichiometry of GDC subunits in rice correlates with a fourfold increase in the photorespiration rate of low light-treated plants compared to those treated with high light (Huang et al., 2013).

Cytoplasmic male sterility is a fundamental part of hybrid rice production and relies on plant lines with pollen-specific defects in mitochondrial function (Eckardt, 2006; Wang et al., 2006). Most CMS-associated genes in rice are chimerics composed of a fragment of a normal mitochondrial gene, encoding small and low abundance mitochondrial membrane proteins, and a novel and disruptive sequence that influences the expression or the function of the gene product (Hanson and Bentolila, 2004; Kubo and Newton, 2008). Quantitative proteomic analysis of CMS-related changes in rice anthers has revealed eight proteins with abundances that are at least twofold lower or higher when comparing YTA (CMS) and YTB (isogenic fertile) lines (Sun et al., 2009). However, none of these were mitochondrially encoded proteins. Further quantitative analysis of the mitochondrial proteomes from 10-day-old rice seedlings has revealed a reduced abundance of specific proteins in mitochondrial complexes, particularly complex V, in the YTA line compared with the YTB line (Liu et al., 2012). Interestingly, a sex determination TASSELSEED-2-like protein (Os07g46920) was found 3.2-fold more abundant in the CMS line (Liu et al., 2012). Analysis of the potential links between the increase in the amount of this protein and jasmonic acid

metabolism has identified a lesion in the jasmonic acid synthesis pathway during the development of microspores in CMS plants (Liu et al., 2012).

FUTURE DIRECTIONS

The plant mitochondrial proteome is a changing entity over time, in different tissues/organs and in response to different environments, as revealed by discoveries made in mitochondrial proteome research of the dicotyledon model plant *Arabidopsis* (Sweetlove et al., 2002; Lee et al., 2008, 2011; Tan et al., 2012). The rice mitochondrial proteome is likely to share these dynamics based on the analysis of rice transcript data for genes encoding mitochondrial proteins (Huang et al., 2009a) as well as the proteome response to anoxia and light as discussed above. Further quantitative analysis of the rice mitochondria proteome will provide an even more detailed picture of the diversification of mitochondrial function at the spatial and developmental levels in this key model monocotyledonous species. A broader understanding of the plasticity of rice mitochondria is particularly important for obtaining more clues on the mechanism of pollen abortion in CMS lines. Furthermore, co-expression analysis will reveal mitochondrial proteins with common functions to provide insights into the regulation of rice mitochondrial biogenesis as well as the respiratory stress response.

ACKNOWLEDGMENTS

This research was funded by the ARC Centre of Excellence in Plant Energy Biology (CE0561495) and The University of Western Australia support for the Centre for Comparative Analysis of Biomolecular Networks (CABiN) to A. Harvey Millar. A. Harvey Millar was funded as an ARC Australian Future Fellow (FT110100242). Rachel N. Shingaki-Wells is supported by Grains Research & Development Corporation (GRDC) and an Australian Postgraduate Award PhD scholarship.

REFERENCES

Agrawal, G. K., and Rakwal, R. (2011). Rice proteomics: a move toward expanded proteome coverage to comparative and functional proteomics uncovers the mysteries of rice and plant biology. *Proteomics* 11, 1630–1649.

Ali, G. M., and Komatsu, S. (2006). Proteomic analysis of rice leaf sheath during drought stress. *J. Proteome Res.* 5, 396–403.

Bardel, J., Louwagie, M., Jaquinod, M., Jourdain, A., Luche, S., Rabilloud, T., et al. (2002). A survey of the plant mitochondrial proteome in relation to development. *Proteomics* 2, 880–898.

Chitteti, B. R., and Peng, Z. (2007). Proteome and phosphoproteome differential expression under salinity stress in rice (*Oryza sativa*) roots. *J. Proteome Res.* 6, 1718–1727.

Couée, I., Defontaine, S., Carde, J.-P., and Pradet, A. (1992). Effects of anoxia on mitochondrial biogenesis in rice shoots: modification of in organello translation characteristics. *Plant Physiol.* 98, 411–421.

Cui, J., Liu, J., Li, Y., and Shi, T. (2011). Integrative identification of *Arabidopsis* mitochondrial proteome and its function exploitation through protein interaction network. *PLoS ONE* 6:e16022. doi: 10.1371/journal.pone.0016022

Eckardt, N. A. (2006). Cytoplasmic male sterility and fertility restoration. *Plant Cell* 18, 515–517.

Eubel, H., Lee, C. P., Kou, J., Meyer, E. H., Taylor, N. L., and Millar, A. H. (2007). Free flow eclectrophoresis for purification of plant mitochondria by surface charge. *Plant J.* 52, 583–594.

Eubel, H., Meyer, E. H., Taylor, N. L., Bussell, J. D., O'Toole, N., Heazlewood, J. L., et al. (2008). Novel proteins, putative membrane transporters, and an integrated metabolic network are revealed by quantitative proteomic analysis of *Arabidopsis* cell culture peroxisomes. *Plant Physiol.* 148, 1809–1829.

Hanson, M. R., and Bentolila, S. (2004). Interactions of mitochondrial and nuclear genes that affect male gametophyte development. *Plant Cell* 16, S154–S169.

Heazlewood, J. L., Howell, K. A., Whelan, J., and Millar, A. H. (2003). Towards an analysis of the rice mitochondrial proteome. *Plant Physiol.* 132, 230–242.

Heazlewood, J. L., Tonti-Filippini, J., Verboom, R. E., and Millar, A. H. (2005). Combining experimental and predicted datasets for determination of the subcellular location of proteins in *Arabidopsis*. *Plant Physiol.* 139, 598–609.

Heazlewood, J. L., Tonti-Filippini, J. S., Gout, A. M., Day, D. A., Whelan, J., and Millar, A. H. (2004). Experimental analysis of the *Arabidopsis* mitochondrial proteome highlights signaling and regulatory components, provides assessment of targeting prediction programs, and indicates plant-specific mitochondrial proteins. *Plant Cell* 16, 241–256.

Hibberd, J. M., Sheehy, J. E., and Langdale, J. A. (2008). Using C4 photosynthesis to increase the yield of rice – rationale and feasibility. *Curr. Opin. Plant Biol.* 11, 228–231.

Howell, K. A., Cheng, K., Murcha, M. W., Jenkin, L. E., Millar, A. H., and Whelan, J. (2007). Oxygen initiation of respiration and mitochondrial biogenesis in rice. *J. Biol. Chem.* 282, 15619–15631.

Huang, S., Jacoby, R. P., Shingaki-Wells, R. N., Li, L., and Millar, A. H. (2013) Differential induction of mitochondrial machinery by light intensity correlates with changes in respiratory metabolism and photorespiration in rice leaves. *New Phytol.* doi: 10.1111/nph.12123 [Epub ahead of print].

Huang, S., Millar, A. H., and Taylor, N. L. (2011). "The plant mitochondrial proteome composition and stress response: conservation and divergence between monocots and dicots," in *Plant Mitochondria, Advances in Plant Biology*, Vol. I, ed. F. Kempken (New York: Springer Science+Business Media), 207–239.

Huang, S., Taylor, N. L., Narsai, R., Eubel, H., Whelan, J., and Millar, A. H. (2009a). Experimental analysis of the rice mitochondrial proteome, its biogenesis, and heterogeneity. *Plant Physiol.* 149, 719–734.

Huang, S., Taylor, N. L., Whelan, J., and Millar, A. H. (2009b). Refining the definition of plant mitochondrial presequences through analysis of sorting signals, N-terminal modifications, and cleavage motifs. *Plant Physiol.* 150, 1272–1285.

Kim, D.-W., Rakwal, R., Agrawal, G. K., Jung, Y.-H., Shibato, J., Jwa, N.-S., et al. (2005). A hydroponic rice seedling culture model system for investigating proteome of salt stress in rice leaf. *Electrophoresis* 26, 4521–4539.

Klodmann, J., Senkler, M., Rode, C., and Braun, H.-P. (2011). Defining the protein complex proteome of plant mitochondria. *Plant Physiol.* 157, 587–598.

Komatsu, S. (2005). Rice Proteome Database: a step towards functional analysis of the rice genome. *Plant Mol. Biol.* 59, 179–190.

Kristensen, B. K., Askerlund, P., Bykova, N. V., Egsgaard, H., and Møller, I. M. (2004). Identification of oxidised proteins in the matrix of rice leaf mitochondria by immunoprecipitation and two-dimensional liquid chromatography-tandem mass spectrometry. *Phytochemistry* 65, 1839–1851.

Kruft, V., Eubel, H., Jansch, L., Werhahn, W., and Braun, H. P. (2001). Proteomic approach to identify novel mitochondrial proteins in *Arabidopsis*. *Plant Physiol.* 127, 383–396.

Kubo, T., and Newton, K. J. (2008). Angiosperm mitochondrial genomes and mutations. *Mitochondrion* 8, 5–14.

Lee, C. P., Eubel, H., and Millar, A. H. (2010). Diurnal changes in mitochondrial function reveal daily optimization of light and dark respiratory metabolism in *Arabidopsis*. *Mol. Cell. Proteomics* 9, 2125–2139.

Lee, C. P., Eubel, H., O'Toole, N., and Millar, A. (2011). Combining proteomics of root and shoot mitochondria and transcript analysis to define constitutive and variable components in plant mitochondria. *Phytochemistry* 72, 1092–1108.

Lee, C. P., Eubel, H., O'Toole, N., and Millar, A. H. (2008). Heterogeneity of the mitochondrial proteome for photosynthetic and non-photosynthetic *Arabidopsis* metabolism. *Mol. Cell. Proteomics* 7, 1297–1316.

Lee, D.-G., Ahsan, N., Lee, S.-H., Kang, K. Y., Bahk, J. D., Lee, I.-J., et al. (2007). A proteomic approach in analyzing heat-responsive proteins in rice leaves. *Proteomics* 7, 3369–3383.

Lee, D.-G., Ahsan, N., Lee, S.-H., Lee, J. J., Bahk, J. D., Kang, K. Y., et al. (2009). Chilling stress-induced proteomic changes in rice roots. *J. Plant Physiol.* 166, 1–11.

Lee, J., Bricker, T. M., Lefevre, M., Pinson, S. R. M., and Oard, J. H. (2006). Proteomic and genetic approaches to identifying defence-related proteins in rice challenged with the fungal pathogen *Rhizoctonia solani*. *Mol. Plant Pathol.* 7, 405–416.

Liu, G., Tian, H., Huang, Y.-Q., Hu, J., Ji, Y.-X., Li, S.-Q., et al. (2012). Alterations of mitochondrial protein assembly and jasmonic acid biosynthesis pathway in Honglian (HL)-type cytoplasmic male sterility rice. *J. Biol. Chem.* 287, 40051–40060.

Millar, A. H., Heazlewood, J. L., Kristensen, B. K., Braun, H.-P., and Møller, I. M. (2005). The plant mitochondrial proteome. *Trends Plant Sci.* 10, 36–43.

Millar, A. H., Sweetlove, L. J., Giege, P., and Leaver, C. J. (2001). Analysis of the *Arabidopsis* mitochondrial proteome. *Plant Physiol.* 127, 1711–1727.

Millar, A. H., Trend, A. E., and Heazlewood, J. L. (2004). Changes in the mitochondrial proteome during the anoxia to air transition in rice focus around cytochrome-containing respiratory complexes. *J. Biol. Chem.* 279, 39471–39478.

Perata, P., and Voesenek, L. A. C. J. (2007). Submergence tolerance in rice requires Sub1A, an ethylene-response-factor-like gene. *Trends Plant Sci.* 12, 43–46.

Salekdeh, G. H., Siopongco, J., Wade, L. J., Ghareyazie, B., and Bennett, J. (2002). Proteomic analysis of rice leaves during drought stress and recovery. *Proteomics* 2, 1131–1145.

Senkler, M., and Braun, H.-P. (2012). Functional annotation of 2D protein maps: the GelMap portal. *Front. Plant Sci.* 3:87. doi: 10.3389/fpls.2012.00087.

Shingaki-Wells, R. N., Huang, S., Taylor, N. L., Carroll, A. J., Zhou, W., and Millar, A. H. (2011). Differential molecular responses of rice and wheat coleoptiles to anoxia reveal novel metabolic adaptations in amino acid metabolism for tissue tolerance. *Plant Physiol.* 156, 1706–1724.

Sun, Q., Hu, C., Hu, J., Li, S., and Zhu, Y. (2009). Quantitative proteomic analysis of CMS-related changes in Honglian CMS rice anther. *Protein J.* 28, 341–348.

Sweetlove, L. J., Heazlewood, J. L., Herald, V., Holtzapffel, R., Day, D. A., Leaver, C. J., et al. (2002). The impact of oxidative stress on *Arabidopsis* mitochondria. *Plant J.* 32, 891–904.

Tan, Y.-F., Millar, A. H., and Taylor, N. L. (2012). Components of mitochondrial oxidative phosphorylation vary in abundance following exposure to cold and chemical stresses. *J. Proteome Res.* 11, 3860–3879.

Taylor, N. L., Howell, K. A., Heazlewood, J. L., Tan, T. Y. W., Narsai, R., Huang, S., et al. (2010). Analysis of the rice mitochondrial carrier family reveals anaerobic accumulation of a basic amino acid carrier involved in arginine metabolism during seed germination. *Plant Physiol.* 154, 691–704.

Turner, S. R., Hellens, R., Ireland, R., Ellis, N., and Rawsthorne, S. (1993). The organization and expression of the gene encoding the mitochondrial glycine decarboxylase complex and serine hydroxymethyltransferase in pea (*Pisum sativum*). *Mol. Gen. Genet.* 236, 402–408.

Vögtle, F. N., Wortelkamp, S., Zahedi, R. P., Becker, D., Leidhold, C., Gevaert, K., et al. (2009). Global analysis of the mitochondrial N-proteome identifies a processing peptidase critical for protein stability. *Cell* 139, 428–439.

Walker, J. L., and Oliver, D. J. (1986). Light-induced increases in the glycine decarboxylase multienzyme complex from pea leaf mitochondria. *Arch. Biochem. Biophys.* 248, 626–638.

Wang, Z., Zou, Y., Li, X., Zhang, Q., Chen, L., Wu, H., et al. (2006). Cytoplasmic male sterility of rice with Boro II cytoplasm is caused by a cytotoxic peptide and is restored by two related PPR motif genes via distinct modes of mRNA silencing. *Plant Cell* 18, 676–687.

Yan, S.-P., Zhang, Q.-Y., Tang, Z.-C., Su, W.-A., and Sun, W.-N. (2006). Comparative proteomic analysis provides new insights into chilling stress responses in rice. *Mol. Cell. Proteomics* 5, 484–496.

Zhang, X. P., and Glaser, E. (2002). Interaction of plant mitochondrial and chloroplast signal peptides with the Hsp70 molecular chaperone. *Trends Plant Sci.* 7,14–21.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 18 December 2012; accepted: 22 January 2013; published online: 07 February 2013.

Citation: Huang S, Shingaki-Wells RN, Taylor NL and Millar AH (2013) The rice mitochondria proteome and its response during development and to the environment. Front. Plant Sci. ***4****:16. doi: 10.3389/fpls.2013.00016*

This article was submitted to Frontiers in Plant Proteomics, a specialty of Frontiers in Plant Science.

Copyright © 2013 Huang, Shingaki-Wells, Taylor and Millar. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.

frontiers in
PLANT SCIENCE

ORIGINAL RESEARCH ARTICLE
published: 15 April 2013
doi: 10.3389/fpls.2013.00084

The mitochondrial complexome of *Medicago truncatula*

Leonard Muriithi Kiirika[1], Christof Behrens[2], Hans-Peter Braun[2] and Frank Colditz[1]*

[1] Department of Plant Molecular Biology, Institute for Plant Genetics, Leibniz University Hannover, Hannover, Germany
[2] Department of Plant Proteomics, Institute for Plant Genetics, Leibniz University Hannover, Hannover, Germany

Edited by:
Nicolas L. Taylor, The University of Western Australia, Australia

Reviewed by:
Ko Noguchi, The University of Tokyo, Japan
Karsten Niehaus, Bielefeld University, Germany

****Correspondence:***
Frank Colditz, Department of Plant Molecular Biology, Leibniz University Hannover, Herrenhäuser Straße 2, 30419 Hannover, Germany.
e-mail: colditz@genetik.uni-hannover.de

Legumes (Fabaceae, Leguminosae) are unique in their ability to carry out an elaborate endosymbiotic nitrogen fixation process with rhizobia proteobacteria. The symbiotic nitrogen fixation enables the host plants to grow almost independently of any other nitrogen source. Establishment of symbiosis requires adaptations of the host cellular metabolism, here foremost of the energy metabolism mainly taking place in mitochondria. Since the early 1990s, the galegoid legume *Medicago truncatula* Gaertn. is a well-established model for studying legume biology, but little is known about the protein complement of mitochondria from this species. An initial characterization of the mitochondrial proteome of *M. truncatula* (Jemalong A17) was published recently. In the frame of this study, mitochondrial protein complexes were characterized using Two-dimensional (2D) Blue native (BN)/SDS-PAGE. From 139 detected spots, the "first hit" (=most abundant) proteins of 59 spots were identified by mass spectrometry. Here, we present a comprehensive analysis of the mitochondrial "complexome" (the "*protein complex proteome*") of *M. truncatula via* 2D BN/SDS-PAGE in combination with highly sensitive MS protein identification. In total, 1,485 proteins were identified within 158 gel spots, representing 467 unique proteins. Data evaluation by the novel GelMap annotation tool allowed recognition of protein complexes of low abundance. Overall, at least 36 mitochondrial protein complexes were found. To our knowledge several of these complexes were described for the first time in *Medicago*. The data set is accessible under http://www.gelmap.de/medicago/. The mitochondrial protein complex proteomes of *Arabidopsis* available at http://www.gelmap.de/arabidopsis/ and *Medicago* are compared.

Keywords: *Medicago truncatula*, mitochondrial complexome, 2D BN/SDS-PAGE, GelMap annotation tool, mitochondrial prohibitins

INTRODUCTION

Mitochondria are of great importance for ATP production in eukaryotic cells. Redox equivalents in the form of NADH and FADH are re-oxidized by the mitochondrial respiratory chain located in the inner mitochondrial membrane. These reactions are conducted particularly by large protein complexes forming the Oxidative Phosphorylation (OXPHOS) system, which transfer electrons to molecular oxygen. Coevally, a proton gradient is generated across the membrane. The backflow of protons into the mitochondrial matrix space mediates phosphorylation of ADP by the ATP synthase complex. A special feature of plant mitochondria is the presence of additional "alternative" oxidoreductases in the OXPHOS system (Heazlewood et al., 2003a; Brugière et al., 2004). Besides OXPHOS, mitochondria carry out additional biochemical functions, like amino acid and nucleotide metabolism, as well as synthesis of cofactors such as heme, biotin, lipoic acid (Dubinin et al., 2011). In plants, mitochondria also carry out some reactions of the photorespiratory pathway (glycolate cycle). The protein complement of *Arabidopsis*, potato, rice, and pea mitochondria have been analyzed extensively by gel-based and gel-free proteomic approaches (Klodmann et al., 2011). Many of the enzymes present in mitochondria are organized in the form of protein complexes.

Two-dimensional (2D) Blue native (BN)/SDS-PAGE is an excellent system for the separation of mitochondrial protein complexes in their native forms and subsequent resolution into their subunits (Klodmann et al., 2011). Using this approach, individual protein complexes of the respiratory chain of plant mitochondria were systematically characterized [e.g., characterization of complex I (Heazlewood et al., 2003b; Meyer et al., 2008; Klodmann et al., 2010; Klodmann and Braun, 2011); characterization of protein complex abundances of complexes I to V in different organs of *Arabidopsis* (Peters et al., 2012)]. By combining 2D BN/SDS-PAGE with sensitive mass spectrometry-based protein identification and subsequent annotation with the novel "*GelMap*" software tool (Senkler and Braun, 2012)[1], a systematic characterization of protein complexes became possible. GelMap allows annotation of proteins according to functional categories, as well as assignment of entire sets of proteins to individual protein spots (Klodmann et al., 2011). GelMap was initially developed to functionally annotate proteins from 2D Isoelectric focusing (IEF)/SDS gels (Rode et al., 2011). For annotation of proteins separated *via* 2D BN/SDS-PAGE, the software was modified (Senkler and Braun, 2012) to allow visualization of protein complexes and their subunits even when they are of low abundance and/or are covered by higher abundant proteins. Thus, GelMap allows the systematic stock take

[1] http://www.gelmap.de/

of the mitochondrial *protein complex proteome*, the *complexome*. In *Arabidopsis*, this led to the identification of 471 distinct mitochondrial proteins and more than 35 different protein complexes (Klodmann et al., 2011).

Legumes frequently interact with soil-borne microbes (Colditz and Braun, 2010). Foremost the legume rhizobia (LR) symbiosis is of high economic value, since LR provides the host legume with independence of other nitrogen sources and helps in the production of protein-rich fruits and seeds. However, it strongly relies on the energy metabolism of the host cells (Dubinin et al., 2011). Since most of the microbial interactions to legumes are located in the rhizosphere, particularly the hosts' root cells are in the focus of molecular research. Differences in the protein patterns of root-derived cell suspension cultures from the model legume *M. truncatula* were observed after inoculation with spores from an oomycete pathogen (Trapphoff et al., 2009). They closely match those of infected plant root cells. Thus, these root-derived cell suspension cultures may be used as an adequate model system for microbe – plant interaction studies. To date, only few studies investigated cellular sub-proteomes from the legume plant family. For example, the response of the pea mitochondrial proteome to abiotic stress conditions was investigated (Taylor et al., 2005). More recently, root plastids from *M. truncatula* were proteomically analyzed (Daher et al., 2010). The first proteomic reference maps (*via* 2D IEF/SDS-PAGE and BN/SDS-PAGE) for purified mitochondrial fractions were established by Dubinin et al. (2011). This study used the "first hit" (=most abundant) proteins from MALDI-TOF MS/MS for each analyzed protein spot.

Recently, the draft sequence of the *M. truncatula* euchromatin was published, covering almost 95% of all predicted genes (Young et al., 2011). A new database for *Medicago* DNA sequences, LegProt db, was established (Lei et al., 2011). As a consequence, chances of identifying proteins based on MS analyses of tryptic peptide mixtures of *Medicago* samples improved considerably and this database was also used for protein identification presented in this study. At the same time, sensitivity of MS systems used for protein analyses increased. Finally, the GelMap software tool for the first time allows extensive annotation of gel-based proteome data. Together, these developments sparked a carefully re-analysis of the mitochondrial proteome of *Medicago*.

MATERIALS AND METHODS

PREPARATION OF MITOCHONDRIA FROM *M. TRUNCATULA*, 2D BN/SDS-PAGE

Mitochondria were isolated from *M. truncatula* ("Jemalong A17") root cell suspension cultures as described by Dubinin et al. (2011). For 2D BN/SDS-PAGE, aliquots of isolated mitochondria equivalent to 1 mg protein were used. BN electrophoretic separation of mitochondrial protein complexes was performed according to Schägger and von Jagow (1991) with modifications (Dubinin et al., 2011) using a Protean II (16 cm × 16 cm) electrophoresis chamber (BioRad), a polyacrylamide concentration gradient of 4.5–16% acrylamide (top to bottom) for BN first dimension gel and 16.5% acrylamide Tricine/SDS-PAGE for the second gel dimension. Gels were stained with Coomassie blue-colloidal (BioRad) overnight and scanned on an UMAX Power Look III Scanner (UMAX Technologies) as described before (Colditz et al., 2007).

MASS SPECTROMETRY

Protein spots of 1.4 mm diameter were cut from Coomassie-stained gels using a GelPal Protein Excision manual spot picker (Genetix, Great Britain) and in-gel digested with Trypsin as described by Klodmann et al. (2010). Tryptic peptides were further analyzed by nanoHPLC (Proxeon, Thermo Scientific) coupled to electrospray ionization quadrupole time of flight MS (micrOQTOF Q II, Bruker Daltonics), using all settings and parameters as described previously (Klodmann et al., 2011). Data processing and protein identification was carried out with ProteinScape 2.0 (Bruker Daltonics) and the MASCOT search engine querying three *Medicago*-specific protein databases [*Mt3.5 ProteinSeq, NCBI Medicago truncatula protein, and Mtf(asta)*[2]] available at the LegProt db (Lei et al., 2011) as well as *Swiss Prot*, using the following parameters: trypsin/P; one missed cleavage allowed; fixed modifications: carbamidomethylation (C), variable modifications: acetylation (N) and oxidation (M); precursor ion mass tolerance, 30 ppm; peptide score >24; charges 1+, 2+, 3+. Protein and peptide assessments with MASCOT scores above 25 were considered. Identified proteins were further analyzed for their sub-cellular localization using their homologous *Arabidopsis* accessions (according to TAIR 10 db) queried against the SUBA III database (Heazlewood et al., 2007)[3].

MITOCHONDRIAL BN REFERENCE MAP *VIA* GelMap

After protein identification, the reference map was visualized using the GelMap platform (see text footnote 2; Senkler and Braun, 2012). For this purpose, spots of a scanned 2D BN/SDS gel were automatically detected and were given consecutive spot numbers with corresponding *x*- and *y*-coordinates by the Delta 2D (4.2) software (Decodon, Greifswald, Germany) (Figures S1–S3 in Supplementary Material). An Excel (Microsoft) file containing this information and the corresponding gel image (.jpg) were then imported into GelMap. MS/MS results were uploaded as well. Detailed information on building a GelMap is available under http://www.gelmap.de/howto.

RESULTS AND DISCUSSION

2D BN/SDS-PAGE OF MITOCHONDRIAL PROTEIN FRACTIONS FROM *M. TRUNCATULA* CELLS

In order to separate mitochondrial proteins from *M. truncatula* root-derived cell suspension cultures, purified mitochondrial fractions were prepared according to an optimized protocol published by Dubinin et al. (2011). Proteins from five independent mitochondrial isolations were then separated by 2D BN/SDS gel electrophoresis. Spot patterns on the gels were highly similar as revealed by Delta 2D analysis (data not shown). From this set of five, a representative gel was selected for MS analyses as well as for online data presentation *via* the GelMap software

[2]http://bioinfo.noble.org/manuscript-support/legumedb/
[3]http://suba.plantenergy.uwa.edu.au/

tool (**Figure 1**)[4]. Using the same gel, a spot coordinate file was generated as described in Section "Materials and Methods" and uploaded simultaneously.

MS-BASED PROTEIN IDENTIFICATION AND ANNOTATION OF *M. TRUNCATULA* MITOCHONDRIAL PROTEINS

All 158 protein spots encircled in **Figure 1** were analyzed *via* nLC ESI-MS measurements. In contrast to the previous analysis of the 2D BN/SDS-PAGE-separated mitochondrial proteome by Dubinin et al. (2011), protein identification was achieved using the *Medicago*-specific protein databases from the LegProt db (Lei et al., 2011), resulting in improved protein identification rates. In total, 1,485 proteins were identified within the selected protein spots, representing 467 unique proteins. For nine proteins, no accessions were found in MtGI. Interestingly, 12 of the uniquely identified proteins in MtGI have no homologs in *Arabidopsis* (spots 47, 75, 76, 77, 85, 98, 101, 106, 113, 116, 132, 133, 137). Among them are three legume-specific proteins involved in symbiosis to Rhizobial bacteria: a legume lectin (ID 101) and two nodulins (nodulin 3, spot 106; nodulin 25, spot 85), as well as a prefoldin protein (spot 77) which is supposed to be also legume-specific. The majority of plant lectins possess a signal peptide and thus are targeted *via* the secretory pathway into the vacuolar and extracellular compartments (Lannoo and Van Damme, 2010). Recently, evidence was given that plants additionally synthesize small amounts of lectins in response to changing environmental conditions or stress factors, which are referred to as "inducible" lectins (Lannoo and Van Damme, 2010). Contrary to the majority of plant lectins, these inducible lectins have been shown to be located to the cytosolic/nuclear compartment, and even their involvement in mitochondrial-induced programed cell death (PCD) has been reported (Van Damme et al., 2004). How these proteins are involved in mitochondrial metabolism should

[4]http://www.gelmap.de/medicago/

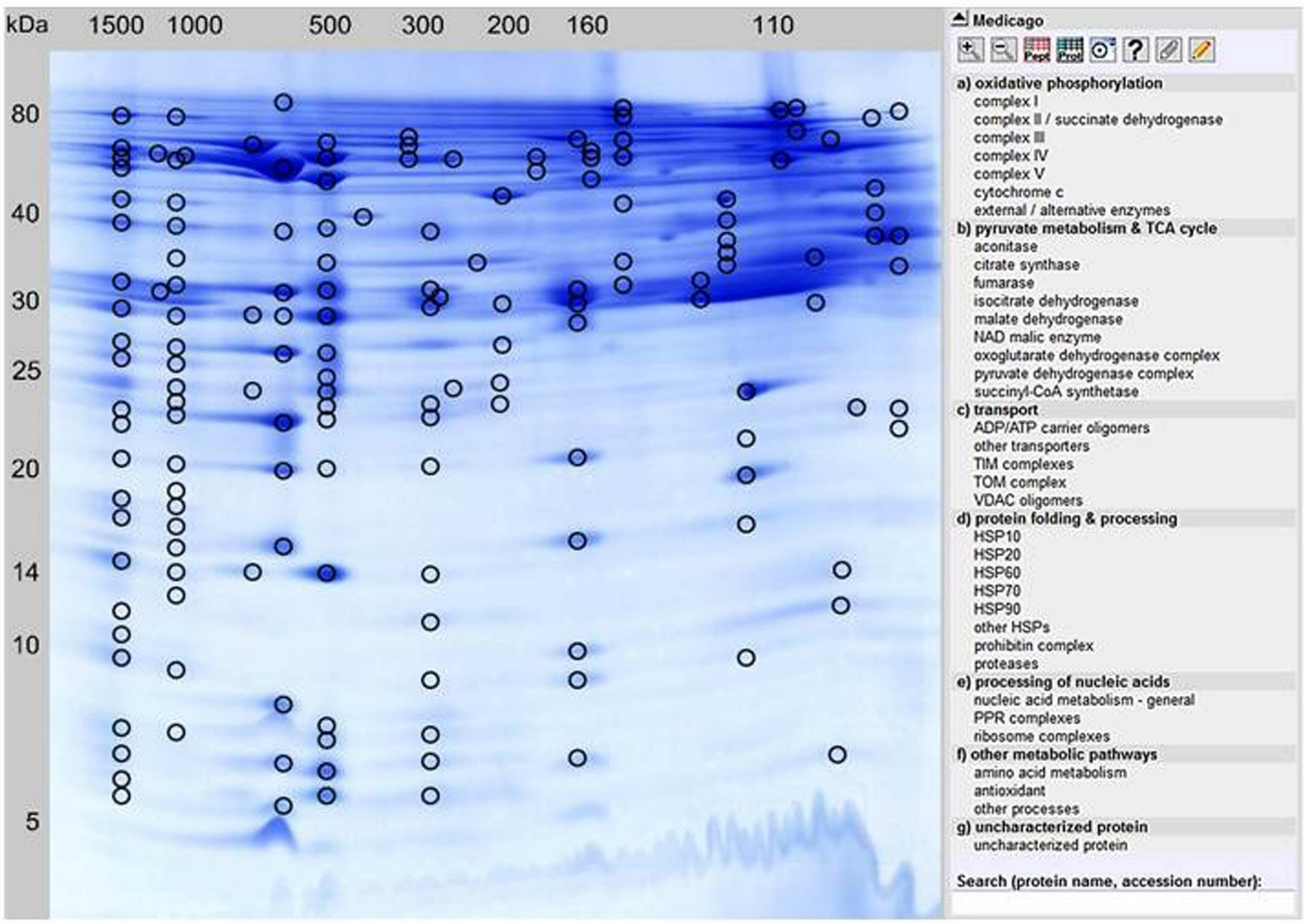


FIGURE 1 | GelMap reference map of the *M. truncatula* mitochondrial protein complexe proteome/complexome (http://www.gelmap.de/medicago/). Hundred and fifty-eight protein spots separated by 2D BN/SDS-PAGE and identified by MS are marked by circles. Most protein spots include multiple protein annotations. By clicking a certain protein spot, all identified proteins within this spot are shown in a pop-up window, beginning with the protein identification of the highest MASCOT score. The menu to the right lists classes of physiological functions for mitochondrial protein complexes. By clicking on the selected protein complex in this menu, accessions of all included individual proteins as well as the corresponding protein spots in the gel image are highlighted. Alternatively, a protein can be found in GelMap by the Search tool at the bottom to the right.

be analyzed in future studies. Nevertheless, we cannot exclude that the identified lectins are contaminants in our mitochondrial fractions. In addition, two *Medicago* hexokinases (hexokinase 7, spots 98a and 113; hexokinase 8, spots 113 and 132) have no homologoues in *Arabidopsis*.

By clicking a spot in **Figure 1**, the description(s) of the identified protein(s) will appear in a pop-up window. These descriptions are hyperlinked and by pointing the mouse cursor at any one of them a detailed information is provided. This includes: spot number, protein name, MS score, calculated and apparent molecular mass (for both gel dimensions), sequence coverage, number of matching peptides, the Tentative Consensus (TC) accession from the *Medicago truncatula Gene Index* [MtGI, Release 11.0 (March 23, 2011), at Dana-Farber Cancer Institute, Harvard School of Public Health, Boston, MA, USA], the TAIR accession from the *Arabidopsis* homologous protein, protein name and origin, the protein database where the protein was identified, its protein complex identity, physiological function, and sub-cellular localization according to the SUBA III database (Heazlewood et al., 2007). Most protein spots shown in **Figure 1** contain several different proteins. Within the pop-up window of each spot, they are sorted according to their relative abundances, as implicated by their respective Mascot scores. Most abundant proteins are listed at the top, least abundant proteins at the bottom. By showing all identified protein hits GelMap promotes the detection of low abundant protein complexes which cannot be found when only the most abundant hits per spot are considered. In case of multiple protein annotation for an individual spot, another mouse-click on the protein of choice opens a new window that includes detailed information. For several proteins, identification in MtGI was not yet possible. In these cases, heterologous protein identifications of the most homologous *Arabidopsis* proteins/accessions are given.

In order to assess the purity of isolated mitochondrial fractions, the sub-cellular localization of all proteins identified was evaluated *via* the Sub-Cellular Proteomic Database (SUBA III, see text footnote 4). Since this database collects experimental data and in silico predictions of the localization of proteins in *Arabidopsis*, the corresponding TAIR homologs of each identified *Medicago* protein were used to assess the intracellular whereabouts of the *Medicago* proteins. At least for the "first hit" identifications, prediction data are available (except for one of the 158 "first hit" proteins). From overall 157 first protein hits, 145 proteins (92%) are assigned to mitochondria. Considering all 467 unique proteins, sub-cellular localization information is available for 413 proteins. The percentage of mitochondrial proteins in this dataset is lower (287 proteins = 69.5%). Twenty-five proteins (=6%) represent cytosolic proteins according to SUBA evaluation. A considerable number of proteins are assigned to other cellular compartments: 6% to plastids, 5% to the nucleus, 3.6% to membrane structures (plasma membrane, endomembrane), and 1.2% to the cells vacuoles. For 17 proteins (4%), no SUBA predictions are available because of a lack of experimental data. These proteins are labeled as "NEW mitochondria" in our GelMap since they represent candidates for mitochondrial proteins. Considering that the "first hit" proteins, 92% of which are of predicted mitochondrial origin, are on average significantly more abundant than the proteins of lower MASCOT scores within each spot, we estimate that the overall purity of our mitochondrial fraction was in the range of 85%.

ANNOTATION OF THE *M. TRUNCATULA* MITOCHONDRIAL COMPLEX PROTEOME/COMPLEXOME

Two-dimensional BN/SDS reference maps generated with GelMap enable annotation and assignment of all proteins identified that belong to one certain functional protein complex, for example the complexes of the OXPHOS system (Klodmann et al., 2011). Systematic evaluation of all apparent protein complexes allows establishing the complexome of the protein sample.

For this purpose, the "physiological function" menu to the right of the GelMap (**Figure 1**) should be used. Here, functional classification of all identified subunits is given, next to their assignment to protein complexes. According to our GelMap evaluation, 36 mitochondrial protein complexes were found in the *Medicago* mitochondrial fractions. Several of these protein complexes were described for the first time in this model legume.

EVALUATION OF THE *MEDICAGO* MITOCHONDRIAL PROTEIN COMPLEX PROTEOME *VIA* GelMap

The GelMap of the *M. truncatula* mitochondrial complex proteome presented here aims to systematically analyze the complexome of this sub-cellular compartment. Since the GelMap annotation portal is web-based, the data set is open to the scientific community and public data evaluation is possible and welcome. The *Medicago* mitochondrial GelMap includes proteins with MASCOT scores $\geq$25 as well as proteins identified by one single peptide in order to provide a maximum of information. Thus, the currently presented data should be treated with caution because false positive identifications are not completely excluded. At the same time, for some proteins, MS-spectra were recorded but no positive identification was possible from the data. To overcome both of these drawbacks we will continuously update this protein reference map when progress in the annotation of *Medicago* genome allows better identification of proteins.

While *Medicago* is still trailing *Arabidopsis* in respect to genome annotations, the data produced in this study nevertheless allow a comparison of the mitochondrial complex proteome of both species. Most complexes found in *Arabidopsis* are also present in *Medicago*, which is not surprising given the importance of mitochondrial function for the energy metabolism of plants. However, some protein complexes of *Medicago* mitochondria seem to lack comparable counterparts in *Arabidopsis*.

A short overview and characterization of the major *Medicago* protein complexes is given below:

- *Complex I* is the biggest respiratory chain complex. It runs at 1000 kDa and 41 of its subunits could be identified, although some subunits seem to be missing since more were found in *Arabidopsis* (Klodmann et al., 2010). Several of its subunits are plant specific, e.g., the gamma carbonic anhydrases (Klodmann et al., 2010). Subcomplexes of complex I are found at MWs of approximately 500, 280, and 140 kDa, which probably represent assembly intermediates of the complex. L-galactono-1,4-lactone dehydrogenase (GLDH), which was recently described to form part of three assembly intermediates of complex I in *Arabidopsis* (Schertl et al., 2012), also does not form part of fully assembled complex I in *Medicago*.
- Complex I interacts with dimeric cytochrome c reductase to form the $I + III_2$ *supercomplex* of 1500 kDa.

- *Complex II* is resolved in its main form at 160 kDa, but an additional version is present at 110 kDa. The occurrence of this smaller form of complex II was described before by Dubinin et al. (2011) for *Medicago*, and by Klodmann et al. (2011) for *Arabidopsis*. Overall, five different complex II subunits were identified, including the two plant specific subunits SDH5 and SDH7-2. According to Dubinin et al. (2011), the relative abundance of the two forms of complex II differs between *Medicago* and *Arabidopsis*, which was confirmed by our new study.
- *Complex III* is a dimer in its active form and runs at 500 kDa. Nine of the expected 10 subunits were found. With complex I, it is involved in the formation of a supercomplex.
- Of *complex IV*, six different subunits were identified.
- *Complex V* (ATP synthase) can be found at 600 kDa on the 2D BN/SDS gel. Thirteen distinct subunits were identified. As described previously by Dubinin et al. (2011), the complex V dimer (V_2) is of higher abundance in mitochondria isolated from *Medicago* cells than in *Arabidopsis*.
- Several of the *alternative oxidoreductases of the plant OXPHOS system* were identified: AOX1a, AOX2, AOX3, NDA1, NDA2, NDB1, and NDB4. The NDA and NDB subunits are present at 160kDa and probably form a protein complex as reported for *Arabidopsis* (Klodmann et al., 2011). AOX is found at many different horizontal positions in the *Medicago* GelMap, possibly indicating its binding to a variety of protein complexes (**Figure 2A**).
- Since *cytochrome c* migrates at low MW (15 kDa) in the second gel dimension, but at much higher MWs in the native first dimension (100–200 kDa), indication for an association with other proteins (complex IV subunits, GLDH) is given.
- Besides the membrane-bound protein complexes, several enzymes of the *citric acid cycle* also are involved in forming protein complexes: aconitase forms a putative dimer at 150 kDa in the native gel dimension, NAD-dependent malic enzymes was described to form a heterohexamer in *Arabidopsis* at 369 kDa (Klodmann et al., 2011), which was also found in *Medicago*. A pyruvate dehydrogenase E1 and E3 subcomplex was detected at 140 kDa, which has also been described in *Arabidopsis* (Klodmann et al., 2011).
- Several *ADP/ATP carrier* oligomers were identified at MWs between 90 and 110 kDa on the native gel dimension.
- Interestingly, several *ABC transporters* identified in the MW range between 160 and 780 kDa were found, which were not described in the *Arabidopsis* mitochondrial proteome (**Figure 2B**).
- *TIM* and *TOM* protein complexes: the TOM complex containing the subunits TOM20-2, TOM20-3, and TOM22-V was found at a MW of 260 kDa. In addition, TOM complex subunits (TOM20-2, TOM40) were found at MWs of 1000 and 1500 kDa. In the same MW range, also TIM subunits were found (TIM17-22, TIM17-2), suggesting the presence of a large TIM/TOM translocon supercomplex in *Medicago*.
- Several *VDAC* oligomers were identified between 90 and 500 kDa, which form either distinct complexes or artificially aggregated during solubilization or electrophoresis of the native gel dimension.
- *HSP*s form protein complexes in *M. truncatula* mitochondria: HSP60 complex is present in its main form at a MW of 600 kDa, where also other HSPs (HSP70, HSP90) were detected.
- Three distinct *prohibitin complexes* were found at 160, 300, and 1200 kDa. Increased abundance of mitochondrial prohibitins in *M. truncatula* was reported previously (Dubinin et al., 2011); presence of varying prohibitin complexes was not found in *Arabidposis* (**Figure 2C**). Inoculation of *Medicago* cells with virulent spores of an oomycete pathogen (according to Trapphoff et al., 2009) resulted in accumulation and increased abundance of the prohibitin protein complex in BN gels (F. Colditz, unpublished). Recent findings indicate that prohibitins are involved in mediating stress tolerance (abiotic stress, pathogen infection, and elicitor signaling) as well as triggering retrograde signals in response to mitochondrial dysfunction (Van Aken et al., 2010).
- *Proteases*: AAA-type ATPase family protein was found at 1500 kDa, together with two cysteine proteinases. ClpA/ClpB protease subunits are present between 150 and 600 kDa, the LON protease at 600 kDa.
- Interestingly, nine different *PPR proteins* involved in nucleic acid metabolism were identified at molecular masses ranging from 90 to 1500 kDa. They form part of so far unknown protein complexes.
- Eight distinct subunits of *ribosomal protein subcomplexes* were identified between 50 and 1000 kDa.
- A large set of *antioxidant proteins* like SOD, glutathione S-transferase and catalase were found in the molecular mass range of 120–1500 kDa on the native gel dimension, indicating their presence within protein complexes.
- Many further enzymes listed under "further metabolic pathways" much likely form part of *other protein complexes*, since they migrate at significantly higher molecular masses on the native (BN) gel dimension than on the denaturing gel dimension.

CONCLUSION

This GelMap was built to systematically define the mitochondrial protein complex proteome of the model legume *M. truncatula*. Generally, our GelMap presents protein candidates that may form protein complexes. It does not provide final proof for the presence of novel complexes: if a protein complex is described here for the first time, its occurrence should be verified by further independent experiments.

Most of the identified protein complexes to be present in *Medicago* mitochondria were already identified and characterized in mitochondria from *Arabidopsis* cell suspension cultures (Klodmann et al., 2011). However, some protein complexes found in *Medicago* mitochondria, such as ABC transporters, TIM/TOM translocon supercomplex, and distinct prohibitin complexes, seem to lack comparable counterparts in *Arabidopsis*.

Since molecular studies with legumes are particularly done to characterize interactions of plants to soil-borne microbes (Colditz and Braun, 2010), the *Medicago* mitochondrial GelMap should promote the analysis of infection-related proteomic alterations at a sub-cellular level. As a next step, analyses of the *Medicago* mitochondrial proteome after microbial infections should be carried out, especially after infection with agronomically important and legume-specific rhizobial bacteria, in order to monitor

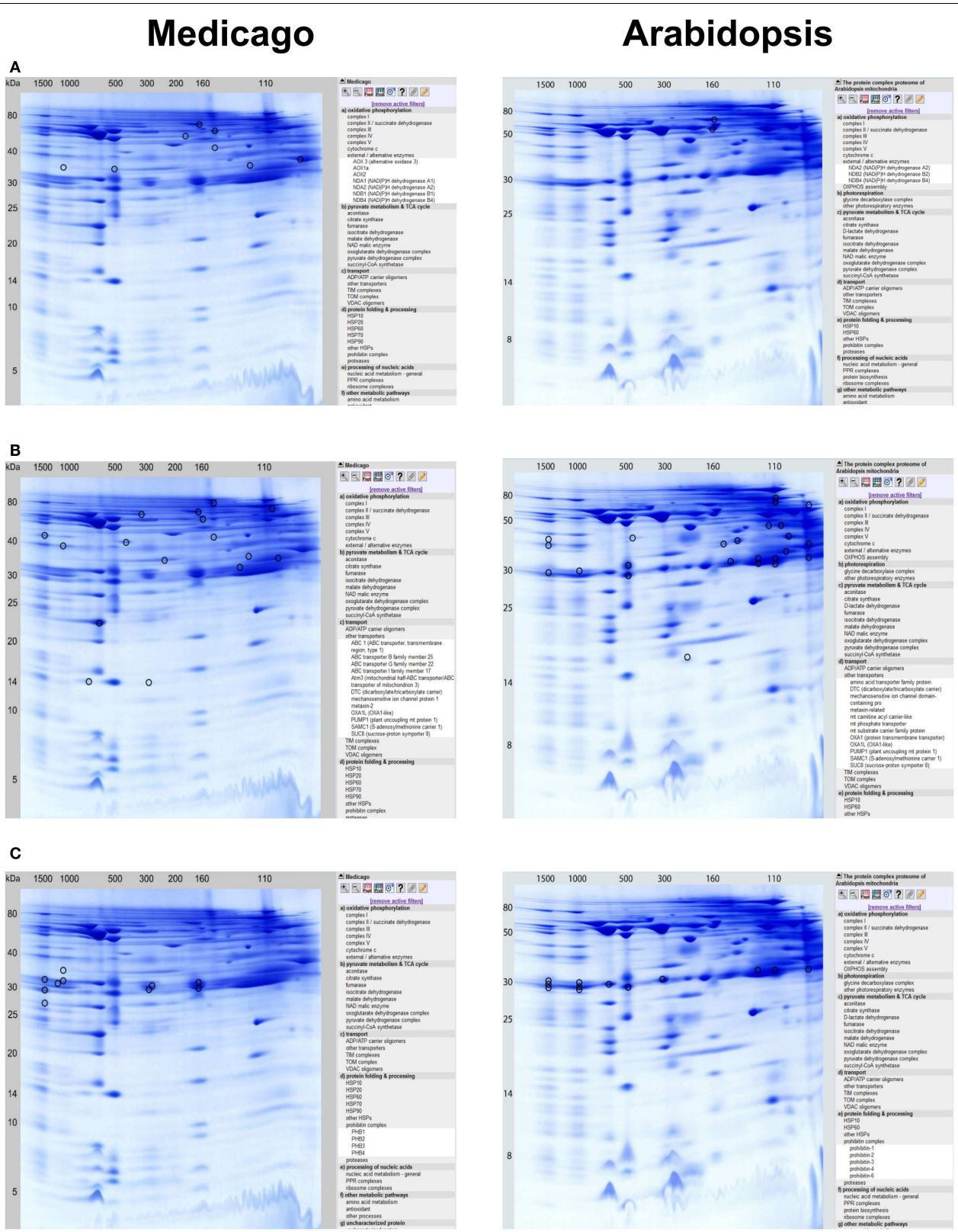


FIGURE 2 | Detailed protein complex visualization in the GelMap reference maps of the *Medicago* mitochondrial complexome (http://www.gelmap.de/medicago/, left row) as compared to the *Arabidopsis* mitochondrial complexome (http://www.gelmap.de/arabidopsis/, right row). Visualization of protein complexes and all of its included individual proteins is exemplarily done for: **(A)** external/alternative enzymes, **(B)** other transporters/ABC transporters, and for **(C)** prohibitin complexes.

adaptive changes in the protein complement of this sub-cellular compartment.

ACKNOWLEDGMENTS

The authors like to thank Michael Senkler, Institute for Plant Genetics, LUH Hannover, for assistance with the creation of the *Medicago truncatula* GelMap. We further thank Katrin Peters, Dagmar Lewejohann, and Haque Eshanuel, all from the Institute for Plant Genetics, LUH Hannover, for experimental assistance in the laboratory. We are grateful to Jennifer Klodmann for fruitful discussions and Holger Eubel, Institute for Plant Genetics, LUH Hannover, for critically reading of the manuscript. We acknowledge support by Deutsche Forschungsgemeinschaft and Open Access Publishing Fund of Leibniz University Hannover.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://www.frontiersin.org/Plant_Proteomics/10.3389/fpls.2013.00084/abstract

Figure S1 | Molecular mass scale for the 2D gel used for calibration and generation of *M. truncatula* mitochondria GelMap.

Figure S2 | Mitochondria protein complexes of *M. truncatula* resolved by 2D blue native/SDS PAGE.

Figure S3 | Spot detection on the 2D BN/SDS gel done automatically using DELTA 2D software package (version 4.3.2).

REFERENCES

Brugière, S., Kowalski, S., Ferro, M., Seigneurin-Berny, D., Miras, S., Salvi, D., et al. (2004). The hydrophobic proteome of mitochondrial membranes from *Arabidopsis* cell suspensions. *Phytochemistry* 65, 1693–1707.

Colditz, F., and Braun, H.-P. (2010). *Medicago truncatula* proteomics. *J. Proteomics* 73, 1974–1985.

Colditz, F., Niehaus, K., and Krajinski, F. (2007). Silencing of PR-10-like proteins in *Medicago truncatula* results in an antagonistic induction of other PR proteins and in an increased tolerance upon infection with the oomycete *Aphanomyces euteiches. Planta* 226, 57–71.

Daher, Z., Recorbet, G., Valot, B., Robert, F., Balliau, T., Potin, S., Schoefs, B., and Dumas-Gaudot, E. (2010). Proteomic analysis *of Medicago truncatula* root plastids. *Proteomics* 10, 2123–2137.

Dubinin, J., Braun, H.-P., Schmitz, U., and Colditz, F. (2011). The mitochondrial proteome of the model legume *Medicago truncatula. Biochim. Biophys. Acta* 1814, 1658–1668.

Heazlewood, J. L., Howell, K. A., and Millar, A. H. (2003a). Mitochondrial complex I from *Arabidopsis* and rice: orthologs of mammalian and fungal components coupled with plant-specific subunits. *Biochim. Biophys. Acta* 1604, 159–169.

Heazlewood, J. L., Howell, K. A., Whelan, J., and Millar, A. H. (2003b). Towards an analysis of the rice mitochondrial proteome. *Plant Physiol.* 132, 230–242.

Heazlewood, J. L., Verboom, R. E., Tonti-Filippini, J., Small, I., and Millar, A. H. (2007). SUBA: the *Arabidopsis* subcellular database. *Nucleic Acids Res.* 35, D213–D218.

Klodmann, J., and Braun, H.-P. (2011). Proteomic approach to characterize mitochondrial complex I from plants. *Phytochemistry* 72, 1071–1080.

Klodmann, J., Senkler, M., Rode, C., and Braun, H.-P. (2011). Defining the *protein complex proteome* of plant mitochondria. *Plant Physiol.* 157, 587–598.

Klodmann, J., Sunderhaus, S., Nimtz, M., Jänsch, L., and Braun, H.-P. (2010). Internal architecture of mitochondrial complex I from *Arabidopsis thaliana. Plant Cell* 22, 797–810.

Lannoo, N., and Van Damme, E. J. M. (2010). Nucleocytoplasmic plant lectins. *Biochim. Biophys. Acta* 1800, 190–201.

Lei, Z., Dai, X., Watson, B. S., Zhao, P. X., and Sumner, L. W. (2011). A legume specific protein database (LegProt) improves the number of identified peptides, confidence scores and overall protein identification success rates for legume proteomics. *Phytochemistry* 72, 1020–1027.

Meyer, E. H., Taylor, N. L., and Millar, A. H. (2008). Resolving and identifying protein components of plant mitochondrial respiratory complexes using three dimensions of gel electrophoresis. *J. Proteome Res.* 2, 786–794.

Peters, K., Nießen, M., Peterhänsel, C., Späth, B., Hölzle, A., Binder, S., Marchfelder, A., and Braun, H. P. (2012). Complex I–complex II ratio strongly differs in various organs of *Arabidopsis thaliana. Plant Mol. Biol.* 79, 273–284.

Rode, C., Senkler, M., Klodmann, J., Winkelmann, T., and Braun, H.-P. (2011). GelMap: a novel software tool for the creation and presentation of proteome reference maps. *J. Proteomics* 74, 2214–2219.

Schägger, H., and von Jagow, G. (1991). Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form. *Anal. Biochem.* 199, 223–231.

Schertl, P., Sunderhaus, S., Klodmann, J., Gergoff, G., Bartoli, C. G., and Braun, H. P. (2012). L-galactono-1,4-lactone dehydrogenase (GLDH) forms part of three subcomplexes of mitochondrial complex I in *Arabidopsis thaliana. J. Biol. Chem.* 287, 14412–14419.

Senkler, M., and Braun, H.-P. (2012). Functional annotation of 2D protein maps: the GelMap portal. *Front. Plant Sci.* 3:87. doi:10.3389/fpls.2012.00087

Taylor, N. L., Heazlewood, J. L., Day, D. A.,and Millar, A. H. (2005). Differential impact of environmental stresses on the pea mitochondrial proteome. *Mol. Cell Proteomics* 4, 1122–1133.

Trapphoff, T., Beutner, C., Niehaus, K., and Colditz, F. (2009). Induction of distinct defense-associated protein patterns in *Aphanomyces euteiches* (oomycota)-elicited and-inoculated *Medicago truncatula* cell-suspension cultures: a proteome and phosphoproteome approach. *Mol. Plant Microbe Interact.* 22, 421–436.

Van Aken, O., Whelan, J., and Van Breusegem, F. (2010). Prohibitins: mitochondrial partners in development and stress response. *Trends Plant Sci.* 15, 276–282.

Van Damme, E. J. M., Lannoo, N., Fouquaert, E., and Peumans, W. J. (2004). The identification of inducible cytoplasmic/nuclear carbohydrate-binding proteins urges to develop novel concepts about the role of plant lectins. *Glycoconj. J.* 20, 449–460.

Young, N. D., Debellé, F., Oldroyd, G. E. D., Geurts, R., Cannon, S. B., Udvardi, M. K., et al. (2011). The *Medicago* genome provides insight into the evolution of rhizobial symbiosis. *Nature* 480, 520–524.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 31 January 2013; paper pending published: 11 March 2013; accepted: 21 March 2013; published online: 15 April 2013.

Citation: Kiirika LM, Behrens C, Braun H-P and Colditz F (2013) The mitochondrial complexome of Medicago truncatula. Front. Plant Sci. ***4****:84. doi: 10.3389/fpls.2013.00084*

This article was submitted to Frontiers in Plant Proteomics, a specialty of Frontiers in Plant Science.

Copyright © 2013 Kiirika, Behrens, Braun and Colditz. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.

frontiers in
PLANT SCIENCE

MINI REVIEW ARTICLE
published: 13 March 2013
doi: 10.3389/fpls.2013.00051

Biochemistry, proteomics, and phosphoproteomics of plant mitochondria from non-photosynthetic cells

Jesper F. Havelund[1], Jay J. Thelen[2] and Ian M. Møller[1]*

[1] Department of Molecular Biology and Genetics, Science and Technology, Aarhus University, Slagelse, Denmark
[2] Department of Biochemistry and Interdisciplinary Plant Group, University of Missouri-Columbia, Columbia, MO, USA

Edited by:
Nicolas L. Taylor, The University of Western Australia, Australia

Reviewed by:
Chun Pong Lee, University of Oxford, UK
Carmela Giglione, Centre National de la Recherche Scientifique, France
Allan Rasmusson, Lund University, Sweden

****Correspondence:***
Ian M. Møller, Department of Molecular Biology and Genetics, Science and Technology, Aarhus University, Forsøgsvej 1, DK-4200 Slagelse, Denmark.
e-mail: ian.max.moller@agrsci.dk

Mitochondria fulfill some basic roles in all plant cells. They supply the cell with energy in the form of ATP and reducing equivalents [NAD(P)H] and they provide the cell with intermediates for a range of biosynthetic pathways. In addition to this, mitochondria contribute to a number of specialized functions depending on the tissue and cell type, as well as environmental conditions. We will here review the biochemistry and proteomics of mitochondria from non-green cells and organs, which differ from those of photosynthetic organs in a number of respects. We will briefly cover purification of mitochondria and general biochemical properties such as oxidative phosphorylation. We will then mention a few adaptive properties in response to water stress, seed maturation and germination, and the ability to function under hypoxic conditions. The discussion will mainly focus on Arabidopsis cell cultures, etiolated germinating rice seedlings and potato tubers as model plants. It will cover the general proteome as well as the posttranslational modification protein phosphorylation. To date 64 phosphorylated mitochondrial proteins with a total of 103 phosphorylation sites have been identified.

Keywords: plant mitochondria, proteomics, mitochondrial isolation, protein phosphorylation

INTRODUCTION

All living plant cells contain mitochondria, the organelle where organic molecules are oxidized to produce energy, in the form of ATP and reducing equivalents, as well as metabolic intermediates for use in biosynthetic processes. However, mitochondria have a number of other important functions depending on the cell and tissue type, as well as environmental conditions. In theory, each of these factors affects the properties of the mitochondria and the composition of its proteome. While this has been demonstrated at the biophysical level (Douce and Neuburger, 1989) and for some enzymes and complexes, we are entering an era where comprehensive protein profiling is possible to quantitatively assess mitochondrial remodeling in response to genetic and environmental cues. In anticipation and preparation for this époque we review critical aspects of the field including: (1) methods for isolation of ultra-pure mitochondria (for proteome interrogation); (2) properties of mitochondria (to characterize function); and (3) the current status of mitochondrial proteome investigations. Due to space constraints, we focus this review on mitochondria from non-photosynthetic tissues, which can be considered the "reference state" for comparative studies. These mitochondria perform the same basic functions and would therefore be expected to contain many of the same proteins as mitochondria from photosynthetic tissues.

PURIFICATION OF MITOCHONDRIA FROM NON-GREEN TISSUES

As is the case for all studies of organellar proteomics, a mitochondrial isolation procedure must remove all contaminants and give a clean preparation containing only mitochondria. Biochemical studies of mitochondrial metabolism often require several mg protein. Therefore, access to large amounts of starting material has been an important factor when selecting the species and tissue to study.

During the 1950–1970's, non-photosynthetic tissues such as etiolated seedlings or storage tissues were usually employed for the isolation of mitochondria by differential centrifugation probably because a major, visible contaminant—thylakoid membranes—was absent. Actually, crude mitochondria from storage tissues are generally contaminated by amyloplast membranes (Neuburger et al., 1982), but they do not interfere with measurements of oxygen consumption, a standard method in the study of mitochondrial metabolism. When rate-zonal, density gradient purification of plant mitochondria was introduced, it was initially applied to potato tubers. Step sucrose gradients were first introduced (Douce et al., 1972) and later Percoll gradients (Neuburger et al., 1982; Struglics et al., 1993; Considine et al., 2003), which had the great advantage over sucrose gradients that the mitochondria were not exposed to large changes in osmolarity. Not only does Percoll gradient centrifugation remove major contaminants like plastids and peroxisomes, it also removes mitochondria with damaged outer and/or inner membranes (Neuburger et al., 1982; Struglics et al., 1993).

Using Percoll gradient centrifugation, purified mitochondria have now been isolated from a range of non-green tissues from a number of species (Moreau and Romani, 1982; Liden and Møller, 1988; Fredlund et al., 1991; Lind et al., 1991; Millar et al., 2001; Bardel et al., 2002; Robson and Vanlerberghe, 2002; Qin et al.,

2009; Lee et al., 2011). This has usually given a very significant improvement in mitochondrial purity and intactness and a concomitant increase in rates of respiration, which indicates that the crude mitochondria previously used had sometimes contained less than 50% undamaged mitochondria, on a protein basis (e.g., Moreau and Romani, 1982). In the case of purified potato tuber mitochondria, contamination by plastid envelope and peroxisomes could be calculated to be <0.5% on a protein basis (Neuburger et al., 1982; Struglics et al., 1993). Percoll-purified mitochondria from Arabidopsis cell cultures, roots and shoots can be further purified by free-flow electrophoresis (Eubel et al., 2007; Lee et al., 2011). For cell culture mitochondria, this method decreases the contamination by peroxisomal and plastidic proteins by an estimated 5- to 10-fold (Eubel et al., 2007).

BIOCHEMICAL PROPERTIES OF MITOCHONDRIA FROM NON-GREEN TISSUES

BASIC PROPERTIES

Mitochondria are semi-autonomous, membrane-bound organelles. Consistent with their endosymbiotic origin, plant mitochondria contain small circular genomes. Varying with species, 30–40 proteins are encoded in the mitochondrial DNA, transcribed and translated in the matrix on bacterial-like ribosomes (Kubo and Newton, 2008). The remaining proteins making up the mitochondrial proteome are imported from the cytosol mainly through two large protein complexes, TIM and TOM (Translocator Inner/Outer Membrane) (Lister et al., 2005).

Isolated mitochondria from non-green cells and tissues oxidize a variety of substrates, including most of the tricarboxylic acid (TCA) cycle intermediates, with good coupling indicating that the TCA cycle is fully functional and that the four electron transport chain (ETC) complexes are present together with ATP synthase (Douce and Neuburger, 1989). In addition, the ETC of mitochondria from non-green tissues contain up to four alternative NAD(P)H dehydrogenases (DH) (Rasmusson et al., 2004), the alternative oxidase (Vanlerberghe and McIntosh, 1997), and has the uncoupling protein associated with it (Vercesi et al., 1995; Zhu et al., 2011). A range of DHs in the matrix feed electrons into the ETC at ubiquinone (UQ)—proline DH, lactate DH, and DHs involved in branched amino acid degradation via electron-transfer flavoprotein and electron-transfer flavoprotein: quinone oxidoreductase. On the outer surface of the inner membrane glycerol-3-phosphate DH also feeds electrons into the UQ pool, while L-galactono-1,4-lactone DH, the last enzyme in ascorbate biosynthesis, feed electrons into cytochrome c. The presence and amount of these enzymes in the mitochondria depends on species, tissue and environmental conditions (reviewed by Rasmusson et al., 2008; Rasmusson and Møller, 2011). To interact with other subcellular compartments within the cell, the inner mitochondrial membrane contains a number of transport proteins for the exchange of metabolites, coenzymes, etc. (Laloi, 1999; Palmieri et al., 2009).

Isolated mitochondria from green tissues do not differ much in the basic properties of the ETC and the TCA cycle. However, "green" mitochondria differ in the balance between the different substrates used. This is especially true for glycine, the product of photorespiration, which is oxidized at high rates by "green" mitochondria, but also their amino acid metabolism differs (e.g., Lee et al., 2008).

RESPONSES TO HYPOXIA, DROUGHT, AND DESICCATION

Plants can be exposed to hypoxic conditions as a result of flooding (Greenway et al., 2006). However, even under normoxic external conditions, the central parts of dense, metabolically active and/or bulky non-photosynthetic tissues can experience hypoxic conditions because the diffusion of oxygen into the tissue cannot keep up with the rate of removal by respiration (Geigenberger, 2003). This leads to metabolic changes such as the induction of ROS-degrading enzymes, presumably to prevent post-anoxic injury (Geigenberger, 2003), and higher activity of the enigmatic formate dehydrogenase (FDH) (Bykova et al., 2003b), one of the most abundant proteins in potato tuber mitochondria (Colas Des Francs-Small et al., 1993). The function of FDH may be to remove formate, but the source of formate is unknown and the reason the enzyme is so abundant is obscure (Igamberdiev, 1999; Ambard-Bretteville et al., 2003).

Drought presents the plant with quite a different problem and production of compatible solutes, such as proline, is one strategy to ameliorate the effects of a low water potential. As mitochondria are involved in proline turnover it is likely that homeostatic balance of this osmolyte requires the coordinated action of multiple organelles (Atkin and Macherel, 2009).

Mitochondria in maturing and germinating seeds experience quite unique conditions. The water content during the late-maturation phase is very low and a late-embryogenesis-abundant (LEA) protein is induced to protect mitochondrial membranes (Macherel et al., 2007; Tolleter et al., 2007). At the same time, the mitochondria presumably have to cope with hypoxic conditions caused by their own activity (Borisjuk and Rolletschek, 2009).

PROTEOMICS OF MITOCHONDRIA FROM NON-GREEN TISSUES

The most comprehensive proteomic studies in non-photosynthetic tissues have been performed on mitochondria from Arabidopsis suspension cells, a de-differentiated cell type that, when grown in darkness with sugar, is non-photosynthetic. For this reason and space considerations, this discussion of proteomics and post-translational modifications will focus mainly on Arabidopsis suspension cells and etiolated rice seedlings.

Bioinformatic analysis of the Arabidopsis genome, using primarily prediction algorithms, has estimated that mitochondria may contain as many as 2000–3000 different proteins (Millar et al., 2005, 2006; Cui et al., 2011). It is surprising then that nearly 20 years into the proteomics era no more than 500–600 different mitochondrial isolation proteins have been experimentally confirmed from any plant mitochondrial source! This is not due to limitations in technology, as it is now routine to not only identify, but quantify, at least 1900 proteins from a single biological sample using standard tandem mass spectrometry (e.g., Balbuena et al., 2012). We therefore expect the gap between the number of experimental and predicted mitochondrial proteins to shrink in the coming years.

The most comprehensive proteomic characterization of purified mitochondria from non-photosynthetic cells has been from

Arabidopsis suspension cells (Heazlewood et al., 2004). A total of 390 unique proteins were qualitatively identified by shotgun, reversed-phase (RP) LC-MS/MS analysis of 15–20 μg of tryptic peptides. Diverse groups of proteins were identified including proteins involved in energy and metabolism, DNA replication, transcription, translation, protein complex assembly, and signaling, as well as approximately 70 proteins of unknown function. Interestingly, various glycolytic enzymes were detected in mitochondrial preparations from Arabidopsis suspension cells (Heazlewood et al., 2004). Subsequent studies verified that these enzymes are indeed associated with mitochondria through scaffold proteins, such as the voltage-dependent anion channel (VDAC) located in the outer membrane (Graham et al., 2007). Additionally, *in vivo* association of cytosolic glycolytic enzymes with mitochondria is a dynamic process allowing for respiration to be supported in a dedicated manner consistent with substrate channeling (Graham et al., 2007). Although many of the individual subunits of the five respiratory complexes were identified by Heazlewood et al. (2004), follow-up proteomic analysis of these complexes was necessary to reveal the full protein complements (Meyer et al., 2008; Klodmann and Braun, 2011). In a study focusing on mitochondrial protein complexes, (Klodmann et al., 2011) identified 471 non-redundant proteins belonging to at least 35 different protein complexes. Subsequent proteomic analysis of purified outer membranes also expanded the compendium of proteins mapped to these mitochondrial subcompartment especially integral membrane proteins (Duncan et al., 2011; Tan et al., 2012).

Besides Arabidopsis roots and suspension cells, mitochondria from etiolated rice shoots have also been characterized at the proteome level (Bardel et al., 2002; Huang et al., 2009). Using both 2D gel electrophoresis and RP-LC-MS/MS analyses a total of 322 non-redundant proteins were identified in a non-quantitative manner. Comparison with Arabidopsis cell culture mitochondria (Heazlewood et al., 2004) revealed a similar cohort of proteins, although 20% of the rice mitochondrial proteins did not produce orthologous matches to the Arabidopsis mitochondrial proteome (Huang et al., 2009). And like the Arabidopsis experimental proteome, approximately 60% of rice mitochondrial proteins were predicted to be targeted to this organelle using various organelle prediction algorithms. Despite the wealth of shotgun and targeted mitochondrial proteomic studies it is obvious that a comprehensive compendium of experimentally-verified plant mitochondrial proteins is currently unavailable. This is apparent when one not only considers the gap between predicted and experimentally identified proteins, but also from perusing the current catalogue of approximately 500 mapped proteins. Missing from this list are many essential plant mitochondrial activities including regulatory proteins, transcription factors, metabolite translocators, and the wealth of tRNA synthases and pentratricopeptide repeat proteins. For example, the two regulatory enzymes of the pyruvate dehydrogenase complex (PDC)—the PDC kinase and phospho-PDC phosphatase, which are activities in mitochondria from both green and non-green tissues (Thelen et al., 1998a,b)—are also undetected in global plant mitochondrial proteomic studies.

While the early stages of non-photosynthetic mitochondrial proteome characterization dealt with protein cataloguing, more recent research has shifted toward comparative analyses to discover dynamic changes in mitochondrial protein expression. Proteomic comparisons of mitochondria from non-photosynthetic Arabidopsis suspension cells and developing photosynthetic shoots (Lee et al., 2008) as well as developing Arabidopsis roots and photosynthetic shoots (Lee et al., 2011) collectively revealed differences in TCA cycle and photorespiratory enzymes. In both instances, most of the component enzymes of the glycine decarboxylase complex, except for the lipoamide dehydrogenase (E3) subunit, which is shared with the PDC, were highly upregulated in green shoot mitochondria compared to the two non-photosynthetic counterparts. Additionally, FDH was highly induced in mitochondria from green shoots compared to roots and suspension cells. In contrast, three subunits of the PDC complex including E3 were more prominently expressed in root and suspension cell mitochondria. These and other changes to TCA and amino acid metabolism between mitochondria from green and non-green sources confirmed previous observations (e.g., upregulation of GDC in photorespiring tissues) but also suggested previously unknown differences in both carbon import/export and oxidative metabolism for this organelle as a direct result of photosynthetic capacity.

Additional comparative proteomic studies have been performed with mitochondria from etiolated rice shoots, analyzing the effect of a hypoxic/anoxic environment (Millar et al., 2004; Howell et al., 2007). In the absence of oxygen, mitochondrial respiration was impaired due to lower activity and expression of respiratory complexes cytochrome bc_1 and cytochrome c oxidase (Millar et al., 2004). Additionally, the E1β subunit of the PDC and a putative succinyl-CoA ligase (GDP-forming) β-chain as well as mitochondrial processing peptidase α- and β-chain subunits (a component of the cytochrome bc_1 complex) were reduced in expression (Howell et al., 2007). In contrast, a TIM subunit was highly induced under anaerobic conditions. It was concluded that under anoxic conditions a direct link between respiratory capacity and protein import could be established at the cytochrome bc_1 complex of the ETC. While total proteome coverage was not attained in either of these studies, the results illustrate the potential for comparative proteomics to elucidate the dynamic properties of mitochondria.

PROTEIN PHOSPHORYLATION IN PLANT MITOCHONDRIA

Several hundred posttranslational modifications (PTMs)—the covalent addition of a chemical group to amino acids in proteins–are known. They often lead to alterations in properties and function of the proteins and are therefore involved in the regulation of large variety of important biological processes (Wold, 1981; Mann and Jensen, 2003). The most well-studied PTM is phosphorylation, which is catalyzed by protein kinases (PKAs) while the dephosphorylation is catalyzed by protein phosphatases. This modification primarily targets the hydroxyl groups of Ser, Thr, and Tyr residues and has long been known to be a key player in signaling. Both protein phosphorylation and other PTMs such as acetylation are important in mammalian mitochondria (Guan and Xiong, 2011; Koc and Koc, 2012), but in plant mitochondria only protein phosphorylation has been studied in any detail.

Table 1 | Phosphoproteins and phosphosites in plant mitochondria.

No	Protein	Plant species	Phosphorylation site	References
ENERGY				
1	Pyruvate dehydrogenase (EC 1.2.4.1), E1α-subunit	Potato	Ser^{298}: YHGH**(S)**MSDP	1, 3, 4[a]
		Arabidopsis	Ser^{292}: YHGH**(S)**MSDP	5[a], 10[a], 9[a], 11[a]
		Pea		14
		Rice	Ser^{110}: AITR**(S)**D(S)II	12[a]
			Ser^{112}: TR(S)D**(S)**IITA	
			Ser^{293}: YHGH**(S)**M(S)D	
			Ser^{295}: GH(S)M**(S)**DPGS	
2	Aconitate hydratase (EC 4.2.1.3)	Potato		3
3	NAD-isocitrate dehydrogenase (EC 1.1.1.41)	Potato		3
4	Succinyl-CoA-ligase (EC 6.2.1.5), α-subunit	Potato		3
		Arabidopsis	Ser^{304}: GAIV**(S)**GGKG	9[a], 12[a]
		Rice	Ser^{289}: GAIV**(S)**GGKG	12[a]
			Ser^{312}: TVVE**(S)**PAKI	
5	Succinyl-CoA-ligase (EC 6.2.1.5), β-subunit	Potato		3
6	Succinate dehydrogenase (EC 1.3.5.1), flavoprotein	Potato		3
		Arabidopsis		5
		Rice	Ser^{161}: FGGQ**(S)**LDFG	12[a]
7	Succinate dehydrogenase (EC 1.3.5.1), subunit 5	Arabidopsis		5
8	Malate dehydrogenase (EC 1.1.1.37)	Potato		3
		Arabidopsis	Ser^{23}: RRSF**(S)(**S)GSV	9[a]
			Ser^{24}: RSF(S)**(S)**GSVP	
			Ser^{314}: LGPL**(S)**DFEK	
		Rice	Ser^{313}: LGQL**(S)**DFEK	12[a]
9	NAD-malic enzyme (EC 1.1.1.38), 62 kDa isoform	Potato		3
10	NAD-malic enzyme (EC 1.1.1.38), 59 kDa isoform	Potato		3
11	Protein NADH-ubiquinone oxidoreductase (EC 1.6.5.3), 39 kDa subunit	Rice	Ser^{397}: PAFG**(S)**(T)V(S)E	12[a]
			Thr^{398}: AFG(S)**(T)**V(S)EK	
			Ser^{400}: G(S)(T)V**(S)**EKIR	
12	Complex III (EC 1.10.2.2), β-mitochondria processing peptidase subunit	Potato		3
		Rice	Ser^{485}: SWFR**(S)**HTY(S)	
			Ser^{489}: (S)HTY**(S)**DDEF	12[a]
13	Protein mitochondrial-processing peptidase (EC 3.4.24.64), α-subunit	Rice	Ser^{234}: HRLD**(S)**(S)ILE	12[a]
			Ser^{235}: RLD(S)**(S)**ILEE	
14	F_1 ATPase (EC 3.6.1.3), δ-subunit	Potato		2
15	ATP synthase (EC 3.6.1.3), α-subunit	Potato		3
16	ATP synthase (EC 3.6.1.3), β-subunit	Rice	Ser^{355}: GRIP**(S)**AVGY	12[a]
			Thr^{362}: GYQP**(T)**LATD	
			Ser^{418}: SRQI**(S)**ELGI	
17	F_0- ATPase (EC 3.6.1.3), b/C/D-subunit	Potato		2
		Rice	Ser^{157}: KAVD**(S)**LVPI	12[a]
		Arabidopsis	Ser^{77}: VENK**(S)**QGSE	9[a]
			Ser^{80}: K(S)QG**(S)**EVLL	
			Ser^{386}: EKSE**(S)**LEEI	
		Arabidopsis	Ser^{248}: QRGV**(T)**LGDP	10[a]

(Continued)

Table 1 | Continued

No	Protein	Plant species	Phosphorylation site	References
18	Formate DH (EC 1.2.1.2)	Potato	Thr^{76}: QYIV**(T)**PDKE Thr^{333}: NQAM**(T)**PHIS	3, 4[a]
		Arabidopsis		5
19	Glutamate DH 1 (EC 1.4.1.2)	Arabidopsis		5
20	Glutamate DH 2 (EC 1.4.1.2)	Arabidopsis		5
21	Mitochondrial lipoamide dehydrogenase 1 (EC 1.8.1.4)	Arabidopsis	Ser^{319}: TPFT**(S)**GLDL	9[a], 12[a]
METABOLISM				
22	Protein cysteine synthase (EC 2.5.1.47)	Rice	Ser^{130}: PAGY**(S)**LDKQ	12[a]
PROTEIN FATE				
23	Heat shock protein 90	Potato		3
		Arabidopsis	Ser^{219}: EKEI**(S)**DDEE	5[a], 9[a], 10[a], 11[a]
24	Heat shock protein 70	Potato		3
		Bean		6
		Rice	Ser^{419}: GLSE**(S)**DIEK Ser^{526}: GSSS**(S)**(S)GGD Ser^{527}: SSS(S)**(S)**GGDQ	12[a]
	Heat shock protein 70-1	Arabidopsis	Ser^{398}: QEIV**(S)**EIFGK Ser^{404}: IFGK**(S)**PCKG Ser^{552}: SGGL**(S)**DDEI	9[a]
	Heat shock protein 70-2	Arabidopsis	Ser^{409}: IFGK**(S)**PSKG	12[a]
25	Chaperonin 60	Potato		3
	60 kDa chaperonin 7	Rice	Ser^{137}: QSYG**(S)**PKV Thr^{522}: DKLP**(T)**ANFD	12[a]
26	Mitochondrial small heat shock protein	Potato		3
	Mitochondrial small heat shock protein, 22 kDa	Maize		7
		Rice	Ser^{58}: GRLL**(S)**LMDD Ser^{101}: ERDE**(S)**DDDS	12[a]
27	26.2 kDa heat shock protein	Rice	Tyr^{71}: PLRR**(Y)**DVVD Ser^{77}: VVDE**(S)**G(T)D(S) Thr^{79}: DE(S)G**(T)**D(S)GD Ser^{81}: (S)G(T)D**(S)**GDE(Y) Tyr^{85}: (S)GDE**(Y)**DA(T)D Thr^{88}: E(Y)DA**(T)**DDGR	12[a]
28	GRPE co-chaperone	Arabidopsis		5
29	Protein GrpE 1	Rice	Ser^{81}: SPEL**(S)**DKEE	12[a]
30	Lon1 protease (EC 3.4.21.53)	Arabidopsis	Ser^{74}: KAVE**(S)**DSEV	5[a], 9[a]
TRANSCRIPTION				
31	Mitochondrial transcription termination factor family protein [At1g62010]	Arabidopsis	Ser^{53}: FKSS**(S)**FLDS	12[a]
32	Transcription termination factor domain-containing protein [AT2G44020]	Arabidopsis	Thr^{488}: GEIV**(T)**DEEE Ser^{495}: EEDE**(S)**DDEV	12[a]

(Continued)

Table 1 | Continued

No	Protein	Plant species	Phosphorylation site	References
33	Mitochondrial transcription termination factor family protein [At2g21710]	Arabidopsis	Tyr[501]: DEVL**(Y)**RRTL Tyr[168]: FGIT**(Y)**A(T)NV Thr[170]: IT(Y)A**(T)**NVTD	9[a]
RNA PROCESSING				
34	Putative mitochondrial RNA helicase (EC 3.6.4.13), helicase 2	Arabidopsis	Ser[603]: GGRS**(S)**FGGF	12[a]
35	Protein ATP-dependent RNA helicase (EC 3.6.4.13), SUV3	Rice	Ser[693]: VEQA**(S)**DDNA	12[a]
PROTEIN SYNTHESIS				
36	Glycine rich RNA-binding protein 3	Arabidopsis		5
37	Ribosomal protein S24/S35	Arabidopsis	Thr[117]: KHAE**(T)**DDEL	5[a], 9[a], 10[a], 12[a]
38	Protein glycyl-tRNA synthetase 2 (EC 6.1.1.14)	Rice	Ser[721]: QKKL**(S)**EFRD	12[a]
39	Protein glycyl-tRNA synthetase 1 (EC 6.1.1.14)	Rice	Ser[671]: EEEA**(S)**E(T).. Thr[673]: EA(S)E**(T)**	12[a]
40	Protein rf1 protein	Rice	Ser[361]: DLMV**(S)**RGCY	12[a]
DEFENSE, STRESS, DETOXIFICATION				
41	Mn superoxide dismutase (1.15.1.1)	Potato		3
COMMUNICATION/SIGNALING				
42	MAM protein	Arabidopsis	Ser[94]: DNID**(S)**DEEM	5
43	MAM 33	Arabidopsis		5
44	Protein calcium-binding mitochondrial protein Anon-60Da	Rice	Ser[609]: ASAS**(S)**VSKE	12[a]
45	Protein protein-tyrosine phosphatase 1 (EC 3.1.3.48)	Rice	Ser[269]: CLIP**(S)**LK..	12[a]
TRANSPORT				
46	Mitochondrial glycoprotein family protein [AT4G32605.1]	Arabidopsis	Ser[76]: SLEG**(S)(**T)GAV Thr[77]: LEG(S)**(T)**GAVL	9[a]
47	Mitochondrial glycoprotein family protein [AT2G39795]	Arabidopsis	Ser[84]: DNID**(S)**DEEM	5[a], 12[a]
48	Protein mitochondrial glycoprotein, expressed [BAD33248.1]	Rice	Ser[88]: RVID**(S)**EIEC Ser[96]: CIVQ**(S)**EEGA	12[a]
49	Protein mitochondrial glycoprotein, expressed [A2ZP51.1]	Rice	Ser[81]: FAEE**(S)**DDHD	12[a]
50	Protein mitochondrial glycoprotein, expressed [BAF25817.1]	Rice	Ser[115]: RRVE**(S)**LERG	12[a]
51	Mitochondrial substrate carrier family protein	Arabidopsis	Ser[64]: RGNN**(S)**FSTQ	12[a]
52	Mitochondrial import inner membrane translocase, subunit TIM17/TIM22/TIM23	Arabidopsis	Ser[15]: DPSS**(S)**PPPI	12[a]
53	Mitochondrial import inner membrane translocase, subunit TIM23	Rice	Ser[8]: PRLF**(S)(**S)G(S)G Ser[9]: RLF(S)**(S)**G(S)GS Ser[11]: F(S)(S)G**(S)**GTRD	12[a]
		Arabidopsis	Ser[126]: AGIE**(S)**GVVA Ser[140]: DVWT**(S)**VVAG	13[a]

(Continued)

Table 1 | Continued

No	Protein	Plant species	Phosphorylation site	References
54	Translocase of outer membrane, subunit TOM20-4	Arabidopsis	Ser^{143}: AGPS**(S)**NSAK	9[a]
55	AMP deaminase (EC 3.5.4.6)	Arabidopsis	Tyr^{70}: VNDQ**(Y)**GR(S)P	12[a]
			Ser^{73}: QYGR**(S)**PASL	9[a], 12[a]
			Ser^{76}: R(S)PA**(S)**LPDA	12[a]
			Ser^{140}: VRPI**(S)**PKSP	9[a], 10[a],12[a]
			Ser^{203}: IRSH**(S)**VSGD	5[a], 12[a]
			Ser^{279}: ISDP**(S)**(T)PKP	12[a]
			Thr^{280}: SDP(S)**(T)**PKPN	12[a]
		Rice	Ser^{94}: KPII**(S)**PAST	12[a]
			Ser^{126}: TIED**(S)**DDDD	
			Ser^{178}: IRSQ**(S)**ATG(S)	
			Ser^{182}: (S)ATG**(S)**LHGA	
			Thr^{256}: TDPS**(T)**PKPN	
56	Mitochondrial outer membrane protein porin 5	Rice	Ser^{96}: KLVT**(S)**VKLP	12[a]
57	Mitochondrial outer membrane protein porin 6	Rice	Ser^{95}: KTSF**(S)**FRVP	12[a]
58	Protein mitochondrial 2-oxoglutarate/malate carrier protein	Rice	Ser^{97}: ARLG**(S)**FRVL	12[a]
59	Protein mitochondrial carrier protein	Rice	Ser^{419}: ARAN**(S)**IEVS	12[a]
60	Mitochondrial carrier protein	Arabidopsis	Ser^{155}: AAPE**(S)**PSSN	5[a]
61	TraB family protein	Arabidopsis	Ser^{64}: DSVF**(S)**GGDG	12[a]
STRUCTURE ORGANIZATION				
62	Protein inner membrane protein OXA1L	Rice	Ser^{422}: SQRF**(S)**DLEN	12[a]
MISCELLANEOUS				
63	Stomatin-like protein	Arabidopsis	Thr^{362}: ALDE**(T)**DLEE	5[a]
64	Prohibitin	Rice		8

This list has been compiled by first manually going through the phosphoproteomic literature on non-photosynthetic plant organisms published before 2008. Secondly, the identified proteins were used as search input in the Plant Protein Phosphorylation Database and Yao et al. (2012)—http://p3db.org/- resulting new phosphosites were included in the list. Lastly, the mitochondrial proteins from Duncan et al. (2011) were searched against P^3DB and the hits included. The blank spaces in the column labeled Phosphorylation site indicate that the phosphoprotein, but not the specific site of phosphorylation was identified.

All phosphorylation sites are highlighted with parentheses. The specific phosphorylation site is highlighted with bold.

[a] *This reference includes phosphosite information.*

1, (Sommarin et al., 1990); 2, (Struglics et al., 1998); 3, (Bykova et al., 2003a); 4, (Bykova et al., 2003b); 5, (Ito et al., 2009); 6, (Vidal et al., 1993); 7, (Lund et al., 2001); 8, (Takahashi et al., 2003); 9, (Sugiyama et al., 2008); 10, (Li et al., 2009); 11, (de la Fuente van Bentem et al., 2008); 12, (Nakagami et al., 2010); 13, (Meyer et al., 2012); 14, (Budde and Randall, 1990).

Traditionally, the methods for detecting phosphoproteins are based on incorporation of radioactive phosphate (^{32}P). This is either done by *in vitro* phosphorylation reactions using protein extracts or isolated organelles or by a more physiologically relevant *in vivo* approach using cells. The phosphorylated proteins are then typically separated by gel electrophoresis and detected by phosphoimaging or autoradiography (Thelen et al., 2000; Bykova et al., 2003a; Berwick and Tavare, 2004). A fluorescent phosphosensor dye called Pro-Q Diamond has also been developed (Schulenberg et al., 2003; Ito et al., 2009). The phosphoproteins can be cut out of the gel and identified by LC-MS/MS (Pappin et al., 1993; Heazlewood et al., 2004; Ito et al., 2009). However, more recently shotgun methods of identifying phosphopeptides and phosphoproteins first uses a phosphopeptide enrichment method such as immobilized metal affinity chromatography (IMAC) (Posewitz and Tempst, 1999; Stensballe et al., 2001) or titanium oxide chromatography (MOC) followed by LC-MS/MS. In this way thousands of phosphorylation sites can be identified without the use of gel electrophoresis (e.g., Engholm-Keller et al., 2012; Meyer et al., 2012).

To date, 64 plant mitochondrial proteins have been reported to be phosphorylated on 103 phosphosites (**Table 1**). The distribution of these on Ser, Thr, and Tyr residues is 81, 14, and 5%, respectively, or close to the distribution in all the known 22,995 eukaryotic phosphoproteins with 100,281 phosphosites (Rao and Møller, 2012) in spite of the fact that plants do not contain canonical Tyr kinases (Yao et al., 2012).

The largest group of phosphorylated proteins is energy and transport (**Table 1**), a trend also observed in yeast (Reinders et al., 2007). Phosphorylation sites have been identified on most of the TCA cycle enzymes, and on complexes III–V, whereas no phosphoproteins have been found in Complexes I and II. It seems likely that oxidative phosphorylation is regulated by reversible protein phosphorylation. Other major processes that may be regulated by phosphorylation are transcription, translation, protein folding, and metabolite transport across the inner mitochondrial membrane. In mammalian mitochondria, translation is regulated by protein phosphorylation (Koc and Koc, 2012).

PDC, which produces the acetyl-CoA entering the TCA cycle, is the only enzyme in plant mitochondria where the detailed regulation by reversible phosphorylation/dephosphorylation has been elucidated (Rubin and Randall, 1977; Budde and Randall, 1990). When PDC products accumulate, they activate the PDC kinase, which phosphorylates and inactivates the PDH component of PDC (Tovar-Mendez et al., 2003). In contrast, PDC substrates inhibit the PDC kinase and stimulate the phospho-PDC phosphatase to dephosphorylate PDH and give a higher activity (Randall et al., 1981; Schuller and Randall, 1990; Moore et al., 1993). FDH phosphorylation appears to be regulated in a similar way by some of the same metabolites, but the effect on FDH activity is still unknown (Bykova et al., 2003b).

In yeast, phosphorylation of TOM22, Mim1, and TOM79 by casein kinase 2 (CK2) and PKA have been shown to regulate the function of TOM. CK2 promotes biogenesis of specific TOM complexes and CK2 inhibits specific TOM receptor activities and thereby the import of mitochondrial metabolite carriers. Together these phosphorylations regulate protein homeostasis (Schmidt et al., 2011). Another study has shown that phosphorylation of subunit *g* of the yeast ATP synthase subunit F_o inhibits dimerization of the ATP synthase and is thus involved in the regulation of the bioenergetic state of mitochondria (Reinders et al., 2007). Phosphorylated TOM and ATP synthase F_o subunits have also been found in plant mitochondria from non-photosynthetic cells (**Table 1**) and the function of these phosphorylations might be similar to that in yeast. Finally, the interaction of chaperone HSP90 with co-chaperones is regulated by phosphorylation in mammalian mitochondria (Mollapour et al., 2011; Johnson, 2012). HSP90 is one of the identified phosphoproteins in plant mitochondria (**Table 1**), but the site of phosphorylation so far identified is different from that of mammalian HSP90 (Mollapour et al., 2011), so the regulatory mechanism may not be the same.

Considering that plant mitochondria contain so many phosphoproteins, we might expect to find a substantial number of PKAs in plant mitochondria as well as a somewhat smaller number of protein phosphatases (Juszczuk et al., 2007). However, even in the most extensive proteomic studies published to date only ten protein kinases have been identified and no protein phosphatases (Heazlewood et al., 2004; Duncan et al., 2011; Taylor et al., 2011). This may indicate that each enzyme is present in very few copies so that they are below the technical detection limit. Alternatively there are actually relatively few kinases and phosphatases each with a relatively broad specificity. Future in depth proteomic studies will no doubt give us more information about such regulatory pathways.

ACKNOWLEDGMENTS

This study was supported by grants from the Danish Council for Independent Research—Natural Sciences (to Ian M. Møller) and a 2012 Sabbatical fellowship (to Jay J. Thelen) from the OECD Cooperative Research Programme: Biological Resource Management for Sustainable Agricultural Systems.

REFERENCES

Ambard-Bretteville, F., Sorin, C., Rebeille, F., Hourton-Cabassa, C., and Colas Des Francs-Small, C. (2003). Repression of formate dehydrogenase in *Solanum tuberosum* increases steady-state levels of formate and accelerates the accumulation of proline in response to osmotic stress. *Plant Mol. Biol.* 52, 1153–1168.

Atkin, O. K., and Macherel, D. (2009). The crucial role of plant mitochondria in orchestrating drought tolerance. *Ann. Bot.* 103, 581–597.

Balbuena, T. S., He, R., Salvato, F., Gang, D. R., and Thelen, J. J. (2012). Large-scale proteome comparative analysis of developing rhizomes of the ancient vascular plant equisetum hyemale. *Front. Plant Sci.* 3:131. doi: 10.3389/fpls.2012.00131

Bardel, J., Louwagie, M., Jaquinod, M., Jourdain, A., Luche, S., Rabilloud, T., et al. (2002). A survey of the plant mitochondrial proteome in relation to development. *Proteomics* 2, 880–898.

Berwick, D. C., and Tavare, J. M. (2004). Identifying protein kinase substrates: hunting for the organ-grinder's monkeys. *Trends Biochem. Sci.* 29, 227–232.

Borisjuk, L., and Rolletschek, H. (2009). The oxygen status of the developing seed. *New Phytol.* 182, 17–30.

Budde, R. J., and Randall, D. D. (1990). Pea leaf mitochondrial pyruvate dehydrogenase complex is inactivated *in vivo* in a light-dependent manner. *Proc. Natl. Acad. Sci. U.S.A.* 87, 673–676.

Bykova, N. V., Egsgaard, H., and Møller, I. M. (2003a). Identification of 14 new phosphoproteins involved in important plant mitochondrial processes. *FEBS Lett.* 540, 141–146.

Bykova, N. V., Stensballe, A., Egsgaard, H., Jensen, O. N., and Møller, I. M. (2003b). Phosphorylation of formate dehydrogenase in potato tuber mitochondria. *J. Biol. Chem.* 278, 26021–26030.

Colas Des Francs-Small, C., Ambard-Bretteville, F., Small, I. D., and Remy, R. (1993). Identification of a major soluble-protein in mitochondria from nonphotosynthetic tissues as nad-dependent formate dehydrogenase. *Plant Physiol.* 102, 1171–1177.

Considine, M. J., Goodman, M., Echtay, K. S., Laloi, M., Whelan, J., Brand, M. D., et al. (2003). Superoxide stimulates a proton leak in potato mitochondria that is related to the activity of uncoupling protein. *J. Biol. Chem.* 278, 22298–22302.

Cui, J. A., Liu, J. H., Li, Y. H., and Shi, T. L. (2011). Integrative identification of arabidopsis mitochondrial proteome and its function exploitation through protein interaction network. *PLoS ONE* 6:e16022. doi: 10.1371/journal.pone.0016022

de la Fuente van Bentem, S., Anrather, D., Dohnal, I., Roitinger, E., Csaszar, E., Joore, J., et al. (2008). Site-specific phosphorylation profiling of Arabidopsis proteins by mass spectrometry and peptide chip analysis. *J. Proteom. Res.* 7, 2458–2470.

Douce, R., Bonner, W. D., and Christen, E. L. (1972). Preparation of intact plant mitochondria. *Biochim. Biophys. Acta* 275, 148–160.

Douce, R., and Neuburger, M. (1989). The uniqueness of plant-mitochondria. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 40, 371–414.

Duncan, O., Taylor, N. L., Carrie, C., Eubel, H., Kubiszewski-Jakubiak, S., Zhang, B. T., et al. (2011). Multiple lines of evidence localize signaling, morphology, and lipid biosynthesis machinery to the mitochondrial outer membrane of arabidopsis. *Plant Physiol.* 157, 1093–1113.

Engholm-Keller, K., Birck, P., Storling, J., Pociot, F., Mandrup-Poulsen, T., and Larsen, M. R. (2012). TiSH - a robust and sensitive global phosphoproteomics strategy employing a combination of TiO(2), SIMAC, and HILIC. *J. Proteomics* 75, 5749–5761.

Eubel, H., Lee, C. P., Kuo, J., Meyer, E. H., Taylor, N. L., and Millar, A.

H. (2007). Free-flow electrophoresis for purification of plant mitochondria by surface charge. *Plant J.* 52, 583–594.

Fredlund, K. M., Rasmusson, A. G., and Møller, I. M. (1991). Oxidation of external Nad(P)H by purified mitochondria from fresh and aged red beetroots (Beta-Vulgaris L). *Plant Physiol.* 97, 99–103.

Geigenberger, P. (2003). Response of plant metabolism to too little oxygen. *Curr. Opin. Plant Biol.* 6, 247–256.

Graham, J. W. A., Williams, T. C. R., Morgan, M., Fernie, A. R., Ratcliffe, R. G., and Sweetlove, L. J. (2007). Glycolytic enzymes associate dynamically with mitochondria in response to respiratory demand and support substrate channeling. *Plant Cell* 19, 3723–3738.

Greenway, H., Armstrong, W., and Colmer, T. D. (2006). Conditions leading to high CO_2 (>5 kPa) in waterlogged-flooded soils and possible effects on root growth and metabolism. *Ann. Bot.* 98, 9–32.

Guan, K. L., and Xiong, Y. (2011). Regulation of intermediary metabolism by protein acetylation. *Trends Biochem. Sci.* 36, 108–116.

Heazlewood, J. L., Tonti-Filippini, J. S., Gout, A. M., Day, D. A., Whelan, J., and Millar, A. H. (2004). Experimental analysis of the Arabidopsis mitochondrial proteome highlights signaling and regulatory components, provides assessment of targeting prediction programs, and indicates plant-specific mitochondrial proteins. *Plant Cell* 16, 241–256.

Howell, K. A., Cheng, K., Murcha, M. W., Jenkin, L. E., Millar, A. H., and Whelan, J. (2007). Oxygen initiation of respiration and mitochondrial biogenesis in rice. *J. Biol. Chem.* 282, 15619–15631.

Huang, S., Taylor, N. L., Narsai, R., Eubel, H., Whelan, J., and Millar, A. H. (2009). Experimental analysis of the rice mitochondrial proteome, its biogenesis, and heterogeneity. *Plant Physiol.* 149, 719–734.

Igamberdiev, A. U. (1999). Foundations of metabolic organization: coherence as a basis of computational properties in metabolic networks. *Biosystems* 50, 1–16.

Ito, J., Taylor, N. L., Castleden, I., Weckwerth, W., Millar, A. H., and Heazlewood, J. L. (2009). A survey of the *Arabidopsis thaliana* mitochondrial phosphoproteome. *Proteomics* 9, 4229–4240.

Johnson, J. L. (2012). Evolution and function of diverse Hsp90 homologs and cochaperone proteins. *Biochim. Biophys. Acta* 1823, 607–613.

Juszczuk, I. M., Bykova, N. V., and Møller, I. M. (2007). Protein phosphorylation in plant mitochondria. *Physiol. Plant.* 129, 90–103.

Klodmann, J., and Braun, H. P. (2011). Proteomic approach to characterize mitochondrial complex I from plants. *Phytochemistry* 72, 1071–1080.

Klodmann, J., Senkler, M., Rode, C., and Braun, H. P. (2011). Defining the protein complex proteome of plant mitochondria. *Plant Physiol.* 157, 587–598.

Koc, E. C., and Koc, H. (2012). Regulation of mammalian mitochondrial translation by post-translational modifications. *Biochim. Biophys. Acta* 1819, 1055–1066.

Kubo, T., and Newton, K. J. (2008). Angiosperm mitochondrial genomes and mutations. *Mitochondrion* 8, 5–14.

Laloi, M. (1999). Plant mitochondrial carriers: an overview. *Cell. Mol. Life Sci.* 56, 918–944.

Lee, C. P., Eubel, H., O'Toole, N., and Millar, A. H. (2008). Heterogeneity of the mitochondrial proteome for photosynthetic and non-photosynthetic Arabidopsis metabolism. *Mol. Cell. Proteom.* 7, 1297–1316.

Lee, C. P., Eubel, H., O'Toole, N., and Millar, A. H. (2011). Combining proteomics of root and shoot mitochondria and transcript analysis to define constitutive and variable components in plant mitochondria. *Phytochemistry* 72, 1092–1108.

Li, H., Wong, W. S., Zhu, L., Guo, H. W., Ecker, J., and Li, N. (2009). Phosphoproteomic analysis of ethylene-regulated protein phosphorylation in etiolated seedlings of Arabidopsis mutant ein2 using two-dimensional separations coupled with a hybrid quadrupole time-of-flight mass spectrometer. *Proteomics* 9, 1646–1661.

Liden, A. C., and Møller, I. M. (1988). Purification, characterization and storage of mitochondria from *Jerusalem artichoke* tubers. *Physiol. Plant.* 72, 265–270.

Lind, C., Hallden, C., and Møller, I. M. (1991). Protein-synthesis in mitochondria purified from roots, leaves and flowers of sugar-beet. *Physiol. Plant.* 83, 7–16.

Lister, R., Hulett, J. M., Lithgow, T., and Whelan, J. (2005). Protein import into mitochondria: origins and functions today (Review). *Mol. Membr. Biol.* 22, 87–100.

Lund, A. A., Rhoads, D. M., Lund, A. L., Cerny, R. L., and Elthon, T. E. (2001). *In vivo* modifications of the maize mitochondrial small heat stress protein, HSP22. *J. Biol. Chem.* 276, 29924–29929.

Macherel, D., Benamar, A., Avelange-Macherel, M. H., and Tolleter, D. (2007). Function and stress tolerance of seed mitochondria. *Physiol. Plant.* 129, 233–241.

Mann, M., and Jensen, O. N. (2003). Proteomic analysis of post-translational modifications. *Nat. Biotech.* 21, 255–261.

Meyer, E. H., Taylor, N. L., and Millar, A. H. (2008). Resolving and identifying protein components of plant mitochondrial respiratory complexes using three dimensions of gel electrophoresis. *J. Proteome Res.* 7, 786–794.

Meyer, L. J., Gao, J., Xu, D., and Thelen, J. J. (2012). Phosphoproteomic analysis of seed maturation in Arabidopsis, rapeseed, and soybean. *Plant Physiol.* 159, 517–528.

Millar, A. H., Heazlewood, J. L., Kristensen, B. K., Braun, H. P., and Møller, I. M. (2005). The plant mitochondrial proteome. *Trends Plant Sci.* 10, 36–43.

Millar, A. H., Sweetlove, L. J., Giege, P., and Leaver, C. J. (2001). Analysis of the Arabidopsis mitochondrial proteome. *Plant Physiol.* 127, 1711–1727.

Millar, A. H., Trend, A. E., and Heazlewood, J. L. (2004). Changes in the mitochondrial proteome during the anoxia to air transition in rice focus around cytochrome-containing respiratory complexes. *J. Biol. Chem.* 279, 39471–39478.

Millar, A. H., Whelan, J., and Small, I. (2006). Recent surprises in protein targeting to mitochondria and plastids. *Curr. Opin. Plant Biol.* 9, 610–615.

Mollapour, M., Tsutsumi, S., Truman, A. W., Xu, W., Vaughan, C. K., Beebe, K., et al. (2011). Threonine 22 phosphorylation attenuates Hsp90 interaction with cochaperones and affects its chaperone activity. *Mol. Cell* 41, 672–681.

Moore, A. L., Gemel, J., and Randall, D. D. (1993). The regulation of *Pyruvate dehydrogenase* activity in pea leaf mitochondria (the effect of respiration and oxidative phosphorylation). *Plant Physiol.* 103, 1431–1435.

Moreau, F., and Romani, R. (1982). Preparation of avocado mitochondria using self-generated percoll density gradients and changes in buoyant density during ripening. *Plant Physiol.* 70, 1380–1384.

Nakagami, H., Sugiyama, N., Mochida, K., Daudi, A., Yoshida, Y., Toyoda, T., et al. (2010). Large-scale comparative phosphoproteomics identifies conserved phosphorylation sites in plants. *Plant Physiol.* 153, 1161–1174.

Neuburger, M., Journet, E. P., Bligny, R., Carde, J. P., and Douce, R. (1982). Purification of plant-mitochondria by isopycnic centrifugation in density gradients of percoll. *Arch. Biochem. Biophys.* 217, 312–323.

Palmieri, F., Rieder, B., Ventrella, A., Blanco, E., Do, P. T., Nunes-Nesi, A., et al. (2009). Molecular identification and functional characterization of *Arabidopsis thaliana* mitochondrial and chloroplastic NAD(+) carrier proteins. *J. Biol. Chem.* 284, 31249–31259.

Pappin, D. J., Hojrup, P., and Bleasby, A. J. (1993). Rapid identification of proteins by peptide-mass fingerprinting. *Curr. Biol.* 3, 327–332.

Posewitz, M. C., and Tempst, P. (1999). Immobilized gallium(III) affinity chromatography of phosphopeptides. *Anal. Chem.* 71, 2883–2892.

Qin, G., Meng, X., Wang, Q., and Tian, S. (2009). Oxidative damage of mitochondrial proteins contributes to fruit senescence: a redox proteomics analysis. *J. Proteome Res.* 8, 2449–2462.

Randall, D. D., Williams, M., and Rapp, B. J. (1981). Phosphorylation-dephosphorylation of pyruvate dehydrogenase complex from pea leaf mitochondria. *Arch. Biochem. Biophys.* 207, 437–444.

Rao, R. S., and Møller, I. M. (2012). Large-scale analysis of phosphorylation site occupancy in eukaryotic proteins. *Biochim. Biophys. Acta* 1824, 405–412.

Rasmusson, A. G., Geisler, D. A., and Møller, I. M. (2008). The multiplicity of dehydrogenases in the electron transport chain of plant mitochondria. *Mitochondrion* 8, 47–60.

Rasmusson, A. G., and Møller, I. M. (2011). "Mitochondrial electron transport and plant stress," in *Plant Mitochondria, Chapter 14*, ed F. Kempken (New York, NY: Springer), 357–381.

Rasmusson, A. G., Soole, K. L., and Elthon, T. E. (2004). Alternative NAD(P)H dehydrogenases of plant mitochondria. *Annu. Rev. Plant Biol.* 55, 23–39.

Reinders, J., Wagner, K., Zahedi, R. P., Stojanovski, D., Eyrich, B., van der Laan, M., et al. (2007). Profiling phosphoproteins of yeast

mitochondria reveals a role of phosphorylation in assembly of the ATP synthase. *Mol. Cell. Proteom.* 6, 1896–1906.

Robson, C. A., and Vanlerberghe, G. C. (2002). Transgenic plant cells lacking mitochondrial alternative oxidase have increased susceptibility to mitochondria-dependent and -independent pathways of programmed cell death. *Plant Physiol.* 129, 1908–1920.

Rubin, P. M., and Randall, D. D. (1977). Regulation of plant pyruvate dehydrogenase complex by phosphorylation. *Plant Physiol.* 60, 34–39.

Schmidt, O., Harbauer, A. B., Rao, S., Eyrich, B., Zahedi, R. P., Stojanovski, D., et al. (2011). Regulation of mitochondrial protein import by cytosolic kinases. *Cell* 144, 227–239.

Schulenberg, B., Aggeler, R., Beechem, J. M., Capaldi, R. A., and Patton, W. F. (2003). Analysis of steady-state protein phosphorylation in mitochondria using a novel fluorescent phosphosensor dye. *J. Biol. Chem.* 278, 27251–27255.

Schuller, K. A., and Randall, D. D. (1990). Mechanism of pyruvate inhibition of plant pyruvate dehydrogenase kinase and synergism with ADP. *Arch. Biochem. Biophys.* 278, 211–216.

Sommarin, M., Petit, P. X., and Møller, I. M. (1990). Endogenous protein phosphorylation in purified plant mitochondria. *Biochim. Biophys. Acta* 1052, 195–203.

Stensballe, A., Andersen, S., and Jensen, O. N. (2001). Characterization of phosphoproteins from electrophoretic gels by nanoscale Fe(III) affinity chromatography with off-line mass spectrometry analysis. *Proteomics* 1, 207–222.

Struglics, A., Fredlund, K. M., Møller, I. M., and Allen, J. F. (1998). Two subunits of the F0F1-ATPase are phosphorylated in the inner mitochondrial membrane. *Biochem. Biophys. Res. Commun.* 243, 664–668.

Struglics, A., Fredlund, K. M., Rasmusson, A. G., and Møller, I. M. (1993). The presence of a short redox chain in the membrane of intact potato-tuber peroxisomes and the association of malate-dehydrogenase with the peroxisomal membrane. *Physiol. Plant.* 88, 19–28.

Sugiyama, N., Nakagami, H., Mochida, K., Daudi, A., Tomita, M., Shirasu, K., et al. (2008). Large-scale phosphorylation mapping reveals the extent of tyrosine phosphorylation in Arabidopsis. *Mol. Syst. Biol.* 4:193. doi: 10.1038/msb.2008.32

Takahashi, A., Kawasaki, T., Wong, H. L., Suharsono, U., Hirano, H., and Shimamoto, K. (2003). Hyperphosphorylation of a mitochondrial protein, prohibitin, is induced by calyculin A in a rice lesion-mimic mutant cdr1. *Plant Physiol.* 132, 1861–1869.

Tan, Y. F., Millar, A. H., and Taylor, N. L. (2012). Components of mitochondrial oxidative phosphorylation vary in abundance following exposure to cold and chemical stresses. *J. Proteom. Res.* 11, 3860–3879.

Taylor, N. L., Heazlewood, J. L., and Millar, A. H. (2011). The *Arabidopsis thaliana* 2-D gel mitochondrial proteome: refining the value of reference maps for assessing protein abundance, contaminants and post-translational modifications. *Proteomics* 11, 1720–1733.

Thelen, J. J., Miernyk, J. A., and Randall, D. D. (1998a). Partial purification and characterization of the maize mitochondrial pyruvate dehydrogenase complex. *Plant Physiol.* 116, 1443–1450.

Thelen, J. J., Muszynski, M. G., Miernyk, J. A., and Randall, D. D. (1998b). Molecular analysis of two pyruvate dehydrogenase kinases from maize. *J. Biol. Chem.* 273, 26618–26623.

Thelen, J. J., Miernyk, J. A., and Randall, D. D. (2000). Pyruvate dehydrogenase kinase from *Arabidopsis thaliana*: a protein histidine kinase that phosphorylates serine residues. *Biochem. J.* 349, 195–201.

Tolleter, D., Jaquinod, M., Mangavel, C., Passirani, C., Saulnier, P., Manon, S., et al. (2007). Structure and function of a mitochondrial late embryogenesis abundant protein are revealed by desiccation. *Plant Cell* 19, 1580–1589.

Tovar-Mendez, A., Miernyk, J. A., and Randall, D. D. (2003). Regulation of pyruvate dehydrogenase complex activity in plant cells. *Eur. J. Biochem.* 270, 1043–1049.

Vanlerberghe, G. C., and McIntosh, L. (1997). Alternative oxidase: from gene to function. *Annu. Rev. Plant Phys.* 48, 703–734.

Vercesi, A. E., Martins, I. S., Silva, M. A. P., Leite, H. M. F., Cuccovia, I. M., and Chaimovich, H. (1995). Pumping plants. *Nature* 375, 24–24.

Vidal, V., Ranty, B., Dillenschneider, M., Charpenteau, M., and Ranjeva, R. (1993). Molecular characterization of a 70 kDa heat-shock protein of bean mitochondria. *Plant J.* 3, 143–150.

Wold, F. (1981). *In vivo* chemical modification of proteins (post-translational modification). *Annu. Rev. Biochem.* 50, 783–814.

Yao, Q., Bollinger, C., Gao, J., Xu, D., and Thelen, J. J. (2012). P(3)DB: an integrated database for plant protein phosphorylation. *Front. Plant Sci.* 3:206. doi: 10.3389/fpls.2012.00206

Zhu, Y., Lu, J. F., Wang, J., Chen, F., Leng, F. F., and Li, H. Y. (2011). Regulation of thermogenesis in plants: the interaction of alternative oxidase and plant uncoupling mitochondrial protein. *J. Integr. Plant Biol.* 53, 7–13.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 03 December 2012; paper pending published: 03 January 2013; accepted: 26 February 2013; published online: 13 March 2013.

Citation: Havelund JF, Thelen JJ and Møller IM (2013) Biochemistry, proteomics, and phosphoproteomics of plant mitochondria from non-photosynthetic cells. Front. Plant Sci. ***4****:51. doi: 10.3389/fpls.2013.00051*

This article was submitted to Frontiers in Plant Proteomics, a specialty of Frontiers in Plant Science.

Copyright © 2013 Havelund, Thelen and Møller. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.

frontiers in
PLANT SCIENCE

REVIEW ARTICLE
published: 26 April 2013
doi: 10.3389/fpls.2013.00100

Comparative analyses of nuclear proteome: extending its function

Kanika Narula, Asis Datta, Niranjan Chakraborty and Subhra Chakraborty*

National Institute of Plant Genome Research, Aruna Asaf Ali Marg, New Delhi, India

Edited by:
Nicolas L. Taylor, The University of Western Australia, Australia

Reviewed by:
Paula Casati, Centro de Estudios Fotosinteticos-Consejo Nacional de Investigaciones Científicas y Técnicas, Argentina
Zhaohua Peng, Mississippi State University, USA

****Correspondence:***
Subhra Chakraborty, National Institute of Plant Genome Research, Aruna Asaf Ali Marg, New Delhi-110067, India.
e-mail: subhrac@hotmail.com

Organeller proteomics is an emerging technology that is critical in determining the cellular signal transduction pathways. Nucleus, the regulatory hub of the eukaryotic cell is a dynamic system and a repository of various macromolecules that serve as modulators of such signaling that dictate cell fate decisions. Nuclear proteins (NPs) are predicted to comprise about 10–20% of the total cellular proteins, suggesting the involvement of the nucleus in a number of diverse functions. Indeed, NPs constitute a highly organized but complex network that plays diverse roles during development and physiological processes. In plants, relatively little is known about the nature of the molecular components and mechanisms involved in coordinating NP synthesis, their action and function. Proteomic study hold promise to understand the molecular basis of nuclear function using an unbiased comparative and differential approach. We identified a few hundred proteins that include classical and non-canonical nuclear components presumably associated with variety of cellular functions impinging on the complexity of nuclear proteome. Here, we review the nuclear proteome based on our own findings, available literature, and databases focusing on detailed comparative analysis of NPs and their functions in order to understand how plant nucleus works. The review also shed light on the current status of plant nuclear proteome and discusses the future prospect.

Keywords: nucleus, comparative proteome, organeller proteome, plant, nuclear proteins

INTRODUCTION

Cell nucleus has a perplex, heterogeneous, self renewable, and dynamic social milieu which can sense signals, deformations, mechano-transduction, biochemical deliberations and many other processes ensuing outside its boundary. Nucleus is enclosed in a phospholipid rich membrane, which has very sensitive ion channels and pores that shuttles biomolecule in and out by conformational and morphological transformations. The nucleus, often referred as the "eukarya," is functionally divided into the nuclear interior, a structurally differentiated and articulated organization surrounded by an envelope which is dynamic but sensitive to outside milieu (Dahl et al., 2008). In lieu, the modular disposition of indispensable, dynamic, and complex morphological feature of nuclear locale administers its function. It paves the way for the role of nuclear architecture in critical regulatory processes. The nucleus is a fundamental component of the microenvironment of both plant and animal cells that has been substantially expanded during evolution and keeps the genetic material separate from other activities of cell (Roix and Misteli, 2002). It is the ultimate exhilaration of gene regulation and proteins directly controlling the gene expression (Wilson and Dawson, 2011). Furthermore, it has been reported that the nucleus plays an important morpho-regulatory role during organogenesis in animal, besides its pivotal role in chloroplast division in plants (Cavalier-Smith, 2006; Dahl et al., 2008). The plant nucleus has biomechanical and morphogenetic functions; it is a viscoelastic solid encompass temerity of protein complexes. The organization of nuclear proteins (NPs) into versatile assemblies provides precise control over the shape, size, and composition of the nucleus, which opens a route toward the construction of sensors, programmable packaging and cargo delivery system within the sub-nuclear compartments as well as between the organelle. Plasticity in the nucleus allows cell differentiation, while rigidity in the nucleus determines its mechanical stiffness (Jiang et al., 2006; Pajerowski et al., 2007). Beyond its paramount importance in the generation of form, nucleus is frequently considered "growth-regulating" (Cavalier-Smith, 2006). The nucleus is evolutionary and inherently bestowed with information that can be both stored and relayed to cell interior via templating processes. It serves as the regulator in cell signaling for perceiving and transmitting extra- and inter-cellular signals in many cellular pathways. Communication between the cytoplasm and the nucleus is necessary and evident because of events such as apoptosis (Broers et al., 2002), mechanical stress (Dahl et al., 2008), environmental perturbation (Cheung and Reddy, 2012) and pathogen infection (Rivas, 2012), which lead to altered biosynthesis and modification of nuclear architecture and downstream cytoplasmic events. In addition, nucleoskeleton acts as a substrate for genome partitioning during mitosis. Further, it has been recognized as a central portal for providing motor centers during chromosome segregation in cell division (De Souza and Osmani, 2009). However, the available data is rather scarce and motor proteins between nucleoskeleton and chromosomes are still not known in higher plants. Throughout the plant kingdom the formation and regulation of the nuclear architecture has been shown to have the potential to influence many conduits of development, epigenetic differentiation, microfabricated patterning and cell senescence,

besides environmental stress response and pathobiology (Vergnes et al., 2004; Constantinescu et al., 2006; Dahl et al., 2008). Also, the nucleus serves a multi-functional role, as a regulator and modulator during cell division, and controller and integrator for fertilization and inheritance. Thus, nucleus plays a critical role as a modulator of cellular phenotype (Franklin et al., 2011). The nucleus must therefore be dynamic as cells divide, modulating its composition and architecture during its formation and after it has been disintegrated. The nuclear function is a multi-step, complex process, and the underlying mechanisms governing these steps are not fully understood.

All eukaryotic lineages are characterized by the loss, gain, expansion, and diversification of gene families (Fritz-Laylin et al., 2010). Understanding protein diversity and shared features can give unprecedented insight into the most fundamental aspects of nuclear structure and protein organization in as diverse kingdoms as plants and animals. Determination of organellar proteomes – the complement of proteins that reside, even if temporarily, in a specific organelle or sub-cellular region is of fundamental importance. Sub-cellular fractionation of tissue and cells in combination with MS/MS analysis has proven to be a powerful approach for the identification of proteins contained in specific organelles, such as the nucleus. Proteome research holds the promise of understanding the molecular basis of the nuclear function using an unbiased comparative and differential approach. Although the field of angiosperm eukaryogenesis has plethora of contradictory ideas, the nature of molecular changes can be reflected from the proteome. The nuptials of proteomics with cell biology have produced extensive inventories of the proteins that inhabit several sub-cellular organelles, including nucleus (Schirmer and Gerace, 2005; Yates et al., 2005; Yan et al., 2008). We and others have identified several hundred plant and animal NPs that include both predicted and non-canonical candidates, presumably associated with a variety of functions; viz., nucleoskeleton structure, development, DNA replication/repair, chromatin assembly/remodeling, signal transduction, mRNA processing, protein folding, transcription and splicing regulation, transport, metabolism, cell defense and rescue; all of which impinge on the complexity of NPs in plant (Pandey et al., 2006; Choudhary et al., 2009) and animal (Henrich et al., 2007). In recent years, reports have also been published focusing on changes in the nuclear proteome in varied cellular events (Bae et al., 2003; Lee et al., 2006; Salzano et al., 2006; Henrich et al., 2007; Buhr et al., 2008; Pandey et al., 2008; Repetto et al., 2008, 2012; Choudhary et al., 2009; Abdalla et al., 2010; Cooper et al., 2011; Varma and Mishra, 2011; Abdalla and Rafudeen, 2012). The identified proteins revealed the presence of complex regulatory networks that function in this organelle. NPs have been shown to account for approximately one-fourth of total proteins in yeast (Moriguchi et al., 2005) and one-fifth in animals (Bickmore and Sutherland, 2002), but the arithmetic estimate in plants is not yet complete. Currently, the focus is on nuclear proteomes in order to understand the nucleus-related processes in plants and animals. Although over the past few years there have been rapid advances in nuclear proteome research, the study on the complexity of NPs remained secondary, despite the fact they correspond to about 10–20% of the total cellular proteins and are comprised of several hundred different molecules with diverse functions. Moreover, a vast array of post-translational modifications to these proteins add diversity to the structure and ligand-binding properties of nuclear components, leading to their differential activity. Therefore, characterization of the nuclear proteome in plant hold the promise of increasing our understanding about the regulation of genes and their function.

Here, we begin by giving updates on the nuclear proteomes and summarizing the essential and unique features of the nucleus. We also discuss recent findings concerning the regulation and biochemistry of it with specific emphasis on the fundamental role of NPs in development, DNA replication/repair, transcriptional regulation, environmental stress, and signaling by analyzing the nuclear proteomes. Furthermore, we report the cross-kingdom comparative analysis of nuclear proteomes toward organism specificity and plant exclusivity based on our own findings, the available literature and databases focusing on NPs in view of the current understanding and perspectives of the nuclear functions.

ORIGIN OF THE NUCLEUS

A landmark event in the evolution of eukaryote was the acquisition of nucleus. Eukarogenesis have evolved albeit independently in plants and animals. Although both are true eukaryotes they have different ancestors. The evolution from prokaryotes to eukaryotes was the most radical change in cell organization. It is known that evolution of complex characters typically involves preadaptation, radical mutational innovation, and different selective forces acting in succession (for review, see Cavalier-Smith, 2006). Physical and mutational mechanisms of origin of the nucleus are seldom considered beyond the longstanding assumption that it involved wrapping pre-existing end membranes around chromatin (Cavalier-Smith, 1988). Evolution of the nucleus starts approximately 850 Million years ago (Cavalier-Smith, 2002), but it was 1833 when Robert Brown discovered the nucleus and said "vim and vigor is sexless devoid of this facet" in a paper to the Linnean Society. Origin of nucleus requires understanding of co-evolution of different nuclear components and their functional interlinking into the fundamentally novel eukaryotic life style. There are two competing theories of eukaryotic evolution. According to the first theory, a subset of bacteria slowly developed nucleus, while in the other, eukaryotes came first, some of them then lost nucleus and gave rise to bacteria. But, woesean revolution highlights that eukaryotes came from archaeal stock. Since, eukaryotes contain both archaeal and bacterial genes and the division of labor arose from the ancient symbiotic partnership between them that gave rise to eukaryotic nucleus. A third option for the nuclear origin revolves around the viruses, but the supporting data are provocative, circumstantial, and controversial (for review, see Pennisi, 2004).

DESCRIPTION OF TOOLS TO STUDY NUCLEAR PROTEOME

An outline of the procedure and the illustration of the data that can be generated with the methodology are shown in **Figure 1**. Each proteomic study is described through a simplified flowchart showing its different steps from experimental material to protein identification. As illustrated in **Figure 1**, density gradient methods can be used to prepare a nuclear fraction with or without DNA

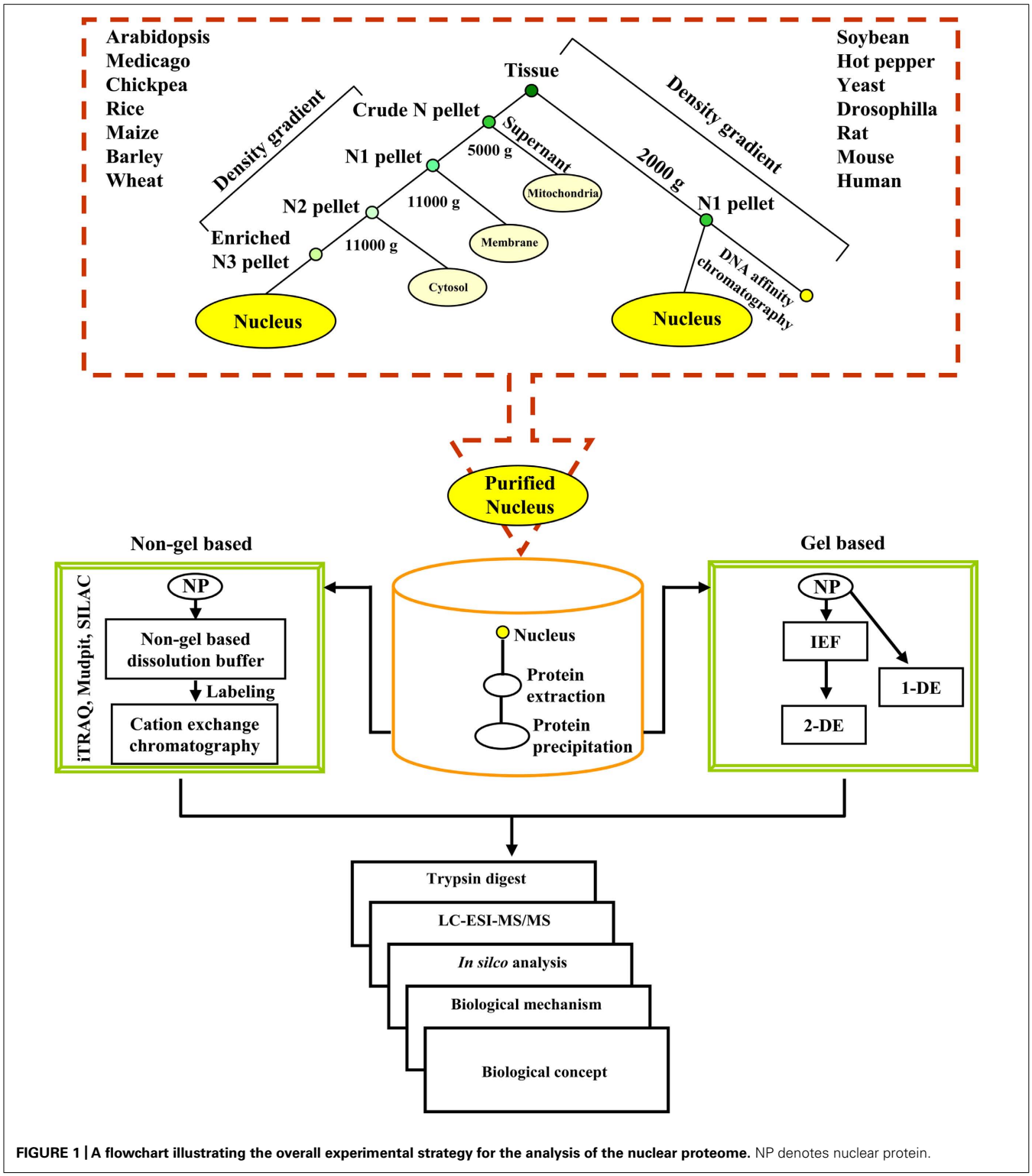


FIGURE 1 | A flowchart illustrating the overall experimental strategy for the analysis of the nuclear proteome. NP denotes nuclear protein.

affinity chromatography. The most efficient means to separate NPs are either two-dimensional gel electrophoresis or cation-exchange chromatography followed by elution of protein fractions with salt gradient. Over the past few years 2-DE coupled MS/MS and LC-MS/MS have extensively been used to study nuclear proteomes in varied organisms (for reviews, see Khan and Komatsu, 2004; Cullen and Mansuy, 2010; Erhardt et al., 2010). In brief, the NP fractions after separation are digested to allow identification of proteins by mass spectrometry. Proteins can be directly submitted to enzymatic digestion with appropriate proteases, such as trypsin

or to chemical treatment to get peptides of appropriate mass (usually between 750 and 4000 Da). Identification of proteins can then be done either by peptide sequencing using liquid chromatography coupled to MS (LC-MS/MS) or by peptide mass mapping using matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF/TOF) followed by *in silico* analyses. Custom NP databases, for example, yeast-NPD[1], human-NPD[2], *Medicago*-NPD[3], TAIR[4], rice nuclear proteome database[5] help improve the identification and post-translational modification of the NPs. To undertake a comprehensive comparison of the plant nucleolar proteomes based on a combined approach of alignment, structure and phylogeny an *Arabidopsis* nucleolar protein database was curetted (Brown et al., 2005). Similarly, comprehensive and well-annotated database of transcription factors may provide a useful resource to check annotations and to study gene regulatory pathways (Guo et al., 2005, 2008; Iida et al., 2005; Gao et al., 2006; Palaniswamy et al., 2006; Zhu et al., 2007; Rushton et al., 2008; Perez-Rodriguez et al., 2010; Romeuf et al., 2010; Wang et al., 2010).

Nuclear proteins are often under-represented in proteomic studies due to their low abundance. The information offered from total nuclear proteome by high-throughput techniques does not illustrate the functional purpose of NPs and compartment structures. Computational modeling, on the other hand, may elucidate functional roles otherwise not captured by any individual existing experimental technology. The predictions used to identify the common organelle-specific sequence features are successful for over-represented proteins but is limited for the low abundant proteins. Thus, this analysis allows the identification of additional proteins sharing the same motif and to estimate the enrichment of the protein motifs in nuclear proteome data set. The estimation of the enrichment of those motifs in the nuclear proteome data set was done by comparing their frequencies in the nuclear data sets and in the target protein databases. Some of the motifs that can be identified in the nucleus are domains of well-known proteins, including histones and helicases. There are many proteins that are known to be imported into the nucleus, but which have no known intra-nuclear compartment association. These proteins may share similar cellular locations or functions, but further experiments are needed for clarification (Gorski and Misteli, 2005).

WHAT HAVE WE LEARNT?

Proteomics has turned out to be an imperative benefactor for studying the acquaintance of plant nuclear structure and function. The field of proteomics is evolving from cataloguing the proteins under static conditions to comparative analyses (Narula et al., 2012). Defining proteins that change in abundance, form, location or other activities may indicate the presence and functional significance of a protein. Whereas comparative nuclear proteome research is quite advanced in animals (Liao et al., 2009) and yeast (Gauci et al., 2009), there is less information in plants. The investigation on plant nuclear proteomes in recent years has raised the following important questions: What are the essential plant NPs? Do NPs show clade specificity in vascular plants? What are those organ-specific NPs, if any? Does the nuclear developmental proteomics of one of the clades yield any astonishing or prolific results? How do NPs remodel during environmental- and/or patho-stress those provide new perspectives? Are some of the NPs unexpected? And, last but not the least, what sort of post-translational modifications have so far been characterized in the nucleus? Here, we analyze and compare the experimental results thus far available on nuclear proteomes to elucidate the dynamics of plant NPs.

DECIPHERING THE ORGANISM-SPECIFIC NUCLEAR PROTEOME DYNAMICS: SOCIAL CLASS VS. DIVERSITY

Proteins evolve at rates differing over many orders of magnitude. As new proteins evolve by gene duplication, evolutionary rates must change dramatically over time. They change systematically among different branches of the evolutionary tree and also episodically (Cavalier-Smith, 2010). In the history of life there are three mega evolutions giving rise to prokaryotes, plants, and animals. Consequently, decoding organism specific nucleoid/nuclear proteomes are of utmost importance to understand the diversity among the protein complement in the three life lines of fruition. The possibility of intra-kingdom and/or cross-kingdom comparison of proteins and cellular regulation with the use of advanced proteomic techniques are of great value. We compared the experimentally determined nuclear proteomes of plants viz., *Arabidopsis thaliana*, *Cicer arietinum, Medicago sativa*, *Glycine max*, *Capsicum frutescens*, *Xerophyta viscose*, and *Oryza sativa* with that of yeast, fruit-fly and animal (human, rat, and mouse; **Table 1**). The *modus operandi* in investigating the nuclear proteomes of available species were the extensive literature search, availability of relevant databases (human, mouse, rat and plant, TIGR, UniProt and Swissprot) and *in silico* analysis. After comparing individual NPs, we investigated the divergence of these proteins among animals and plants to understand the integration and coordination of nuclear functions. Further, we calculated the percentage of proteins that was found to be unique to each proteome by calculating the number of proteins estimated from matches to SWISS-PROT as described in Skovgaard et al. (2001), Semple et al. (2003), and Bhushan et al. (2006). The NPs identified in these studies were classified into different functional categories. This classification is only tentative, since the biological role of many of the proteins identified has not been established experimentally. Furthermore, we applied a cross-species comparison on the available datasets. When analyzing proteomes within the specified group of plants, a logical strategy was used to maximize efficiency and the overall comparative results. Thus, it was imperative to first evaluate the available nuclear proteome maps, followed by an analysis of stimulus-specific proteomes of the above mentioned organisms. We then moved on to assess the stress-responsive plant nuclear proteomes in order to understand the overlap and specificity amongst different environmental- and patho-stress. These comparative studies were customized for specific protein families. It is to be noted that protein consensus can be obtained across any combination of proteomes based on the type of extraction procedure.

[1] www.pin.mskcc.org
[2] www.npd.hgu.mrc.ac
[3] www.masc.proteomics-org
[4] www.arabidopsis.org
[5] http://gene64.dna.affrc.go.jp/RPD

Table 1 | A comprehensive list of plant nuclear proteomes.

S. No.	Organism	No. of proteins identified	Biological material used for fractionation	Technique used	Experimental conditions	Reference
1	*Chlamydomonas reinhardtii*	672	Cell suspension culture of algae	nanoLC-MS/MS	Cell	Winck et al. (2011)
2	*Arabidopsis thaliana*	154	Leaves	2DE/LC-MS/MS	cold	Bae et al. (2003)
3	*Medicago truncatula*	143	Seed	nano-liquid chromatography–nano-LC-MS/MS	Development	Repetto et al. (2008)
4	*Cicer arietinum*	150	Leaves	2DE/LC-MS/MS	Tissue	Pandey et al. (2006)
5	*Cicer arietinum*	150	Leaves	2DE/LC-MS/MS	Dehydration	Pandey et al., 2008
6	*Xerophyta viscosa*	18	Leaves	2-DE/ MALDI-TOF-TOF	Dehydration	Abdalla et al. (2010)
7	*Xerophyta viscosa*	121	Leaves	2-DE/ ESI- MS/MS	Dehydration	Abdalla and Rafudeen (2012)
8	*Glycine max*	292	Leaves	nano-liquid chromatography–nano-LC-MS/MS	Rust fungus	Cooper et al. (2011)
9	*Capsicum annuum*	56	Leaves	2-DE/MALDI-TOF-MS	Tobacco mosaic virus	Lee et al. (2006)
10	*Linum usitatissimum*	292	Immature seed coat at 16daf (Torpedo stage)	1-DE, WB, Dot blot, gel shift	Development	Renouard et al. (2012)
11	*Oryza sativa*	468	Endosperm	LC-MS/MS	Tissue	Li et al. (2008)
12	*Oryza sativa*	109	Leaves	2DE/LC-MS/MS	Dehydration	Choudhary et al. (2009)
13	*Oryza sativa*	307	Leaves	Percoll gradient centrifugation-nano LC-MS/MS	Glucose responsive	Aki and Yanagisawa (2009)
14	*Zea mays*	Not identified	Endosperm	1-DE, WB	Development	Ferreira et al. (2006)

Nuclear protein composition was found to differ between two major Kingdoms viz., plant and animal. Results were predominantly obtained with human (lymphoma and myeloid lines; Salzano et al., 2006; Henrich et al., 2007), mouse (Buhr et al., 2008), rat (McClatchy et al., 2011), *Drosophila* (Varma and Mishra, 2011), model plants *Arabidopsis* (Bae et al., 2003), *Medicago* (Repetto et al., 2008, 2012), and crop plants rice (Choudhary et al., 2009), hot pepper (Lee et al., 2006), soybean (Cooper et al., 2011), *Xerophyta* (Abdalla et al., 2010; Abdalla and Rafudeen, 2012), and chickpea (Pandey et al., 2006, 2008). The stunning findings from these comparisons suggest that until now only 1868 NPs are identified in humans, while 1548 in mouse, 842 in rats, 282 in *Drosophila*, 328 in yeast, and 1510 in plants contributing to the large repertoires of the nuclear proteome database and prophram (**Figure 2A**). A more accurate vision of animal nuclear proteome illustrate that approximately 2000 NPs from lymphoid and myeloid tissues of humans symbolizing around one-third of their estimated nuclear proteome. The two branches of angiosperm, monocot, and eudicot, firmly set up compelling evidence from the available nuclear proteomes that some NPs are unique, while some are shared. Extended nuclear proteome research is required for monocot family with only 312 NPs reported thus far than dicots that reported 1856 NPs (**Figure 2A**). Contextual information on nuclear proteomes of eudicots revealed that until now 521 NPs have been identified in *Arabidopsis* representing about one-third of its estimated nuclear proteome (Bae et al., 2003), while 406, 282, 219, 82, and 133 NPs were identified in *Medicago*, soybean, chickpea, hot pepper, and *Xerophyta*, respectively. To explore early messages arising from comparison of the content of monocot and dicot proteomes address key consequences of research for dicot comparative proteomics. The recurring observation that monocot proteome research centers on rice proves factual for nuclear proteome. Until recently, 212 NPs have been identified in rice, whereas only 50 and 51 NPs are known in wheat and barley, respectively (**Figure 2A**). Our comparative analyses of different species in relation to their function showed that high percentage of proteins to be unique to each proteome: 89% in animal (human, rat, and mouse), 81% in human (lymphoma and myeloid lines), 71% in mouse, 68% in rat, 84% in yeast, and 74% in *Drosophila*; whereas plant proteomes show 85% in *Arabidopsis*, 78% in soybean, 81% in chickpea, 71% in *Medicago*, 84% in rice, and 54% in hot pepper with only actin and 26S proteasome being the social class of proteins present ubiquitously in all. The available nuclear proteomes of nine plants compared in **Figure 2A** varied in molecular weight from 9.1 to 150 kDa and had a spread of p*I* values from 3.6 to 10; while yeast shows 15 to 110 kDa, 3.1 to 12.0 pI; *Drosophila* shows 12 to 140 kDa, 3.1 to 11.4 pI and animals show 9.4 to 150 kDa, 3.0 to 12.0 pI. Most of the NPs were basic in nature concordant with the acidic environment of this organelle.

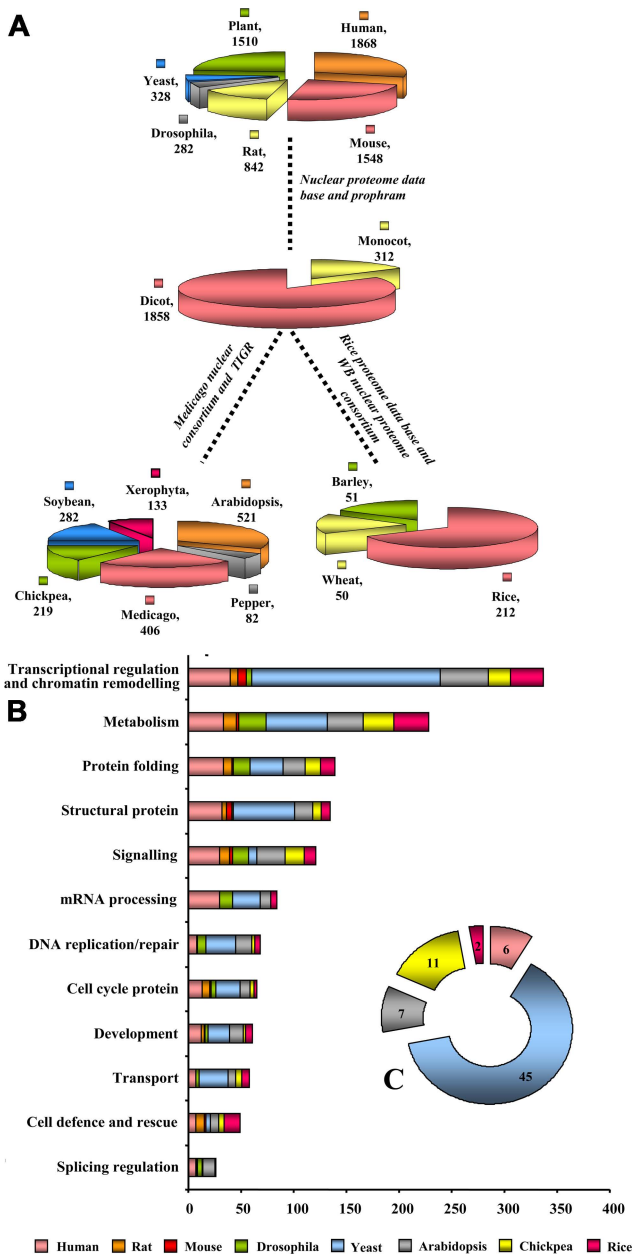


FIGURE 2 | Benchmarking nuclear proteomes. (A) Cross-kingdom-clade-species comparison of nuclear proteomes. **(B)** Functional classification of the identified proteins was made according to the biological processes. The length of the bar indicates the number of proteins present in a particular species, such as, *Arabidopsis*, chick pea, rice, human, rat, mouse, *Drosophila*, yeast. **(C)** The pie chart inset represents the fraction of unknown proteins in each of these organisms.

Functional categorization of nuclear proteomes reported till date revealed an imperative corollary, which shows overall proteins belonging to transcriptional regulation and chromatin remodeling contribute radically to nuclear proteomes of yeast (54%), plants (29%), animals (16.5%), and *Drosophila* (1.2%). 54% of proteins in this category in yeast represent more than half of the total NP implying thereby that yeast is a dynamically dividing organism having plurality of transcription regulators. Of these, 42% represents nuclear structural proteins involved in cell division. Furthermore, plants are deskbound, therefore to establish, maintain and alter global and local level of nucleic acid they require rapid turnover of DNA and RNA metabolizing proteins, DNA replication/repair proteins, and splicing regulation proteins toward acclimatization in the environment. Indeed, identified

plant NPs from the available reference proteomes showed 42% NPs belong to metabolism category whereas 46% confers splicing regulation. NPs behave like a network scaffold and acts as an entry point to ensure smoother regulation of different cellular processes that require rapid protein turn over. Comparison of plant nuclear proteomes with other organisms revealed that protein folding and turnover category contributes to 35%, which is in close correlation to 30% NPs from animals. Plant protein networks revealed the predominance of the development specific proteins (36%) and cell cycle proteins (24%). It is to be noted that the NP extraction protocol used in the animal and plant are different. Animal nuclear proteome research spotlight sub-nuclear compartmentalization, whereas, plant nuclear proteomes except *A. thaliana* (Calikowski et al., 2003; Pendle et al., 2005) and *O. sativa* (Tan et al., 2007, 2010) tranquil total nucleus. Therefore, the chromatin assembly/remodeling proteins are identified much less in all organisms as they can be isolated best in high salt buffer concentration which is not usually used for total NP extraction. A dramatic rearrangement of the nuclear structure takes place during mitosis and meiosis, which dynamically changes sub-nuclear proteomes (Beven et al., 1995; Holmes-Davis and Comai, 1998). Major role of interphase chromatin is in transcription, while mitotic chromatin contributes to cell division and meiotic chromatin is engaged in pairing, cross-over, and chromosome segregation. Compositional divergence in protein complement of stationary and dividing nucleus is thus a call to study proteomes at different phases of nuclear division.

Between plant and animal, the gene families and members are not related but functionally appear to be similar. Mosaic comparison of nuclear proteomes revealed that chimeric evolution was the main cause of proteome diversity in animals. For example, in human nucleophosmin protein has diverse protein members, whereas in mouse it is conserved. Likewise, when rat and mouse were compared DEAD, septin, lamin B2 box are some of the diverse class of proteins found in rodents. Also, when human, rodents and *Drosophila* were evaluated lactate dehydrogenase, nicotinamine synthase1, tubulin beta, and 26S proteasome represented the social class. Furthermore, when, human, rodent (mouse, rat), insect (*Drosophila*) and fungus (yeast) were compared actin and nicotinamine synthase 1 correspond to the social class. Whereas, comparison between monocots and dicots, showed that *Arabidopsis* is better explored than rice and therefore comparison of their proteomes may not yield the postulated results as defined by genome analysis. But the proteins revealed an evolutionary divergence in plant as well as dicot vs. monocot specificity, with few conserved proteins (**Figure 2B**). When the *Medicago* nuclear proteome was compared with that of *Arabidopsis*, results revealed an evolutionary divergence as well as tissue specificity, with few conserved proteins (**Figure 2B**). Comparison of the functional classes of NPs amongst dicot species like *Arabidopsis*, *Medicago*, soybean, hot pepper, *Xerophyta*, and chickpea confirm the dynamic and heterogeneous nature of nucleus as exemplified by the presence of only actin in all dicots. Another protein namely 26S proteasome may be considered as social class except its absence in *Xerophyta*. The presence of chaperone 60 in *Cicer* and HSP71in *Medicago* illustrate that nature invented vastly different solutions to a common problem viz., protein folding. When the studies on the legumes like *Cicer* and *Medicago* were compared to *Arabidopsis* belonging to the Brassicaceae family (**Figure 2**), it can be readily observed that the splicing regulation in the nucleus for activating splicing enzymes is diverse between the two families as well as between the members of the same family, leguminosae. The protein network of rice revealed the predominance of the chromatin assembly/remodeling proteins, for example, histone deacetylase, histone 2A, histone 2B, histone 3, histone 4, while the *Arabidopsis* protein network was found to be rich in splicing regulation proteins and structural protein as transcriptional regulators.

It may be assumed that the divergence in the resulting proteomes of the vascular plants is due to the presence of the different nuclear architecture based on the protein and nucleic acid compositions, suggesting the occurrence of clade-specific NPs that would bind to their cognitive biomolecules to bring out specific functions both spatially and temporally. Most intriguing are the remaining 10–18% of plant NPs that do not have any similarity to the known proteins in other organisms. The challenge is to elucidate their biological role within the cell nucleus.

EXPLORING THE SINK AND LINK IN NUCLEUS

Ubiquitously present, except in RBC, the nucleus is composed of different molecules with diverse functions to meet the specialized requirements of different organs and tissues. Nuclear functional compartmentalization is a paradigm of molecular machines necessary for biogenesis and functionality (Strouboulis and Wolffe, 1996). It is a dynamic milieu having a reservoir for bioactive molecules, such as carbohydrates, nucleic acid, and proteins which is necessary for assembly and also for communication with the other parts of the cell. For decades, cell nucleus has been a black box in biology. The determination of comprehensive chemical differences between plant and animal nucleus is still difficult to understand, but the switching of cellular programs by NPs mediated chemical networking is tightly linked to the regulation of gene expression in both the kingdoms. However, the distributions of transcription sites in chromosome territories are conserved in plants and animals. It is the heterochromatic centers which makes the difference in nuclear processes in both these kingdoms (van Driel and Fransz, 2004). Being a store house of nucleic acid and proteinaceous domain, nucleus contains distinct structural and functional compartments (Misteli, 2005). Proteinaceous domain include nucleolus containing rRNA binding protein and splicing proteins, the cajal body having snRNA forming and binding proteins; whereas the nucleic acid domain encompass euchromatin and heterochromatin. Euchromatin is the reservoir of histones and histone binding proteins, while heterochromatin consists of heterochromatin binding unknown proteins. In these two domains nucleic acid occurs three dimensionally (Dundr and Misteli, 2001). Nuclear bodies are functionally and/or morphologically discrete accommodating usually distinct resident proteins. Paradigm includes the nucleoli (site of rRNA transcription), nuclear speckles (site for splicing) and splicing factor compartment (store-house for cajal body and PML body; Takizawa and Meshorer, 2008). Dramatic developments in high-resolution live-cell imaging have revealed the cell nucleus as a highly heterogeneous and complex organelle, and the global genome and proteome architecture

changes during processes such as differentiation and development (Misteli, 2001; Spector, 2003; Misteli, 2005). It is, therefore, relevant that different family members show highly regulated and specific patterns of the expression of nuclear components in an evolutionary context. Similarities in nuclear design may be apparent as it is likely that ancient functional protein domains and nucleic acid backbones have been used in a variety of arrangements and combinations to affect the function of convergent biological structures. Nucleus serves as the self organizing mediator. Most proteins are in constant motion, and their residence time within a compartment is very low, being at most 1 min (Gonzalez-Melendi et al., 2000). This mobility ensures that proteins find their targets by energy-independent passive diffusion (Pederson, 2000). In addition to protein heterogeneity and the presence of various regulators, mediators, transducers as well as linkers, RNA and chromatin compositions can vary between cell types and even within a given cell in different time (Hetzer et al., 2005), suggesting that the nucleus serves as a sink of variability in terms of macromolecules or microelements. Regulated trafficking of proteins, RNAs, RNA-protein complexes, and other molecules in and out of the nucleus is important in diverse processes. The nucleus serves as the end line culminator in cell signaling to perceive and transmit extra- and intercellular signals in many cellular pathways. NPs not only constitute more than just a structural scaffold, but also play various roles in development, cell cycle, defense against environmental stresses and in the tight regulation of gene expression.

THE NUCLEAR PROTEIN SINK: A DYNAMIC FRAMEWORK FOR MULTIPLE FUNCTIONS

Eukaryotic NPs are complex with plurifunctional role, evolutionary tinkering, and subtle modifications evoked repeatedly and independently among different taxa. A macromolecular machine in the form of nuclear pore allows a protein or protein complex up to approx. 500 kDa to traverse the nucleus. NPs roam through the nucleus in search of a high-affinity binding site where they can exert their functions (Dundr and Misteli, 2001; Mans et al., 2004). The specific domain and architecture of NP contain information of biological importance and evolutionary value.

Altogether, NPs include those which are highly mobile viz., transcription factors, pre-mRNA splicing factors, rRNA processing enzymes and 3α-processing factors, DNA repair enzymes, chromatin-binding proteins and apoptotic caspases; while immobilized NPs encompass DNA replication factors, intermediate filament proteins, and histones H1 (Dundr and Misteli, 2001). Plant and animal show least homology as far as nuclear intermediate filament proteins are concerned. However, proteins belonging to DNA replication/repair are found to be orthologous (Moriguchi et al., 2005). In plants, elongation factor thermo unstable (EF-Tu), Zinc finger protein, glycine rich RNA binding, histone 2B, histone 3, glycine dehydrogenase, peptidyl prolyl isomerase, 26S proteasome, 60 kDa chaperone, glyceraldehyde 3-phosphate dehydrogenase, malate dehydrogenase, peroxiredoxin, transaldolase, calcium protein kinase, PHO1 like protein, ά expansin, actin, 14-3-3, and 40S ribosomal protein SA are consistently represented in thus far studied nuclear proteomes that play diverse and crucial roles in nuclear function. Most predominant class of NPs reported are the transcription regulators in which TF2A, RNA polymerase have been optimized during eukaryotic evolution for acting in post-transcriptional gene regulation. The linear representation of promoter elements provides competency for physiological responsiveness within the contexts of development, cell cycle, and phenotype-dependent regulation as transcription factors can bind to these cis-acting elements dictating where and when a gene to be active. Chromatin binding proteins and nucleosome organization protein viz., PolII, MADS box, RCC2 protein, and HEAT box reduce distances between independent regulatory elements providing a basis for integrating components of transcriptional control. It is known that the nuclear matrix proteins support gene expression by imposing physical constraints on chromatin related to the three-dimensional genomic organization. In addition, the nuclear matrix proteins facilitate gene localization besides the concentration and targeting of transcription factors. Histone deacetylase 6 and DNA methyltransferase physically interact; together they mediate histone acetylation and modulate DNA methylation status, silencing the transposable element (Casati et al., 2008; Casati, 2012). Transcriptional reprograming by WRKY, ERF, TGA, Whirly, and MYB factors is thought to cause alteration in transcript level, which in turn regulates various physiological processes like growth, development, and pathogen perturbation (Mayrose et al., 2006; Yu et al., 2011; Feng et al., 2012). Among others, it is ascertain that G5bf protein and TF rough sheath 2 are embodied persistently in dicot nuclear proteome in customary environment. Perhaps, the protein most expected to be similar to their metazoan counterpart in the plant nucleus is DNA ligase, which have been shown to regulate transcription (Truncaite et al., 2006). RF2B, SPT2-chromatin binding domain, RING zinc finger protein, and gypsy-like retroposon are exclusively present in monocot nuclear proteomes (Li et al., 2008; Aki and Yanagisawa, 2009; Choudhary et al., 2009). Aforesaid, transcription regulators of two clades have solitary similarity that they are regulated by circadian rhythm and have multivariate decision to find motif combination (Cavalier-Smith, 2010). Proteome data indicate that BABY BOOM, AP2/EBEBP2, and syringolide induced proteins are leguminosae allied transcriptional regulators having role in development, cell/organ identity and fate; while ribosomal recycling factor, CHP rich zinc finger protein, nucleolin, RuvB, BRI KD interacting protein, WPP domain protein, pescadillo protein, and MYB transcription factor are solanaceae associated transcriptional regulators (Boutilier et al., 2002; Abe et al., 2003). Each of these proteins reported in leguminosae and solanaceae have been shown to be involved in diverse cellular functions, viz. development, embryogenesis, and signaling pathways. This further highlights the technical challenges when attempting to isolate high purity nucleus and resolution of proteins using proteomic technology. Proteins involved in the metabolism are customary in the case of any nuclear proteome. Indeed, methylenetetrahydrofolate reductase, homocysteine methyl transferase, methyltransferase, and ornithine aminotransferase are proteins belonging to this category, which play pivotal role in RNA and DNA metabolism. Astounding result obtained from the comparative analysis of nuclear proteomes in plants with that of animals, suggests the

presence of many metabolism related animal orthologs in dicot (*Arabidopsis*) proteome whose role in plants have not yet been defined viz., biliverdin reductase A1, LROS1 acyl transferase, and KES1 oxysterol binding protein. Nuclear structural proteins are ubiquitous in both the kingdom but shows more divergence in plants (Meagher et al., 1999). Actin and myosin are the ancient component of this category that form a platform for all three DNA-dependent RNA polymerases, mediate RNA export from the nucleus, and are required for the long-range movement of specific loci within the nucleus (Caudron-Herger and Rippe, 2012). Recent evidence suggests that proteins such as actin, myosin, tubulin, NuMA, Annexin A1, Annexin A2, viscialin, spectrins, and titin are recognized as having fundamental roles in nuclear structure and genome function in living eukaryotes (Wilson and Dawson, 2011). Coiled-coil protein, an orthologue of lamin, the building block of the NPC complex in plants (Mans et al., 2004) has many candidates, namely disease resistance proteins, NUF1, NUP82, NUP88. Nucleoporins, which anchor intermediate filament proteins during scaffold formation play a crucial role in chromosome scaffolding and mRNA export. The nucleoskeleton and nucleopore complex protein present ubiquitously in plants are expansins and NUPs. They maintain the mechanostatic and load bearing properties of the nucleus (Dahl et al., 2008). In other words, nucleostructural dynamics in plant cell is a team effort of multiple proteins orchestrating this very fast-paced game. Nucleus is an evolutionary chimera of cell cycle related proteins. Progress made in the area of plant cell cycle regulation has resulted in recognition of NP candidates, including chromobox proteins, RCC2 proteins, and BUB3 having role in cell division. Additionally, CDC 5, one of the cell cycle proteins has a varied role in mitosis, ciliary motility and trafficking. Another component of cell cycle regulation namely, Ran cycle is represented by nuclear Ran GTPase, mago nashi protein, Ras GTPase, and Ran. Ribosome subunit export system of nucleus involved in cell cycle regulation focus on the three most appealing candidates: Nops, Nugs, GTPases, besides recently added AAA-ATPase and exportin in animals. However, analyses of plant nuclear proteomes do not show the presence of these proteins. Perhaps another protein most expected to be similar to their metazoan counterparts in the plant cell nucleus is karyopherins, which has a role in nuclear trafficking. Nucleus includes numerous enzymes viz., rRNA processing enzyme, polymerase, ligase, gyrase, and number of helicase that alter DNA conformation, replication, degradation, and chromatin modifiers. Histone variants, one of the important classes of NPs in eukaryotes are important component which play a key role in genome maintenance and stability. In the nucleus, RNA is an architectural factor for shaping the genome and its nuclear environment, besides being an effector molecule in maintaining the chromatin structure (Ma et al., 2011). We find many mRNA processing proteins that include nucleolar RNA-associated protein (NRAP), LSM2, paraspeckle protein 1, and Non-POU domain containing protein in this category. It is well known that protein folding supports diverse but specific signal transducers and lies at the interface of several developmental pathways (Caudron-Herger and Rippe, 2012). Likewise, different chaperones, HSP71, proteasome subunit alpha types, DnaJ, protein disulfide isomerase, HSP20, glutathione-S-transferase, and HSP70 reported in plant nuclear proteomes might maintain protein homeostasis by providing stability to other nuclear resident proteins. Involvement of some of these chaperones with the class of developmental NPs viz., DEAD box, DUX3, von wilberand factor, HOMEO BOX, and U box have already been reported (Barthelery et al., 2008; Su and Li, 2008). A chronic theme proverbial to the class of nucleoskeleton linker proteins of plant cells is that these mechano-transducing transmembrane molecules communicate and interact preferentially with the intermediate filament on the nuclear side of the nuclear membrane. Our analyses suggest, several attributes of NP contribute to cross-talk in gene regulation and cellular phenotype.

NUCLEAR INVENTORIES FOR *IN SILICO* PROTEIN PROFILING OF COMPARATIVE STRESS PROTEOME

Nucleus senses and physiologically responds to environmental stress via signaling pathways. Signaling events are clearly not linear and induce many different reactions, including stress-related processes that crosstalk with hormone signaling pathways. Most signaling pathways culminate in the nucleus leading to regulation of expression of specific genes whose products are necessary for eliciting a signal specific response like nuclear localization of pathogen effectors, R proteins, and other host defense proteins that modulate stress response. Here, we have customized the comparative analyses for specific protein families. For example, when the environmental stress-responsive proteomes were compared, the parallel analysis of the proteomes of different clades of vascular plants were performed, viz., chick pea vs. *Xerophyta* vs. rice for dehydration, *Arabidopsis* for cold response, and *Medicago* for seed filling that mimic the dehydration response. Similarly, in case of patho-stress, soybean, and hot-pepper proteomes were compared.

We analyzed the nuclear proteomes of *A. thaliana* in response to cold- stress (Bae et al., 2003) and dehydration-responsive nuclear proteomes of *Cicer arietinum* and *O. sativa* (Pandey et al., 2008; Choudhary et al., 2009). Interestingly, a great level of divergence in the protein classes amongst these organisms was observed (**Figure 3**). To our surprise, except development category all of the NPs were found to have members common in all organisms under both kind of abiotic stresses studied. Families of development related proteins, viz., embryonic flower 1- like protein, copia-like, and Hd3a protein have been found in dehydration responsive proteome of rice, while chickpea DRPs exclude most of the nuclear structural proteins such as cellulose synthase like, alpha amylase, and beta-expansin otherwise abundantly present in dehydration-responsive proteome of rice. It is intriguing to note that cold-responsive NPs under all functional categories of *Arabidopsis* were present in dehydration-responsive proteomes of rice and chickpea. Another important finding was the presence of cyc3 protein in high abundance during cold-stress in *Arabidopsis*. Whereas zinc finger, ring finger, and RNA glycine rich proteins were predominantly found during dehydration response but were absent in response to cold-stress. Various kinases known to mediate the stress-induced synthesis of NPs, such as PHO1, galectin, thioredoxin peroxidase were present both in monocot and dicot

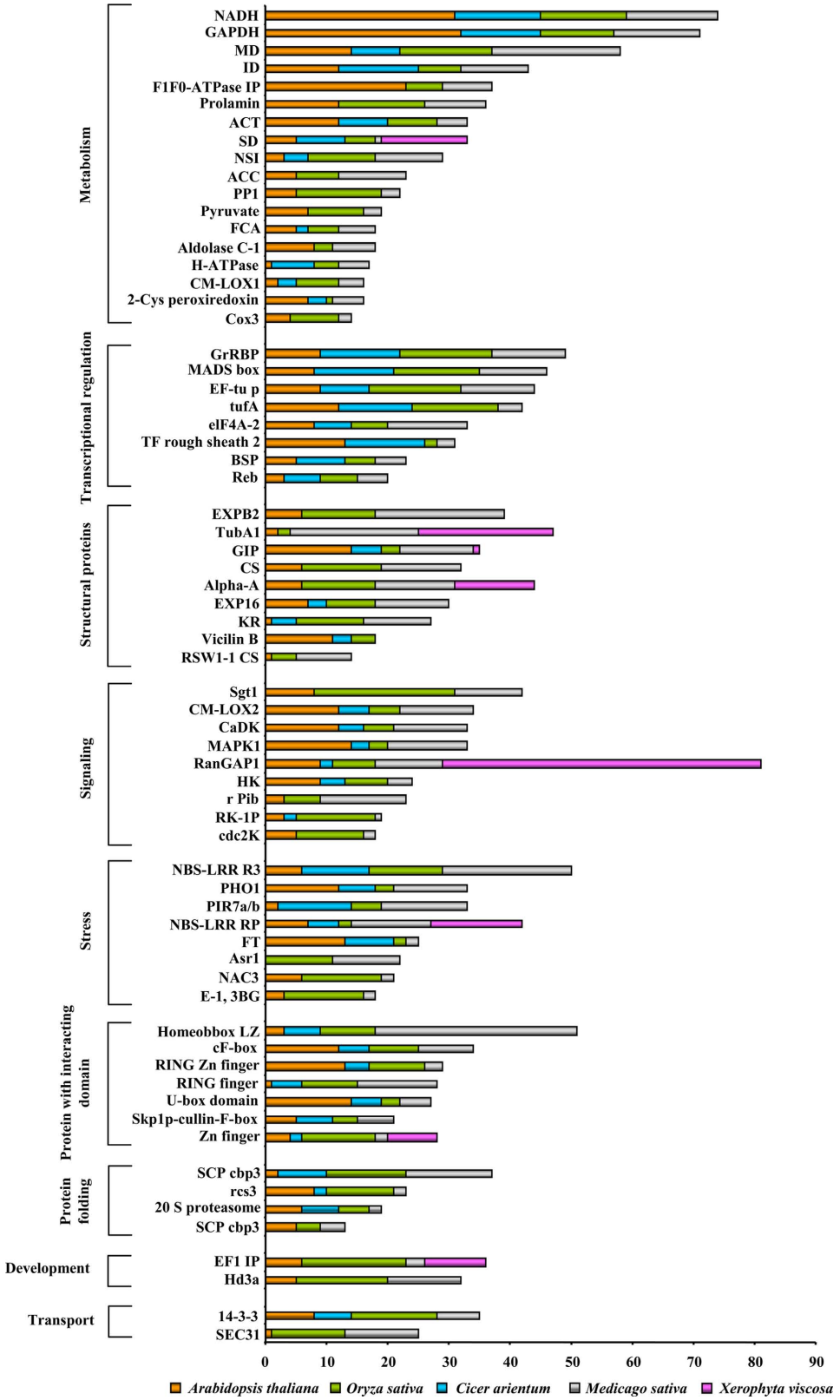


FIGURE 3 | Comparative abiotic stress-responsive nuclear proteome: The functional classification of the identified proteins was according to the biological processes in which they are involved. The length of the bars indicates the number of proteins present in a particular stress.

under varied stresses. Our analyses revealed the presence of monocot and dicot cdc-2k, SEC31, TubA1 having specific protein sequences that clearly demonstrate the diversity of the identical NPs in two divisions of angiosperm. This may be attributed to the evolution of orthologs vs. paralogs.

Responses to various patho-stresses largely depend on the plant's capacity to modulate rapidly but specifically its proteome. External signals are translocated into the nucleus in a stress-type dependent manner to activate transcription factors, resulting in the increased expression of particular sets of defense-related genes. During evolution, mutual recognition between plants and pathogens has resulted in development of fascinating variety of molecular strategies in the nucleus of the host against the invader. Some pathogens have been shown to directly activate transcription (Lev et al., 2005). It is now well accepted that modulation of chromatin configuration is an additional strategy employed by pathogen to subvert plant immune response (Ma et al., 2011). Nevertheless, plants also dispose an array of proteins in the nucleus that act as a scrutiny scheme to allow the early detection of an impending pathogen assault. We analyzed the nuclear proteomes of soybean and hot pepper in response to fungal (Cooper et al., 2011), and viral (Lee et al., 2006) stresses, respectively (**Figure 4**).

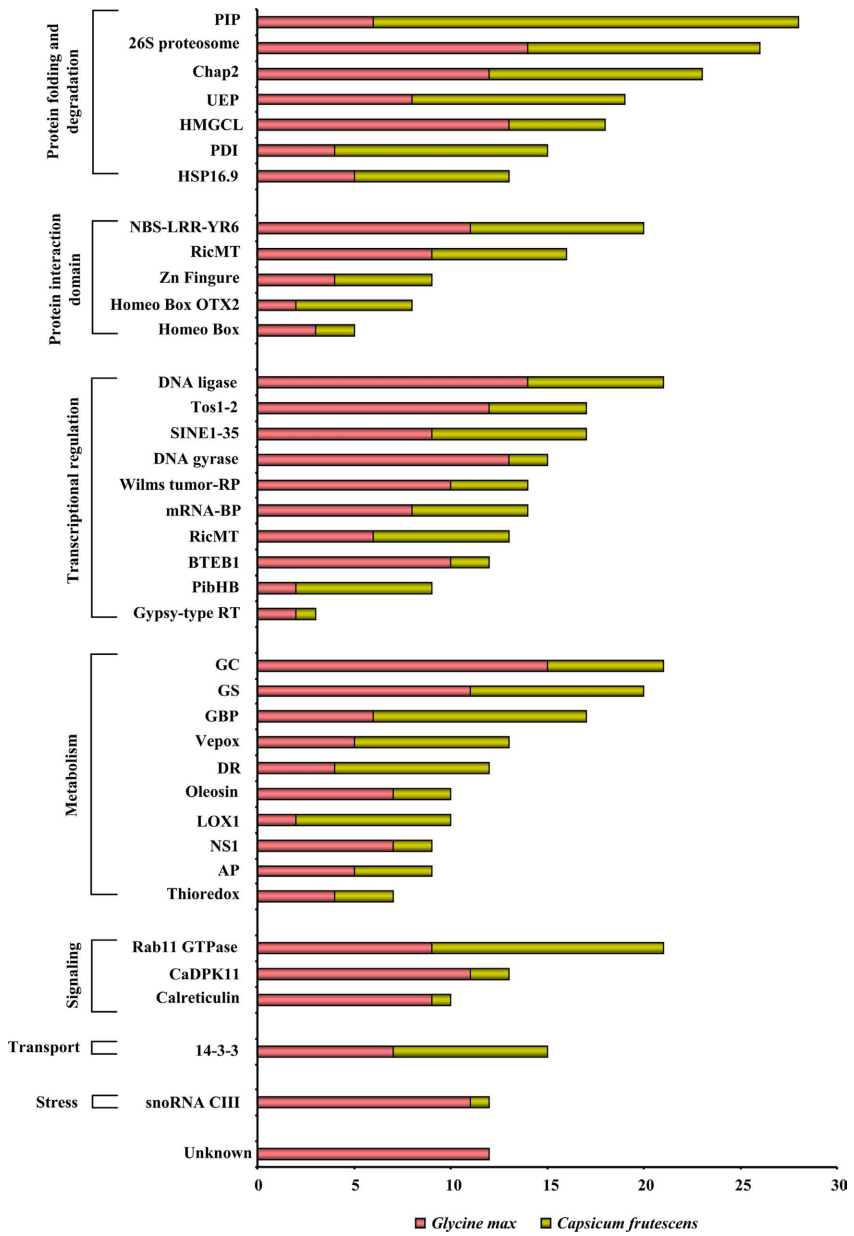


FIGURE 4 | Comparative patho-stress responsive nuclear proteome: The functional classification of the identified proteins was according to the biological processes in which they are involved. The length of the bars indicates the number of proteins present in a particular stress.

The widespread NPs identified in fungal and viral stresses belong to the category of protein folding and degradation. On the contrary, it was interesting enough to observe that there was not a single protein to be exclusive in case of either soybean-rust interaction or hot pepper–tobacco mosaic virus (TMV) interaction. During these host–pathogen interactions complex architecture of nucleus might respond differently against two different pathogens but using same set of NPs. Fungal stress and viral stress both might induce fundamental machinery of the nucleus to correctly target expressed proteins in a diverse but adaptation-related pathway thereby barricade the pathogens. However, NPs belonging to protein folding and degradation, transcription regulation, and metabolism categories toward patho-stress needs further consideration to understand the fungal-viral difference or specificity.

CONCLUSION

Since its existence was first discovered almost 180 years ago, the nucleus has been a central focus of biological research. Initially it was assumed that nucleus is a static organelle. Progress over the years has gradually changed this view, and more recently, the importance of the NPs in chromatin organization, gene regulation, and signal transduction has become evident. In this study, cross-kingdom, cross-species as well as cross-condition comparisons of nuclear proteomes in vascular plants and animals illustrates the divergence in protein profiles within only a few social classes. *In silico* experimental analyses of the nuclear interior revealed a morphologically structured yet dynamic mix of NPs. Major nuclear events depend on the functional integrity of protein species and their timely interaction. Yet, unknown drivers of protein ensure that they are in the right place at the time when they are needed. Furthermore, the incessant unrest of proteins can be captured by the comparative nuclear proteome study under various regulatory events. As expected, the proteins involved in transcriptional regulation and chromatin remodeling were found to be the most predominant across all conditions. Nonetheless, a large number of proteins were unique or novel to each of the clades and under different stresses. It may be thought, the ubiquitously present protein classes are essential for sustenance, while the unique classes bring out the condition-specific special function. The differences in terms of protein pattern and protein function appear to encompass both genetic and physiological information. It may be speculated that the differential proteome is shaped by the cellular environment and the ecological niche of the corresponding organism. The divergence may arise due to codon bias, amino acid composition, and protein length. A much more comprehensive survey of the nuclear proteomes in several plants will ultimately draw a more complete picture of the social class vs. protein diversity in this organelle.

MARCHING AHEAD: NEXT FIVE YEARS

We are witnessing a significant but inadequate progress in understanding the nuclear proteomes of various crops of agricultural importance. Our understanding of nuclear composition, organization, and homeostasis has been greatly enhanced through targeted biochemical and genetic approaches. Unbiased "discovery" methods, such as proteomics, have only recently gained traction in the field of regulation biology. To date, a key word search using "Plant nuclear proteome" retrieves only 116 results in a pubmed search, emphasizing the need for in-depth study in the field. Although our knowledge of nuclear proteome and NPs has greatly increased, many open ended questions remain to be answered. It is to be noted that few thousands NPs identified in the nuclear proteomes have not been functionally characterized. Thus, for this new and emerging field, we predict that the potential for an accelerated pace of future discoveries in nuclear cell biology is tremendously high. The future scientific interest should center around the diverse roles NPs play in regulating cell division, growth, differentiation, aging, disease, and environmental perturbations. The comparative analysis of organism, clade-specific and stress-responsive plant nuclear proteomes revealed the presence of certain proteins that were unexpected, either in their abundance, form, number or else localization. These unexpected or non-canonical proteins suggest the constant remodeling of nuclear proteomes. The exact function and specificity of these candidates can only be comprehended once they are functionally characterized. Furthermore, role of PTMs on gene expression and NP-interactome dynamics remains as two important but challenging facets. Our future efforts will focus on the development and analysis of comparative nuclear proteomes toward an understanding of crop- and genotype-specific adaptation as an important amendment for the determination of protein networks influenced by the internal and external cues associated with the complex cellular, biochemical and physiological process that bring about phenome variation.

ACKNOWLEDGMENTS

This research work was supported by grants from the Department of Biotechnology (DBT), Ministry of Science and Technology, Govt. of India (grant no. BT/PR10796/BRB/10/621/2008) and the National Institute of Plant Genome Research, India to S.C. K.N. is the recipient of pre-doctoral fellowship from Council of Scientific and Industrial Research (CSIR), Govt. of India. Authors thank Mr. Jasbeer Singh for illustrations and graphical representation in the manuscript.

REFERENCES

Abdalla, K. O., Baker, B., and Rafudeen, M. S. (2010). Proteomic analysis of nuclear proteins during dehydration of the resurrection plant *Xerophyta viscose*. *Plant Growth Regul.* 62, 279–292.

Abdalla, K. O., and Rafudeen, M. S. (2012). Analysis of the nuclear proteome of the resurrection plant *Xerophyta viscosa* in response to dehydration stress using iTRAQ with 2DLC and tandem mass spectrometry. *J. Proteomics* 18, 2361–2374.

Abe, H., Urao, T., Seiki, M., and Shinozaki, Y. (2003). *Arabidopsis* AtMYC2 and AtMYB2 function as transcription regulator in abscisic acid signaling. *Plant Cell* 15, 68–73.

Aki, T., and Yanagisawa, S. (2009). Application of rice nuclear proteome analysis to the identification of evolutionarily conserved and glucose-responsive nuclear proteins. *J. Proteome Res.* 8, 3912–3924.

Bae, M. S., Cho, E. J., Choi, E. Y., and Park, O. K. (2003). Analysis of the *Arabidopsis* nuclear proteome and its response to cold stress. *Plant J.* 36, 652–663.

Barthelery, M., Salli, U., and Vrana, K. E. (2008). Enhanced nuclear proteomics. *Proteomics* 8, 1832–1838.

Beven, A. F., Simpson, G. G., Brown, J. W., and Shaw, P. J. (1995). The organization of spliceosomal components in the nuclei of higher plants. *J. Cell Sci.* 108, 509–518.

Bhushan, D., Pandey, A., Chattopadhyay, A., Choudhary, M. K., Chakraborty, S., Datta, A., et al. (2006). Extracellular matrix proteome of chickpea (*Cicer arietinum* L.) illustrates pathway abundance, novel protein functions and evolutionary perspect. *J Proteome Res.* 5, 1711–1720.

Bickmore, W. A., and Sutherland, H. G. E. (2002). Addressing protein localization within the nucleus. *EMBO J.* 21, 1248–1254.

Boutilier, K., Offringa, R., Sharma, V. K., Kieft, H., Ouellet, T., Zhang, L. M., et al. (2002). Ectopic expression of BABY BOOM triggers a conversion from vegetative to embryonic growth. *Plant Cell* 14, 1737–1749.

Broers, J. L., Bronnenberg, N. M., Kuijpers, H. J., Schutte, B., Hutchison, C. J., and Ramaekers, F. C. (2002). Partial cleavage of A-type lamins concurs with their total disintegration from the nuclear lamina during apoptosis. *Eur. J. Cell Biol.* 81, 677–691.

Brown, J. W. S., Shaw, P. J., Shaw, P., and Marshall, D. F. (2005). *Arabidopsis* nucleolar protein database (AtNoPDB). *Nucleic Acids Res.* 33, 633–636.

Buhr, N., Carapito, C., Schaeffer, C., Kieffer, E., Van Dorsselaer, A., and Viville, S. (2008). Nuclear proteome analysis of undifferentiated mouse embryonic stem and germ cells. *Electrophoresis* 29, 2381–2390.

Calikowski, T. T., Meulia, T., and Meier, I. (2003). A proteomic study of the *Arabidopsis* nuclear matrix. *J. Cell. Biochem.* 90, 361–378.

Casati, P. (2012). Recent advances in maize nuclear proteomic studies reveal histone modifications. *Front. Plant Sci.* 3:278. doi: 10.3389/fpls.2012.00278

Casati, P., Campi, M., Chu, F., Suzuki, N., Maltby, D., Guan, S., et al. (2008). Histone acetylation and chromatin remodeling are required for UV-B-dependent transcriptional activation o fregulated genes in maize. *Plant Cell* 20, 827–842.

Caudron-Herger, M., and Rippe, K. (2012). Nuclear architecture by RNA. *Curr. Opin. Genet. Dev.* 22, 179–187.

Cavalier-Smith, T. (1988). Origin of the cell nucleus. *BioEssays* 9, 72–78.

Cavalier-Smith, T. (2002). The phagotrophic origin of eukaryotes and phylogenetic classification of protozoa. *Int. J. Syst. Evol. Microbiol.* 52, 297–354.

Cavalier-Smith, T. (2006). Cell evolution and earth history: stasis and revolution. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 361, 969–1006.

Cavalier-Smith, T. (2010). Origin of the cell nucleus, mitosis and sex: roles of intracellular coevolution. *Biol. Direct* 5, 7–14.

Cheung, A. Y., and Reddy, A. S. N. (2012). Nuclear architecture and dynamics: territories, nuclear bodies, and nucleocytoplasmic trafficking. *Plant Physiol.* 158, 23–25.

Choudhary, M. K., Basu, D., Datta, A., Chakraborty, N., and Chakraborty, S. (2009). Dehydration-responsive nuclear proteome of rice (*Oryza sativa* L.) illustrates protein network, novel regulators of cellular adaptation and evolutionary perspective. *Mol. Cell. Proteomics* 8, 1579–1598.

Constantinescu, D., Gray, H. L., Sammak, P. J., Schatten, G. P., and Csoka, A. B. (2006). Lamin A/C expression is a marker of mouse and human embryonic stem cell differentiation. *Stem Cells* 24, 177–185.

Cooper, B., Campbell, K. B., Feng, J., Garrett, W. M., and Frederick, R. (2011). Nuclear proteomic changes linked to soybean rust resistance. *Mol. Bisys.* 3, 773–783.

Cullen, Y. T., and Mansuy, I. M. (2010). Towards a better understanding of nuclear processes based on proteomics. *Amino Acids* 9, 1117–1130.

Dahl, K. N., Ribeiro, A. J. S., and Lammerding, J. (2008). Nuclear shape, mechanics, and mechanotransduction. *Circ. Res.* 102, 1307–1318.

De Souza, C. P., and Osmani, S. A. (2009). Double duty for nuclear proteins – the price of more open forms of mitosis. *Trends Genet.* 25, 545–554.

Dundr, M., and Misteli, T. (2001). Functional architecture in the cell nucleus. *Biochem. J.* 356, 297–310.

Erhardt, M., Adamska, I., and Luiz Franco, O. (2010). Plant nuclear proteomics – inside the cell maestro. *FEBS J.* 16, 3295–3307.

Feng, B., Yang, S., Du, H., Hou, X., Zhang, J., Liu, H., et al. (2012). Molecular characterization and functional analysis of plant WRKY genes. *African J. Biotechnol.* 11, 13606–13613.

Ferreira, B. S., Branco, A. T., deOliveira, M. A., Pereira, M. G., and de Souza Filho, G. A. (2006). Methodological improvements on extraction of nuclear proteins and its preliminary analysis during the maize (*Zea mays* L.) endosperm development. *Protein Pept. Lett.* 13, 981–984.

Franklin, S., Zhang, M. J., Chen, H., Paulsson, A. K., Jordan, S. S. M., Li, Y., et al. (2011). Specialized compartments of cardiac nuclei exhibit distinct proteomic anatomy. *Proteomics* 10, 1074–1015.

Fritz-Laylin, L. K., Prochnik, S. E., Ginger, M. L., Dacks, J. B., Carpenter, M. L., Field, M. C., et al. (2010). The genome of Naegleria gruberi illuminates early eukaryotic versatility. *Cell* 140, 631–642.

Gao, G., Zhong, Y., Guo, A., Zhu, Q., Tang, W., Zheng, W., et al. (2006). DRTF: a database of rice transcription factors. *Bioinformatics* 15, 1286–1287.

Gauci, S., Veenhoff, L. M., Heck, A. J. R., and Krijgsveld, J. (2009). Orthogonal separation techniques for the characterization of the yeast nuclear proteome. *J. Proteome Res.* 8, 3451–3463.

Gonzalez-Melendi, P., Beven, A., Boudonck, K., Abranches, R., Wells, B., Dolan, L., et al. (2000). The nucleus: a highly organized but dynamic structure. *J. Microsc.* 198, 199–207.

Gorski, S., and Misteli, T. (2005). Systems biology in the cell nucleus. *J. Cell Sci.* 118, 4083–4092.

Guo, A., He, K., Liu, D., Bai, S., Gu, X., Wei, L., et al. (2005). DATF: a database of *Arabidopsis* transcription factors. *Bioinformatics* 15, 2568–2569.

Guo, A. Y., Chen, X., Gao, G., Zhang, H., Zhu, Q. H., Liu, X. C., et al. (2008). PlantTFDB: a comprehensive plant transcription factor database. *Nucleic Acids Res.* 36, 966–969.

Henrich, S., Cordwell, S. J., Crossett, B., Baker, M. S., and Christopherson, R. I. (2007). The nuclear proteome and DNA-binding fraction of human Raji lymphoma cells. *Biochim. Biophys. Acta* 1774, 413–432.

Hetzer, M. W., Walther, T. C., and Mattaj, I. W. (2005). Pushing the envelope: structure, function, and dynamics of the nuclear periphery. *Annu. Rev. Cell Dev. Biol.* 21, 347–380.

Holmes-Davis, R., and Comai, L. (1998). Nuclear matrix attachment regions and plant gene expression. *Trends Plant Sci.* 3, 91–97.

Iida, K., Seki, M., Sakurai, T., Satou, M., Akiyama, T. T., Konagaya, A., et al. (2005). RARTF: database and tools for complete sets of *Arabidopsis* transcription factors. *DNA Res.* 12, 247–256.

Jiang, G., Huang, A. H., Cai, Y., Tanase, M., and Sheetz, M. P. (2006). Rigidity sensing at the leading edge through alphavbeta3 integrins and RPTPalpha. *Biophys. J.* 90, 1804–1809.

Khan, M. K. K., and Komatsu, S. (2004). Rice proteomics: recent developments and analysis of nuclear proteins. *Phytochemistry* 65, 1671-1681

Lee, B. J., Kwon, S. J., Kim, S. K., Kim, K. J., Park, C. J., Kim, Y. J., et al. (2006). Functional study of hot pepper 26S proteasome subunit RPN7 induced by tobacco mosaic virus from nuclear proteome analysis. *Biochem. Biophy. Res. Comm.* 351, 405–411.

Lev, S., Hadar, R., Amedeo, P., Baker, S. E., Yoder, O. C., and Horwitz, B. A. (2005). Activation of an AP1-Like transcription factor of the maize pathogen *Cochliobolus heterostrophus* in response to oxidative stress and plant signals. *Eukaryot. Cell* 4, 443–454.

Li, G., Nallamilli, B. R., Tan, F., and Peng, Z. (2008). Removal of high abundance proteins for nuclear subproteome studies in rice (*Oryza sativa*) endosperm. *Electrophoresis* 29, 604–617.

Liao, L., McClatchy, D. B., and Yates, J. R. (2009). Shotgun proteomics in neuroscience. *Neuron* 63, 12–26.

Ma, K. W., Flores, C., and Ma, W. (2011). Chromatin configuration as a battlefield in plant-bacteria interaction. *Plant Physiol.* 157, 535–543.

Mans, B. J., Anantharaman, V., Aravind, L., and Koonin, E. V. (2004). Comparative genomics, evolution and origins of the nuclear envelope and nuclear pore complex. *Cell Cycle* 3, 1612–1637.

Mayrose, M., Ekengren, S. K., Melech-Bonfil, S., Martin, G. B., and Sessa, G. (2006). A novel link between tomato GRAS genes, plant disease resistance and mechanical stress response. *Mol Plant Pathol.* 7, 593–604.

McClatchy, D. B., Liao, L., Park, S. K., Xu, T., Lu, B., and Yates, J. R. (2011). Differential proteomic analysis of mammalian tissues using SILAM. *PLoS ONE* 6:e16039. doi: 10.1371/journal.pone.0016039

Meagher, R. B., McKinney, E. C., and Kandasamy, M. K. (1999). Isovariant dynamics expand and buffer the responses of complex systems: the diverse plant actin gene family. *Plant Cell* 11, 995–1005.

Misteli, T. (2001). Protein dynamics: implications for nuclear architecture and gene expression. *Science* 291, 843–847.

Misteli, T. (2005). Concepts in nuclear architecture. *BioEssays* 27, 477–487.

Moriguchi, K., Suzuki, T., Ito, Y., Yamazaki, Y., Niwa, Y., and Kurataa, N. (2005). Functional isolation of novel nuclear proteins showing a variety of subnuclear localizations. *Plant Cell* 17, 389–403.

Narula, K., Elagamey, E., Datta, A., Chakraborty, N., and Chakraborty, S. (2012). "Comparative analyses of extracellular matrix proteome:

an under-explored area in plant research," in *Crop Plants*, ed. A. Goyal (Janeza Tradine, Croatia: InTech), 145–166.

Pajerowski, J. D., Dahl, K. N., Zhong, F. L., Sammak, P. J., and Discher, D. E. (2007). Physical plasticity of the nucleus in stem cell differentiation. *Proc. Natl. Acad. Sci. U.S. A.* 104, 15619 –15624.

Palaniswamy, S. K., James, S., Sun, H., Lamb, R. S., Davuluri, R. V., and Grotewold, E. (2006). AGRIS and AtRegNet. *A* platform to link cis-regulatory elements and transcription factors into regulatory networks. *Plant Physiol.* 140, 818–829.

Pandey, A., Chakraborty, S., Datta, A., and Chakraborty, N. (2008). Proteomics approach to identify dehydration responsive nuclear proteins from chickpea (*Cicer arietinum* L.). *Mol. Cell. Proteomics* 7, 88–107.

Pandey, A., Choudhary, M. K., Bhushan, D., Chattopadhyay, A., Chakraborty, S., Datta, A., et al. (2006). The nuclear proteome of chickpea (*Cicer arietinum* L.) reveals predicted and unexpected proteins. *J. Proteome Res.* 5, 3301–3311.

Pederson, T. (2000). Diffusional protein transport within the nucleus: a message in the medium. *Nat. Cell Biol.* 2, E73–E74.

Pendle, A. F., Clark, G. P., Boon, R., Lewandowska, D., Lam, Y. W., Andersen, J., et al. (2005). Proteomic analysis of the *Arabidopsis* nucleolus suggests novel nucleolar functions. *Mol. Biol. Cell* 16, 260–269.

Pennisi, E. (2004). The birth of the nucleus. *Science* 305, 766–768.

Perez-Rodriguez, P., Riano-Pachon, D. M., Correa, L. G., Rensing, S. A., Kersten, B., and Roeber, M. B. (2010). PlnTFDB: updated content and new features of the plant transcription factor database. *Nucleic Acids Res.* 38, 822–827.

Renouard, S., Cyrielle, C., Lopez, T., Lamblin, F., Lainé, E., and Hano, C. (2012). Isolation of nuclear proteins from flax (*Linumusitatissimum* L.) seed coats for gene expression regulation studies. *BMC Res. Notes* 5:15. doi: 10.1186/1756-0500-5-15

Repetto, O., Rogniaux, H., Firnhaber, C., Zuber, H., Küster, H., Larré, C., et al. (2008). Exploring the nuclear proteome of *Medicago truncatula* at the switch towards seed filling. *Plant J.* 56, 398–410.

Repetto, O., Rogniaux, H., Larré, C., Thompson, R., and Gallardo, K. (2012). The seed nuclear proteome. *Front Plant Sci.* 3:289. doi: 10.3389/fpls.2012.00289

Rivas, S. (2012). Nuclear dynamics during plant innate immunity. *Plant Physiol.* 158, 87–94.

Roix, J., and Misteli, T. (2002). Genomes, proteomes, and dynamic networks in the cell nucleus. *Histochem. Cell Biol.* 118, 105–116.

Romeuf, I., Tessier, D., Dardevet, M., Branlard, G., Charmet, G., and Ravel, C. (2010). wDBTF: an integrated database resource for studying wheat transcription factor families. *BMC Genomics* 11:185. doi: 10.1186/1471-2164-11-185

Rushton, P. J., Bokowiec, M. T., Laudeman, T. W., Brannock, J. F., Chen, X., and Timko, M. P. (2008). TOBFAC: the database of tobacco transcription factors. *BMC Bioinformatics* 9:53. doi: 10.1186/1471-2105-9-53

Salzano, A. M., Paron, I., Pines, A., Bachi, A., Talamo, F., Bivi, N., et al. (2006). Differential proteomic analysis of nuclear extracts from thyroid cell lines. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 833, 41–50.

Schirmer, E. C., and Gerace, L. (2005). The nuclear membrane proteome: extending the envelope. *Trends Biochem. Sci.* 30, 551–558.

Semple, C. A., Riken, G. E. R., and Members, G. S. L. (2003). The comparative proteomics of ubiquitination in mouse. *Genome Res.* 13,1389–1394.

Skovgaard, M., Jensen, L. J., Brunak, S., Ussery, D., and Krogh, A. (2001). On the total number of genes and their length distribution in complete microbial genomes. *Trends Genet.* 17, 425–428.

Spector, D. L. (2003). The dynamics of chromosome organization and gene regulation. *Annu. Rev. Biochem.* 72, 573–608.

Strouboulis, J., and Wolffe, A. P. (1996). Functional compartmentalization of the nucleus. *J. Cell Sci.* 109, 1991–2000.

Su, P. A., and Li, H. M. (2008). *Arabidopsis* stromal 70-kD heat shock proteins are essential for plant development and important for thermotolerance of germinating seeds. *Plant Physiol.* 146, 1231–1241.

Takizawa, T., and Meshorer, E. (2008). Chromatin and nuclear architecture in the nervous system. *Trends Neurosci.* 31, 345–355.

Tan, F., Li, G., Chitteti, B. R., and Peng, Z. (2007). Proteome and phosphoproteome analysis of chromatin associated proteins in rice (*Oryza sativa*). *Proteomics* 7, 4511–4527.

Tan, F., Zhang, K., Mujahid, H., Verma, D. P., and Peng, Z. (2010). Differential histone modification and protein expression associated with cell wall removal and regeneration in rice (*Oryza sativa*). *J. Proteome Res.* 10, 551–563.

Truncaite, L., Zajanckauskaite, A., Arlauskas, A., and Nivinskas, R. (2006). Transcription and RNA processing during expression of genes preceding DNA ligase gene 30 in T4-related bacteriophages. *Virology* 344, 378–390.

van Driel, R., and Fransz, P. (2004). Nuclear architecture and genome functioning in plants and animals: what can we learn from both? *Exp. Cell Res* 296, 86–90.

Varma, P., and Mishra, R. K. (2011). Dynamics of nuclear matrix proteome during embryonic development in *Drosophila melanogaster*. *J. Biosci.* 36, 439–459.

Vergnes, L., Peterfy, M., Bergo, M. O., Young, S. G., and Reue, K. (2004). Lamin B1 is required for mouse development and nuclear integrity. *Proc. Natl. Acad. Sci. U.S.A.* 101, 10428–10433.

Wang, Z., Libault, M., Joshi, T., Valliyodan, B., Nguyen, H. T., Xu, D., et al. (2010). SoyDB: a knowledge database of soybean transcription factors. *BMC Plant Biol.* 10:14. doi: 10.1186/1471-2229-10-14

Wilson, K. L., and Dawson, S. C. (2011). Functional evolution of nuclear structure. *J. Cell Biol.* 195, 171–181.

Winck, F. V., Riano-Pacho, D. M., Sommer, F., Rupprecht, J., and Roeber, B. M. (2011). The nuclear proteome of the green alga Chlamydomonas reinhardtii. *Proteomics* 12, 95–100.

Yan, W., Hwang, D., and Aebersold, R. (2008). Quantitative proteomic analysis to profile dynamic changes in the spatial distribution of cellular proteins. *Methods Mol. Biol.* 432, 389–401.

Yates, J. R., Gilchrist, A., Howell, K. E., and Bergeron, J. J. (2005). Proteomics of organelles and large cellular structures. *Nat. Rev. Mol. Cell Biol.* 6, 702–714.

Yu, Y., Xu, W., Wang, S., Xu, Y., Li, H., Wang, Y., et al. (2011). VpRFP1, a novel C4C4-type RING finger protein gene from Chinese wild Vitis pseudoreticulata, functions as a transcriptional activator in defence response of grapevine. *J Exp. Bot.* 62, 671–682.

Zhu, Q. H., Guo, A. Y., Gao, G., Zhong, Y. F., Xu, M., Huang, M., and Luo, J. (2007). DPTF: a database of poplar transcription factors. *Bioinformatics* 15, 1307–1308.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 15 January 2013; paper pending published: 01 February 2013; accepted: 30 March 2013; published online: 26 April 2013.

Citation: Narula K, Datta A, Chakraborty N and Chakraborty S (2013) Comparative analyses of nuclear proteome: extending its function. Front. Plant Sci. ***4****:100. doi: 10.3389/fpls.2013.00100*

This article was submitted to Frontiers in Plant Proteomics, a specialty of Frontiers in Plant Science.

Copyright © 2013 Narula, Datta, Chakraborty and Chakraborty. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.

frontiers in
PLANT SCIENCE

MINI REVIEW ARTICLE
published: 12 December 2012
doi: 10.3389/fpls.2012.00278

Recent advances in maize nuclear proteomic studies reveal histone modifications

Paula Casati*

Centro de Estudios Fotosintéticos y Bioquímicos, Universidad Nacional de Rosario, Rosario, Santa Fe, Argentina

Edited by:
Nicolas L. Taylor, The University of Western Australia, Australia

Reviewed by:
Brian Mooney, University of Missouri, USA
Ghasem Hosseini Salekdeh, Agricultural Biotechnology Research Institute of Iran, Iran

****Correspondence:***
Paula Casati, Centro de Estudios Fotosintéticos y Bioquímicos, Universidad Nacional de Rosario, Suipacha 531, Rosario, Santa Fe 2000, Argentina.
e-mail: casati@cefobi-conicet.gov.ar

The nucleus of eukaryotic organisms is highly dynamic and complex, containing different types of macromolecules including DNA, RNA, and a wide range of proteins. Novel proteomic applications have led to a better overall determination of nucleus protein content. Although nuclear plant proteomics is only at the initial phase, several studies have been reported and are summarized in this review using different plants species, such as *Arabidopsis thaliana,* rice, cowpea, onion, garden cress, and barrel clover. These include the description of the total nuclear or phospho-proteome (i.e., *Arabidopsis*, cowpea, onion), or the analysis of the differential nuclear proteome under different growth environments (i.e., *Arabidopsis*, rice, cowpea, onion, garden cress, and barrel clover). However, only few reports exist on the analysis of the maize nuclear proteome or its changes under various conditions. This review will present recent data on the study of the nuclear maize proteome, including the analysis of changes in posttranslational modifications in histone proteins.

Keywords: *Zea mays*, nuclei, histones, posttranslational modification, mass spectroscopy

INTRODUCTION

The eukaryotic nucleus is highly dynamic and complex, it has several subcompartments, types of DNA and RNA, and a wide range of proteins. Different novel proteomic applications have led to a better overall determination of nucleus protein content, enabling researchers to analyze protein–protein interactions, structures, activities, and even posttranslational modifications (Erhardt et al., 2010).

Although nuclear plant proteomics is only at the initial phase, several studies have been reported using different plants species. However, there are a low number of publications in the field using maize (*Zea mays*). These low numbers of publications have also mainly focused on a small number of nuclear proteins, the histones. Because insights from any species can trigger ideas for research in other plants, and due to the lack of research in this area at present on maize nuclear proteomics and its limited scope focused around histones, in this review, examples of experiments completed using different plant species are presented, including recent data on the study of the nuclear maize proteome, and in particular the histones.

EXAMPLES OF PLANT NUCLEAR PROTEOME ANALYSIS

One example of an analysis of changes in the nuclear proteome, and particularly in the phospho-proteome of a plant species was done using onion nuclei (González-Camacho and Medina, 2004). To detect variations associated with cell proliferation, two-dimensional proteomes from soluble fractions of onion nuclei isolated from actively proliferating meristematic and non-meristematic root cells were compared (González-Camacho and Medina, 2004). Interestingly, the nucleolin-like protein NopA100 was significantly increased in proliferating cells, and Western blots with anti-NopA100 antibody demonstrated 26 spots in the meristematic sample corresponding to this protein. All the spots detected were clustered at 100 kDa, suggesting NopA100 was differentially phosphorylated, and that the protein was more highly phosphorylated in cycling cells (González-Camacho and Medina, 2004).

The *Arabidopsis* nuclear proteomes of plants under different growth conditions has also been analyzed. In response to cold stress, 54 out of 184 protein spots were changed in response to cold treatment in two-dimensional gels (2D-PAGE; Bae et al., 2003); while in the presence of oligogalacturonides, elicitors of plant defense responses, significant changes in protein abundance for 19 proteins were reported (Casasoli et al., 2007). Proteins responding to the oligogalacturonide treatment were involved in the protein translation machinery and regulation, suggesting a general reprogramming of the plant cell metabolism in response to oligogalacturonides (Casasoli et al., 2007). The proteome of the *Arabidopsis* nucleoli was also analyzed. The eukaryotic nucleolus is involved in ribosome biogenesis and a range of other RNA metabolism and cellular functions; in this compartment 217 proteins were identified by proteome analysis (Pendle et al., 2005). The comparison of the proteomes of the arabidopsis and the human nucleoli identified many common proteins, plant-specific proteins, proteins of unknown function in both proteomes, and proteins that were nucleolar in plants but non-nucleolar in human, suggesting that in plants, nucleoli may have additional functions in mRNA export or surveillance (Pendle et al., 2005).

A proteome reference map of a legume, chickpea, was completed using 2D-PAGE (Pandey et al., 2006). Approximately, 600 protein spots were detected and LC-ESI-MS/MS analyses led to the identification of 150 proteins that have been implicated in different cellular functions. These included proteins involved in signaling, gene regulation, DNA replication, and transcription (Pandey et al., 2006). Besides, the nuclear proteome of chickpea seedlings under dehydration conditions was compared to that of control plants

using 2D-PAGE (Pandey et al., 2007). MS analysis allowed the identification of 147 differentially expressed proteins involved in various functions, including gene transcription and replication, molecular chaperones, cell signaling, and chromatin remodeling (Pandey et al., 2007). A similar study was done using a draught tolerant rice variety, identifying 150 proteins that showed changes in their levels (Choudhary et al., 2009). The proteomic analysis led to the identification of differentially regulated proteins involved in transcriptional regulation and chromatin remodeling, signaling and gene regulation, cell defense and rescue, and protein degradation. Furthermore, a comparison between the dehydration responsive nuclear proteome of rice and that of chickpea, showed an evolutionary divergence in dehydration response, with only a few conserved proteins (Choudhary et al., 2009). Using rice, a nuclear proteome analysis was used to search for novel nuclear proteins that could play evolutionarily conserved roles in the sugar response in plants (Aki and Yanagisawa, 2009). Five hundred sixty-three different proteins were identified by nanoLC/ESI/MS/MS analysis of extracts from rice nuclei that were purified by Percoll density gradient centrifugation, whereas 307 different proteins were identified with nucleic acid-associated proteins that were enriched by DNA affinity chromatography (Aki and Yanagisawa, 2009). Among them, transcription and splicing factors were identified, but also a mediator of sugar signaling in plants, hexokinase.

The nuclear proteome of *Medicago truncatula* 12 days after pollination (dap) was also analyzed; this stage marks the switch toward seed filling (Repetto et al., 2008). Nano-liquid chromatography–tandem mass spectrometry analysis of nuclear protein bands excised from 1D SDS-PAGE identified 179 polypeptides, providing an insight into the complexity and distinctive feature of the seed nuclear proteome, and highlighting new plant nuclear proteins with possible roles in the biogenesis of ribosomal subunits or nucleocytoplasmic trafficking (Repetto et al., 2008). To identify proteins that contribute to disease resistance in soybean, the nuclear proteome from a susceptible cultivar was compared to that of a resistant inbred isoline (Cooper et al., 2011). About 4975 proteins from nuclear preparations of leaves were detected using a high-throughput liquid chromatography-mass spectrometry method. Statistics of summed spectral counts revealed proteins with differential accumulation changes between susceptible and resistant plants; however, these protein accumulation changes were compared to previously reported gene expression changes and very little overlap was found. Thus, it appears that numerous proteins are posttranslationally affected in the nucleus after infection (Cooper et al., 2011). Finally, the nuclear proteome of the unicellular green alga *Chlamydomonas reinhardtii* was also analyzed (Winck et al., 2012). Using LC-MS/MS, 672 proteins from nuclei isolates were identified. Well-known proteins like histones, transcription factors and other transcriptional regulators were identified (Winck et al., 2012).

Only few reports exist on the analysis of the maize nuclear proteome or its changes under various conditions. Next, we will present recent data on the study of the nuclear maize proteome, including the analysis of changes in posttranslational modifications in histone proteins.

MAIZE NUCLEAR PROTEOME STUDIES

A comparison of the maize nuclear proteomes after a UV-B light treatment was done using maize lines that differ in UV-B tolerance by 2D-PAGE (Casati et al., 2008). Eight maize lines with documented differences in UV-B tolerance were employed. The W23 UV-B light-sensitive line, deficient in flavonoid sunscreens, and high-altitude Confite Puneño and Mishca lines selected in their natural environment for UV-B tolerance were compared. In addition, four hypersensitive lines expressing RNAi constructs to reduce the expression of predicted chromatin remodeling genes (*chc101*, *nfc102*, *sdg102*, and *mbd101*) were compared to the B73 line. Fluorescently labeled proteins were resolved by isoelectric focusing on a 3–10 pH gradient and in the second dimension by molecular weight using PAGE. Approximately 500 proteins were resolved; most protein spots were present in all genotypes; there were only a few cases in which spots were present in some lines and not in others (Casati et al., 2008). Differential accumulation of chromatin proteins, particularly histones, constituted the largest class identified by mass spectrometry; other DNA- and chromatin-associated proteins, and several ribosomal proteins were also identified.

UV-B-tolerant landraces and the B73 inbred line showed twice as many protein changes as the UV-B-sensitive W23 line and transgenic maize expressing RNAi constructs directed against chromatin factors (Casati et al., 2008). Although many changes were line-specific, reflecting the distinctive germplasm of the two high-altitude lines and of W23, it was clear that UV-B-tolerant lines exhibit more nuclear proteome changes than do sensitive lines. For example, the high-altitude Confite and Mischa showed 42 and 31 protein changes, respectively, while the sensitive W23 line showed 21 protein spots changed by the treatment. Paralleling the conclusion based on all proteins, more changes in histone proteins were found in the high-altitude lines than in W23. Similarly, more histone isotypes were differentially accumulated in B73 than in the near-isogenic RNAi transgenic lines.

ANALYSIS OF HISTONES AND HISTONE COVALENT MODIFICATIONS IN MAIZE NUCLEI

For some histones, the same protein was identified in different spots in one gel (Casati et al., 2008). The presence of the same protein type at multiple spots could reflect either differential expression of loci encoding different proteins or posttranslational regulation of the same gene product. Histones are subject to numerous covalent modifications, such as acetylation, methylation, phosphorylation, and ubiquitination, and these modifications control many aspects of chromatin function mediated by histones (Kouzarides, 2007). To analyze histone composition systematically by MS, histones were acid extracted from UV-B-treated or control B73 leaves and then the four core histones were separated by reverse phase HPLC. A direct comparison of histones H2B, H2A, H4, and H3 at the protein level did not reveal any noticeable difference between the UV-B-treated and control samples. Nevertheless, substantial changes were observed in some acetylated peptides at the N-terminal tails of H4 and H3. The acetylated H4 N-terminal tail, for example, was approximately doubled in UV-B-exposed samples compared with the control (Casati et al., 2008). Tryptic peptides of histone H3 were analyzed in a similar manner;

an N-terminal tail acetylated peptide was observed to be considerably increased in UV-B-treated samples as well (**Figure 1A**). In contrast, the level of detected methylations remained essentially unchanged after UV-B treatment (**Figure 1A**). Thus, for both H3 and H4, the differential intensities of various isoforms observed by 2D-PAGE were explained at least in part by post-translational modification levels that change in response to UV-B treatment, and acetylation in the N-terminal tail of H3 and H4 is the most significantly altered epigenetic mark after UV-B treatment. These acetylated histones were enriched in the promoter and transcribed regions of two UV-B-upregulated genes examined; radiation-sensitive lines lack this enrichment (Casati et al., 2008).

The increase in histone acetylation by UV-B was also demonstrated by Western blot analysis (**Figure 1B**; Campi et al., 2012). Using antibodies against acetylated histone H3 in the N-terminal domain, maize plants from the B73 genotype showed increased acetylation of this histone after a UV-B treatment, similarly as shown by MS analysis (**Figure 1A**; Casati et al., 2008). However, the RNAi *chc101*, *nfc102*, *sdg102*, and *mbd101* chromatin remodeling deficient plants showed lower levels of H3 acetylation after the UV-B treatment than wild-type plants (**Figure 1B**). Together, these experiments demonstrated that chromatin remodeling, and in particular histone acetylation, are important for UV-B responses in maize. In particular, UV-B radiation can also induce different chromatin remodeling events in the promoter regions of *Mutator* transposons (Qüesta et al., 2010). Increased transcript abundance of the *mudrA* transposase and *mudrB*, an unknown gene encoded in *MuDR*, the master *Mutator* transposon, is accompanied by an increase in histone H3 acetylation and by decreased H3K9me2 methylation (Qüesta et al., 2010). To date, only radiation treatments such as UV-B have reactivated silenced *Mutator*. Therefore, transposon reactivation by UV-B requires epigenetic changes, suggesting that early changes in H3 methylation and chromatin remodeling contribute directly to transposon reactivation by UV-B in maize (Qüesta et al., 2010).

So far, changes in histone covalent modifications have been the most extensively studied posttranslational modifications in maize nuclei. A method for the reliable and sensitive detection of specific chromatin modifications on selected genes has been described (Jaskiewicz et al., 2011). The technique is based on the crosslinking of modified histones and DNA with formaldehyde, extraction and sonication of chromatin, chromatin immunoprecipitation with modification-specific antibodies, de-crosslinking of histone-DNA complexes, and gene-specific real-time quantitative PCR. This approach has proven useful for detecting specific histone modifications associated with C4 photosynthesis in maize (Jaskiewicz et al., 2011). In addition, the relative abundance of eight different histone modifications was tested at various regions in several imprinted maize genes using a chromatin immunoprecipitation protocol coupled with quantitative allele-specific single nucleotide polymorphism assays (Haun and Springer, 2008). Imprinting is an epigenetically controlled form of gene regulation in which the expression of a gene is based on its parent of origin. This epigenetic regulation is likely to involve allele-specific DNA or histone modifications. In this work, histone H3 lysine-27 di- and tri-methylation were paternally enriched at three imprinted loci. In contrast, acetylation of histones H3 and H4 and H3K4 dimethylation were enriched at the maternal alleles of these genes. Di- and tri-methylation of H3 lysine-9, which is generally associated with constitutively silenced chromatin, was not enriched at either allele of imprinted loci. These patterns of enrichment were specific to tissues that exhibit imprinting; in addition, the enrichment of these modifications was dependent upon the parental origin of an allele (Haun and Springer, 2008). Changes in the covalent modification of histones were also demonstrated during paramutation (Haring et al., 2010). Paramutation is the transfer of epigenetic information between alleles that leads to a heritable change in expression of one of these alleles. Paramutation at the tissue-specifically expressed

A

Peptide	Fold increase (UV-B/C)
$^{9}K_{ac}STGGK_{ac}APR^{17}$	2.25
$^{57}STELLIR^{63}$	1
$^{27}KSAPATGGVK^{36}$	1
$^{27}K_{me1}SAPATGGVK^{36}$	1
$^{27}K_{me2}SAPATGGVK^{36}$	1

B

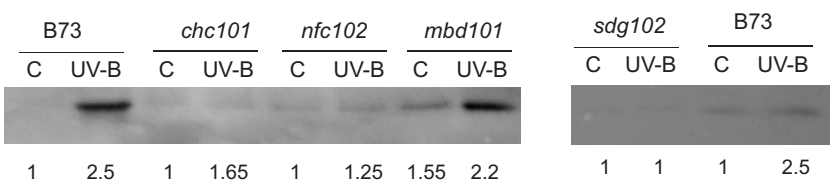


FIGURE 1 | Analysis of histone H3 modifications in maize after a UV-B treatment. (A) Average ratios of peak areas integrated from LC-MS runs of representative peptides detected from the duplicates of UV-B-treated and the control in the absence of UV-B histone H3. The peptide 9-KacSTGGKacAPR-17 is more abundant in the UV-B-exposed sample, while non-covalent modified or methylated peptides are not changed by the treatment. **(B)** Western blot analysis of histone extracts from maize plants revealed using antibodies against acetylated H3 in the N-terminal domain. Quantification of the bands determined by densitometrical analysis of the western blots is shown below each band. Data adapted from Casati et al. (2008) and Campi et al. (2012).

maize *b1* locus involves the low-expressing *B′* and high-expressing *B-I* allele. A hepta-repeat located 100 kb upstream of the *b1* coding region was required for paramutation, and nucleosome occupancy, H3 acetylation, and H3K9 and H3K27 methylation were mainly involved in tissue-specific regulation of the hepta-repeat (Haring et al., 2010).

A different example of a proteomic analysis of histones was done using maize kernels (Kalamajka et al., 2010). In maize kernel development, the onset of grain-filling represents a major developmental switch that correlates with a massive reprogramming of gene expression. In this study, the linker histones of developing maize kernel tissue were compared. Chromosomal linker histones from developing maize kernels before (11 dap) and after (16 dap) initiation of storage synthesis were isolated, and six linker histone gene products were identified by MALDI-TOF mass spectrometry. The acid soluble histones were separated by 2D-PAGE, several of the spots corresponding to the linker histone bands were excised from the gels, in gel digested with trypsin, and examined by MALDI-TOF MS. Peptide fragmentation data were obtained by MS/MS analysis of intense peaks in the MS spectra (Kalamajka et al., 2010). The linker histones HON101, HON102, HON103, HON104, HON106, and HON110 were identified; interestingly, the majority of the linker histones from the 11 dap endosperm were found to migrate in the 2D gels with a lower p*I* than those from the 16 dap endosperm sample. Since the same gene products were identified in gel spots from 11 to 16 dap, the difference in linker histone p*I*s is most likely to be due to differential posttranslational modification(s) of the proteins during kernel development (Kalamajka et al., 2010).

As previously shown, histone modifications have also been analyzed by immunological techniques. To investigate the mitosis-dependent cross-talk between histone H4 tetra-acetylation, DNA methylation, and H3K9 dimethylation, specific antibody immunostaining in Western blot analysis and in *in situ* chromatin immunostaining were used to detect and compare H4ac, H3K9me2, and DNA methylation patterns during mitosis in maize root meristems (Yang et al., 2010). Treatment with trichostatin A, which inhibits histone deacetylases, resulted in increased histone H4 acetylation accompanied by the decondensation of interphase chromatin and a decrease in both global H3K9 dimethylation and DNA methylation during mitosis. These observations suggest that histone acetylation may affect DNA and histone methylation during mitosis. Treatment with 5-azacytidine, a cytosine analog that reduces DNA methylation, caused chromatin decondensation and mediated an increase in H4 acetylation, in addition to reduced DNA methylation and H3K9 dimethylation during interphase and mitosis, suggesting that decreased DNA methylation causes a reduction in H3K9 dimethylation and an increase in H4 acetylation (Yang et al., 2010). Using a DNA fiber-fluorescence *in situ* hybridization approach to study individual maize centromeres, the association of a specific centromeric histone H3 (CENH3) was visualized in centromeres (Jin et al., 2004). This analysis revealed that CENH3 is always associated with centromere-specific satellite repeats, but that not all these sequences are associated with CENH3 (Jin et al., 2004). Finally, fluorescence *in situ* hybridization analysis of a reciprocal translocation in maize between chromosomes 1 and 5 revealed the presence of an inactive centromere at or near the breakpoints of the two chromosomes (Gao et al., 2011). To confirm the active and inactive states of two sets of centromeric sequences, biochemical features of active centromeres were examined in root tip metaphase spreads of this material. CENP-C is an inner kinetochore protein that is characteristic of all active centromeres (Dawe et al., 1999). CENP-C was only located at the site of the primary constriction and was not detectable at the second cluster of centromere sequences. A second biochemical feature of active centromeres is the phosphorylation of Serine-10 on histone H3 (Houben et al., 2007). Immunolocalizations combined with FISH using antibodies against this histone modification revealed that only the set of centromeric sequences at the primary constriction were detectably labeled. Together, the results showed that this centromere does not exhibit any of the tested biochemical features of active centromeres, and it extends the evidence for an epigenetic component to centromere function in plants (Gao et al., 2011).

PERSPECTIVES

Although plant nuclear proteomics is only at the initial phase, several studies have been reported using different species that include the analysis of the total nuclear or phospho-proteomes, and differential nuclear proteomes; however, few reports exist on the analysis of the maize nuclear proteome or its changes under various environmental conditions. In particular, most experiments have been done on the study of histones and their covalent modifications. The reason for this is probably due to the number of experiments done in epigenetics using this crop as a model, in particular in paramutation, imprinting, and transposon silencing. Thus, the low number of publications available provides most of the sources for this review. Therefore, the research presented here from other plant species can provide ideas for potential future experiments in maize. On the other hand, experiments using analysis by mass spectrometry, but also by immunological techniques will be very helpful to increase our knowledge of maize nuclear proteomics.

ACKNOWLEDGMENTS

This work was supported by FONCyT grants PICT-2007-00711 and PICT-2010-00105. Paula Casati is a member of the Research Career of the CONICET of Argentina.

REFERENCES

Aki, T., and Yanagisawa, S. (2009). Application of rice nuclear proteome analysis to the identification of evolutionarily conserved and glucose-responsive nuclear proteins. *J. Proteome Res.* 8, 3912–3924.

Bae, M. S., Cho, E. J., Choi, E.-Y., and Park, O. K. (2003). Analysis of the *Arabidopsis* nuclear proteome and its response to cold stress. *Plant J.* 36, 652–663.

Campi, M., D'Andrea, L., Emiliani, J., and Casati, P. (2012). Participation of chromatin-remodeling proteins in the repair of ultraviolet-B-damaged. *Plant Physiol.* 158, 981–995.

Casasoli, M., Meliciani, I., Cervone, F., De Lorenzo, G., and Mattei, B. (2007). Oligogalacturonide-induced changes in the nuclear proteome of *Arabidopsis thaliana*. *Int. J. Mass Spectrom.* 268, 277–283.

Casati, P., Campi, M., Chu, F., Suzuki, N., Maltby, D., Guan, S., et al. (2008). Histone acetylation and chromatin remodeling are required for UV-B-dependent transcriptional activation of regulated genes in maize. *Plant Cell* 20, 827–842.

Choudhary, M. K., Basu, D., Datta, A., Chakraborty, N., and Chakraborty, S. (2009). Dehydration-responsive nuclear proteome of rice (*Oryza sativa* L.) illustrates protein network,

novel regulators of cellular adaptation, and evolutionary perspective. *Mol. Cell. Proteomics* 8, 1579–1598.

Cooper, B., Campbell, K. B., Feng, J., Garrett, W. M., and Frederick, R. (2011). Nuclear proteomic changes linked to soybean rust resistance. *Mol. BioSyst.* 7, 773–783.

Dawe, R. K., Reed, L. M., Yu, H., Muszynski, M. G., and Hiatt, E. N. (1999). A maize homolog of mammalian CENPC is a constitutive component of the inner kinetochore. *Plant Cell* 11, 1227–1238.

Erhardt, M., Adamska, I., and Franco, O. L. (2010). Plant nuclear proteomics – inside the cell maestro. *FEBS J.* 277, 3295–3307.

Gao, Z., Fu, S., Dong, Q., Han, F., and Birchler, J. A. (2011). Inactivation of a centromere during the formation of a translocation in maize. *Chromosome Res.* 19, 755–761.

González-Camacho, F., and Medina, F. J. (2004). Identification of specific plant nucleolar phosphoproteins in a functional proteomic analysis. *Proteomics* 4, 407–417.

Haring, M., Bader, R., Louwers, M., Schwabe, A., van Driel, R., and Stam, M. (2010). The role of DNA methylation, nucleosome occupancy and histone modifications in paramutation. *Plant J.* 63, 366–378.

Haun, W. J., and Springer, N. M. (2008). Maternal and paternal alleles exhibit differential histone methylation and acetylation at maize imprinted genes. *Plant J.* 56, 903–912.

Houben, A., Demidov, D., Caperta, A. D., Karimi, R., Agueci, F., and Vlasenko, L. (2007). Phosphorylation of histone H3 in plants – a dynamic affair. *Biochim. Biophys. Acta* 1769, 308–315.

Jaskiewicz, M., Peterhansel, C., and Conrath, U. (2011). Detection of histone modifications in plant leaves. *J. Vis. Exp.* 55, e3096.

Jin, W., Melo, J. R., Nagaki, K., Talbert, P. B., Henikoff, S., Dawe, R. K., et al. (2004). Maize centromeres: organization and functional adaptation in the genetic background of oat. *Plant Cell* 16, 571–581.

Kalamajka, R., Finnie, C., and Grasser, K. D. (2010). Onset of grain filling is associated with a change in properties of linker histone variants in maize kernels. *Planta* 231, 1127–1135.

Kouzarides, T. (2007). Chromatin modifications and their function. *Cell* 128, 693–705.

Pandey, A., Chakraborty, S., Datta, A., and Chakraborty, N. (2007). Proteomics approach to identify dehydration responsive nuclear proteins from chickpea (*Cicer arietinum* L.). *Mol. Cell. Proteomics* 7, 88–107.

Pandey, A., Choudhary, M. K., Bhushan, D., Chattopadhyay, A., Chakraborty, S., Datta, A., et al. (2006). The nuclear proteome of chickpea (*Cicer arietinum* L.) reveals predicted and unexpected proteins. *J. Proteome Res.* 5, 3301–3311.

Pendle, A. F., Clark, G. P., Boon, R., Lewandowska, D., Lam, Y. W., Andersen, J., et al. (2005). Proteomic analysis of the *Arabidopsis nucleolus* suggests novel nucleolar functions. *Mol. Biol. Cell* 16, 260–269.

Qüesta, J. I., Walbot, V., and Casati, P. (2010). Mutator transposon activation after UV-B involves chromatin remodeling. *Epigenetics* 5, 1–12.

Repetto, O., Rogniaux, H., Firnhaber, C., Zuber, H., Küster, H., Larré, C., et al. (2008). Exploring the nuclear proteome of *Medicago truncatula* at the switch towards seed filling. *Plant J.* 56, 398–410.

Winck, F. V., Riaño-Pachón, D. M., Sommer, F., Rupprecht, J., and Mueller-Roeber, B. (2012). The nuclear proteome of the green alga *Chlamydomonas reinhardtii. Proteomics* 12, 95–100.

Yang, F., Zhang, L., Li, J., Huang, J., Wen, R., Lu, M., et al. (2010). Trichostatin A and 5-azacytidine both cause an increase in global histone H4 acetylation and a decrease in global DNA and H3K9 methylation during mitosis in maize. *BMC Plant Biol.* 10:178. doi:10.1186/1471-2229-10-178

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 04 October 2012; paper pending published: 31 October 2012; accepted: 24 November 2012; published online: 12 December 2012.

Citation: Casati P (2012) Recent advances in maize nuclear proteomic studies reveal histone modifications. Front. Plant Sci. ***3****:278. doi: 10.3389/fpls.2012.00278*

This article was submitted to Frontiers in Plant Proteomics, a specialty of Frontiers in Plant Science.

Copyright © 2012 Casati. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.

frontiers in
PLANT SCIENCE

MINI REVIEW ARTICLE
published: 20 December 2012
doi: 10.3389/fpls.2012.00289

The seed nuclear proteome

Ombretta Repetto[1†], Hélène Rogniaux[2], Colette Larré[2], Richard Thompson[1] and Karine Gallardo[1]*

[1] UMR1347 Agroécologie, Institut National de la Recherche Agronomique, Dijon, France
[2] UR1268 Biopolymers, Interactions, Assemblies, Institut National de la Recherche Agronomique, Nantes, France

Edited by:
Nicolas L. Taylor, The University of Western Australia, Australia

Reviewed by:
Jozef Samaj, Centre of the Region Hana for Biotechnological and Agricultural Research, Palacky University Olomouc, Czech Republic
Tiago S. Balbuena, State University of Campinas, Brazil

***Correspondence:**
Karine Gallardo, UMR1347 Agroécologie, Institut National de la Recherche Agronomique, 17 rue de Sully, BP 86510, Dijon, France.
e-mail: gallardo@dijon.inra.fr

†Present address:
Ombretta Repetto, Proteomics Core Facility, Experimental and Clinical Pharmacology, Centro di Riferimento Oncologico - Istituto di Ricovero e Cura a Carattere Scientifico, Aviano, Italy.

Understanding the regulatory networks coordinating seed development will help to manipulate seed traits, such as protein content and seed weight, in order to increase yield and seed nutritional value of important food crops, such as legumes. Because of the cardinal role of the nucleus in gene expression, sub-proteome analyses of nuclei from developing seeds were conducted, taking advantage of the sequences available for model species. In this review, we discuss the strategies used to separate and identify the nuclear proteins at a stage when the seed is preparing for reserve accumulation. We present how these data provide an insight into the complexity and distinctive features of the seed nuclear proteome. We discuss the presence of chromatin-modifying enzymes and proteins that have roles in RNA-directed DNA methylation and which may be involved in modifying genome architecture in preparation for seed filling. Specific features of the seed nuclei at the transition between the stage of cell divisions and that of cell expansion and reserve deposition are described here which may help to manipulate seed quality traits, such as seed weight.

Keywords: seeds, development, nuclei, proteomics, regulation

INTRODUCTION

Because seeds, such as those of legumes and cereals, are a source of nutrients for animal and human nutrition, breeding objectives include improving seed quality and yield and/or stabilizing these traits under fluctuating environmental conditions. To develop an understanding of the genetic factors controlling these traits, omics studies of seed development were performed from the year 2000 onward exploiting the availability of genome sequence for several species, including Arabidopsis, rice, and *Medicago truncatula*. This last species was adopted in 2001 as a model for legumes because of its small genome size compared to other legumes (Bell et al., 2001). Genomics resources were then developed in this species (Young et al., 2011) and extensively exploited, notably to study seed biology. Proteomics has been used to identify candidate proteins with roles in seed development (Thompson et al., 2009). While only abundant soluble proteins were identified by proteomics targeted to entire seed tissues, transcriptome studies provided information about low-abundance expression of some genes (Thompson et al., 2009). By comparing the timing of appearance of the proteins with that of their corresponding transcripts during seed development, divergent patterns were found for 50% of the proteins detected in the *M. truncatula* seed proteome (Gallardo et al., 2007). This indication of major post-transcriptional events highlighted the need to choose a proteomics approach to identify the regulatory mechanisms governing seed development. Targeted to the nucleus, proteomics allowed identification of regulatory proteins in leaves, suspension cells, or seedlings from various species, including Arabidopsis, rice (*Oryza sativa*), maize (*Zea mays*), and chickpea (*Cicer arietinum*; Bae et al., 2003; Khan and Komatsu, 2004; Ferreira et al., 2006; Pandey et al., 2006; Tan et al., 2007; Li et al., 2008). To provide a list of nuclear proteins with potential regulatory role(s) in developing *M. truncatula* seeds, the approach of combining nuclei isolation with proteomics was applied at 12 days after pollination (dap; Repetto et al., 2008). This key stage is characterized by the switch from an embryogenesis-oriented program, with frequent cell divisions, to a filling program associated with embryo cell expansion and reserve accumulation. In a parallel study, a nuclear proteomics approach was applied to the filial tissue of rice seeds (i.e., the endosperm) at 9 dap (Li et al., 2008). At this stage, the embryo is differentiated and the reserves start to accumulate (Luo et al., 2011). Because understanding the processes underlying the embryogenesis/filling transition might help greatly to modulate both seed size and storage capacities, after outlining the strategies used to identify nuclear proteins from developing seeds, we describe the specificities of the seed nuclear proteome, and discuss the proteins that might play key roles in controlling this transition.

SEED NUCLEI PURIFICATION AND PROTEIN EXTRACTION

Nuclear isolation methods based on density gradients were applied to immature seeds or seed tissues in flax (*Linum usitatissimum*), *M. truncatula*, rice, and maize (**Table 1**), with the objective of obtaining nuclei of sufficient yield and quality for protein profiling (Ferreira et al., 2006), proteomics (Li et al., 2008; Repetto et al., 2008), or gel shift experiments (Renouard et al., 2012). Castillo et al. (2000) also succeeded in isolating nuclei from ungerminated

Table 1 | Isolation of seed nuclei for protein identification or gel shift experiments.

Plant	Tissues/organs	Objective(s)	Extraction procedure	Approach for protein isolation and/or identification	Results	Reference
Linum usitatissimum cv. Barbara	Immature seed coat at 16 daf (torpedo stage)	Establishment of a protocol for isolation of nuclear proteins from flax seed coats	Nuclei isolation (sucrose-ficoll) after cell wall and mucilage digestion	1-DE protein separation; WB and dot blot to confirm enrichment in nuclear proteins; Gel shift assay to analyze DNA–protein interactions	Isolation of nuclear proteins from flax seed coat without contaminants for their use in gel shift experiments	Renouard et al. (2012)
Medicago truncatula cv. Jemalong, line A17	Whole seeds at 12 dap (embryo genesis – seed filling transition)	Exploration of the nuclear proteome at the switch toward seed filling	CelLytic PN extraction kit (Sigma-Aldrich) with some modifications	1-DE, then WB to confirm enrichment in nuclear proteins; MS analyses: C18 RP-LC nano-ESI MS/MS (Waters Q-TOF Global); extended to gene expression profiling (qRT-PCR and microarray data exploitation)	179 polypeptides identified (143 different proteins), including ribosomal proteins, HD2A-type histone deacetylases, Pescadillo-like protein, RNA polymerase IV, argonaute 4, EBP1, Alba protein-like	Repetto et al. (2008)
Oryza sativa cv. Nipponbare	Endosperm at 9 dap (milky stage)	Establishment of a protocol to remove starch grains and recover the low abundant proteins for nuclear sub-proteome studies in rice endosperm	Method adapted from Lee and Lin (2005); based on the use of sucrose	MS analyses on: (a) gel-free protein samples, SCX, C18 RP-LC and ESI-MS/MS (Thermo Finnigan LCQ Deca XP ion trap); (b) spots, MALDI-TOF/TOF (Applied Biosystems 4700 Proteomics Analyzer)	468 proteins identified, including HD2A-type histone deacetylases, Pescadillo-like protein, bZIP and bHLH transcription factors, and many hypothetical proteins	Li et al. (2008)
Zea mays var. UENF 506-6	Endosperm 8–35 dap	Establishment of a method to profile nuclear proteins during maize endosperm development	Method adapted from Kodrzycki et al. (1989); based on the use of sucrose	1-DE, then WB to assess the quality of the nuclear protein extracts	Abundant nuclear proteins differentially expressed during seed development were revealed in the 1-DE gels. The proteins remain to be identified	Ferreira et al. (2006)
Pisum sativum cv. Lincoln	embryonic axes (ungerminated)	Characterization of a protein of 16 kDa (p16) induced by dehydration	According to Ull and Franco (1986); Percoll	1-DE, then WB; Edman degradation (Waters Prosequencer 6625); cDNA library and DNA sequencing	Isolation and partial sequencing of a nuclear protein of the vicilin superfamily with possible roles in the protection of chromatin structure against desiccation in seeds	Castillo et al. (2000)

Some proteins identified were listed and underlined are those identified in at least two studies. 1-DE, one dimensional electrophoresis; bHLH, basic helix-loop-helix; bZIP, basic leucine zipper; daf, days after flowering; dap, days after pollination; DIP2, DNA-binding domain interacting protein; EBP1, epidermal growth factor receptor binding protein; ESI, electro spray ionization; LC, liquid chromatography; MALDI, matrix-assisted laser desorption ionization; MS, mass spectrometry; Q-TOF, quadrupole time of flight; RP, reverse phase; SCX, strong cation exchange; TOF, time of flight; WB, western blotting.

pea embryonic axes to purify and sequence a nuclear protein induced by dehydration (**Table 1**). The isolation of nuclei from developing seeds is challenging due to the presence of storage compounds such as globulins, oils, and carbohydrates (Gallardo et al., 2008). In *M. truncatula*, we tested several nuclear separation procedures from seeds collected at different developmental stages, including flow cytometry, sucrose or percoll density gradients, before adopting a sucrose-based "semi-pure" nuclear preparation of the CelLytic plant nuclei isolation kit (Sigma-Aldrich) to which we have made some modifications described in Repetto et al. (2008). At the 12 dap stage, the *M. truncatula* seed possesses nuclei of 5–15 μm diameter with low DNA *C*-value (0.48 pg; Arumuganathan and Earle, 1991). Observations of nuclei preparations from *M. truncatula* seeds at later stages reveal few and larger seed nuclei, along with many starch granules probably originating from the seed coats (Abirached-Darmency et al., 2005). Optimizations are necessary to obtain high-purity nuclei at these stages, which differ in the number of contaminants (e.g., protein bodies, starch granules), average nuclear size, and DNA content. Interestingly, a cotton filtration step was set up by Li et al. (2008) for starch grain removal from rice endosperm at 9 dap, and a protocol allowing the removal of mucilage and phenolic compounds from seed coats before nuclei isolation was developed by Renouard et al. (2012); (**Table 1**).

Two of the nuclei isolation methods presented in **Table 1** were combined with mass spectrometry (MS) for sub-proteome analyses. In Repetto et al. (2008), the nuclei-containing pellets obtained from 12 dap *M. truncatula* seeds were directly resuspended in a high salt concentration buffer (1 M NaCl), and then sonicated to destroy the nuclear membranes. After validating the enrichment for nuclear proteins by western blotting with antibodies for histone H1 and for proteins specific for other subcellular compartments, the resulting protein extract was directly separated by mono-dimensional gel electrophoresis (1-DE) and the whole lane was sequentially cut into 36 portions for MS analyses (**Figure 1**). A different approach was used by Li et al. (2008). They first removed the highly abundant bands corresponding to storage proteins from the 1-DE profile by excision, and then crushed the rest of the gel to extract the low abundance proteins using a phenol extraction buffer. After precipitation, the protein pellet was dissolved in 6 M urea with 100 mM Tris–Cl for MS analyses.

IDENTIFICATION OF SEED NUCLEAR PROTEINS

In Repetto et al. (2008), the in-gel trypsin-digested peptides were separated by liquid nano-chromatography (nanoLC) and further measured and fragmented (MS/MS experiments) in a hybrid quadrupole-time-of-flight mass spectrometer. A search in both a wide databank (UniRef100) and a targeted databank made of expressed sequence tags from *M. truncatula* (the TIGR MtGI release 8 database) was realized from the mass data. The databank search program was MASCOT 2.2 and proteins were identified when at least two of their peptides matched the databank entry with a *p*-value <0.05. We succeeded in identifying 179 polypeptides, corresponding to 143 distinct proteins, using this approach. Sequence annotations were manually checked or completed by (cross-) BLAST "parameters" searches against the NCBI non-redundant database. The proteins were functionally classified according to the MapMan ontology (Usadel et al., 2005) as well as to a manual assignment not limited to homologs as described in Gallardo et al. (2007). A complete list of proteins is available in Repetto et al. (2008) that remains to date the most comprehensive description of the *M. truncatula* nuclear proteome.

In a parallel study, a shotgun proteomics approach was used by Li et al. (2008) to characterize the rice nuclear proteome. The complex peptide mixtures derived from trypsin digestion were subjected to 2-D liquid chromatography coupled to an ESI-IT (electro spray ionization-ion trap) mass spectrometer. A search in the rice non-redundant protein database (NCBInrPDB) was done from the mass data, and proteins were identified when at least two of their peptides matched the databank entry with a *p*-value <0.01. This approach identified 468 proteins from the nuclear enriched fractions of rice endosperm. A nuclear localization was assigned for 47% of these proteins by searching the Gene Ontology (GO) database (http://www.geneontology.org/). It should be noted that prediction of nuclear localization of proteins is far from being easy and entirely reliable. In fact, the nucleo-cytoplasmic protein shuttling through the nuclear pore complex (NPC) is a highly dynamic and complex system (Grünwald and Singer, 2012), and for many proteins (e.g., ribosomal and cytoskeletal) there is a consistent evidence for multiple locations. Moreover, only a fraction of the proteins localized in nuclei possess nuclear localization signals for NPC-mediated transport into the nucleus. Therefore, the prediction of nuclear localization based on the presence of signal peptides (e.g., PSORT; Nakai and Horton, 1999) is usually coupled with homology-based GO annotations, and must ideally be confirmed by further experiments, for example using fluorescent protein fusions or specific antibodies.

Among the proteins identified in the *M. truncatula* seed nucleus that may be multifunctional and might display different organelle functions and localizations, are certain enzymes of intermediary metabolism. Previous studies also reported the presence of these enzymes in the nucleus although no obvious nuclear localization signal was found in their sequences (Yamamoto et al., 1997; Markova et al., 2006; Li et al., 2008; Lee et al., 2012). As an example, sulfite reductase, a key plastid enzyme involved in sulfur reduction in plants, was identified in the *M. truncatula* seed nucleus. This enzyme was shown to bind to DNA in the chloroplast, and thus to repress genomic activity (i.e., transcription) through DNA compaction (Sato et al., 2003; Sekine et al., 2007). Although further experiments are needed to confirm their nuclear localization, the presence of such proteins raises the possibility of a regulation of transcriptional activities in seeds through nuclear targeting of metabolic enzymes. They may be able to monitor metabolic status in response to various stimuli by transmitting the changes to the transcriptional apparatus.

SPECIFICITIES OF THE SEED NUCLEAR PROTEOME

A comparison of nuclear proteomes from different organs and species might help to decipher the level of conservation of nuclear proteins and to identify tissue- or species-specific nuclear functions. With the aim to identify specific nuclear features in seeds, we compared the nucleus proteome of *M. truncatula* seeds with that of the rice endosperm at a milky stage (Li et al., 2008), and that of chickpea seedlings (Pandey et al., 2006) and Arabidopsis leaves

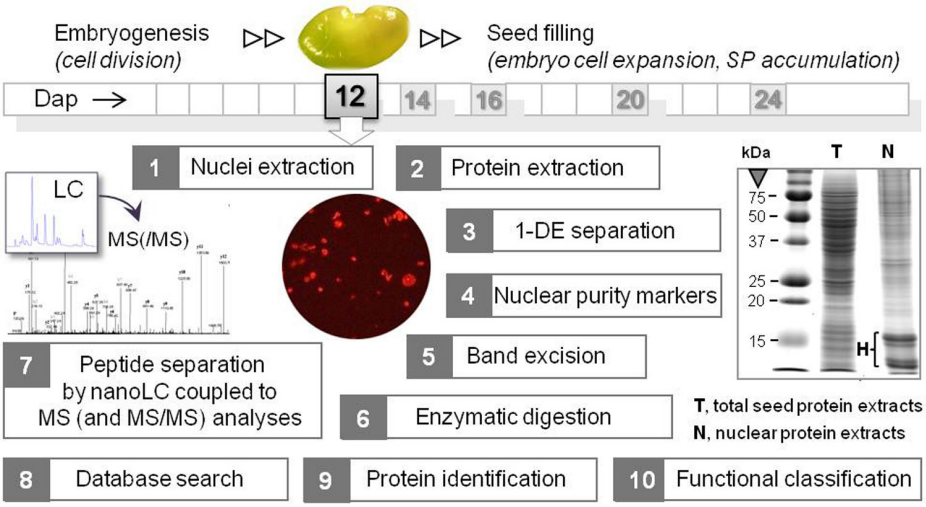


FIGURE 1 | Workflow of the nuclear proteomics approach applied to *M. truncatula* seeds at a key stage between embryogenesis and seed filling. Organelles were purified from 12 dap seeds (1 seed = 1.5 mm length), and the proteins were extracted and separated by mono-dimensional gel electrophoresis (1-DE). A typical 1-DE profile is shown with intense bands at about 10–15 kDa corresponding to histones (H). After assessing the purity of the nuclear protein fraction by Western blotting using antibodies against proteins specific of different cell compartments, the in-gel digested peptides were analyzed by nanoLC coupled to MS and MS/MS analyses. The peptide mass data were subjected to a database search for putative protein identification, and the proteins were functionally classified after a search for nuclear peptide signals. SP, storage proteins. dap, days after pollination.

(Bae et al., 2003). Interestingly, two protein classes were particularly enriched in the *M. truncatula* seed nucleus at a stage preparing for reserve deposition: RNA processing and ribosome biogenesis. In particular, an abundant pool of proteins (22% of the proteins identified) was found that are members of the ribosomal protein families comprising the 40S and 60S subunits synthesized within the nucleolus in eukaryotes. The abundance of their transcripts decreased sharply at the beginning of seed filling (i.e., 14–16 dap; Gallardo et al., 2007). A salient feature of 12 dap *M. truncatula* seeds is therefore the storage of a large pool of ribosomal proteins within the nucleus, that can presumably be further readily used for storage protein synthesis during seed filling. This may contribute to our understanding of the mechanisms allowing legume seeds to synthesize large amounts of storage proteins while entering into a quiescent state. It also raises an important question of whether the stored ribosomal proteins could be involved in the intricate control of homeostasis of protein amount per seed under challenging environmental conditions. Interestingly, a PESCADILLO-like protein that may play a role in the biogenesis of ribosomal subunits was identified in the nuclear proteome of both the rice endosperm and *M. truncatula* seeds (**Table 1**). This protein is not functionally characterized in plants but implicated in rRNA precursor processing and ribosomal subunit assembly in human and mammalian cells (Andersen et al., 2002; Lerch-Gaggl et al., 2002).

In the nuclear proteomes of both the *M. truncatula* seed and rice endosperm the proportion of functionally annotated proteins belonging to the DNA metabolism class (12% in *M. truncatula* and 29% in rice) exceeded that found in chickpea seedlings (Pandey et al., 2006) and Arabidopsis leaves (Bae et al., 2003). Some of these proteins are involved in the epigenetic regulation of the genome (Li et al., 2008; Repetto et al., 2008). There is increasing evidence that some components of the chromatin modification machinery play a significant role in developing seeds. Recent surveys demonstrated that genomic imprinting primarily occurs in the endosperm in both rice and Arabidopsis, and that gene-specific imprinting in the embryo also exists in maize (Ikeda, 2012 and references therein). By comparing candidate imprinted genes from rice and Arabidopsis, Luo et al. (2011) found a low degree of conservation, suggesting that imprinting targets have evolved independently in dicots and monocots. In seeds, the epigenetic regulation of the genome, which modulates chromatin structure to limit the expression of genes to a particular tissue at a specific developmental stage, could play a crucial role in the developmental switch of the dicot embryo cells from division to expansion and filling (**Figure 1**). In legumes, final seed weight is largely determined by the number of cotyledon cells (Munier-Jolain and Ney, 1998). Therefore, identifying the epigenetic components of legume seeds that regulate the timing of the transition between cell division and cell expansion might help to manipulate final seed weight.

Among the epigenetic components detected in the *M. truncatula* seed nuclei were histone deacetylases HD2A that are good candidates for regulating the transition from an embryonic program to a filling mode. HD2A are plant-specific chromatin-remodeling factors participating in transcriptional repression *via* the modification of gene accessibility (Li et al., 2002). Interestingly, these proteins were also identified in the filial tissue of rice (**Table 1**). HD2A are expressed strongly in embryonic tissues and their ectopic expression under the control of the 35S promoter resulted in developmental abnormalities, including aborted seed development (Zhou et al., 2004). Importantly, Tanaka et al. (2008) demonstrated that histone deacetylases are involved in the repression of embryonic properties upon germination *via* direct or indirect repression of embryo-specific transcription factors. It is therefore possible that HD2A also plays a role in

regulating the switch from embryogenesis to seed filling in eudicots and monocots. Although this hypothesis requires experimental confirmation, it holds promise to resolve the presently unclear mechanisms shifting the seed developmental program to reserve deposition (**Figure 1**).

The histone modifications induced by HD2A may be associated with other chromatin modifications, such as DNA methylation, to silence gene expression in response to developmental stimuli. Interestingly, two proteins needed for RNA-directed DNA methylation (i.e., DNA methylation guided by 24 nucleotide small interfering RNAs) were identified in the *M. truncatula* seed nucleus: a subunit of the plant-specific RNA polymerase IV, and argonaute 4 (AGO4). These proteins were not identified in nuclei from rice endosperm (**Table 1**), chickpea seedlings (Pandey et al., 2006), or Arabidopsis leaves (Bae et al., 2003), suggesting a specific role in legume seeds and/or in immature embryos. RNA polymerase IV is required for the biogenesis of a major class of 24-nucleotide small interfering RNAs, which are predominantly expressed in the developing endosperm of Arabidopsis seeds (Lu et al., 2012). Li et al. (2006) showed that the C-terminal domain of a RNA polymerase IV subunit interacts with AGO4 within nucleolus-associated bodies (i.e., Cajal bodies), which have been proposed to be a site for the generation of siRNA/protein complexes acting in RNA-directed DNA methylation. The detection of these proteins in the *M. truncatula* seed nucleus suggests they may interact in 12 dap seeds in concert with HD2A to repress the expression of genes *via* chromatin remodeling. To elucidate the mechanism of repression, it will be necessary to identify the target genes, some putative candidates could be described in the following section.

PROTEINS IMPLICATED IN TRANSCRIPTIONAL REGULATION

When targeted to the nucleus, proteomics offers the opportunity to identify regulatory factors controlling cell development, differentiation, and cell growth by binding to DNA and regulating gene expression. In seeds, there is great interest in identifying such factors to manipulate seed size and weight. A putative transcriptional regulator which was found specifically in the *M. truncatula* seed nucleus may control cell division but its function in seeds has not yet been characterized. This protein, named EBP1 (epidermal growth factor receptor binding protein), recruits histone deacetylase activity in human cells to mediate the transcriptional repression of E2 promoter binding factors (E2F) controlling cell cycle progression (Zhang et al., 2003). In potato and Arabidopsis, Horváth et al. (2006) demonstrated that EBP1 regulates organ size through cell growth and proliferation: elevating or decreasing EBP1 levels in transgenic plants resulted in a dose-dependent increase or reduction in leaf surface area, respectively. In the same study, they showed that EBP1 is required for expression of cell cycle genes in an auxin-dependent manner. This is likely to occur through the repression of RBR1 (retinoblastoma binding protein-like) that blocks cell cycle progression by inhibiting E2F-dependent transcription, which is required for expression of many genes involved in S-phase and cell cycle progression (Lai et al., 1999). The presence of EBP1 in the *M. truncatula* seed nucleus suggests this protein could play a key role in the control of cell division during seed development. Various other regulatory proteins were specifically detected in the rice endosperm nucleus (e.g., basic leucine zipper and basic helix-loop-helix transcription factors) or in the *M. truncatula* seed nucleus (e.g., DNA-binding domain interacting protein DIP2, Alba protein-like; **Table 1**). In plants, the exact functions of some of these proteins remain to be defined. The Alba protein has been proposed to control chromatin structure through interaction with histone deacetylase in Archaea and could also have a function in RNA metabolism (Bell et al., 2002; Aravind et al., 2003). The DIP2 protein displays similarities with the animal transcriptional coactivator ALY, suggesting it could be involved in transcriptional regulation. In plants, DIP2 interacts with the DNA-binding domain of plant poly(ADP-ribose) polymerases possibly implicated in chromosome dynamics and modifying proteins involved in different signaling pathways from DNA damage to energy metabolism (Babiychuk et al., 2001; Storozhenko et al., 2001).

CONCLUSION

The availability of data from next-generation technologies, now used for *de novo* sequencing of genomes in crops, such as pigeonpea (Varshney et al., 2011), will facilitate the identification of seed nuclear proteins in these species. Ethyl methanesulfonate (EMS) and TnT1 insertion mutant populations have been developed in *M. truncatula* and EMS mutants in pea and rice (Dalmais et al., 2008; Tadege et al., 2008; Le Signor et al., 2009; Wang et al., 2012; Cooper et al., 2013). Both reverse and forward genetics can be applied to study mutants from these collections. Moreover, a series of EMS mutations could be identified by TILLING in candidate genes for regulating the embryogenesis-filling transition, which in addition to providing mutants for functional studies, could reveal favorable alleles to be used in selection for seed quality improvements.

ACKNOWLEDGMENTS

The work on *M. truncatula* was supported by an INRA postdoctoral fellowship and by grants #B05796 from the Regional Council of Burgundy. We sincerely thank Steven P. C. Groot and Jan Bergervoët (Plant Research International, Wageningen, Netherlands) for advice and help regarding nuclei preparations.

REFERENCES

Abirached-Darmency, M., Abdelgawwad, M. R., Conejero, G., Verdeil, J. L., and Thompson, R. (2005). *In situ* expression of two storage protein genes in relation to histodifferentiation at mid-embryogenesis in *Medicago truncatula* and *Pisum sativum* seeds. *J. Exp. Bot.* 56, 2019–2028.

Andersen, J. S., Lyon, C. E., Fox, A. H., Leung, A. K., Lam, Y. W., Stehen, H., et al. (2002). Directed proteomic analysis of the human nucleolus. *Curr. Biol.* 12, 1–11.

Aravind, L., Iyer, L. M., and Anantharaman, V. (2003). The two faces of Alba: the evolutionary connection between proteins participating in chromatin structure and RNA metabolism. *Genome Biol.* 4, R64.

Arumuganathan, K., and Earle, E. D. (1991). Nuclear DNA content of some important plant species. *Plant Mol. Biol. Rep.* 9, 208–218.

Babiychuk, E., Van Montagu, M., and Kushnir, S. (2001). N-terminal domains of plant poly(ADP-ribose) polymerases define their association with mitotic chromosomes. *Plant J.* 28, 245–255.

Bae, M. S., Cho, E. J., Choi, E., and Prk, O. K. (2003). Analysis of the Arabidopsis nuclear proteome and its response to cold stress. *Plant J.* 36, 652–663.

Bell, C. J., Dixon, R. A., Farmer, A. D., Flores, R., Inman, J., Gonzales, R. A., et al. (2001). The Medicago

Genome Initiative: a model legume database. *Nucleic Acids Res.* 29, 114–117.

Bell, S. D., Botting, C. H., Wardleworth, B. N., Jackson, S. P., and White, M. F. (2002). The interaction of Alba, a conserved archaeal chromatin protein, with Sir2 and its regulation by acetylation. *Science* 296, 148–151.

Cooper, J. L., Henikoff, S., Comai, L., and Till, B. J. (2013). TILLING and Ecotilling for rice. *Methods Mol. Biol.* 956, 39–56.

Castillo, J., Rodrigo, M. I., Márquez, J. A., Zúñiga, A., and Franco, L. (2000). A pea nuclear protein that is induced by dehydration belongs to the vicilin superfamily. *Eur. J. Biochem.* 267, 2156–2165.

Dalmais, M., Schmidt, J., Le Signor, C., Moussy, F., Burstin, J., Savois, V., et al. (2008). UTILLdb, a *Pisum sativum in silico* forward and reverse genetics tool. *Genome Biol.* 9, R43.

Ferreira, B. S., Branco, A. T., de Oliveira, M. A., Pereira, M. G., and de Souza Filho, G. A. (2006). Methodological improvements on extraction of nuclear proteins and its preliminary analysis during the maize (*Zea mays* L.) endosperm development. *Protein Pept. Lett.* 13, 981–984.

Gallardo, K., Firnhaber, C., Zuber, H., Héricher, D., Belghazi, M., Henry, C., et al. (2007). A combined proteome and transcriptome analysis of developing *Medicago truncatula* seeds: evidence for metabolic specialization of maternal and filial tissues. *Mol. Cell. Proteomics* 6, 2165–2179.

Gallardo, K., Thompson, R., and Burstin, J. (2008). Reserve accumulation in legume seeds. *C. R. Biol.* 331, 755–762.

Grünwald, D., and Singer, R. H. (2012). Multiscale dynamics in nucleocytoplasmic transport. *Curr. Opin. Cell Biol.* 24, 100–106.

Horváth, B. M., Magyar, Z., Zhang, Y., Hamburger, A. W., Bakó, L., Visser, R. G., et al. (2006). EBP1 regulates organ size through cell growth and proliferation in plants. *EMBO J.* 25, 4909–4920.

Ikeda, Y. (2012). Plant imprinted genes identified by genome-wide approaches and their regulatory mechanisms. *Plant Cell Physiol.* 53, 809–816.

Khan, M. M. K., and Komatsu, S. (2004). Rice proteomics: recent developments and analysis of nuclear proteins. *Phytochemistry* 65, 1671–1681.

Kodrzycki, R., Boston, R. S., and Larkins, B. A. (1989). The opaque-2 mutation of maize differentially reduces zein gene transcription. *Plant Cell* 1, 105–114.

Lai, A., Marcellus, R. C., Corbeil, H. B., and Branton, P. E. (1999). RBP1 induces growth arrest by repression of E2F-dependent transcription. *Oncogene* 18, 2091–2100.

Lee, H., and Lin, T. Y. (2005). Isolation of plant nuclei Suitable for flow cytometry from recalcitrant tissue by use of a filtration column. *Plant Mol. Biol. Rep.* 23, 53–58.

Lee, S., Doxey, A. C., McConkey, B. J., and Moffatt, B. A. (2012). Nuclear targeting of methyl-recycling enzymes in *Arabidopsis thaliana* is mediated by specific protein interactions. *Mol. Plant* 5, 231–248.

Lerch-Gaggl, A., Haque, J., Li, J., Ning, G., Traktman, P., and Duncan, S. A. (2002). Pescadillo is essential for nucleolar assembly, ribosome biogenesis, and mammalian cell proliferation. *J. Biol. Chem.* 47, 45347–45355.

Le Signor, C., Savois, V., Aubert, G., Verdier, J., Nicolas, M., Pagny, G., et al. (2009). Optimizing TILLING populations for reverse genetics in *Medicago truncatula. Plant Biotechnol. J.* 7, 430–441.

Li, G., Hall, T. C., and Holmes-Davis, R. (2002). Plant chromatin: development and gene control. *Bioessays* 24, 234–243.

Li, C. F., Pontes, O., El-Shami, M., Henderson, I. R., Bernatavichute, Y. V., Chan, S. W., et al. (2006). An ARGONAUTE4-containing nuclear processing center colocalized with Cajal bodies in *Arabidopsis thaliana. Cell* 126, 93–106.

Li, G., Nallamilli, B. R., Tan., F., and Peng, Z. (2008). Removal of high-abundance proteins for nuclear subproteome studies in rice (*Oryza sativa*) endosperm. *Electrophoresis* 29, 604–617.

Lu, J., Zhang, C., Baulcombe, D. C., and Chen, Z. J. (2012). Maternal siRNAs as regulators of parental genome imbalance and gene expression in endosperm of Arabidopsis seeds. *Proc. Natl. Acad. Sci. U.S.A.* 109, 5529–5534.

Luo, M., Taylor, J. M., Spriggs, A., Zhang, H., Wu, X., Russell, S., et al. (2011). A genome-wide survey of imprinted genes in rice seeds reveals imprinting primarily occurs in the endosperm. *PLoS Genet.* 7: e1002125. doi: 10.1371/journal.pgen.1002125

Markova, N. G., Pinkas-Sarafova, A., and Simon, M. (2006). A metabolic enzyme of the short-chain dehydrogenase/reductase superfamily may moonlight in the nucleus as a repressor of promoter activity. *J. Invest. Dermatol.* 126, 2019–2031.

Munier-Jolain, N., and Ney, B. (1998). Seed growth rate in grain legumes II. Seed growth rate depends on cotyledon cell number. *J. Exp. Bot.* 49, 1971–1976.

Nakai, K., and Horton, P. (1999). PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization. *Trends Biochem. Sci.* 24, 34–36.

Pandey, A., Choudhary, M. K., Bhushan, D., Chattopadhyay, A., Chakraborty, S., Datta, A., et al. (2006). The nuclear proteome of chickpea (*Cicer arietinum* L.) reveals predicted and unexpected proteins. *J. Proteome Res.* 5, 3301–3311.

Renouard, S., Cyrielle, C., Lopez, T., Lamblin, F., Lainé, E., and Hano, C. (2012). Isolation of nuclear proteins from flax (*Linum usitatissimum* L.) seed coats for gene expression regulation studies. *BMC Res. Notes* 5:15. doi: 10.1186/1756-0500-5-15

Repetto, O., Rogniaux, H., Firnhaber, C., Zuber, H., Küster, H., Larré, C., et al. (2008). Exploring the nuclear proteome of *Medicago truncatula* at the switch towards seed filling. *Plant J.* 56, 398–410.

Sato, N., Terasawa, K., Miyajima, K., and Kabeya, Y. (2003). Organization, developmental dynamics, and evolution of plastid nucleoids. *Int. Rev. Cytol.* 232, 217–262.

Sekine, K., Fujiwara, M., Nakayama, M., Takao, T., Hase, T., and Sato N. (2007). DNA binding and partial nucleoid localization of the chloroplast stromal enzyme ferredoxin:sulfite reductase. *FEBS J.* 274, 2054–2069.

Storozhenko, S., Inzé, D., Van Montagu, M., and Kushnir, S. (2001). Arabidopsis coactivator ALY-like proteins, DIP1 and DIP2, interact physically with the DNA-binding domain of the Zn-finger poly(ADP-ribose) polymerase. *J. Exp. Bot.* 52, 1375–1380.

Tadege, M., Wen, J. Q., He, J., Tu, H. D., Kwak, Y., Eschstruth, A., et al. (2008). Large-scale insertional mutagenesis using the Tnt1 retrotransposon in the model legume *M. truncatula. Plant J.* 54, 335–347.

Tan, F., Li, G., Chitteti, B. R., and Peng, Z. (2007). Proteome and phosphoproteome analysis of chromatin associated proteins in rice (*Oryza sativa*). *Proteomics* 7, 4511–4527.

Tanaka, M., Kikuchi, A., and Kamada, H. (2008). The Arabidopsis histone deacetylases HDA6 and HDA19 contribute to the repression of embryonic properties after germination. *Plant Physiol.* 146, 149–161.

Thompson, R., Burstin, J., and Gallardo, K. (2009). Post-genomics studies of developmental processes in legume seeds. *Plant Physiol.* 151, 1023–1029.

Ull, M. A., and Franco, L. (1986). The nucleosomal repeat length of pea (*Pisum sativum*) chromatin changes during germination. *Plant Mol. Biol.* 7, 25–31.

Usadel, B., Nagel, A., Thimm, O., Redestig, H., Blaesing, O. E., Palacios-Rojas, N., et al. (2005). Extension of the visualization tool MapMan to allow statistical analysis of arrays, display of corresponding genes, and comparison with known responses. *Plant Physiol.* 138, 1195–1204.

Varshney, R. K., Chen, W., Li, Y., Bharti, A. K., Saxena, R. K., Schlueter, J. A., et al. (2011). Draft genome sequence of pigeonpea (*Cajanus cajan*), an orphan legume crop of resource-poor farmers. *Nat. Biotechnol.* 30, 83–89.

Wang T. L., Uauy C., Robson F., and Till B. (2012). TILLING in extremis. *Plant Biotechnol. J.* 10, 761–772.

Yamamoto, Y. T., Zamski, E., Williamson, J. D., Conkling, M. A., and Pharr, D. M. (1997). Subcellular localization of celery mannitol dehydrogenase: a cytosolic metabolic enzyme in nuclei. *Plant Physiol.* 115, 1397–1403.

Young, N. D., Debellé, F., Oldroyd, G. E., Geurts, R., Cannon, S. B., Udvardi, M. K., et al. (2011). The Medicago genome provides insight into the evolution of rhizobial symbioses. *Nature* 480, 520–524.

Zhang, Y., Woodford, N., Xia X., and Hamburger, A. W. (2003). Repression of E2F1-mediated transcription by the ErbB3 binding protein Ebp1 involves histone deacetylases. *Nucleic Acids Res.* 15, 2168–2177.

Zhou, C., Labbe, H., Sridha, S., Wang, L., Tian, L., Latoszek-Green, M., et al. (2004). Expression and function of HD2-type histone deacetylases in Arabidopsis development. *Plant J.* 38, 715–724.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 22 October 2012; paper pending published: 11 November 2012; accepted:

05 December 2012; published online: 20 December 2012.
*Citation: Repetto O, Rogniaux H, Larré C, Thompson R and Gallardo K (2012) The seed nuclear proteome. Front. Plant Sci. **3**:289. doi: 10.3389/fpls.2012.00289*
This article was submitted to Frontiers in Plant Proteomics, a specialty of Frontiers in Plant Science.

Copyright © 2012 Repetto, Rogniaux, Larré, Thompson and Gallardo. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.

frontiers in
PLANT SCIENCE

REVIEW ARTICLE
published: 24 April 2013
doi: 10.3389/fpls.2013.00101

Arabidopsis peroxisome proteomics

John D. Bussell[1]*, Christof Behrens[2], Wiebke Ecke[2] and Holger Eubel[2]*

[1] Australian Research Council Centre of Excellence in Plant Energy Biology, The University of Western Australia, Crawley, WA, Australia
[2] Institute for Plant Genetics, Leibniz Universität Hannover, Hannover, Germany

Edited by:
Harvey Millar, The University of Western Australia, Australia

Reviewed by:
Harvey Millar, The University of Western Australia, Australia
Birgit Kersten, Johann Heinrich von Thünen Institute, Institute of Forest Genetics, Germany

***Correspondence:**
John D. Bussell, Australian Research Council Centre of Excellence in Plant Energy Biology, The University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia.
e-mail: john.bussell@uwa.edu.au;
Holger Eubel, Institute for Plant Genetics, Leibniz Universität Hannover, Herrenhäuser Str. 2, 30419 Hannover, Germany.
e-mail: heubel@genetik.uni-hannover.de

The analytical depth of investigation of the peroxisomal proteome of the model plant *Arabidopsis thaliana* has not yet reached that of other major cellular organelles such as chloroplasts or mitochondria. This is primarily due to the difficulties associated with isolating and obtaining purified samples of peroxisomes from *Arabidopsis*. So far only a handful of research groups have been successful in obtaining such fractions. To make things worse, enriched peroxisome fractions frequently suffer from significant organellar contamination, lowering confidence in localization assignment of the identified proteins. As with other cellular compartments, identification of peroxisomal proteins forms the basis for investigations of the dynamics of the peroxisomal proteome. It is therefore not surprising that, in terms of functional analyses by proteomic means, peroxisomes are lagging considerably behind chloroplasts or mitochondria. Alternative strategies are needed to overcome the obstacle of hard-to-obtain organellar fractions. This will help to close the knowledge gap between peroxisomes and other organelles and provide a full picture of the physiological pathways shared between organelles. In this review, we briefly summarize the status quo and discuss some of the methodological alternatives to classic organelle proteomic approaches.

Keywords: peroxisome, subcellular localization, protein:protein interaction, free-flow electrophoresis, functional proteomics, targeted quantitation of proteins

INTRODUCTION

Microbodies were discovered in the mid 1950s as a particular structure visible in electron micrographs of mouse kidney and rat liver cells (de Duve and Baudhin, 1966). The organelles were initially characterized by co-precipitating enzymatic activities and were consequently named peroxisomes due to the co-precipitation of oxidases and hydrogen peroxide metabolism (catalase) with the isolated structures (de Duve and Baudhin, 1966). Peroxisomes were discovered in plants in the late 1960s due to their association with enzymes of photorespiration (Tolbert et al., 1968, 1969) and subsequently found in most eukaryotic organisms (de Duve, 1969a). These early studies showed that peroxisomes primarily housed reactions that yielded reactive oxygen species (ROS; including oxidase reactions of β-oxidation and purine metabolism) and enzymes (e.g., catalase) to detoxify ROS. In addition, enzymes of the photorespiratory pathway were found in plant peroxisomes (**Figure 1**).

Up until about 10 years ago understanding of the basic function of peroxisomes in plants had not changed significantly from those early discoveries. Primarily, enzyme activities that completed already known peroxisomal pathways were identified. These included, for example, peroxisomal thiolase in the β-oxidation pathway of plants (Cooper and Beevers, 1969) which previously had only acyl-CoA oxidase (ACX), β-hydroxyacyl-CoA dehydrogenase (multi-functional protein, MFP) and enoyl-CoA hydratase (MFP) activities defined (de Duve, 1969a). More recently, the availability of the genomic sequence of *Arabidopsis thaliana* (Arabidopsis Genome Initiative, 2000) provided a significant boost to investigations of peroxisome function. For the first time, a whole genome could be interrogated using algorithms designed for prediction of protein subcellular localization signals (as discussed below). Moreover, comparative genomic tools have meant that homologs from unrelated species could be sought in the *Arabidopsis* genome sequence.

Experimental approaches aimed at elucidating the protein content of a cellular compartment require its isolation. Compared to other plant compartments, chloroplasts and mitochondria were relatively easy to obtain in fractions of good purity and detailed subcellular proteomes were obtained from these organelles soon after annotation and publication of the *Arabidopsis* genome. To date, around 800 proteins have been reported from proteome

Abbreviations: 1D, one-dimensional; 2D, two-dimensional; 2,4-D, 2,4-dichloroacetic acid; 2,4-DB, 2,4-dichlorobutyric acid; ACX, acyl-CoA oxidase; BA, benzoic acid; BSA, bovine serum albumin; CDS, coding sequence; co-IP, co-immunoprecipitation; DRP, dynamin-related protein; ER, endoplasmic reticulum; ERPIC, ER/peroxisome intermediate compartment; ESI, electrospray ionization; EST, expressed sequence tag; FFE, free-flow electrophoresis; FIS, fertilization-independent seed; GFP, green fluorescent protein; GPK, glyoxysomal protein kinase; IAA, indole-3-acetic acid; IBA, indole-3-butyric acid; ICAT, isotope coded affinity tags; ICL, isocitrate lyase; IDHP, NADP-dependent isocitrate dehydrogenase; IEF, iso-electric focusing; iTRAQ, isobaric tags for relative and absolute quantitation; LACS, long chain acyl-CoA synthetase; MALDI, matrix-assisted laser desorption/ionization; MFP, multi-functional protein; MRM, multiple reaction monitoring; MS, mass spectrometry; MVA, mevalonic acid; NADK, NAD kinase; PAGE, polyacrylamide gel electrophoresis; PMF, peptide mass fingerprint; PTS, peroxisome targeting signal; RNAi, ribonucleic acid interference; RNS, reactive nitrogen species; ROS, reactive oxygen species; SA, salicylic acid; SDS, sodium dodecyl sulfate; SILAC, stable isotope labeling by amino acids in cell culture; SRM, selected reaction monitoring; TAP, tandem affinity purification; TMT, tandem mass tags

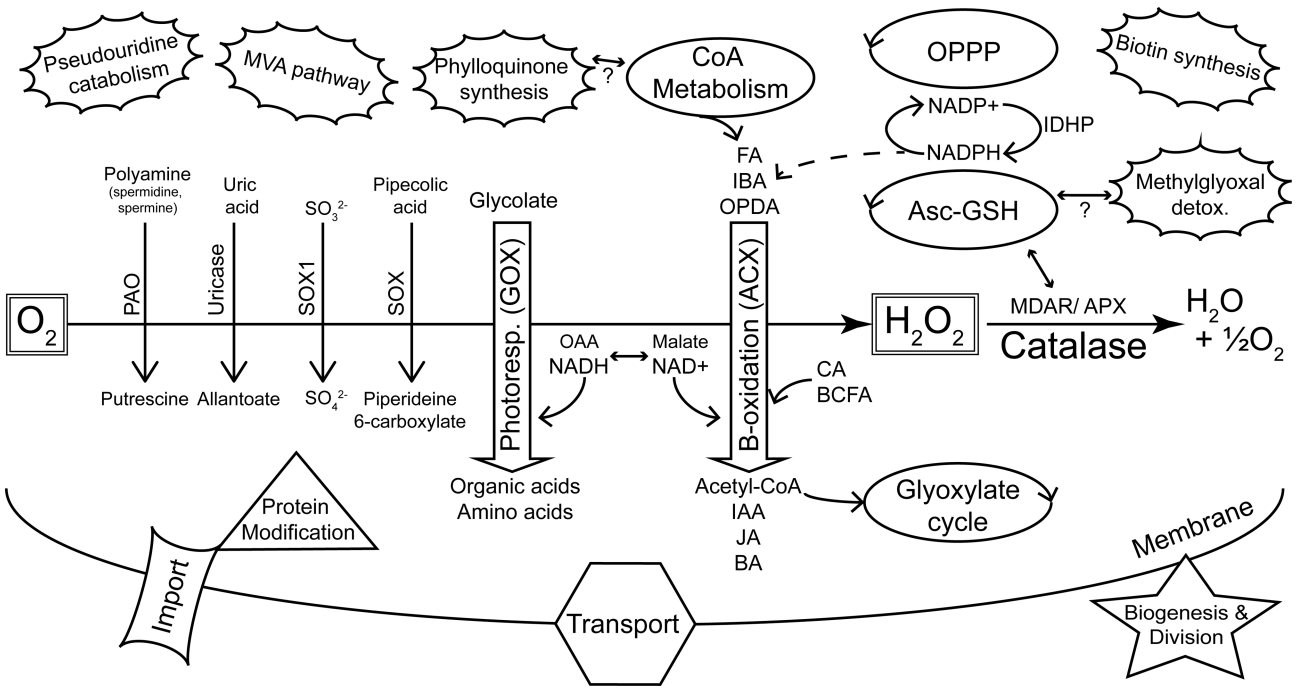


FIGURE 1 | Summary of major metabolic pathways and processes of plant peroxisomes as identified by proteomics. The peroxisome plays a key role in sequestering reactions that evolve reactive oxygen species (de Duve, 1969a). This is highlighted by the diverse oxidases (downward pointing arrows) that generate H_2O_2, and the detoxification of ROS by catalase and other enzymes. Photorespiration and β-oxidation are emphasized in the center of the figure as the major pathways. The dehydrogenases of β-oxidation are primarily dependent on NAD^+ as an electron acceptor. Hydroxypyruvate reductase (HPR) of photorespiration requires NADH and this shift in redox requirements is depicted by the interchange between OAA/NADH on the one hand and malate/NAD^+ on the other. NADPH is also required by β-oxidation for pathways including unsaturated FA catabolism and JA synthesis. Peroxisome-localized pathways (e.g., glyoxylate cycle, CoA metabolism, OPPP, Asc-GSH cycle) that can be linked with classic peroxisome metabolism are indicated as regular ellipses. The unconnected (pseudouridine catabolism, MVA pathway, biotin synthesis) or tentatively connected (phylloquinone synthesis, methylglyoxal detoxification) clouds are new additions to the list of known or proposed peroxisome-localized pathways (Reumann, 2011) and are not readily related to core peroxisome metabolism. *KEY*: *Oxidases*: PAO, polyamine oxidase; SOX, sarcosine oxidase; SOX1, sulfite oxidase; GOX, glycolate oxidase; ACX, acyl-CoA oxidase. *Substrate classes*: FA, fatty acid; IBA, indole-3-butyric acid; OPDA, 12-oxo-phytodienoic acid; CA, cinnamic acid; BCFA, branched chain fatty acid; *Product classes*: IAA, indole-3-acetic acid; JA, jasmonic acid; BA, benzoic acid; OAA, oxaloacetic acid. *Other enzymes and pathways*: MDAR, monodehydroascorbate reductase; APX, ascorbate peroxidase; Asc-GSH, ascorbate-glutathione cycle; OPPP, oxidative pentose phosphate pathway; MVA, mevalonate; IDHP, NADP-dependent isocitrate dehydrogenase.

studies to reside in mitochondria (source: SUBA3[1]; Tanz et al., 2013). Similarly, more than 2100 plastid proteins have been reported (source: SUBA3, queried on November 23, 2012). After the initial cataloging of inventory, recent years have seen mitochondria and chloroplasts being subjected to increasingly detailed functional proteomics with the emphasis shifting from discovery to dynamics (reviewed in Braun and Eubel, 2012). Although the predicted peroxisomal proteome of up to 670 proteins (see below) is clearly simpler than that of chloroplasts (>6000 predicted by ChloroP 1.1; Emanuelsson et al., 1999) or mitochondria (>4000 by TargetP 1.1; Emanuelsson et al., 2000), the difficulties involved in obtaining a pure fraction have severely limited the progress in identifying its true components by MS. It was not until 2007 with significant refinements in peroxisome fractionation techniques and the resultant improvement in the quality of the fractions that large-scale proteomics experiments involving peroxisomes became possible (Reumann et al., 2007, 2009; Eubel et al., 2008).

[1]http://suba.plantenergy.uwa.edu.au/ (queried on November 23, 2012)

In this review, we will document peroxisome proteome methodologies before arguing that the basic inventory of the peroxisomal proteome is now reasonably well covered, allowing the move toward quantitative and functional studies.

PROTEOMIC STUDIES OF THE *Arabidopsis* PEROXISOME

APPROACHES

In total, five studies have been published with the specified aim of identifying peroxisomal proteins of *Arabidopsis* (Fukao et al., 2002, 2003; Reumann et al., 2007, 2009; Eubel et al., 2008). The experimental strategies involved in the isolation of organelles and identification of proteins are summarized in **Figure 2**. Together, these efforts produced a non-redundant list of 204 proteins (source: SUBA, November 23, 2012) but many more are predicted to be located in peroxisomes (Reumann et al., 2004; Reumann, 2011). Conversely, it is likely that at least some of those that have been identified are contaminants.

Nishimura and colleagues' pioneering studies in *Arabidopsis* proteomics (Fukao et al., 2002, 2003) reflect the state-of-the-art of proteomics at the beginning of twenty-first century. Using

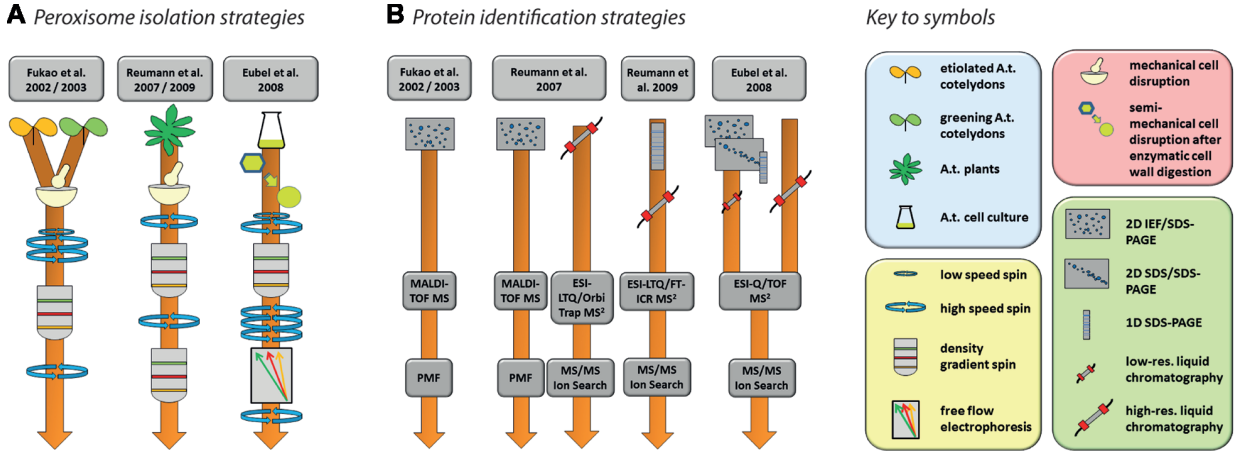


FIGURE 2 | Strategies for the purification of *Arabidopsis* peroxisomes (A) and the identification of proteins (B) as performed by Fukao et al. (2002, 2003), Reumann et al. (2007, 2009), and Eubel et al. (2008). Please refer to the original publication for more detailed information on materials and methods.

peptide mass fingerprinting (PMF) of proteins separated by two-dimensional iso-electric focusing/sodium dodecyl sulfate polyacrylamide gel electrophoresis (2D IEF/SDS-PAGE), 29 proteins were identified from leaf peroxisomes of greening *Arabidopsis* cotyledons (Fukao et al., 2002), while 19 proteins were found in glyoxysomes of etiolated cotyledons (Fukao et al., 2003). The latter study identified glyoxysomal protein kinase 1 (GPK1), a peroxisome-localized protein kinase. The authors were able to show that a number of glyoxysomal proteins are potentially regulated by phosphorylation events (see Section "Protein Modifications" for further details). Interestingly, the overlap between the two studies consisted of only three proteins, suggesting considerable differences in the protein content of *Arabidopsis* leaf peroxisomes and glyoxysomes. In total, these two studies identified less than 50 proteins. This number includes contaminants as assessed by the combination of high abundances of these proteins in other organelles, the absence of peroxisome targeting signals (PTS; see below), and no obvious relation to expected metabolic activities in peroxisomes.

Since the use of 2D IEF/SDS-PAGE for protein separation actively selects against the identification of membrane proteins for technical reasons, subsequent studies also employed alternative approaches capable of identifying membrane proteins. Reumann et al. (2007) used 2D IEF/SDS-PAGE complimented by shotgun MS to increase the analytical depth of their study of *Arabidopsis* rosette leaf peroxisomes. In addition, they also established a new purification method employing two successive gradients (**Figure 2A**). The first gradient was unusual for organelle separations in that the lysate was spun through a zone of Percoll placed on top of three density layers in which the Percoll concentration was successively reduced toward the bottom while sucrose concentration increased concomitantly. The peroxisomes were retrieved from the bottom of this gradient and then further purified through a second gradient made of seven layers of sucrose solutions with increasing density. The peroxisomes formed a clearly visible white band in the lower part of this gradient. The same isolation procedure on the same tissue was used in a follow-up study (Reumann et al., 2009). In this second study, identification of new proteins was facilitated by first running the peroxisomal proteins in a single lane of a SDS gel, then cutting the lane into 16 slices, each of which was submitted to high-resolution tandem MS (**Figure 2B**). A large proportion of the newly identified putative peroxisomal proteins were then tested for subcellular localization by yellow fluorescent protein fusions (Reumann et al., 2009).

In contrast to the two Reumann studies on *Arabidopsis* leaf peroxisomes, Eubel et al. (2008) approached elucidation of the proteome of plant peroxisomes by choosing non-green *Arabidopsis* cell suspension cultures, thereby eliminating chloroplasts as a major source of contamination in the peroxisomal fraction. With the aim of increasing organelle purity even further, a Percoll gradient was followed by free-flow electrophoresis (FFE; reviewed by Islinger et al., 2010; **Figure 2A**). By using subfractionation of peroxisomes as well as gel (1D and 2D) and non-gel approaches, further novel proteins, including hydrophobic proteins, were discovered by tandem MS (**Figure 2B**).

CHALLENGES AND PITFALLS

Arabidopsis has been the preferred option for subcellular plant proteomics. Isolates of moderate to good quality can readily be obtained for plastids, mitochondria, plasma membrane, and other organelles and compartments from green tissue or non-green suspension cell cultures. Unfortunately isolating good peroxisomal fractions is notoriously difficult in *Arabidopsis*. The reasons for this are likely to be due to some or all of the following: (1) the high content of secondary metabolites found in Brassicaceae potentially interferes with isolation (Kaur and Hu, 2011; Reumann, 2011); (2) peroxisomes are considered to be present in much lower numbers in cells than, for example, mitochondria or chloroplasts (e.g., see Germain et al., 2001; Palma et al., 2009); (3) peroxisomes are thought to be particularly fragile *in vitro* (Palma et al., 2009; Reumann, 2011); (4) losses can be expected to occur due to stresses imposed on the organelles during the isolation procedure; and (5) peroxisomes often physically interact with other organelles (e.g., chloroplasts; Reumann, 2011) and, because of

their fragility, using enough force to break these associations may lead to additional peroxisome damage. As such, highly specialized methods were developed for the isolation of organelles from this species that require sophisticated equipment (e.g., FFE) and/or highly optimized procedures as well as detailed knowledge about the stumbling blocks associated with them. Low yields are typical for peroxisome preparations from *Arabidopsis*; this further reduces the error margins for successful preparation of peroxisomes and compromises the purity of peroxisome fractions.

The advantages inherent to *Arabidopsis* (genetic resources, short generation time, easy cultivation, research history) at least partly outweigh the difficulties associated with isolating its peroxisomes and persistence has eventually resulted in acceptable proteomic outcomes. We suggest that a big step forward for peroxisomal proteomics would be if a more generally accessible and well-described isolation technique for *Arabidopsis* were available. Also, we envisage that species such as spinach, pumpkin, soy, and castor oil bean that were all used as early plant models for peroxisome studies will again rise to be of prominence in peroxisome proteomic studies due the increasing availability of sequenced, annotated genomes and their amenability to be used as model species.

THE *Arabidopsis* PEROXISOMAL PROTEOME

The Arabidopsis 2010 peroxisome project[2] compiled a "parts list" of 133 confirmed *Arabidopsis* peroxisomal proteins, most of them validated by localization of fluorescent reporter fusion proteins. Although curation of this list has ended, an updated version comparing the *Arabidopsis* proteome with that of rice has recently been published (Kaur and Hu, 2011). This list[3] includes 163 proteins that fulfill at least two of the criteria of (a) having been identified by MS, and having had their peroxisomal location supported by either (b) the presence of a PTS or (c) localization of fluorescent reporter fusion proteins. As might be expected, this list contains proteins involved in β-oxidation of fatty acids, auxiliary β-oxidation pathways, photorespiration, ROS detoxification, the glyoxylate cycle, branched chain amino acid metabolism, and peroxisome proliferation. It also contains a considerable number of "less traditional" peroxisomal proteins and others with unknown functions, but surprisingly few matrix protein import components are represented. These groups of proteins are summarized in **Figure 1** and will be discussed below in terms of their proteomic coverage and the scope to improve this coverage.

It is particularly noteworthy that the main peroxisomal pathways of *Arabidopsis* (**Figure 1**), as well as almost all of the secondary pathways represented by these 163 proteins, are also among the putative rice peroxisome proteins, at least as evidenced by the presence of rice homologs possessing peroxisome targeting signals (Kaur and Hu, 2011). Such conservation between plant genomes as diverged as rice and *Arabidopsis* suggests that this set of proteins can be taken to comprise the core plant peroxisomal proteome. Indeed it is fair to say that (occasional future surprises notwithstanding) the basic function and proteome of plant peroxisomes is now well established and that this fundamental knowledge will foster more advanced proteomic studies in the future.

PEROXISOME EVOLUTION, BIOGENESIS, AND PROTEIN IMPORT

Debate on the origin of peroxisomes has centered on two competing hypotheses. Thus, peroxisomes could either have originated as discrete (single-) membrane-bound structures in primitive eukaryotes or alternatively as an engulfed endosymbiont similar to nascent mitochondria and plastids (de Duve, 1969b). The evolution of peroxisomes remained a hotly debated field well into the 2000s. It was only the recent discovery in baker's yeast (*Saccharomyces cerevisiae*) of *de novo* peroxisome biogenesis from endoplasmic reticulum (ER) that provided strong evidence against an endosymbiont hypothesis for their origin (Hoepfner et al., 2005). As well as budding directly from ER (van der Zand et al., 2012), peroxisomes can proliferate by division of existing organelles and they frequently receive vesicles from the ER that add to the peroxisomal membrane and carry proteins destined for the peroxisomes. The current model of peroxisome biogenesis in plants considers both options of organelle genesis and, accordingly, is referred to as the "ER semi-autonomous peroxisome maturation and replication model" (Mullen and Trelease, 2006). According to this model, in addition to the direct import of proteins synthesized in the cytosol, proteins can also be transported to peroxisomes directly from the ER via the proposed ER/peroxisome intermediate compartments (ERPICs). Understanding of the basics of protein import machinery in plants followed discoveries in mammalian and yeast systems and was largely dictated by the availability of genome sequences that facilitated gene mining for homologs from these systems (Baker and Sparkes, 2005). Genes encoding the essential import and biogenesis machinery are largely conserved across kingdoms indicating that these processes are ancient evolutionary innovations. Peroxisome *de novo* biogenesis involves peroxins PEX3, PEX16, and PEX19 that are all at some time associated with the peroxisomal membrane. Division in *Arabidopsis* involves five different isoforms of PEX11(a–e), at least two fertilization-independent seed (FIS) proteins and three dynamin-related proteins (DRPs) which again are at some time peroxisome-localized (Lingard and Trelease, 2006; Lingard et al., 2008; Hu et al., 2012).

The majority of peroxisome matrix proteins are imported directly from the cytosol by import machinery that recognizes PTS on the proteins. Two types of PTS peptides are responsible for protein targeting to peroxisomes. About 75% of peroxisome-targeted proteins have a so-called PTS1: a C-terminal tripeptide comprised of a non-polar residue in position-1, a basic amino acid in position-2 and a small and uncharged amino acid in position-3 (Reumann, 2004; Lingner et al., 2011). Serine–lysine–leucine (SKL) is the canonical PTS1 but many possible amino acid combinations can function as a PTS1. Improvements in proteomic techniques (e.g., by increased sensitivity of mass spectrometers which produce higher proteome coverage) and confirmation of targeting by other methods (such as targeting of fluorescent fusion proteins) have led to considerable refinement of the definition of the plant PTS1, recognition of the importance of the sequence context of the tripeptide and expansion in the range

[2]http://www.peroxisome.msu.edu/

[3]http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3355810/table/T1/

of permissible PTS1 sequences (Reumann et al., 2012). The second type of peroxisome targeting signal, PTS2, is a nonapeptide usually located within the 20–30 N-terminal residues of proteins that utilize this signal (Reumann, 2004). The consensus PTS2 sequence consists of two pairs of conserved amino acids separated by five non-conserved amino acids and in plants is R[ILQ]x5HL (Kato et al., 1998). As is the case for PTS1s, more recent studies have refined and expanded the range of import-capable PTS2 peptides (e.g., Simkin et al., 2011). PTS2 sequences may also in rare instances be found in the body of the protein (Reumann et al., 2007; Lamberto et al., 2010) and include alternative residues at the second and ninth positions (Kaur and Hu, 2011).

PTS1 tripeptides are recognized by the PEX5 protein that docks with its cargo and is then imported by membrane-bound assembly of PEX13 and PEX14, which form the core import channel. PTS2 import utilizes PEX7 that must first interact with PEX5 before import via the same channels (reviewed in Hu et al., 2012). In both cases, the PEX5/PEX7-cargo complex is imported as a whole into the peroxisomal matrix, with PEX5/PEX7 recycled to the cytosol by a mechanism involving ubiquitination of PEX5 and that utilizes membrane-bound proteins PEX22, PEX2, PEX10, PEX12, and APEM9 as well as other PEX proteins (PEX4, PEX6, and PEX1) that are tethered on the cytosolic side to various protein components of the receptor recycling machinery. After import into peroxisomes, PTS2 sequences are cleaved by the trypsin-like DEG15 protease (Schuhmann et al., 2008); such modifications do not occur with PTS1 signals.

Proteomic analysis by MS has to date been almost singularly unsuccessful at detecting peroxisome-localized components of biogenesis, division and import pathways in *Arabidopsis*. Of the at least 12 protein import-related proteins, only PEX14 was identified in recent *Arabidopsis* studies (Reumann et al., 2007, 2009; Eubel et al., 2008). Similarly, with the exception of PEX11 isoforms (see Reumann et al., 2007, 2009; Eubel et al., 2008), peroxisome biogenesis and division related proteins have not been found by MS-proteomics. Possible reasons for the lack of success in isolating most of the membrane-bound proteins involved in peroxisome biogenesis and protein import in plants are their very low abundances, life stage specific expression, transient associations with peroxisomes, difficulties in the isolation and analysis of peroxisome membranes, and selection against the identification of membrane proteins due to technical limitations (e.g., the use of 2D IEF/SDS-PAGE).

PROTEINS OF THE CORE PEROXISOME METABOLISM

Peroxisome metabolism is important to all stages of plant growth, including seed development, germination, general growth, and senescence. Peroxisomes can be seen essentially as organelles that sequester reactions producing reactive oxygen species and facilitate benign detoxification of ROS. Other reactions that occur in peroxisomes have almost always been shown to comprise a "service-industry" that recycles reaction intermediates and cofactors. Thus, peroxisomes are primarily β-oxidation (ACX) and photorespiration (glycolate oxidase) machines (**Figure 1**) and the metabolism from these pathways is highly integrated internally and with the rest of the cell. Peroxisomes also house a number of other oxidases that catalyze reactions that are less obviously integrated with "the model peroxisome" metabolism (e.g., sarcosine oxidase, sulfite oxidase, polyamine oxidase, copper amine oxidase, hydroxy acid oxidase and uricase (urate oxidase) (**Figure 1**).

β-Oxidation provides the primary metabolism remobilizing fatty acids from the oil bodies of oil seed species (such as *Arabidopsis*) to supply energy to seedlings in the initial phases of growth before the onset of photosynthetic energy supply. The sucrose dependence of many β-oxidation pathway mutants is testimony to this (Graham, 2008). The pathway is also responsible for the recycling of fatty acids in senescent plant tissue (Kunz et al., 2009). Peroxisomal β-oxidation is also employed in the modification of other compounds, phytohormones in particular. *De novo* jasmonic acid (JA) synthesis must pass through peroxisomes and precursors of this lipid-derived hormone undergo three cycles of β-oxidation before being exported to the cytosol for conversion to more active forms (Baker et al., 2006). Similarly, indole-3-butyric acid (IBA), regarded as a storage form of auxin, passes through a single round of β-oxidation to form the bioactive indole-3-acetic acid (IAA; Baker et al., 2006). There are numerous peroxisome mutants that display resistance to pro-auxins as a consequence of interruption to the pathway (e.g., Wiszniewski et al., 2009; Strader et al., 2011). Remobilizing IBA to IAA provides at various stages of plant development for a rapid or controlled increase in the pool of bioactive IAA, for example, in germinating seedlings (Strader et al., 2011). Finally, previous suggestions (Boatright et al., 2004; Baker et al., 2006; Kliebenstein et al., 2007) of a peroxisomal contribution to benzoic acid (BA; and as a consequence salicylic acid, SA) synthesis have found new support in the observation of reduced accumulation of BA and BA-CoA in knockouts or ribonucleic acid interference (RNAi) lines of peroxisome-targeted β-oxidation enzymes (Klempien et al., 2012; Lee et al., 2012; Qualley et al., 2012). Given this emphasis on β-oxidation for the organelle, and the increasing diversity of molecules that are altered by the pathway, it is perhaps not surprising that of the total peroxisomal proteome, at least 25% of the predicted (Pracharoenwattana and Smith, 2008) and confirmed (Kaur and Hu, 2011) peroxisomal proteins are directly involved in β-oxidation. This includes glyoxylate cycle enzymes [isocitrate lyase (ICL) and malate synthase], core β-oxidation activities such as acyl-activating enzymes (AAE), ACX, MFP, 3-ketoacyl-CoA thiolase (KAT), and the many subsidiary enzymes in these multigene families such as single-function hydratases, epimerases, and short chain dehydrogenases. This figure of 25% does not include enzymes outside the core pathway such as those of ROS detoxification, malate dehydrogenase (MDH), peroxisomal adenine nucleotide carrier (ANT), enzymes of CoA metabolism, and others that are necessary to sustain β-oxidation and that collectively would clearly significantly increase this proportion.

Once seedlings are established, the primary role of plant peroxisomes is their participation in the photorespiratory pathway. This process serves the detoxification of 2-phosphoglycolate produced by the oxygenase reaction of RubisCO and the salvage of the bulk of its carbon atoms. While the photorespiration pathway is spatially split between chloroplasts, mitochondria, and peroxisomes, the majority of the enzymes directly involved in the recovery of phosphoglycolate (after removal of the Pi group in the chloroplast)

are peroxisomal. Six enzymes of photorespiration (including catalase) are located in the peroxisome and at least six different substrates have to be transported across the peroxisome membrane for photorespiration to function properly. This renders peroxisomes the central hub in a biochemical pathway that is so important for the central carbon metabolism of plants that, in terms of carbon flux, it is surpassed only by photosynthesis (Bauwe et al., 2010) and it is no surprise that these proteins are prominently represented in proteomic studies. Despite substantial efforts to reduce the ostensibly wasteful process of photorespiration (e.g., Kebeish et al., 2007; Peterhansel and Maurino, 2011), it is becoming increasingly clear that it is a necessary prerequisite for photosynthesis in C3 plants, that limiting photorespiration inevitably limits photosynthesis itself (Heineke et al., 2001) and that increasing the abundance of photorespiratory enzymes leads to higher rates of photosynthetic carbon fixation and accelerated plant growth (Timm et al., 2012).

PROTEINS OF OTHER PEROXISOME METABOLIC PATHWAYS

While peroxisomes carry out many hazardous reactions (usually by the actions of ROS-producing oxidases and, to a far lower extent, by reactive nitrogen species (RNS; reviewed in Palma et al., 2009) that need to be shielded away from the rest of the cell, newer data indicate that a range of non-hazardous reactions and pathways also operate in peroxisomes. Potential peroxisomal functions such as protection from herbivore and pathogen attack (Reumann, 2011) are subject of ongoing research.

Aside from ROS-evolving reactions/pathways, peroxisomes have been implicated in co-factor metabolism (CoA), methylglyoxal detoxification, pseudouridine catabolism, and phylloquinone biosynthesis (reviewed in Reumann, 2011). Also, nicotinamide adenine dinucleotide phosphate (NADPH) recycling [including by IDHP (NADP-dependent isocitrate dehydrogenase), oxidative pentose phosphate pathway, the glutathione cycle and NADPH synthesis by NADK3 reviewed in Kaur and Hu, 2011], the initial step of biotin synthesis (Kaur and Hu, 2011) and the mevalonic acid (MVA) pathway of isoprenoid synthesis (Sapir-Mir et al., 2008; Simkin et al., 2011) may take place in peroxisomes (**Figure 1**). Peroxisomal localization of proteins in these pathways has been predicted by informatics and in many cases their targeting potentials were confirmed by green fluorescent protein (GFP) studies. Moreover, with the exception of the MVA pathway enzymes, many of them have also been resolved in proteomic studies (Kaur and Hu, 2011). Since for these non-hazardous pathways distribution to different compartments does not make as much sense as for the ROS-producing processes, one could speculate that the bulk of the reactions carried out in these pathways might also be confined to peroxisomes. At least those proteins that have so far not been positively assigned to any other organelle can be regarded as potential candidates for peroxisome localizations.

Exactly how these processes contribute to peroxisome function and metabolism, or for that matter why they are localized to peroxisomes remains an open question. For example, two enzymes of pseudouridine catabolism (pxPfkB/At1g49350 and IndA/At1g50510) have been found in two different proteome studies, and have been confirmed to have functional PTS peptides (see Reumann, 2011). However, there is no obvious reason why such metabolism should occur in peroxisomes. The substrates and products of the reactions would require transport into the organelle and they do not play any other role in known peroxisomal metabolism. Reumann (2011) suggested that the peroxisome might provide a venue for RNA catabolism away from the actual function and synthesis of RNA. It seems likely that until peroxisome metabolism is better understood, this kind of speculation provides a working model for peroxisomal localization for some of the "non-toxic" pathways.

PROTEIN MODIFICATIONS

Peroxisomes also modify imported proteins. Such modifications include DEG15 dependent cleavage of PTS2 peptides (Helm et al., 2007; Schuhmann et al., 2008), phosphorylation via the GPK1 kinase (Fukao et al., 2003) and degradation of enzymes such as malate synthase and ICL of the glyoxylate cycle during the transition of peroxisomes from fatty acid degrading to photorespiratory organelles (Lingard et al., 2009). Moreover, the *Arabidopsis* LON2 protease appears to be involved in maintaining matrix protein import by a yet to be determined mechanism (Lingard and Bartel, 2009). It would be particularly interesting to determine the phosphoproteome of peroxisomes and then follow the regulation of phosphatase/kinase activities in the peroxisome. The aforementioned GPK1 and the prediction and observation of a number of other proteases and kinases localized to peroxisomes (Kaur and Hu, 2011) suggest dynamic phosphorylation and dephosphorylation of peroxisomal proteins. At the time of writing, 12 confirmed peroxisomal proteins (At1g04710, At1g06290, At1g07180, At1g20480, At1g20620, At1g20630, At1g23310, At1g49350, At1g54340, At1g65520, At2g06050, and At1g49670) are listed in the PhosPhAt database[4]. Furthermore, PMP38/PXN (AT2G39970, a peroxisome membrane-localized NAD^+ transporter; Bernhardt et al., 2012), was also found to be phosphorylated (Eubel et al., 2008). We suggest that targeted proteomics could play an important role in analyzing mutants in these protein modification and phosphorylation/kinase genes to determine substrates and extent of processing.

PEROXISOME PROTEOMICS IN THE FUTURE

With the knowledge gained from previous studies, *Arabidopsis* peroxisome proteomics can be expected to venture in the following directions: (1) search for novel peroxisomal proteins and (2) detailed analyses of peroxisomal proteins in respect to dynamic changes of protein abundances, modifications of peroxisomal proteins, and their potential interactions with other proteins to form temporary or stable protein complexes. Several routes with the potential to further populate the list of plant peroxisomal proteins are discussed and these approaches may be tied to functional proteomic characterization as outlined below.

IDENTIFICATION OF NOVEL PEROXISOMAL PROTEINS

Refinement of peroxisome isolation methodology

In general, allocating proteins to the peroxisomal compartment by experimental data (MS or other methods such as localization of fluorescent protein fusions) provides greater confidence

[4]http://phosphat.mpimp-golm.mpg.de/phosphat.html (queried on March 8, 2013)

than predictions alone. However, often MS data and even fluorescent reporter proteins produce false-positive results. In the case of MS, the presence of some of the proteins in the data is likely to be due to contamination of the peroxisomal fraction with other organelles. Usually, proteins from contaminating organelles are present in low numbers in the target fraction, but nonetheless they may swamp the analytical sensitivity needed to isolate low abundance proteins of the target organelle. Alternative means are therefore required to isolate organelles to a higher level of purity, both to reduce contamination and to increase the chance of finding rare proteins. One such strategy could be the use of FFE. FFE has been used to further purify mitochondria from yeast and plants, as well as rat peroxisomes (reviewed in Islinger et al., 2010). It has also been successfully employed for the isolation of peroxisomes from *Arabidopsis* cell suspension cultures. Starting from material heavily contaminated with mitochondria, FFE was able to reduce the mitochondrial content by a factor of 5 (based on oxygen electrode assays) while increasing the concentration of peroxisomes threefold (Eubel et al., 2008). Although FFE will face a second major source of contamination in fractions prepared from green leaves (plastids in addition to mitochondria), this approach might also improve the purity of peroxisomes prepared from green plant tissue. In combination with the gradient(s) established for leaf material (Reumann et al., 2007), FFE has the potential to deliver peroxisomes with very low levels of contamination by mitochondria or plastids.

While FFE might be able to reduce contamination, it cannot compensate for the low yields typical of plant peroxisome preparations. The use of more starting material in existing protocols will most probably not produce more isolate since time seems to be an especially critical factor for peroxisome isolation. More starting material means longer preparations and therefore also higher losses. A higher abundance of peroxisomes in the plant material could support higher yields without the need to modify existing procedures. Abiotic stress has been reported to increase the number of peroxisomes per cell (Zhu, 2002) and, while such treatments will inevitably alter the physiology of the plant cell and its organelles, they may represent a viable option to increase yield if the integrity of the organelles is not affected. With these considerations in mind, perhaps the best aim is to counteract peroxisomal breakdown during the isolation procedure. Conditions that stabilize the organelles such as high osmotic strength in the isolation buffers and avoidance of pelleting steps have been reported to be successful (Reumann et al., 2007, 2009). In addition, special care is usually taken to reduce, inhibit, or divert protease activity away from the target proteins in organelle preparations destined for proteomic analyses. However, the same cannot be said for lipid-degrading enzymes set free during cell disruption, which can directly contribute to membrane degradation especially during the first stages of the isolation procedure. This may compromise integrity of the organelles and, finally, the yield of the preparation. Broad-band inhibitors for phospholipases are commercially available but largely omitted in organelle preparations. The use of protease inhibitors in combination with sacrificial phospholipase substrates (such as choline and ethanolamine; Scherer and Morre, 1978) may therefore help to maintain organelle integrity and consequently may also improve peroxisome yield for *Arabidopsis* peroxisome proteomic studies. However, choline and ethanolamine may interfere with subsequent FFE.

Bioinformatic approaches to predict novel peroxisomal proteins

Bioinformatics has provided considerable insight into likely upper limits to the size of the *Arabidopsis* peroxisomal proteome. Up to 542 proteins are predicted to contain a PTS1 sequence (SUBA3, November 23, 2012, queried using the PredPlantPTS1 search algorithm of Reumann et al., 2012) and around 110 additional proteins are potentially targeted to peroxisomes by a PTS2, membrane PTS or other means (see Reumann et al., 2004; Kaur and Hu, 2011). The number of potential peroxisomal proteins has expanded considerably in the last 10 years. Reumann et al. (2004) provided the first comprehensive database of putative *Arabidopsis* peroxisome proteins (Araperox[5]), listing 284 proteins based on prediction of PTS1 and PTS2 targeting peptides. Araperox was updated in 2008 to 440 proteins, including another 110 proteins with the newly demonstrated PTS1 signals SSL, SSI, ASL, and AKI (Lisenbee et al., 2005; Reumann et al., 2007), PEX proteins, demonstrated membrane proteins and proteins that are imported using non-standard targeting peptides (e.g., catalase and sarcosine oxidase; Goyer et al., 2004; Oshima et al., 2008). The number of predicted proteins has further increased by refinements to (in particular) PTS1 prediction algorithms (Ma and Reumann, 2008; Lingner et al., 2011; Chowdhary et al., 2012). These studies have taken particular consideration of the context (upstream) of the putative PTS1s and thus identified weak, non-canonical PTS sequences that supported import of proteins into plant peroxisomes. Extensive testing of 23 newly predicted PTS1 motifs suggested unforeseen diversity in plant peroxisome-import competent C-terminal tripeptides (Lingner et al., 2011). The majority of these were tested by fusing enhanced yellow fluorescent protein (EYFP) to the 10 terminal amino acid residues of the predicted proteins or, in a few cases, to the N-terminus of full-length proteins.

While it is possible that the peroxisomal proteome might contain more matrix proteins than are currently predicted, it is equally clear that some predicted peroxisomal proteins are unlikely to localize to the organelle *in vivo*. As a trivial example, the plastid genome encoded rpoC2 (RNA pol) has a strong PTS1 like sequence (SRI) at its C-terminus. In total, 204 proteins have been assigned to the peroxisomal compartment by MS but only 97 of these are found in the list of ~670 proteins predicted (or confirmed by other means) to reside in peroxisomes. This small overlap (<15% of the predicted and <50% of the MS-detected proteins) implies that both the prediction and detection of peroxisomal proteins suffer from false-positive results. It also seems likely that there will be more surprises in the form of unexpected proteins to be assigned to this organelle. Further research and refinement of prediction algorithms (especially for PTS2 sequences) will yield more candidate proteins for the peroxisomal proteome. These bioinformatic works represent major advances in setting the outer boundaries for the plant peroxisomal proteome.

An alternative bioinformatic approach has been to mine the wealth of publicly available transcriptome data for genes that are

[5]http://www3.uis.no/araperoxv1/

co-transcribed with genes encoding proteins for core peroxisome functions. This approach was taken in Wiszniewski et al. (2009) to show that numerous previously uncharacterized PTS-encoding β-oxidation genes followed similar patterns of transcriptional expression to other, well-characterized β-oxidation genes. Such data mining for proteins involved in other areas of peroxisome biology could help to provide clues for timing of expression and the function of novel or thus-far undetected putative peroxisomal proteins.

Genetic resources

Two classic screens for mutants affected in peroxisome function have been used extensively and both screens can reveal mutants compromised in β-oxidation (Hayashi et al., 1998; Zolman et al., 2000). The first uses the peroxisome-localized conversion of pro-auxins [IBA or 2,4-dichlorobutyric acid (2,4-DB)] into active forms [IAA or 2,4-dichloroacetic acid (2,4-D)]. Mutants in these pathways exhibit root elongation when grown on media containing IBA or 2,4-DB, whereas growth of wild-type roots is inhibited. The second screen utilizes the requirement for β-oxidation to release carbon from fatty acids to fuel germination and seedling establishment. Lesions in this pathway result in seedlings that either do not germinate or fail to establish unless they are grown on media that is supplemented with a sugar carbon source. These early forward genetics studies identified a number of single gene mutants of large effect in these pathways (CTS/PXA1/PED3, KAT2/PED1, etc.). However, many β-oxidation proteins are encoded by gene families that exhibit functional redundancy, and this necessitates targeted (reverse genetic) generation of double- and potentially higher order mutants [e.g., ACX, long chain acyl-CoA synthetase (LACS) families].

Two studies have taken a brute force approach to genetic characterization of putative peroxisome genes. In the first (Wiszniewski et al., 2009), all available mutants for 16 newly predicted (or otherwise uncharacterized) genes with similarity to known β-oxidation genes (as reported in Araperox; Reumann et al., 2004) were screened for response to growth on IBA, 2,4-DB, and sugar-free media. This study yielded new genes in auxin metabolism pathways, but suggested that there were no new single gene knockouts of peroxisome proteins that would display a sugar dependence phenotype. The study also showed that the new auxin pathway genes followed a transcriptional pattern common to many β-oxidation genes. Secondly, Kaur and Hu (2011) hint at a large, unpublished study that took a similar approach with about 50 predicted peroxisome genes and that involved various biochemical, physiological, and cell biological assays aimed at documenting the role of these genes in embryogenesis, peroxisomal protein import, and defense response.

Identification of proteins interacting with bona fide peroxisomal proteins

An alternative experimental approach to defining localization is to identify interacting proteins, since these are very likely to be localized in the same compartment. By using an approach not biased by previously defined parameters for peroxisomal targeting, the verification of peroxisome localization of interacting proteins potentially leads to the discovery of so far unknown plant peroxisome proteins and new definitions of functional PTS sequences.

A suite of techniques is available to isolate protein complexes. Antibody- or affinity-tag-based techniques such as co-immunoprecipitation (co-IP) or co-precipitation have been flagged to be of likely utility in expanding proteomic knowledge (Yates, 2000). By these methods, a protein of known localization is used as bait and incubated with protein extracts from fractionated samples to permit interactions and formation of complexes. Subsequently, an antibody against the known peroxisomal protein (e.g., bound to the resin of an affinity chromatography column) is used to pull down the complex, which can be denatured and run on a gel or subject directly to MS detection. An extension of these methods, tandem affinity purification (TAP), has shown particular promise in proteomic applications. TAP involves a two-step affinity purification method that may significantly reduce the incidence of non-specific binding, thus resulting in greater purity of the isolate (Rigaut et al., 1999). The potential for these methods to be used in expanding documentation of the plant peroxisomal proteome is discussed briefly below.

Co-IP has been used in mammalian cells (rat liver) to isolate and confirm interaction of a matrix protein (L-bifunctional enzyme, L-BFE) with catalase (Makkar et al., 2006). Likewise, it was used to demonstrate interactions between a subset of yeast (*S. cerevisiae*) PEX proteins (Eckert and Johnsson, 2003). However, these clearly do not represent high-throughput discovery studies. In a more generic approach, antibodies against a PTS could be used in pull-down experiments. For example, anti-luciferase-PTS1 antibody detects multiple proteins on western blots of purified peroxisome fractions from rat livers (Gould et al., 1990). Using such an antibody in co-IP studies could conceivably yield new proteins, but the antibody is likely to be specific to the particular C-terminal tripeptide (SKL in this case; Gould et al., 1990) and the growing diversity of PTS1 sequences as documented above may limit this approach. Isolation of PTS2-containing proteins would require another suite of antibodies.

TAP has been promoted as a high-throughput method for protein complex discovery. An early application of TAP was a large-scale analysis of protein complexes in yeast (*S. cerevisiae*) in which TAP tags were directly attached to the C-termini of 1739 proteins (Gavin et al., 2002). In total, 589 tagged proteins were purified, 78% of which were associated with potential interaction partners. The experiment involved whole cell extracts, but in principle the interactions occurred *in vivo* and most can therefore be expected to be compartment specific. Thus, proteins and interacting partners were assigned to subcellular compartments, but peroxisomes were not represented amongst them. Likewise, a later study of *S. cerevisiae* purified 2357 tagged proteins and identified over 4000 different proteins involved in interactions, but very few of these were peroxisome-localized (Krogan et al., 2006). These results may reflect (a) that peroxisomal proteins do not form complexes, (b) relatively low abundance of peroxisomal proteins compared to those successfully isolated, or (c) that C-terminal TAP attachment used in both studies masks PTS1 signals with the result that the proteins are not imported and normal complex formation is precluded.

S. cerevisiae is well suited to such TAP analysis because homologous recombination can be used to generate large libraries of strains expressing the tagged proteins at approximately native levels (under control of endogenous promoters), and N-terminal tagging could just as well be used to preserve endogenous C-terminal targeting signals. Unfortunately, translating such approaches to *Arabidopsis* and other plants has been problematic, not least because there is no method for homologous recombination of constructs into the genome. TAP tagging in plants thus requires *Agrobacterium*-mediated transformation of individually cloned constructs into wild-type plants, or into mutants in which the native gene has been knocked out to preclude competition of the untagged endogenous protein for binding. Nevertheless, TAP tagging has been promoted for protein complex discovery in plants and the method adapted for plant-specific application (Van Leene et al., 2008). Plant cell cultures have been successfully used with TAP methodology to identify cell cycle component interacting proteins but the process had numerous drawbacks including its complexity, susceptibility to false negatives (due to low abundance, transient expression, or absence of likely interactors in cell cultures) and false positives to highly "sticky" non-specific interactions (Van Leene et al., 2011).

Data on protein:protein interactions of peroxisomal proteins can also be found in recent global interactome studies. Databases such as the IntAct molecular interaction database (Lee et al., 2010; Kerrien et al., 2012) or the third version of the Arabidopsis Subcellular Database (SUBA3, Tanz et al., 2013), that now incorporates protein interactions, can be queried to detect interaction partners of peroxisomal proteins. Kaur and Hu (2011) have summarized the currently confirmed *Arabidopsis* peroxisomal proteome into a list of 163 proteins (see above). By interrogating SUBA3 (November 23, 2012) with this list, we identified a non-redundant set of 133 proteins with claims for interactions with the original set (**Figure 3A**; **Table 1**). For 35 of these, the subcellular location was already deduced from fluorescent reporter fusion proteins. Nine of them were exclusively or non-exclusively assigned to peroxisomes. MS assigned 30 other proteins to their respective intracellular locations, and nine of these were also found in peroxisomes. For the remaining 68 proteins, no experimental GFP and/or MS data are stored in SUBA: these proteins are thus candidates for peroxisomal localization. Eight of them were predicted to be peroxisomal by PredPlantPTS1 (Reumann et al., 2012) or Multiloc2 (Blum et al., 2009) and are therefore very likely imported into peroxisomes. On this basis, they should now no longer be considered as candidates but as established peroxisomal proteins (**Table 1**). The remaining 60 proteins were used as queries to interrogate SUBA3 for their interaction partners. Using the putative subcellular localization of the returned interaction partners, as given by the consensus of all localization data for each protein stored in the SUBA3 (SUBAcon), we calculated the percentage of the interacting proteins that are known to be peroxisomal compared to those localizing to other cellular compartments. For 13 candidate proteins, at least a third of the interaction partners consisted of peroxisomal proteins, while this number shrank to nine if a cut-off value of 50% was applied (**Figure 3B**; **Table 1**). Of these, six proteins had only a single interacting partner and thus by definition had a 100% interaction rate with other peroxisomal proteins: the "bait" was the only confirmed peroxisomal protein. Depending on the level of stringency applied (33, 50, or 100% peroxisomal interactors), these proteins represent the strongest candidates for peroxisomal targeting by this approach and checking their intracellular localization by other means could be a worthwhile undertaking.

FUNCTIONAL CHARACTERIZATION OF THE PEROXISOMAL PROTEOME

There is more to proteomics than a mere stocktake of the protein content of an organelle. The studies by Fukao et al. (2002, 2003) have clearly shown that we can expect differences in peroxisomes isolated from cotyledons performing autotrophic and heterotrophic metabolism. Except for these early studies (that have more of a qualitative than quantitative character), comparative studies of the plant peroxisomal proteome are, to our knowledge, non-existent. Clearly, obtaining results on changing protein abundances would have a strong impact on our current view of peroxisomes as cellular organelles and their reactions to changing physiological conditions. MS-based comparative studies are traditionally performed by using quantitative approaches such as isobaric labeling [isobaric tags for relative and absolute quantitation (iTRAQ), tandem mass tag (TMT)], heavy mass tags (isotope coded affinity tags, ICAT) and, to a lesser degree, stable isotope labeling (stable isotope labeling by amino acids in cell culture, SILAC). These techniques are quite sensitive and allow quantitation of several thousand proteins at a time. However, due to the relatively low number of peroxisomes within cells and the resulting low average abundance of peroxisomal proteins in cell lysates, coverage of peroxisomal proteins present only in low copy numbers still presents a challenge to quantitative shotgun proteomics. Therefore, isolating peroxisomes from two or more plant populations is still deemed necessary to achieve a better coverage of peroxisomal proteins in shotgun comparative approaches.

A promising alternative to this classical approach may be the use of targeted quantitation of proteins. Using targeted quantitation, low abundance proteins are detectable in tryptic digests of, for example, leaf homogenates without the requirement of first obtaining peroxisome isolates. Targeted absolute or relative quantitation is most commonly performed by employing selected reaction monitoring (SRM; Gerber et al., 2003; DeSouza et al., 2008; Zhi et al., 2011), a technique that originated from the targeted analysis of small molecules and has also become available for peptide analysis. In SRM, only a few peptides specific for the target protein are considered in the MS analysis, resulting in short duty cycles which save analytical time when compared to data-dependent analyses approaches. Thus the elution peaks of the target peptides can be monitored more closely, resulting in higher accuracy quantitation, especially for low abundance peptides. Modern mass spectrometers and knowledge of the retention times of the selected peptides allow the quantitation of up to several hundred transitions in one event and for this reason the technique is often also referred to as multiple reaction monitoring (MRM). Unfortunately, establishing transitions for MRM is a labor-intensive process. However, once this has been

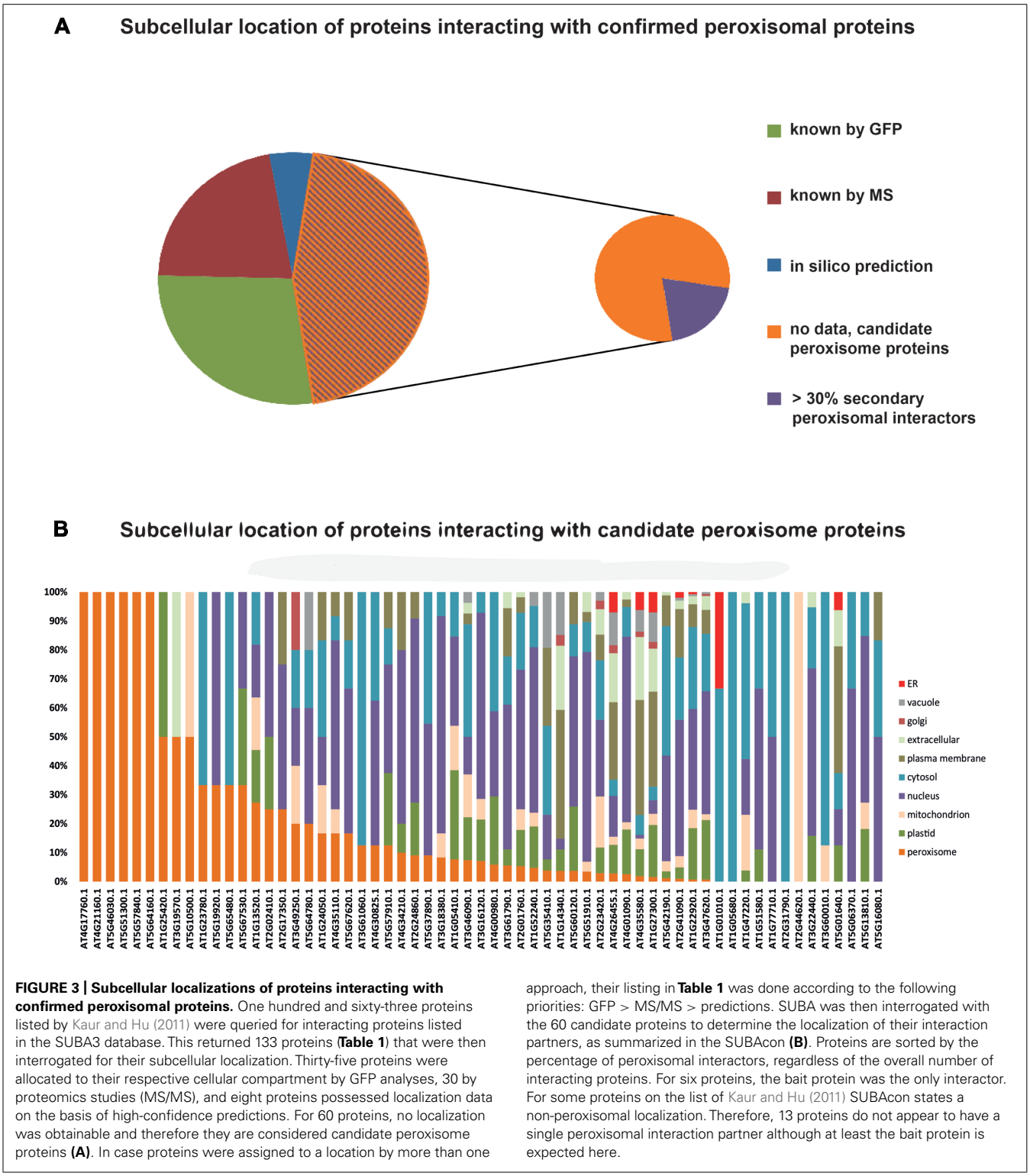


FIGURE 3 | Subcellular localizations of proteins interacting with confirmed peroxisomal proteins. One hundred and sixty-three proteins listed by Kaur and Hu (2011) were queried for interacting proteins listed in the SUBA3 database. This returned 133 proteins (**Table 1**) that were then interrogated for their subcellular localization. Thirty-five proteins were allocated to their respective cellular compartment by GFP analyses, 30 by proteomics studies (MS/MS), and eight proteins possessed localization data on the basis of high-confidence predictions. For 60 proteins, no localization was obtainable and therefore they are considered candidate peroxisome proteins **(A)**. In case proteins were assigned to a location by more than one approach, their listing in **Table 1** was done according to the following priorities: GFP > MS/MS > predictions. SUBA was then interrogated with the 60 candidate proteins to determine the localization of their interaction partners, as summarized in the SUBAcon **(B)**. Proteins are sorted by the percentage of peroxisomal interactors, regardless of the overall number of interacting proteins. For six proteins, the bait protein was the only interactor. For some proteins on the list of Kaur and Hu (2011) SUBAcon states a non-peroxisomal localization. Therefore, 13 proteins do not appear to have a single peroxisomal interaction partner although at least the bait protein is expected here.

achieved, samples can be analyzed in a high-throughput manner allowing the rapid quantitation of proteins in many complex mixtures. The peroxisomal proteome lends itself well to this kind of analysis because protein diversity is rather low compared to other major organelles of the plant cell. This allows relatively good coverage of the peroxisomal proteome with just a single or very few MRM runs. Especially for peroxisomes, establishing good MRM transitions is a worthwhile target, owing to the trials and tribulations associated with their isolation from plant material.

Table 1 | SUBA3 localization data of proteins interacting with confirmed peroxisomal proteins according to Kaur and Hu (2011).

Interacting proteins without experimental localization data, without predicted peroxisomal localization, and >30% peroxisomal interacting proteins
AT1G23780.1, AT1G25420.1, AT4G17760.1, AT4G21160.1, AT5G10500.1, AT5G19920.1, AT5G46030.1, AT5G51300.1, AT5G57840.1, AT5G64160.1, AT5G65480.1, AT5G67530.1
Interacting proteins without experimental localization data, without predicted peroxisomal localization, and <30% peroxisomal interacting proteins
AT1G01010.1, AT1G05410.1, AT1G05680.1, AT1G13520.1, AT1G14340.1, AT1G22920.1, AT1G24050.1, AT1G27300.1, AT1G47220.1, AT1G51580.1, AT1G52240.1, AT1G77710.1, AT2G01760.1, AT2G02410.1, AT2G17350.1, AT2G23420.1, AT2G24860.1, AT2G31790.1, AT2G41090.1, AT2G44620.1, AT3G16120.1, AT3G18380.1, AT3G22440.1, AT3G46090.1, AT3G47620.1, AT3G49250.1, AT3G60010.1, AT3G61060.1, AT3G61790.1, AT4G00980.1, AT4G01090.1, AT4G26455.1, AT4G30825.1, AT4G34210.1, AT4G35110.1, AT4G35580.1, AT5G01640.1, AT5G06370.1, AT5G13810.1, AT5G16080.1, AT5G35410.1, AT5G37890.1, AT5G42190.1, AT5G51910.1, AT5G57910.1, AT5G60120.1, AT5G64780.1, AT5G67620.1
Interacting proteins without experimental localization data and with predicted peroxisomal localization
AT2G01950.1, AT2G33520.1, AT3G03490.1, AT3G58740.1, AT4G14440.1, AT5G56220.1, AT5G65683.1
Interacting proteins with MS/MS-based localization data
AT1G03130.1, AT1G06290.1, AT1G12920.1, AT1G20950.1, AT1G43560.1, AT1G68010.1, AT2G13360.1, AT2G26230.1, AT2G35500.1, AT2G42790.1, AT3G12110.1, AT3G14415.1, AT3G14420.1, AT3G21865.1, AT3G26900.1, AT3G58750.1, AT3G60600.1, AT4G02770.1, AT4G22240.1, AT4G28440.1, AT4G35250.1, AT5G11450.1, AT5G25760.1, AT5G38420.1, AT5G46570.1, AT5G55190.1, AT5G56630.1, AT5G65940.1, AT5G66510.1
Interacting proteins with GFP-based localization data
AT1G02140.1, AT1G12520.1, AT1G13030.1, AT1G14830.1, AT1G48320.1, AT1G75950.1, AT1G76150.1, AT1G78300.1, AT2G14120.1, AT2G26350.1, AT2G26800.1, AT2G42490.1, AT3G01910.1, AT3G02150.1, AT3G04460.1, AT3G06720.1, AT3G07560.1, AT3G16310.1, AT3G18780.1, AT3G19570.1, AT3G21720.1, AT3G50070.1, AT3G56900.1, AT4G02150.1, AT4G22220.1, AT4G26450.1, AT4G33650.1, AT5G22290.1, AT5G25440.1, AT5G27600.1, AT5G27620.1, AT5G42980.1, AT5G44560.1, AT5G48230.1, AT5G56290.1, AT5G58220.1, AT5G63610.1

Apart from the peroxisome-specific reasons to employ MRM for the quantitation of proteins, it might also prove advantageous for other organelles. Isolation of the target compartment of a cell is a process that usually takes several hours, during which time unforeseen changes to metabolites, membranes and also to proteins may occur. Additionally, because centrifugation is often used, organellar subpopulations characterized by extreme densities or sizes might constitute the final isolate, differing somewhat from the situation found *in vivo*. Again, due to their fragility, this might affect peroxisomes more severely than other, more stable compartments. Thus, the development of these newer methods to reduce the isolation time and organellar stress prior to analysis will only serve to enhance the value of and possibilities for dynamic proteome studies.

CONCLUSION

Peroxisomal proteomics in *Arabidopsis* is seriously hampered by limited access to isolated organelles. Due to the low number of peroxisomes in plant cells, most peroxisomal proteins do not rank among the high-abundant proteins. This prevents good coverage of the peroxisomal proteome in shotgun proteomics. Increasingly good predictions of peroxisomal location by different routes will become available in the future. However, experience gained in the past on false-positive allocation of proteins to this organelle has shown that predictions are only suitable for generating candidates for the peroxisomal proteome and that experimental validation of predicted results is still necessary and will most likely remain so. Most of the experimental evidence will be obtained by fluorescent reporter fusion protein assays, but other approaches also lend themselves for this purpose. Laboratories geared toward proteomic studies could use targeted quantitative analysis of candidate proteins from isolated peroxisomes and fractions purified to lesser degrees of homogeneity in order to show enrichment of candidate proteins in peroxisomes.

In order for peroxisomal proteomics to reach the standard of the studies performed on plastids or mitochondria, technical improvements are of paramount importance. Such improvements will be obtained either in the purity and yield of peroxisomal isolates, or by alternative analytical methods to increase proteome coverage. Moreover, investigation of the dynamic changes occurring in the plant peroxisomal proteome may require a completely different approach. Since MS-technology is developing rapidly and advances in the isolation of peroxisomes are comparatively slow, recent and future developments in MS in combination with alternative strategies for identifying members of the plant peroxisomal proteome will be key to a better understanding of the functions and dynamics of this important compartment of the plant cell.

ACKNOWLEDGMENTS

This work was supported by the Australian Research Council (grant number CE0561495). We acknowledge support by Deutsche Forschungsgemeinschaft and Open Access Publishing Fund of Leibniz Universität Hannover.

REFERENCES

Arabidopsis Genome Initiative. (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408, 796–815.

Baker, A., Graham, I. A., Holdsworth, M., Smith, S. M., and Theodoulou, F. L. (2006). Chewing the fat: beta-oxidation in signalling and development. *Trends Plant Sci.* 11, 124–132.

Baker, A., and Sparkes, I. A. (2005). Peroxisome protein import: some answers, more questions. *Curr. Opin. Plant Biol.* 8, 640–647.

Bauwe, H., Hagemann, M., and Fernie, A. R. (2010). Photorespiration: players, partners and origin. *Trends Plant Sci.* 15, 330–336.

Bernhardt, K., Wilkinson, S., Weber, A. P., and Linka, N. (2012). A peroxisomal carrier delivers NAD(+) and contributes to optimal fatty acid degradation during storage oil mobilization. *Plant J.* 69, 1–13.

Blum, T., Briesemeister, S., and Kohlbacher, O. (2009). MultiLoc2: integrating phylogeny and Gene Ontology terms improves subcellular protein localization prediction. *BMC Bioinformatics* 10:274. doi: 10.1186/1471-2105-10-274

Boatright, J., Negre, F., Chen, X., Kish, C. M., Wood, B., Peel, G., et al. (2004). Understanding in vivo benzenoid metabolism in petunia petal tissue. *Plant Physiol.* 135, 1993–2011.

Braun, H. P., and Eubel, H. (2012). "Organellar proteomics: close insights into the spatial breakdown and functional dynamics of plant primary metabolism" in *Genomics of Chloroplasts and Mitochondria*, Vol. 35, *Advances in Photosynthesis and Respiration*, eds R. Bock and V. Knoop (Dordrecht: Springer), 357–378.

Chowdhary, G., Kataya, A. R., Lingner, T., and Reumann, S. (2012). Non-canonical peroxisome targeting signals: identification of novel PTS1 tripeptides and characterization of enhancer elements by computational permutation analysis. *BMC Plant Biol.* 12:142. doi: 10.1186/1471-2229-12-142

Cooper, T. G., and Beevers, H. (1969). Beta oxidation in glyoxysomes from castor bean endosperm. *J. Biol. Chem.* 244, 3514–3520.

de Duve, C. (1969a). Evolution of the peroxisome. *Ann. N. Y. Acad. Sci.* 168, 369–381.

de Duve, C. (1969b). The peroxisome: a new cytoplasmic organelle. *Proc. R. Soc. Lond. B Biol. Sci.* 173, 71–83.

de Duve, C., and Baudhin, P. (1966). Peroxisomes (microbodies and related particles). *Physiol. Rev.* 46, 323–357.

DeSouza, L. V., Taylor, A. M., Li, W., Minkoff, M. S., Romaschin, A. D., Colgan, T. J., et al. (2008). Multiple reaction monitoring of mTRAQ-labeled peptides enables absolute quantification of endogenous levels of a potential cancer marker in cancerous and normal endometrial tissues. *J. Proteome Res.* 7, 3525–3534.

Eckert, J. H., and Johnsson, N. (2003). Pex10p links the ubiquitin conjugating enzyme Pex4p to the protein import machinery of the peroxisome. *J. Cell Sci.* 116, 3623–3634.

Emanuelsson, O., Nielsen, H., Brunak, S., and von Heijne, G. (2000). Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J. Mol. Biol.* 300, 1005–1016.

Emanuelsson, O., Nielsen, H., and von Heijne, G. (1999). ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Sci.* 8, 978–984.

Eubel, H., Meyer, E. H., Taylor, N. L., Bussel, J. D., O'Toole N., Heazlewood, J. L., et al. (2008). Novel proteins, putative membrane transporters, and an integrated metabolic network are revealed by quantitative proteomic analysis of *Arabidopsis* cell culture peroxisomes. *Plant Physiol.* 148, 1809–1829.

Fukao, Y., Hayashi, M., Nishimura, I. H., and Nishimura, M. (2003). Novel glyoxysomal protein kinase, GPK1, identified by proteomic analysis of glyoxysomes in etiolated cotyledons of *Arabidopsis thaliana*. *Plant Cell Physiol.* 44, 1002–1012.

Fukao, Y., Hayashi, M., and Nishimura, M. (2002). Proteomic analysis of leaf peroxisomal proteins in greening cotyledons of *Arabidopsis thaliana*. *Plant Cell Physiol.* 43, 689–696.

Gavin, A. C., Bosche, M., Krause, R., Grandi, P., Marzioch, M., Bauer, A., et al. (2002). Functional organization of the yeast proteome by systematic analysis of protein complexes. *Nature* 415, 141–147.

Gerber, S. A., Rush, J., Stemman O., Kirschner, M. W., and Gygi, S. P. (2003). Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc. Natl. Acad. Sci. U.S.A.* 100, 6940–6945.

Germain, V., Rylott, E. L., Larson, T. R., Sherson, S. M., Bechtold, N., Carde, J. P., et al. (2001). Requirement for 3-ketoacyl-CoA thiolase-2 in peroxisome development, fatty acid beta-oxidation and breakdown of triacylglycerol in lipid bodies of *Arabidopsis* seedlings. *Plant J.* 28, 1–12.

Gould, S. J., Krisans, S., Keller, G. A., and Subramani, S. (1990). Antibodies directed against the peroxisomal targeting signal of firefly luciferase recognize multiple mammalian peroxisomal proteins. *J. Cell Biol.* 110, 27–34.

Goyer, A., Johnson, T. L., Olsen, L. J., Collakova, E., Shachar-Hill, Y., Rhodes, D., et al. (2004). Characterization and metabolic function of a peroxisomal sarcosine and pipecolate oxidase from *Arabidopsis*. *J. Biol. Chem.* 279, 16947–16953.

Graham, I. A. (2008). Seed storage oil mobilization. *Annu. Rev. Plant Biol.* 59, 115–142.

Hayashi, M., Toriyama, K., Kondo, M., and Nishimura, M. (1998). 2,4-Dichlorophenoxybutyric acid-resistant mutants of *Arabidopsis* have defects in glyoxysomal fatty acid beta-oxidation. *Plant Cell* 10, 183–195.

Heineke, D., Bykova, N., Gardestrom, P., and Bauwe, H. (2001). Metabolic response of potato plants to an antisense reduction of the P-protein of glycine decarboxylase. *Planta* 212, 880–887.

Helm, M., Luck, C., Prestele, J., Hierl, G., Huesgen, P. F., Frohlich, T., et al. (2007). Dual specificities of the glyoxysomal/peroxisomal processing protease Deg15 in higher plants. *Proc. Natl. Acad. Sci. U.S.A.* 104, 11501–11506.

Hoepfner, D., Schildknegt, D., Braakman, I., Philippsen, P., and Tabak, H. F. (2005). Contribution of the endoplasmic reticulum to peroxisome formation. *Cell* 122, 85–95.

Hu, J., Baker, A., Bartel, B., Linka, N., Mullen, R. T., Reumann, S., et al. (2012). Plant peroxisomes: biogenesis and function. *Plant Cell* 24, 2279–2303.

Islinger, M., Eckerskorn, C., and Völkl, A. (2010). Free-flow electrophoresis in the proteomic era: a technique in flux. *Electrophoresis* 31, 1754–1763.

Kato, A., Takeda-Yoshikawa, Y., Hayashi, M., Kondo, M., Hara-Nishimura, I., and Nishimura, M. (1998). Glyoxysomal malate dehydrogenase in pumpkin: cloning of a cDNA and functional analysis of its presequence. *Plant Cell Physiol.* 39, 186–195.

Kaur, N., and Hu, J. (2011). Defining the plant peroxisomal proteome: from *Arabidopsis* to rice. *Front. Plant Sci.* 2:103. doi: 10.3389/fpls.2011.00103

Kebeish, R., Niessen, M., Thiruveedhi, K., Bari, R., Hirsch, H. J., Rosenkranz, R., et al. (2007). Chloroplastic photorespiratory bypass increases photosynthesis and biomass production in *Arabidopsis thaliana*. *Nat. Biotechnol.* 25, 593–599.

Kerrien, S., Aranda, B., Breuza, L., Bridge, A., Broackes-Carter, F., Chen, C., et al. (2012). The IntAct molecular database in 2012. *Nucleic Acids Res.* 40, D841–D846.

Klempien, A., Kaminaga, Y., Qualley, A., Nagegowda, D. A., Widhalm, J. R., Orlova, I., et al. (2012). Contribution of CoA ligases to benzenoid biosynthesis in petunia flowers. *Plant Cell* 24, 2015–2030.

Kliebenstein, D. J., D'Aauria, J. C., Behere, A. S., Kim, J. H., Gunderson, K. L., Breen, J. N., et al. (2007). Characterization of seed-specific benzoyloxyglucosinolate mutations in *Arabidopsis thaliana*. *Plant J.* 51, 1062–1076.

Krogan, N. J., Cagney, G., Yu, H., Zhong, G., Guo, X., Ignatchenko, A., et al. (2006). Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae*. *Nature* 440, 637–643.

Kunz, H. H., Scharnewski, M., Feussner, K., Feussner, I., Flugge, U. I., Fulda, M., et al. (2009). The ABC transporter PXA1 and peroxisomal beta-oxidation are vital for metabolism in mature leaves of *Arabidopsis* during extended darkness. *Plant Cell* 21, 2733–2749.

Lamberto, I., Percudani, R., Gatti, R., Folli, C., and Petrucco, S. (2010). Conserved alternative splicing of *Arabidopsis* transthyretin-like determines protein localization and S-allantoin synthesis in peroxisomes. *Plant Cell* 22, 1564–1574.

Lee, K. Y., Thorneycroft, D., Achuthan, P., Hermjakob, H., and Ideker, T. (2010). Mapping plant interactomes using literature curated and predicted protein–protein interaction data sets. *Plant Cell* 22, 997–1005.

Lee, S., Kaminaga, Y., Cooper, B., Pichersky, E., Dudareva, N., and Chapple, C. (2012). Benzoylation and sinapoylation of glucosinolate R-groups in *Arabidopsis*. *Plant J.* 72, 411–422.

Lingard, M. J., and Bartel, B. (2009). Arabidopsis LON2 is necessary for peroxisomal function and sustained matrix protein import. *Plant Physiol.* 151, 1354–1365.

Lingard, M. J., Gidda, S. K., Bingham, S., Rothstein, S. J., Mullen, R. T., and Trelease, R. N. (2008). *Arabidopsis* PEROXIN11c–e, FISSION1b, and DYNAMIN-RELATED PROTEIN3A cooperate in cell cycle-associated replication of peroxisomes. *Plant Cell* 20, 1567–1585.

Lingard, M. J., Monroe-Augustus, M., and Bartel, B. (2009). Peroxisome-associated matrix protein

degradation in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* 106, 4561–4566.

Lingard, M. J., and Trelease, R. N. (2006). Five *Arabidopsis* peroxin 11 homologs individually promote peroxisome elongation, duplication or aggregation. *J. Cell Sci.* 119, 1961–1972.

Lingner, T., Kataya, A. R., Antonicelli, G. E., Benichou, A., Nilssen, K., Chen, X.-Y., et al. (2011). Identification of novel plant peroxisomal targeting signals by a combination of machine learning methods and in vivo subcellular targeting analysis. *Plant Cell* 23, 1556–1572.

Lisenbee, C. S., Lingard, M. J., and Trelease, R. N. (2005). *Arabidopsis* peroxisomes possess functionally redundant membrane and matrix isoforms of monodehydroascorbate reductase. *Plant J.* 43, 900–914.

Ma, C., and Reumann, S. (2008). Improved prediction of peroxisomal PTS1 proteins from genome sequences based on experimental subcellular targeting analyses as exemplified for protein kinases from *Arabidopsis*. *J. Exp. Bot.* 59, 3767–3779.

Makkar, R. S., Contreras, M. A., Paintlia, A. S., Smith, B. T., Haq, E., and Singh, I. (2006). Molecular organization of peroxisomal enzymes: protein–protein interactions in the membrane and in the matrix. *Arch. Biochem. Biophys.* 451, 128–140.

Mullen, R. T., and Trelease, R. N. (2006). The ER–peroxisome connection in plants: development of the "ER semi-autonomous peroxisome maturation and replication" model for plant peroxisome biogenesis. *Biochim. Biophys. Acta* 1763, 1655–1668.

Oshima, Y., Kamigaki, A., Nakamori, C., Mano, S., Hayashi, M., Nishimura, M., et al. (2008). Plant catalase is imported into peroxisomes by Pex5p but is distinct from typical PTS1 import. *Plant Cell Physiol.* 49, 671–677.

Palma, J. S., Corpas, F. J., and del Rio, L. A. (2009). Proteome of plant peroxisomes: new perspectives on the role of these cell organelles in cell biology. *Proteomics* 9, 2301–2312.

Peterhansel, C., and Maurino, V. G. (2011). Photorespiration redesigned. *Plant Physiol.* 155, 49–55.

Pracharoenwattana, I., and Smith, S. M. (2008). When is a peroxisome not a peroxisome? *Trends Plant Sci.* 13, 522–525.

Qualley, A. V., Widhalm, J. R., Adebesin, F., Kish, C. M., and Dudareva, N. (2012). Completion of the core beta-oxidative pathway of benzoic acid biosynthesis in plants. *Proc. Natl. Acad. Sci. U.S.A.* 109, 16383–16388.

Reumann, S. (2004). Specification of the peroxisome targeting signals type 1 and type 2 of plant peroxisomes by bioinformatics analyses. *Plant Physiol.* 135, 783–800.

Reumann, S. (2011). Toward a definition of the complete proteome of plant peroxisomes: where experimental proteomics must be completed by bioinformatics. *Proteomics* 11, 1764–1779.

Reumann, S., Babujee, L., Ma, C., Wienkoop, S., Siemens, T., Antonicelli, G. E., et al. (2007). Proteome analysis of *Arabidopsis* leaf peroxisomes reveals novel targeting peptides, metabolic pathways, and defense mechanisms. *Plant Cell* 19, 3170–3193.

Reumann, S., Buchwald, D., and Lingner, T. (2012). PredPlantPTS1: a web server for the prediction of plant peroxisomal proteins. *Front. Plant Sci.* 3:194. doi: 10.3389/fpls.2012.00194

Reumann, S., Ma, C., Lemke, S., and Babujee, L. (2004). AraPerox. A database of putative *Arabidopsis* proteins from plant peroxisomes. *Plant Physiol.* 136, 2587–2608.

Reumann, S., Quan, S., Aung, K., Yang, P., Manandhar-Shrestha, K., Holbrook, D., et al. (2009). In-depth proteome analysis of *Arabidopsis* leaf peroxisomes combined with in vivo subcellular targeting verification indicates novel metabolic and regulatory functions of peroxisomes. *Plant Physiol.* 150, 125–143.

Rigaut, G., Shevchenko, A., Rutz, B., Wilm, M., Mann, M., and Seraphin, B. (1999). A generic protein purification method for protein complex characterization and proteome exploration. *Nat. Biotechnol.* 17, 1030–1032.

Sapir-Mir, M., Mett, A., Belausov, E., Tal-Meshulam, S., Frydman, A., Gidoni, D., et al. (2008). Peroxisomal localization of *Arabidopsis* isopentenyl diphosphate isomerases suggests that part of the plant isoprenoid mevalonic acid pathway is compartmentalized to peroxisomes. *Plant Physiol.* 148, 1219–1228.

Scherer, G. F., and Morre, D. J. (1978). Action and inhibition of endogenous phospholipases during isolation of plant membranes. *Plant Physiol.* 62, 933–937.

Schuhmann, H., Huesgen, P., Gietl, C., and Adamska, I. (2008). The DEG15 serine protease cleaves peroxisomal targeting signal 2-containing proteins in *Arabidopsis*. *Plant Physiol.* 148, 1847–1856.

Simkin, A. J., Guirimand, G., Papon, N., Courdavault, V., Thabet, I., Ginis, O., et al. (2011). Peroxisomal localisation of the final steps of the mevalonic acid pathway in planta. *Planta* 234, 903–914.

Strader, L. C., Wheeler, D. L., Christensen, S. E., Berens, J. C., Cohen, J. D., Rampey, R. A., et al. (2011). Multiple facets of *Arabidopsis* seedling development require indole-3-butyric acid-derived auxin. *Plant Cell* 23, 984–999.

Tanz, S. K., Castleden, I., Hooper, C. M., Vacher, M., Small, I., and Millar, A. H. (2013). SUBA3: a database for integrating experimentation and prediction to define the SUBcellular location of proteins in *Arabidopsis*. *Nucleic Acids Res.* 41, D1185–D1191.

Timm, S., Florian, A., Arrivault, S., Stitt, M., Fernie, A., and Bauwe, H. (2012). Glycine decarboxylase controls photosynthesis and plant growth. *FEBS Lett.* 586, 3692–3696.

Tolbert, N. E., Oeser, A., Kisaki, T., Hageman, R. H., and Yamazaki, R. K. (1968). Peroxisomes from spinach leaves containing enzymes related to glycolate metabolism. *J. Biol. Chem.* 243, 5179–5184.

Tolbert, N. E., Oeser, A., Yamazaki, R. K., Hageman, R. H., and Kisaki, T. (1969). A survey of plants for leaf peroxisomes. *Plant Physiol.* 44, 135–147.

van der Zand, A., Gent, J., Braakman, I., and Tabak, H. F. (2012). Biochemically distinct vesicles from the endoplasmic reticulum fuse to form peroxisomes. *Cell* 149, 397–409.

Van Leene, J., Boruc, J., De Jaeger, G., Russinova, E., and De Veylder, L. (2011). A kaleidoscopic view of the *Arabidopsis* core cell cycle interactome. *Trends Plant Sci.* 16, 141–150.

Van Leene, J., Witters, E., Inze, D., and De Jaeger, G. (2008). Boosting tandem affinity purification of plant protein complexes. *Trends Plant Sci.* 13, 517–520.

Wiszniewski, A. A., Zhou, W., Smith, S. M., and Bussell, J. D. (2009). Identification of two *Arabidopsis* genes encoding a peroxisomal oxidoreductase-like protein and an acyl-CoA synthetase-like protein that are required for responses to proauxins. *Plant Mol. Biol.* 69, 503–515.

Yates, J. R. III. (2000). Mass spectrometry. From genomics to proteomics. *Trends Genet.* 16, 5–8.

Zhi, W., Wang, M., and She, J.-X. (2011). Selected reaction monitoring (SRM) mass spectrometry without isotope labeling can be used for rapid protein quantification. *Rapid Commun. Mass Spectrom.* 25, 1583–1588.

Zhu, J. K. (2002). Salt and drought stress signal transduction in plants. *Annu. Rev. Plant Biol.* 53, 247–273.

Zolman, B. K., Yoder, A., and Bartel, B. (2000). Genetic analysis of indole-3-butyric acid responses in *Arabidopsis thaliana* reveals four mutant classes. *Genetics* 156, 1323–1337.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 05 December 2012; accepted: 02 April 2013; published online: 24 April 2013.

Citation: Bussell JD, Behrens C, Ecke W and Eubel H (2013) Arabidopsis peroxisome proteomics. Front. Plant Sci. ***4****:101. doi: 10.3389/fpls.2013.00101*

This article was submitted to Frontiers in Plant Proteomics, a specialty of Frontiers in Plant Science.

Copyright © 2013 Bussell, Behrens, Ecke and Eubel. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.

frontiers in
PLANT SCIENCE

REVIEW ARTICLE
published: 11 April 2013
doi: 10.3389/fpls.2013.00086

Advancements in the analysis of the *Arabidopsis* plasma membrane proteome

Koste A. Yadeta†, J. Mitch Elmore† and Gitta Coaker*

Department of Plant Pathology, University of California Davis, Davis, CA, USA

Edited by:
Nicolas L. Taylor, The University of Western Australia, Australia

Reviewed by:
Scott C. Peck, University of Missouri, USA
Myriam Ferro, Commisariat à l'Energie Atomique et aux Energies Alternatives, France

****Correspondence:***
Gitta Coaker, Department of Plant Pathology, University of California Davis, One Shields Avenue, 576 Hutchison Hall, Davis, CA 95616, USA.
e-mail: glcoaker@ucdavis.edu

† Koste A. Yadeta and J. Mitch Elmore have contributed equally to this work.

The plasma membrane (PM) regulates diverse processes essential to plant growth, development, and survival in an ever-changing environment. In addition to maintaining normal cellular homeostasis and plant nutrient status, PM proteins perceive and respond to a myriad of environmental cues. Here we review recent advances in the analysis of the plant PM proteome with a focus on the model plant *Arabidopsis thaliana*. Due to membrane heterogeneity, hydrophobicity, and low relative abundance, analysis of the PM proteome has been a special challenge. Various experimental techniques to enrich PM proteins and different protein and peptide separation strategies have facilitated the identification of thousands of integral and membrane-associated proteins. Numerous classes of proteins are present at the PM with diverse biological functions. PM microdomains have attracted much attention. However, it still remains a challenge to characterize these cell membrane compartments. Dynamic changes in the PM proteome in response to different biotic and abiotic stimuli are highlighted. Future prospects for PM proteomics research are also discussed.

Keywords: plasma membrane, proteomics, mass spectrometry, membrane proteins, *Arabidopsis*

INTRODUCTION

The plasma membrane (PM) is the cellular interface that regulates the exchange of molecules and information between cells and their environment. The PM is involved in a range of plant physiological processes including growth and development, ion and metabolite transport, perception of environmental changes, and disease resistance (Marmagne et al., 2004; Mongrand et al., 2010). At the cellular level, PM proteins maintain the electrochemical gradients required for membrane transport and play a critical role in osmoregulation of the cell (Schulz, 2011). In addition, the PM plays an essential role in sensing and responding to biotic and abiotic stresses. PM transporters control the distribution and movement of plant hormones and thus mediate short- and long-distance signaling processes (Kerr et al., 2011). Various plant hormone receptors for auxin, brassinosteroids (BR), and abscisic acid are also localized to the PM (Wang et al., 2001; Pandey et al., 2009; Robert et al., 2010). In addition, many plant innate immune receptors and defense response regulators are integrally or peripherally associated with the PM (Dodds and Rathjen, 2010; Monaghan and Zipfel, 2012), highlighting the importance of the PM in regulating numerous aspects of plant growth, development, and adaptation to changing environments.

The PM is composed of a lipid bilayer and associated proteins. Plant cell membrane lipids consist primarily of glycerophospholipids, sphingolipids, and sterols (Furt et al., 2011; Cacas et al., 2012). Membrane proteins can be directly embedded within the lipid bilayer or undergo lipid modification which impacts their localization and membrane association. For many years, the biological membrane was considered as a dynamic two-dimensional fluid composed of homogenously distributed lipids and proteins (Singer and Nicolson, 1972). However, now it is clear that distinct membrane microdomains of various sizes and mobilities exist in eukaryotic cells (Kusumi et al., 2012). In plants, PM microdomains have been implicated in various cellular processes including cell wall attachment, protein sorting and trafficking, signal transduction, and plant–microbe interactions (Mongrand et al., 2010; Simon-Plas et al., 2011; Urbanus and Ott, 2012). In addition to membrane compartmentalization, post-translational protein modifications (PTMs) of PM proteins impacts their activity and signaling capabilities. Currently, over 300 different PTMs have been reported, with protein phosphorylation being the most intensively studied (Zhao and Jensen, 2009; Kline-Jonakin et al., 2011). PTMs can modulate protein activity through changes in protein conformation, localization, stability, and protein–protein interactions. Global surveys and functional analysis of protein PTMs during signaling events are now possible through advancements in proteomic approaches (Zhao and Jensen, 2009).

In eukaryotes, roughly 30% of the genome encodes membrane proteins (Tan et al., 2008). However, in many studies PM proteins are often underrepresented due to their physiochemical heterogeneity, hydrophobicity, and low abundance (Marmagne et al., 2007). Several technological advances have been developed that overcome some of the challenges afflicting PM proteomics analyses. Various label-based and label-free methods exist for the quantification of peptides, proteins, and PTMs in plant tissue extracts (as reviewed in Schulze and Usadel, 2010; Kline-Jonakin et al., 2011; Kota and Goshe, 2011). Here we briefly summarize recent advancements in PM proteomics with a focus on the model plant *Arabidopsis thaliana*.

EVOLUTION OF PLASMA MEMBRANE PROTEOMIC STRATEGIES: ENRICH, SOLUBILIZE, AND ANALYZE

Due to the unique roles of the PM in cellular function, identification, and functional characterization of the plant PM proteome has been critical for understanding of how plants grow, develop, and respond to their environment. The low relative abundance of PM proteins in whole tissue extracts has necessitated the development of various strategies to enrich for proteins specific to the cell membrane before proteomic analysis. Even after isolation of PM fractions, due to the complexity of protein species and the large dynamic range of protein abundance, it is necessary to employ various protein and/or peptide separation techniques to achieve a comprehensive survey of the PM proteome (**Figure 1**).

Various techniques have been used for PM separation and enrichment from total microsomal fractions including: density gradient centrifugation, free-flow electrophoresis, and phase polymer systems (Larsson et al., 1987; Dunkley et al., 2006; Santoni, 2007). Aqueous two-phase partitioning is one of the most common and effective techniques for enrichment of PM vesicles to high purity (Larsson et al., 1987). Above a critical concentration, the polymers polyethylene glycol (PEG) and dextran will form distinct phases when mixed in aqueous solution. Biomembranes associate with one phase or the other based on the membrane surface charge (Larsson et al., 1987). PM vesicles preferentially partition in the PEG-enriched upper phase while other cellular membranes associate with the dextran-enriched lower phase. Repeated cycles of phase partitioning can result in highly enriched PM fractions estimated to be over 90% pure based on enzyme marker assays (Larsson et al., 1987; Palmgren et al., 1990). Recently, a relatively simple approach for enrichment of PM proteins has been described (Zhang and Peck, 2011). Although the resulting PM fractions are not as pure as those derived from two-phase partitioning, this method is rapid and requires less sample handling, making it advantageous in situations where many samples must be processed at the same time (Zhang and Peck, 2011).

Because subcellular compartments isolated using biochemical approaches are never 100% pure, it is of interest to reduce and then evaluate contaminating organelles/proteins when isolating PMs. Testing PM enrichment and purity relative to total microsomal fractions usually involves enzymatic assays of the PM H+-ATPase or various immunological markers (Larsson et al., 1987; Marmagne et al., 2007). Because PM vesicles isolated by two-phase partitioning are predominantly apoplastic-side out, the non-ionic detergent Brij-58 is commonly used to invert PM vesicles inside-out and release organelles and/or cytosolic proteins that are trapped within the vesicles during tissue disruption (Palmgren

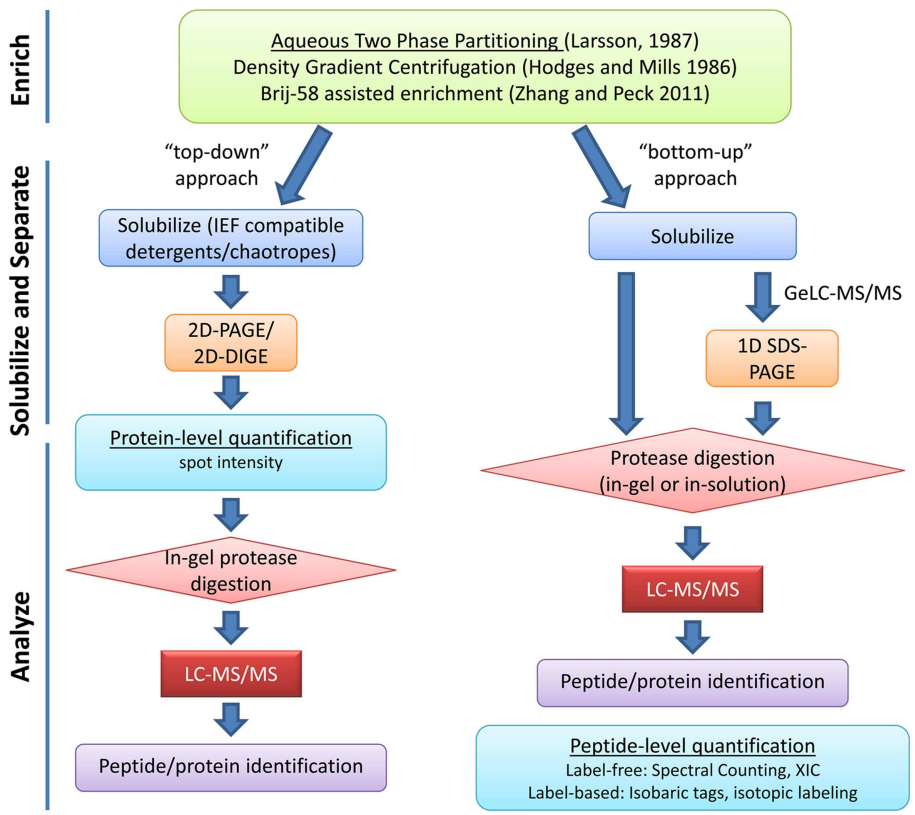


FIGURE 1 | Overview of typical plasma membrane proteomics experiments. Enrichment of PM proteins is usually achieved through aqueous two-phase partitioning. Strategies for solubilization and separation of PM proteins are chosen based on downstream applications. "Top-down" approaches use protein-level quantification while "bottom-up" approaches use peptide-level quantification. Various strategies for label-based and label-free peptide quantification exist and will influence experimental design (reviewed in Schulze and Usadel, 2010; Kota and Goshe, 2011). XIC, extracted ion chromatogram.

et al., 1990; Johansson et al., 1995; Zhang and Peck, 2011). If the goal is to analyze primarily integral membrane proteins (IMPs), high pH and/or high salt treatments of inverted vesicles can be used to remove peripheral and loosely associated cytosolic proteins from the membrane (Santoni et al., 1999; Santoni, 2007; Marmagne et al., 2007). These treatments have facilitated the identification of a larger number of hydrophobic proteins after PM enrichment (Marmagne et al., 2007).

In order to evaluate contaminating proteins in PM preparations, quantitative isotopic labeling has been used for fractions enriched by density gradient centrifugation (Dunkley et al., 2006) and aqueous two-phase partitioning (Nelson et al., 2006). By comparing the degree of enrichment of known PM protein markers (e.g., PM H^+-ATPase) in a given fraction relative to other subcellular markers, it was possible to confidently assign PM localization to a number of previously unknown PM proteins in *Arabidopsis* (Dunkley et al., 2006; Nelson et al., 2006). Using this approach, it was estimated that over 25% of proteins identified in PM-enriched fractions could be considered biological contaminants (Nelson et al., 2006). However, often in functional proteomics investigations, the goal is not to achieve absolutely pure PM fractions, but to enrich PM proteins in order to study the behavior of the PM during a physiological process or under stress conditions (Zhang and Peck, 2011). Thus it is less important to unequivocally assign a subcellular location to a protein, than to reproducibly identify and quantify its behavior under a specific condition (Elmore et al., 2012). Nevertheless, several excellent resources are available for the analysis and validation of proteins identified from a PM proteomics experiment (**Table 1**).

After PM isolation, proteomics analyses typically involve both gel-based and gel-free methods for separation of proteins or peptides prior to identification by mass spectrometry (MS). Early efforts in *Arabidopsis* utilized two-dimensional gel electrophoresis (2DGE) but later it became clear that 2DGE was not an ideal technique for separation of membrane proteins (Santoni et al., 1998, 2000; Prime et al., 2000). Most hydrophobic proteins have limited solubility in buffers required for the first dimension isoelectric focusing (IEF) step of 2DGE (Wilkins et al., 1998; Santoni et al., 2000). The low abundance, hydrophobicity, generally large molecular weight, and generally alkaline nature of

Table 1 | Web-based resources for protein analyses and validation.

Database	Description	Website	Reference
SUBCELLULAR LOCATION/TRANSMEMBRANE (TM) PREDICTION			
ARAMEMNON	Consensus prediction of TM domains, lipid modification, signal peptides, and subcellular location; contains link to all prediction programs	http://aramemnon.botan ik.uni-koeln.de/	Schwacke et al. (2003)
SUBA3	Combines subcellular prediction programs with experimental data (GFP localization, MS/MS, etc)	http://suba.plantenergy.uwa.edu.au/	Heazlewood et al. (2007)
ExPASy	Collection of multiple tools for the prediction of post-translational modifications and protein localization	http://www.expasy.org/p roteomics/post-translational_modification	Artimo et al. (2012)
PROTEIN PHOSPHORYLATION			
PhosPhAt	Aggregation of phosphorylation sites identified by mass spectrometry collected from over 20 studies in *Arabidopsis*. Can visualize raw MS/MS spectrum data	http://phosphat.mpimp-golm.mpg.de/	Durek et al. (2010)
P3DB	Plant protein phosphorylation site database. Can visualize raw MS/MS spectrum data	http://www.p3db.org/	Yao et al. (2012)
PlantsP	Functional genomics database focusing on protein kinases and phosphatases	http://plantsp.genomics.purdue.edu/ index.html	Tchieu et al. (2003)
DATA AGGREGATION/GENERAL PROTEOMICS			
MaSCP gator	Constantly updated data aggregation portal that retrieves proteomics information from several actively curated databases	http://gator.masc-proteomics.org/	Joshi et al. (2011)
pep2pro	Searchable mass spectrum library of experimentally identified peptides in MS/MS studies	http://fgcz-pep2pro.uzh.ch/	Hirsch-Hoffmann et al. (2012)
MetNet	Systems biology tool for combined analysis of protein, gene expression, and metabolite profiling data	http://www.metnetonline.org/	Sucaet et al. (2012)
PROTEIN–PROTEIN INTERACTIONS			
MIND	Membrane protein–protein interaction dataset using yeast split-ubiquitin system	http://www.associomics.org/ Associomics/MIND.html	Lalonde et al. (2010)
ANAP	Aggregates gene and protein interaction data from diverse sources. Web portal for cytoscape visualization	http://gmdd.shgmo.org/Computational-Biology/ANAP/ANAP_V1.1/	Wang et al. (2012)

PM proteins have all led to the poor performance of 2DGE in PM proteomics (Santoni et al., 2000; Gilmore and Washburn, 2010). Nevertheless, various chaotropes and detergents have been used with improvements in solubility and resolution of some membrane proteins in 2DGE and this "top-down" approach has been used with success to study hormone signaling at the PM (Santoni et al., 1999; Luche et al., 2003; Tang et al., 2008a) (**Figure 1**).

For most researchers, liquid chromatography-tandem MS (LC-MS/MS) "bottom-up" shotgun proteomics has emerged as the method of choice for large-scale identification and quantification of proteins, especially membrane proteins (**Figure 1**). PM protein samples are first solubilized, digested with a protease to cleave polypeptide chains into shorter peptide fragments, and then these fragments are separated by LC prior to ionization and MS/MS analysis. Various PM protein solubilization strategies prior to in-solution or in-gel digestion have been used to increase coverage of PM proteins in LC-MS/MS analysis (Marmagne et al., 2004; Mitra et al., 2007).

After digestion, peptides can be separated in one or more dimensions, typically involving reverse phase (RP) and/or strong cation exchange (SCX) chromatography for increased resolution and improved detection of low abundance peptides (Washburn et al., 2001; Fournier et al., 2007; Gilmore and Washburn, 2010). The use of different proteases with varying cleavage specificities has also increased the representation of membrane proteins in MS/MS datasets (Wu et al., 2003; Fischer and Poetsch, 2006). Recent reviews focus on advances in protein and peptide separation strategies for MS-based membrane proteomics (Fournier et al., 2007; Komatsu, 2008; Gilmore and Washburn, 2010).

One method gaining popularity is Gel-enhanced LC-MS/MS (GeLC-MS/MS), where extracted proteins are first subjected to one dimensional SDS-PAGE to separate by size and then regions of the gel lane are excised, digested, and subjected to LC-MS/MS separately (Alexandersson et al., 2004; Marmagne et al., 2007; Gilmore and Washburn, 2010). GeLC-MS/MS has been shown to outperform other separation techniques in terms of reproducibility and total number of protein identifications (Fang et al., 2010; Piersma et al., 2010). Another advantage of the GeLC-MS/MS approach is that PM fractions can be efficiently solubilized in strong detergents and/or chaotropes prior to SDS-PAGE, then digested in-gel to yield peptides suitable for MS/MS analysis.

MAJOR CLASSES OF PROTEINS IN THE *ARABIDOPSIS* PM AND THEIR BIOLOGICAL FUNCTIONS

The PM consists of structurally and functionally diverse proteins. The composition of the PM proteome varies with plant cell-type, developmental stage, and environmental conditions (Alexandersson et al., 2004). PM proteins can be classified into three main categories depending on the type of membrane association: IMPs, peripheral membrane proteins (PMPs), and glycosylphosphatidylinositol (GPI)-anchored membrane proteins. Many resources exist for the prediction of PM localization, transmembrane (TM) domains, lipid-based modifications, and GPI-anchors in proteins identified from PM fractions (Schwacke et al., 2003; Heazlewood et al., 2007) (**Table 1**).

INTEGRAL MEMBRANE PROTEINS

Integral membrane proteins are composed of one or more hydrophobic TM domains that span the lipid bilayer of the membrane. The majority of IMPs span the lipid bilayer with an α-helical structure, although some IMP domains exhibit β-barrel structure (Marmagne et al., 2007; Tan et al., 2008). Many IMPs contain a N-terminal signal peptide for secretion and membrane targeting through the ER and Golgi. Most active and passive membrane transport processes are controlled by a variety of IMP pumps, channels, and carriers (Schulz, 2011). One of the most abundant proteins in the plant PM, the PMH^+-ATPase, is the primary proton pump responsible for the establishment of the electrochemical gradient across the membrane that drives secondary transport processes. Other highly abundant PM proteins include the PM intrinsic protein (PIP) or aquaporin family which function mainly as water channels but can transport other small molecules (Schulz, 2011). Various other ion, hormone, and nutrient transporters exist at the PM as IMPs. ATP-binding cassette (ABC) transporters use ATP hydrolysis to drive the efflux or influx of a variety of substances including auxin, ABA, heavy metals, and antimicrobial compounds (Schulz, 2011) (**Figure 3**).

In addition to membrane transport activities, other IMPs are involved in the perception of extracellular signals and activation of downstream responses. One of the largest classes of signaling proteins in plants is the Receptor-like kinase (RLK) family, whose members can be relatively abundant on the PM (Santoni et al., 2003; Alexandersson et al., 2004; Marmagne et al., 2004). Characterized *Arabidopsis* RLKs function in a variety of processes including cell division and differentiation, hormone perception, meristem maintenance, pathogen recognition, and cell death control (De Smet et al., 2009; Monaghan and Zipfel, 2012).

PERIPHERAL MEMBRANE PROTEINS

Peripheral membrane proteins lack a membrane spanning domain but are membrane-associated either through covalent lipid modifications or non-covalent protein–protein interactions (Marmagne et al., 2007; Tan et al., 2008). Lipid modifications such as N-myristoylation, S-palmitoylation, or prenylation are common in PMPs and these modifications can control protein localization, sorting, and function (Testerink and Munnik, 2011). Proteins involved in vesicular membrane trafficking such as Rho of plants (ROPs) and Soluble N-ethylmaleimide sensitive factor attachment protein receptors (SNAREs) are commonly targeted to the PM via lipid modification (Sanderfoot et al., 2000; Testerink and Munnik, 2011). In addition to these lipid PTMs, many proteins associate with the PM via protein–protein interactions. These types of PMPs are often involved in signaling events by relaying messages from the PM to the rest of the cell.

GPI-ANCHORED MEMBRANE PROTEINS

Glycosylphosphatidylinositol-anchored membrane proteins are post-translationally modified to carry a C-terminal GPI-anchor that mediates their association with the membrane. The GPI-anchor is synthesized in endoplasmic reticulum and subsequently attached to a protein, which is transported to PM via the Golgi (Elortza et al., 2003; Fujita and Kinoshita, 2012). Unlike most PMPs, which localize to the cytosolic side of the PM, GPI-anchored

membrane proteins are mostly found attached to the outer surface of the PM. Many enzymes associated with cell wall processes (e.g., β-1,3-glucanases, pectinesterases, and polygalacturonases) are among the GPI-anchored proteins identified in *Arabidopsis* PM proteome (Borner et al., 2003; Elortza et al., 2003). Accordingly, GPI-anchored proteins are implicated in biological processes such as directional cell expansion, cellulose deposition, cell wall attachment and remodeling, and plant immunity (Elortza et al., 2006; Fujita and Kinoshita, 2012). GPI-anchored proteins are often found enriched in membrane microdomain preparations (discussed below), suggesting that cell wall maintenance hubs are compartmentalized within the membrane (Kierszniowska et al., 2009).

THE DYNAMIC PLASMA MEMBRANE

Proteins are constantly associating and disassociating from the membrane during endocytic, secretion, and signaling events. Enzyme activity, signal transduction, and transport regulation are all influenced by post-translational modifications that can affect the function and/or localization of proteins at or within the PM without changing their overall abundance. Even within a cell, polarized distribution of proteins involved in auxin transport has been readily observed. Dynamic focal accumulation of PM proteins involved in the plant immune response has been documented at sites of pathogen infection (Frey and Robatzek, 2009). Proteomic analysis of the PM during diverse signaling events has led to a greater appreciation of plant PM complexity and plasticity (Simon-Plas et al., 2011; Urbanus and Ott, 2012). Label-based and label-free approaches can be employed for protein quantification at the level of proteins or peptides (**Figure 1**). The various PM proteome quantification strategies have been recently reviewed (Komatsu, 2008; Schulze and Usadel, 2010; Kline-Jonakin et al., 2011; Kota and Goshe, 2011).

Advances in quantitative fluorescence microscopy have also improved our understanding of PM protein movement within the membrane. Single molecule analysis of the PIP2;1 aquaporin has revealed disparate localizations and lateral mobilities in both non-stressed and salt-stressed cells indicating that this water channel is under complex regulation even under normal conditions (Li et al., 2011). Another recent study observed a range of different diffusion rates for a representative set of *Arabidopsis* PM proteins (Martinière et al., 2012). Furthermore, in contrast with other eukaryotes, the cytoskeleton and lipid microdomains had little effect on the mobilities of the proteins studied (Kusumi et al., 2012; Martinière et al., 2012). Interestingly, the plant cell wall was found to restrict the movement of proteins with domains projecting to the outer surface of the cell, suggesting that the plant cell wall can play a major role in the organization and mobilities of PM proteins (Martinière et al., 2012).

MEMBRANE MICRODOMAINS

Dynamic, compositionally distinct regions exist within the PM that are implicated in the lateral compartmentalization of specialized signaling hubs and biological response pathways (Zappel and Panstruga, 2008; Simon-Plas et al., 2011; Urbanus and Ott, 2012). These membrane microdomains are enriched in sphingolipids and sterols relative to the rest of the membrane, which create a liquid-ordered phases distinct from the liquid-disordered membrane regions enriched in phospholipids (**Figure 2**). PM microdomains tend to contain characteristic proteins but are not static; their lipid and protein composition can be modulated during various signaling events (as recently reviewed in Mongrand et al., 2010; Simon-Plas et al., 2011; Cacas et al., 2012; Urbanus and Ott, 2012) (**Figure 2**). We should note that PM microdomain isolation using detergent insoluble membrane (DIM) preparations is prone to artifacts derived from isolation conditions and it is unlikely that DIMs preparations are equivalent to pre-existing microdomains *in vivo* (Tanner et al., 2011). Nevertheless, the utility of this technique in analyzing dynamic protein re-localization to DIMs during biological stimulus has been recently demonstrated (Minami et al., 2009; Keinath et al., 2010; Tanner et al., 2011). While plant PM microdomains are expected to have major roles in plant cell function and stress signaling, caution should be used when analyzing and interpreting proteins identified in DIM preparations.

ABIOTIC STRESS AND HORMONE SIGNALING

The PM proteome mediates many cellular responses to environmental changes and hormone signaling. Numerous physiological adaptations of plants to cold stress occur at the PM (Kawamura and Uemura, 2003; Minami et al., 2009; Li et al., 2012). Substantial changes in the abundance of PM proteins were detected after cold or ABA treatment of *Arabidopsis* suspension cell cultures using label-free ion intensity quantification (Li et al., 2012). There was a significant overlap in protein regulation during cold stress and ABA treatment, suggesting that ABA signaling mediates cold tolerance (Li et al., 2012). Another study of DIM composition during cold acclimation found that the proteins and sterols present in DIMs are modulated when plants are exposed to freezing conditions, pointing to possible mechanisms of cell survival (Minami et al., 2009). Other studies have used ^{15}N-metabolic labeling to study the effects of cadmium toxicity on PM protein regulation (Lanquar et al., 2007).

Brassinosteroids regulate a variety of plant growth and developmental processes. The PM-localized receptor kinase BRI1 directly binds BR at its extracellular domain and activates intracellular signaling (Tang et al., 2010; Clouse, 2011) (**Figure 3**). Proteomic examination of *Arabidopsis* BR responses at the PM identified proteins that change in abundance and/or phosphorylation status after BR treatment (Tang et al., 2008a,b, 2010; Wang et al., 2008; Karlova et al., 2009). Extensive phosphoproteomic analyses have identified specific regulatory sites in the somatic embryogenesis receptor-like kinase (SERK) family, whose members play diverse roles in mediating BR-signaling and immunity (Wang et al., 2005, 2008; Karlova et al., 2009; Tang et al., 2010). Phosphorylated forms of the BRI1-associated kinase BAK1 (SERK3) and the novel BR-signaling kinases BSK1 and BSK2 were detected by two-dimensional difference gel electrophoresis (2D-DIGE) shortly after BR treatment (Tang et al., 2008a,b). BRI1 phosphorylates BSK1 directly which releases it from the BRI1 PM complex and promotes its interaction with downstream cytoplasmic signaling components (**Figure 3**) (Tang et al., 2008b, 2010; Clouse, 2011). These studies highlight the advantages of using proteomic approaches to dissect complex signaling pathways and identify important, but genetically redundant signal mediators.

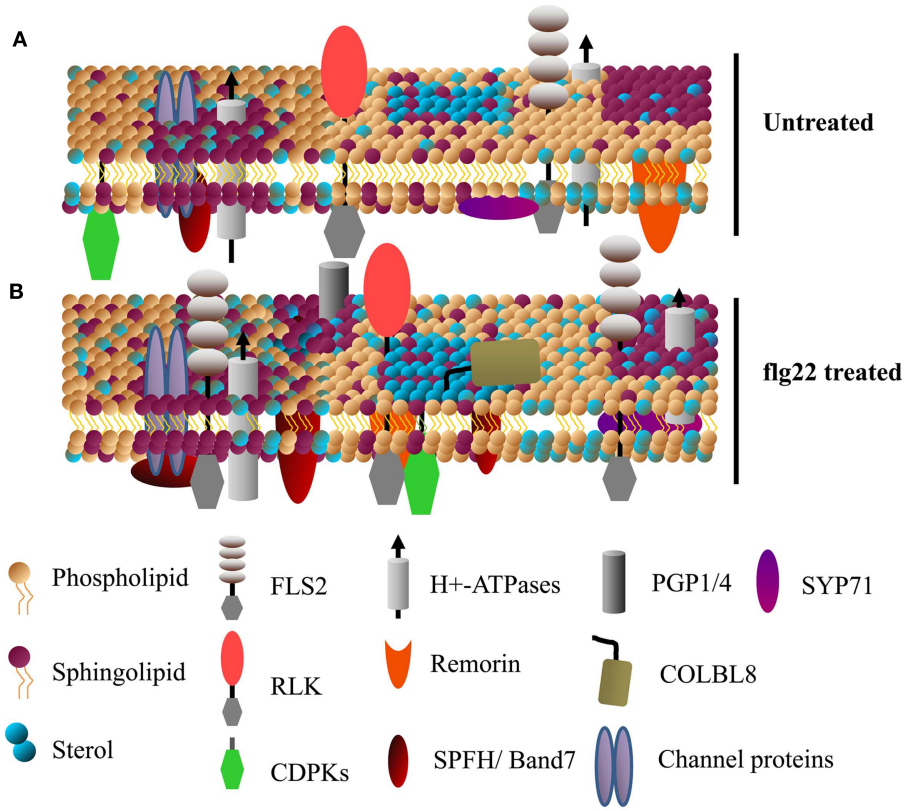


FIGURE 2 | Specific membrane proteins are enriched in detergent insoluble membranes upon activation of plant innate immunity. A model showing the dynamics of detergent insoluble membrane (DIM) protein composition upon activation of plant innate immunity. **(A,B)** Show DIMs in untreated and flg22-treated *Arabidopsis* cell suspensions, respectively. Various receptor-like kinases (RLKs, including FLS2) and calcium dependent protein kinases (CDPKs) known to play a key role in plant defense singling are enriched in DIMs. Sytaxins (e.g., SYP71), transporter proteins (e.g., AHA1), ABC transporters (e.g., PGP1/4), the GPI-anchored protein COLBL8 (COBRA-like protein 8 precursor) are also some of the proteins rapidly associated with DIM after flg22 treatment (for details, see Keinath et al., 2010). Band 7 domain-containing proteins, known to be involved in membrane microdomain/vesicle formation, are among the proteins that rapidly associate with DIMs.

BIOTIC STRESS

Many proteins that function in plant immune responses reside on or associate with the PM. Several studies have analyzed PM dynamics during pathogen perception and immune signaling. Protein phosphorylation has an extensive role in immune signaling and quantitative proteomics of phosphopeptides enriched from PM fractions isolated from tissue treated with pathogen-associated molecular patterns (PAMPs) has uncovered novel modes of protein regulation during immunity (Benschop et al., 2007; Nuhse et al., 2007). Plant defense response regulators RBOHD, SYP121, and PM H^+-ATPase were differentially phosphorylated after PAMP application (Benschop et al., 2007; Nuhse et al., 2007). A subset of these phosphosites were also demonstrated to affect protein activity (Nuhse et al., 2007) (**Figure 3**). Thus, analysis of PTMs during pathogen recognition events has contributed to a mechanistic understanding of how immune regulators are activated.

Besides post-translational modifications, the local membrane environment of PM proteins is likely to affect enzyme activity, protein complex constituents, and signal transduction events. A study of PAMP-induced changes in $^{15}N/^{14}N$-labeled *Arabidopsis* suspension cell cultures identified over 60 proteins that showed significant enrichment in DIM fractions within 15 min of flg22 (a 22 amino acid epitope of bacterial flagellin) treatment (Keinath et al., 2010). Among these, the flg22 receptor FLS2 abundance increased in DIMs, suggesting that rapid lateral compartmentalization of this receptor plays an important role in activation of downstream signaling. FLS2 undergoes endocytosis shortly after flg22 perception, and increased association with membrane microdomains could play a role in receptor endocytosis (Robatzek et al., 2006). In addition to FLS2, various other receptor kinases, PM H^+-ATPases, Ca^{2+}-ATPases, transporters, and characterized DIM-associated remorin and band 7 proteins showed enrichment in DIMs after flg22 treatment (Raffaele et al., 2009; Keinath et al., 2010; Qi et al., 2011) (**Figure 2**). The upregulation of known DIM markers in PM fractions after activation of plant immune receptors suggest that membrane microdomains have a significant role in plant disease resistance (Elmore et al., 2012).

While rapid protein re-localization to DIMs and post-translational modifications like phosphorylation can quickly

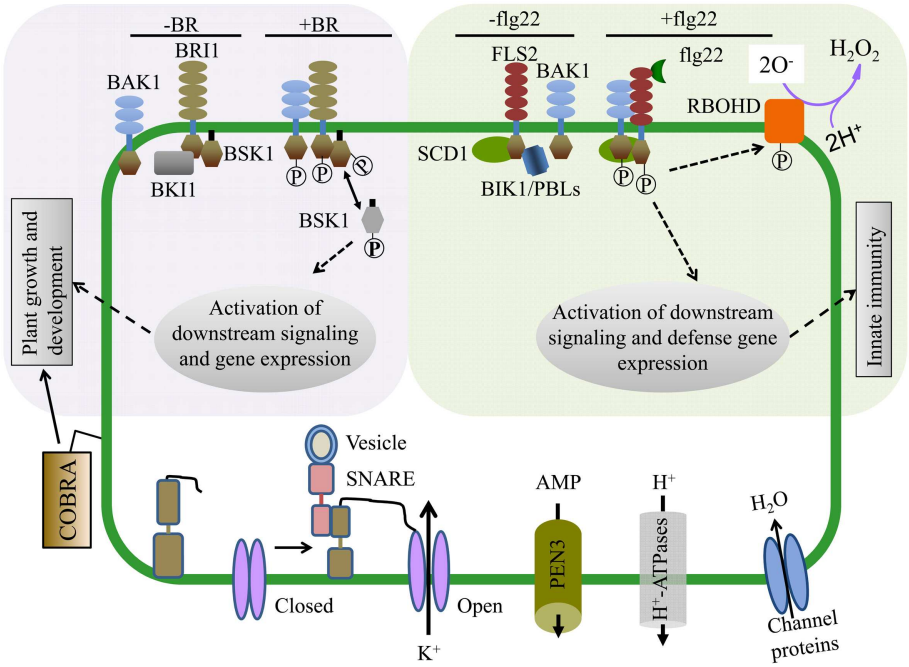


FIGURE 3 | Proteomics approaches have enabled mechanistic insight into hormone and pathogen perception as well as identified proteins required for cellular function. (Upper Left) Brassinosteroids (BR) regulate plant growth and development and are perceived by the hormone receptor BRI1 (BR-insensitive 1). In the absence of BR, BKI1 inhibits BRI1 and its downstream signaling components. In the presence of BR, BRI1 associates with its co-receptor BAK1 and phosphorylates BSK1. BSK1 then disassociates from the BR receptor complex and plays key roles in phosphorylation dependent downstream signaling leading to transcriptional changes affecting plant growth and development. (Upper Right) The FLS2 (Flagellin Sensing 2) innate immune receptor recognizes a 22 amino acid epitope of the bacterial PAMP flagellin (flg22). In the presence of flg22, FLS2 interacts with its co-receptor BAK1 and multiple transphosphorylation events occur between the kinase domains of FLS2, BAK1, and BIK1/PBLs, leading to the activation of plant innate immunity and disease resistance. Within minutes of flg22 perception, the NADPH oxidase RBOHD is activated, potassium and calcium ion fluxes occur, and the apoplastis alkalinized. (Bottom) PM proteomics studies have also identified many proteins essential for both normal cellular homeostasis as well as signaling. The abundant GPI-anchored protein COBRA controls orientational cell expansion. Multiple integral PM proteins are ion transporters, ABC transporters (e.g., PEN3, transporting antimicrobial peptides), and water transporters. PM proteins can also dynamically interact with proteins from other compartments. For example, SNAREs like SYP121 (SNARE domain-containing syntaxin) play an important role in membrane fusion and shuttling of proteins between organelles. SYP121 mediates the association between itself, an R-SNARE and the PM potassium inward rectifying channel, leading to the opening of the potassium channel and transport across the membrane.

modulate the plant immune response, the entire complement of PM proteins can change drastically over time during the execution of immunity. One study examined PM changes upon activation of the plant disease resistance protein RPS2, a signaling event that culminates in a form of programed cell death termed the hypersensitive response (HR) (Elmore et al., 2012). Relative protein quantification using spectral counting revealed that nearly 20% of the proteins identified in PM fractions significantly changed in abundance after RPS2 activation, revealing a striking alteration in PM composition during HR-associated immune responses (Elmore et al., 2012). Taken together, these studies highlight the dynamic nature of the plant PM during abiotic and biotic stress signaling and demonstrate the utility of PM proteomics approaches to study diverse biological processes.

Proteomic approaches have also been instrumental in identifying *Arabidopsis* immune-related PM protein complexes. Affinity purification-mass spectrometry (AP-MS) experiments have been instrumental in identifying interacting partners of the PAMP receptors FLS2 and EFR, the nucleotide binding-leucine repeat immune receptor RPS2, and the immune regulator RIN4 (Heese et al., 2007; Liu et al., 2009, 2011; Qi and Katagiri, 2009; Qi et al., 2011; Roux et al., 2011). It is likely that certain proteins exist in compositionally distinct complexes in different cell-types or even within the same cell. Future work using cell-type specific promoters driving expression of epitope-tagged proteins will facilitate the analysis of cell-type specific protein complexes. Thus, AP-MS experiments are an excellent tool for identifying potential interacting partners in PM protein complexes and provide a means to dissect how protein complexes are modulated under diverse signaling conditions.

CONCLUSION AND FUTURE PROSPECTS

Analysis of the *Arabidopsis* PM proteome over the last 15 years has uncovered many new insights into plant cell membrane structure and function. Recent studies have greatly advanced our understanding of PM microdomain behavior and receptor kinase-mediated signaling in *Arabidopsis* (Tang et al., 2010; Simon-Plas et al., 2011). Both top-down and bottom-up proteomics studies have been instrumental in the large-scale analysis of protein phosphorylation events during hormone and stress

signaling, which otherwise would be impossible to study using alternative experimental approaches (Kline-Jonakin et al., 2011). Many other post-translational modifications control protein function, and we are only beginning to understand the intricacies of protein regulation. The development of proteomics approaches to study PTMs outside of phosphorylation will undoubtedly uncover additional layers of complexity in plant signaling networks.

Increases in the speed and sensitivity of the mass spectrometer instrument will soon facilitate virtually complete analysis of the PM proteome in a single experiment. Combining LC-MS/MS with cell biology approaches to survey PM responses to diverse stimuli will undoubtedly play an integral role in systems biology approaches for understanding complex cellular signaling events. Furthermore, combining quantitative proteomics with transcriptomics, metabolomics, and protein–protein interaction datasets will generate a wealth of testable models that will contribute to a holistic view of cell function.

ACKNOWLEDGMENTS

We thank members of the Coaker lab for critical reading of the manuscript. This work is supported by a National Science Foundation Grant MCB-1054298. Research in the Coaker lab is supported by grants from the National Science Foundation MCB-1054298, the National Institutes of Health RO1GM092772, and the USDA National Institute of Food and Agriculture 2010-65108-20527 awarded to Gitta Coaker. J. Mitch Elmore is supported by a by National Science Foundation CREATE-IGERT Graduate Research Training Program Grant DGE-0653984.

REFERENCES

Alexandersson, E., Saalbach, G., Larsson, C., and Kjellbom, P. (2004). *Arabidopsis* plasma membrane proteomics identifies components of transport, signal transduction and membrane trafficking. *Plant Cell Physiol.* 45, 1543–1556.

Artimo, P., Jonnalagedda, M., Arnold, K., Baratin, D., Csardi, G., De Castro, E., et al. (2012). ExPASy: SIB bioinformatics resource portal. *Nucleic Acids Res.* 40, W597–W603.

Benschop, J. J., Mohammed, S., O'Flaherty, M., Heck, A. J., Slijper, M., and Menke, F. L. (2007). Quantitative phosphoproteomics of early elicitor signaling in *Arabidopsis*. *Mol. Cell. Proteomics* 6, 1198–1214.

Borner, G. H. H., Lilley, K. S., Stevens, T. J., and Dupree, P. (2003). Identification of glycosylphosphatidylinositol-anchored proteins in *Arabidopsis*. A proteomic and genomic analysis. *Plant Physiol.* 132, 568–577.

Cacas, J.-L., Furt, F., Le Guédard, M., Schmitter, J.-M., Buré, C., Gerbeau-Pissot, P., et al. (2012). Lipids of plant membrane rafts. *Prog. Lipid Res.* 51, 272–299.

Clouse, S. D. (2011). Brassinosteroid signal transduction: from receptor kinase activation to transcriptional networks regulating plant development. *Plant Cell* 23, 1219–1230.

De Smet, I., Vosz, U., Jurgens, G., and Beeckman, T. (2009). Receptor-like kinases shape the plant. *Nat. Cell Biol.* 11, 1166–1173.

Dodds, P. N., and Rathjen, J. P. (2010). Plant immunity: towards an integrated view of plant-pathogen interactions. *Nat. Rev. Genet.* 11, 539–548.

Dunkley, T. P. J., Hester, S., Shadforth, I. P., Runions, J., Weimar, T., Hanton, S. L., et al. (2006). Mapping the *Arabidopsis* organelle proteome. *Proc. Natl. Acad. Sci. U.S.A.* 103, 6518–6523.

Durek, P., Schmidt, R., Heazlewood, J. L., Jones, A., MacLean, D., Nagel, A., et al. (2010). PhosPhAt: the *Arabidopsis thaliana* phosphorylation site database. An update. *Nucleic Acids Res.* 38, D828–D834.

Elmore, J. M., Liu, J., Smith, B., Phinney, B., and Coaker, G. (2012). Quantitative proteomics reveals dynamic changes in the plasma membrane during *Arabidopsis* immune signaling. *Mol. Cell. Proteomics* 11. doi:10.1074/mcp.M111.014555

Elortza, F., Mohammed, S., Bunkenborg, J., Foster, L. J., Nühse, T. S., Brodbeck, U., et al. (2006). Modification-specific proteomics of plasma membrane proteins: identification and characterization of glycosylphosphatidylinositol-anchored proteins released upon phospholipase D treatment. *J. Proteome Res.* 5, 935–943.

Elortza, F., Nühse, T. S., Foster, L. J., Stensballe, A., Peck, S. C., and Jensen, O. N. (2003). Proteomic analysis of glycosylphosphatidylinositol-anchored membrane proteins. *Mol. Cell. Proteomics* 2, 1261–1270.

Fang, Y., Robinson, D. P., and Foster, L. J. (2010). Quantitative analysis of proteome coverage and recovery rates for upstream fractionation methods in proteomics. *J. Proteome Res.* 9, 1902–1912.

Fischer, F., and Poetsch, A. (2006). Protein cleavage strategies for an improved analysis of the membrane proteome. *Proteome Sci.* 4, 2.

Fournier, M. L., Gilmore, J. M., Martin-Brown, S. A., and Washburn, M. P. (2007). Multidimensional separations-based shotgun proteomics. *Chem. Rev.* 107, 3654–3686.

Frey, N. F. D., and Robatzek, S. (2009). Trafficking vesicles: pro or contra pathogens? *Curr. Opin. Plant Biol.* 12, 437–443.

Fujita, M., and Kinoshita, T. (2012). GPI-anchor remodeling: potential functions of GPI-anchors in intracellular trafficking and membrane dynamics. *Biochim. Biophys. Acta* 1821, 1050–1058.

Furt, F., Simon-Plas, F., and Mongrand, S. (2011). "Lipids of the plant plasma membrane," in *The Plant Plasma Membrane*, eds A. S. Murphy, B. Schulz, and W. Peer (Berlin: Springer), 3–30.

Gilmore, J. M., and Washburn, M. P. (2010). Advances in shotgun proteomics and the analysis of membrane proteomes. *J. Proteomics* 73, 2078–2091.

Heazlewood, J. L., Verboom, R. E., Tonti-Filippini, J., Small, I., and Millar, A. H. (2007). SUBA: the *Arabidopsis* subcellular database. *Nucleic Acids Res.* 35, D213–D218.

Heese, A., Hann, D. R., Gimenez-Ibanez, S., Jones, A. M., He, K., Li, J., et al. (2007). The receptor-like kinase SERK3/BAK1 is a central regulator of innate immunity in plants. *Proc. Natl. Acad. Sci. U.S.A.* 104, 12217–12222.

Hirsch-Hoffmann, M., Gruissem, W., and Baerenfaller, K. (2012). pep2pro: the high-throughput proteomics data processing, analysis and visualization tool. *Front. Plant Sci.* 3:123. doi:10.3389/fpls.2012.00123

Johansson, F., Olbe, M., Sommarin, M., and Larsson, C. (1995). Brij 58, a polyoxyethylene acyl ether, creates membrane vesicles of uniform sidedness. A new tool to obtain inside-out (cytoplasmic side-out) plasma membrane vesicles. *Plant J.* 7, 165–173.

Joshi, H. J., Hirsch-Hoffmann, M., Baerenfaller, K., Gruissem, W., Baginsky, S., Schmidt, R., et al. (2011). MASCP gator: an aggregation portal for the visualization of *Arabidopsis* proteomics data. *Plant Physiol.* 155, 259–270.

Karlova, R., Boeren, S., Van Dongen, W., Kwaaitaal, M., Aker, J., Vervoort, J., et al. (2009). Identification of in vitro phosphorylation sites in the *Arabidopsis thaliana* somatic embryogenesis receptor-like kinases. *Proteomics* 9, 368–379.

Kawamura, Y., and Uemura, M. (2003). Mass spectrometric approach for identifying putative plasma membrane proteins of *Arabidopsis* leaves associated with cold acclimation. *Plant J.* 36, 141–154.

Keinath, N. F., Kierszniowska, S., Lorek, J., Bourdais, G., Kessler, S. A., Asano, H., et al. (2010). PAMP (pathogen-associated molecular pattern)-induced changes in plasma membrane compartmentalization reveal novel components of plant immunity. *J. Biol. Chem.* 285, 39140–39149.

Kerr, I., Carrier, D., and Twycross, J. (2011). "Hormone transport," in *The Plant Plasma Membrane*, eds A. S. Murphy, B. Schulz, and W. Peer (Berlin: Springer), 379–397.

Kierszniowska, S., Seiwert, B., and Schulze, W. X. (2009). Definition of *Arabidopsis* sterol-rich membrane microdomains by differential treatment with methyl-β-cyclodextrin and quantitative proteomics. *Mol. Cell. Proteomics* 8, 612–623.

Kline-Jonakin, K. G., Barrett-Wilt, G. A., and Sussman, M. R. (2011). Quantitative plant phosphoproteomics. *Curr. Opin. Plant Biol.* 14, 507–511.

Komatsu, S. (2008). Plasma membrane proteome in *Arabidopsis* and rice. *Proteomics* 8, 4137–4145.

Kota, U., and Goshe, M. B. (2011). Advances in qualitative and quantitative plant membrane proteomics. *Phytochemistry* 72, 1040–1060.

Kusumi, A., Fujiwara, T. K., Chadda, R., Xie, M., Tsunoyama, T. A., Kalay, Z., et al. (2012). Dynamic organizing principles of the plasma membrane that regulate signal transduction: commemorating the fortieth anniversary of singer and Nicolson's fluid-mosaic model. *Annu. Rev. Cell Dev. Biol.* 28, 215–250.

Lalonde, S., Sero, A., Pratelli, R., Pilot, G., Chen, J., Sardi, M. I., et al. (2010). A membrane protein/signaling protein interaction network for *Arabidopsis* version AMPv2. *Front. Physiol.* 1:24. doi:10.3389/fphys.2010.00024

Lanquar, V., Kuhn, L., Lelièvre, F., Khafif, M., Espagne, C., Bruley, C., et al. (2007). 15N-Metabolic labeling for comparative plasma membrane proteomics in *Arabidopsis* cells. *Proteomics* 7, 750–754.

Larsson, C., Widell, S., and Kjellbom, P. (1987). "Preparation of high-purity plasma membranes," in *Methods in Enzymology*, ed. R. D. Lester Packer (San Diego: Academic Press), 558–568.

Li, B., Takahashi, D., Kawamura, Y., and Uemura, M. (2012). Comparison of plasma membrane proteomic changes of *Arabidopsis* suspension-cultured cells (T87 Line) after cold and ABA treatment in association with freezing tolerance development. *Plant Cell Physiol.* 53, 543–554.

Li, X., Wang, X., Yang, Y., Li, R., He, Q., Fang, X., et al. (2011). Single-molecule analysis of PIP2;1 dynamics and partitioning reveals multiple modes of *Arabidopsis* plasma membrane aquaporin regulation. *Plant Cell* 23, 3780–3797.

Liu, J., Elmore, J. M., Fuglsang, A. T., Palmgren, M., Staskawicz, B., and Coaker, G. (2009). RIN4 functions with the plasma membrane H+-ATPase to regulate stomatal apertures during pathogen attack. *PLoS Biol.* 7:e1000139. doi:10.1371/journal.pbio.1000139

Liu, J., Elmore, J. M., Lin, Z. J., and Coaker, G. (2011). A receptor-like cytoplasmic kinase phosphorylates the host target RIN4, Leading to the Activation of a Plant Innate Immune Receptor. *Cell Host Microbe* 9, 137–146.

Luche, S., Santoni, V., and Rabilloud, T. (2003). Evaluation of non-ionic and zwitterionic detergents as membrane protein solubilizers in two-dimensional electrophoresis. *Proteomics* 3, 249–253.

Marmagne, A., Ferro, M., Meinnel, T., Bruley, C., Kuhn, L., Garin, J., et al. (2007). A high content in lipid-modified peripheral proteins and integral receptor kinases features in the *Arabidopsis* plasma membrane proteome. *Mol. Cell. Proteomics* 6, 1980–1996.

Marmagne, A., Rouet, M.-A., Ferro, M., Rolland, N., Alcon, C., Joyard, J., et al. (2004). Identification of new intrinsic proteins in *Arabidopsis* plasma membrane proteome. *Mol. Cell. Proteomics* 3, 675–691.

Martinière, A., Lavagi, I., Nageswaran, G., Rolfe, D. J., Maneta-Peyret, L., Luu, D.-T., et al. (2012). Cell wall constrains lateral diffusion of plant plasma-membrane proteins. *Proc. Natl. Acad. Sci. U.S.A.* 109, 12805–12810.

Minami, A., Fujiwara, M., Furuto, A., Fukao, Y., Yamashita, T., Kamo, M., et al. (2009). Alterations in detergent-resistant plasma membrane microdomains in *Arabidopsis thaliana* during cold acclimation. *Plant Cell Physiol.* 50, 341–359.

Mitra, S. K., Gantt, J. A., Ruby, J. F., Clouse, S. D., and Goshe, M. B. (2007). Membrane proteomic analysis of *Arabidopsis thaliana* using alternative solubilization techniques. *J. Proteome Res.* 6, 1933–1950.

Monaghan, J., and Zipfel, C. (2012). Plant pattern recognition receptor complexes at the plasma membrane. *Curr. Opin. Plant Biol.* 15, 349–357.

Mongrand, S., Stanislas, T., Bayer, E. M. F., Lherminier, J., and Simon-Plas, F. (2010). Membrane rafts in plant cells. *Trends Plant Sci.* 15, 656–663.

Nelson, C. J., Hegeman, A. D., Harms, A. C., and Sussman, M. R. (2006). A quantitative analysis of *Arabidopsis* plasma membrane using trypsin-catalyzed 18O labeling. *Mol. Cell. Proteomics* 5, 1382–1395.

Nuhse, T. S., Bottrill, A. R., Jones, A. M., and Peck, S. C. (2007). Quantitative phosphoproteomic analysis of plasma membrane proteins reveals regulatory mechanisms of plant innate immune responses. *Plant J.* 51, 931–940.

Palmgren, M. G., Askerlund, P., Fredrikson, K., Widell, S., Sommarin, M., and Larsson, C. (1990). Sealed inside-out and right-side-out plasma membrane vesicles: optimal conditions for formation and separation. *Plant Physiol.* 92, 871–880.

Pandey, S., Nelson, D. C., and Assmann, S. M. (2009). Two novel GPCR-type G proteins are abscisic acid receptors in *Arabidopsis*. *Cell* 136, 136–148.

Piersma, S. R., Fiedler, U., Span, S., Lingnau, A., Pham, T. V., Hoffmann, S., et al. (2010). Workflow comparison for label-free, quantitative secretome proteomics for cancer biomarker discovery: method evaluation, differential analysis, and verification in serum. *J. Proteome Res.* 9, 1913–1922.

Prime, T. A., Sherrier, D. J., Mahon, P., Packman, L. C., and Dupree, P. (2000). A proteomic analysis of organelles from *Arabidopsis thaliana*. *Electrophoresis* 21, 3488–3499.

Qi, Y., and Katagiri, F. (2009). Purification of low-abundance *Arabidopsis* plasma-membrane protein complexes and identification of candidate components. *Plant J.* 57, 932–944.

Qi, Y., Tsuda, K., Nguyen, L. V., Wang, X., Lin, J., Murphy, A. S., et al. (2011). Physical association of *Arabidopsis* hypersensitive induced reaction proteins (HIRs) with the immune receptor RPS2. *J. Biol. Chem.* 286, 31297–31307.

Raffaele, S., Bayer, E., Lafarge, D., Cluzet, S., German Retana, S., Boubekeur, T., et al. (2009). Remorin, a *Solanaceae* protein resident in membrane rafts and plasmodesmata, impairs potato virus X movement. *Plant Cell* 21, 1541–1555.

Robatzek, S., Chinchilla, D., and Boller, T. (2006). Ligand-induced endocytosis of the pattern recognition receptor FLS2 in *Arabidopsis*. *Genes Dev.* 20, 537–542.

Robert, S., Kleine-Vehn, J., Barbez, E., Sauer, M., Paciorek, T., Baster, P., et al. (2010). ABP1 mediates auxin inhibition of clathrin-dependent endocytosis in *Arabidopsis*. *Cell* 143, 111–121.

Roux, M., Schwessinger, B., Albrecht, C., Chinchilla, D., Jones, A., Holton, N., et al. (2011). The *Arabidopsis* leucine-rich repeat receptor-like kinases BAK1/SERK3 and BKK1/SERK4 are required for innate immunity to hemibiotrophic and biotrophic pathogens. *Plant Cell* 23, 2440–2455.

Sanderfoot, A. A., Assaad, F. F., and Raikhel, N. V. (2000). The *Arabidopsis* genome. An abundance of soluble N-ethylmaleimide-sensitive factor adaptor protein receptors. *Plant Physiol.* 124, 1558–1569.

Santoni, V. (2007). "Plant plasma membrane protein extraction and solubilization for proteomic analysis," in *Plant Proteomics*, eds H. Thiellement, M. Zivy, C. Damerval, and V. Méchin (New York: Humana Press), 93–109.

Santoni, V., Molloy, M., and Rabilloud, T. (2000). Membrane proteins and proteomics: un amour impossible? *Electrophoresis* 21, 1054–1070.

Santoni, V., Rabilloud, T., Doumas, P., Rouquié, D., Mansion, M., Kieffer, S., et al. (1999). Towards the recovery of hydrophobic proteins on two-dimensional electrophoresis gels. *Electrophoresis* 20, 705–711.

Santoni, V., Rouquié, D., Doumas, P., Mansion, M., Boutry, M., Degand, H., et al. (1998). Use of a proteome strategy for tagging proteins present at the plasma membrane. *Plant J.* 16, 633–641.

Santoni, V., Vinh, J., Pflieger, D., Sommerer, N., and Maurel, C. (2003). A proteomic study reveals novel insights into the diversity of aquaporin forms expressed in the plasma membrane of plant roots. *Biochem. J.* 373, 289–296.

Schulz, B. (2011). "Functional classification of plant plasma membrane transporters," in *The Plant Plasma Membrane*, eds A. S. Murphy, B. Schulz, and W. Peer (Berlin: Springer), 131–176.

Schulze, W. X., and Usadel, B. (2010). Quantitation in mass-spectrometry-based proteomics. *Annu. Rev. Plant Biol.* 61, 491–516.

Schwacke, R., Schneider, A., Van Der Graaff, E., Fischer, K., Catoni, E., Desimone, M., et al. (2003). ARAMEMNON, a novel database for *Arabidopsis* integral membrane proteins. *Plant Physiol.* 131, 16–26.

Simon-Plas, F., Perraki, A., Bayer, E., Gerbeau-Pissot, P., and Mongrand, S. (2011). An update on plant membrane rafts. *Curr. Opin. Plant Biol.* 14, 642–649.

Singer, S. J., and Nicolson, G. L. (1972). The fluid mosaic model of the structure of cell membranes. *Science* 175, 720–731.

Sucaet, Y., Wang, Y., Li, J., and Wurtele, E. (2012). MetNet online: a novel integrated resource for plant systems biology. *BMC Bioinformatics* 13:267. doi:10.1186/1471-2105-13-267

Tan, S., Tan, H. T., and Chung, M. C. M. (2008). Membrane proteins and membrane proteomics. *Proteomics* 8, 3924–3932.

Tang, W., Deng, Z., Oses-Prieto, J. A., Suzuki, N., Zhu, S., Zhang, X., et al. (2008a). Proteomics studies

of brassinosteroid signal transduction using prefractionation and two-dimensional DIGE. *Mol. Cell. Proteomics* 7, 728–738.

Tang, W., Kim, T.-W., Oses-Prieto, J. A., Sun, Y., Deng, Z., Zhu, S., et al. (2008b). BSKs mediate signal transduction from the receptor kinase BRI1 in *Arabidopsis*. *Science* 321, 557–560.

Tang, W., Deng, Z., and Wang, Z.-Y. (2010). Proteomics shed light on the brassinosteroid signaling mechanisms. *Curr. Opin. Plant Biol.* 13, 27–33.

Tanner, W., Malinsky, J., and Opekarová, M. (2011). In plant and animal cells, detergent-resistant membranes do not define functional membrane rafts. *Plant Cell* 23, 1191–1193.

Tchieu, J. H., Fana, F., Fink, J. L., Harper, J., Nair, T. M., Niedner, R. H., et al. (2003). The plantsP and plantsT functional genomics databases. *Nucleic Acids Res.* 31, 342–344.

Testerink, C., and Munnik, T. (2011). Molecular, cellular, and physiological responses to phosphatidic acid formation in plants. *J. Exp. Bot.* 62, 2349–2361.

Urbanus, S. L., and Ott, T. (2012). Plasticity of plasma membrane compartmentalization during plant immune responses. *Front. Plant Sci.* 3:181. doi:10.3389/fpls.2012.00181

Wang, C., Marshall, A., Zhang, D., and Wilson, Z. A. (2012). ANAP: an integrated knowledge base for *Arabidopsis* protein interaction network analysis. *Plant Physiol.* 158, 1523–1533.

Wang, X., Goshe, M. B., Soderblom, E. J., Phinney, B. S., Kuchar, J. A., Li, J., et al. (2005). Identification and functional analysis of in vivo phosphorylation sites of the *Arabidopsis* brassinosteroid-insensitive1 receptor kinase. *Plant Cell* 17, 1685–1703.

Wang, X., Kota, U., He, K., Blackburn, K., Li, J., Goshe, M. B., et al. (2008). Sequential transphosphorylation of the BRI1/BAK1 receptor kinase complex impacts early events in brassinosteroid signaling. *Dev. Cell* 15, 220–235.

Wang, Z. Y., Seto, H., Fujioka, S., Yoshida, S., and Chory, J. (2001). BRI1 is a critical component of a plasma-membrane receptor for plant steroids. *Nature* 410, 380–383.

Washburn, M. P., Wolters, D., and Yates, J. R. (2001). Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotech.* 19, 242–247.

Wilkins, M. R., Gasteiger, E., Sanchez, J.-C., Bairoch, A., and Hochstrasser, D. F. (1998). Two-dimensional gel electrophoresis for proteome projects: the effects of protein hydrophobicity and copy number. *Electrophoresis* 19, 1501–1505.

Wu, C. C., MacCoss, M. J., Howell, K. E., and Yates, J. R. (2003). A method for the comprehensive proteomic analysis of membrane proteins. *Nat. Biotech.* 21, 532–538.

Yao, Q., Bollinger, C., Gao, J., Xu, D., and Thelen, J. J. (2012). P3DB: an integrated database for plant protein phosphorylation. *Front. Plant Sci.* 3:206. doi:10.3389/fpls.2012.00206

Zappel, N. F., and Panstruga, R. (2008). Heterogeneity and lateral compartmentalization of plant plasma membranes. *Curr. Opin. Plant Biol.* 11, 632–640.

Zhang, Z. J., and Peck, S. C. (2011). Simplified enrichment of plasma membrane proteins for proteomic analyses in *Arabidopsis thaliana*. *Proteomics* 11, 1780–1788.

Zhao, Y., and Jensen, O. N. (2009). Modification-specific proteomics: strategies for characterization of post-translational modifications using enrichment techniques. *Proteomics* 9, 4632–4641.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 04 December 2012; paper pending published: 17 February 2013; accepted: 22 March 2013; published online: 11 April 2013.

Citation: Yadeta KA, Elmore JM and Coaker G (2013) Advancements in the analysis of the Arabidopsis plasma membrane proteome. Front. Plant Sci. ***4****:86. doi: 10.3389/fpls.2013.00086*

This article was submitted to Frontiers in Plant Proteomics, a specialty of Frontiers in Plant Science.

Copyright © 2013 Yadeta, Elmore and Coaker. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.

frontiers in
PLANT SCIENCE

ORIGINAL RESEARCH ARTICLE
published: 06 March 2013
doi: 10.3389/fpls.2013.00033

Developmental distribution of the plasma membrane-enriched proteome in the maize primary root growth zone

Zhe Zhang*[1,2,3†‡], *Priyamvada Voothuluru*[3,4‡], *Mineo Yamaguchi*[3,4], *Robert E. Sharp*[3,4] *and Scott C. Peck*[1,2,3]

[1] *Division of Biochemistry, University of Missouri, Columbia, MO, USA*
[2] *Christopher S. Bond Life Sciences Center, University of Missouri, Columbia, MO, USA*
[3] *Interdisciplinary Plant Group, University of Missouri, Columbia, MO, USA*
[4] *Division of Plant Sciences, University of Missouri, Columbia, MO, USA*

Edited by:
Harvey Millar, The University of Western Australia, Australia

Reviewed by:
Berit Ebert, Lawrence Berkeley Lab Joint BioEnergy Institute, USA
Ning LI, The Hong Kong University of Science and Technology, China

****Correspondence:***
Scott C. Peck, Department of Biochemistry, Christopher S. Bond Life Science Center, University of Missouri, 271H Bond Life Sciences Center, 1201 Rollins Street, Columbia MO65211, USA.
e-mail: pecks@missouri.edu

†Present address:
Zhe Zhang, Department of Hematology, Mayo Clinic, Rochester, MN, USA.

‡Zhe Zhang and Priyamvada Voothuluru have contributed equally to this work.

Within the growth zone of the maize primary root, there are well-defined patterns of spatial and temporal organization of cell division and elongation. However, the processes underlying this organization remain poorly understood. To gain additional insights into the differences amongst the defined regions, we performed a proteomic analysis focusing on fractions enriched for plasma membrane (PM) proteins. The PM is the interface between the plant cell and the apoplast and/or extracellular space. As such, it is a key structure involved in the exchange of nutrients and other molecules as well as in the integration of signals that regulate growth and development. Despite the important functions of PM-localized proteins in mediating these processes, a full understanding of dynamic changes in PM proteomes is often impeded by low relative concentrations relative to total proteins. Using a relatively simple strategy of treating microsomal fractions with Brij-58 detergent to enrich for PM proteins, we compared the developmental distribution of proteins within the root growth zone which revealed a number of previously known as well as novel proteins with interesting patterns of abundance. For instance, the quantitative proteomic analysis detected a gradient of PM aquaporin proteins similar to that previously reported using immunoblot analyses, confirming the veracity of this strategy. Cellulose synthases increased in abundance with increasing distance from the root apex, consistent with expected locations of cell wall deposition. The similar distribution pattern for Brittle-stalk-2-like protein implicates that this protein may also have cell wall related functions. These results show that the simplified PM enrichment method previously demonstrated in *Arabidopsis* can be successfully applied to completely unrelated plant tissues and provide insights into differences in the PM proteome throughout growth and development zones of the maize primary root.

Keywords: plasma membrane, proteomics, growth zone, maize, roots, development

INTRODUCTION

Unlike animals, plants continue to grow and increase in size throughout their lifecycle. Plant growth, however, does not occur indiscriminately but rather is distributed through defined regions of roots and shoots that are referred to as the growth zones (Erickson and Silk, 1979). Within these regions of growth, there is considerable variability in spatial and temporal cell production and cell elongation, two processes defining the overall growth rate (Peters and Bernstein, 1997; Beemster and Baskin, 1998). Therefore, these growth zones have been used extensively to study the processes contributing to cell production and elongation (Pilet et al., 1983; MacAdam et al., 1992; Bernstein et al., 1993; Liang et al., 1997).

Among plant organs, the root has a relatively simple organization of the growth zone, making it a good model for studying various aspects of plant growth (Beemster and Baskin, 1998; Sharp et al., 2004; Brady et al., 2005). The meristematic region with dividing cells is near the tip of the root, and the elongation region consisting of cells at various stages of expansion progresses away from the zone of division. Cell elongation accelerates as cells move out of the meristem, reaching a maximal rate in the middle of the growth zone, and decelerates thereafter such that cells reach their final length by the end of the growth zone. This spatial distribution of cell production and cell elongation has been exploited for numerous studies of growth related processes, revealing gradients within these zones of metabolites, phytohormones, pH, transcription factor proteins, and cell wall-modifying proteins (Mulkey and Evans, 1980; Pilet et al., 1983; Sharp et al., 1990; Baskin et al., 2004; Brady et al., 2005; Wolters and Jurgens, 2009; Yamaguchi and Sharp, 2010).

Because the plasma membrane (PM) is the interface between a cell and the apoplast/environment, it is a key structure involved in integrating signals and responses involved in cell growth. For instance, PM-localized proteins such as ion and water channels are

necessary for regulating turgor and cell elongation in roots (Kiegle et al., 2000; Hachez et al., 2006). The acidification of apoplastic pH by the PM-localized proton ATPase has been shown to be important for wall loosening and cell elongation in different plant species (Pilet et al., 1983; Cosgrove, 1989). Additionally, cells at various stages of elongation must constantly modify their cell walls, and several metabolites needed for cell wall deposition are transported across the PM (Wightman and Turner, 2010; Endler and Persson, 2011).

Despite the important functions of PM proteins in the regulation of growth, these proteins are often underrepresented in proteomic analyses because they are found in low concentrations relative to the total cellular protein content (Ephritikhine et al., 2004; Morel et al., 2006; Zhang and Peck, 2011). To increase depth of coverage, aqueous two-phase partitioning is often utilized which, although ideal for yielding highly purified PM fractions, is both time-consuming and technically challenging, often requiring significant technical optimization for each new tissue or species examined (Albertsson et al., 1987; Komatsu, 2008). Although crude microsomal fractionation does enrich for PM proteins over total protein, the large degree of contamination from other organellar proteins, particularly from the endoplasmic reticulum (ER), still limits the depth of coverage of PM proteomes. Using *Arabidopsis* suspension cell cultures as starting material, Zhang and Peck (2011) recently reported a simple method for decreasing the representation of organellar proteins from crude microsomal fractions to obtain greater than threefold enrichment of PM proteins from *Arabidopsis* culture cells for proteomic analyses while decreasing contamination with ER proteins by sevenfold. Although this method is not applicable to assigning definitive location of a protein to the PM because it does not yield samples as pure as those from aqueous two-phase partitioning, the strategy is useful to enrich sufficiently for the PM fraction to allow for meaningful quantitative comparisons. In the current study, we evaluated the applicability of this simplified PM enrichment method using the growth zone of maize primary roots grown under well-watered conditions. We demonstrate that the strategy is easily transferred to this new tissue and species, and we report the region-specific distribution of proteins, including many PM proteins, associated with the spatial growth pattern.

MATERIALS AND METHODS

CHEMICALS

All chemicals used in this study were ultrapure grade (obtained from Sigma-Aldrich Co, St. Louis, MO, USA; Fisher Scientific, Pittsburgh, PA, USA; Promega Corp., Madison, WI, USA). HPLC-water was obtained from the Millipore Synthesis system (Millipore Corp. Billerica, MA, USA).

PLANT GROWTH AND TISSUE COLLECTION

B73 X MO17 hybrid maize seed were used in all experiments. Seeds were surface sterilized in 5% NaClO solution for 15 min, rinsed with deionized water for 15 min, and imbibed in aerated 1 mM $CaSO_4$ solution for 24 h. The imbibed seeds were germinated between sheets of germination paper moistened with 1 mM $CaSO_4$ solution at 29°C and near-saturating humidity in the dark. Seedlings with primary roots of 10–20 mm in length were transplanted against the interior surface of Plexiglass containers filled with vermiculite (no. 2A, Therm-O-Rock East Inc., New Eagle, PA, USA) which was moistened to the drip point with 1 mM $CaSO_4$ solution. The seedlings were then grown at 29°C and near-saturating humidity in the dark for 48 h. Primary root elongation was monitored by periodically marking the position of the root apices on the Plexiglass. The apical 20 mm of the primary roots were harvested and divided into four regions (all distances are from the root apex including the root cap); the 0–3 mm region (R1), the 3–7 mm region (R2), the 7–12 mm region (R3), and the 12–20 mm region (R4) (**Figure 1A**). The harvested root segments were collected by position, transferred to tubes containing liquid nitrogen, and stored at −80°C. In each of three replicate experiments, root segments were harvested from >150 seedlings for proteomic analysis. Transplanting, root elongation measurements, and harvesting were performed using a green "safe" light (Saab et al., 1990).

KINEMATIC ANALYSIS OF DISPLACEMENT VELOCITY AND RELATIVE ELONGATION RATE PROFILES

The spatial distributions of displacement velocity and relative elongation rate in the primary root growth zone were obtained from cell length profiles and root elongation rates as described by Silk et al. (1989). Briefly, 10 seedlings were grown for 48 h and their root elongation monitored periodically. Elongation rates were steady after ∼15 h from transplanting as required for accurate determinations of elongation rate profiles from anatomical records. Three to four roots that were straight and had an elongation rate similar to the mean of the population were then harvested for cell length measurements. The apical 15 mm was sectioned longitudinally using a vibratome (Lancer series 1000 Vibratome, St Louis, MO, USA) and a 125 μm thick section of each root was stained with 1 mg mL^{-1} of Calcofluor (Sigma-Aldrich, St. Louis, MO, USA) for 15 min to visualize the cell walls. The stained sections were imaged by confocal microscopy as described by Yamaguchi et al. (2010). Average cell lengths at various positions from the root apex were calculated by measuring the cell lengths of 4–12 cortical cells. The final cell lengths were calculated by averaging the four most distal measurement positions (12–15 mm from the root apex). Displacement velocities were calculated using the relationship $L_A/L_F = V_A/V_F$, as described in Silk et al. (1989) where

L_A = the mean cell length at position A
L_F = final cell length
V_A = displacement velocity at position A
V_F = the final displacement velocity (equal to the root elongation rate).

The cell length method cannot be used to calculate accurate displacement velocities in the meristematic region (Silk et al., 1989). Therefore, displacement velocities were calculated starting at the distal end of the meristem, which was estimated to occur at a cell length of 2.5 times the length of the shortest cells (2.5 mm from the root apex; Erickson, 1961). The mean displacement velocities were plotted and a fifth-order polynomial curve was fitted to the

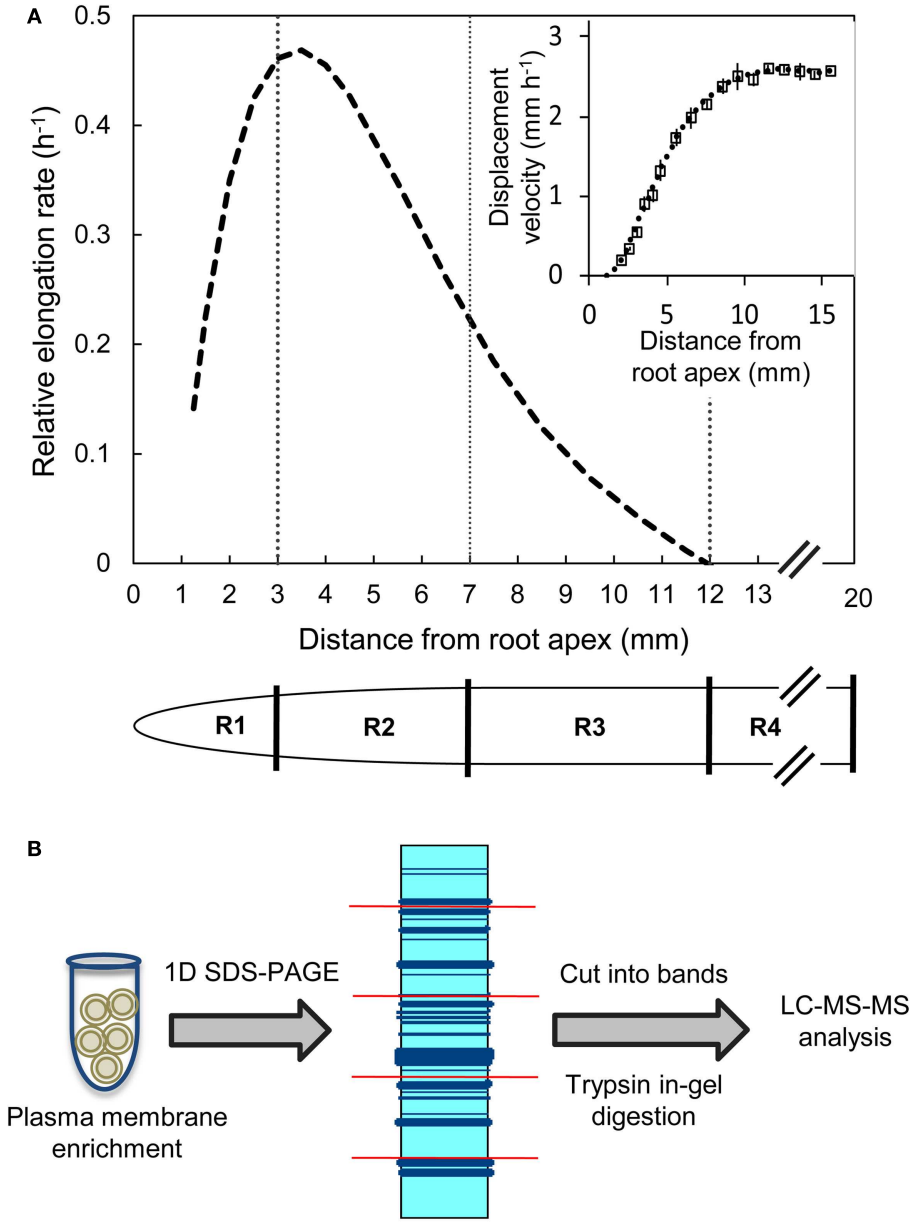


FIGURE 1 | (A) Relative elongation rate profile and displacement velocity profile (inset) as a function of distance from the root apex (including the root cap) of the primary root of maize hybrid line B73 X MO17 grown under well-watered conditions. The spatial distribution of displacement velocity (mm h^{-1}) from the root apex was calculated from root elongation rate and cell length profile data obtained at 48 h after transplanting (data are means ± SE, $n = 3$). A fifth-order polynomial curve was fitted to the displacement velocity profile and the derivative of displacement velocity against position was used to obtain the relative elongation rate profile (h^{-1}). Tissues from region 1 (R1, 0–3 mm), region 2 (R2, 3–7 mm), region 3 (R3, 7–12 mm), and region 4 (R4, 12–20 mm) were harvested. **(B)** Label-free proteomics work flow shown for plasma membrane-enriched samples as protein input followed by in-gel tryptic digestion and LC-MS/MS analysis.

data (**Figure 1A** inset). The derivative of the resulting curve was used to obtain the relative elongation rate profile (**Figure 1A**; Silk et al., 1989).

PROTEIN EXTRACTION AND PLASMA MEMBRANE ENRICHMENT

Frozen root tissues were ground to a fine powder and proteins were extracted as described by Zhang and Peck (2011). Briefly, proteins were extracted from the tissues using 1 mL of ice-cold buffer H (330 mM sucrose, 50 mM HEPES/KOH pH 7.5, 50 mm $Na_4P_2O_7$, 25 mm NaF, 5% glycerol, 0.5% polyvinyl pyrrolidone, 10 mm EDTA, 1 mM Na_2MoO_4, 1 mM PMSF, 10 mM leupeptin A, 1 nM calyculin A, and 3 mM DTT) per gram fresh weight of tissue. The samples were subsequently centrifuged for 10 min (10,000 × g at 4°C) to remove cell debris. The supernatant was

used to obtain crude microsomal pellets by ultracentrifugation for 30 min (100,000 × *g* at 4°C). The microsomal pellets were washed with buffer H without DTT, incubated in 2 μL of buffer B (buffer H without DTT but with 0.02% w/v Brij-58) per μg of crude microsomal protein on ice for 45 min and centrifuged for 30 min at 100,000 × *g*. The resulting pellets were washed with 10 mM Na_2CO_3 (pH 11–12) in buffer H without DTT to yield the PM-enriched protein fraction.

IMMUNOBLOTTING

Protein samples from the soluble fraction, crude microsomes, and PM-enriched fraction were separated by SDS-PAGE and transferred to Immobilon P membrane for 2 h at 70 V in Towbin's transfer buffer (39 mM glycine, 48 mM Tris-base, 0.037% SDS, and 20% methanol) at 4°C. The membranes were incubated first in blocking buffer (Tris-buffered saline at pH 7.8 containing 5% non-fat milk and 0.05% Tween-20) for 1 h at room temperature, then incubated with primary antibodies in blocking buffer at 4°C overnight, and finally incubated in blocking buffer containing the secondary antibody horseradish peroxidase-labeled anti-rabbit IgG (Sigma-Aldrich Co, St. Louis, MO, USA) for 1 h at room temperature. Supersignal Femto or Pico substrate (Thermo Scientific, IL, USA) was used for chemiluminescence detection. The primary antibodies (antibodies were obtained from Agrisera, Sweden) used in the study were α-AHA (H^+-ATPase, PM marker) and α-SMT-1 (sterol methyltransferase-1, ER marker).

SDS-PAGE ANALYSIS AND TRYPTIC DIGESTION

For each biological replicate, equal concentrations (∼30 μg) of PM-enriched protein from each region of the growth zone were solubilized in sample loading buffer, heated to 75°C, separated by 8% SDS-PAGE and stained with colloidal Coomassie G-250 overnight, as described by Neuhoff et al. (1985). The following day, gels were destained in distilled water, and each gel lane was cut into eight slices with a razor blade. The proteins from the gel slices were reduced with 50 mM TCEP-HCl, alkylated with 50 mM iodoacetamide and digested within the gel overnight using 1:20 w/w trypsin in 100 mM ammonium bicarbonate buffer (pH 8.3) at 37°C. After digestion, the peptides were eluted twice with 1% trifluoroacetic acid with 60% acetonitrile, and the eluted mixture was lyophilized overnight to obtain dried peptides. The dry peptides were stored at −80°C until LC-MS/MS analysis.

LC-MS/MS ANALYSIS

Lyophilized peptides dissolved in 0.1% formic acid were applied to a 10 cm prepacked column (Picotips with 75 μm inner diameter and 15 μm tip, obtained from New Objective, Woburn, MA, USA) and eluted into the nanoelectrospray ion source of a LTQ-Orbitrap LC-MS/MS mass spectrometer (Thermoelectron Corp., Rockford, IL, USA) that was controlled by XCalibur version 2.2.1. The mass spectrometer operating in data-dependent mode was used to carry out a fully automated chromatography run using 1% formic acid and 99.9% acetonitrile, 0.1% formic acid with a 1% per min incremental gradient for the first 45 min and 11% per min for the final 5 min. The mass spectrometer measurements were obtained with the specifications as described in Zhang and Peck (2011).

PEPTIDE AND PROTEIN IDENTIFICATION

Mascot Distiller version 2.0 (Matrix Science, London, UK) was used to deconvolute the tandem mass spectra. However, deisotoping was not performed. Mascot (server version 2.3, Matrix Science, London, UK) and X! Tandem (version 2007.01.01.1)[1] were used to analyze the MS/MS spectra by searching an in-house database created by downloading all *Zea mays* proteins from NCBI[2] and filtering for duplicate entries. The MS/MS based peptide and protein identities were validated by Scaffold (version Scaffold_3_00_08, Proteome Software Inc., Portland, OR, USA). The specified variable modifications were oxidized methionine and iodoacetamide derivative of cysteine. The identities of peptides (at greater than 95% probability) were accepted using the Peptide Prophet algorithm (Keller et al., 2002), and identities of proteins with at least two peptides identified (at greater than 99% probability) were accepted using the Protein Prophet algorithm (Nesvizhskii et al., 2003), respectively. If proteins containing similar peptides could not be differentiated based on MS/MS analysis alone, they were grouped to satisfy the principles of parsimony.

DATA ANALYSIS AND BIOINFORMATICS

Spectral counts were normalized within a biological replicate using the mean of total spectral counts from all four regions (the average deviation within an experiment was ∼3%). For each region, the mean and standard deviation were calculated using the spectral counts from all three biological experiments. Proteins for which the fold difference between the means was greater than the two times the coefficient of variation (CV) determined for a technical replicate determined using these samples (CV = 0.34) was considered for pattern analysis.

To assign the distribution of proteins within the different regions, the protein data were analyzed according to the following steps. Proteins identified in at least two out of the three biological replicates within a region were considered reproducible. Reproducible proteins were used to generate a four-way Venn diagram (**Figure 2C**) using the algorithm at http://bioinfogp.cnb.csic.es/tools/venny/index.html. Proteins with a region-specific distribution pattern (i.e., with zero spectral counts in one or more region) were grouped into the three classes "single region present," "single region absent," and "two-region present" (**Table 2**). For proteins that were present in all regions, five major patterns of distribution were identified that were based on the gradient starting from R1 to R4, such as "decreasing," "increasing," "R1-lowest," "R2-highest," and "R3-highest" (**Table 1**). Proteins that were not present in all regions were grouped into three classes: "single region present," "single region absent," and "two regions present/absent" (**Table 2**).

RESULTS AND DISCUSSION

SPATIAL DISTRIBUTION OF RELATIVE ELONGATION RATE IN THE PRIMARY ROOT GROWTH ZONE

Kinematic analysis showed that the primary root growth zone encompassed the apical 12 mm (**Figure 1A**). The relative elongation rate increased as the cells were displaced away from the

[1] http://www.thegpm.org
[2] http://www.ncbi.nlm.nih.gov/protein

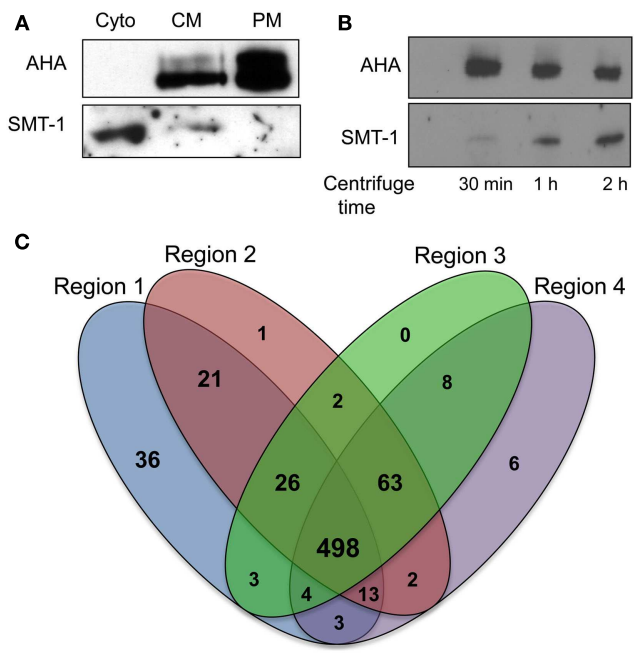


FIGURE 2 | Application of plasma membrane enrichment strategy to analysis of maize primary roots. (A) Proteins from steps of the enrichment procedure were analyzed by immunoblotting with organelle specific markers (Cyto, cytosolic fraction; CM, crude microsomal fraction; PM, plasma membrane-enriched fraction). Equal amounts of proteins from each fraction were analyzed with antibodies specific for a plasma membrane H^+-ATPase (AHA) or an endoplasmic reticulum integral membrane sterol methyltransferase (SMT-1). **(B)** Increasing time of ultracentrifugation increases ER proteins in crude microsomal pellets. Proteins were prepared for immunoblot analysis as for **(A)**, except that the ultracentrifugation at 100,000 × g was performed for 30 min, 1 h, or 2 h. Equal amounts of proteins from each crude microsomal fraction were analyzed with antibodies specific for PM H^+-ATPases (AHA) or an endoplasmic reticulum integral membrane sterol methyltransferase (SMT-1). **(C)** Four-way Venn diagram showing number of proteins identified from the four regions (created using algorithm from: http://bioinfogp.cnb.csic.es/tools/venny/index.html).

apex, reaching a maximum between 3–4 mm followed by a gradual decrease to reach zero at about 12 mm. Therefore, the apical 20 mm was divided into four contiguous regions based on the relative elongation rate profile: the 0–3 mm region (R1) that showed acceleration in elongation rate, the 3–7 mm region (R2) that exhibited the initial phase of decelerating elongation rate ($>0.2\ h^{-1}$), the 7–12 mm region (R3) of decelerating elongation rate ($<0.2\ h^{-1}$), and the 12–20 mm region (R4) that showed no elongation (**Figure 1A**). These tissues were then used for comparisons of PM-enriched proteomes.

PLASMA MEMBRANE ENRICHMENT AND LC-MS/MS EFFICACY

A simplified method for PM enrichment based on treatment of microsomal fractions with Brij-58 detergent was recently reported using *Arabidopsis* cell cultures (Zhang and Peck, 2011). The relative efficacy of this PM enrichment strategy was tested in maize root samples by conducting immunoblot experiments with the PM marker AHA (a family of integral PM proton ATPases) and the ER membrane marker SMT-1 (integral ER sterol methyltransferase). Equal amounts of protein from the soluble protein fraction (obtained from the first ultracentrifugation step), crude microsomal fraction, and PM-enriched fraction were separated by SDS-PAGE and blotted with antibodies of AHA1 and SMT-1 (**Figure 2A**). Comparisons of the CM and PM fractions clearly showed an increase in signal corresponding to AHA proteins in the putative PM-enriched fraction while the abundance of SMT-1 decreased. The presence of the SMT-1 marker in the soluble fraction was unexpected, but we speculated that perhaps the relatively short duration (30 min) of ultracentrifugation was not sufficient to pellet all of the ER fraction. Thus, we investigated if increasing the duration of ultracentrifugation affected the representation of SMT-1 in the microsomal pellet. Indeed, we found that increasing the duration of ultracentrifugation to longer than 30 min increased the proportion of SMT-1 relative to AHA in the microsomal fraction (**Figure 2B**). Therefore, the 30 min ultracentrifugation was used in all subsequent experiments to decrease ER contamination, thereby increasing the proportion of PM proteins in the sample. These results show that the PM enrichment strategy using Brij-58 detergent contained the highest levels of AHA1 and lowest levels of SMT-1, indicating that the PM fraction was enriched for PM proteins while depleting proteins from the ER. Therefore, the simplified PM enrichment protocol previously described for use in *Arabidopsis* suspension cell cultures (Zhang and Peck, 2011) also is applicable to studies in unrelated tissues such as the maize primary root.

Equal amounts of the PM-enriched protein fractions from the four contiguous regions in the root growth zone were loaded and separated by 1D SDS-PAGE. After separation of samples from a biological replicate, gel slices were excised and digested with trypsin, and the resultant peptide fractions were analyzed by LS-MS/MS analysis for comparisons of PM-enriched proteomes in the different regions of the growth zone (**Figure 1B**; summary of all raw and processed LC-MS/MS data from each biological replicate is found in Table S1 in Supplementary Material).

PROTEIN DISTRIBUTION PATTERNS

For analysis of quantitative protein distributions, proteins were considered to be reproducibly identified if they were found in two of the three biological replicates. Comparisons of this subset of reproducible proteins showed that there was a 72% (498/686) overlap between proteins identified from the four contiguous root regions (**Figure 2B**). R1, which contained the meristem and zone of accelerating elongation rate, had the most unique proteins compared to the other regions (36 in R1 versus 1 in R2, 0 in R3, and 6 in R4). Contiguous regions (R1–3 or R2–4) were more similar than non-contiguous regions (R1 and R4), indicating that the proteomes reflect the developmental gradient of root elongation.

We next analyzed the distribution of proteins in terms of relative abundance throughout the different regions. Of the 686 proteins, 83% (574) did not show any specific pattern of distribution, whereas 6% (43) showed region-distributed patterns (**Table 1**) and 11% (74) were present in only one or two regions (**Table 2**). The protein distribution patterns between the four different regions were grouped into several categories (**Table 1**). For example, when

Table 1 | Proteins present in all regions with distinct distribution patterns in the primary root growth zone.

Protein name	Accession number	Mol. weight (kDa)	Mean NSC				Standard error				Pattern category	R1 R2 R3 R4
			R1	R2	R3	R4	R1	R2	R3	R4		
Cellulose synthase-2	gi|162459760	121	13.32	45.07	40.11	35.34	3.76	10.55	1.86	10.49	R1-lowest	
Cellulose synthase-4	gi|162461169	121	5.97	34.07	25.09	26.85	1.15	11.10	3.33	7.48		
Cellulose synthase-8	gi|162460924	123	4.35	46.87	41.68	36.66	1.51	11.48	4.69	11.88		
Fasciclin-like Arabinogalactan protein 7 precursor	gi|195612412	27	3.63	8.83	8.32	8.32	0.62	0.61	1.10	1.43		
Brittle stalk-2-like protein 3	gi|119034638	50	3.96	9.65	11.63	10.17	0.94	1.48	1.86	1.57		
Glycoside transferase, six-hairpin, subgroup	gi|195614916	69	3.33	11.82	14.36	13.59	0.70	3.81	1.80	4.32		
Transmembrane 9 superfamily protein member 2 precursor	gi|195612428	74	16.67	19.34	16.67	14.34	2.73	1.33	4.84	1.86	R2-highest	
Unknown	gi|194689992	72	11.19	15.45	11.09	8.98	2.82	2.95	3.64	1.38		
Unknown	gi|223948977	89	35.71	44.98	26.07	12.45	9.20	14.07	5.96	2.14		
Unknown	gi|223975481	72	3.67	5	4.34	1.05	0.67	0.58	0.88	0.12		
Unknown	gi|219884637	48	20.48	35.8	23.44	19.14	6.02	7.67	5.58	5.68		
Ubiquitin fusion protein	gi|902527	18	12.67	20.67	28.67	26	4.06	7.54	7.32	9.30	R3-highest	
Unknown	gi|194702312	54	43.36	51.06	59.13	54.49	4.64	3.66	2.25	8.67		
Unnamed protein product	gi|219978130	81	11.85	27.48	29.7	23.66	3.85	9.57	8.03	7.25		
AIR12	gi|195654711	27	13.08	6.08	7.69	2.99	4.68	0.32	1.26	0.46	Down	
ATPase subunit 1	gi|102567957	55	197.31	136.131	105.38	92.93	9.14	19.84	4.39	7.07		
Cytochrome P450 CYP72A123	gi|195615656	60	9.87	6.08	2.69	2.26	1.93	0.84	0.38	0.79		
Dolichyl-diphosphooligosaccharide	gi|195648146	70	9.29	6.72	7.95	6.64	0.79	0.58	2.17	0.49		
IQ calmodulin-binding motif family protein	gi|195627120	46	83.63	47.71	38.45	30.95	2.59	1.37	11.15	7.71		
Mitochondrial chaperonin 60	gi|162460375	61	32.59	14.97	12.83	11.44	7.90	3.61	4.26	3.74		
Hypothetical protein	gi|195649497	72	64.36	37.7	35.42	24.72	5.17	6.25	6.38	6.20		
Unknown	gi|219888281	54	39.38	28.96	26.51	25.26	2.97	4.96	3.92	4.64		
Unknown	gi|194702738	47	62.6	52.5	35.7	35.01	6.97	10.70	8.97	8.15		
Unknown	gi|238006974	48	18.55	17.24	14.58	12.36	1.47	4.28	3.21	3.19		
Unknown	gi|194706574	44	13.19	4.94	5.69	4.74	1.43	1.56	0.44	1.35		
Unknown	gi|223947901	76	18.77	13.08	15.16	10.02	5.33	4.56	3.82	2.18		
Unnamed protein product	gi|219902525	61	33	16.34	13.34	11.34	7.64	3.76	4.26	3.34		
Unnamed protein product	gi|158058388	43	16.5	12.36	10.06	6.74	2.12	1.65	1.17	2.21		
Unnamed protein product	gi|219749304	115	18.23	12.07	12.34	10.41	3.15	4.33	1.75	2.92		

(Continued)

Table 1 | Continued

Protein name	Accession number	Mol. weight (kDa)	Mean NSC				Standard error				Pattern category	R1	R2	R3	R4
			R1	R2	R3	R4	R1	R2	R3	R4					
Actin-1	gi|195639446	42	33.48	45.11	51.95	56.78	1.33	2.81	6.63	7.04	Up				
Lipoxygenase	gi|162461114	98	18	18.67	15.67	26.67	4.94	5.37	1.77	8.46					
Lipoxygenase	gi|162462799	98	11.34	10.67	9.34	17	1.77	3.18	0.33	5.77					
Ribosomal proteinS6	gi|162463315	29	18.14	25.99	28.79	43.48	2.05	3.03	7.68	9.17					
Unknown	gi|219888401	105	22.7	92.651	25.55	161.71	3.67	10.59	14.90	44.88					
Unknown	gi|224029983	78	20.68	39.31	44.17	52.94	4.71	1.85	3.42	18.97					
Unnamed protein product	gi|219748040	105	30.97	72.73	87.33	114.21	5.05	4.09	6.62	34.54					
Unnamed protein product	gi|147223487	32	21	29.02	30.91	41.42	4.13	7.34	5.73	8.57					
Vacuolar ATP synthase catalytic subunit A	gi|195658441	68	32.5	38.38	51.98	55.91	3.69	6.39	9.95	13.16					

Categories of patterns were assigned based on the means of normalized spectral counts (NSC). Means and standard errors were calculated using the NSC from three biological replicates for each region.

a protein showed lowest expression in R1 and similar levels in R2–4, it was grouped in the category "R1-lowest." In contrast, if a protein showed highest expression in R1 and decreasing abundance thereafter, it was categorized into the group "down" (see **Table 1** for details). Because of the incomplete annotation of the maize genome, a large number of the proteins with defined distribution patterns are listed as "unknown." However, for some of the annotated proteins, the distribution pattern appears to be consistent with possible biological functions in the growing root, and a few of these examples are discussed below.

REGION-SPECIFIC CHANGES IN AQUAPORIN PROFILES – ROLE IN WATER MOVEMENT

The movement of water in plant tissues occurs by both an apoplastic pathway and a symplastic pathway (Steudle, 2000). The PM intrinsic proteins (PIPs), or aquaporins, are involved in symplastic movement of water across cellular membranes (Maurel and Chrispeels, 2001). Their relative contribution in water movement increases as tissues mature and develop apoplastic barriers such as suberized/lignified cell wall deposits, which reduce apoplastic water movement (Hachez et al., 2006). The results from the present study show a developmental gradient in the abundance of numerous aquaporin proteins (**Figure 3A**). The protein level of ZmPIP2-2 was lowest in R1 and increased thereafter, reaching a maximum in R4. ZmPIP2-3, 4 and ZmMIP proteins were not detected in R1, whereas there was a progressive increase in their abundance from R2 to R4. The increased aquaporin protein abundance in R4, which represented the maturation region beyond the growth zone (including suberization/lignification of the endodermis and development of the casparian strip), is likely involved in increased symplastic water transport. The abundance pattern of aquaporins observed in this study is similar to the aquaporin abundance patterns observed previously using immunoblot analyses in maize primary roots by Hachez et al. (2006). Therefore, our quantitative PM proteomic comparison is consistent with known spatial gradients of PM proteins, supporting the validity of this strategy.

SPATIAL DISTRIBUTION OF AIR12 PROTEINS – POTENTIAL ROLE IN REDOX SIGNALING

Two apparent maize orthologs of *Arabidopsis thaliana* AIR12 (for auxin induced in root cultures) were identified in all four regions of the root (**Figure 3B**). One of the ZmAIR12 proteins (gi 195608915) was ubiquitously distributed across all four regions, whereas another ZmAIR12 (gi 195654711) was more abundant in R1 compared to R2–4. An ortholog of AtAIR12 in soybean was recently identified as the major PM-localized b-type cytochrome that is fully reduced by ascorbate and fully oxidized by monodehydroascorbate radicals (Preger et al., 2009). The AtAIR12 and GmAIR12 proteins were found to be highly glycosylated and contained a glycosylphosphatidylinositol-anchor that positioned the proteins on the external side of the PM *in vivo* (Borner et al., 2003; Preger et al., 2005, 2009). AIR12 is physically associated with other redox signaling proteins and is suggested to be a link between the apoplast and cytoplasm in redox signaling (Lefebvre et al., 2007; Preger et al., 2009). Therefore, higher levels of ZmAIR12 (gi 195654711) in R1 of the maize primary root may

Table 2 | Proteins only present in specific regions in the primary root growth zone.

Pattern	Identified proteins	Accession number	Molecular weight	R1	R2	R3	R4
Single region present	50S ribosomal protein L1	gi|195625038	44 kDa	+			
	Nucleolar protein 10	gi|195655253	80 kDa	+			
	Phosphoribosylanthranilate transferase	gi|195652617	89 kDa	+			
	Phytochrome A2	gi|129560387	125 kDa	+			
	ATPase family AAA domain-containing protein 1	gi|195637516	41 kDa	+			
	GDP-mannose 4,6 dehydratase 2	gi|195624368	42 kDa	+			
	Guanine nucleotide-binding protein-like 3	gi|195656991	67 kDa	+			
	Hypothetical protein	gi|195647050	69 kDa	+			
	Pescadillo	gi|223945377	71 kDa	+			
	Unknown	gi|194688842	51 kDa	+			
	Unknown	gi|224031377	73 kDa	+			
	Unknown	gi|224030191	48 kDa	+			
	Unknown	gi|238014338	33 kDa	+			
	Unknown	gi|238014320	29 kDa	+			
	Unknown	gi|194689536	80 kDa	+			
	Unknown	gi|219885489	74 kDa	+			
	Unknown	gi|223949449	61 kDa	+			
	Unknown	gi|194690694	59 kDa	+			
	Unknown	gi|194694732	48 kDa	+			
	Unknown	gi|194689188	40 kDa	+			
	Plasma membrane intrinsic protein	gi|162460423	30 kDa				+
	Unknown	gi|226505982	66 kDa				+
	Unknown	gi|194703914	42 kDa				+
	Unnamed protein product	gi|219766771	71 kDa				+
Single region absent	Tubulin beta-1 chai n	gi|135449	50 kDa				–
	Squalene synthase 1	gi|162463458	46 kDa				–
	Hypothetical protein	gi|195612288	60 kDa				–
	Unknown	gi|238011696	60 kDa				–
	Unknown	gi|194707602	41 kDa				–
	5-Methyltetrahydropteroyltriglutamate	gi|195657041	84 kDa	–			
	Beta3-glucuronyl transferase	gi|195612932	37 kDa	–			
	Glycoside transferase, six-hairpin, subgroup	gi|195616232	69 kDa	–			
	Glycosyl transferase	gi|195645114	52 kDa	–			
	HGA4	gi|195625278	58 kDa	–			
	Patellin-5	gi|195624944	58 kDa	–			
	Plasma membrane integral protein ZmPIP2-3	gi|13447805	30 kDa	–			
	Plasma membrane integral protein ZmPIP2-4	gi|13447807	30 kDa	–			
	Plasma membrane MIP protein	gi|162458078	31 kDa	–			
	Proline oxidase	gi|195612286	52 kDa	–			
	Putative auxin efflux carrier	gi|111378666	65 kDa	–			
	Transmembrane protein	gi|162458279	31 kDa	–			
	Transmembrane 9 super family protein member 4	gi|195647402	74 kDa	–			
	Uclacyanin-2 precursor	gi|195610584	20 kDa	–			
	Unknown	gi|194697168	31 kDa	–			
	Unknown	gi|194690676	68 kDa	–			
	Unknown	gi|194690580	25 kDa	–			
	Unknown	gi|223947829	62 kDa	–			
	Unknown	gi|194703616	30 kDa	–			
	Unknown	gi|223949153	35 kDa	–			
	Unknown	gi|219884351	70 kDa	–			
	Unknown	gi|194688782	56 kDa	–			

(Continued)

Table 2 | Continued

Pattern	Identified proteins	Accession number	Molecular weight	R1	R2	R3	R4
	Unknown	gi|224033221	40 kDa	–			
	Unknown	gi|223942749	44 kDa	–			
	Unknown	gi|194690662	43 kDa	–			
	Unknown	gi|194690360	58 kDa	–			
	Unknown	gi|224033649	66 kDa	–			
	Unknown	gi|194689110	50 kDa	–			
	Unknown	gi|219885505	53 kDa	–			
	Unnamed protein product	gi|219900697	79 kDa	–			
	unnamed protein product	gi|218405930	64 kDa	–			
Two-region presence	Ankyrin protein kinase-like	gi|195654649	68 kDa	+	+		
	ATP binding protein	gi|195622982	56 kDa	+	+		
	Unknown	gi|223945493	115 kDa	+	+		
	Unknown	gi|219885707	82 kDa	+	+		
	Unknown	gi|194703992	62 kDa	+	+		
	Unknown	gi|194702080	41 kDa	+	+		
	Unknown	gi|223949925	52 kDa	+	+		
	Unnamed protein product	gi|29892201	21 kDa	+	+		
	Unnamed protein product	gi|219731540	91 kDa	+	+		
	Unknown	gi|194689072	49 kDa		+	+	
	Phospho-2-dehydro-3-deoxyheptonate aldolase 1	gi|195642466	59 kDa	+			+
	Unknown	gi|224030621	66 kDa	+			+
	Unknown	gi|194704320	57 kDa			+	+
	Unknown	gi|194706396	45 kDa			+	+

Region-specific patterns were assigned based on the means of normalized spectral count (NSC). Means and standard errors were calculated using the NSC from three biological replicates for each region.

indicate increased redox signaling, which would be consistent with increased apoplastic superoxide production in this region that has been suggested to play a role in cell wall loosening activities (Liszkay et al., 2004). In addition, increased superoxide production in the apical region of primary roots of *Arabidopsis* was suggested to be involved in regulating meristem size (Tsukagoshi et al., 2010). If ZmAIR12 activity is involved in regulating cell expansion, it will be interesting to understand how this process is regulated as one member of the protein family is uniformly distributed in all regions whereas another shows region-distributed accumulation.

SPATIAL DISTRIBUTION OF CELL WALL BIOSYNTHESIS RELATED PROTEINS

Cellulose synthases

Cellulose, the main constituent of plant cell walls, is composed of parallel unbranched glucan chains referred to as cellulose microfibrils (Anderson et al., 2010). These microfibrils are synthesized by large multimeric complexes of cellulose synthase proteins at the PM (Somerville, 2006). As cells expand, there is increased deposition of cellulose synthase complexes at the PM, and these complexes in turn deposit cellulose necessary for maintaining rigidity and directional growth of the cells (Baskin, 2005). Several cellulose synthases were identified in the PM-enriched fraction from the different regions of the maize primary root (**Table 1**). The abundance of cellulose synthases was lowest in R1, increased in R2, and remained high in R3–4 (**Figure 4A**). Because the integrated expansion of the cells in the maize primary root (as determined by the area under the relative elongation rate curve) is maximal in R2–3, the increased abundance of cellulose synthases in these regions is consistent with increased cell wall deposition in the elongating cells. As R4 corresponds to the region of maturation and secondary cell wall synthesis (**Figure 1A**), an increased abundance of cellulose synthases in this region is consistent with increased cell wall maturation (**Figure 4A**).

Glycosyltransferases

Several glycosyltransferases were differentially distributed throughout the regions of the root (**Figure 4B**). Glycosyltransferases are enzymes that transfer a sugar moiety from activated donor to acceptor molecules forming glycosidic bonds (Zhong and Ye, 2003; Lim and Bowles, 2004). Plant glycosyltransferases are involved in biosynthesis of several cell wall components such as polysaccharides, hemicelluloses, and pectins, and these enzymes may also be involved in transferring sugar molecules to proteins, hormones, and secondary metabolites (Zhong and Ye, 2003; Liepman et al., 2010). Therefore, the increased abundance of various glycosyltransferases particularly in the regions of maximal cell expansion (R2–3; **Figure 1A**) and secondary cell wall formation (R4) indicates that the glycosyltransferases identified in this study may be involved in cell wall synthesis (**Figure 4B**). Although some studies have suggested that glycosyltransferases are associated

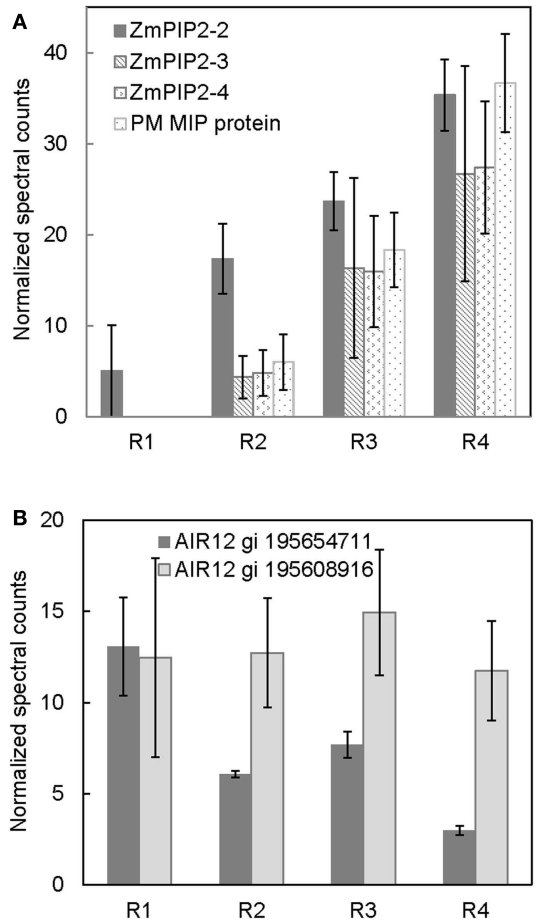


FIGURE 3 | Spatial distribution of (A) aquaporin protein abundances, and (B) ZmAIR12 protein abundances in R1–4 of the maize primary root harvested 48 h after transplanting to well-watered conditions. Data are means ± SE of normalized spectral counts from three replicate experiments.

with endomembranes (Keegstra, 2010), the localization of cellulose synthases, a major family of glycosyltransferases, at the PM suggests that glycosyltransferases can be associated with the PM (Scheible and Pauly, 2004; Somerville, 2006). Additionally, a recent study of PM proteomics in poplar in which the PM fraction was obtained from two-phase partitioning of microsomal membranes identified several glycosyltransferases (Nilsson et al., 2010). Therefore, it is likely that different classes of glycosyltransferases are localized to different compartments dependent on their function(s).

Brittlestalk-2-like protein 3 and fasciclin-like arabinogalactan proteins

The orientation of microfibrils is known to regulate cell shape. In isotropic expansion, the microfibrils are oriented in all directions, whereas in anisotropic expansion they are oriented perpendicular to the direction of expansion (Baskin, 2005). Although expansion is anisotropic in the root growth zone, the degree of anisotropy increases as the cells move away from the root apex (Liang et al., 1997; Baskin et al., 2004). These changes potentially involve changes in cellulose microfibril orientation and patterning. Several studies have shown that the cellulose microfibril orientation can be modified by PM-localized proteins such as COBRA and fasciclin-like arabinogalactan proteins in *Arabidopsis* and an ortholog of COBRA in maize, brittlestalk-2 (Roudier et al., 2005; Ching et al., 2006; MacMillan et al., 2010).

Brittlestalk-2 (*bk2*) was identified in a screen for maize mutants defective in cellulose biosynthesis; the mutant stalks had reduced mechanical strength (Ching et al., 2006). Analysis of the brittle phenotype found that disruption in the *bk2* gene interferes with the pattern of cellulose microfibril deposition, which leads to reduced cellulose and increased lignin in the *bk2* mutant shoot tissues (Ching et al., 2006; Sindhu et al., 2007). In *Arabidopsis*, the ortholog of *bk2*, *COBRA*, encodes a GPI-anchored PM-localized protein that is suggested to regulate cellulose microfibril deposition and anisotropic expansion in both primary root and hypocotyl tissues (Roudier et al., 2005). The protein abundance pattern of brittlestalk-2-like protein (**Figure 4C**) is similar to the mRNA expression pattern of *COBRA* in the primary root (Roudier et al., 2005), suggesting that brittlestalk-2-like protein may be involved in cellulose microfibril patterning and anisotropic expansion in the maize primary root.

Similarly, fasciclin-like arabinogalactan proteins are suggested to be involved in cellulose microfibril orientation and secondary cell wall synthesis in *Arabidopsis* and *Eucalyptus* (MacMillan et al., 2010). These proteins belong to a large protein family containing the cell adhesion fasciclin domain, which is also conserved in cell adhesion proteins in bacteria, algae, fungi, and animals (Johnson et al., 2003). In plants, the fasciclin-like arabinogalactan proteins are suggested to be PM-localized either with or without a GPI-anchor (Borner et al., 2003; Johnson et al., 2003; Lefebvre et al., 2007). However, the functions of the various fasciclin-like arabinogalactan proteins in plants are still largely unknown. Recent analysis revealed that mutations in two stem-specific fasciclin-like-arabinogalactan proteins lead to altered cellulose deposition and reduced stem stiffness in *Arabidopsis* stems (MacMillan et al., 2010). Thus, the increased accumulation of fasciclin-like-arabinogalactan protein in R2–4 (**Figure 4D**) hints at a role for this protein in modification of cellulose microfibril deposition.

CONCLUSION

The results in the present study demonstrate that the simplified PM enrichment method previously shown to work in *Arabidopsis* also can be used to characterize the developmental distribution of PM proteins in the maize primary root. Therefore, it appears that this technical strategy may be more broadly applicable to PM protein studies in diverse plant species. In addition, we found that shorter durations of ultracentrifugation decrease the representation of the ER marker, SMT-1, in the microsomal fraction. We hypothesize that this difference is caused because of the differences in buoyant densities of PM vs. ER-containing vesicles in maize root extracts. However, it should be noted that this method does not replace more stringent, but challenging, methods such as two-phase partitioning if the goal of the study is to more conclusively demonstrate that a protein specifically localizes to the PM.

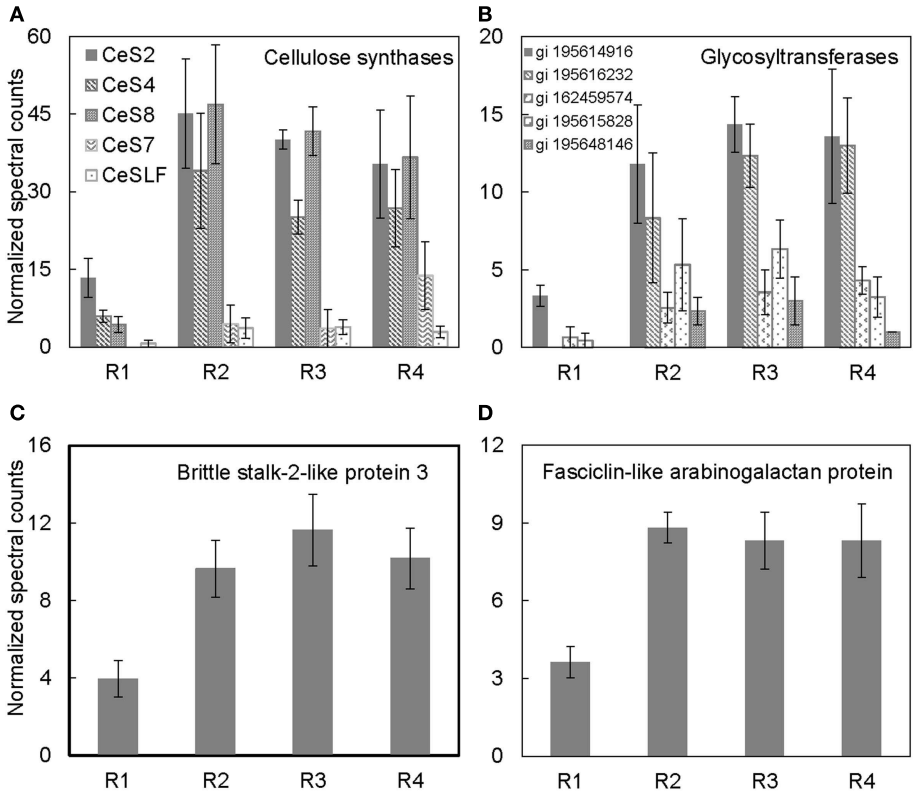


FIGURE 4 | Spatial distribution patterns of protein abundances for (A) cellulose synthases, (B) glycosyltransferases, (C) brittle stalk-2-like protein 3, and (D) fasciclin-like arabinogalactan proteins in R1–4 of the maize primary root harvested 48 h after transplanting to well-watered conditions. Data are means ± SE of normalized spectral counts from three replicate experiments.

Rather, the present method serves as a simple and robust method to increase the representation of PM proteins in proteomic studies.

Of the proteins showing defined developmental distribution patterns in the primary root of maize, a number of them, such as the aquaporins and cellulose synthases, appear logical in relation to their roles in cell expansion. Therefore, other proteins of unknown function such as the glycosyltransferases and the fasciclin-like arabinogalactans serve as candidates for involvement in growth regulation. Of course, further characterization studies are needed to determine the functional significance of their differential distribution in the growth zone of the maize primary root.

ACKNOWLEDGMENTS

We acknowledge Dr. Jeffrey Anderson for his useful suggestions and discussions about this work.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at http://www.frontiersin.org/Plant_Proteomics/10.3389/fpls.2013.00033/abstract

Table S1 | Complete list of all raw and normalized spectral counts for proteins identified in this study.

REFERENCES

Albertsson, P. A., Cajarville, A., Brooks, D. E., and Tjerneld, F. (1987). Partition of proteins in aqueous polymer two-phase systems and the effect of molecular weight of the polymer. *Biochim. Biophys. Acta* 926, 87–93.

Anderson, C. T., Carroll, A., Akhmetova, L., and Somerville, C. (2010). Real-time imaging of cellulose reorientation during cell wall expansion in *Arabidopsis* roots. *Plant Physiol.* 152, 787–796.

Baskin, T. I. (2005). Anisotropic expansion of the plant cell wall. *Annu. Rev. Cell Dev. Biol.* 21, 203–222.

Baskin, T. I., Beemster, G. T. S, Judy-March, J. E., and Marga, F. (2004). Disorganization of cortical microtubules stimulates tangential expansion and reduces the uniformity of cellulose microfibril alignment among cells in the root of *Arabidopsis thaliana*. *Plant Physiol.* 135, 2279–2290.

Beemster, G. T., and Baskin, T. I. (1998). Analysis of cell division and elongation underlying the developmental acceleration of root growth in *Arabidopsis thaliana*. *Plant Physiol.* 116, 1515–1526.

Bernstein, N., Lauchli, A., and Silk, W. K. (1993). Kinematics and dynamics of sorghum (*Sorghum bicolor* L.) leaf development at various Na/Ca salinities. *Plant Physiol.* 103, 1107–1114.

Borner, G. H. H., Lilley, K. S., Stevens, T. J., and Dupree, P. (2003). Identification of glycosylphosphatidylinositol-anchored proteins in *Arabidopsis*. A proteomic and genomic analysis. *Plant Physiol.* 132, 568–577.

Brady, S. M., David, A., Orlando, D. A., Lee, J., Wang, J. Y., Koch, J., et al. (2005). A high-resolution root spatiotemporal map reveals dominant expression patterns. *Science* 318, 801–806.

Ching, A., Dhugga, K. S., Appenzeller, L., Meeley, R., Bourett, T. M., Howard, R. J., et al. (2006). Brittle stalk2 encodes a putative glycosylphosphatidylinositol-anchored protein that affects mechanical strength of maize tissues by altering the composition and structure of secondary cell walls. *Planta* 224, 1174–1184.

Cosgrove, D. J. (1989). Characterization of long-term extension of isolated cell walls from growing cucumber hypocotyls. *Planta* 177, 121–130.

Endler, A., and Persson, S. (2011). Cellulose synthases and synthesis in *Arabidopsis*. *Mol. Plant* 4, 199–211.

Ephritikhine, G., Ferro, M., and Rolland, N. (2004). Plant membrane proteomics. *Plant Physiol. Biochem.* 42, 943–962.

Erickson, R. O. (1961). Probability of division of cells in the epidermis of the *Phleum* root. *Am. J. Bot.* 48, 268–274.

Erickson, R. O., and Silk, W. K. (1979). The kinematics of plant growth. *Sci. Am.* 242, 134–151.

Hachez, C., Moshelion, M., Zelazny, E., Cavez, D., and Chaumont, F. (2006). Localization and quantification of plasma membrane aquaporin expression in maize primary root: a clue to understanding their role as cellular plumbers. *Plant Mol. Biol.* 62, 305–323.

Johnson, K. L., Jones, B. J., Bacic, A., and Schultz, C. J. (2003). The fasciclin-like arabinogalactan proteins of *Arabidopsis*. A multigene family of putative cell adhesion molecules. *Plant Physiol.* 133, 1911–1925.

Keegstra, K. (2010). Plant cell walls. *Plant Physiol.* 154, 483–486.

Keller, A., Nesvizhskii, A. I., Kolker, E., and Aebersold, R. (2002). Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal. Chem.* 74, 5383–5392.

Kiegle, E., Gilliham, M., Haseloff, J., and Tester, M. (2000). Hyperpolarization-activated calcium currents found only in cells from the elongation zone of *Arabidopsis thaliana* roots. *Plant J.* 21, 225–229.

Komatsu, S. (2008). Plasma membrane proteome in *Arabidopsis* and rice. *Proteomics* 8, 4137–4145.

Lefebvre, B., Furt, F., Hartmann, M. A., Michaelson, L. V., Carde, J. P., Sargueil-Boiron, F., et al. (2007). Characterization of lipid rafts from *Medicago truncatula* root plasma membranes: a proteomic study reveals the presence of a raft-associated redox system. *Plant Physiol.* 144, 402–418.

Liang, B. M., Sharp, R. E., and Baskin, T. I. (1997). Regulation of growth anisotropy in well-watered and water-stressed maize roots. I. Spatial distribution of longitudinal, radial and tangential expansion rates. *Plant Physiol.* 115, 101–111.

Liepman, A. H., Wightman, R., Geshi, N., Turner, S. R., and Scheller, H. V. (2010). *Arabidopsis*-a powerful model system for plant cell wall research. *Plant J.* 61, 1107–1121.

Lim, E. K., and Bowles, D. J. (2004). A class of plant glycosyltransferases involved in cellular homeostasis. *EMBO J.* 23, 2915–2922.

Liszkay, A., van der Zalm, E., and Schopfer, P. (2004). Production of reactive oxygen intermediates (O_2^-, H_2O_2, and OH) by maize roots and their role in wall loosening and elongation growth. *Plant Physiol.* 136, 3114–3123.

MacAdam, J. W., Sharp, R. E., and Nelson, C. J. (1992). Peroxidase activity in the leaf elongation zone of tall fescue. *Plant Physiol.* 99, 879–885.

MacMillan, C. P., Mansfield, S. D., Stachurski, Z. H., Evans, R., and Southerton, S. G. (2010). Fasciclin-like arabinogalactan proteins: specialization for stem biomechanics and cell wall architecture in *Arabidopsis* and *Eucalyptus*. *Plant J.* 62, 689–703.

Maurel, C., and Chrispeels, M. J. (2001). Aquaporins. A molecular entry into plant water relations. *Plant Physiol.* 125, 135–138.

Morel, J., Claverol, S., Mongrand, S., Furt, F., Fromentin, J., Bessoule, J. J., et al. (2006). Proteomics of plant detergent-resistant membranes. *Mol. Cell Proteomics* 5, 1396–1411.

Mulkey, T. J., and Evans, M. L. (1980). Geotropism in corn roots: evidence for its mediation by differential acid efflux. *Science* 212, 70–71.

Nesvizhskii, A. I., Keller, A., Kolker, E., and Aebersold, R. (2003). A statistical model for identifying proteins by tandem mass spectrometry. *Anal. Chem.* 75, 4646–4658.

Neuhoff, V., Stamm, R., and Eibl, H. (1985). Clear background and highly sensitive protein staining with Coomassie Blue dyes in polyacrylamide gels: a systematic analysis. *Electrophoresis* 6, 427–448.

Nilsson, R., Bernfur, K., Gustavsson, N., Bygdell, J., Wingsle, G., and Larsson, C. (2010). Proteomics of plasma membranes from poplar trees reveals tissue distribution of transporters, receptors, and proteins in cell wall formation. *Mol. Cell Proteomics* 9, 368–387.

Peters, W. S., and Bernstein, N. (1997). The determination of relative elemental growth rate profiles from segmental growth rates. *Plant Physiol.* 113, 1395–1404.

Pilet, P. E., Versel, J. M., and Mayor, G. (1983). Growth distribution and surface pH patterns along maize roots. *Planta* 158, 398–402.

Preger, V., Scagliarini, S., Pupillo, P., and Trost, P. (2005). Identification of an ascorbate-dependent cytochrome b of the tonoplast membrane sharing biochemical features with members of the cytochrome b561 family. *Planta* 220, 365–375.

Preger, V., Tango, N., Marchand, C., Lemaire, S. D., Carbonera, D., Valentin, M. D., et al. (2009). Auxin-responsive genes AIR12 code for a new family of plasma membrane b-type cytochromes specific to flowering plants. *Plant Physiol.* 150, 606–620.

Roudier, F., Fernandez, A. G., Fujita, M., Himmelspach, R., Borner, G. H. H., Schindelman, G., et al. (2005). COBRA, an *Arabidopsis* extracellular glycosyl-phosphatidyl inositol-anchored protein, specifically controls highly anisotropic expansion through its involvement in cellulose microfibril orientation. *Plant Cell* 17, 1749–1763.

Saab, I. N., Sharp, R. E., Pritchard, J., and Voetberg, G. S. (1990). Increased endogenous abscisic acid maintains primary root growth and inhibits shoot growth of maize seedlings at low water potentials. *Plant Physiol.* 93, 1329–1336.

Scheible, W. R., and Pauly, M. (2004). Glycosyltransferases and cell wall biosynthesis: novel players and insights. *Curr. Opin. Plant Biol.* 7, 285–295.

Sharp, R. E., Hsiao, T. C., and Silk, W. K. (1990). Growth of the maize primary root at low water potentials. II. Role of growth and deposition of hexose and potassium in osmotic adjustment. *Plant Physiol.* 93, 1337–1346.

Sharp, R. E., Poroyko, V., Hejlek, L. G., Spollen, W. G., Springer, G. K., and Bohnert, H. J., et al. (2004). Root growth maintenance during water deficits: physiology to functional genomics. *J. Exp. Bot.* 55, 2343–2351.

Silk, W. K., Lord, E. M., and Eckard, K. J. (1989). Growth patterns inferred from anatomical records. *Plant Physiol.* 90, 708–713.

Sindhu, A., Langewisch, T., Olek, A., Multani, D. S., McCann, M. C., Vermerris, W., et al. (2007). Maize Brittle stalk2 encodes a COBRA-like protein expressed in early organ development but required for tissue flexibility at maturity. *Plant Physiol.* 145, 1444–1459.

Somerville, C. (2006). Cellulose synthesis in higher plants. *Annu. Rev. Cell Dev. Biol.* 22, 53–78.

Steudle, E. (2000). Water uptake by roots: effects of water deficit. *J. Exp. Bot.* 51, 1531–1542.

Tsukagoshi, H., Busch, W., and Benfey, P. N. (2010). Transcriptional regulation of ROS controls transition from proliferation to differentiation in the root. *Cell* 143, 606–616.

Wightman, R., and Turner, S. (2010). Trafficking of the plant cellulose synthase complex. *Plant Physiol.* 153, 427–432.

Wolters, H., and Jurgens, G. (2009). Survival of the flexible: hormonal growth control and adaptation in plant development. *Nat. Rev. Genet.* 10, 305–317.

Yamaguchi, M., and Sharp, R. E. (2010). Complexity and coordination of root growth at low water potentials: recent advances from transcriptomic and proteomic analyses. *Plant Cell Environ.* 33, 590–603.

Yamaguchi, M., Valliyodan, B., Zhang, J., LeNoble, M. E., Yu, O., Rogers, E. E., et al. (2010). Regulation of growth response to water stress in the soybean primary root. I. Proteomic analysis reveals region-specific regulation of phenylpropanoid metabolism and control of free iron in the elongation zone. *Plant Cell Environ.* 33, 223–243.

Zhang, Z., and Peck, S. C. (2011). Simplified enrichment of plasma membrane proteins for proteomic analyses in *Arabidopsis thaliana*. *Proteomics* 11, 1780–1788.

Zhong, R., and Ye, Z. (2003). Unraveling the functions of glycosyltransferase family 47 in plants. *Trends Plant Sci.* 12, 565–568.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 31 October 2012; accepted: 10 February 2013; published online: 06 March 2013.

Citation: Zhang Z, Voothuluru P, Yamaguchi M, Sharp RE and Peck SC (2013) Developmental distribution of the plasma membrane-enriched proteome in the maize primary root growth zone. Front. Plant Sci. ***4****:33. doi: 10.3389/fpls.2013.00033*

This article was submitted to Frontiers in Plant Proteomics, a specialty of Frontiers in Plant Science.

Copyright © 2013 Zhang, Voothuluru, Yamaguchi, Sharp and Peck. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.

frontiers in
PLANT SCIENCE

MINI REVIEW ARTICLE
published: 22 February 2013
doi: 10.3389/fpls.2013.00027

Detergent-resistant plasma membrane proteome to elucidate microdomain functions in plant cells

Daisuke Takahashi[1], Yukio Kawamura[1,2] and Matsuo Uemura[1,2]*

[1] United Graduate School of Agricultural Sciences, Iwate University, Morioka, Japan
[2] Cryobiofrontier Research Center, Faculty of Agriculture, Iwate University, Morioka, Japan

Edited by:
Nicolas L. Taylor, The University of Western Australia, Australia

Reviewed by:
Stefanie Wienkoop, University of Vienna, Austria
Natalia V. Bykova, Memorial University of Newfoundland, Canada

****Correspondence:***
Matsuo Uemura, Cryobiofrontier Research Center, Faculty of Agriculture, Iwate University, 3-18-8 Ueda, Morioka, Iwate 020-8550, Japan.
e-mail: uemura@iwate-u.ac.jp

Although proteins and lipids have been assumed to be distributed homogeneously in the plasma membrane (PM), recent studies suggest that the PM is in fact non-uniform structure that includes a number of lateral domains enriched in specific components (i.e., sterols, sphingolipids, and some kind of proteins). These domains are called as microdomains and considered to be the platform of biochemical reaction center for various physiological processes. Microdomain is able to be extracted as detergent-resistant membrane (DRM) fractions, and DRM fractions isolated from some plant species have been used for proteome and other biochemical characterizations to understand microdomain functions. Profiling of sterol-dependent proteins using a putative microdomain-disrupting agent suggests specific lipid–protein interactions in the microdomain. Furthermore, DRM proteomes dynamically respond to biotic and abiotic stresses in some plant species. Taken together, these results suggest that DRM proteomic studies provide us important information to understand physiological functions of microdomains that are critical to prosecute plant's life cycle successfully in the aspect of development and stress responses.

Keywords: detergent-resistant plasma membrane (DRM), microdomain, lipid raft, proteome, biotic stress, abiotic stress

INTRODUCTION

The plasma membrane (PM) is a typical cellular membrane with selective permeability and surrounds all organelles and cellular substances. Therefore, the PM is thought to be the most important cellular membrane due to relationships to various important cellular processes including cell division, differentiation, and biotic/abiotic stress adaptation. The PM contains a variety of proteins associated with transport, signaling, cytoskeleton construction, metabolism, and stress protection in the form of transmembrane, peripheral, and lipid modified types.

Lateral distribution of these membrane proteins has been described by diffusion of each lipid and protein molecule which is proposed as fluid mosaic model (Singer and Nicolson, 1972). In addition to this hypothesis, Simons and Ikonen (1997) proposed functional microdomain of the PM. In PM microdomain hypothesis, it is considered that microdomain is organized with highly hydrophobic lipids such as sterols and sphingolipids, and specific proteins with defined functions (Brown and London, 1998, 2000; London and Brown, 2000). In animal cells, one of microdomain function is considered to be a scaffold in association with signaling complex, membrane trafficking, and transport (Simons and Ikonen, 1997; Simons and Toomre, 2000; Lingwood and Simons, 2010). Experimentally, microdomain can be obtained as non-ionic detergent-resistant membrane (DRM) fraction due to their own hydrophobic properties (Schroeder et al., 1994; Simons and Ikonen, 1997; Brown and London, 1998).

Peskan et al. (2000) reported for the first time isolation of DRM fractions from plant materials using tobacco leaves. After this report, the isolation of DRM fractions have been reported with other plant species such as tobacco, *Arabidopsis thaliana*, leek, *Medicago truncatula*, *Solanum tuberosum*, rice, oat, and rye (Mongrand et al., 2004; Borner et al., 2005; Morel et al., 2006; Laloi et al., 2007; Lefebvre et al., 2007; Krügel et al., 2008; Fujiwara et al., 2009; Minami et al., 2009; Takahashi et al., 2012). Some physiological studies showed possibilities that microdomain is involved in pollen tube tip growth, intracellular virus movement, and clathrin-independent endocytotic pathway (Liu et al., 2009; Raffaele et al., 2009; Li et al., 2012). In addition to these functions, PM microdomain may have roles in cell wall polysaccharide synthesis in hybrid aspen (Bessueille et al., 2009).

There have been attempts to identify microdomain-associated proteins for elucidation of novel microdomain-dependent regulatory mechanisms on cellular physiological processes in plant. Most of these studies were 2D or 1D electrophoresis gel-based proteomics or nano-LC-MS/MS-based shotgun proteomics using microdomain-enriched DRM fraction. In addition to DRM fraction, methyl-β-cyclodextrin (mβCD), which is known as a sterol chelator and, hence, a sterol-dependent microdomain disrupter, was used to characterize how protein was associated with the primary microdomain lipid, sterol (Kierszniowska et al., 2009). Comprehensive analyses of DRM proteomes may contribute to demonstrate the importance of lateral segregation of proteins in plant PM microdomains. Ultimately, these results may lead to new findings of plant cellular homeostasis system such as signaling machinery, transport regulation, and novel response system against perception of biotic stress such as fungal infection and abiotic stress such as drought, salt, light, nutrition, and temperature.

DETERGENT-RESISTANT MEMBRANE FRACTION AS A BIOCHEMICAL SAMPLE FOR OBTAINING INFORMATION ASSOCIATED WITH PLASMA MEMBRANE MICRODOMAIN

To analyze biochemical properties, the extraction of DRM fractions from the PM is considered to be the only way to prepare microdomain samples (**Figure 1**). DRM fractions were isolated from a number of plant species and tissues as described above, and the preparation protocols of DRM fractions are in general quite similar regardless of plant species. First, a highly pure PM fraction is prepared using a two-phase partition system and then treated with 1% (w/v) Triton X-100 detergent at low temperature (on ice or 4°C) for 30 min. Next, treated membrane fraction is subjected to sucrose density gradient centrifugation. After centrifugation, white band appeared at the interface of sucrose layers recovered and collected by centrifugation. Precipitated membrane fraction is suspended in a proper buffer as DRM fraction. Because unknown artificial effects might be caused due to detergent treatment at low temperature, some researchers concerned that intact microdomains that function *in vivo* are not extracted by the widely adapted preparation protocols (Tanner et al., 2011). Nevertheless, DRM fraction is a useful tool for estimating microdomain functions associated with specific components. Many microdomain-related phenomena have been elucidated in DRM and non-DRM fractions, and experiments with DRM fraction is apparently one of the most effective ways to determine specific functions in relation to microdomains in PM (Lingwood and Simons, 2007).

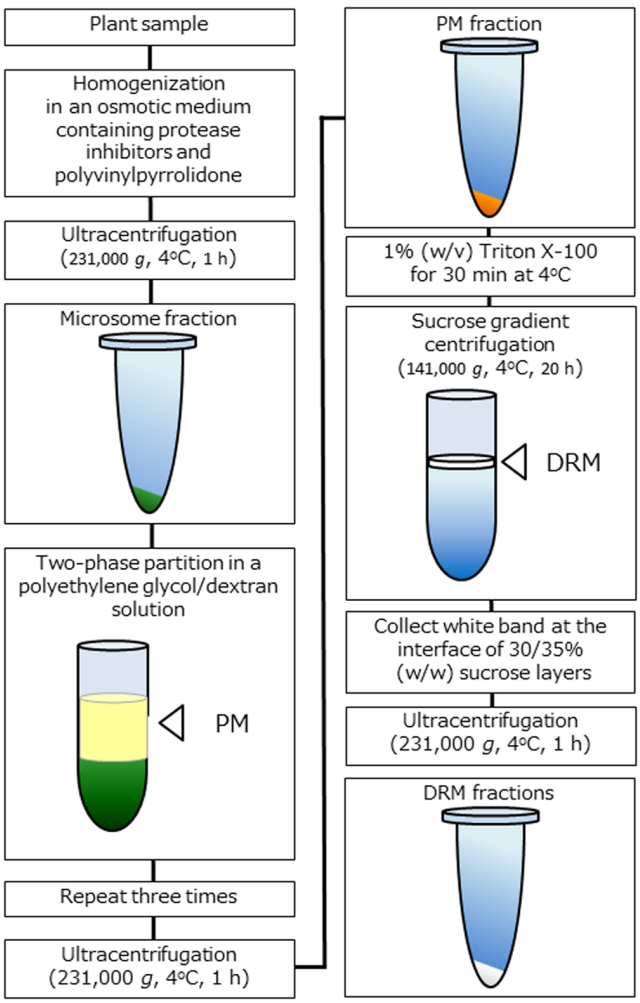


FIGURE 1 | Schematic representation of DRM extraction in plants. Overview of the DRM extraction procedure. DRM fractions are obtained from purified PM fractions by the 1% (w/v) Triton X-100 treatment and subsequent sucrose density gradient centrifugation.

Proteomics approaches of DRM proteins is well-conducted in various organisms and, further, quantification of DRM and non-DRM proteins is also reported in some plant species using isotope labeling, 2D difference gel electrophoresis (2D-DIGE), and label-free quantification software (Kierszniowska et al., 2009; Minami et al., 2009; Takahashi et al., 2012). However, there are still difficulties in quantitative determination of a large number of proteins correctively using proteomic approached. This is in part because solubilization of membrane proteins including those localized in the PM as well as DRM may not be consistent in a series of experiments due to hydrophobic characteristics of the proteins and assignment of peptide fragments to the appropriate protein may not be accurate in some species for which we have not yet completed genome sequencing. It is necessary to combine another approaches (such as immunochemical and biochemical approaches) to obtain the amount of proteins in the membrane accurately.

FUNCTION OF THE PM MICRODOMAIN

Detergent-resistant membrane proteomes have been determined in some plant species (Mongrand et al., 2004; Shahollari et al., 2004; Borner et al., 2005; Morel et al., 2006; Lefebvre et al., 2007; Fujiwara et al., 2009; Minami et al., 2009; Stanislas et al., 2009; Takahashi et al., 2012). Comparisons of DRM proteomes from these plant species indicated that DRM protein functions are very similar among plant species: DRM fractions contain many transporters, proteins associated with membrane vesicle trafficking processes and cytoskeleton such as H^+-ATPases, aquaporins, clathrins, actins, and tubulins. Further, microscopic observations and biochemical analyses of DRM fractions or intact plant cells implied that microdomains play some functional roles in the physiological aspects. **Table 1** summarizes proteins that were found in common in some plant species on papers published so far. Localization or function of some of these proteins in distinct regions in the PM was further confirmed by additional approaches either morphologically or biochemically.

As an example of functional involvement in developmental process, Liu et al. (2009) reported the involvement of microdomain in pollen tube tip growth. Using sterol-enriched microdomains in pollen tube using one of microdomain-staining lipophilic styryl dyes, di-4-ANEPPDHQ, they clearly revealed localization of NADPH oxidase in microdomain. From the results, they suggested that one of predicted microdomain properties (i.e., clustering of specific, hydrophobic lipids and proteins) is required for NADPH oxidase activity and polarization of sterol-enriched microdomain regulates NADPH oxidase-dependent reactive oxygen species signaling. Ultimately, polar growth of pollen tube tip

Table 1 | Proteins identified as DRM-enriched proteins in plants by proteomic and other studies

Proteins	Deduced function	Reference	Evidence for association with microdomain
P-type ATPase	Proton transport	*Arabidopsis thaliana* (Shahollari et al., 2004; Borner et al., 2005; Minami et al., 2009) Tobacco (Mongrand et al., 2004; Morel et al., 2006; Stanislas et al., 2009) *Medicago truncatula* (Lefebvre et al., 2007) Rice (Fujiwara et al., 2009) Oat and rye (Takahashi et al., 2012)	
Aquaporin	Water channel	*Arabidopsis thaliana* (Shahollari et al., 2004; Borner et al., 2005; Minami et al., 2009) Tobacco (Mongrand et al., 2004; Morel et al., 2006; Stanislas et al., 2009) *Medicago truncatula* (Lefebvre et al., 2007) Rice (Fujiwara et al., 2009) Oat and rye (Takahashi et al., 2012)	
Leucine-rich repeat receptor like kinase	Signaling receptor	*Arabidopsis thaliana* (Shahollari et al., 2004; Minami et al., 2009) Tobacco (Morel et al., 2006; Stanislas et al., 2009) *Medicago truncatula* (Lefebvre et al., 2007) Rice (Fujiwara et al., 2009) Oat and rye (Takahashi et al., 2012)	
Remorin	Unknown	*Arabidopsis thaliana* (Shahollari et al., 2004; Minami et al., 2009) Tobacco (Mongrand et al., 2004; Morel et al., 2006; Stanislas et al., 2009) *Medicago truncatula* (Lefebvre et al., 2007) Oat and Rye (Takahashi et al., 2012)	Involvement of microdomain and remorin in Potato virus X movement in tomato (Raffaele et al., 2009)
Tubulin	Cytoskeleton component	*Arabidopsis thaliana* (Minami et al., 2009) Tobacco (Mongrand et al., 2004; Morel et al., 2006; Stanislas et al., 2009) *Medicago truncatula* (Lefebvre et al., 2007) Rice (Fujiwara et al., 2009) Oat and rye (Takahashi et al., 2012)	
NADPH oxidase (Rboh)	Response to stress	Tobacco (Mongrand et al., 2004; Morel et al., 2006; Stanislas et al., 2009) Rice (Fujiwara et al., 2009) Oat and rye (Takahashi et al., 2012)	Regulations of pollen tube tip growth through microdomain-dependent NADPH oxidase polarization in *Picea meyeri* (Liu et al., 2009)
Hypersensitive-induced reaction protein (Band 7 family proteins, flotillins)	Response to stress	*Arabidopsis thaliana* (Shahollari et al., 2004; Borner et al., 2005; Minami et al., 2009) Tobacco (Morel et al., 2006; Stanislas et al., 2009) *Medicago truncatula* (Lefebvre et al., 2007) Rice (Fujiwara et al., 2009) Oat and rye (Takahashi et al., 2012)	Relationship with clathrin-independent endocytosis pathway in *Arabidopsis thaliana* (Li et al., 2012)
Glucan synthase	Cell wall metabolism	*Arabidopsis thaliana* (Minami et al., 2009) Tobacco (Morel et al., 2006; Stanislas et al., 2009) Rice (Fujiwara et al., 2009) Oat and rye (Takahashi et al., 2012)	Enzymatic regulations in DRM isolated from hybrid aspen cells (Bessueille et al., 2009)

may be modulated by the localization of proteins in microdomain. In addition to plant pollen tip, polarization of microdomain in hyphal tip of *Candida albicans* was also observed (Martin and Konopka, 2004). These data together suggest that characteristics of microdomain are common in the function on cell polarization among various species not only plants but also microorganisms.

A recent study also suggested that microdomain is related to intracellular membrane trafficking. *Arabidopsis* Flot1 is a DRM-associated protein that was identified in DRM proteome (Borner et al., 2005). Li et al. (2012) observed that Flot1 showed patch-like localization on PM using electron microscopic technique. They further showed that Flot 1 is participated in endocytic vesicle formation but gold-conjugated antibody of Flot 1 does not co-localize with clathrin light chain. It means that Flot1 plays some roles in a microdomain-associated but clathrin-independent endocytosis pathway. Considering that RNA interference of Flot1 results in the defect of seedling development, microdomain-Flot1 mediated vesicle trafficking has important implications for seedling development such as root hair elongation regulated by vesicle trafficking (Ovecka et al., 2010).

According to protein clustering in microdomain, proteomic and subsequent enzymatic characterizations of DRM fraction from hybrid aspen cells strongly suggested the involvement of DRM in cell wall polysaccharide synthesis (Bessueille et al., 2009). DRM from hybrid aspen was enriched in glucan synthases such as callose and cellulose synthase, and, surprisingly, 73% of total glucan synthase activities of PM were detected in DRM. They concluded that microdomain is functional platform for cell wall component synthesis and controls cell morphogenesis.

Detailed analysis of *M. truncatula* DRM showed considerable differences in DRM fraction and the total PM fraction (Lefebvre et al., 2007). This study showed that free sterols, sphingolipids, and steryl glycosides are highly enriched in DRM fractions. These results are consistent with previous studies with tobacco and *A. thaliana* (Mongrand et al., 2004; Borner et al., 2005). In addition to lipids, global survey of DRM proteins were performed and revealed that signaling-, transport-, redox-, cytoskeleton-, trafficking-, and cell wall-related proteins were enriched in DRM, most of which were also found in early works of plant DRM protein identification (Mongrand et al., 2004; Shahollari et al., 2004; Borner et al., 2005; Morel et al., 2006). Proteome profiling of *M. truncatula* DRM further indicated the possible presence of microdomain-dependent redox regulation system and microdomain platform for signaling.

As described above, sterols are one of the primary components of microdomain-enriched DRM fractions in both animal and plant cells. Kierszniowska et al. (2009) applied mβCD to isolated *Arabidopsis* DRM fractions to analyze sterol-dependent enrichment of DRM proteins. mβCD is a sterol-removing cyclic oligosaccharide and mβCD treatment disrupts the organization of membrane microdomain. Proteomic analysis of the mβCD-treated and untreated DRM fractions revealed that cell wall-related and glycosylphosphatidylinositol anchored proteins (a class of lipid-modified proteins) were changed by sterol depletion. Thus, these results strongly suggest that sterol is an important factor for segregation of specific proteins into DRM fraction and PM is "phase-separated" to form specific domains (i.e., sterol-enriched microdomains, Xu et al., 2001). As shown in these studies, proteome analysis has been used for estimating microdomain functions in plant cells for the past decades and, therefore, greatly contributed to elucidation of microdomain-associated physiological functions in plant cells.

DRM PROTEOME ON BIOTIC STRESS RESPONSE

Plant proteomic studies for elucidating microdomain function have been carried out intensively in the research area of plant–pathogen interactions. The possibility of lipid microdomain-pathogen interactions was first reported by Bhat et al. (2005). The authors suggest that fungal pathogen (*Blumeria graminis* f.sp. *hordei*) recognizes barley *mildew resistance locus o* (*Mlo*) that seems to re-localize as microdomain-like structure at pathogen invasion site. In addition to pathogen infection, plant immune responses against biotic stress may be supported by functional microdomain. Fujiwara et al. (2009) successfully identified 192 proteins from DRM proteome analysis in rice suspension cultured cells that were pre-transformed with constitutively active OsRac. OsRac1 is one of the Rac/Rop GTPase family proteins and regulates rice immunity as a key regulator (Kawasaki et al., 1999; Ono et al., 2001; Wong et al., 2004; Lieberherr et al., 2005; Kim et al., 2012). Shift of OsRac1 to DRM fractions was found after elicitor treatment. At the same time, DRM proteome suggests that microdomain exists as platform for rice innate immunity. Actually, *receptor-like kinases* (*RLK*), disease resistance proteins and band7 family proteins, members of disease-related proteins, were detected in rice DRM fractions as well as some other plant species (Borner et al., 2005; Morel et al., 2006; Fujiwara et al., 2009; Minami et al., 2009). Interactions between OsRac1 and those proteins may occur during initial immunity process against biotic stimuli.

Mongrand et al. (2004) also suggest from proteomics of DRM that microdomain isolated as DRM fractions has important functions in plant defense responses because tobacco DRM proteome contains a variety of defense-related proteins such as remorin, NtrbohD, and Ntrac5. Some physiological studies further indicated that DRM-enriched proteins are associated with plant–pathogen interactions. Remorin is the most characterized and a representative DRM protein and Raffaele et al. (2009) reported interesting results that remorin is associated with intercellular virus movement. Solanaceae remorin was fractionated into DRM fraction, which is also reported in tobacco DRM proteomics (Mongrand et al., 2004; Morel et al., 2006; Stanislas et al., 2009) as well as oat and rye proteomics (Takahashi et al., 2012). Interestingly, the distribution of remorin on the PM was represented as patch-like patterns and disappeared when mβCD was added to the sample. These results strongly suggest that DRM fractions partly reflect intact microdomain. Raffaele et al. (2009) also showed that remorin is localized in plasmodesmata and its accumulation levels affect cell-to-cell transfer of *Potato virus X* (PVX) through plasmodesmata. Detailed analysis of DRM proteome against elicitor signaling in tobacco BY-2 cells revealed that the DRM enrichment of cell trafficking related proteins (dynamins) and a signaling protein (14-3-3 protein) altered after cryptogein treatment (Stanislas et al., 2009). These studies clearly indicate that DRM proteomics has potential to find new factors of elicitor signaling pathway and

their functions in plants, and DRM proteome has methodological significance in approach for findings of novel microdomain functions on plant pathology.

DRM PROTEOME ON ABIOTIC STRESS RESPONSE

Abiotic stress response and adaptation mechanism in association with microdomain is not well characterized. The only study showing changes of DRM compositions in response to abiotic stimuli was with *Arabidopsis* leaves reported by Minami et al. (2009). They performed *Arabidopsis* DRM proteomic analysis to find the possibilities of microdomain functions for adaptation to freezing temperature. Plants can increase survival at severe freezing temperatures by sensing non-freezing low temperature and subsequently reconstituting cellular processes (called as cold acclimation; Guy, 1990; Sharma et al., 2005). Although there are a number of papers revealing considerable changes of PM compositions during cold acclimation (Uemura and Yoshida, 1984; Lynch and Steponkus, 1987; Webb et al., 1994; Kawamura and Uemura, 2003; Uemura et al., 2006), analysis of DRM compositions during cold acclimation was conducted in very few studies. Using a combination of 1D sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), 2D-DIGE and Liquid chromatography-tandem mass spectrometry (LC-MS/MS) and Matrix-assisted laser desorption/ionisation-time of flight mass spectrometry (MALDI-TOF/MS) analysis, Minami et al. (2009) demonstrated that proteomic profiles of DRM fractions altered significantly during cold acclimation. The cold acclimation-responsive proteins include synaptotagmin protein homolog, tubulin and P-type ATPase. Each protein is considered to have important roles in cold acclimation process from previous studies. For example, synaptotagmin homolog SYT1 was identified in DRM and increased after cold acclimation. SYT1 is related to calcium-dependent PM resealing (or repairing) process when PM destruction occurs due to freeze-induced mechanical stress imposed by extracellular ice formation (Yamazaki et al., 2008). In addition to SYT1, other membrane fusion-related proteins such as syntaxin were identified in oat and rye DRM (Takahashi et al., 2012). Disassembly of microtubule consisting of tubulins is suggested to be important for inducing cold acclimation process (Abdrakhamanova et al., 2003). Enhancement of ATPase activity in PM during cold acclimation is one of the well-known reactions in some plant species (Ishikawa and Yoshida, 1985; Martz et al., 2006).

How interactions between these proteins and microdomain properties affect cold acclimation processes, however, is still to be elucidated. We need to conduct additional physiological and microscopic experiments to understand responsiveness of microdomain and/or DRM proteins to cold acclimation. We have evidence from proteomic studies that there are several interesting abiotic stress-related proteins in DRM fractions such as RLKs, aquaporins, heat shock proteins, actins, and clathrins in various plants (Mongrand et al., 2004; Borner et al., 2005; Morel et al., 2006; Takahashi et al., 2012). To elucidate their contribution to abiotic stress sensing, signaling, and response, comprehensive proteomic analyses such as protein–protein interactions and post-translational modifications of the proteins would be necessary and expected.

FUTURE PERSPECTIVE

Proteomic analyses of DRM fractions have been conducted and provided information for suggestive but important functions of PM microdomain in plants. Several physiological studies using both intact cells and isolated membrane fractions supported implications derived from proteomic analyses with DRM and added further interesting information on the roles of membrane microdomains. However, evidence of functional roles of microdomains in the PM is in a large part lacking. Now we are entering in next phase for elucidating microdomain characteristics and functions in plants. We need to consider morphology and dynamics of microdomains, physical and chemical state of PM proteins in microdomains from the perspective of post-translational modifications and molecular ultrastructure, and ultimately functional significance of microdomains in various events in plant's life. Development of microscopic and biochemical techniques, such as single-molecule tracking and artificial membrane system, will help us to understand physiological roles of microdomain in plant cells.

Plants are immobile and, thus, perception and response to environmental stimuli are quite important for plant's life. The PM is thought to be the primary cellular compartment of these reactions because it surrounds intracellular organelles and the cytoplasm and transduces extracellular stimuli to the specific components in the cell. Microdomain is expected to play important roles in these processes. Thus, proteomic approaches will further provide useful information for understanding plant physiological responses and microdomain significance in the future.

ACKNOWLEDGMENTS

We thank Drs. Fukao and Masayuki Fujiwara (Nara Institute of Science and Technology) and Setsuko Komatsu (National Agriculture and Food Research Organization) for valuable discussions. This study was in part supported by Grants-in-Aids for Scientific Research from Ministry of Education, Culture, Sports, Science and Technology in Japan (#22120003) and from Japanese Society for the Promotion of Sciences (#24·7373 and #24370018).

REFERENCES

Abdrakhamanova, A., Wang, Q. Y., Khokhlova, L., and Nick, P. (2003). Is microtubule disassembly a trigger for cold acclimation? *Plant Cell Physiol.* 44, 676–686.

Bessueille, L., Sindt, N., Guichardant, M., Djerbi, S., Teeri, T. T., and Bulone, V. (2009). Plasma membrane microdomains from hybrid aspen cells are involved in cell wall polysaccharide biosynthesis. *Biochem. J.* 420, 93–103.

Bhat, R. A., Miklis, M., Schmelzer, E., Schulze-Lefert, P., and Panstruga, R. (2005). Recruitment and interaction dynamics of plant penetration resistance components in a plasma membrane microdomain. *Proc. Natl. Acad. Sci. U.S.A.* 102, 3135–3140.

Borner, G. H. H., Sherrier, D. J., Weimar, T., Michaelson, L. V., Hawkins, N. D., Macaskill, A., et al. (2005). Analysis of detergent-resistant membranes in *Arabidopsis*: evidence for plasma membrane lipid rafts. *Plant Physiol.* 137, 104–116.

Brown, D. A., and London, E. (1998). Function of lipid rafts in biological membranes. *Annu. Rev. Cell Dev. Biol.* 14, 111–136

Brown, D. A., and London, E. (2000). Structure and function of sphingolipid- and cholesterol-rich membrane rafts. *J. Biol. Chem.* 275, 17221–17224.

Fujiwara, M., Hamada, S., Hiratsuka, M., Fukao, Y., Kawasaki, T., and Shimamoto, K. (2009). Proteome analysis of detergent-resistant membranes (DRMs) associated with OsRac1-mediated innate immunity in rice. *Plant Cell Physiol.* 50, 1191–1200.

Guy, C. (1990). Cold acclimation and freezing stress tolerance: role of protein metabolism. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 41, 187–223.

Ishikawa, M., and Yoshida, S. (1985). Seasonal changes in plasma membranes and mitochondria isolated from Jerusalem artichoke tubers: possible relationship to cold hardiness. *Plant Cell Physiol.* 26, 1331–1344.

Kawamura, Y., and Uemura, M. (2003). Mass spectrometric approach for identifying putative plasma membrane proteins of *Arabidopsis* leaves associated with cold acclimation. *Plant J.* 36, 141–154.

Kawasaki, T., Henmi, K., Ono, E., Hatakeyama, S., Iwano, M., Satoh, H., et al. (1999). The small GTP-binding protein rac is a regulator of cell death in plants. *Proc. Natl. Acad. Sci. U.S.A.* 96, 10922–10926.

Kierszniowska, S., Seiwert, B., and Schulze, W. X. (2009). Definition of *Arabidopsis* sterol-rich membrane microdomains by differential treatment with methyl-beta-cyclodextrin and quantitative proteomics. *Mol. Cell. Proteomics* 8, 612–623.

Kim, S.-H., Oikawa, T., Kyozuka, J., Wong, H. L., Umemura, K., Kishi-Kaboshi, M., et al. (2012). The bHLH Rac immunity1 (RAI1) is activated by OsRac1 via OsMAPK3 and OsMAPK6 in rice immunity. *Plant Cell Physiol.* 53, 740–754.

Krügel, U., Veenhoff, L. M., Langbein, J., Wiederhold, E., Liesche, J., Friedrich, T., et al. (2008). Transport and sorting of the *Solanum tuberosum* sucrose transporter SUT1 is affected by post-translational modification. *Plant Cell* 20, 2497–2513.

Laloi, M., Perret, A.-M., Chatre, L., Melser, S., Cantrel, C., Vaultier, M.-N., et al. (2007). Insights into the role of specific lipids in the formation and delivery of lipid microdomains to the plasma membrane of plant cells. *Plant Physiol.* 143, 461–472.

Lefebvre, B., Furt, F., Hartmann, M.-A., Michaelson, L. V., Carde, J.-P., Sargueil-Boiron, F., et al. (2007). Characterization of lipid rafts from *Medicago truncatula* root plasma membranes: a proteomic study reveals the presence of a raft-associated redox system. *Plant Physiol.* 144, 402–418.

Li, R., Liu, P., Wan, Y., Chen, T., Wang, Q., Mettbach, U., et al. (2012). A membrane microdomain-associated protein, *Arabidopsis* Flot1, is involved in a clathrin-independent endocytic pathway and is required for seedling development. *Plant Cell* 24, 2105–2122.

Lieberherr, D., Thao, N. P., Nakashima, A., Umemura, K., and Kawasaki, T. (2005). A sphingolipid elicitor-inducible mitogen-activated protein kinase is regulated by the small GTPase OsRac1 and heterotrimeric G-protein in rice. *Plant Physiol.* 138, 1644–1652.

Lingwood, D., and Simons, K. (2007). Detergent resistance as a tool in membrane research. *Nat. Protoc.* 2, 2159–2165.

Lingwood, D., and Simons, K. (2010). Lipid rafts as a membrane-organizing principle. *Science* 327, 46–50.

Liu, P., Li, R.-L., Zhang, L., Wang, Q.-L., Niehaus, K., Baluska, F., et al. (2009). Lipid microdomain polarization is required for NADPH oxidase-dependent ROS signaling in *Picea meyeri* pollen tube tip growth. *Plant J.* 60, 303–313.

London, E., and Brown, D. A. (2000). Insolubility of lipids in Triton X-100: physical origin and relationship to sphingolipid/cholesterol membrane domains (rafts). *Biochim. Biophys. Acta* 1508, 182–195.

Lynch, D. V., and Steponkus, P. L. (1987). Plasma membrane lipid alterations associated with cold acclimation of winter rye seedlings (*Secale cereale* L. cv Puma). *Plant Physiol.* 83, 761–767.

Martin, S. W., and Konopka, J. B. (2004). Lipid raft polarization contributes to hyphal growth in *Candida albicans*. *Eukaryot. Cell* 3, 675–684.

Martz, F., Sutinen, M.-L., Kiviniemi, S., and Palta, J. P. (2006). Changes in freezing tolerance, plasma membrane H^+-ATPase activity and fatty acid composition in *Pinus resinosa* needles during cold acclimation and de-acclimation. *Tree Physiol.* 26, 783–790.

Minami, A., Fujiwara, M., Furuto, A., Fukao, Y., Yamashita, T., Kamo, M., et al. (2009). Alterations in detergent-resistant plasma membrane microdomains in *Arabidopsis thaliana* during cold acclimation. *Plant Cell Physiol.* 50, 341–359.

Mongrand, S., Morel, J., Laroche, J., Claverol, S., Carde, J.-P., Hartmann, M.-A., et al. (2004). Lipid rafts in higher plant cells: purification and characterization of Triton X-100-insoluble microdomains from tobacco plasma membrane. *J. Biol. Chem.* 279, 36277–36286.

Morel, J., Claverol, S., Mongrand, S., Furt, F., Fromentin, J., Bessoule, J.-J., et al. (2006). Proteomics of plant detergent-resistant membranes. *Mol. Cell. Proteomics* 5, 1396–1411.

Ono, E., Wong, H. L., Kawasaki, T., Hasegawa, M., Kodama, O., and Shimamoto, K. (2001). Essential role of the small GTPase Rac in disease resistance of rice. *Proc. Natl. Acad. Sci. U.S.A.* 98, 759–764.

Ovecka, M., Berson, T., Beck, M., Derksen, J., Samaj, J., Baluska, F., et al. (2010). Structural sterols are involved in both the initiation and tip growth of root hairs in *Arabidopsis thaliana*. *Plant Cell* 22, 2999–3019.

Peskan, T., Westermann, M., and Oelmüller, R. (2000). Identification of low-density Triton X-100-insoluble plasma membrane microdomains in higher plants. *Eur. J. Biochem.* 267, 6989–6995.

Raffaele, S., Bayer, E., Lafarge, D., Cluzet, S., German Retana, S., Boubekeur, T., et al. (2009). Remorin, a solanaceae protein resident in membrane rafts and plasmodesmata, impairs potato virus X movement. *Plant Cell* 21, 1541–1555.

Schroeder, R., London, E., and Brown, D. (1994). Interactions between saturated acyl chains confer detergent resistance on lipids and glycosylphosphatidylinositol (GPI)-anchored proteins: GPI-anchored proteins in liposomes and cells show similar behavior. *Proc. Natl. Acad. Sci. U.S.A.* 91, 12130–12134.

Shahollari, B., Peskan-Berghofer, T., and Oelmuller, R. (2004). Receptor kinases with leucine-rich repeats are enriched in Triton X-100 insoluble plasma membrane microdomains from plants. *Physiol. Plant.* 122, 397–403.

Sharma, P., Sharma, N., and Deswal, R. (2005). The molecular biology of the low-temperature response in plants. *Bioessays* 27, 1048–1059.

Simons, K., and Ikonen, E. (1997). Functional rafts in cell membranes. *Nature* 387, 569–572.

Simons, K., and Toomre, D. (2000). Lipid rafts and signal transduction. *Nat. Rev. Mol. Cell Biol.* 1, 31–39.

Singer, S. J., and Nicolson, G. L. (1972). The fluid mosaic model of the structure of cell membranes. *Science* 175, 720–731.

Stanislas, T., Bouyssie, D., Rossignol, M., Vesa, S., Fromentin, J., Morel, J., et al. (2009). Quantitative proteomics reveals a dynamic association of proteins to detergent-resistant membranes upon elicitor signaling in tobacco. *Mol. Cell. Proteomics* 8, 2186–2198.

Takahashi, D., Kawamura, Y., Yamashita, T., and Uemura, M. (2012). Detergent-resistant plasma membrane proteome in oat and rye: similarities and dissimilarities between two monocotyledonous plants. *J. Proteome Res.* 11, 1654–1665.

Tanner, W., Malinsky, J., and Opekarová, M. (2011). In plant and animal cells, detergent-resistant membranes do not define functional membrane rafts. *Plant Cell* 23, 1191–1193.

Uemura, M., Tominaga, Y., Nakagawara, C., Shigematsu, S., Minami, A., and Kawamura, Y. (2006). Responses of the plasma membrane to low temperatures. *Physiol. Plant.* 126, 81–89.

Uemura, M., and Yoshida, S. (1984). Involvement of plasma membrane alterations in cold acclimation of winter rye seedlings (*Secale cereale* L. *cv* Puma). *Plant Physiol.* 75, 818–826.

Webb, M. S., Uemura, M., and Steponkus, P. L. (1994). A comparison of freezing injury in oat and rye: two cereals at the extremes of freezing tolerance. *Plant Physiol.* 104, 467–478.

Wong, H. L., Sakamoto, T., Kawasaki, T., Umemura, K., and Shimamoto, K. (2004). Down-regulation of metallothionein, a reactive oxygen scavenger, by the small GTPase OsRac1 in rice. *Plant Physiol.* 135, 1447–1456.

Xu, X., Bittman, R., Duportail, G., Heissler, D., Vilcheze, C., and London, E. (2001). Effect of the structure of natural sterols and sphingolipids on the formation of ordered sphingolipid/sterol domains (rafts): comparison of cholesterol to plant, fungal, and disease-associated sterols and comparison of sphingomyelin, cerebrosides, and ceramide. *J. Biol. Chem.* 276, 33540–33546.

Yamazaki, T., Kawamura, Y., Minami, A., and Uemura, M. (2008). Calcium-dependent freezing tolerance in *Arabidopsis* involves membrane resealing via synaptotagmin SYT1. *Plant Cell* 20, 3389–3404.

Conflict of Interest Statement: The authors declare that the research was

conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 28 December 2012; paper pending published: 15 January 2013; accepted: 05 February 2013; published online: 22 February 2013.

Citation: Takahashi D, Kawamura Y and Uemura M (2013) Detergent-resistant plasma membrane proteomes to elucidate microdomain functions in plant cells. Front. Plant Sci. ***4****:27. doi: 10.3389/fpls.2013.00027*

This article was submitted to Frontiers in Plant Proteomics, a specialty of Frontiers in Plant Science.

Copyright © 2013 Takahashi, Kawamura and Uemura. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.

frontiers in
PLANT SCIENCE

ORIGINAL RESEARCH ARTICLE
published: 06 February 2013
doi: 10.3389/fpls.2013.00011

Defining the core proteome of the chloroplast envelope membranes

Stefan Simm[1], Dimitrios G. Papasotiriou[2†], Mohamed Ibrahim[1†], Matthias S. Leisegang[1], Bernd Müller[3†], Tobias Schorge[2], Michael Karas[2,4,5], Oliver Mirus[1], Maik S. Sommer[1] and Enrico Schleiff[1,4,5]*

[1] Institute of Molecular Cell Biology of Plants, Goethe University, Frankfurt, Germany
[2] Institute of Pharmaceutical Chemistry, Goethe University, Frankfurt, Germany
[3] Department of Biology I, Ludwig-Maximilians-University, Munich, Germany
[4] Center of Membrane Proteomics, Goethe University, Frankfurt, Germany
[5] Cluster of Excellence 'Macromolecular Complexes', Goethe University, Frankfurt, Germany

Edited by:
Harvey Millar, The University of Western Australia, Australia

Reviewed by:
Lutz Andreas Eichacker, University of Stavanger, Norway
Sascha Rexroth, Ruhr-University Bochum, Germany
Markus Teige, University of Vienna, Austria

****Correspondence:***
Enrico Schleiff, Center of Membrane Proteomics, Cluster of Excellence 'Macromolecular Complexes', Institute of Molecular Cell Biology of Plants, Goethe University, Max-von-Laue Strasse 9, Frankfurt 60438, Germany.
e-mail: schleiff@bio.uni-frankfurt.de

†Current address:
Dimitrios G. Papasotiriou, Jealott's Hill International Research Centre, Bracknell, Berkshire, UK; Mohamed Ibrahim, Botany Department, Faculty of Science, Ain Shams University, Cairo, Egypt; Bernd Müller, AB SCIEX Germany GmbH, Darmstadt, Germany.

High-throughput protein localization studies require multiple strategies. Mass spectrometric analysis of defined cellular fractions is one of the complementary approaches to a diverse array of cell biological methods. In recent years, the protein content of different cellular (sub-)compartments was approached. Despite of all the efforts made, the analysis of membrane fractions remains difficult, in that the dissection of the proteomes of the envelope membranes of chloroplasts or mitochondria is often not reliable because sample purity is not always warranted. Moreover, proteomic studies are often restricted to single (model) species, and therefore limited in respect to differential individual evolution. In this study we analyzed the chloroplast envelope proteomes of different plant species, namely, the individual proteomes of inner and outer envelope (OE) membrane of *Pisum sativum* and the mixed envelope proteomes of *Arabidopsis thaliana* and *Medicago sativa*. The analysis of all three species yielded 341 identified proteins in total, 247 of them being unique. 39 proteins were genuine envelope proteins found in at least two species. Based on this and previous envelope studies we defined the core envelope proteome of chloroplasts. Comparing the general overlap of the available six independent studies (including ours) revealed only a number of 27 envelope proteins. Depending on the stringency of applied selection criteria we found 231 envelope proteins, while less stringent criteria increases this number to 649 putative envelope proteins. Based on the latter we provide a map of the outer and inner envelope core proteome, which includes many yet uncharacterized proteins predicted to be involved in transport, signaling, and response. Furthermore, a foundation for the functional characterization of yet unidentified functions of the inner and OE for further analyses is provided.

Keywords: membrane proteome, plant proteomics, chloroplast membrane proteins, mass spectrometry, envelope membrane proteome approach comparison

INTRODUCTION

The characterization of a single protein function is associated with an enumeration of different features. Some of these features are the subcellular localization of the protein, its interaction with other proteins, co- or post-translational modifications as well as its (enzymatic) activity. With the growing number of sequenced genomes, the "proteome," as sum of all proteins in an entire cell or cellular (sub-)compartment, becomes important for the understanding of cellular function (Wilkins et al., 1996; James, 1997). The mass spectrometric analysis of complete cellular proteomes still remains difficult, especially in the highly compartmentalized eukaryotic cells. Furthermore, proteomes are dynamic and change in response to different stimuli. They include different splice forms and post-translationally modified proteins in different abundances. Thus, different technical approaches have been developed to accommodate the complexity of a proteome (e.g., Karas and Hillenkamp, 1988; Aebersold and Mann, 2003), especially to study the subcellular localization of membrane proteins as a complementary approach to the complete cell proteome analyses (van Wijk, 2000; Millar et al., 2009).

This complexity of eukaryotic cells leads us to focus on the proteome of chloroplast, which are organelles essential for different metabolic pathways like photosynthesis, and biosynthesis of fatty acids or amino acids. These organelles contain several thousand different proteins and the majority of which is cytosolically synthesized and has to be translocated across the envelope membranes (Leister, 2003; Schleiff and Becker, 2011).Thereby, the proteome of the organelle as such (Zabrouskov et al., 2003; Kleffmann et al., 2004) or of subfractions like the thylakoid lumen (Peltier et al., 2002), the thylakoid membranes (Eichacker et al., 2004; Friso et al., 2004), the stroma (Goulas et al., 2006; Peltier et al., 2006), plastoglobules (Ytterberg et al., 2006), or the envelope membranes (Schleiff et al., 2003b; Bräutigam and Weber, 2009) have been analyzed in the past. The current knowledge on the

proteomic content of chloroplasts has been deposited in several independent databases like PLPROT (Kleffmann et al., 2006) or AT_CHLORO (Ferro et al., 2010). However, especially the analysis of the envelope and more specifically the inner envelope (IE) and outer envelope (OE) membrane proteome is still a challenging task due to the hydrophobicity of membrane proteins (Eichacker et al., 2004). More specifically, the dissection of the IE and OE membrane proteome is still very poorly supported by direct proteomic studies (Ferro et al., 2003; Schleiff et al., 2003b).

The determination of a protein's localization is a very important tool for experimental guidance. In here, we aimed at the determination of a reliable proteome of the OE and IE membranes of chloroplasts. To this end, we comparatively analyzed the overall envelope proteomes of the model species *Arabidopsis thaliana* and *Medicago sativa*. To substantiate our findings, we individually analyzed IE and OE membranes of *Pisum sativum*, the only plant to date, for which the separation of both can be achieved (Ferro et al., 2003; Schleiff et al., 2003b). We chose the genetic model *A. thaliana* by its comprehensive genome and transcriptome data available (see, e.g., The Arabidopsis Information Resource, TAIR10; Lamesch et al., 2012). The legume *P. sativum* was chosen, as it is the model plant for biochemical analyses of chloroplast function (see, e.g., Franssen et al., 2012). Due to the paucity of data, the recently sequenced and closely related *M. sativa* was used to substantiate our findings for *P. sativum*.

The identified proteins in these plant species were compared to each other and to the publicly available datasets of previous studies. We identified a total of 247 different proteins, of which – based on comparisons with other studies – 191 were assigned as putative envelope proteins. To our surprise, only 27 of these were found in all studies. Based on intersection and cross-contamination analysis of available previous studies, we were able to reliably assign 50/49 proteins as outer/inner membrane-localized, while at least 37 additional proteins in the mixed envelope fractions can be assigned as envelope proteins as well, but not reliably to a specific membrane.

RESULTS AND DISCUSSION

CHLOROPLAST PROTEOME ANALYSES

We analyzed the chloroplast proteomes with focus on the envelope membranes from three model plant species, namely *A. thaliana*, *P. sativum*, and *M. sativa*. We chose *A. thaliana* because of the availability of a comprehensive genome and many existing transcriptome data (e.g., The Arabidopsis Information Resource, TAIR10; Lamesch et al., 2012). Thus, the well annotated genome of *Arabidopsis* provides a solid base for the assignment of the identified inner and OE proteins. In turn, the legumes *P. sativum* and *M. sativa* are model plants for biochemical analyses of chloroplast function (e.g., Franssen et al., 2012), as well as crop plants. Using envelopes of different plant species allows the detection of proteins with different abundances. The varying achievable purity of the samples allows the detection of an additional different set of peptides.

We isolated and subfractionated chloroplasts to analyze the envelope proteomes (**Figures 1A,B**). The enrichment of the

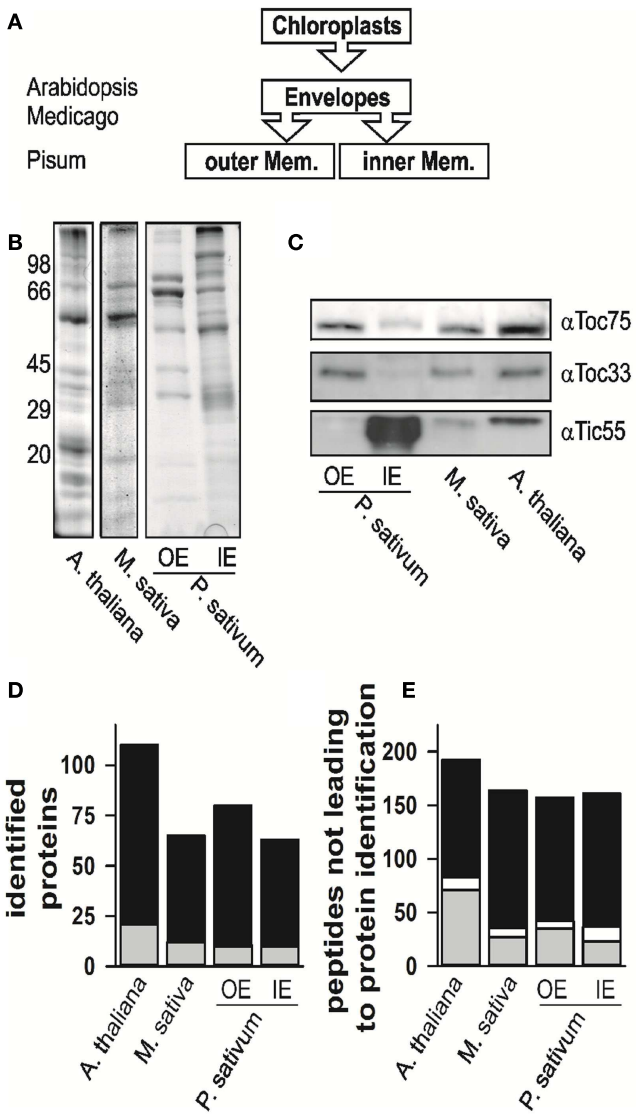


FIGURE 1 | The proteome analysis. (A) Schematic representation of which fractions were isolated and analyzed. The different species are indicated for the envelope fraction results of six independent replicates, three after trypsin and three after elastase digestion were combined. **(B)** The fractions of mixed envelope of *A. thaliana* and *M. sativa* as well as the outer (OE) and inner envelope (IE) membrane of *P. sativum* were subjected to SDS-PAGE analysis followed by Coomassie Blue staining. The migration of the molecular weight standard is indicated on the left. **(C)** The purity of the fractions in **(B)** was assessed by Western blotting using indicated antibodies. **(D)** Numbers of proteins identified in the according fractions by MALDI nano-LC-MS/MS and the two digestion methods indicated. Gray indicates the portion for which more than one AGI was assigned for one protein family, in white the portion where more than one isoform was specifically identified for one protein, black indicates the portion for which one AGI was assigned. **(E)** Numbers of peptides not assigned by MALDI nano-LC-MS/MS and BLAST assignment. Gray indicates the portion of peptides, which were assigned to one amino acid sequence only, whereas white indicates the portion of peptides, which were assigned to various proteins, black indicates the portion of peptides, which were not assigned at all.

obtained fractions was assessed by Western blotting using specific antibodies (**Figure 1C**). The analysis confirmed the enrichment of inner and outer membrane proteins in the mixed envelope fractions of *A. thaliana* and *M. sativa*, the mixed envelope fractions could not be further separated. In contrast, separation of envelope membranes in the IE and OE from *P. sativum* chloroplasts has been established previously (e.g., Schleiff et al., 2003a,b). Subsequently, the distinct fractions were analyzed by mass spectrometry.

The proteomes of all envelope membranes were analyzed by MALDI nLC-MS/MS (Table S14 in Supplementary Material) yielding in total 110 proteins in *A. thaliana* (**Figure 1D**, *Arabidopsis* EM, three independent isolations; Table S1 in Supplementary Material). In parallel, we identified 71 proteins in *M. sativa* (**Figure 1D**, *Medicago* EM, three independent isolations; Table S2 in Supplementary Material) and 124 different proteins in both membranes (87 IE; 73 OE) of *P. sativum* (**Figure 1D**, *Pisum* IE and *Pisum* OE, three independent isolations; Tables S3 and S4 in Supplementary Material).

Our peptide-based assignment relies on a stringent BLAST search, where an identity >95% and no mismatch or gap was allowed. Only a single amino acid substitution with a residue of similar properties or a single undefined amino acid position was accepted (for details see Experimental Section). The BLAST search was combined with a bidirectional best BLAST hit search to assign the homologous sequences in *A. thaliana* to the proteins identified in *P. sativum* or *M. sativa* to render the assignment from different species comparable. To confirm that the peptide-based assignment is consistent with the expected chloroplast localization, we analyzed the expression of the corresponding genes with respect to leaves and roots (e.g., Vojta et al., 2004). Indeed, almost all genes coding for the identified proteins including those identified by a single peptide only are highly expressed in leaf tissue (**Figure A1** in Appendix). AT3G45360 is the only exception identified by more than one peptide with an expression value smaller than 10 in leaves. However, this gene is annotated as a transposable element. Furthermore, almost all genes are equally high or higher expressed in leaves in comparison to roots. The only gene with a significantly higher expression in roots than in leaves is AT3G09260 identified in *A. thaliana*. It encodes a β-glucosidase annotated as Pyk10, which was identified in ER-bodies (Matsushima et al., 2003). Although the protein most likely represents a contamination of the sample, its overall expression pattern supports the peptide-based protein assignment approach.

While analyzing the data, a large number of the obtained peptides did not lead to an identification of a protein (**Figure 1E**, Tables S5–S7 in Supplementary Material). About 15–30% of these peptides mapped uniquely to a single sequence (in gray), while few peptides mapped to multiple sequences (in white). The large portion of peptides which remained unassigned (in black) might have three different reasons: (i) The choice of too stringent search parameters, (ii) contaminations of the samples, or (iii) the existence of natural variances of sequences in form of unknown splice variants or nucleotide polymorphisms of genes leading to alternative amino acid sequences. The analysis of this phenomenon, however, goes beyond the scope of this work.

COMPARISON TO OTHER ENVELOPE MEMBRANE PROTEOMIC APPROACHES

To establish a core envelope proteome we unified results of our and previous studies (Ferro et al., 2003, 2010; Froehlich et al., 2003; Bräutigam et al., 2008; Bräutigam and Weber, 2009). For that, we first assigned the *Arabidopsis* Genome Initiative (AGI) number of the closest homolog of *A. thaliana* to each of the proteins found in *M. sativa* and *P. sativum*. Combining our four data sets, we obtained 247 different proteins in total. The globally unified protein pool contains a total of 911 different proteins. Ferro et al. (2010) assigned their identified proteins according to the suborganellar (stroma, thylakoid, and envelope) localization, which we have used to assess the quality of our data (cross-contaminations from thylakoid and stroma). We defined four different categories (**Table 1**): Category I are proteins that were found in at least two studies but not in the stroma and thylakoid according to Ferro et al. (2010). Category II unites proteins, which were found in at least three studies but also in the stroma or thylakoid. Category III are proteins found in one study only, but exclusively in the envelope, and category IV are proteins found in less than three studies, but also in the stroma or thylakoid. The selection of three independent studies for category II as criterion takes into account that two studies each come from Bräutigam et al. (2008), Bräutigam and Weber (2009), and Ferro et al. (2003, 2010). For better visualization of the impact of our study we have marked the identified proteins of the categories as identified in here (a) or in previous studies (b).

From our point of view the list of proteins of category I is most reliable, because there are no cross-contaminations via thylakoid and stroma and they are supported by previous studies. Proteins of categories II and III have to be confirmed experimentally first and proteins of category IV are considered to be not reliably assigned.

We noticed that only 30 proteins were identified in all six studies (categories Ia, IIa, **Table 1**), of which three have been identified in the stroma or thylakoid as well. In total, we found 231 proteins of category I. Additionally, we found 346 proteins of category III according to Ferro et al. (2010), which are not cross-contaminations of the stroma or thylakoid (Ferro et al., 2010). Hence, they might represent envelope proteins as well. However, as stated above, this conclusion should be challenged by biochemical approaches. The latter holds true for particularly 72 proteins of category II, which have been identified in envelope and in stroma or thylakoid. However, 262 proteins have been assigned to category IV.

Based on the PPDB and SubaII databases, we next analyzed whether proteins have been previously assigned to the mitochondrion, peroxisome, nucleus, ER, golgi, plasma membrane or cytosol, and not to the plastid (**Table 1**; Heazlewood et al., 2007; Sun et al., 2009). Accordingly, 31/12 proteins of category I were assigned to other cellular localizations according to PPDB/SubaII, respectively. In category II we found 2/0 proteins and in category III 74/56 proteins, respectively, which have been identified in cellular compartments other than chloroplasts. Thus, about 10 and 20% of the proteins assigned to category I or category III are found in other cellular compartments than the chloroplast. The low abundance of mislocalized proteins in category II might reflect that the proteome of the stroma and thylakoid (Ferro et al.,

Table 1 | Categories for the classification of envelope membrane proteins.

Category	Mixed envelope fraction		Thylakoid or Stroma	Proteins identified in the given number of studies							
	This study	Others	Ferro et al. (2010)	6	5	4	3	2	1	SUM	OTH
Ia	+	At least one	–	27	12	25	17	18	–	99	11/8
Ib	–	At least two	–	–	5	11	41	75	–	132	20/4
IIa	+	At least two	+	3	14	11	16	–	–	44	0/0
IIb	–	At least three	+	–	2	9	17	–	–	28	2/0
IIIa	+	None	–	–	–	–	–	–	48	48	6/10
IIIb	–	One	–	–	–	–	–	–	298	298	68/46
IVa	+	Less than two	+	–	–	–	–	35	21	56	0/0
IVb	–	Less than three	+	–	–	–	–	53	153	206	16/0

Given is the category defined in the text (column 1), protein identification by us (+, column 2) or by any other proteomic study (column 3 defines the required number of identifications), identification in the thylakoid or stroma (+, Ferro et al., 2010; column 4) and the number of proteins identified in 6, 5, 4, 3, 2, or 1 study (column 5–10) as well as the number of proteins in each category (column 11), and the number of proteins, which have been identified in other cellular fractions than chloroplasts as well based on PPDB (Sun et al., 2009)/SubAll (Heazlewood et al., 2007; column 12).

2010) has been established quite well. Nevertheless, the assignment of proteins in other organellar fractions does not necessarily mark them as false positive chloroplast proteins as (i) chloroplasts are the major organelles of plant cells and thus, contaminations of other fractions might exist and (ii) an increasing number of proteins are found to be dually localized (Carrie and Small, 2013).

COMPARISON OF THE IDENTIFIED ENVELOPE PROTEOMES OF THE DIFFERENT SPECIES

Next, we compared our envelope proteomes obtained for the different analyzed plants with focus on proteins assigned to categories I–III (Table S8 in Supplementary Material). The 191 of total proteins assigned included 48 proteins identified in *M. truncatula*, 68 proteins in *A. thaliana*, and 127 proteins in *P. sativum*. Thirty-nine proteins were identified in at least two plant species, 13 of which were found in all three (**Figure 2A**; **Table 2**). Dissecting the protein set of *P. sativum* into OE and IE localized, revealed a total of 46 OE and 60 IE proteins. Twenty-one proteins were found in both fractions. We compared the OE and IE proteins separately with the identified envelope proteins of the other two plants (**Figure 2B**). This analysis shows that all 13 proteins identified in all species were also found in the IE, while 7 of them are also found in the OE. Similarly, all proteins found in the overlap between *P. sativum* and *A. thaliana* are found in the IE fraction (11), while the overlap with the *M. sativa* envelope contains four proteins (AT2G01320, AT4G32250, Toc64-III, and Toc132) specifically found in the OE of *P. sativum*.

The set of proteins found in all three species include amongst others solute transporters like LptD and Iep37 and as part of the IE/OE preprotein translocases Toc75-III, Toc159, and Tic55-II. Remarkably, only a single protein with unknown function was identified in all envelope fractions, namely At5g08540. Additionally, seven proteins of category II are detected in all species including the photosynthesis proteins LHCB6, PSAD-2, and ATPB. Furthermore, three proteins involved in signaling and response (CA1, RCA, and FNR1) and SDX1 of the lipid biosynthesis are identified. Remarkably, we could identify only one protein of category I in the envelope fractions of *A. thaliana* and *M. sativa*, which is the dually targeted (mitochondria and chloroplast) *S*-adenosylmethionine carrier 1 (SamC1; Palmieri et al., 2006). The category I proteins involved in transport (Oep16, NAP8), preprotein import (CJD1, Tic110), and signaling (MDH) could be detected in the envelope fractions of *A. thaliana* and *P. sativum* (**Table 2**). It appears that subfractionation of IE and OE membranes in case of the samples from *P. sativum* lead to an increased detection of preprotein import (Toc120, Tic55-IV, and Tic40) and transporter (Oep37, NAP14, MEX1, KEA2, DiT1, and DiT2.1) proteins. For the envelope fractions of *A. thaliana* and *M. sativa* only the preprotein import protein Toc75-V (*M. sativa*) and the transport proteins KEA1, TIP1.1, PIP2A, and PCaP1 in *A. thaliana* and Oep16-2 in *M. sativa* could be identified.

THE OUTER AND INNER ENVELOPE MEMBRANE PROTEOME

Next, we inspected the individual proteomes of the OE and IE membrane of *P. sativum*, respectively. We only assigned proteins of categories I and III, which have been identified by at least two peptides. Due to the high uncertainty, proteins of category II were omitted (see above). Taking these criteria into account we could assign 30 proteins of known function to the OE (**Table 3**), and 34 proteins to the IE membrane (**Table 4**) and additional 22 proteins could not be clearly assigned (**Table 5**). In addition, we assigned 50 proteins of unknown function, 15 of them to the IE, 20 to the OE (fraction), and 15 to the envelope in general (**Table 6**). Thus, in total we were able to clearly assign 50 OE and 49 IE proteins (**Figure 3**) and will explain them in detail in the following sections.

Outer envelope proteins

We identified homologs to known OE proteins such as components of the TOC complex (Schleiff and Becker, 2011), like Toc75-III, Toc34, Toc159, Toc120 and Toc132, and Toc64-III which have been previously reported (Schleiff et al., 2003a; Ladig et al., 2011). The latter three were exclusively found in the OE membrane. Remarkably, we were not able to detect Toc75-V, except in the envelope fraction of *M. sativa* (**Table 5**). Further identified proteins with confirmed OE localization were Oep37, Oep21, and Oep16 (Schleiff et al., 2003a), SENSITIVE TO FREEZING 2

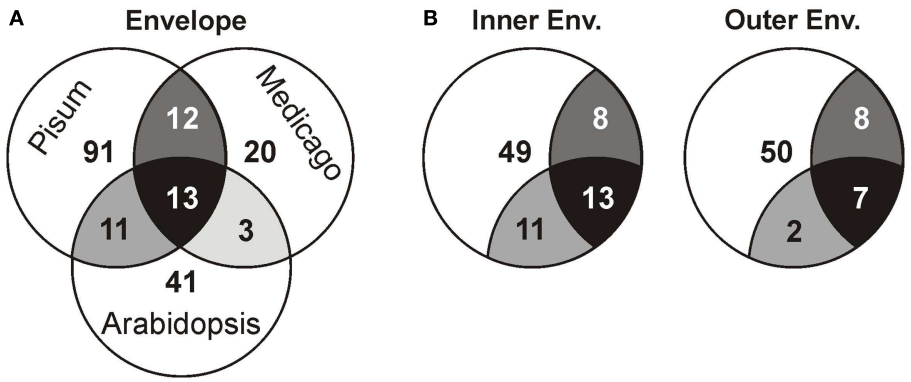


FIGURE 2 | Analysis of the different proteomes with respect to the putative localization. (A) The overlap of the proteome determined for *A. thaliana*, *M. sativa*, and *P. sativum* mixed envelope was analyzed by MALDI nano-LC-MS/MS. The overlap is displayed by the shared regions of the circles in the Venn diagram. **(B)** The overlap of the proteome determined for *A. thaliana*, *M. sativa* mixed envelope and *P. sativum* inner (left) and outer (right) envelope was analyzed and the color code is taken from **(A)**.

Table 2 | Proteins in the envelope fraction of at least two species.

AGI	Abbr.	*A. thaliana*	*M. sativa*	*P. sativum*			AGI	Abbr.	*A. thaliana*	*M. sativa*	*P. sativum*		
				OE	IE	M					OE	IE	M
At1g03130	PSAD-2	X	X		X		At3g26740	CCL	X				X
At1g06950	Tic110	X			X		At3g46740	Toc75-III	X	X			X
At1g08640	CJD1	X			X		At3g47520	MDH	X			X	
At1g15820	LHCB6	X	X		X		At3g63410	Iep37	X	X		X	
At1g55670	PSAG		X		X		At4g02510	Toc159	X	X			X
At1g65260	VIPP1		X		X		At4g15440	HPL1		X		X	
At1g67090	RBCS1A		X			X	*At4g20360*	RAB8d	X			X	
At1g74470	Unknown	X			X		At4g25450	NAP8	X			X	
At1g77590	LACS9		X			X	At4g32250	Unknown		X	X		
At2g01320	Unknown		X	X			*At4g32260*	Unknown	X	X			
At2g16640	Toc132		X	X			*At4g32770*	SDX1	X	X		X	
At2g24820	Tic55-II	X	X		X		At4g33350	Unknown		X			X
At2g28900	Oep16	X				X	At4g39460	SAMC1	X	X			
At2g39730	RCA	X	X			X	At5g05000	Toc34		X			X
At2g44640	LptD	X	X			X	At5g08540	Unknown	X	X		X	
At3g01500	CA1	X	X			X	At5g24650	Unknown		X		X	
At3g04340	emb2458	X			X		*At5g50920*	HSP93-V	X			X	
At3g16950	LPD1	X			X		*At5g66190*	FNR1	X	X			X
At3g17970	Toc64-III		X	X			*AtCg00480*	ATPB, PB	X	X			X
At3g23400	FIB4	X	X										

Given is the Arabidopsis Genome Initiative (AGI) number (italic indicates category II), the short name and aliases, the identification in A. thaliana mixed envelope, M. sativa mixed envelope, P. sativum outer envelope, inner envelope, or mixed envelope. Identification of the protein is marked by an X in the column.

protein (Sfr2), a galactolipid-remodeling enzyme (Fourrier et al., 2008; Moellering et al., 2010) and CRUMPLED LEAF protein (Crl) and PDV2, which are both involved in plastid division (Asano et al., 2004; Glynn et al., 2008).

Additionally, we included proteins, for which significantly more peptides were found in the OE than in the IE fraction, albeit their exact localization is unclear. The long-chain acyl-CoA synthetase Lacs9 (Schnurr et al., 2002), the ABC-type transporter WBC7 (Ferro et al., 2003), and the paralog of TGD4 (Xu et al., 2008) encoded by AT2G44640 (LptD; Haarmann et al., 2010) were shown to be localized in the envelope membranes, before (Ferro et al., 2003; Froehlich et al., 2003). Similarly, the kinase CoaE was identified in the chloroplast proteome, but experimental data on the localization does not exist (Zybailov et al., 2008).

Table 3 | Outer envelope proteins with known function.

	AGI	Abbr.	Name and function	TM fold	Other Loc.	Studies	Cat.
Preprotein import	At2g16640	Toc132	GTP-binding chloroplast preprotein receptor	Unknown	–	3	I
	At3g16620	Toc120	GTP-binding chloroplast preprotein receptor	Unknown	n.d.	2	I
	At3g17970	Toc64-III	Chloroplast preprotein receptor	α-Helical TM	–	3	I
	At3g46740	Toc75-III	Translocon channel	β-Barrel	–	6	I
	At4g02510	Toc159	GTP-binding chloroplast preprotein receptor	Unknown	–	6	I
	At5g05000	Toc34	GTP-binding chloroplast preprotein receptor	1 α-Helical TM	–	6	I
Lipid biosyn.	At1g77590	LACS9	Long-chain acyl-CoA synthetase	1 α-Helical TM	–	6	I
	At3g06510	SFR2	Beta-glucosidase	2 α-Helical TM	–	6	I
	At4g31780	MGD1	Type A monogalactosyldiacylglycerol synthase	None	–	5	I
Transport	At1g20816	Oep21	Outer envelope channel	None	n.d./–	3	I
	At1g45170	Oep24	Outer envelope channel	None	–/n.d.	1	III
	At2g01320	WBC7	Putative subfamily G ABC-type transporter	4 α-Helical TM	–	3	I
	At2g28900	Oep16-I	Outer envelope protein	2 α-Helical TM	–	6	I
	At2g43950	Oep37	Outer membrane ion channel	β-Barrel	–	6	I
	At2g44640	LptD	Lipopolysaccharide-assembly protein D	β-Barrel	–	6	I
Others	At2g16070	PDV2	Plastid division machinery	1 α-Helical TM	–	3	I
	At2g17390	AKR2B	AKR2-like protein	None	n.d./–	1	III
	At2g27490	COAE	Putative dephospho-CoA kinase	None	X	2	I
	At3g27820	MDAR4	Membrane-associated monodehydroascorbate reductase	2 α-Helical TM	X	2	I
	At4g05050	UBQ11	Polyubiquitin	None	n.d.	1	III
	At4g29130	HXK1	Glucose-responsive sensor hexokinase	1 α-Helical TM	X	4	I
	At5g17770	CBR1	NADH:cytochrome b5 reductase	1 α-Helical TM	n.d./X	2	I
	At5g51020	CRL	Affects pattern of plastid division	1 α-Helical TM	–	2	I
	At5g58140	NPL1	Multifunctional blue-light-responsive photoreceptor	None	X	2	I
Other Organelles	At1g27390	Tom20-2	Putative mitochondrial outer membrane translocase component	1 α-Helical TM	n.d./X	3	I
	At3g46030	HTB11	Putative H2B-type histone	None	X/n.d.	1	III
	At4g14430	ECHIb	Putative enoyl-CoA hydratase/isomerase	1 α-Helical TM	X	1	III
	At4g35000	APX3	Putative peroxisomal ascorbate peroxidase	1 α-Helical TM	X	5	I
	At4g38920	VHA-C3	c-Type subunit of vacuolar H(+)-ATPase membrane V0 subcomplex	4 α-Helical TM	n.d./X	1	III
	At5g43070	WPP1	Nuclear envelope-targeted protein involved in mitotic activity	None	n.d./X	2	I

Given is the functional pathway or the organellar compartment, the AGI number, the short name (Abbr.), the (putative) function, the transmembrane anchor architecture, other localization by the PPDB (Sun et al., 2009) and SUBAII (Heazlewood et al., 2007; n.d. not defined, –, no other localization, X, other localization), the number of studies where the protein was identified (our study; Ferro et al., 2003; Froehlich et al., 2003; Bräutigam et al., 2008; Bräutigam and Weber, 2009; Ferro et al., 2010), and the category from our study. Lipid biosynthesis (Lipid biosyn.).

The protein encoded by AT5G27330 is annotated as Prefold in chaperone subunit family protein and was predicted to be localized in the endoplasmic reticulum (Dunkley et al., 2006). Likewise, ascorbate peroxidase Apx3 (Narendra et al., 2006) was previously assigned to peroxisomal membranes, while Cbr1 (Fukuchi-Mizutani et al., 1999) was described as a protein of the microsomal electron-transfer system. Remarkably, both proteins were identified as substrates of the Akr2a-dependent transport (Shen et al., 2010), which is also involved in the transport of Oep7 to the chloroplast OE membrane (Bae et al., 2008). Furthermore, Apx3 was previously identified in the chloroplast proteome (Zybailov et al., 2008). Although unclear, these proteins are most likely dually localized to both, peroxisomes or ER and chloroplasts.

In contrast, we identified a couple of proteins, which indicate a slight impurity of the sample, namely Mdar4 (Lisenbee et al., 2005), which was clearly assigned to the peroxisomal membrane, Wpp1 (Patel et al., 2004) and Hxk1 (Moore et al., 2003), which are nuclear proteins, the mitochondrial proteins Tom20 and AT4G16450 (Lister et al., 2007; Klodmann et al., 2010) vacuolar protein AVA-P3, and IE protein MGD1, the MGDG synthase (Awai et al., 2001; Ladig et al., 2011). Further, LHCB1.4 is a thylakoid protein. AT4G27680, AT3G52230, AT2G32240, AT3G53560, AT2G24440, AT1G09920, AT3G49350, and AT1G68680 are unknown, while for the protein kinase encoded by AT4G32250 a stromal localization was proposed (Friso et al., 2004; Zybailov et al., 2009).

Inner envelope proteins

Analyzing the IE proteome of *P. sativum*, we realized that it was in contrast to the OE fraction heavily contaminated with proteins of

Table 4 | Inner envelope proteins with known function.

	AGI	Abbr.	Name and function	TM fold	Other Loc.	Studies	Cat
Preprotein import	At1g06950	Tic110	Inner envelope translocon component	1 α-Helical TM	–	6	I
	At1g08640	CJD1	DnaJ-like membrane protein	3 α-Helical TM	–	5	I
	At2g24820	Tic55-II	Inner envelope Rieske iron-sulfur protein	3 α-Helical TM	–	5	I
	At4g23420	Tic32-IVb	NAD- or NADP-dependent oxidoreductase	1 α-Helical TM	n.d.	2	I
	At4g25650	Tic55-IV	Inner envelope Rieske iron-sulfur protein	2 α-Helical TM	–	4	I
	At4g33350	Tic22-IV	Inner envelope translocon component	None	–	6	I
	At5g16620	Tic40	Inner envelope translocon component	1 α-Helical TM	–	6	I
Lipid biosyn.	At4g15440	HPL1	Membrane-associated hydroperoxide lyase	2 α-Helical TM	–	4	I
	At4g31500	SUR2	Cytochrome P450 monooxygenase	2 α-Helical TM	X	1	III
	At5g01220	SQD2	UDP-sulfoquinovose:DAG sulfoquinovosyltransferase	1 α-Helical TM	–	4	I
	At5g05580	FAD8	chloroplast omega-3 fatty acid desaturase	3 α-Helical TM	–	3	I
Transport	At1g80300	NTT1	Plastidic ATP/ADP antiporter	11 α-Helical TM	–	5	I
	At3g20320	TGD2	Putative subfamily I ABC protein	1 α-Helical TM	–	6	I
	At3g63410	IEP37	37 kDa chloroplast inner envelope protein	2 α-Helical TM	–	6	I
	At4g00630	KEA2	Putative potassium cation efflux antiporter	14 α-Helical TM	–	4	I
	At4g25450	NAP8	Putative subfamily B ABC-type transporter	5 α-Helical TM	–	6	I
	At5g12860	DiT1	Plastidic 2-oxoglutarate/malate-translocator	13 α-Helical TM	–	6	I
	At5g14100	NAP14	Putative subfamily I ABC protein	None	–	3	I
	At5g17520	MEX1	Putative maltose translocator	9 α-Helical TM	–	3	I
	At5g24650	PRAT2.2	Putative dual-targeted mitochondrial and plastidial membrane translocase	4 α-Helical TM	–	6	I
	At5g64290	DiT2.1	Plastidic glutamate/malate-translocator	11 α-Helical TM	–	6	I
SR	At1g32080	LRGB	LrgB-like membrane protein	12 α-Helical TM	–	5	I
	At3g47520	MDH	NAD-malate dehydrogenase	1 α-Helical TM	–	4	I
	At5g23040	CDF1	Cell growth defect factor	3 α-Helical TM	–	5	I
Proteases	At1g79560	FtsH12	ATP-dependent metalloprotease	2 α-Helical TM	–	4	I
	At5g53170	FtsH11	ATP-dependent metalloprotease	1 α-Helical TM	X	4	I
	At5g64580	FtsHi4	Putative ATP-dependent metalloprotease	1 α-Helical TM	–	4	I
Embryon. develop.	At1g10510	Emb2004	Putative membrane protein	1 α-Helical TM	–	6	I
	At3g04340	Emb2458	Putative membrane protein	3 α-Helical TM	–	4	I
	At3g52590	UBQ1	Ubiquitin extension protein	None	X	1	III
	At5g22640	Emb1211	Putative membrane protein	None	–	4	I
	At5g53860	Emb2737	Putative membrane protein	None	–	3	I
Others	At1g65260	VIPP1	Membrane-associated vesicle-inducing prot.	None	–	6	I
	At2g37860	LCD1	Mutant lcd1-1 exhibits pale phenotype	2 α-Helical TM	–	4	I

Given is the functional pathway or the organellar compartment, the AGI number, the short name (Abbr.), the (putative) function, the transmembrane anchor architecture, other localization by the PPDB (Sun et al., 2009) and SUBAII (Heazlewood et al., 2007; n.d., not defined, –, no other localization, X, other localization), the number of studies where the protein was identified (our study; Ferro et al., 2003; Froehlich et al., 2003; Bräutigam et al., 2008; Bräutigam and Weber, 2009; Ferro et al., 2010), and the category from our study. Lipid biosynthesis (Lipid biosyn.); signaling and response (SR); embryonic development (Embryon. develop.).

the stroma and the OE. First, with clearly annotated OE proteins like Toc75-III, Toc159, Toc34, Lacs9, and Oep21. Second, with stromal proteins like the small subunit of ribulose bisphosphate carboxylase Rbcs1A, Rbcl and the ATP-dependent RuBisCO activase (RCA), the malate dehydrogenase (MDH), and subunit PsaG of photosystem I complex as prominent stromal contaminations. For Emb1211, PsaD-2, the beta-subunit of ATP synthase (ATPB), AT1G33810, and the geranyl reductase AT1G74470 a thylakoid localization was determined (Peltier et al., 2004).

As expected, we identified proteins of the preprotein translocon of the inner membrane (TIC; Soll and Schleiff, 2004), namely Tic110, Tic55-II, Tic55-IV, Tic40, and Tic32-IVb as major components of the IE fraction. Although assigned to category II we identified IE membrane-associated cpHsp70 (two peptides; Su and Li, 2010) and CPN60 (two peptides; Stürzenbaum et al., 2005), which were two chaperones previously discussed to be involved in preprotein import. We also detected two peptides for intermembrane space localized Tic22-IV. Remarkably, the chloroplast-targeted

Table 5 | Mixed envelope proteins with known function.

	AGI	Abbr.	Name and function	TM fold	Other Loc.	Studies	Cat.
Preprotein import	At5g19620	Toc75-V	Protein translocation channel at OEM	β-Barrel	–	5	I
Lipid biosyn.	AtCg00500	ACCD	Carboxyltransferase beta-subunit of acetyl-CoA carboxylase complex	None	–	1	III
Transport	At1g01790	KEA1	Putative potassium cation efflux antiporter	13 α-Helical TM	–	6	I
	At2g36830	TIP1.1	Putative tonoplast intrinsic protein	6 α-Helical TM	n.d./–	3	I
	At3g53420	PIP2A	Putative plasma membrane intrinsic protein 2a	6 α-Helical TM	–	2	I
	At4g16160	Oep16-2	Putative plastid outer envelope protein	β-Barrel	–	1	III
	At4g20260	PCaP1	Mediates hypocotyls cell elongation	None	–	2	I
	At4g39460	SamC1	*S*-Adenosylmethionine transporter	5 α-Helical TM	–	5	I
	At5g13450	ATP5	Mitochondrion MF1-ATP synthase subunit	None	X	3	I
SR	At1g15690	AVP-3	Type I proton-translocating pyrophosphatase	14 α-Helical TM	–	4	I
	At1g55020	LOX1	Lipooxygenase	1 α-Helical TM	n.d.	1	III
	At3g09260	PYK10	Beta-glucosidase	1 α-Helical TM	X	1	III
	At3g14210	ESM1	Putative GDSL-type lipase	1 α-Helical TM	n.d./–	2	I
Others	At1g55860	UPL1	Putative ubiquitin-protein ligase 1	1 α-Helical TM	n.d./X	1	III
	At1g80370	CYCA2_4	A-type cyclin	1 α-Helical TM	n.d./X	1	III
	At2g38040	CAC3	Alpha subunit of acetyl-CoA carboxylase complex	None	–	6	I
	At4g19170	CCD4	Putative carotenoid cleavage dioxygenase	None	–	1	III
	At4g22710	CYP706A2	Cytochrome P450 monooxygenase	1 α-Helical TM	n.d./–	1	III
	At5g25980	TGG2	Thioglucoside glucohydrolase	1 α-Helical TM	n.d./–	2	I
Other Organelle	At1g78900	VHA-A	A-type subunit of vacuolar H(+)-ATPase peripheral V1 subcomplex	None	n.d./–	4	I
	At2g38670	PECT1	Phosphoethanolamine cytidylyltransferase	1 α-helical TM	n.d./X	2	I
	At5g15920	SMC5	Putative SMC5-like component of chromosome metabolism	None	n.d.	1	III

Given is the functional pathway or the organellar compartment, the AGI number, the short name (Abbr.), the (putative) function, the transmembrane anchor architecture, other localization by the PPDB (Sun et al., 2009) and SUBAII (Heazlewood et al., 2007; n.d., not defined, –, no other localization, X, other localization), the number of studies where the protein was identified (our study; Ferro et al., 2003; Froehlich et al., 2003; Bräutigam et al., 2008; Bräutigam and Weber, 2009; Ferro et al., 2010), and the category from our study. Signaling and response (SR).

ferredoxin-NADP(+)-oxidoreductase FNR1 (**Table 4**), which was found to be associated with the IE via interaction with Tic62 before (Küchler et al., 2002), was clearly detected, whereas Tic62 was identified by only one peptide. Similarly, for Tic20 we found only a single peptide as well. The absence or the low coverage of the membrane-inserted TIC proteins might reflect the problems of analyzing membrane proteins in general (Eichacker et al., 2004).

Besides the TIC components, we identified Iep37, which is described as an IE protein involved in Polyquinone biosynthesis (Dreses-Werringloer et al., 1991). Similarly, the cell growth defect factor Cdf1 (Kawai-Yamada et al., 2005), which is able to induce apoptosis when expressed in yeast, was found to be localized in the IE of chloroplasts (Ladig et al., 2011). Sulfoquinovosyldiacylglycerol (SQDG) synthesis occurs in envelope membranes (Seifert and Heinz, 1992) and here identified SQDG synthase (SQD2; Yu et al., 2002) was localized in chloroplasts. Similarly, we detected the stromal FAD8 (Matsuda et al., 2005) involved in lipid desaturation, and TGD2 involved in transport of lipids from the ER to chloroplasts (three peptides; Awai et al., 2006).

Further, we detected the ATP/ADP antiporter of the IE (NTT1; Neuhaus et al., 1997), the preprotein and amino acid transporter family protein Prat2.2 (Murcha et al., 2007) and the potassium cation efflux antiporter KEA2 (Zybailov et al., 2008) with at least four peptides. In addition, one peptide each was found for the putative magnesium cation transporter MGT10 (Froehlich et al., 2003), for the triose-phosphate/phosphate translocator TPT (Schneider et al., 2002), for the mitoferrin-like carrier MFL1 (Tarantino et al., 2011), for the plastidial sodium-dependent pyruvate transporter BAT1 (Furumoto et al., 2011), and two peptides for the plastidic glutamate/malate-translocator (DIT2; Renné et al., 2003), the putative sugar transporter encoded by AT5G59250 (Froehlich et al., 2003), as well as three peptides for the plastidic 2-oxoglutarate/malate-translocator (DIT1; Weber et al., 1995), and for the maltose transporter Mex1 (Niittylä et al., 2004).

The beta-carbonic anhydrase (CA1; Fabre et al., 2007) of category II and the three metalloproteases (category I; FtsH4i, FtsH11, and FtsH12), detected in the IE fraction, were previously allocated to the stroma (Sakamoto et al., 2003), but they might be associated with the IE as well as suggested for Emb2458

Table 6 | Proteins from Category I and III with unknown function.

Loc.	AGI	Putative function[A] or Closest homologue[B] or GO annotation[C]	Putative TM fold	Studies	Cat.
Inner envelope membrane	At1g33810	TIM phosphate binding super family[A]	1 α-Helical TM	5	I
	At1g42960	Glutaredoxin2 C[A]	1 α-Helical TM	6	I
	At2g35800	Mitochondrial glutamate carrier (*M. truncatula*)[B]	2 α-Helical TM	4	I
	At2g36570	Leucine-rich repeat receptor-like proteinkinase (Tyrosin kinase)[A,B]	1 α-Helical TM	1	III
	At2g38550	Transmembrane proteins 14C (*H. sapiens*)[B]	4 α-Helical TM	6	I
	At3g02900	YCF1.2 protein[A]	1 α-Helical TM	4	I
	At3g10840	Putative alpha/beta-fold-type hydrolase[A,B]	2 α-Helical TM	4	I
	At3g32930	Pterin 4 alpha carbinolamine dehydratase[A]	None	4	I
	At3g54390	Putative DNA-binding protein (SWI3, ADA2, N-CoR, and TFIIIB)[A]	None	1	III
	At4g13590	PF27 (small family from bacteria and eukaryotes) belongs to the lysine exporter superfamily)[A]	7 α-Helical TM	4	I
	At5g03900	Iron-sulfur cluster biosynthesis family protein[A]	2 α-Helical TM	4	I
	At5g08540	Oligonucleotide/oligosaccharide binding DNA ligase family[A]	1 α-Helical TM	6	I
	At5g12470	DUF3411 and Glycine rich proteien family[A]	4 α-Helical TM	6	I
	At5g59250	Putative sugar transporter[A/]D-xylose-proton symporter-like protein[B]	11 α-Helical TM	6	I
	AtCg01130	Ycf1 protein[A/]Ycf1 (*B. napus*)[B]	8 α-Helical TM	3	I
Outer envelope membrane	At1g07930	Putative elongation factor Tu GTP-binding protein[A,B]	None	1	III
	At1g09920	Putative PRLI-interacting factor K[A]	1 α-Helical TM	1	III
	At1g27300	Protein binding[C]	1 α-Helical TM	1	III
	At1g68680	Putative mitochondrial respiratory chain complex I[C]	2 α-Helical TM	1	III
	At1g70480	Putative steroidogenic acute regulatory protein; lipid transfer protein[A]	None	1	III
	At2g24440	Putative protein that binds to UDP-glucose:glycoprotein glucosyltransferase[A]/selT/selW/selH selenoprotein (*Z. mays*)[B]	None	1	III
	At2g32240	Putative SMC protein that bind DNA and act in organizing and segregating chromosomes for partition[A]	1 α-Helical TM	2	I
	At3g26740	CCL putative light regulated protein[A,B]	None	1	III
	At3g49350	Putative GTPase activator protein of Rab-like small GTPases[A,B]	None	1	III
	At3g52230	Chloroplast outer envelope 24 kD protein like (omp24)[B]	1 α-Helical TM	6	I
	At3g53560	Putative protein protein interaction function[A]	None	2	I
	At3g63170	Putative chalcone-flavonone isomerase[A,B]	1 α-Helical TM	4	I
	At4g16450	Putative NADH-ubiquinone oxidoreductase complex I[A,B]	1 α-Helical TM	3	I
	At4g27680	Putative ATPases associated [A/]Spastin (*M. truncatula*)[B]	1 α-Helical TM	1	III
	At4g27990	YLMG1-2 (YGGT family protein) YGGT repeat found in conserved hypothetical integral membrane proteins[A]	3 α-Helical TM	5	I
	At4g32250	Putative Serine/Theronine protein kinase[A/]G protein-coupled receptor kinase (*M. truncatula*)[B]	1 α-Helical TM	3	I
	At5g16870	Putative peptidyl-tRNA hydrolase (PTH2)[A,B]	1 α-Helical TM	1	III
	At5g21920	YLMG2 (YGGT family protein) YGGT repeat found in conserved hypothetical integral membrane proteins[A]	2 α-Helical TM	1	III
	At5g27330	Putative SMC protein that bind DNA and act in organizing and segregating chromosomes for partition[A]	1 α-Helical TM	2	I
	At5g64816	Not determined	1 α-Helical TM	3	I
Mixed envelope membrane	At1g16790	Ribosomal protein related[C]	None	5	I
	At1g51400	Photosystem II reaction center subunit T[A,B]	None	1	III
	At1g75690	Putative chaperone protein dnaJ-related[A]	1 α-Helical TM	1	III
	At1g76030	V-type (H +)-ATPase V1 like protein[A,B]	None	2	I
	At1g76405	Putative Cupin-like protein[A]/Outer envelope pore protein of 21 kDa (*P. sativum*)	None	3	I
	At2g45460	FHA Putative nuclear signaling domain (FHA)[A]	None	1	III
	At2g47840	Chloroplast protein import component Tic20/Ycf60 famly protein[A,B]	3 α-Helical TM	6	I
	At4g14500	Putative steroidogenic acute regulatory lipid transfer protein[A,B]	3 α-Helical TM	1	III
	At4g24750	Rhodanese homology domain[A]	1 α-Helical TM	1	III

(Continued)

Table 6 | Continued

Loc.	AGI	Putative function[A] or Closest homologue[B] or GO annotation[C]	Putative TM fold	Studies	Cat.
	At4g26910	Putative 2-oxoglutarate dehydrogenase E2 component[A,B]	None	1	III
	At5g02940	DUF1012[A/]Os03g163100 (*O. sativa*)[B]	3 α-Helical TM	4	I
	At5g16660	Apolipoprotein domain[A/]Os12g0583400 (*O. sativa*)[B]	1 α-Helical TM	4	I
	At5g23890	Os03g0862100 (*O. sativa*)[B]	1 α-Helical TM	6	I
	At5g37360	Ammonium transporter AES61175.1 (*M. truncatula*)[B]	2 α-Helical TM	1	III
	At5g44960	Putative F-box like protein associated in nuclear processes[A]	None	1	III

Given is the fraction the protein was localized in (outer/inner/mixed envelope membrane), the AGI number, putative function or functional domain indicated by [A] or closest homolog indicated by [B] or function annotated in GeneOntology (GO) [C], the transmembrane architecture, the number of studies where the protein was identified with (our study; Ferro et al., 2003; Froehlich et al., 2003; Bräutigam et al., 2008; Bräutigam and Weber, 2009; Ferro et al., 2010), and the category.

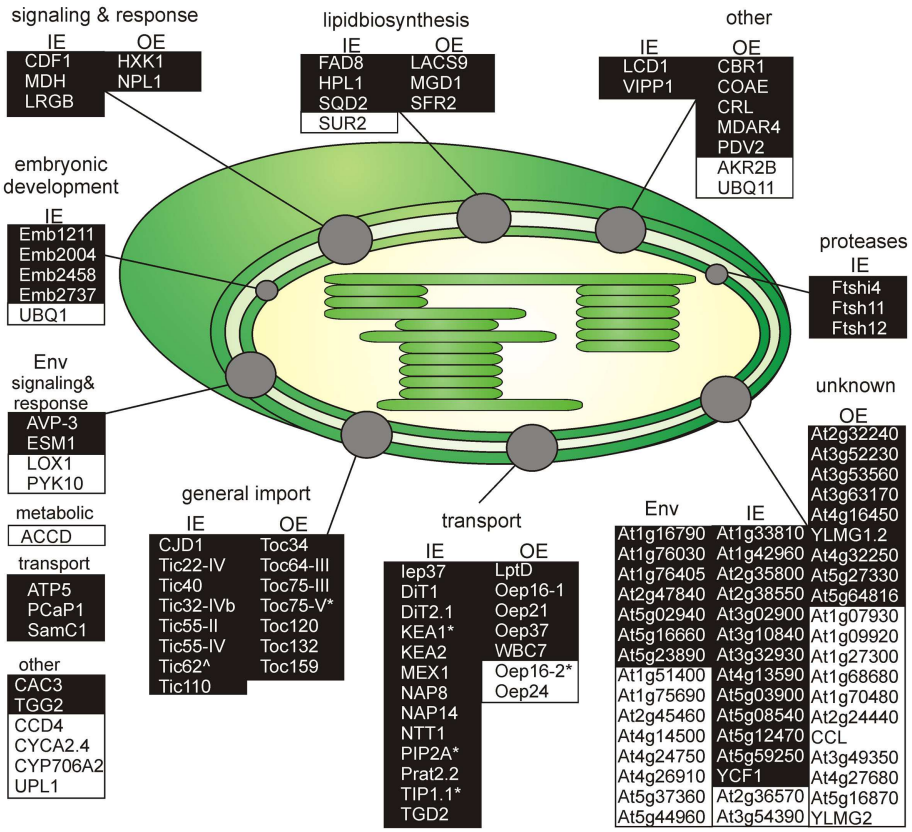


FIGURE 3 | Functional categorization of detected outer or inner envelope proteins. Shown are all proteins clustered by their known functions (**Tables 3–5**). The proteins of each cluster are sorted by the localization in the inner (IE) or in the outer envelope (OE) fraction. Proteins with an unclear localization (ENV) are listed in a separated cluster. Also proteins with unknown function (**Table 6**) are listed in a separated cluster. Black boxes indicate proteins of category I and white boxes proteins of category III (**Table 1**). * Indicates not found in ENV but clearly localized to IE or OE known from literature. ^Indicates localized to IE or OE but only detected with one peptide.

(Froehlich et al., 2003). The same holds true for the DnaJ-like membrane protein of unknown function (CJD1; Zybailov et al., 2008). The tocopherol cyclase SXD1 (category II; Provencher et al., 2001) is chloroplast-localized and is involved in tocopherol synthesis, which takes place in the IE membrane (Lichtenthaler et al., 1981). Thus, it is most likely that these six proteins are membrane-associated and correctly assigned to the IE membrane. The proteins encoded by AT1G33810, AT1G42960, AT2G35800, AT2G38550, AT3G02900, AT3G10840, AT3G32930, AT4G13590, AT5G03900, AT5G08540, and AT5G12470 are assigned as (inner) envelope proteins (Ferro et al., 2003, 2010; Froehlich et al., 2003; Bräutigam et al., 2008; Bräutigam and Weber, 2009), but their function remains to be explored. We further confirmed the IE localization of the plastid-encoded Ycf1.2 (Ladig et al., 2011). The latter might be inserted by the recently identified Sec translocon (Skalitzky et al., 2011).

Non-assignable and unknown proteins

Next to the proteins with known functions that could be clearly assigned to the OE/IE membrane in *P. sativum*, we identified two additional classes of proteins. The first are proteins that have a known function but could not clearly be allocated to either of the membranes (**Table 5**), because these proteins were found only in the mixed envelope of *A. thaliana* and/or *M. sativa*. Most of these proteins function as transporters like KEA1, TIP1.1, PIP2A, Oep16-2, PCaP1, and SamC1 or signaling and response (AVP-3, LOX1, PYK10, and ESM1). Toc75-V (Schleiff et al., 2003a) was the only preprotein import protein, which could be identified in the mixed envelope fraction but not in the OE or IE membrane of *P. sativum*.

The second are proteins of which neither function nor localizations are known yet (**Table 6**). These proteins were assigned concerning their identification in OE or IE membrane of *P. sativum*. Two of 15 IE-assigned proteins of unknown function (At2g36570, At3g54390) were only detected in our study, whereas ∼50% of OE-assigned unknown proteins are of category III. To characterize the unknown proteins of the two groups and support them as potential new IE/OE envelope proteins we used TOPCONS single (**Figure A2** in Appendix, Hennerdal and Elofsson, 2011) and Aramemnon (Schwacke et al., 2003) for secondary structure prediction. Eighty-five percent of the unknown IE proteins possess at least one predicted transmembrane helix (**Table 6**) and might therefore be anchored or embedded into the IE membrane. None of the unknown OE proteins are found to be β-barrel structures, which would have been an argument for an OE localization (Schleiff et al., 2003a) However, it has to be taken into account, the prediction of eukaryotic β-barrel proteins is not as reliable as of helical proteins (Mirus and Schleiff, 2005). Also, the putative function via Pfam (Finn et al., 2010) and CDD (Marchler-Bauer and Bryant, 2004) and the closest homolog via reciprocal best BLAST hit search were predicted to allocate the proteins correctly (**Table 6**). Interestingly, most of the unknown proteins assigned to the IE are localized via PPDB and SUBAII to the plastid except of At2g36570 (other localization) and At3g54390 (not determined), whereas most of the OE-assigned proteins are not determined at least by one database and only six proteins are localized in the plastid (At3g26740, At3g52230, At3g53560, At3g63170, At4g27990, and At4g32250).

EXPERIMENTAL SECTION

ISOLATION AND FRACTIONATION OF CHLOROPLASTS

Arabidopsis thaliana

Chloroplasts were isolated from 20-day-old *A. thaliana* plants (Col-0 ecotype Columbia; 8 h light/16 h dark photoperiod of 120 μmol m^{-2} s^{-1}; 25°C). Plants were harvested before light onset and all procedures were carried out at 4°C. Leaves were cut and homogenized in 450 mM Sorbitol, 20 mM Tricin-KOH pH 8.4, 10 mM EDTA, 5 mM $NaHCO_3$, 1 mM PMSF, using a waring blender (four pulses: low speed 3 s; medium speed 3 s; high speed 2 s; low speed 4 s). The homogenate was filtered through four layers of cheesecloth and one layer of miracloth and centrifuged for 5 min at 1,500 × *g* and 4°C. The pellet was resuspended using a paintbrush in 300 mM Sorbitol, 20 mM Tricin-KOH pH 7.6, 5 mM $MgCl_2$, 2.5 mM EDTA, 1 mM PMSF (resuspension buffer), placed on top of percoll gradients by underlying 12 ml of 45% (v/v) Percoll™ with 8 ml of 85% (v/v) Percoll™, and centrifuged for 10 min at 10,000 × *g*. Intact chloroplasts between 40 and 80% (v/v) Percoll™ were collected after removal of broken chloroplasts by water jet pump. Intact chloroplasts were washed twice by centrifugation for 5 min at 1,500 × *g* in resuspension buffer and collected.

Chloroplasts were lysed by resuspension in 10 mM Tris-HCl pH 8.0, 1 mM EDTA, 1 mM PMSF (TE buffer) to a final concentration of 2 mg chlorophyll/ml. The suspension was placed on top of a sucrose step-gradient (2.4 ml 1.2 M; 4 ml 1.0 M; 4 ml 0.45 M sucrose in TE buffer) and centrifuged for 2 h at 125,000 × *g* and 4°C. Chloroplast fractions were recovered by Pasteur pipettes, diluted 1:3 in TE buffer, centrifuged, pooled, and immediately frozen in liquid nitrogen and stored in −80°C.

Pisum sativum

Chloroplast isolation was adapted from Schleiff et al. (2003a,b). Pea (*P. sativum* cv. Arvika) plants were grown for 8 days in a greenhouse (8 h dark/16 h light, 70 μmol m^{-2} s^{-1}; 25°C). Pea leaves were harvested and homogenized in the 330 mM Sorbitol, 13 mM Tris, 20 mM MOPS, 0.1 mM $MgCl_2$, 0.02% (w/v) BSA, 1 mM β-ME, 0.3 mM PMSF using a waring blender (five pulses, low/medium/high/low/medium, all 2 s). The suspension was filtered through four layers of cheesecloth and one layer of miracloth and centrifuged for 5 min at 1,500 × *g* and 4°C. The pellet was resuspended in the remaining buffer, transferred with cut 5 ml-pipette tip on top of Percoll gradients prepared by underlaying 13 ml of 40% (v/v) Percoll™ with 8 ml of 80% (v/v) Percoll™, centrifuged for 10 min at 10,000 × *g* and 4°C. Intact chloroplasts were collected from the phase between 40 and 80% Percoll™ and washed twice in 330 mM Sorbitol, 1 mM β-ME, and 0.3 mM PMSF.

Chloroplasts were osmotically shocked by adding 2.4 M sucrose solution to a final concentration of 0.6 M sucrose and incubation for 10 min in dark, followed by mechanical disruption with 50 strokes in a dounce homogenizer. Solution was mixed with 2.4 M sucrose solution to a final concentration of 1.35 M, overlayed with 10 ml 1.1 M, 10 ml 1.0 M, and 8 ml 0.45 sucrose solutions, respectively. Chloroplast sub-compartments were recovered after centrifugation for 18 h at 125,000 × *g* and 4°C, resuspended 10 mM Tris-HCl pH 8.0, 1 mM EDTA, 1 mM PMSF, and stored in −80°C.

Medicago sativa

Chloroplast isolation, and subsequent fractionation, from Alfalfa seedlings was performed as described for pea chloroplast with the following modifications. Seedlings were grown for 20 days and leaves were homogenized in a waring blender (2 × 3 pulses at low speed for 3 s; at medium speed for 3 s at high speed for 2 s). Further, Percoll™ gradients were prepared by underlying 13 ml of 42% (v/v) Percoll™ with 8 ml of 82% (v/v) Percoll™.

PROTEOME ANALYSIS BY MALDI NANO-LC-MS/MS

Preparation for enzymatic digestion

An amount of 120 μg membranes were washed using 25 mM NH_4HCO_3 pH 8.0 and carbamidomethylated prior to digestion. After 2 min of centrifugation at 12,000 × *g* the supernatant was removed, and the pellet was gently resuspended in 100% (v/v) methanol. Sample reduction with DTT was performed at 56°C for 45 min and 10 μl of a 500 mM iodoacetamide in 25 mM

NH_4HCO_3 solution was used for sulfhydryl alkylation. Following a 10 min period of sonication, the methanol was diluted to 60% (v/v) using 25 mM NH_4HCO_3 buffer. The proteolytic digestion was performed by adding either 2 μg of trypsin (three biological replicates each organism and envelope fraction) or 10 μg of elastase (three biological replicates each organism and envelope fraction) for 16 h at 36°C. Prior to storing at −20°C the peptide-containing sample was centrifuged at 12,000 × *g* for 2 min in order to remove all undigested membranes and finally the supernatant was concentrated to 15 μl.

Mass spectrometry

Extracted peptides were subjected to MALDI nLC-MS/MS. Specifically, extracted peptides were injected into an Easy-nLC from Proxeon Systems (Thermo Fisher Scientific, Dreieich, Germany) using solvent A [8% (v/v) acetonitrile, 0.1% (v/v) trifluoric acid]. Separation was performed on a thermostatic (40°C) custom made C_{18} column (X-Bridge™ BEH 180 C_{18} 300 Å 3.5 μm, 75 μm × 150 mm) at a flow rate of 300 nl/min with increasing acetonitrile concentrations. The linear-gradient profile was used for tryptic peptide digests started with 8–90% solvent B [95% (v/v) acetonitrile, 0.1%(v/v) trifluoric acid] in 75 min, a stagnation at this level for 8 min, followed by a quick decline to 8% in 5 min and finally, an additional 2 min at 8% for column equilibration. In the case of elastase generated peptide mixtures, the linear-gradient profile duration was increased to 105 min. The separated peptides were then mixed on a tee (Upchurch Scientific) with matrix solution supplied by an auxiliary pump (flow rate, 1.0 μl/min). This solution contained 3 mg/ml α-Cyano-4-hydroxycinnamic acid (α-CHCA; Bruker Daltonics, Germany) dissolved in 70% (v/v) acetonitrile, 30% (v/v) H_2O, and 0.1% (v/v) trifluoric acid. The final mixture was directly spotted every 20 s on a blank 123 mm × 81 mm Opti-TOF™ LC/MALDI insert metal target. Subsequent MALDI-TOF/TOF measurements were carried out using the 4800 TOF/TOF Analyzer (Applied Biosystems, Germany). All peptides used for calibration were taken from the Sequazyme™ Peptide Mass Standards kit (Applied Biosystems, Germany). Spectra were acquired in the positive reflector mode between 700 and 4000 *m*/*z* with fixed laser intensity. A total of 750 laser shots per spot were accumulated. The precursor selection for MS/MS was carried out via the software of the instrument to avoid unnecessary multiple selections of identical precursor peptides. Up to 10 precursors per spot were selected for fragmentation each requiring a minimum signal-to-noise ratio of 30. The fragmentation of the selected precursors was performed at collision energy of 1 kV using air as collision gas at a pressure of 1×10^{-6} torr. Depending on the spectral quality, 1250–2500 laser shots were recorded. Potential matrix cluster signals were removed from precursor selection by excluding all masses in the range from 700 to 1400 *m*/*z* having values of 0.030 ± 0.1 *m*/*z* as well as the internal calibrant mass.

DATA ANALYSIS AND PRESENTATION

Format parsing

Mascot generic format (mgf) files were retrieved from each nLC-MALDI MS/MS run (three biological replicates each organism and envelope fraction; Table S14 in Supplementary Material) using the built-in Peaks2Mascot feature, exporting up to 65 peaks per MS/MS spectrum, each requiring a minimum signal-to-noise of 5. MS/MS queries were processed using the Mascot database search engine v2.2.03 (Matrix Science Ltd.). Data were analyzed using the following settings: below 60 ppm MS precursor mass tolerance (except for the OE of *P. sativum* in combination with trypsin which was 90 ppm due to a technical problem with the instrument that day) and below 0.5 Da MS/MS mass tolerance for MALDI-TOF/TOF. For all database searches, the post-translational modifications carbamidomethylation of cysteins and oxidation of methionines were both selected as variable. When tryptic searches were performed, up to three missed cleavages were taken into consideration in combination with a specific cleavage after K and R and not before P. In all elastase searches, the number of missed cleavages was set to the maximum value of 9 and enzyme specificity was set to A, V, L, I, S, and T, but not before P. For all samples, a custom Viridiplantae database was generated from UniProtKB containing 887,260 entries as of March 02, 2011. Additionally, for *P. sativum* and *M. sativa* samples, customized databases containing 79,106 and 47,532 sequences were provided by the EST-library (Franssen et al., 2012) and MT3.0 of the IMGAG[1], respectively. False discovery rates (FDR, Table S14 in Supplementary Material) given are those originating from the internal Mascot decoy database search function. For each nLC-MALDI-MS/MS run and each sample, the ions score cut off was calculated individually as $-10 \log(p)$ with $p = 0.05$ (95% confidence level; Table S14 in Supplementary Material). The Mascot analyses were described in the paper of Rietschel et al. (2009). For multiple fragmentations of identical precursors, due to the reappearance in repetitions, only data from the highest scoring peptide were kept. Significant proteins present in all three triplicates were taken and summarized in one table for each type of experiment. Afterward, these tables of elastase and trypsin treatments, containing non-identical hits and peptides, were fused.

Peptide assignment

Depending on the source the peptides identified by Mascot or Sequest were afterward aligned either to the protein database of TAIR9 (*A. thaliana*[2]), the protein database of MIPS (*M. truncatula*[3]), or the data file of contigs und singlets (*P. sativum*, data file from Franssen et al. (2012) using a standalone version of Blast from NCBI (substitution matrix BLOSUM62 with linear gap penalty). Following criteria were applied: peptides were only assigned to proteins in the database, if (i) they were aligned with an identity of >95% (determined via blastp), (ii) they had no gaps or mismatches except for (iii) a single substitution with amino acid residues with similar qualities (defined by the substitution matrix) or a single undefined amino acid position (declared by X). Short peptides (<11 aa), which were already covered by assigned peptides, were not subject to the previously mentioned criteria. Those short peptides were assigned to the protein, although they were

[1] ftp://ftpmips.gsf.de/plants/medicago/MT_3_0/, International Medicago Genome Annotation Group
[2] ftp://ftp.arabidopsis.org/home/tair/
[3] ftp://ftpmips.gsf.de/plants/medicago/MT_3_0/

not aligned with BLAST, which is insufficiently accurate regarding the assignment of peptides shorter than 11 amino acids. This method was used to reduce redundancy and as a more stringent criterion for the detection of proteins via the predicted peptides of Mascot. Also, we used a single method to assign the different species and databases in the same way under the same parameter settings of BLAST. Additionally, we searched in parallel for the closest homolog of *A. thaliana* in the other species.

The peptides allocated to *P. sativum* or *M. truncatula* are also allocated to the possible orthologs in *A. thaliana*. On the basis of the *A. thaliana* gene identifiers and their allocated peptides, the splice variants of the proteins were merged to a single gene identifier. The next step to reduce the abundance was connecting all gene identifiers with exactly the same allocated peptides. These gene identifiers were summed up and given the name of the gene identifier with the most allocated peptides or the shortest amino acid sequence by identity. In the end gene identifiers with an overlap of allocated peptides were also combined to one gene identifier. The name of the gene identifier was chosen on the basis of the number of uniquely allocated peptides or the length of the amino acid sequence. All proteins with only one allocated peptide were handled as not significant and are listed in Tables S5–S7 in Supplementary Material.

Prediction of outer/inner envelope membrane proteins

All gene identifiers including splice variants and proteins, which could be identified with the allocated peptides were used to predict the envelope membrane proteins. Two different experimental approaches were applied for *P. sativum*. The first approach for mass spectrometry analyses contained purified OE proteins. The other approach contained purified IE proteins. The peptides detected by MS were blasted against a database of contigs and singlets of *P. sativum*. For classification of the detected contigs and singlets to outer or IE proteins, we first had to find orthologs in *A. thaliana*. The contigs of the *P. sativum* database were blasted against the *A. thaliana* protein database and subsequently the best hit was reblasted against the *P. sativum* contigs database to verify the *A. thaliana* protein. The dedicated *A. thaliana* gene identifiers were used for the prediction of the OE and IE membrane proteins. All gene identifiers with at least four assigned peptides were used for the analysis of the membrane protein prediction.

Also the identified gene identifiers were allocated to the subcompartments in the chloroplasts. For this the Plant Proteome Database (Sun et al., 2009) was used, which includes the experimentally annotated localizations of the *A. thaliana* gene identifiers. In the end, the amino acid sequences of the identified proteins in the envelope pools were used to predict transmembrane α-helices via TOPCONS single[4] (Hennerdal and Elofsson, 2011).

Database comparison

The proteins of the three different organisms detected in our envelope studies were compared to previous envelope studies including proteomic data for the membrane envelope of plastids by Bräutigam et al. (2008), Bräutigam and Weber (2009), Ferro et al. (2003, 2010), and Froehlich et al. (2003). Also the detected proteins are categorized concerning their occurrence in the different studies and stroma or thylakoid in this study or the study of Ferro et al. (2010).

Domain and homolog searches, structural predictions

First, the function and the name of the protein represented by the gene identifiers of **Tables 3–6** were looked up in Aramemnon rel. 7.0[5] (Schwacke et al., 2003). Afterwards, the predicted transmembrane fold was annotated. If Aramemnon predicts transmembrane β-barrel structures the sequences of the gene identifiers were used to build 3D models of respective amino acid sequence with the help of alignments to known protein structures via the protein fold recognition server Phyre2 (Kelley and Sternberg, 2009). For the gene identifiers of unknown function, the putative domains were searched using the *P*rotein *fam*ilies database (Pfam; Finn et al., 2010) and the Conserved Domain Database (CDD; Mitra et al., 2007).

CONCLUSION

The determination of subcellular and suborganellar proteomes or alterations thereof (due to, e.g., environmental changes) by mass spectrometry is still limited in respect to protein abundance and sample purity (**Figure 1**), but most likely not by bioinformatic methods used for protein assignment (**Figure A1** in Appendix). The assignment of peptides depends in general on their length and the false positive rate can be regulated by mapping criteria. Unassigned peptides usually observed in such studies can in parts be explained by the stringency of the mapping criteria, but point also toward natural variances at the protein level.

In the study at hand, we performed proteomic analyses of chloroplast envelope membranes from three different plant species. The necessity to sustain proteomic studies on the analyses of different species was formerly shown by the unexpected high diversity of soluble chloroplast proteomes, when comparing data from *A. thaliana* and *P. sativum* (Bayer et al., 2011). The comparison of envelope fractions from different plant species in our study increased the number of detected proteins but did not result in a large intersection of these envelope proteins (**Figure 2**; **Table 2**).

Furthermore, when comparing our findings with previous proteomic envelope approaches, we were able to refine the available proteome data and assign a reliable, comprehensive core proteome. Contrary to expectations, intersection of proteins identified in these studies was rather small (**Table 1**). Altogether, we identified 191 potential envelope proteins (categories I–III). After detecting putative cross-contaminations of stromal and thylakoid proteins the remaining 136 envelope proteins were clustered according to their predicted/confirmed localization and cellular function (**Figure 3**). To this end 35 IE, 24 OE, and 19 known non-assignable envelope proteins were identified. Amongst these UBQ1 and SUR2 as well as AKR2B, UBQ11, Oep16-2, and Oep24 were newly assigned to IE and OE, respectively.

Moreover, we identified 21 new potential envelope proteins of category III of unknown function which might give rise to further analyses. Finally, we observed differences concerning the predicted localizations in the independent studies which point

[4]http://single.topcons.net/

[5]http://aramemnon.uni-koeln.de/index.ep

toward a possible membrane-association or a possible dual or multi-sublocalization inside the chloroplast or cell.

ACKNOWLEDGMENTS

We are grateful to Markus T. Bohnsack for support. The work was supported by grants from the Deutsche Forschungsgemeinschaft SFB807-P17 and from the Volkswagenstiftung to Enrico Schleiff.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at http://www.frontiersin.org/Plant_Proteomics/10.3389/fpls.2013.00011/abstract

Table S1 | Proteins identified in the *A. thaliana* envelope membrane fraction. The first column gives the AGI number, the second the number of identified splice variants, the third the AGI of the splice variants, the fourth the AGI code of similar proteins detected, the fifth column the number of peptides assigned to the protein only, the sixth column the number of peptides additionally assigned to other proteins, and the seventh column a short description of the protein. In the second sheet the AGI number and all identified peptides are listed. Every peptide is identified by MS/MS.

Table S2 | The proteins identified in the *M. sativa* envelope membrane fraction. The first column gives the AGI number, the second column the *Medicago* specific ID, the third the number of identified splice variants, the fourth the AGI of the splice variants, the fifth the AGI code of similar proteins detected, the sixth column the number of peptides assigned to the protein only, the seventh column the number of peptides additionally assigned to other proteins, and the eight column a short description of the protein. In the second sheet the *Medicago* ID number and all identified peptides are listed. Every peptide is identified by MS/MS.

Table S3 | The proteins identified in the *P. sativum* outer envelope membrane fraction. The first column gives the AGI number, the second column the *Pisum* specific ID, the third the number of identified splice variants, the fourth the AGI of the splice variants, the fifth the AGI code of similar proteins detected, the sixth column the number of peptides assigned to the protein only, the seventh column the number of peptides additionally assigned to other proteins, and the eight column a short description of the protein. In the second sheet the *Pisum* ID number and all identified peptides are listed. Every peptide is identified by MS/MS.

Table S4 | The proteins identified in the *P. sativum* inner envelope membrane fraction. The first column gives the AGI number, the second column the *Pisum* specific ID, the third the number of identified splice variants, the fourth the AGI of the splice variants, the fifth the AGI code of similar proteins detected, the sixth column the number of peptides assigned to the protein only, the seventh column the number of peptides additionally assigned to other proteins, and the eight column a short description of the protein. In the second sheet the *Pisum* ID number and all identified peptides are listed. Every peptide is identified by MS/MS.

Table S5 | Peptides identified by analysis of *A. thaliana* fractions not assigned to a protein. The peptide, the type of digestion yielding the peptide and the fraction(s) the peptide was identified in is given in sheet one. In sheet two the *Arabidopsis* ID, the peptide, the type of digestion yielding the peptide, the fraction(s) the peptide was identified in, and the short description of the protein is given for all proteins identified by a single peptide only. In sheet three the Arabidopsis IDs, the peptide, the type of digestion yielding the peptide, and the fraction(s) the peptide was identified in is given for all peptides leading to the identification of multiple proteins.

Table S6 | Peptides identified by analysis of *P. sativum* fractions not assigned to a protein. The peptide, the type of digestion yielding the peptide, and the fraction(s) the peptide was identified in is given in sheet one. In sheet two the *Arabidopsis* ID, the *Pisum* ID, the peptide, the type of digestion yielding the peptide, the fraction(s) the peptide was identified in, and the short description of the protein is given for all proteins identified by a single peptide only. In sheet three the *Arabidopsis* IDs, the *Pisum* IDs, the peptide, the type of digestion yielding the peptide, and the fraction(s) the peptide was identified in is given for all peptides leading to the identification of multiple proteins.

Table S7 | Peptides identified by analysis of *M. sativa* fractions not assigned to a protein. The peptide, the type of digestion yielding the peptide, and the fraction(s) the peptide was identified in is given in sheet one. In sheet two the *Arabidopsis* ID, the *Medicago* ID, the peptide, the type of digestion yielding the peptide, the fraction(s) the peptide was identified in, and the short description of the protein is given for all proteins identified by a single peptide only. In sheet three the *Arabidopsis* IDs, the *Medicago* IDs, the peptide, the type of digestion yielding the peptide, and the fraction(s) the peptide was identified in is given for all peptides leading to the identification of multiple proteins.

Table S8 | List of all identified proteins. The *Arabidopsis* IDs of all proteins identified in this study including those with only one peptide matching are listed. The first column gives the ID, the second column the predicted compartment the protein is supposed to be localized in, the column 3 the Arabidopsis fraction, columns 6 and 7 the two Pisum fractions, and column 8 the Medicago fraction; the last column indicates whether the protein is identified in at least one fraction by more than one peptide (norm) or whether identification occurred by one peptide match only (onehit). The fraction the protein was identified in is marked by X.

Table S9–S12 | List of all proteins in category I. The first column is the AGI identifier, the second column the name and aliases of the protein, and the third column the number of studies, where the protein was identified. Category Ia are proteins found in our study and at least one other study and category Ib are proteins identified not in our study but at least two other studies. Category IIa are proteins found in our study and at least two other studies but also in the stromal or thylakoid fraction. Category IIb are proteins found in three other studies and also in the stromal or thylakoid fraction. Category IIIa are proteins only identified in our study and category IIIb are proteins found only in one study excluding our study. Category IVa and IVb contains proteins identified in the stromal or thylakoid fraction and only in our and less than two other studies (IVa) or in less than three other studies (IVb).

Table S13 | List of overlapping and not overlapping proteins in the Venn diagram. The first column gives the AGI identifier, the second column the name and aliases of the protein, the third column the number of studies where the protein was identified, the fourth column the category of the protein, the columns 5–9 show in which envelope fractions and plant species the proteins could be identified. X, identified; –, not identified.

Table S14 | List of the ions score cutoff and FDR for nLC-MALDI MS/MS. The first column gives the used MS method, the second column the organism and fraction, the third column the restriction enzyme, the fourth column the used database for searching, the fifth column the number of repetition, the sixth column the ions score cutoff in $-10\log(p)$ by $p = 0.05$, and the seventh column the false discovery rate (FDR). The used databases are the UniProtKB, the MT3.0 from IMGAG for *Medicago truncatula*, and the EST-library by Franssen et al. (2012) for *Pisum sativum*.

Tables S15–S57 | Raw data measured by nLC-MALDI MS/MS. Each excel sheet is grouped in two levels. The first level contains information for each identified accession ID. The first column gives the accession (UniProtKB, IMGAG, or EST-library by Franssen et al., 2012), the second column the coverage, the third column the number of peptide spectrum matches (#PSMs), the fourth column the number of peptides, the fifth column the number of amino acids (#AAs), the sixth column the molecular weight (MW in kDa), the seventh column the isoelectric point (pI), the eighth column the score, and the ninth column the description. The second level contains all peptide information for each accession ID. The second column gives the confidence icon (Low; Medium; High), the third column the peptide sequence, the fourth column the protein accessions, the fifth column the number of proteins, the

sixth column the number of protein groups, the seventh column the activation type (Collision Induced Dissociation, CID), the eighth column the modifications, the ninth column the ion score, the 10th column the expectation value (exp. value), the 11th column the delta score (Δscore), the 12th column the rank, the 13th column the identity High, the 14th column the homology threshold, the 15th column the charge, the 16th column the mass to charge ratio in daltons (*m*/*z*), the 18th column the delta mass (ΔM, difference between the theoretical mass of the peptide and the experimental mass of the precursor ion), the 19th column the matched ions, and the 20th column the spectrum file.

REFERENCES

Aebersold, R., and Mann, M. (2003). Mass spectrometry-based proteomics. *Nature* 422, 198–207.

Asano, T., Yoshioka, Y., Kurei, S., Sakamoto, W., and Machida, Y. (2004). A mutation of the CRUMPLED LEAF gene that encodes a protein localized in the outer envelope membrane of plastids affects the pattern of cell division, cell differentiation, and plastid division in Arabidopsis. *Plant J.* 38, 448–459.

Awai, K., Maréchal, E., Block, M. A., Brun, D., Masuda, T., Shimada, H., et al. (2001). Two types of MGDG synthase genes, found widely in both 16:3 and 18:3 plants, differentially mediate galactolipid syntheses in photosynthetic and nonphotosynthetic tissues in Arabidopsis thaliana. *Proc. Natl. Acad. Sci. U.S.A.* 98, 10960–10965.

Awai, K., Xu, C., Tamot, B., and Benning, C. (2006). A phosphatidic acid-binding protein of the chloroplast inner envelope membrane involved in lipid trafficking. *Proc. Natl. Acad. Sci. U.S.A.* 103, 10817–10822.

Bae, W., Lee, Y. J., Kim, D. H., Lee, J., Kim, S., Sohn, E. J., et al. (2008). AKR2A-mediated import of chloroplast outer membrane proteins is essential for chloroplast biogenesis. *Nat. Cell Biol.* 10, 220–227.

Bayer, R. G., Stael, S., Csaszar, E., and Teige, M. (2011). Mining the soluble chloroplast proteome by affinity chromatography. *Proteomics* 11, 1287–1299.

Bräutigam, A., Shrestha, R. P., Whitten, D., Wilkerson, C. G., Carr, K. M., Froehlich, J. E., et al. (2008). Low-coverage massively parallel pyrosequencing of cDNAs enables proteomics in non-model species: comparison of a species-specific database generated by pyrosequencing with databases from related species for proteome analysis of pea chloroplast envelopes. *J. Biotechnol.* 136, 44–53.

Bräutigam, A., and Weber, A. P. M. (2009). Proteomic analysis of the proplastid envelope membrane provides novel insights into small molecule and protein transport across proplastid membranes. *Mol Plant* 2, 1247–1261.

Carrie, C., and Small, I. (2013). A reevaluation of dual-targeting of proteins to mitochondria and chloroplasts. *Biochim. Biophys. Acta* 1833, 253–259.

Dreses-Werringloer, U., Fischer, K., Wachter, E., Link, T. A., and Flügge, U. I. (1991). cDNA sequence and deduced amino acid sequence of the precursor of the 37-kDa inner envelope membrane polypeptide from spinach chloroplasts. Its transit peptide contains an amphiphilic alpha-helix as the only detectable structural element. *Eur. J. Biochem.* 195, 361–368.

Dunkley, T. P. J., Hester, S., Shadforth, I. P., Runions, J., Weimar, T., Hanton, S. L., et al. (2006). Mapping the Arabidopsis organelle proteome. *Proc. Natl. Acad. Sci. U.S.A.* 103, 6518–6523.

Eichacker, L. A., Granvogl, B., Mirus, O., Müller, B. C., Miess, C., and Schleiff, E. (2004). Hiding behind hydrophobicity. Transmembrane segments in mass spectrometry. *J. Biol. Chem.* 279, 50915–50922.

Fabre, N., Reiter, I. M., Becuwe-Linka, N., Genty, B., and Rumeau, D. (2007). Characterization and expression analysis of genes encoding alpha and beta carbonic anhydrases in Arabidopsis. *Plant Cell Environ.* 30, 617–629.

Ferro, M., Brugiere, S., Salvi, D., Seigneurin-Berny, D., Court, M., Moyet, L., et al. (2010). AT_CHLORO, a comprehensive chloroplast proteome database with subplastidial localization and curated information on envelope proteins. *Mol. Cell. Proteomics* 9, 1063–1084.

Ferro, M., Salvi, D., Brugière, S., Miras, S., Kowalski, S., Louwagie, M., et al. (2003). Proteomics of the chloroplast envelope membranes from Arabidopsis thaliana. *Mol. Cell Proteomics* 2, 325–345.

Finn, R. D., Mistry, J., Tate, J., Coggill, P., Heger, A., Pollington, J. E., et al. (2010). The Pfam protein families database. *Nucleic Acids Res.* 38, D211.

Fourrier, N., Bédard, J., Lopez-Juez, E., Barbrook, A., Bowyer, J., Jarvis, P., et al. (2008). A role for SENSITIVE TO FREEZING2 in protecting chloroplasts against freeze-induced damage in Arabidopsis. *Plant J.* 55, 734–745.

Franssen, S. U., Shrestha, R. P., Bräutigam, A., Bornberg-Bauer, E., and Weber, A. P. M. (2012). Comprehensive transcriptome analysis of the highly complex Pisum sativum genome using next generation sequencing. *BMC Genomics* 12:227. doi:10.1186/1471-2164-12-227

Friso, G., Giacomelli, L., Ytterberg, A. J., Peltier, J.-B., Rudella, A., Sun, Q., et al. (2004). In-depth analysis of the thylakoid membrane proteome of Arabidopsis thaliana chloroplasts: new proteins, new functions, and a plastid proteome database. *Plant Cell* 16, 478–499.

Froehlich, J. E., Wilkerson, C. G., Ray, W. K., McAndrew, R. S., Osteryoung, K. W., Gage, D. A., et al. (2003). Proteomic study of the Arabidopsis thaliana chloroplastic envelope membrane utilizing alternatives to traditional two-dimensional electrophoresis. *J. Proteome Res.* 2, 413–425.

Fukuchi-Mizutani, M., Mizutani, M., Tanaka, Y., Kusumi, T., and Ohta, D. (1999). Microsomal electron transfer in higher plants: cloning and heterologous expression of NADH-cytochrome b5 reductase from Arabidopsis. *Plant Physiol.* 119, 353–362.

Furumoto, T., Yamaguchi, T., Ohshima-Ichie, Y., Nakamura, M., Tsuchida-Iwata, Y., Shimamura, M., et al. (2011). A plastidial sodium-dependent pyruvate transporter. *Nature* 476, 472–475.

Glynn, J. M., Froehlich, J. E., and Osteryoung, K. W. (2008). Arabidopsis ARC6 coordinates the division machineries of the inner and outer chloroplast membranes through interaction with PDV2 in the intermembrane space. *Plant Cell* 20, 2460–2470.

Goulas, E., Schubert, M., Kieselbach, T., Kleczkowski, L. A., Gardeström, P., Schröder, W., et al. (2006). The chloroplast lumen and stromal proteomes of Arabidopsis thaliana show differential sensitivity to short- and long-term exposure to low temperature. *Plant J.* 47, 720–734.

Haarmann, R., Ibrahim, M., Stevanovic, M., Bredemeier, R., and Schleiff, E. (2010). The properties of the outer membrane localized lipid A transporter LptD. *J. Phys. Condens. Matter* 22, 454124.

Heazlewood, J. L., Verboom, R. E., Tonti-Filippini, J., Small, I., and Millar, A. H. (2007). SUBA: the Arabidopsis subcellular database. *Nucleic Acids Res.* 35, D213–D218.

Hennerdal, A., and Elofsson, A. (2011). Rapid membrane protein topology prediction. *Bioinformatics* 27, 1322–1323.

James, P. (1997). Protein identification in the post-genome era: the rapid rise of proteomics. *Q. Rev. Biophys.* 30, 279–331.

Karas, M., and Hillenkamp, F. (1988). Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal. Chem.* 60, 2299–2301.

Kawai-Yamada, M., Saito, Y., Jin, L., Ogawa, T., Kim, K.-M., Yu, L.-H., et al. (2005). A novel Arabidopsis gene causes Bax-like lethality in Saccharomyces cerevisiae. *J. Biol. Chem.* 280, 39468–39473.

Kelley, L. A., and Sternberg, M. J. E. (2009). Protein structure prediction on the Web: a case study using the Phyre server. *Nat. Protoc.* 4, 363–371.

Kleffmann, T., Hirsch-Hoffmann, M., Gruissem, W., and Baginsky, S. (2006). plprot: a comprehensive proteome database for different plastid types. *Plant Cell Physiol.* 47, 432–436.

Kleffmann, T., Russenberger, D., von Zychlinski, A., Christopher, W., Sjölander, K., Gruissem, W., et al. (2004). The Arabidopsis thaliana chloroplast proteome reveals pathway abundance and novel protein functions. *Curr. Biol.* 14, 354–362.

Klodmann, J., Sunderhaus, S., Nimtz, M., Jansch, L., and Braun, H. P. (2010). Internal architecture of mitochondrial complex I from Arabidopsis thaliana. *Plant Cell* 22, 797–810.

Küchler, M., Decker, S., Hörmann, F., Soll, J., and Heins, L. (2002). Protein import into chloroplasts involves redox-regulated proteins. *EMBO J.* 21, 6136–6145.

Ladig, R., Sommer, M. S., Hahn, A., Leisegang, M. S., Papasotiriou, D. G., and Ibrahim, M., et al. (2011). A high-definition native polyacrylamide gel electrophoresis system for the analysis of membrane complexes. *Plant J.* 67, 181–194.

Lamesch, P., Berardini, T. Z., Li, D., Swarbreck, D., Wilks, C., Sasidharan, R., et al. (2012). The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. *Nucleic Acids Res.* 40, D1202–D1210.

Leister, D. (2003). Chloroplast research in the genomic age. *Trends Genet.* 19, 47–56.

Lichtenthaler, H. K., Prenzel, U., Douce, R., and Joyard, J. (1981). Localization of prenylquinones in the envelope of spinach chloroplasts. *Biochim. Biophys. Acta* 641, 99–105.

Lisenbee, C. S., Lingard, M. J., and Trelease, R. N. (2005). Arabidopsis peroxisomes possess functionally redundant membrane and matrix isoforms of monodehydroascorbate reductase. *Plant J.* 43, 900–914.

Lister, R., Carrie, C., Duncan, O., Ho, L. H. M., Howell, K. A., Murcha, M. W., et al. (2007). Functional definition of outer membrane proteins involved in preprotein import into mitochondria. *Plant Cell* 19, 3739–3759.

Marchler-Bauer, A., and Bryant, S. H. (2004). CD-Search: protein domain annotations on the fly. *Nucleic Acids Res.* 32, W327–331.

Matsuda, O., Sakamoto, H., Hashimoto, T., and Iba, K. (2005). A temperature-sensitive mechanism that regulates post-translational stability of a plastidial omega-3 fatty acid desaturase (FAD8) in Arabidopsis leaf tissues. *J. Biol. Chem.* 280, 3597–3604.

Matsushima, R., Kondo, M., Nishimura, M., and Hara-Nishimura, I. (2003). A novel ER-derived compartment, the ER body, selectively accumulates a (β-glucosidase with an ER-retention signal in Arabidopsis. *Plant J.* 33, 493–502.

Millar, A. H., Carrie, C., Pogson, B., and Whelan, J. (2009). Exploring the function-location nexus: using multiple lines of evidence in defining the subcellular location of plant proteins. *Plant Cell* 21, 1625–1631.

Mirus, O., and Schleiff, E. (2005). Prediction of (β-barrel membrane proteins by searching for restricted domains. *BMC Bioinformatics* 6:254. doi:10.1186/1471-2105-6-254

Mitra, S. K., Gantt, J. A., Ruby, J. F., Clouse, S. D., and Goshe, M. B. (2007). Membrane proteomic analysis of Arabidopsis thaliana using alternative solubilization techniques. *J. Proteome Res.* 6, 1933–1950.

Moellering, E. R., Muthan, B., and Benning, C. (2010). Freezing tolerance in plants requires lipid remodeling at the outer chloroplast membrane. *Science* 330, 226–228.

Moore, B., Zhou, L., Rolland, F., Hall, Q., Cheng, W.-H., Liu, Y.-X., et al. (2003). Role of the Arabidopsis glucose sensor HXK1 in nutrient, light, and hormonal signaling. *Science* 300, 332–336.

Murcha, M. W., Elhafez, D., Lister, R., Tonti-Filippini, J., Baumgartner, M., Philippar, K., et al. (2007). Characterization of the preprotein and amino acid transporter gene family in Arabidopsis. *Plant Physiol.* 143, 199–212.

Narendra, S., Venkataramani, S., Shen, G., Wang, J., Pasapula, V., Lin, Y., et al. (2006). The Arabidopsis ascorbate peroxidase 3 is a peroxisomal membrane-bound antioxidant enzyme and is dispensable for Arabidopsis growth and development. *J. Exp. Bot.* 57, 3033–3042.

Neuhaus, H. E., Thom, E., Mohlmann, T., Steup, M., and Kampfenkel, K. (1997). Characterization of a novel eukaryotic ATP/ADP translocator located in the plastid envelope of Arabidopsis thaliana L. *Plant J.* 11, 73–82.

Niittylä, T., Messerli, G., Trevisan, M., Chen, J., Smith, A. M., and Zeeman, S. C. (2004). A previously unknown maltose transporter essential for starch degradation in leaves. *Science* 303, 87–89.

Palmieri, L., Arrigoni, R., Blanco, E., Carrari, F., Zanor, M. I., Studart-Guimaraes, C., et al. (2006). Molecular identification of an Arabidopsis S-adenosylmethionine transporter. Analysis of organ distribution, bacterial expression, reconstitution into liposomes, and functional characterization. *Plant Physiol.* 142, 855–865.

Patel, S., Rose, A., Meulia, T., Dixit, R., Cyr, R. J., and Meier, I. (2004). Arabidopsis WPP-domain proteins are developmentally associated with the nuclear envelope and promote cell division. *Plant Cell* 16, 3260–3273.

Peltier, J.-B., Cai, Y., Sun, Q., Zabrouskov, V., Giacomelli, L., Rudella, A., et al. (2006). The oligomeric stromal proteome of Arabidopsis thaliana chloroplasts. *Mol. Cell Proteomics* 5, 114–133.

Peltier, J.-B., Emanuelsson, O., Kalume, D. E., Ytterberg, J., Friso, G., Rudella, A., et al. (2002). Central functions of the lumenal and peripheral thylakoid proteome of Arabidopsis determined by experimentation and genome-wide prediction. *Plant Cell* 14, 211–236.

Peltier, J.-B., Ytterberg, A. J., Sun, Q., and van Wijk, K. J. (2004). New functions of the thylakoid membrane proteome of Arabidopsis thaliana revealed by a simple, fast, and versatile fractionation strategy. *J. Biol. Chem.* 279, 49367–49383.

Provencher, L. M., Miao, L., Sinha, N., and Lucas, W. J. (2001). Sucrose export defective1 encodes a novel protein implicated in chloroplast-to-nucleus signaling. *Plant Cell* 13, 1127–1141.

Renné, P., Dreßen, U., Hebbeker, U., Hille, D., Flügge, U.-I., Westhoff, P., et al. (2003). The Arabidopsis mutant dct is deficient in the plastidic glutamate/malate translocator DiT2. *Plant J.* 35, 316–331.

Rietschel, B., Arrey, T. N., Meyer, B., Bornemann, S., Schuerken, M., Karas, M., et al. (2009). Elastase digests: new ammunition for shotgun membrane proteomics. *Mol. Cell. Proteomics* 8, 1029–1043.

Sakamoto, W., Zaltsman, A., Adam, Z., and Takahashi, Y. (2003). Coordinated regulation and complex formation of yellow variegated1 and yellow variegated2, chloroplastic FtsH metalloproteases involved in the repair cycle of photosystem II in Arabidopsis thylakoid membranes. *Plant Cell* 15, 2843–2855.

Schleiff, E., and Becker, T. (2011). Common ground for protein translocation: access control for mitochondria and chloroplasts. *Nat. Rev. Mol. Cell Biol.* 12, 48–59.

Schleiff, E., Eichacker, L. A., Eckart, K., Becker, T., Mirus, O., Stahl, T., et al. (2003a). Prediction of the plant beta-barrel proteome: a case study of the chloroplast outer envelope. *Protein Sci.* 12, 748–759.

Schleiff, E., Soll, J., Küchler, M., Kühlbrandt, W., and Harrer, R. (2003b). Characterization of the translocon of the outer envelope of chloroplasts. *J. Cell Biol.* 160, 541–551.

Schneider, A., Hausler, R. E., Kolukisaoglu, U., Kunze, R., van der Graaff, E., Schwacke, R., et al. (2002). An Arabidopsis thaliana knock-out mutant of the chloroplast triose phosphate/phosphate translocator is severely compromised only when starch synthesis, but not starch mobilisation is abolished. *Plant J.* 32, 685–699.

Schnurr, J. A., Shockey, J. M., de Boer, G.-J., and Browse, J. A. (2002). Fatty acid export from the chloroplast. Molecular characterization of a major plastidial acyl-coenzyme A synthetase from Arabidopsis. *Plant Physiol.* 129, 1700–1709.

Schwacke, R., Schneider, A., van der Graaff, E., Fischer, K., Catoni, E., Desimone, M., et al. (2003). ARAMEMNON, a novel database for Arabidopsis integral membrane proteins. *Plant Physiol.* 131, 16–26.

Seifert, U., and Heinz, E. (1992). Enzymatic characteristics of UDP-sulfoquinovose-diacylglycerolsulfoquinovosyl transferase from chloroplast envelopes. *Botan. Acta* 285, 197–205.

Shen, G., Kuppu, S., Venkataramani, S., Wang, J., Yan, J., Qiu, X., et al. (2010). ANKYRIN REPEAT-CONTAINING PROTEIN 2A is an essential molecular chaperone for peroxisomal membrane-bound ASCORBATE PEROXIDASE3 in Arabidopsis. *Plant Cell* 22, 811–831.

Skalitzky, C. A., Martin, J. R., Harwood, J. H., Beirne, J. J., Adamczyk, B. J., Heck, G. R., et al. (2011). Plastids contain a second Sec translocase system with essential functions. *Plant Physiol.* 155, 354–369.

Soll, J., and Schleiff, E. (2004). Plant cell biology: protein import into chloroplasts. *Nat. Rev. Mol. Cell Biol.* 5, 198–208.

Stürzenbaum, S. R., Arts, M. S. J., and Kammenga, J. E. (2005). Molecular cloning and characterization of Cpn60 in the free-living nematode Plectus acuminatus. *Cell Stress* 10, 79.

Su, P.-H., and Li, H.-M. (2010). Stromal Hsp70 is important for protein translocation into pea and Arabidopsis chloroplasts. *Plant Cell* 22, 1516–1531.

Sun, Q., Zybailov, B., Majeran, W., Friso, G., Olinares, P. D. B., and van Wijk, K. J. (2009). PPDB, the Plant Proteomics Database at Cornell. *Nucleic Acids Res.* 37, D969.

Tarantino, D., Morandini, P., Ramirez, L., Soave, C., and Murgia, I. (2011). Identification of an Arabidopsis mitoferrin like carrier protein involved in Fe metabolism. *Plant Physiol. Biochem.* 49, 520–529.

van Wijk, K. J. (2000). Proteomics of the chloroplast: experimentation and prediction. *Trends Plant Sci.* 5, 420–425.

Vojta, A., Alavi, M., Becker, T., Hörmann, F., Küchler, M., Soll, J., et al. (2004). The protein translocon of the plastid envelopes. *J. Biol. Chem.* 279, 21401–21405.

Weber, A., Menzlaff, E., Arbinger, B., Gutensohn, M., Eckerskorn, C., and Fluegge, U.-I. (1995). The 2-oxoglutarate/malate translocator of chloroplast envelope membranes: molecular cloning of a transporter containing a 12-helix motif and expression of the functional protein in yeast cells. *Biochemistry* 34, 2621–2627.

Wilkins, M. R., Pasquali, C., Appel, R. D., Ou, K., Golaz, O., Sanchez, J.-C., et al. (1996). From proteins to proteomes: large scale protein identification by two-dimensional electrophoresis and arnino Acid Analysis. *Nat. Biotechnol.* 14, 61–65.

Xu, C., Fan, J., Cornish, A. J., and Benning, C. (2008). Lipid trafficking between the endoplasmic reticulum and the plastid in Arabidopsis requires the extraplastidic TGD4 protein. *Plant Cell* 20, 2190–2204.

Ytterberg, A. J., Peltier, J.-B., and van Wijk, K. J. (2006). Protein profiling of plastoglobules in chloroplasts and chromoplasts. A surprising site for differential accumulation of metabolic enzymes. *Plant Physiol.* 140, 984–997.

Yu, B., Xu, C., and Benning, C. (2002). Arabidopsis disrupted in SQD2 encoding sulfolipid synthase is impaired in phosphate-limited growth. *Proc. Natl. Acad. Sci. U.S.A.* 99, 5732–5737.

Zabrouskov, V., Giacomelli, L., van Wijk, K. J., and McLafferty, F. W. (2003). A new approach for plant proteomics: characterization of chloroplast proteins of Arabidopsis thaliana by top-down mass spectrometry. *Mol. Cell. Proteomics* 2, 1253–1260.

Zybailov, B., Rutschow, H., Friso, G., Rudella, A., Emanuelsson, O., Sun, Q., et al. (2008). Sorting signals, N-terminal modifications and abundance of the chloroplast proteome. *PLoS ONE* 3:e1994. doi:10.1371/journal.pone.0001994

Zybailov, B., Sun, Q., and van Wijk, K. J. (2009). Workflow for large scale detection and validation of peptide modifications by RPLC-LTQ-orbitrap: application to the Arabidopsis thaliana leaf proteome and an online modified peptide library. *Anal. Chem.* 81, 8015–8024.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 09 October 2012; accepted: 15 January 2013; published online: 06 February 2013.

Citation: Simm S, Papasotiriou DG, Ibrahim M, Leisegang MS, Müller B, Schorge T, Karas M, Mirus O, Sommer MS and Schleiff E (2013) Defining the core proteome of the chloroplast envelope membranes. Front. Plant Sci. ***4****:11. doi: 10.3389/fpls.2013.00011*

This article was submitted to Frontiers in Plant Proteomics, a specialty of Frontiers in Plant Science.

Copyright © 2013 Simm, Papasotiriou, Ibrahim, Leisegang, Müller, Schorge, Karas, Mirus, Sommer and Schleiff. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.

APPENDIX

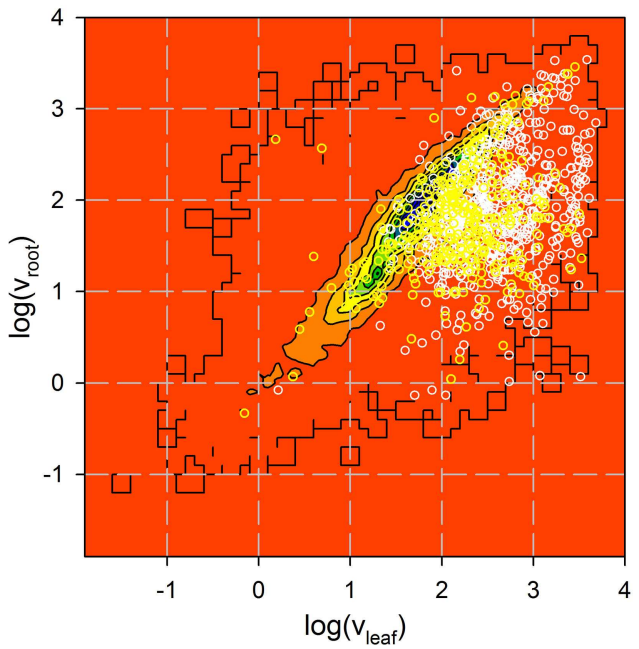


FIGURE A1 | Analysis of expression of genes coding for identified proteins. Expression of 735 genes coding for proteins identified by more than one peptide (white) and 322 genes coding for proteins identified by one peptide only (yellow) in roots and leaves was analyzed by Affymetrix (Vojta et al., 2004). Shown is the Affymetrix value in logarithmic scale for all genes previously analyzed as contour plot (from low (red) to high (blue) number of genes identified with a certain expression distribution) and the distribution of the genes coding for the proteins identified as circles.

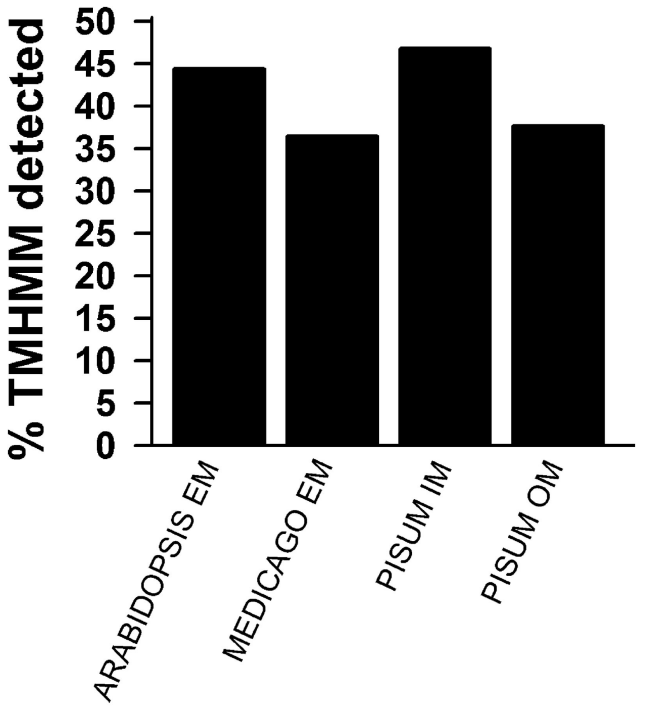


FIGURE A2 | Prediction of membrane localization. TOPCONS single was used to predict transmembrane helices. The percentage of identified proteins for the different pools as described in **Figures 1D,E** is presented. Shown are the results for all proteins identified by more than one peptide (black) and for all identified proteins.

frontiers in
PLANT SCIENCE

ORIGINAL RESEARCH ARTICLE
published: 19 March 2013
doi: 10.3389/fpls.2013.00054

The hydrogen peroxide-sensitive proteome of the chloroplast *in vitro* and *in vivo*

Meenakumari Muthuramalingam[1], Andrea Matros[2], Renate Scheibe[3], Hans-Peter Mock[2] and Karl-Josef Dietz[1]*

[1] *Biochemistry and Physiology of Plants, Faculty of Biology – W5-134, Bielefeld University, Bielefeld, Germany*
[2] *Applied Biochemistry, Institute of Plant Genetics and Crop Plant Research, Gatersleben, Germany*
[3] *Plant Physiology, Faculty of Biology and Chemistry, University of Osnabrück, Osnabrück, Germany*

Edited by:
Harvey Millar, The University of Western Australia, Australia

Reviewed by:
Martin Hajduch, Slovak Academy of Sciences, Slovakia
Georgia Tanou, Aristotle University of Thessaloniki, Greece

Correspondence:
Karl-Josef Dietz, Biochemistry and Physiology of Plants, Faculty of Biology – W5-134, Bielefeld University, 33501 Bielefeld, Germany. e-mail: karl-josef.dietz@uni-bielefeld.de

Hydrogen peroxide (H_2O_2) evolves during cellular metabolism and accumulates under various stresses causing serious redox imbalances. Many proteomics studies aiming to identify proteins sensitive to H_2O_2 used concentrations that were above the physiological range. Here the chloroplast proteins were subjected to partial oxidation by exogenous addition of H_2O_2 equivalent to 10% of available protein thiols which allowed for the identification of the primary targets of oxidation. The chosen redox proteomic approach employed differential labeling of non-oxidized and oxidized thiols using sequential alkylation with *N*-ethylmaleimide and biotin maleimide. The *in vitro* identified proteins are involved in carbohydrate metabolism, photosynthesis, redox homeostasis, and nitrogen assimilation. By using methyl viologen that induces oxidative stress *in vivo*, mostly the same primary targets of oxidation were identified and several oxidation sites were annotated. Ribulose-1,5-bisphosphate (RubisCO) was a primary oxidation target. Due to its high abundance, RubisCO is suggested to act as a chloroplast redox buffer to maintain a suitable redox state, even in the presence of increased reactive oxygen species release. 2-cysteine peroxiredoxins (2-Cys Prx) undergo redox-dependent modifications and play important roles in antioxidant defense and signaling. The identification of 2-Cys Prx was expected based on its high affinity to H_2O_2 and is considered as a proof of concept for the approach. Targets of Trx, such as phosphoribulokinase, glyceraldehyde-3-phosphate dehydrogenase, transketolase, and sedoheptulose-1,7-bisphosphatase have at least one regulatory disulfide bridge which supports the conclusion that the identified proteins undergo reversible thiol oxidation. In conclusion, the presented approach enabled the identification of early targets of H_2O_2 oxidation within the cellular proteome under physiological experimental conditions.

Keywords: chloroplast proteome, hydrogen peroxide, methyl viologen, ribulose-bisphosphate carboxylase, redox regulation

INTRODUCTION

Chloroplasts are essential organelles in plant cells with a wide range of metabolic functions. The redox cascades of the light-driven photosynthetic electron transport chain provide the driving force for metabolism, but they also conditionally generate oxidizing power in the form of reactive oxygen species (ROS). ROS levels increase due to several environmental factors influencing the photosynthetic efficiency, which in turn changes the redox state of the plastid (Foyer and Noctor, 2003). The redox state is also instrumental in regulating the chloroplast metabolic activities but also plastid and nuclear gene expression (Allen, 2003; Pfannschmidt, 2003; Pogson et al., 2008). Thus, chloroplasts serve as an excellent model for better understanding of the redox system which enhances the plant tolerance to environmental stresses. Because of their central role in plant cell signaling, chloroplasts are also considered to function as sensors of environmental fluctuations. According to this scenario, the redox status of chloroplasts is crucial in biological stress response and helps the plant to cope with environmental changes (Scheibe and Dietz, 2012).

Cysteine (Cys) residues in proteins harbor thiol side chains that are highly reactive toward oxidants and can undergo various redox-based modifications. The oxidation of sensitive Cys may cause intra- or intermolecular disulfides. The concomitant conformational changes often regulate the activity and protect the critical thiols against irreversible oxidation (Brandes et al., 2009; König et al., 2012). Reactive thiols can form higher oxidation states like sulfenic acid (SOH) and sulfinic acid (SO_2H), which are reversed by thiol-specific cellular reductants like glutathione, thioredoxin, or sulfiredoxin (Liu et al., 2006). Hyperoxidation describes two forms of Cys oxidation, namely the SO_2H and sulfonic acid (SO_3H) states, the latter one being irreversible to our present-day knowledge. Cys can also form mixed-disulfides

Abbreviations: Cys, cysteine; 2-Cys Prx, 2-cysteine peroxiredoxin; 2DE, two-dimensional gel electrophoresis; DTNB, 5,5-dithio-bis(2-nitrobenzoic acid); DTT, dithiothreitol; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; HCF, high chlorophyll fluorescence; MALDI-TOF, matrix-assisted laser desorption/ionization-time of flight; NEM, *N*-ethylmaleimide; ROS, reactive oxygen species; RubisCO, ribulose-1,5-bisphosphate carboxylase/oxygenase; TCA, trichloroacetic acid.

with glutathione (glutathionylation) or is *S*-nitrosylated. Both modifications receive increasing attention as important redox regulatory mechanism in biology (Mieyal and Chock, 2012). Through these different post-translational modifications Cys appears to be involved in virtually all cellular activities, including immediate metabolic regulation and control of transcriptional and translational activities in development and defense.

Analytical methods have been developed to detect reversible thiol oxidation and they employ combinations of labeling and blocking strategies. The challenge is to identify a few oxidized disulfides among the mass of reduced thiols of a healthy cell. Generally this method includes a saturating blockage of free thiols with thiol reactive reagents, followed by reduction of the disulfides (Muthuramalingam et al., 2010). Subsequently, the newly exposed thiol groups from the reduction step are labeled with detectable thiol-specific reagents. Depending on the specificity of the thiol reductant, this method can be used to identify all reversible thiol modifications (Leichert and Jakob, 2004). Novel labeling techniques based on isotope-coded affinity tags (ICAT) enable quantification of differential protein expression (Sethuraman et al., 2004). The combination of labeling strategies and advanced proteomic methodologies led to the identification of redox proteins which are regulated by thiol-disulfide transitions (Motohashi et al., 2001; Buchanan and Balmer, 2005; Rouhier et al., 2005; Bartsch et al., 2008; Ströher and Dietz, 2008). Often these studies employ rather extreme oxidizing condition. Thus a major open issue concerns the question as to which of the redox-regulated proteins are the primary targets of oxidation and how an initial redox imbalance is sensed in the cells.

Hydrogen peroxide (H_2O_2) is a by-product of normal metabolism and has a sufficient half-life to allow its spreading throughout the entire cell (Bhattachrjee, 2005). H_2O_2 is involved in a number of signaling cascades (Neill et al., 2002) and also in programmed cell death in plants (Levine et al., 1994). A recent study has shown that aquaporins facilitate the movement of H_2O_2 across the membrane (Bienert et al., 2007). As thiols play major roles in ROS-mediated signaling pathways, identification of thiols that are most sensitive to H_2O_2 will help to understand the redox-signaling pathways. Several proteomics studies have addressed the effects of H_2O_2 treatment of seedlings, roots and shoot on proteome composition and carbonylation state of proteins (Tanou et al., 2010; Barba-Espín et al., 2011; Zhou et al., 2011). These studies provide important insight into non-redox effects of H_2O_2 stress and downstream events of H_2O_2-dependent signaling in plants.

In this context, the present study focuses on the identification of chloroplast stroma proteins which are most sensitive to H_2O_2. *Arabidopsis thaliana* stroma proteins were subjected to partial oxidation by exogenous addition of limited amounts of H_2O_2 in order to observe the global response of the chloroplast redox network to an oxidizing stimulus. Initial targets of oxidation were identified by mass spectrometry (MS). In order to confirm the response of identified proteins to oxidation *in vivo*, plants were subjected to methyl viologen (MV) treatment. MV is a redox-active herbicide that accepts electrons at the photosystem I site and produces superoxide through reduction of oxygen within the chloroplasts (Halliwell and Gutteridge, 1989; Jacob and Dietz, 2009). The approach was successfully established and a first list of ROS-sensitive stroma proteins could be provided. Surprisingly, ribulose-1,5-bisphosphate carboxylase oxygenase (RubisCO) proved to be a prominent target *in vitro* and *in vivo*, allowing us to hypothesize on its redox-buffering function during episodes of transient oxidative stress.

MATERIALS AND METHODS

GROWTH OF *ARABIDOPSIS THALIANA* AND CHLOROPLAST ISOLATION

Arabidopsis thaliana (ecotype Columbia) was grown in soil culture with 10 h light/14 h darkness at 22/18°C, respectively, and a photosynthetic photon fluence rate of 120 μmol quanta m^{-2} s^{-1}. 6 week old plants were used for the chloroplast isolation. Leaves were harvested and homogenized in buffer containing 0.3 M sorbitol, 20 mM Tricine/KOH (pH 8.4), 5 mM ethylenediaminetetraacetic acid (EDTA) and 2 mM ascorbic acid. The homogenate was filtered through eight layers of muslin cloth and nylon mesh. The debris was removed by centrifugation at 3000 rpm and 4°C for 2 min. The sedimented chloroplasts were resuspended in isolation buffer, containing 0.33 M sorbitol, 5 mM $MgCl_2$, 20 mM HEPES/KOH (pH 7.9), 2 mM EDTA with freshly added ascorbate. The resuspended chloroplasts were loaded on top of a Percoll step gradient consisting of layers with 40 and 80% Percoll medium containing 0.02 g Ficoll and 0.1 g PEG. The gradient was centrifuged at 3000 rpm for 30 min without brakes. Intact chloroplasts were collected from the interphase between the Percoll layers and washed twice by spinning at 3000 rpm for 2 min. The stroma proteins were extracted following lysis and RubisCO was partially removed according to Ströher and Dietz (2008). The purity of stromal protein preparation was verified by using organelle specific enzymatic and antibody assay (**Figure 1**). The cytosolic marker enzyme UDP-glucose pyrophosphorylase (UGPase) activity was measured according to Zrenner et al. (1993). Using UDP-glucose and pyrophosphate as substrates, Glc-1-P released by UGPase was converted to glucose-6-phosphate (Glc-6-P) which was quantified by coupling to $NADP^+$ reduction by Glc-6-P dehydrogenase. Mitochondrial type II peroxiredoxin F (AtPrxIIF) was used as a marker for mitochondrial contaminations. Equal amounts of total plant and stromal proteins (25 μg) were loaded and separated on reducing SDS-PAGE gels. Western blot analysis with antibodies raised against heterologous expressed AtPrxII F was performed as described in Finkemeier et al. (2005).

DTNB-BASED QUANTIFICATION OF THIOL GROUPS

Total sulfhydryl contents of stroma proteins were determined as described by Tietze (1969). Proteins were precipitated in 3% trichloroacetic acid (TCA) and recovered after a brief centrifugation. To expose buried thiol groups the resulting pellet was dissolved in denaturing buffer containing 100 mM Tris-HCl (pH 8.0), 6 M guanidinium HCl or 1% SDS. The free thiol groups were quantified spectrophotometrically at 412 nm using 6 mM 5,5-dithio-bis(2-nitrobenzoic acid; DTNB) as substrate.

SAMPLE PREPARATION (*IN VITRO* AND *IN VIVO* OXIDATION TREATMENT)

Partial oxidation *in vitro* was performed in 20 mM Tris-HCl (pH 7.8) buffer by adding H_2O_2 in varying stoichiometric quantities equivalent to 1, 2.5, 5, and 10% of the protein thiol content

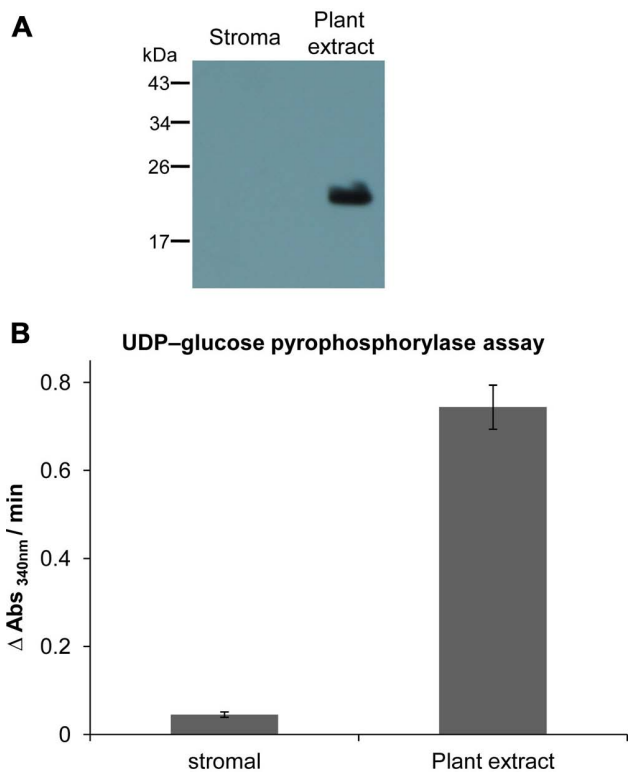


FIGURE 1 | Estimation of cross-contamination of purified chloroplast fraction. (A) Western blot analysis of stroma and plant extract proteins using specific antibody against AtPrxII F (Finkemeier et al., 2005). Equal amounts of protein (25 μg) were loaded in each lane. AtPrxII F was detected as monomer around 21 kDa and was only found in total leaf extract. **(B)** Activity of UGPase as cytosolic marker enzyme, represented as rate of NADPH+H^+ formation in plant extract and stroma fraction. The assay was performed three times. Data express means ± SD.

determined by DTNB assay in the respective fraction. The range of ratios which was usually equivalent to 1–10 μM concentrations was selected to identify preferred sites for oxidation. For *in vivo* oxidation, 50 μM MV supplemented with 0.1% Tween-20 was sprayed on whole plants that were harvested after 10, 30 and 60 min. Three independent experiments were performed for both *in vitro* and *in vivo* oxidation treatments and the identified target proteins are representative of respective replicate experiments.

DETECTION OF REDOX-REGULATED PROTEINS VIA A SEQUENTIAL LABELING STRATEGY

All buffers were depleted from dissolved O_2 by bubbling with argon gas at room temperature (RT). The stroma protein fraction was completely reduced in the presence of 25 mM dithiothreitol (DTT). The reaction was performed inside a closed microaerobiosis chamber continuously flushed with nitrogen gas produced in a nitrogen generator to maintain the oxygen content at less than 0.4%. Excess DTT was removed by desalting using 10 ml desalting columns. The proteins in 20 mM Tris-HCl (pH 7.8) buffer with about 100 μM thiols were treated with H_2O_2 (about 10 μM) for 5 min with gentle shaking to identify the initial targets of oxidation. Remaining cysteinyl thiols were alkylated using 100 mM *N*-ethylmaleimide (NEM) in darkness for 1 h to prevent oxidation. Excess NEM was removed by TCA precipitation according to Muthuramalingam et al. (2010). The washed precipitate was solubilized in denaturing buffer containing 200 mM Bis-Tris (pH 6.5), 6 M urea, 0.5% (w/v) SDS and 10 mM EDTA supplemented with 100 mM DTT to allow full reduction of oxidized thiols. Excess DTT was removed by repeated TCA precipitation. The labeling of the rereduced, previously oxidized thiols was achieved with 25 mM biotin maleimide in the dark for 90 min under constant shaking. Excess labeling reagent was removed either by TCA precipitation or with PD-10 desalting columns (GE Healthcare) depending on the downstream processing.

To monitor the *in vivo* redox status of proteins, proteins from MV-treated plants were extracted in the presence of 100 mM NEM. Furtheron, the proteins were reduced with DTT and subsequently labeled with biotin maleimide as described above. Biotinylated proteins were separated either by one or two-dimensional gel electrophoresis (2DE) and subsequently transferred onto nitrocellulose membrane using the semidry blotter Fastblot B44 (Whatman/Biometra, Germany). After a blocking step with 1% fish gelatin (Sigma, Germany) for 2 h at RT, the membrane was probed with anti-biotin antibody (Clone BN-34 from Sigma–Aldrich, St. Louis, USA). Then the membrane was incubated for 1 h with horseradish peroxidase-conjugated anti-mouse antibody (Sigma–Aldrich, St. Louis, USA) and developed using the enhanced chemiluminescence method (Thermo Scientific, Germany).

STREPTAVIDIN AFFINITY PURIFICATION

The protein samples were desalted to remove labeling solution. This was needed since denaturing reagents used for solubilizing TCA-precipitated protein pellets inhibited binding of biotinylated polypeptides to the streptavidin column. Biotinylated proteins were enriched using streptavidin agarose. The biotinylated sample was incubated with streptavidin agarose equilibrated with phosphate buffered saline (PBS) buffer, pH 7.4, at 4°C with constant shaking overnight. Nonspecifically bound proteins were washed with 1× PBS until the absorbance at 280 nm reached zero. Proteins were incubated with elution buffer containing 1% SDS, 30 mM biotin (pH 12) for 15 min at RT, followed by heating at 96 °C for 15 min.

IDENTIFICATION OF PROTEINS USING MALDI MS ANALYSIS

Proteins purified by streptavidin agarose were resolved by one-dimensional SDS-PAGE and stained with silver nitrate according to Blum et al. (1987). Spots of interest were excised from the gel and placed in U-shaped microtiter plate wells (Greiner Bio-one, Germany). To remove the silver, the excised gel spots were de-stained using Farmer's reducing reagent containing 30 mM potassium ferricyanide (III) and 100 mM sodium thiosulfate. Then the gel spots were washed several times with ultrapure H_2O until the gel slices became transparent. Protein spots were washed twice with 30% (v/v) acetonitrile in 0.1 M ammonium hydrogen carbonate and subsequently dried in the speedvac. The gel pieces were rehydrated in the presence of 0.01 μg trypsin/μl (Promega, Mannheim, Germany) at RT for 30 min followed by overnight incubation at 37°C according to manufacturer's protocol. The gel

slices were vacuum dried, and the peptides were extracted with 50% acetonitrile and 0.1% trifluoroacetic acid for MS analysis. Acquisition of peptide mass fingerprint data and corresponding LIFT spectra was performed using an ultrafleXtreme matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) device (Bruker Daltonics, Bremen, Germany) equipped with a Smartbeam-II laser with a repetition rate of 1000 Hz. The spectra were calibrated using external calibration and subsequent internal mass correction. For databank searching, Biotools 3.2 software (Bruker Daltonics) with the implemented MASCOT search engine (Matrix Science) was used, searching for *A. thaliana* in the non-redundant National Center for Biotechnology Information database (26/07/2010, 55602 sequences). Search parameters were as follows: monoisotopic mass accuracy; 50 ppm tolerance; fragment tolerance of 0.3 Da; missed cleavages 1; and the allowed variable modifications were oxidation (Met), propionamide (Cys), and carbamidomethyl (Cys). Proteins were identified from all three independent experiments applying MASCOT significance scores of 60 (protein level) and 32 (peptide level). Proteins found in just one of the experiments are indicated in table legends (**Tables 1–3**).

BIOTIN QUANTIFICATION ASSAY

The extent of biotinylation was quantified using the HABA-avidin assay developed by Green (1975). HABA forms a red complex with avidin that can be monitored spectrophotometrically at 500 nm. Due to its higher affinity, biotin displaces HABA, accompanied by the decrease in absorbance at 500 nm. The biotinylated protein samples were desalted to remove the excess of biotin maleimide reagent before performing the assay. The assay mixture consisted of 1× PBS buffer containing HABA-avidin reagent (Sigma, Germany) and 10 μg of biotinylated protein sample. After 2 min incubation the absorbance at 500 nm was recorded. The change in absorbance at 500 nm is proportional to the amount of biotin in the assay. A standard curve was generated using free biotin and used to estimate the number of moles of biotin incorporated after biotinylating the protein.

RESULTS

PURITY OF THE CHLOROPLAST FRACTION

Chloroplasts were isolated and lysed to obtain a stromal protein fraction which was checked for contaminations by other cellular constituents. Type II peroxiredoxin F (AtPrxII F) was used as a marker for mitochondrial contamination using Western blot analysis. In total plant protein extract AtPrxII F was detected at the expected size of about 21 kDa, while it was absent in the stromal fraction (**Figure 1A**). As shown in **Figure 1B** the plant protein extract exhibited high rates of nicotinamide adenine dinucleotide phosphate (NADPH) formation at 340 nm, representative for high UGPase activity, while its enzymatic activity in the stromal fraction was minimal with less than 6% relative contamination.

TOTAL PROTEIN THIOL DETERMINATION

The present study aimed to identify the primary protein targets of H_2O_2 oxidation in the chloroplast *in vitro* and *in vivo*. Percoll-purified intact chloroplasts were lysed, and the stroma

Table 1 | *Arabidopsis thaliana* stroma proteins containing H_2O_2sensitive thiols identified by MALDI-TOF/MS.

No.	Protein	Accession number	MM (kDa)	Cys number	Functional role
1	Fd-GOGAT	AT5G04140	165.3	24	Nitrogen assimilation
2	Transketolase	AT3G60750	73.1	8	Photosynthesis
3	RubisCO large subunit	ATCG00490	52.9	9	Photosynthesis
4	Fructose-bisphosphate aldolase-2	AT4G38970	37.9	2	Photosynthesis
5	GAPDH subunit B	AT1G42970	39.3	7	Photosynthesis
6	GAPDH subunit A-2	AT1G12900	37.7	5	Photosynthesis
7	Ferredoxin-NADP(+) oxidoreductase 1	AT5G66190	35.2	5	Photosynthesis
8	Carbonic anhydrase 1	AT3G01500	25.6	6	Photosynthesis
9	2-Cys peroxiredoxin	AT3G11630	22.4	2	Detoxification
10	Cyclophilin 20-3	AT3G62030	19.9	4	Protein folding
11	RubisCO small subunit	AT1G67090	14.7	4	Photosynthesis
12*	Phosphoribulokinase	AT1G32060	39.2	4	Photosynthesis
13*	STN7 kinase	AT1G68830	58.5	5	Serine/threonine kinase
14*	Glutamine synthetase 2	AT5G35630	42.5	6	Ammonia assimilation cycle
15*	ATP synthase beta subunit	ATCG00480	52.5	1	ATP synthesis
16*	HCF 136	AT5G23120	35.8	0	PSII stability complex
17*	DRT 112	AT1G20340	10.5	1	Electron carrier

Each protein is annotated by its name, accession number, molecular weight (without predicted transit peptide), and the biological function. The number of cysteines theoretically present in the mature protein is given. The table compiles the results from three independent experiments. Proteins 1–11 were identified in each experiment. Asterisk "" denotes that these proteins were identified in single experiments where 10% molar equivalence to protein thiols corresponded to 15 μM H_2O_2 concentration.*

Table 2 | Identification of cysteines modified upon oxidation.

Protein identified	Peptide sequence	Predicted mass		Observed mass in H_2O_2 treated
		NEM	Biotin-mal	
RubisCO large subunit	AVYEC*LR	978	1304.5	1304.6*
	VALEAC*VQAR	1184.2	1510.7	1510.7*
Fd-GOGAT	IC*NVDR	843.8	1170.3	1170.6*
	FCTGGMSLGAISR	1424.5	1751	1424.7
	ALYYLCEAADDAVR	1697.7	2024.2	1697.8
	IGFVPEEATIVGNTCLYGATGGQIFAR	2910.1	3234.2	2909.5
Ferredoxin-NADP(+) oxidoreductase 1	C*LLNTK	815.9	1142.4	1142.6
	TVSLCVK	873.9	1200.4	873.5
	GVC*SNFLCDLK	1304.5	1631	1630.8*
GAPDH subunit B	GILDVC*DAPLVSVDFR	1843.9	2170.5	2170*
RubisCO small subunit	LPLFGCTDSAQVLK	1616.6	1943.3	1616.9
	QVQC*ISFIAYKPPSFTG	2011.2	2337.7	2337.2*
	WIPC*VEFELEHGFVYR	2149.3	2475.8	2475.2*

Each entry indicates the peptide sequence and the predicted mass in the presence of the alkylating agents either NEM (+125 Da) or biotin maleimide (+451.5 Da). The mass observed in MALDI-TOF/MS of the H_2O_2-treated sample is shown. Mass of NEM-labeled Cys matches to experimental mass value suggesting that these Cys are not modified during oxidation. C marks Cys in the peptide sequence which is preferentially labeled with biotin maleimide, since the predicted mass matches to the measured value. Here, the predicted mass of biotin maleimide-labeled peptide was not found in the corresponding control sample. These data suggest that the approach used can identify proteins sensitive to oxidation and in addition locates the Cys modification.*

protein fraction was recovered by centrifugation. Total thiol contents of stroma protein extract was determined using the DTNB assay in order to adjust the amount of H_2O_2 to be added for oxidation to 10% of total protein thiols. Protein thiols were quantified under reducing and denaturing conditions to obtain an average amount of thiols to be used as a conversion factor for future experiments. Low molecular weight thiol metabolites such as glutathione were removed by TCA precipitation, followed by a centrifugation. Both denaturing methods, namely guanidinium hydrochloride- and SDS-treatment gave a highly similar result of 57.1 ± 4.3 and 58.3 ± 3.1 μmol/g protein, respectively (mean ± SD of $n = 3$), corresponding to an average of 3 Cys per 50 kDa protein.

In order to optimize the workflow (**Figure 2A**) and to check for reaction specificity, 2-cysteine peroxiredoxin (2-Cys Prx) was used as test protein, since it is a well characterized redox-regulated protein (**Figure 2B**). 2-Cys Prx has two cysteinyl residues and forms an intermolecular disulfide bond upon oxidation. Thus the oxidized form runs as a dimer on non-reducing SDS-PAGE. To check the specificity of the labeling strategy, proteins were directly labeled with biotin maleimide after each step of the work flow and detected immunologically with antibody against biotin. As shown in **Figure 2B**, free thiols were not available for biotin maleimide labeling after oxidation of 2-Cys Prx as indicated by the absence of signal in the Western blot (lane 2). A similar result was observed after blocking of free thiols with NEM, an alkylating agent to prevent thiol-disulfide exchange reactions (lane 3). After reduction, Cys were efficiently labeled with biotin maleimide and strong bands were detected in the blot (lane 1 and 4). The band detected around 48 kDa corresponds to the half-oxidized dimer, since each dimer contains two catalytic sites each of which can form a disulfide bridge.

EFFECT OF H_2O_2-MEDIATED OXIDATION ON STROMA THIOL PROTEINS

Stroma proteins were subjected to H_2O_2 oxidation, subsequently reduced and labeled with biotin maleimide. Oxidation was performed by adding H_2O_2 at amounts of varying stoichiometry relative to protein thiol contents (1, 2.5, 5, and 10%) which corresponded to 1 to 10 μM concentration. Under optimal conditions for photosynthesis, H_2O_2 concentrations are considered to be below 1 μM. H_2O_2 accumulates under stress. 10 μM H_2O_2 inhibits the Calvin cycle in isolated chloroplast by half (Kaiser, 1976; Asada, 1999; Polle, 2001). Remaining free thiols were blocked with NEM followed by reduction of reversibly oxidized proteins with DTT. The newly recovered thiol groups were then labeled with biotin maleimide. Hence biotinylated proteins corresponded to Cys-containing proteins that had been reversibly oxidized by the added H_2O_2. The increase in biotin label corresponding to increased oxidation is shown in **Figure 3A**. In the control reaction, the proteins were reduced and directly blocked with NEM without exposure to H_2O_2. Complete reduction and immediate blocking of free thiols in the control sample resulted in only minor incorporation of biotin maleimide into proteins (**Figure 3A**; lane 1). The labeling degree increased with increasing H_2O_2 concentrations.

The amount of biotin maleimide in the labeled samples was quantified with the HABA-avidin assay. Equal amounts of protein from different H_2O_2-treated and control samples were mixed with HABA-avidin reagent. The assay displayed a decrease in the absorbance that is proportional to the amount of biotin maleimide present in the sample. The degree of biotinylation was calculated as

Table 3 | Oxidation-susceptible proteins in *A. thaliana* treated with methyl viologen, subjected to differential labeling and identified by MS.

No.	Protein name	Accession number	Molecular weight (kDa)	Number of Cys	Localization	Function
1	RubisCO large subunit	ATCG00490	52.9	9	Chloroplast	Photosynthesis
2	Myrosinase	AT5G26000	61.1	9	Vacuole	Defense
3	Chaperonin 60 subunit alpha 1	AT2G28000	62.1	3	Chloroplast	Chaperonin
4	Fructose-bisphosphate aldolase-2	AT4G38970	43.0	2	Chloroplast	Photosynthesis
5	2-Cys peroxiredoxin	AT3G11630	29.1	2	Chloroplast	Detoxification
6	Sedoheptulose-1,7-bisphosphatase	AT3G55800	42.4	7	Chloroplast	Photosynthesis
7	Annexin D1	AT1G35720	36.2	2	Cytoplasm/ Membranes	Detoxification
8	PS II oxygen-evolving complex 1 (PsbO-1)	AT5G66570	35.1	4	Chloroplast	Photosynthesis
9	PS II oxygen-evolving complex 23K protein (PsbP-1)	AT1G06680	28.1	3	Chloroplast	Photosynthesis
10	Glutathione S-transferase F2 (AtGSTF2)	AT4G02520	24.1	0	Cytoplasm	Detoxification
11	Glutathione S-transferase TAU19 (AtGSTU19)	AT1G78380	25.7	1	Cytoplasm	Detoxification
12	Glutathione S-transferase TAU20 (AtGSTU20)	AT1G78370	25.0	2	Cytoplasm	Detoxification
13	Putative plastocyanin (DRT112)	AT1G20340	17.0	1	Chloroplast	Photosynthesis
14	PS II oxygen-evolving enhancer (PSBQ-2)	AT4G05180	24.6	0	Chloroplast	Photosynthesis
15	PS II oxygen-evolving enhancer (PSBQ-1)	AT4G21280	23.9	0	Chloroplast	Photosynthesis
16	RubisCO small subunit	AT1G67090	20.2	5	Chloroplast	Photosynthesis
17	Thioredoxin M1	AT1G03680	19.7	4	Chloroplast	Photosynthesis
18	Putative peroxisomal (S)-2-hydroxy-acid oxidase 2	AT3G14420	40.3	1	Peroxisome	Oxidoreductase
19	NAD(P)-binding Rossmann-fold-containing protein	AT2G37660	34.9	2	Chloroplast	Oxidoreductase
20	Glutathione S-transferase DHAR1	AT1G19570	23.6	2	unclear	Detoxification
21	Ferredoxin-NADP(+) oxidoreductase 1 (FNR1)	AT5G66190	40.3	6	Chloroplast	Photosynthesis
22	Kinesin-5	AT4G05190	89.2	11	Cytoplasm	Microtubule motor
23	Nucleoside diphosphate kinase 2	AT5G63310	25.6	5	Chloroplast	Metabolism
24	Unknown protein (partial)		7.0	4		Probable peroxin

Each protein is annotated by its name, accession number, molecular weight, subcellular localization, and biological function as given in the plant protein database (PPDB). The number of cysteines theoretically present in the mature protein is given. The table compiles the results from three independent experiments. Proteins 3, 7, 10, 12, 18, 20, 22, 23, and 24 were significantly identified in one experiment only.

micromol biotin-labeled thiol per gram protein. As expected the increasing amounts of H_2O_2 and subsequent reduction allowed the incorporation of more biotin, representing the extent of thiol oxidation (**Figure 3B**).

Two-dimensional gel electrophoresis was performed in order to get insight into the complexity of the H_2O_2-mediated oxidative changes. One hundred micrograms of biotin maleimide-labeled control and H_2O_2-treated proteins were separated by 2DE and subsequently visualized by silver staining (**Figure 4A**). The corresponding Western blot membranes probed with anti-biotin antibody are shown in **Figure 4B**. The patterns of biotinylated polypeptides differed strongly between control and treated samples, while the patterns from the silver-stained gel detecting total proteins revealed a similar spot pattern despite the fact that apparently slightly more protein had been solubilized in the H_2O_2-treated sample. The direct comparison of the blots and the silver-stained gels for identifying and excising the proteins of interest appeared unreliable due to the expected background of unlabeled polypeptides. Therefore the biotinylated proteins were further enriched by purification via streptavidin agarose column chromatography.

PURIFICATION AND IDENTIFICATION OF BIOTINYLATED PROTEINS FOLLOWING H_2O_2-MEDIATED OXIDATION

Biotin-labeled control and H_2O_2-treated samples were purified by streptavidin agarose chromatography to separate the H_2O_2-sensitive biotinylated proteins from the complex protein mixture. Affinity-purified proteins were precipitated with TCA (10% w/v) to remove excess biotin from elution. The protein samples were resolved by one-dimensional SDS-PAGE analysis and visualized by silver staining (**Figure 5A**). Both control and H_2O_2-treated proteins exhibited a similar pattern on the gel, which is explained by loading equal protein amounts. However, immuno-reactive signals only appeared on the blot from H_2O_2-treated samples, confirming that the biotin labeling was linked to protein oxidation by H_2O_2 (**Figure 5B**). To identify proteins containing redox sensitive Cys thiols the indicated gel sections were excised from the gel. Results of MALDI-TOF/MS analysis are summarized in

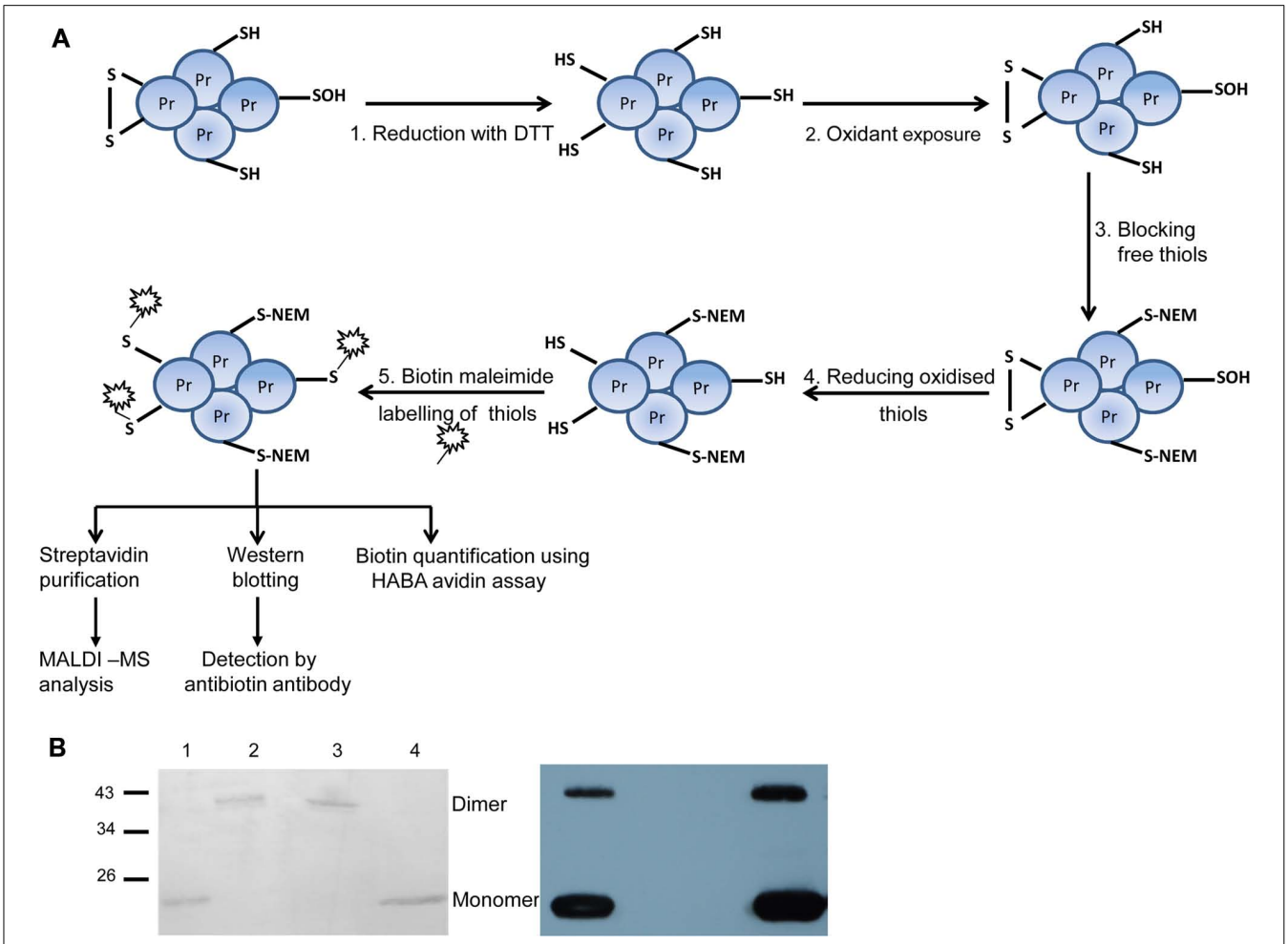


FIGURE 2 | Overview of the procedure for labeling reversibly oxidized cysteines and test experiment with 2-Cys Prx. (A) Work flow used for identifying protein targets sensitive to low H_2O_2 concentrations consisted of five steps: step 1 includes complete reduction of protein thiols using the artificial reductant DTT. Step 2: After desalting to remove excess DTT, proteins were exposed to H_2O_2 for 5 min. In step 3, free thiol groups were blocked using NEM under denaturing conditions. Reversibly oxidized thiols were reduced with DTT in step 4. The newly exposed thiol groups were labeled using the thiol specific reagent biotin maleimide in step 5. As a control reaction, proteins were not subjected to oxidation (step 2) while all other steps were done as described above. From steps 1 to 3, the samples were treated inside the microaerobic chamber with <0.4% O_2. Subsequent steps were performed in ambient air. **(B)** Test of experimental workflow with 2-Cys Prx. Each subunit of the 2-Cys Prx dimer has two cysteines. Biotin labeling of reduced 2-Cys Prx (lane 1); oxidized (lane 2); reduced and blocked with NEM (lane 3), and reduced, oxidized and re-reduced sample (lane 4). On the left is the Ponceau S-stained membrane which shows equal protein amounts loaded. The corresponding blot developed with antibody against biotin is shown on the right and is representative for three experiments.

Table 1. In total 17 proteins were identified from affinity purified H_2O_2-treated samples. All identified proteins are located to the chloroplast. Five proteins function in the Calvin cycle (protein # 2, 5, 6, 11, and 12). The other identified proteins have various functions, such as nitrogen assimilation (proteins #1 and 14), adenosine triphosphate (ATP) synthesis (protein #15) and electron transport (protein # 17) among others. The analysed amino acid composition of these proteins revealed that except for the photosystem II (PSII) stability/assembly factor HCF136 (protein #16) one or more Cys are present in all identified proteins. To confirm oxidation-mediated Cys modification in H_2O_2-treated samples, the mass lists of unmatched peptides were compared with the predicted mass of *in silico* trypsin-digested and biotinylated peptides (**Table 2**). This approach allowed us to confirm two peptides of the large subunit (LSU) and small subunit (SSU) of RubisCO, single peptides of ferredoxin-dependent glutamate synthase (Fd-GOGAT), subunit B of GAPDH and ferredoxin-NADP oxidoreductase (FNR).

PURIFICATION AND IDENTIFICATION OF BIOTINYLATED PROTEINS FOLLOWING MV-MEDIATED OXIDATION

To determine whether and which proteins are oxidized *in vivo*, 6 week old plants were sprayed with MV that induces photo-oxidative stress. The MV treatment of plants revealed a slight increase in biotinylated proteins after different times (**Figure 6A**). Under oxidizing conditions the catalytic Cys of 2-Cys Prx form an intermolecular disulfide bridge, and it runs as dimer at about 43 kDa, whereas the fully reduced form runs as monomer of

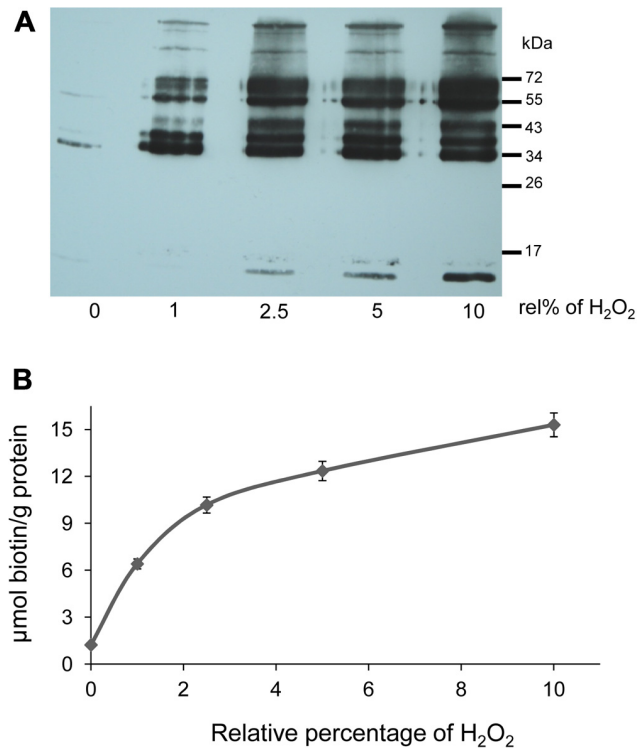


FIGURE 3 | H_2O_2-dependent oxidation of stroma proteins. (A) Western blot analysis with anti-biotin antibody to detect oxidized stroma proteins. "rel%" stands for the amount of H_2O_2 added relative to the thiol contents of the protein sample (100%). The bands demonstrate H_2O_2-dependent oxidation of stroma proteins. Same samples were used for biotin quantification. The experiment with similar results was performed three times. **(B)** Degree of biotinylation in H_2O_2-treated samples as calculated from HABA displacement assay. For biotin quantification, 10 µg of proteins were used. Data are means ± SD, $n = 3$.

22 kDa. At higher oxidant concentrations 2-Cys Prx is prone to overoxidation, which also results in a monomer (data not shown). After 30 min of MV treatment 2-Cys Prx was fully oxidized as shown in the immunoblot analysis (**Figure 6B**). At the later time points the protein was found as monomer again, suggesting overoxidation. Based on the redox behavior of one of the early target proteins, namely 2-Cys Prx, the 30-min exposure time was selected for further experiments.

After 30 min exposure to MV, proteins were extracted in NEM-containing buffer. Extracted and alkylated proteins were reduced, free thiols labeled with biotin maleimide and subsequently purified via streptavidin column. After elution, the proteins were resolved by SDS-PAGE and identified using MS (**Table 3**). In total 24 proteins were repeatedly identified from affinity purified MV-treated samples. Most identified proteins are located to the chloroplast, while some are located to the cytoplasm (protein #7, 10, 11, 12, and 22), the vacuole (protein #2), the peroxisome (protein #18) and the mitochondrion (protein # 20). Some target proteins such as RubisCO, myrosinase, and fructose-bisphosphate aldolase-2 were found both in control and MV-treated plants. Proteins that were differentially oxidized by MV treatment are discussed below. Among these were 2-Cys Prx, sedoheptulose-1,7-bisphosphatase (SBPase), subunits of the water-oxidizing complex and FNR1. Six *in vivo* identified proteins were common with the *in vitro* H_2O_2-treated sample, namely RubisCO LSU and SSU, fructose-bisphosphate aldolase-2, 2-Cys Prx, plastocyanin (DRT 112) as well as FNR. Most identified chloroplast proteins function in photosynthesis (proteins # 8, 9, 14, and 15) and redox homeostasis (proteins # 5, 19, 20, 21, and 24), while others are involved in photorespiration (protein #1), Calvin cycle (protein #6 and 16) and electron transport (proteins #13 and 17).

Except for the glutathione S-transferase (GST) F2 (protein #10), PSBQ-2 (protein #14), and PSBQ-1 (protein #15), one or more Cys are theoretically present in all identified proteins when analyzing the amino acid composition. The oxidation-mediated Cys modifications in MV-treated samples were identified by comparing the predicted mass list of *in silico* trypsin-digested and biotinylated peptides (**Table 4**). This approach allowed us to confirm single peptides of RubisCO LSU, myrosinase, NAD(P)-binding Rossmann-fold-containing protein and FNR.

DISCUSSION

In order to identify stroma protein targets sensitive to oxidation by H_2O_2, this work adopted a strategy similar to the "biotin switch"-method used to detect post-translational *S*-nitrosylation or glutathionylation (Jaffrey and Snyder, 2001; Lind et al., 2002). The present study relies on differential labeling of reduced thiols and formerly oxidized thiols, using two different alkylation reagents (NEM and biotin maleimide) of distinct molecular mass, which eased the preferential identification of the H_2O_2-sensitive proteins. Often redox proteomic studies identify redox-sensitive proteins by direct labeling of proteins during cell extraction, possibly leading to Cys oxidation during the lysis and labeling steps causing false positive results. The cross contamination assays indicate that the isolated chloroplast fractions were highly pure since the mitochondrial protein AtPrxII F was below detection limit (**Figure 1A**) and the cytosolic UGPase constituted less than 6% on a protein basis (**Figure 1B**). In the *in vitro* part of this work, the stroma proteins were completely reduced and then oxidized with low amounts of H_2O_2 to isolate only the most oxidant-sensitive protein thiols. It has been suggested that about 4 µmol H_2O_2 m^{-2}leaf $area^{-1}$ is formed in the chloroplast during photosynthesis under normal conditions (Foyer and Noctor, 2003). This value corresponds to about 300 µmol H_2O_2 L^{-1} stroma produced every second, assuming a leaf chlorophyll content of 300 $mg \cdot m^{-2}$ and a stroma volume of 40 µL mg^{-1} chlorophyll. The H_2O_2-detoxification capacity of the chloroplast ascorbate- and peroxiredoxin-dependent water-water cycles is high, and models predict that resting H_2O_2 concentrations are low as long as reductants are available (Polle, 2001). Here a low H_2O_2 concentration of about 10 µM equivalent to 10% of total protein thiols was used to identify primary targets of oxidation and can be considered to represent physiologically relevant conditions. In previous studies much higher concentrations of oxidants ranging from 1 mM up to 10 mM in non-stoichiometric amounts have been used in plant and animal redox proteomic studies. The results from such studies should be considered with care, since the employed concentrations are

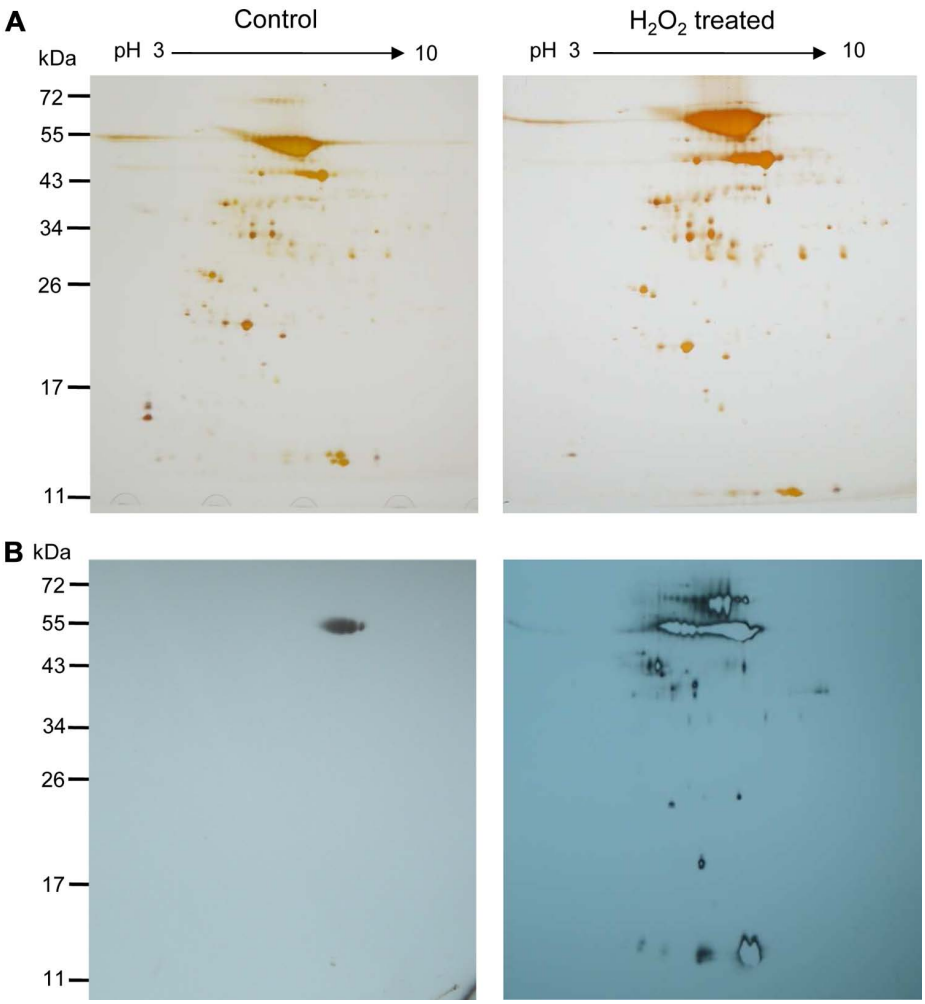


FIGURE 4 | H_2O_2-mediated oxidation of stroma proteins. **(A)** Two-dimensional map of the *A. thaliana* stroma proteins after differential labeling. Separation was performed on 18 cm Immobiline DryStrip (pH 3–10 NL) and by 12% SDS-PAGE as second-dimension. Silver-stained gels of stroma proteins (100 μg) that remained untreated (left) and were treated with H_2O_2 at about 10 μM for 5 min (right) are presented. **(B)** Two-dimensional Western blot map after decoration with anti-biotin antibody. Control and H_2O_2-treated samples of the respective gels are shown. Oxidized protein thiols were labeled with biotin maleimide as described. The figure is representative of three independent experiments. Molecular weight markers, in kDa, are indicated on the left.

often outside the reasonable physiological range (Kim et al., 2000; Baty et al., 2002; Marx et al., 2003; Sethuraman et al., 2004; Winger et al., 2007).

The degree of labeling with biotin-maleimide increased in the H_2O_2-treated sample indicating the occurrence of oxidation-mediated Cys modification in our experiments (**Figure 3A**). Two-dimensional-immunoblots provided insight into the response of the H_2O_2-mediated oxidative changes. Linking our approach with MALDI-TOF analysis enabled the identification of proteins which are most sensitive to oxidation. Most proteins identified possess oxidation-susceptible Cys and are known to undergo dithiol-disulfide exchange reactions as reported in previous studies (Meyer et al., 2005; Ströher and Dietz, 2008; Lindahl and Kieselbach, 2009). All the identified proteins had functional annotations. Based on their function they can be generally categorized as enzymes involved in carbohydrate metabolism, photosynthesis, redox homeostasis, and nitrogen assimilation (**Table 1**).

FUNCTIONAL CLASSIFICATION OF IDENTIFIED PROTEINS

Enzymes such as phosphoribulokinase, GAPDH, and transketolase were identified as primary targets responding to reduction or oxidation of regulatory thiols. They were previously reported to be redox-regulated targets of Trx (Motohashi et al., 2001; Marchand et al., 2004). RubisCO was identified as one of the primary targets of oxidation. RubisCO plays a major role in photosynthesis, hence it is regulated by several mechanisms and redox-dependent modulation is one of them. Based on our MALDI-MS data analysis (**Table 2**) two Cys of RubisCO LSU C192 and C427 are predicted to be oxidized. In *Chlamydomonas reinhardtii* C192 was previously shown to be inactivated by arsenite (Moreno and Spreitzer, 1999). However, site-directed mutagenesis suggested that

C192 lacks a role in disulphide-mediated inactivation. Rather C449 and 459 are assumed to be involved in redox-dependent catalytic inactivation and to trigger increased proteolysis of the protein (Moreno et al., 2008). In this study the identified C427 is conserved in 91% of the analyzed photosynthetic organisms (**Figure A1** in Appendix). In three-dimensional structures C427, 449, and 459 are in close proximity with basic rich amino acids, which could be essential for oxidative activity of the protein. Due to the high protein abundance in the millimolar range, RubisCO can be assumed to act as redox buffer to maintain a suitable redox state of the chloroplast in the presence of transiently increased ROS release. RubisCO SSU was identified as thioredoxin target by several studies (Motohashi et al., 2001; Balmer et al., 2003). In this work C41 and 117 were found to be modified by biotin maleimide due to oxidation, although the distance between both Cys is too far from each other (11–25 Å) to form a disulfide bond (Taylor and Andersson, 1997). But these Cys might be a target for SOH oxidation or mixed-disulfide formation and function as oxidant sensor. In *Chlamydomonas*, Zaffagnini et al. (2012) recently reported 225 glutathionylated proteins, among them many Calvin cycle enzymes, indicating a role of glutathionylation in protecting and regulating chloroplast carbohydrate metabolism. We can exclude glutathionylation, since the proteins were desalted following the reduction with DTT (see Materials and Methods). This step also eliminated any glutathione from the extracts.

In higher plant chloroplasts, Fd-GOGAT is a major enzyme for glutamate synthesis, involved in the conversion of glutamine and 2-oxoglutarate to glutamate. Thioredoxin-mediated redox regulation of Fd-GOGAT was addressed *in vitro* by several studies (Lichter and Häberlein, 1998; Motohashi et al., 2001; Balmer et al., 2003). The amino acid sequence of the protein shows 24 Cys residues in the mature form, three of which were alkylated with NEM (**Table 2**). FNR catalyzes the electron transfer between ferredoxin and NADPH, producing reducing equivalents for chloroplast metabolism and thus represents a crucial enzyme for various pathways requiring reductants. The predicted mass of a peptide containing a Cys residue matched the measured mass value (±0.5 Da) suggesting, that this particular Cys is sensitive to oxidation (**Table 2**). GAPDH is subjected to post-translational modifications which involve thiol-disulfide transitions of regulatory Cys, complex formation with ribulose-5-phosphate kinase and a regulatory protein named CP12 (Scheibe et al., 2002). This mechanism allows the coordination of GAPDH redox regulation with availability of its substrate 1,3-bisphosphoglycerate.

Proof of concept is also provided by the identification of 2-Cys Prx and cylophilin 20-3. 2-Cys Prx is among the top 20 most abundant stroma proteins (Peltier et al., 2006), functions as high-affinity thiol-peroxidase (König et al., 2003), and undergoes large redox-dependent conformational changes linked to functional switches (Dietz, 2012). Here the recombinant 2-Cys Prx protein was a convincing system to test and validate the various steps of oxidation, blocking, labeling, and detection in the work flow (**Figure 2B**), and its recovery *in vitro* and *in vivo* shows also, that the early oxidizable proteins are indeed trapped with this method (**Tables 1–3**). Likewise cylophilin 20-3 is a known target of thiol/disulfide transition (Laxa et al., 2007) and this redox switch affects its peptidylprolyl-cis/trans isomerase activity and probably also its capability for protein/protein interactions.

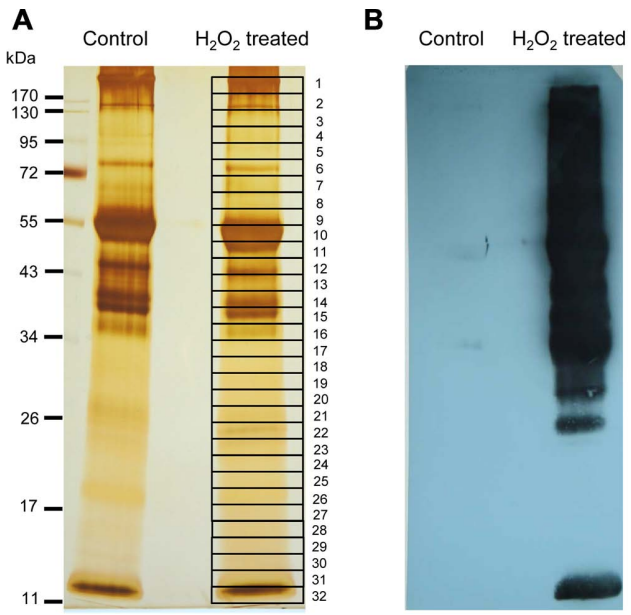


FIGURE 5 | Identification of early targets of H_2O_2-mediated protein oxidation. (A) Silver-stained SDS-PAGE gel showing the biotinylated proteins purified via streptavidin-agarose affinity column. Equal amounts of eluted protein were loaded from control and H_2O_2-treated samples. Band numbered on the gel corresponds to the excised gel sections which were trypsin-digested and analyzed by MALDI-TOF MS. Proteins identified with significant MASCOT score >75 are shown in **Table 1**. At least two peptides were identified per protein. **(B)** Western blot visualization of biotinylated proteins with anti-biotin antibody. Equal amounts of protein (100 μg) were loaded in case of the untreated control and H_2O_2-treated samples. The result is representative of three independent experiments.

REDOX MODIFICATION OF PROTEINS UPON OXIDATIVE STRESS *IN VIVO*

The *in vivo* redox status of protein thiols was monitored using MV-mediated oxidative stress. The majority of the identified proteins are involved in antioxidant defense (**Table 3**). GST catalyzes hydroperoxide detoxification in the presence of glutathione. GST tau is also known to detoxify herbicides and plays a role in signal transduction (Neuefeind et al., 1997; Dixon et al., 2003). Thiol-mediated light/dark regulation of Calvin cycle enzymes is a well known process for years (Buchanan, 1980). In this line SBPase and fructose-bisphosphate aldolase were identified. SBPase is activated by disulfide reduction. The involved Cys have been identified by site-directed mutagenesis (Dunford et al., 1998). Interestingly overexpression of SBPase stimulates growth during early development and under stress suggesting that SBPase activity controls fluxes in the Calvin cycle to a major extent (Lefebvre et al., 2005; Feng et al., 2007). The sensitivity of SBPase to oxidation *in vivo* may provide an explanation why SBPase plays a particular role under stress with increased ROS production.

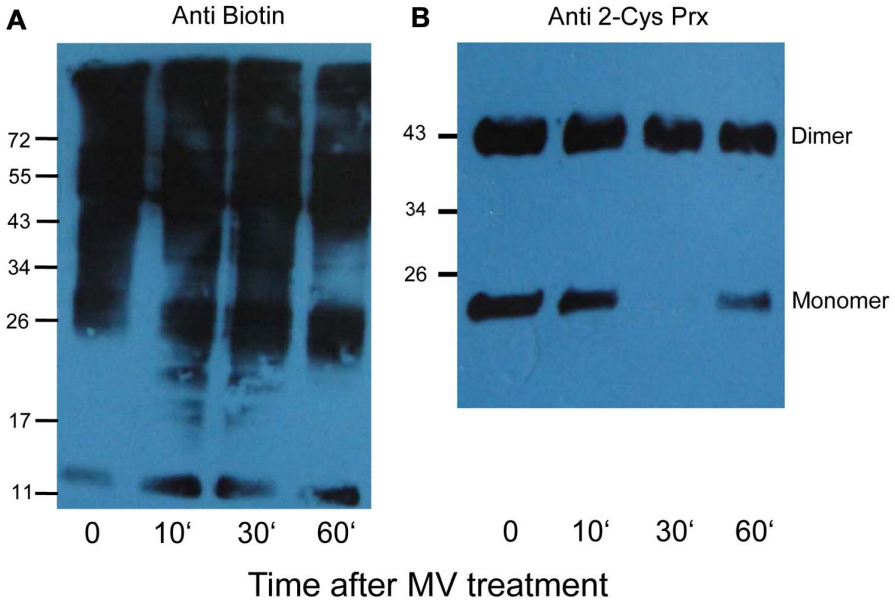


FIGURE 6 | Biotinylated proteins from *A. thaliana* leaves after methyl viologen treatment. (A) The immunoblot shows increased labeling in protein extracts after treating plants with 50 μM MV for different time periods. Control plants (0) were sprayed with water. All samples were extracted in the presence of NEM to block free thiols, then reduced with DTT, and finally labeled with biotin maleimide. **(B)** The same samples were probed with 2-Cys Prx antibody to determine the redox state of 2-Cys Prx. The oxidized dimer and reduced monomers are indicated.

Table 4 | Cysteines modified upon methyl viologen-mediated oxidation.

No.	Protein name	Peptide sequence	Predicted mass		Observed mass in MV treatment
			NEM	Biotin maleimide	
1	RubisCO large subunit	AVYEC*LR	978.0	1304.5	1304.61*
2	Myrosinase	C*SPKIDVR	1113.5	1440.0	1439.69*
19	NAD(P)-binding Rossmann-fold-containing protein	SLVSDSTSICGPSKFTGK	1938.9	2265.4	1938.1
21	Ferredoxin-NADP(+) oxidoreductase 1 (FNR1)	C*LLNTK	815.9	1142.4	1142.57*

The table shows the peptide sequence and the predicted mass in the presence of the alkylating agents either NEM or biotin maleimide. The mass observed in MALDI-TOF MS of the MV-treated sample is shown. NEM-labeled cysteine (C) mass matches to experimental mass value suggesting that these Cys are not modified during oxidation. C denotes Cys in the peptide sequence which is labeled with biotin maleimide, since the predicted mass matched to the experimentally observed value. The underlined peptides were addressed to be modified also in in vitro oxidation treatment (refer **Table 2**).*

Regulators of PS II oxygen-evolving complex (OEC) such as PsbO-1, PsbP-1, and PsbQ were also identified as oxidation-sensitive proteins. A recent study has proposed the presence of an intra-molecular disulfide bond in PsbO-1 and PsbP-1 using a diagonal two-dimensional gel (Ströher and Dietz, 2008). PsbQ indeed has no Cys but could be associated and coeluted with other PS II OEC proteins. Biochemical studies showed that *A. thaliana* nucleoside diphosphate kinase 2 (NDPK2) is associated with MAPK-mediated H_2O_2 signaling in plants (Moon et al., 2003). Being one of the early oxidation targets, NDPK2 thiols might play a major role in the ROS-mediated signaling pathway. The *in vivo* experimental strategy should be improved in the future, e.g., by partial removal of RubisCO, followed by enrichment of less abundant proteins. The here applied method does not distinguish between direct oxidation by H_2O_2 and proximity-based oxidation mechanism where thiol proteins such as peroxiredoxins first react with H_2O_2 to form a SOH intermediate, which then oxidizes a thiol of another protein (König et al., 2012).

CONCLUSION

In conclusion, the here developed approach provides insight into early targets of oxidation by H_2O_2 within the stroma proteome under physiologically relevant conditions. This method is applicable to identify redox-dependent protein modifications *in vitro* and *ex vivo*, e.g., redox changes upon exogenous MV treatment but also in response to other stresses. The study identified expected early targets such as 2-Cys peroxiredoxin which has been classified as redox sensor and many known targets of redox regulation such as SBPase and FNR. Redox input elements, transmitters, targets, and sensors form the cellular redox regulatory

network (Dietz, 2008). This network needs to be expanded by the functional element redox buffer proteins. RubisCO is composed of eight LSU and eight SSU. LSU contains 9, SSU 4-Cys residues in *A. thaliana*. With a concentration of about 500 μM RubisCO, the RubisCO thiols represent the largest thiol pool of the stroma that exceeds any other chloroplast compound (Jacquot et al., 2013). A future challenge will be to monitor the diverse thiol modifications such as S-nitrosylation, glutathionylation, and intra- or interpeptide disulfide formation in parallel and to dissect the spatial and functional specificity of these modifications under conditions of environmental changes.

ACKNOWLEDGMENTS

This work was supported by the NRW International Graduate School for Bioinformatics and Genome Research (to Meenakumari Muthuramalingam), the DFG (Di346, FOR 804, FOR 944), and Bielefeld University. We are grateful to Annegret Wolf for technical assistance. Authors also thank Dr. Peter Klein for his scientific discussions.

REFERENCES

Allen, J. F. (2003). State transitions – a question of balance. *Science* 299, 1530–1532.

Asada, K. (1999). The water-water cycle in chloroplasts: scavenging of active oxygens and dissipation of excess photons. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 50, 601–639.

Balmer, Y., Koller, A., del Val, G., Manieri, W., Schürmann, P., and Buchanan, B. B. (2003). Proteomics gives insight into the regulatory function of chloroplast thioredoxins. *Proc. Natl. Acad. Sci. U.S.A.* 100, 370–375.

Barba-Espín, G., Diaz-Vivancos, P., Job, D., Belghazi, M., Job, C., and Hernández, J. A. (2011). Understanding the role of H2O2 during pea seed germination: a combined proteomic and hormone profiling approach. *Plant Cell Environ.* 34, 1907–1919.

Bartsch, S., Monnet, J., Selbach, K., Quigley, F., Gray, J., von Wettstein, D., et al. (2008). Three thioredoxin targets in the inner envelope membrane of chloroplasts function in protein import and chlorophyll metabolism. *Proc. Natl. Acad. Sci. U.S.A.* 105, 4933–4938.

Baty, J. W., Hampton, M. B., and Winterbourn, C. C. (2002). Detection of oxidant sensitive thiol proteins by fluorescence labeling and two-dimensional electrophoresis. *Proteomics* 2, 1261–1266.

Bhattachrjee, S. (2005). Reactive oxygen species and oxidative burst: roles in stress, senescence and signal transduction in plant. *Curr. Sci.* 89, 1113–1121.

Bienert, G. P., Moller, A. L., Kristiansen, K. A., Schulz, A., Møller, I. M., Schjørring, J. K., et al. (2007). Specific aquaporins facilitate the diffusion of hydrogen peroxide across membranes. *J. Biol. Chem.* 282, 1183–1192.

Blum, H., Beier, H., and Gross, H. J. (1987). Improved silver staining of plant proteins, RNA and DNA in polyacrylamide gels. *Electrophoresis* 8, 93–99.

Brandes, N., Schmitt, S., and Jakob, U. (2009). Thiol-based redox switches in eukaryotic proteins. *Antiox. Redox Signal.* 11, 997–1014.

Buchanan, B. B. (1980). Role of light in the regulation of chloroplast enzymes. *Annu. Rev. Plant Physiol.* 31, 341–374.

Buchanan, B. B., and Balmer, Y. (2005). Redox regulation: a broadening horizon. *Annu. Rev. Plant Biol.* 56, 187–220.

Dietz, K. J. (2008). Redox signal integration: from stimulus to networks and genes. *Physiol. Plant.* 133, 459–468.

Dietz, K. J. (2012). Peroxiredoxins in plants and cyanobacteria. *Antioxid. Redox Signal.* 15, 1129–1159.

Dixon, D. P., McEwen, A. G., Lapthorn, A. J., and Edwards, R. (2003). Forced evolution of a herbicide detoxifying glutathione transferase. *J. Biol. Chem.* 278, 23930–23935.

Dunford, R. P., Durrant, M. C., Catley, M. A., and Dyer, T. A. (1998). Location of the redox-active cysteines in chloroplast sedoheptulose-1,7-bisphosphatase indicates that its allosteric regulation is similar but not identical to that of fructose-1,6-bisphosphatase. *Photosyn. Res.* 58, 221–230.

Feng, L., Wang, K., Li, Y., Tan, Y. P., Kong, J., Li, H., et al. (2007). Overexpression of SBPase enhances photosynthesis against high temperature stress in transgenic rice plants. *Plant Cell Rep.* 26, 1635–1646.

Finkemeier, I., Goodman, M., Lamkemeyer, P., Kandlbinder, A., Sweetlove, L. J., and Dietz, K. J. (2005). The mitochondrial type II peroxiredoxin F is essential for redox homeostasis and root growth of *Arabidopsis thaliana* under stress. *J. Biol. Chem.* 280, 12168–12180.

Foyer, C. H., and Noctor, G. (2003). Redox sensing and signaling associated with reactive oxygen in chloroplasts, peroxisomes and mitochondria. *Physiol. Plant.* 119, 355–364.

Green, N. M. (1975). Avidin. *Adv. Protein Chem.* 29, 85–133.

Halliwell, B., and Gutteridge, J. M. C. (1989). *Free Radicals in Biology and Medicine*, 2nd Edn. Oxford: Clarendon Press.

Jacob, S., and Dietz, K. J. (2009). "Systematic analysis of superoxide-dependent signaling in plant cells: usefulness and specificity of methyl viologen application," in *Plant Stress Biology: From Genomics to Systems Biology*, ed. H. Hirt (Weinheim: Wiley-VCH Verlag), 179–196.

Jacquot, J. P., Dietz, K. J., Rouhier, N., Meux, E., Lallement, P. A., Selles, B., et al. (2013). "Redox regulation in plants: glutathione and "redoxin" related families," in *Oxidative Stress and Redox Regulation*, eds U. Jakob and D. Reichmann (Dordrecht: Springer Science and Business Media). doi: 10.1007/978-94-007-5787-5 (in press).

Jaffrey, S. R., and Snyder, S. H. (2001). The biotin switch method for the detection of *S*-nitrosylated proteins. *Sci. STKE* 86, pl1.

Kaiser, W. M. (1976). The effect of hydrogen peroxide on CO_2-fixation of isolated chloroplasts. *Biochim. Biophys. Acta* 440, 476–482.

Kim, J. R., Yoon, H. W., Kwon, K. S., Lee, S. R., and Rhee, S. G. (2000). Identification of proteins containing cysteine residues that are sensitive to oxidation by hydrogen peroxide at neutral pH. *Anal. Biochem.* 283, 214–221.

König, J., Lotte, K., Plessow, R., Brockhinke, A., Baier, M., and Dietz, K. J. (2003). Reaction mechanism of plant 2-Cys peroxiredoxin. Role of the C terminus and the quaternary structure. *J. Biol. Chem.* 278, 24409–24420.

König, J., Muthuramalingam, M., and Dietz, K. J. (2012). Mechanisms and dynamics in the thiol/disulfide redox regulatory network: transmitters, sensors and targets. *Curr. Opin. Plant Biol.* 15, 261–268.

Laxa, M., König, J., Dietz, K. J., and Kandlbinder, A. (2007). Role of the cysteine residues in *Arabidopsis thaliana* cyclophilin CYP20-3 in peptidyl-prolyl cis-trans isomerase and redox-related functions. *Biochem. J.* 401, 7–27.

Lefebvre, S., Lawson, T., Zakhleniuk, O. V., Lloyd, J. C., and Raines, C. A. (2005). Increased sedoheptulose-1,7-bisphosphatase activity in transgenic tobacco plants stimulates photosynthesis and growth from an early stage in development. *Plant Physiol.* 138, 451–460.

Leichert, L. I., and Jakob, U. (2004). Protein thiol modifications visualized in vivo. *PLoS Biol.* 2:e333. doi: 10.1371/journal.pbio.0020333

Levine, A., Tenhaken, R., Dixon, R., and Lamb, C. (1994). H_2O_2 from the oxidative burst orchestrates the plant hypersensitive disease resistance response. *Cell* 79, 583–593.

Lichter, A., and Häberlein, I. (1998). A light-dependent redox-signal participates in the regulation of ammonia fixation in chloroplast of higher plants – ferredoxin:glutamate synthase is a thioredoxin-dependent enzyme. *J. Plant Physiol.* 153, 83–90.

Lind, C., Gerdes, R., Hamnell, Y., Schuppe-Koistinen, I., von Lowenhielm, H. B., Holmgren, A., et al. (2002). Identification of S-glutathionylated cellular proteins during oxidative stress and constitutive metabolism by affinity purification and proteomic analysis. *Arch. Biochem. Biophys.* 406, 229–240.

Lindahl, M., and Kieselbach, T. (2009). Disulphide proteomes and interactions with thioredoxin on the track towards understanding redox regulation in chloroplasts and cyanobacteria. *J. Proteomics* 72, 416–438.

Liu, X. P., Liu, X. Y., Zhang, J., Xia, Z. L., Liu, X., Qin, H. J., et al. (2006). Molecular and functional characterization of sulfiredoxin homologs from higher plants. *Cell Res.* 16, 287–296.

Marchand, C., Le Marechal, P., Meyer, Y., Miginiac-Maslow, M., Issakidis-Bourguet, E., and Decottignies, P. (2004). New targets of *Arabidopsis* thioredoxins revealed by proteomic analysis. *Proteomics* 4, 2696–2706.

Marx, C., Wong, J. H., and Buchanan, B. B. (2003). Thioredoxin and germinating barley: targets and protein redox changes. *Planta* 216, 454–460.

Meyer, Y., Reichheld, J. P., and Vignols, F. (2005). Thioredoxins in *Arabidopsis* and other plants. *Photosynth. Res.* 86, 419–433.

Mieyal, J. J., and Chock, P. B. (2012). Posttranslational modification of cysteine in redox signaling and oxidative stress: focus on S-glutathionylation. *Antioxid. Redox Signal.* 16, 471–475.

Moon, H., Lee, B., Choi, G., Shin, D., Prasad, D. T., Lee, O., et al. (2003). NDP kinase 2 interacts with two oxidative stress-activated MAPKs to regulate cellular redox state and enhances multiple stress tolerance in transgenic plants. *Proc. Natl. Acad. Sci. U.S.A.* 100, 358–363.

Moreno, J., Garcia-Murria, M. J., and Marin-Navarro, J. (2008). Redox modulation of Rubisco conformation and activity through its cysteine residues. *J. Exp. Bot.* 59, 1605–1614.

Moreno, J., and Spreitzer, R. J. (1999). C172S substitution in the chloroplast-encoded large subunit affects stability and stress-induced turnover of ribulose-1,5-bisphosphate carboxylase/oxygenase. *J. Biol. Chem.* 274, 26789–26793.

Motohashi, K., Kondoh, A., Stumpp, M. T., and Hisabori, T. (2001). Comprehensive survey of proteins targeted by chloroplast thioredoxin. *Proc. Natl. Acad. Sci. U.S.A.* 98, 11224–11229.

Muthuramalingam, M., Dietz, K. J., and Ströher, E. (2010). Thiol–disulfide redox proteomics in plant research. *Methods Mol. Biol.* 639, 219–238.

Neill, S. J., Desikan, R., Clarke, A., Hurst, R. D., and Hancock, J. T. (2002). Hydrogen peroxide and nitric oxide as signalling molecules in plants. *J. Exp. Bot.* 53, 1237–1247.

Neuefeind, T., Huber, R., Reinemer, P., Knäblein, J., Prade, L., Mann, K., et al. (1997). Cloning, sequencing, crystallization and X ray structure of glutathione *S*-transferase-III from *Zea mays* var. mutin: a leading enzyme in detoxification of maize herbicides. *J. Mol. Biol.* 274, 577–587.

Peltier, J. B., Cai, Y., Sun, Q., Zabrouskov, V., Giacomelli, L., Rudella, A., et al. (2006). The oligomeric stromal proteome of *Arabidopsis thaliana* chloroplasts. *Mol. Cell. Proteomics* 5, 114–133.

Pfannschmidt, T. (2003). Chloroplast redox signals: how photosynthesis controls its own genes. *Trends Plant Sci.* 8, 33–41.

Pogson, B. J., Woo, N. S., Forster, B., and Small, I. D. (2008). Plastid signalling to the nucleus and beyond. *Trends Plant Sci.* 13, 602–609.

Polle, A. (2001). Dissecting the superoxide dismutase-ascorbate-glutathione-pathway in chloroplasts by metabolic modeling. Computer simulations as a step towards flux analysis. *Plant Physiol.* 126, 445–462.

Rouhier, N., Villarejo, A., Srivastava, M., Gelhaye, E., Keech, O., Droux, M., et al. (2005). Identification of plant glutaredoxin targets. *Antioxid. Redox Signal.* 7, 919–929.

Scheibe, R., and Dietz, K. J. (2012). Reduction-oxidation network for flexible adjustment of cellular metabolism in photoautotrophic cells. *Plant Cell Environ.* 35, 202–216.

Scheibe, R., Wedel, N., Vetter, S., Emmerlich, V., and Sauermann, S. M. (2002). Co-existence of two regulatory NADP-glyceraldehyde 3-P dehydrogenase complexes in higher plant chloroplasts. *Eur. J. Biochem.* 269, 5617–5624.

Sethuraman, M., McComb, M. E., Huang, H., Huang, S., Heibeck, T., Costello, C. E., et al. (2004). Isotope-coded affinity tag (ICAT) approach to redox proteomics: identification and quantitation of oxidant-sensitive cysteine thiols in complex protein mixtures. *J. Proteome Res.* 3, 1228–1233.

Ströher, E., and Dietz, K. J. (2008). The dynamic thiol-disulphide redox proteome of the *Arabidopsis thaliana* chloroplast as revealed by differential electrophoretic mobility. *Physiol. Plant.* 133, 566–583.

Tanou, G., Job, C., Belghazi, M., Molassiotis, A., Diamantidis, G., and Job, D. (2010). Proteomic signatures uncover hydrogen peroxide and nitric oxide cross-talk signaling network in citrus plants. *J. Proteome Res.* 9, 5994–6006.

Taylor, T. C., and Andersson, I. (1997). Structure of a product complex of activated spinach ribulose-bisphosphate carboxylase/oxygenase. *Biochemistry* 36, 4041–4046.

Tietze, F. (1969). Enzymic method for quantitative determination of nanogram amounts of total and oxidized glutathione: applications to mammalian blood and other tissues. *Anal. Biochem.* 27, 502–522.

Winger, A. M., Taylor, N. L., Heazlewood, J. L., Day, D. A., and Millar, A. H. (2007). Identification of intra- and intermolecular disulphide bonding in the plant mitochondrial proteome by diagonal gel electrophoresis. *Proteomics* 7, 4158–4170.

Zaffagnini, M., Bedhomme, M., Groni, H., Marchand, C. H., Puppo, C., Gontero, B., et al. (2012). Glutathionylation in the photosynthetic model organism *Chlamydomonas reinhardtii*: a proteomic survey. *Mol. Cell. Proteomics* 11, M111.01414.

Zhou, L., Bokhari, S. A., Dong, C., and Liu, J. (2011). Comparative proteomics analysis of the root apoplast of rice seedlings in response to hydrogen peroxide. *PLoS ONE* 6:e16723. doi: 10.1371/journal.pone.0016723

Zrenner, R., Willmitzer, L., and Sonnewald, U. (1993). Analysis of the expression of potato uridine diphosphate-glucose pyrophosphorylase and its inhibition by antisense RNA. *Planta* 190, 247–252.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 21 December 2012; accepted: 28 February 2013; published online: 19 March 2013.

Citation: Muthuramalingam M, Matros A, Scheibe R, Mock H-P and Dietz K-J (2013) The hydrogen peroxide-sensitive proteome of the chloroplast in vitro and in vivo. Front. Plant Sci. ***4****:54. doi: 10.3389/fpls.2013.00054*

This article was submitted to Frontiers in Plant Proteomics, a specialty of Frontiers in Plant Science.

Copyright © 2013 Muthuramalingam, Matros, Scheibe, Mock and Dietz. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.

APPENDIX

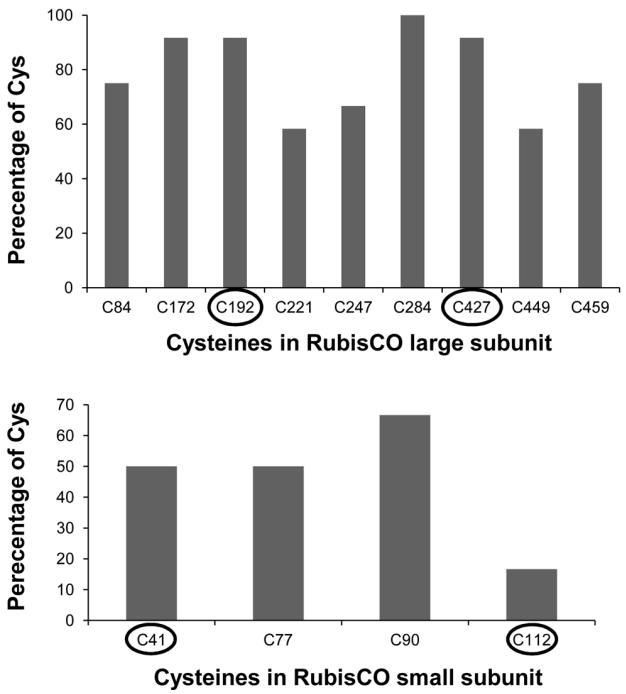


FIGURE A1 | Percentage of cysteine conservation in *A. thaliana* RubisCO large and small subunit protein in comparison to the following organisms: Rice, maize, Physcomitrella, Chlamydomonas, Chlorella, Euglena, Synechocystis, Nostoc, Hydrogenovibrio, *R. capsulatus* and *R. sphaeroides*. Encircled are the identified redox sensitive cysteines.

frontiers in
PLANT SCIENCE

REVIEW ARTICLE
published: 18 March 2013
doi: 10.3389/fpls.2013.00052

Functional proteomics of barley and barley chloroplasts – strategies, methods and perspectives

Jørgen Petersen, Adelina Rogowska-Wrzesinska and Ole N. Jensen*

Department of Biochemistry and Molecular Biology, University of Southern Denmark, Odense, Denmark

Edited by:
Nicolas L. Taylor, The University of Western Australia, Australia

Reviewed by:
Wolfgang P. Schröder, Umeå University, Sweden
Clark Nelson, University of Western Australia, Australia

****Correspondence:***
Ole N. Jensen, Department of Biochemistry and Molecular Biology, University of Southern Denmark, Campusvej 55, DK-5230 Odense, Denmark.
e-mail: jenseno@bmb.sdu.dk

Barley (*Hordeum vulgare*) is an important cereal grain that is used in a range of products for animal and human consumption. Crop yield and seed quality has been optimized during decades by plant breeding programs supported by biotechnology and molecular biology techniques. The recently completed whole-genome sequencing of barley revealed approximately 26,100 open reading frames, which provides a foundation for detailed molecular studies of barley by functional genomics and proteomics approaches. Such studies will provide further insights into the mechanisms of, for example, drought and stress tolerance, micronutrient utilization, and photosynthesis in barley. In the present review we present the current state of proteomics research for investigations of barley chloroplasts, i.e., the organelle that contain the photosynthetic apparatus in the plant. We describe several different proteomics strategies and discuss their applications in characterization of the barley chloroplast as well as future perspectives for functional proteomics in barley research.

Keywords: barley, *Hordeum vulgare*, proteomics, chloroplast, mass spectrometry, 2D gel electrophoresis

INTRODUCTION

Barley (*Hordeum vulgare*) is one of the earliest domesticated cereals and it is the fourth most important crop world-wide in terms of total dry production, only exceeded by maize, rice, and wheat. Barley is mainly used in the brewing industry and as animal feed, but in certain areas of the world it is an important food source for humans (Schulte et al., 2009). The increasing demand for food due to the growing world population has propelled the implementation of plant breeding programs and biomolecular plant research to improve sustainable crop production. Prioritized areas include research in plant resistance to abiotic stress such as soil salinity, temperature, drought, nutrient uptake (Saeed et al., 2012), and biotic stress caused by other living organisms and pathogens (Dreher and Callis, 2007). Barley is by nature diploid, has a low chromosome number ($2n = 14$) and a large genome size (5.1 Gb), is easy to cross-breed and is able to grow under various climatic conditions. These abilities and the fact that barley is an extremely important crop makes it desirable to identify genes responsible for specific beneficial traits in order to improve crop production and sustainability (Saisho and Takeda, 2011). The recently completed whole-genome sequencing of the barley genome (Mayer et al., 2012), gave rise to several interesting observations. A total of 26,159 high confidence genes with gene-family similarity to other plant genomes, and 53,220 genes with lack of homology denoted low confidence genes were identified. By comparison to *Arabidopsis thaliana*, the barley genome was estimated to encompass 30,400 genes. RNA sequencing data indicated extensive alternative splicing of the coding regions of the high confidence genes (Mayer et al., 2012), this adds to protein diversity and may play a role in protein regulation and gene expression (Syed et al., 2012). These data opens new opportunities for pursuing in-depth studies of barley biology by using genomics, transcriptomics, metabolomics, and proteomics approaches.

The chloroplast is one of the specialized plastids in the plant cell and it conducts important processes such as photosynthesis and biosynthesis of amino acids, starch, and vitamins. The chloroplast contains its own genome, but most of the estimated 2000–3000 chloroplast proteins are encoded by the nuclear genome. Targeting of proteins to the chloroplast often requires N-terminal pre-sequences called chloroplast transit peptides (cTPs), which to some extend can be predicted from the genome by using computational methods such as chloroP, targeP, WoLF PSORT, iPSORT, predotar, or Protein Prowler (Emanuelsson et al., 1999; Bannai et al., 2002; Small et al., 2004; Boden and Hawkins, 2005; Horton et al., 2007).

Functional proteomics is a rapidly evolving scientific discipline that is driven by advancements in a series of bioanalytical and computational technologies to enable increasingly detailed studies of complex protein mixtures derived from cells, tissues, and organisms (Aebersold and Mann, 2003; Cravatt et al., 2007; Bensimon et al., 2012). The main methods used in proteomics are: (1) protein and peptide separation techniques; (2) mass spectrometry; (3) biological sequence databases and computational query tools (summarized in **Boxes 1** and **2**).

Proteomics technologies are now extensively used in plant biology, particularly in studies of the model plants and the most important food crops (Jorrin et al., 2007). Proteomics, i.e., the systematic study and characterization of proteins in a cell type, tissue, or a whole organism, encompasses the mapping of protein composition and abundance, protein interactions and protein localization, as well as dynamic events in protein regulatory networks, including signaling mechanisms, metabolism, and transcription (de Hoog and Mann, 2004). A majority of such studies in plants were carried out in *A. thaliana* and rice where completely sequenced genomes are available (Kaul et al., 2000; Goff et al., 2002). Proteome analysis of plant organelles, including

BOX 1 | Mass spectrometry.

Mass spectrometry enables unambiguous identification of proteins by accurate mass measurements of gas-phase protein and peptide ions and peptide fragment ions. Mass spectrometers using matrix-assisted laser desorption ionization (MALDI) are preferred for simple peptide mixtures derived by in-gel digestion of proteins obtained from 2D gel spots (Gevaert and Vandekerckhove, 2000). Electrospray ionization (ESI) mass spectrometers are frequently interfaced directly to nanoliter-flow HPLC systems, thereby providing separation, mass determination, and amino acid sequencing in one analytical setup (LC-MS/MS) (Aebersold and Mann, 2003). Besides being able to identify thousands of proteins in one single LC-MS/MS analysis, modern proteomics workflows also provides rather accurate protein quantification and capability to identify PTMs (Larsen et al., 2006; van Bentem et al., 2006; Ytterberg and Jensen, 2010; Mithoe and Menke, 2011). These features make MALDI and ESI mass spectrometry indispensable in proteomics research for the characterization and quantification of complex protein mixtures.

BOX 2 | Quantitative proteomics.

2D gel electrophoresis is the preferred method for comparative quantitative proteomics in studies of organisms for which only incomplete gene annotation is available, e.g., for carrots and cabbage (Nawrocki et al., 2011). The advantage of using 2D gel electrophoresis is the one spot – one protein premise that makes it relative easy to make sequence homology searches, de-novo sequencing of fragmented peptides or protein isoform characterization (Jacob and Turck, 2008; Moller et al., 2011b). Mass spectrometry driven quantitative proteomics methods can be categorized into "label-free" approaches based on peptide intensity or peptide counting and "stable isotope labeling" methods where proteins and/or peptides are metabolically or chemically encoded by heavy stable isotopes of, e.g., carbon, nitrogen, and oxygen (13-C, 15-N, 18-O; Ong and Mann, 2005; Thelen and Peck, 2007; Bantscheff et al., 2012). Commonly used metabolic labeling methods in plant proteomics include stable isotope labeling by 15-N (Nelson et al., 2007; Bindschedler et al., 2008; Gouw et al., 2008) and by amino acids in cell culture [stable isotope labeling by amino acids in cell culture (SILAC); Ong et al., 2002], although the latter is not easily implemented in plants due to their amino acid metabolism (Gruhler et al., 2005). Chemical methods for stable isotope labeling are generically applicable in plant proteomics and include iTRAQ (Ross et al., 2004; Wiese et al., 2007) and isotope-coded protein labeling (ICPL; Schmidt et al., 2005). Examples include phosphoproteomics (Jones et al., 2006; Melo-Braga et al., 2012), global protein regulation in response to stress (Neilson et al., 2011; Abdalla and Rafudeen, 2012) or as a consequence of genotypic differences (Chen et al., 2009; Ng et al., 2012).

The advantage of label-free approaches is that they are rather straightforward to implement, however, their robustness and accuracy relies on multiple replicate runs and comparative data analysis is often rather complex. Nevertheless, recent improvements in software and statistics for label-free proteomics make this a very attractive approach. The main advantages of stable isotope labeling techniques are their accuracy of quantification and the ability to perform multiplex experiments. iTRAQ allows up to eight-plex analysis in one LC-MS/MS experiment (Bantscheff et al., 2007; Pottiez et al., 2012).

chloroplasts, have been reported (Kleffmann et al., 2004). For example, proteomics strategies were used to elucidate the influence of various biotic and abiotic stresses on chloroplasts proteins.

The recently completed sequencing of the barley genome now provides a foundation for more detailed functional proteomics studies of barley biology. We therefore foresee an increased effort in barley proteomics using state-of-the-art mass spectrometry based strategies for qualitative and quantitative characterization of barley proteins, organelles and regulatory networks. Proteomics will likely play a major role in further improvements of barley cultivars, e.g., by identifying the underlying mechanisms of biotic and abiotic stress. In the following sections we provide an overview of proteomics strategies and techniques and the current state of barley chloroplast proteomics.

GENERAL CONSIDERATIONS AND PROTEOMICS STRATEGIES

Several factors affect the outcome of a proteomics experiment, and need to be included in the experimental planning phase, like for example proteome complexity and protein concentration (summarized in **Box 3**). This section covers two classical proteomics strategies and highlights things to consider before starting a chloroplast-targeted proteomics experiment.

Purification: The first step toward success in organelle or sub proteomic experiment is the quality and purity of the sample. Contaminating proteins or unwanted cellular debris can obscure the results with respect to assignment of organelle specific proteins and their quantification (Agrawal et al., 2011). Highly purified chloroplasts or mitochondria can be obtained using a Percoll gradient centrifugation step (Neuburger et al., 1982; Aronsson and Jarvis, 2002; van Wijk, 2004; Millar et al., 2005). Endomembrane organelles such as Golgi apparatus, endoplasmic reticulum, vacuoles, and vesicles are more difficult to purify without cross-contamination from other organelles. Gentle rupture of the intact chloroplasts enables further purification of four subcompartments (1) the inner and outer envelope membranes, (2) the stroma, (3) the thylakoid membrane, (4) the thylakoid lumen (Kieselbach et al., 1998, 2000; Peltier et al., 2000, 2006; Schubert et al., 2002; Ferro et al., 2003). The above mentioned extractions method were used for diverse plant species, and it is important to have in mind that protocols developed for a specific plant species, not necessarily works for other species. Typically intact chloroplasts are obtained using Percoll gradient centrifugation. This is by far the best way to obtain pure chloroplasts, but the yield is rather low. Less pure chloroplast can be obtained in high yields using low speed centrifugation. It is possible to obtain thylakoid, stroma, and envelope fractions using a sucrose gradient of osmotic shocked intact chloroplasts. Soluble luminal thylakoid proteins can be isolated from the thylakoid preparation using yeda press rupture of the membranes (Hall et al., 2011).

How much material is needed? It is possible to make quantitative proteomics experiments with less than 20 μg of extracted protein. The number of identified proteins from such an experiment depends not only on the complexity and dynamics of the proteome but also the in-house instrumentation (Eriksson and Fenyo, 2010). In sub proteomic work the amount of starting material might exceed several grams to extract a few micro grams of

BOX 3 | Proteome complexity and protein concentrations.

Due to the high complexity and wide concentration range of proteins within proteomes, large scale proteome analysis is often executed at the sub-proteome level (James, 1997; Kuntz and Rolland, 2012) where specific cellular or tissue fractions are isolated and analyzed. For example, enrichment strategies can be used to isolate sub-proteome consisting of, e.g., kinases, or proteins containing specific modifications (e.g., phosphorylation or glycosylation), body or tissue fluids (e.g., sap) or organelles such as cell nuclei, mitochondria, Golgi apparatus, or chloroplasts. The need for fractionation into sub-proteomes becomes obvious when considering that the potential number of different proteins from a single genome coding for 20,000–30,000 genes, might be as high as 200,000–2 million when considering genomic recombination, splice variants, differential initiation/termination of transcripts and protein processing and covalent modifications (Ayoubi and Van De Ven, 1996; Lander et al., 2001). In addition, the concentration ranges of proteins in eukaryotic cells typically span five–six orders of magnitude and in some sub-proteomes as high as 10 orders of magnitude. In some plants it has been estimated that RuBisCO makes up 40% of the total protein content, making the stroma in the chloroplast a very challenging protein matrix to analyze (Patterson and Aebersold, 2003; Bindschedler and Cramer, 2011). By reducing protein complexity by sub-proteome fractionation it is possible to identify low abundant proteins in the proteome of an organism.

a desired proteome. As an example, from 100 g of soil grown *A. thaliana* plants it is possible to extract approximately 1000 mg leaf protein, 100 mg thylakoid proteins, and only 0.4 mg envelope membrane protein (Froehlich et al., 2003).

What buffers should I use? There is no universal buffer composition to be used in proteomics experiments. Depending on the targeted tissue or sub-cellular compartment different protein extraction and sample preparation buffers are used (Fido et al., 2004; Mano et al., 2008). However, there are few universal rules that should be taken into consideration. Always add protease inhibitors, but be aware of the lifetime of the inhibitors, it might be short under certain conditions, or use strong denaturing buffers [e.g., 8 M urea or sodium dodecyl sulfate (SDS)] to inactivate potential proteolytic activity of enzymes present in the sample. Use metal chelating agents, e.g., ethylenediaminetetraacetic acid (EDTA) to trap free metal ions from the sample to prevent unwanted spontaneous protein oxidation – this is particularly important when working with organelles such as chloroplasts and mitochondria. Most buffers used in biological experiments contains components which are not compatible with liquid chromatography-mass spectrometry (LC-MS) but if the proteins are separated by polyacrylamide gel electrophoresis (PAGE) all buffers are allowed, due to the excellent washing ability of gel plugs. For non-gel based strategies some compounds such as detergents (SDS, Triton X-100, etc.) or ampholytes compromise nanoliter-flow LC or mass spectrometry and they need to be avoided or removed prior to analysis (Xiao et al., 2004; Yeung and Stanley, 2010).

Can high abundant proteins be removed? In photosynthetic tissue the predominant protein is the carbon fixation protein ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco). In some cases more than 50% of the total leaf protein content consist of Rubisco (Metodiev and Demirevskakepova, 1992). Such highly abundant protein will hamper both gel and non-gel based proteome analysis because this highly abundant protein will obscure other proteins and suppress their detection. In gel based studies it will dominate the gel pattern eclipsing low abundant proteins with similar physico-chemical properties. In non-gel based peptides generated from this abundant protein will saturate high performance liquid chromatography (HPLC) columns and suppress the signal from lower abundant proteins. This problem can be partially solved by removing the highly abundant protein by fractionation, antibody based spin columns, or using the relative newly developed ProteoMiner beads (Boschetti and Righetti, 2008; Frohlich et al., 2012). Removal of highly abundant proteins can also result in removal of the associated low abundant proteins (Cellar et al., 2008; Krishnan and Natarajan, 2009). Another way to reduce the complexity and the dynamics of protein sample is to perform organelle or sub-organelle fractionation. Isolation of mitochondria or a thylakoid preparation from chloroplast will exclude the majority of Rubisco protein from the analysis.

How to proceed after proteome extraction? Proteins, both for gel and non-gel based strategies (see below) need to be digested into peptides prior to mass spectrometry analysis. The aim is to generate ionizable peptides in the mass range 700–2500 Da, which is the optimal range for most biological mass spectrometers. Disulfide bridges (Cys-Cys) in proteins are typical reduced and alkylated using dithiothreitol (DTT) and iodacetamide (IAA) prior to digestion. Denaturation of the proteins improves digestion efficiency, thus contributing to the overall protein identification rate. Proteins separated by SDS-PAGE are inherently denatured and are typically cut out of the gel, reduced, *S*-alkylated and digested by trypsin. This is a well-established "in-gel digestion" technique routinely used by most proteomics laboratories (Shevchenko et al., 1996, 2006). In solution based digestion is a more delicate procedure. Keeping the proteins in solution, denatured and available for trypsin digestion can be facilitated by buffers containing the commercially available surfactant RapiGest, urea buffers or detergents such as sodium deoxycholate (SDC) that, in contrast to SDS can relatively easy be removed from the sample prior to the mass spectrometry analysis (Speers and Wu, 2007; Norrgran et al., 2009; Lin et al., 2012). In-solution digestion protocols where the digestion is performed within a spin filter device has become popular and is highly recommended for the digestion of protein amounts exceeding 100 μg. The filter enables washing of the sample and retention of large unwanted structures on the filter (Manza et al., 2005; Wisniewski et al., 2009).

How do I evaluate the quality of the experiment? Proteomics experiments often aim to detect differential regulated proteins between groups. This can be accomplished using a statistical test based on hypotheses about characteristics of both the biological samples that represent the population, and the variability of the technical measurements (Podwojski et al., 2012).

If possible, evaluate the protein extract by electrophoresis; this gives an overall picture of the extract. Non-gel based approaches can benefit using an internal spike-in protein standard. The protein standard is digested together with the extract, and by comparing sequence coverage and peptide intensities of the spiked-in

standard among samples, the digest efficiency can be evaluated. This can be archived using selected reaction monitoring (SRM) or other label-free quantification methods. Absolute quantification can be archived using spiked-in peptides that act as internal standards (Gerber et al., 2003; Silva et al., 2006).

PROTEOMICS STRATEGIES

The choice of proteomic strategy depends on several factors such as the overall aim of the proteomics experiment, protein sample complexity and protein amount, number of samples to analyze, mass spectrometry instrument considerations, sequence database availability and whether protein quantification is necessary (**Figure 1**).

2D gel electrophoresis is a separation technique that is based on isoelectric focusing of the proteins followed by separation of the proteins according to their molecular mass. It has been used in proteomics for more than 30 years. Although a number of its limitations have been recognized (reviewed in Issaq and Veenstra, 2008; Chevalier, 2010) it is an effective strategy for the separation and quantitation of intact protein mixtures, including protein isoforms and modified proteins. A variation of the classical denaturing 2D PAGE is blue native (BN) 2D PAGE (Reisinger and Eichacker, 2007). This technique has been used in several membrane proteins studies (Krause, 2006), and is also one of the preferred ways for characterization of protein complexes. Protein separated by electrophoresis are visualized by staining, isotope or fluorescent labeling (Patton, 2002). Often only the differential regulated proteins are selected for spot picking, protein digestion, and protein identification (Berth et al., 2007). The advantage using the 2D gel strategy is the one spot – one protein premise, which allows for relatively easy *de novo* annotation of peptide fragment spectra and homolog search.

The combination of SDS-PAGE and LC-MS is very efficient for proteome profiling. The combination is often called GeLC-MS/MS, and is excellent for proteome profiling due to the unbiased solubilization of all protein groups including membrane proteins. For quantitative measurements it can be used with metabolically incorporated stable isotopes, isobaric tags for relative and absolute quantitation (iTRAQ) and semi quantitative approaches such as spectral counting (Sachon et al., 2006; Wienkoop et al., 2006).

Recently, 2D LC-MS/MS strategies have become more widespread and robust. The orthogonality between the two LC separation dimensions is often obtained by using strong cation exchange chromatography (SCX) in the first dimension and reverse phase (RP) chromatography in the second dimension, separating the peptides according to charge and then according to hydrophobicity (Washburn et al., 2001). Other types of resin, e.g., hydrophilic interaction liquid chromatography (HILIC) and size-exclusion chromatography (SEC) have also been used in proteomic studies (Gilar et al., 2005a). More recently, RP–RP HPLC systems using high pH and low pH mobile phases in the first and second separation dimensions, respectively, have proved to be excellent and robust for proteomics work (Gilar et al., 2005b). This set up can be fully automated and is suitable for proteomics work where several biological replicates are needed. It can be combined with both label based and label-free quantification methods. It is also possible to achieve absolute quantification of the identified

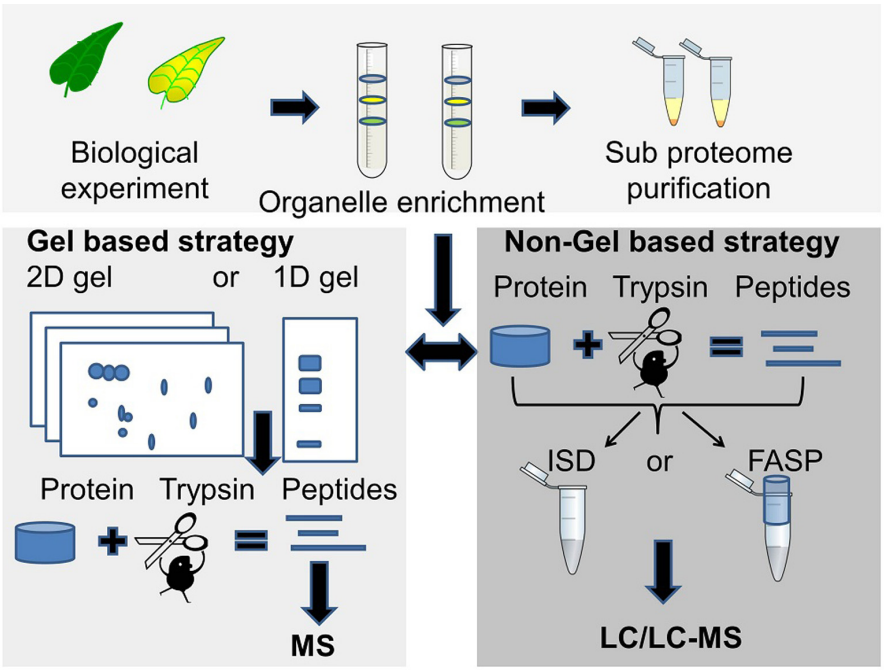


FIGURE 1 | Two commonly used proteomics strategies. The gel based strategy (lower left panel), where protein bands or spots are cut out of the gel followed by trypsin digestion and MS analysis of peptides, and the MS based strategy (lower right panel) where the unseparated or partially fractionated protein sample is "in-solution digested" (ISD) with trypsin or digested in a spin filter [filter-aided sample preparation (FASP)] followed by LC/LC-MS/MS for peptide separation and sequencing.

proteins by spiking in known amounts of digested protein standards (Silva et al., 2006). Separation using only one dimension is also possible, but for complex samples or samples with high dynamic range, the number of protein identifications will be limited due to lower peak capacity compared to 2D LC strategies where two orthogonally retention mechanisms are used.

Mass spectrometry data contains peptide information at the MS and at the MS/MS level. For protein identification the MS and MS/MS data can be searched using commercial or publicly available search engines such as Sequest, Mascot, OMSSA, or X!tandem (Cottrell, 2011). Software designed for handling large proteomics datasets integrates multiple features such as identification, quantification, visualization, statistics, and reporting. These include packages such as Phenyx, Trans-Proteomic Pipeline (TPP) MaxQuant, and Peaks (Lemeer et al., 2012).

CURRENT STATUS OF BARLEY PROTEOMICS

The areas where barley proteomics has been used can be divided into (a) industry driven biotechnology, including seed germination and maturation, beer proteomes, and malting proteomes and (b) biology driven proteomics covering plant adaptation to abiotic stress and organelle function including the chloroplast that is the focus of this review.

Biotechnology driven proteomics: Understanding the mechanisms involved in seed germination and maturation processes are important aspects in the malting industry where, e.g., enzyme amount such as amylase in different cultivars influences the conversion of starch into fermentable sugars. The work with proteome analysis of different barley seed cultivars and proteomes from different developmental stages of germinating barley started in year 2002 (Finnie et al., 2002; Ostergaard et al., 2002). 2D gels were used as a protein profiling tool. The proteins were extracted using a low salt buffer, favoring the extraction of water soluble seed proteins such as amylases and chitinases, and minimized extraction of high abundant storage proteins such as hordeins that otherwise would dominate the protein profile in the 2D gel. The TrEMBL database at that time only contained 546 barley protein sequences, so therefore most of the protein identifications were based on cross-species protein annotation using other cereals, such as rice, maize, and wheat. The strength of 2D gel electrophoresis was also pointed out in these studies, since the same protein was identified in multiple protein spots, maybe as a consequence of post-translational modifications (PTM) or multiple alleles with almost identical protein sequences.

Hynek et al. (2009) reported the enrichment of hydrophobic membrane proteins from the barley plasma membrane fraction, which may play a key role in the germination process, by using two-phase partitioning and RP chromatography. The enrichment of the membrane fraction was validated using western blotting against H^+-ATPase, a protein located in the membrane. Sixty-one barley proteins were identified after SDS-PAGE by using electrospray tandem mass spectrometry (ESI-MS/MS).

Protein profiles of different beers are diverse due to differences in the barley cultivar, the malting process and the brewing yeast. 2D gel maps of different beer proteomes representing different cultivars and malting types have been created. The maps can be used as quality control step in the brewing industry and as a tool to detect and identify beer type specific proteins or protein isoforms that might represent taste, flavor, or texture. In the long term this will potentially enable manipulation of, e.g., flavor proteins (Fasoli et al., 2010; Iimure et al., 2010). The industrial induced protein modification called Maillard reactions has also been monitored and characterized and is important for color, taste, and flavor and include thermal stability of proteins and the non-enzymatic glycation of proteins (Perrocheau et al., 2005; Okada et al., 2008; Petry-Podgorska et al., 2010)

Biology driven proteomics: 2D gel electrophoresis was the preferred method to study the proteome of barley plants exposed to salinity stress and adaptation (Fatehi et al., 2012). Barley plants, a tolerant and a salt-sensitive genotype, were exposed to 0 (control) or 300 mM NaCl. More than 500 reproducible protein spots were detected of which 44 appeared to be regulated. The regulated proteins were involved in several biological processes such as reactive oxygen species scavenging, signal transduction, and protein processing. The advantage of this 2D gel strategy for studying a non-sequenced organism was pointed out – only the regulated proteins needed to be analyzed and identified by mass spectrometry. A similar procedure was used in a nitrogen use efficiency study of barley, where proteomes from barley shoots and roots were analyzed using 2D gels. Comparative proteome analysis of plants grown with a nitrogen source and plants grown under nitrogen deficiency revealed 67 and 49 differentially regulated protein spots in roots and shoots, respectively (Moller et al., 2011a). Proteins associated with drought have also been analyzed using 2D gel proteomics (Wendelboe-Nelson and Morris, 2012). In a comparative study of barley, extracted leave and root proteomes from boron tolerant and boron intolerant barley plants were studied using an iTRAQ based method and peptide fractionation by 2D LC prior to mass spectrometry analysis. A total of 138 proteins were identified from leaf tissue and 341 were identified from root tissues. Only 11 out of 1038 peptides from the root tissue were regulated in the boron tolerant barley plant. Interestingly seven of these peptides identified three proteins involved in iron deficiency response (Patterson et al., 2007).

Protein modifications such as acetylation, glycosylation, and phosphorylation are important regulators of a wide range of biological processes in plants (Ytterberg and Jensen, 2010). In barley only a handful of proteomics studies deal with protein modifications. These include protein characterization in seeds during maturation using 2D gels (Finnie et al., 2006; Laugesen et al., 2007), where spot "trains" of the same proteins appeared during maturation as a consequence of small amino acids sequence differences, processing and differences in the degree of protein glycosylation. Phosphoprotein studies in tonoplasts revealed a total of 65 phosphopeptides, and provide a first view into the regulation of several metabolic pathways in tonoplast (Endler et al., 2009). Phosphoproteomics in plants were recently reviewed (Kline-Jonakin et al., 2011).

THE BARLEY CHLOROPLAST PROTEOME

Only a few studies concerning the barley chloroplast proteome have been published, and a comprehensive list of barley chloroplast proteins is yet to be reported. In contrast, global proteomics in *Arabidopsis* has been a reality for more than 20 years due to

the complete sequencing of the *A. thaliana* genome at the beginning of this millennium. (Kaul et al., 2000; Wortman et al., 2003). Chloroplast proteome work in barley, wheat, and *A. thaliana* will be discussed below. **Figure 2** compares the number of proteins identified in the chloroplast sub-compartments from these three species.

The envelope membrane: The envelope membrane of the chloroplast is the site of several important functions such as biosynthesis of glycerolipids, fatty acid export, metabolite transport, and protein import. In *A. thaliana*, Ferro et al. (2010) reported 644 proteins to be associated with the membrane envelope using both in-gel and in-solution digestion of proteins. Earlier studies of the envelope membrane using both geLC-MS/MS and 2D LC-MS/MS produced fewer identifications (Ferro et al., 2003; Froehlich et al., 2003).

The thylakoid membrane: The thylakoid membrane contains the photosynthetic machinery, but also proteins involved in regulation and maintenance of this machinery. In thylakoid preparations from *A. thaliana* the number of identified proteins sums up to 242 using geLC-MS/MS and LC-MS/MS and 154 proteins using 2D gels (Friso et al., 2004; Peltier et al., 2004). A total of 198 thylakoid luminal proteins have been identified combing data from several studies (Peltier et al., 2002; Giacomelli et al., 2006). Some of these are believed to be up to 10,000-fold less abundant than photosynthetic proteins, and can only be identified by sub-proteome isolation. Other studies on luminal proteins report less proteins (Schubert et al., 2002), which might reflect differences in purification and proteomics strategies.

The thylakoid membrane of barley was investigated by the use of BN 2D PAGE, with the aim to compare the photosynthetic machinery of barley with that of other higher plants (Ciambella et al., 2005). The number of barley thylakoid proteins identified was 45, of these 17 proteins from photosystem II (PSII), 16 from PSI, 7 proteins from cytochrome B6, and 5 from the ATP synthase. The same number of barley thylakoid proteins was reached in another study (Granvogl et al., 2006). One recent study from 2011 (Ploscher et al., 2011) compares protein complexes from etioplast and chloroplast. This is at the moment the most comprehensive chloroplast proteome study in barley. Etioplasts develop in the absence of light but can mature into chloroplasts by illumination. By using 2D BN/SDS-PAGE to separate the protein membrane complexes from etioplast and chloroplast, they found eight etioplast/chloroplast shared protein complexes, among those with high number of subunit representation were the ATPase, cytochrome b6, and the NAD(P)H dehydrogenase complex, whereas the PSI and PSII complexes were only present in the chloroplast. The use of BN gels made it possible to quantify and distinguish between monomeric, dimeric, and multimeric forms of the photosynthetic protein machinery, and to distinguish between the different subunits present in the protein complexes, making assumptions of assembly and maturation of protein complexes possible. Both automated and manually inspected fragment spectra were generated from the mass spectrometry based analysis where both online protein identification of tryptic digested proteins and off-line identification of intact small proteins extraction from gel were identified. In an earlier study by the same group (Ploscher et al., 2009), intact low molecular weight proteins from PSII were identified using off-line ESI MS.

The stroma: The stroma contains the genetic material and important metabolic enzymes including those involved in the Calvin cycle. Using the geLC-MS/MS approach a total of 590 *A. thaliana* proteins were identified (Zybailov et al., 2008). Less protein identifications were obtained in an attempt to identify paralogs using 2D native gels (Peltier et al., 2006). For barley no stromal proteome studies have appeared to date, but four proteins from the above mentioned preparations (Ciambella et al., 2005; Ploscher et al., 2011) are supposedly targeted to the stroma.

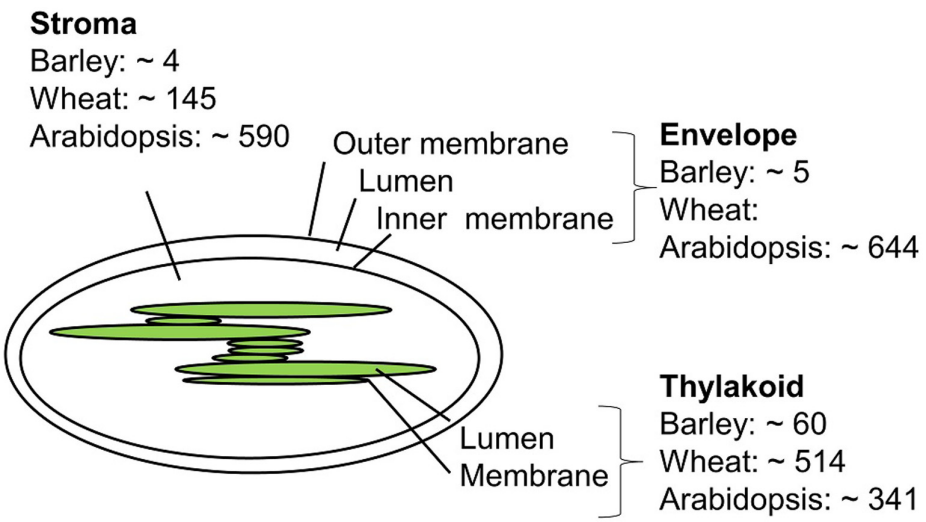


FIGURE 2 | Approximate number of chloroplast proteins identified in barley (Ciambella et al., 2005; Bartsch et al., 2008; Ploscher et al., 2011), wheat (Kamal et al., 2012), and *Arabidopsis* (Peltier et al., 2004; Giacomelli et al., 2006; Zybailov et al., 2008; Ferro et al., 2010).

In a recent chloroplast proteomic study in wheat, which shares sequence similarity with barley, the geLC-MS/MS strategy was used, and a total of 607 chloroplast proteins were identified. Of these, 145 were from stroma, 342 were from the thylakoid membrane, 163 from the lumen, and 166 proteins were integral membrane proteins (Kamal et al., 2012).

Armbruster et al. (2011) summarizes all proteomics work on chloroplast and comes up with a total number of nucleus encoded proteins to be 1741, 63% with predicted cTP.

THE FUTURE FOR PROTEOMICS OF BARLEY AND BARLEY CHLOROPLASTS

Proteomics work in barley has to date been hampered by the lack of complete genomic sequence. But by the complete sequencing of the barley genome the goal to identify all of the predicted 2000–3000 chloroplast protein is within reach. The shift in analytical methods in proteomics from 2D gels toward 2D LC-MS/MS based strategies, due to completely sequenced genomes, improved nano-LC systems and faster and more sensitive tandem mass spectrometers has over the years increased the output of proteomics data. We foresee that new robust in-solution digestion protocol coupled with fast online 2D LC-MS/MS systems will enable the next major step in barley proteomics by decreasing workload and increasing the throughput, identification rate and accuracy of quantitation of the proteomics technologies.

In the near future we expect to see more quantitative proteomics studies of barley, e.g., for molecular analysis of abiotic stress, where sensitive versus non-sensitive barley genotypes are compared, with the aim of identifying protein biomarker involved in a certain genotypic trait. Ultimately this would couple proteomics and other technologies into the multidisciplinary systems biology platform in the pursuit of sustainable crop production.

ACKNOWLEDGMENT

This work was funded by The Danish Council for Strategic Research grant number 10-093498 for the research project "NutriEfficient".

REFERENCES

Abdalla, K. O., and Rafudeen, M. S. (2012). Analysis of the nuclear proteome of the resurrection plant *Xerophyta viscosa* in response to dehydration stress using iTRAQ with 2DLC and tandem mass spectrometry. *J. Proteomics* 75, 2361–2374.

Aebersold, R., and Mann, M. (2003). Mass spectrometry-based proteomics. *Nature* 422, 198–207.

Agrawal, G. K., Bourguignon, J., Rolland, N., Ephritikhine, G., Ferro, M., Jaquinod, M., et al. (2011). Plant organelle proteomics: collaborating for optimal cell function. *Mass Spectrom. Rev.* 30, 772–853.

Armbruster, U., Pesaresi, P., Pribil, M., Hertle, A., and Leister, D. (2011). Update on chloroplast research: new tools, new topics, and new trends. *Mol. Plant* 4, 1–16.

Aronsson, H., and Jarvis, P. (2002). A simple method for isolating import-competent *Arabidopsis* chloroplasts. *FEBS Lett.* 529, 215–220.

Ayoubi, T. A. Y., and Van De Ven, W. J. M. (1996). Regulation of gene expression by alternative promoters. *FASEB J.* 10, 453–460.

Bannai, H., Tamada, Y., Maruyama, O., Nakai, K., and Miyano, S. (2002). Extensive feature detection of N-terminal protein sorting signals. *Bioinformatics* 18, 298–305.

Bantscheff, M., Lemeer, S., Savitski, M. M., and Kuster, B. (2012). Quantitative mass spectrometry in proteomics: critical review update from 2007 to the present. *Anal. Bioanal. Chem.* 404, 939–965.

Bantscheff, M., Schirle, M., Sweetman, G., Rick, J., and Kuster, B. (2007). Quantitative mass spectrometry in proteomics: a critical review. *Anal. Bioanal. Chem.* 389, 1017–1031.

Bartsch, S., Monnet, J., Selbach, K., Quigley, F., Gray, J., von Wettstein, D., et al. (2008). Three thioredoxin targets in the inner envelope membrane of chloroplasts function in protein import and chlorophyll metabolism. *Proc. Natl. Acad. Sci. U.S.A.* 105, 4933–4938.

Bensimon, A., Heck, A. J., and Aebersold, R. (2012). Mass spectrometry-based proteomics and network biology. *Annu. Rev. Biochem.* 81, 379–405.

Berth, M., Moser, F. M., Kolbe, M., and Bernhardt, J. (2007). The state of the art in the analysis of two-dimensional gel electrophoresis images. *Appl. Microbiol. Biotechnol.* 76, 1223–1243.

Bindschedler, L. V., and Cramer, R. (2011). Quantitative plant proteomics. *Proteomics* 11, 756–775.

Bindschedler, L. V., Palmblad, M., and Cramer, R. (2008). Hydroponic isotope labelling of entire plants (HILEP) for quantitative plant proteomics; an oxidative stress case study. *Phytochemistry* 69, 1962–1972.

Boden, M., and Hawkins, J. (2005). Prediction of subcellular localization using sequence-biased recurrent networks. *Bioinformatics* 21, 2279–2286.

Boschetti, E., and Righetti, P. G. (2008). The ProteoMiner in the proteomic arena: a non-depleting tool for discovering low-abundance species. *J. Proteomics* 71, 255–264.

Cellar, N. A., Kuppannan, K., Langhorst, M. L., Ni, W., Xu, P., and Young, S. A. (2008). Cross species applicability of abundant protein depletion columns for ribulose-1,5-bisphosphate carboxylase/oxygenase. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 861, 29–39.

Chen, M. J., Mooney, B. P., Hajduch, M., Joshi, T., Zhou, M. Y., Xu, D., et al. (2009). System analysis of an *Arabidopsis* mutant altered in de novo fatty acid synthesis reveals diverse changes in seed composition and metabolism. *Plant Physiol.* 150, 27–41.

Chevalier, F. (2010). Highlights on the capacities of "Gel-based" proteomics. *Proteome Sci.* 8, 23.

Ciambella, C., Roepstorff, P., Aro, E. M., and Zolla, L. (2005). A proteomic approach for investigation of photosynthetic apparatus in plants. *Proteomics* 5, 746–757.

Cottrell, J. S. (2011). Protein identification using MS/MS data. *J. Proteomics* 74, 1842–1851.

Cravatt, B. F., Simon, G. M., and Yates, J. R. III. (2007). The biological impact of mass-spectrometry-based proteomics. *Nature* 450, 991–1000.

de Hoog, C. L., and Mann, M. (2004). Proteomics. *Annu. Rev. Genomics Hum. Genet.* 5, 267–293.

Dreher, K., and Callis, J. (2007). Ubiquitin, hormones and biotic stress in plants. *Ann. Bot.* 99, 787–822.

Emanuelsson, O., Nielsen, H., and Von Heijne, G. (1999). ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Sci.* 8, 978–984.

Endler, A., Reiland, S., Gerrits, B., Schmidt, U. G., Baginsky, S., and Martinoia, E. (2009). In vivo phosphorylation sites of barley tonoplast proteins identified by a phosphoproteomic approach. *Proteomics* 9, 310–321.

Eriksson, J., and Fenyo, D. (2010). "Modeling experimental design for proteomics," in *Computational Biology*, ed. D. Fenyo (Totowa: Humana Press Inc.), 223–230.

Fasoli, E., Aldini, G., Regazzoni, L., Kravchuk, A. V., Citterio, A., and Righetti, P. G. (2010). Les maitres de l'Orge: the proteome content of your beer mug. *J. Proteome Res.* 9, 5262–5269.

Fatehi, F., Hosseinzadeh, A., Alizadeh, H., Brimavandi, T., and Struik, P. C. (2012). The proteome response of salt-resistant and salt-sensitive barley genotypes to long-term salinity stress. *Mol. Biol. Rep.* 39, 6387–6397.

Ferro, M., Brugiere, S., Salvi, D., Seigneurin-Berny, D., Court, M., Moyet, L., et al. (2010). AT_CHLORO, a comprehensive chloroplast proteome database with subplastidial localization and curated information on envelope proteins. *Mol. Cell. Proteomics* 9, 1063–1084.

Ferro, M., Salvi, D., Brugiere, S., Miras, S., Kowalski, S., Louwagie, M., et al. (2003). Proteomics of the chloroplast envelope membranes from *Arabidopsis thaliana*. *Mol. Cell. Proteomics* 2, 325–345.

Fido, R. J., Mills, E. N., Rigby, N. M., and Shewry, P. R. (2004). Protein extraction from plant tissues. *Methods Mol. Biol.* 244, 21–27.

Finnie, C., Bak-Jensen, K. S., Laugesen, S., Roepstorff, P., and Svensson, B. (2006). Differential appearance of

isoforms and cultivar variation in protein temporal profiles revealed in the maturing barley grain proteome. *Plant Sci.* 170, 808–821.

Finnie, C., Melchior, S., Roepstorff, P., and Svensson, B. (2002). Proteome analysis of grain filling and seed maturation in barley. *Plant Physiol.* 129, 1308–1319.

Friso, G., Giacomelli, L., Ytterberg, A. J., Peltier, J. B., Rudella, A., Sun, Q., et al. (2004). In-depth analysis of the thylakoid membrane proteome of *Arabidopsis thaliana* chloroplasts: new proteins, new functions, and a plastid proteome database. *Plant Cell* 16, 478–499.

Froehlich, J. E., Wilkerson, C. G., Ray, W. K., McAndrew, R. S., Osteryoung, K. W., Gage, D. A., et al. (2003). Proteomic study of the *Arabidopsis thaliana* chloroplastic envelope membrane utilizing alternatives to traditional two-dimensional electrophoresis. *J. Proteome Res.* 2, 413–425.

Frohlich, A., Gaupels, F., Sarioglu, H., Holzmeister, C., Spannagl, M., Durner, J., et al. (2012). Looking deep inside: detection of low-abundance proteins in leaf extracts of *Arabidopsis* and phloem exudates of pumpkin. *Plant Physiol.* 159, 902–914.

Gerber, S. A., Rush, J., Stemman, O., Kirschner, M. W., and Gygi, S. P. (2003). Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc. Natl. Acad. Sci. U.S.A.* 100, 6940–6945.

Gevaert, K., and Vandekerckhove, J. (2000). Protein identification methods in proteomics. *Electrophoresis* 21, 1145–1154.

Giacomelli, L., Rudella, A., and van Wijk, K. J. (2006). High light response of the thylakoid proteome in *Arabidopsis* wild type and the ascorbate-deficient mutant vtc2-2. A comparative proteomics study. *Plant Physiol.* 141, 685–701.

Gilar, M., Olivova, P., Daly, A. E., and Gebler, J. C. (2005a). Orthogonality of separation in two-dimensional liquid chromatography. *Anal. Chem.* 77, 6426–6434.

Gilar, M., Olivova, P., Daly, A. E., and Gebler, J. C. (2005b). Two-dimensional separation of peptides using RP-RP-HPLC system with different pH in first and second separation dimensions. *J. Sep. Sci.* 28, 1694–1703.

Goff, S. A., Ricke, D., Lan, T. H., Presting, G., Wang, R. L., Dunn, M., et al. (2002). A draft sequence of the rice genome (*Oryza sativa* L. ssp *japonica*). *Science* 296, 92–100.

Gouw, J. W., Tops, B. B., Mortensen, P., Heck, A. J., and Krijgsveld, J. (2008). Optimizing identification and quantitation of 15N-labeled proteins in comparative proteomics. *Anal. Chem.* 80, 7796–7803.

Granvogl, B., Reisinger, V., and Eichacker, L. A. (2006). Mapping the proteome of thylakoid membranes by de novo sequencing of intermembrane peptide domains. *Proteomics* 6, 3681–3695.

Gruhler, A., Schulze, W. X., Matthiesen, R., Mann, M., and Jensen, O. N. (2005). Stable isotope labeling of *Arabidopsis thaliana* cells and quantitative proteomics by mass spectrometry. *Mol. Cell. Proteomics* 4, 1697–1709.

Hall, M., Mishra, Y., and Schröder, W. P. (2011). "Preparation of stroma, thylakoid membrane, and lumen fractions from *Arabidopsis thaliana* chloroplasts for proteomic analysis." in *Chloroplast Research in Arabidopsis*, ed. R. Paul Jarvis (New York: Humana Press), 207–222.

Horton, P., Park, K. J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C. J., et al. (2007). WoLF PSORT: protein localization predictor. *Nucleic Acids Res.* 35, W585–W587.

Hynek, R., Svensson, B., Jensen, O. N., Barkholt, V., and Finnie, C. (2009). The plasma membrane proteome of germinating barley embryos. *Proteomics* 9, 3787–3794.

Iimure, T., Nankaku, N., Hirota, N., Zhou, T. S., Hoki, T., Kihara, M., et al. (2010). Construction of a novel beer proteome map and its use in beer quality control. *Food Chem.* 118, 566–574.

Issaq, H., and Veenstra, T. (2008). Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE): advances and perspectives. *Biotechniques* 44, 697–698, 700.

Jacob, A. M., and Turck, C. W. (2008). Detection of post-translational modifications by fluorescent staining of two-dimensional gels. *Methods Mol. Biol.* 446, 21–32.

James, P. (1997). Protein identification in the post-genome era: the rapid rise of proteomics. *Q. Rev. Biophys.* 30, 279–331.

Jones, A. M. E., Bennett, M. H., Mansfield, J. W., and Grant, M. (2006). Analysis of the defence phosphoproteome of *Arabidopsis thaliana* using differential mass tagging. *Proteomics* 6, 4155–4165.

Jorrin, J. V., Maldonado, A. M., and Castillejo, M. A. (2007). Plant proteome analysis: a 2006 update. *Proteomics* 7, 2947–2962.

Kamal, A. M., Cho, K., Komatsu, S., Uozumi, N., Choi, J. S., and Woo, S. H. (2012). Towards an understanding of wheat chloroplasts: a methodical investigation of thylakoid proteome. *Mol. Biol. Rep.* 39, 5069–5083.

Kaul, S., Koo, H. L., Jenkins, J., Rizzo, M., Rooney, T., Tallon, L. J., et al. (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408, 796–815.

Kieselbach, T., Bystedt, M., Hynds, P., Robinson, C., and Schroder, W. P. (2000). A peroxidase homologue and novel plastocyanin located by proteomics to the *Arabidopsis* chloroplast thylakoid lumen. *FEBS Lett.* 480, 271–276.

Kieselbach, T., Hagman, Andersson, B., and Schroder, W. P. (1998). The thylakoid lumen of chloroplasts. Isolation and characterization. *J. Biol. Chem.* 273, 6710–6716.

Kleffmann, T., Russenberger, D., von Zychlinski, A., Christopher, W., Sjolander, K., Gruissem, W., et al. (2004). The *Arabidopsis thaliana* chloroplast proteome reveals pathway abundance and novel protein functions. *Curr. Biol.* 14, 354–362.

Kline-Jonakin, K. G., Barrett-Wilt, G. A., and Sussman, M. R. (2011). Quantitative plant phosphoproteomics. *Curr. Opin. Plant Biol.* 14, 507–511.

Krause, F. (2006). Detection and analysis of protein-protein interactions in organellar and prokaryotic proteomes by native gel electrophoresis: (Membrane) protein complexes and supercomplexes. *Electrophoresis* 27, 2759–2781.

Krishnan, H. B., and Natarajan, S. S. (2009). A rapid method for depletion of Rubisco from soybean (*Glycine max*) leaf for proteomic analysis of lower abundance proteins. *Phytochemistry* 70, 1958–1964.

Kuntz, M., and Rolland, N. (2012). "Subcellular and sub-organellar proteomics as a complementary tool to study the evolution of the plastid proteome," *Organelle Genetics: Evolution of Organelle Genomes and Gene Expression*, ed. C. E. Bullerwell (New York: Springer).

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., et al. (2001). Initial sequencing and analysis of the human genome. *Nature* 409, 860–921.

Larsen, M. R., Trelle, M. B., Thingholm, T. E., and Jensen, O. N. (2006). Analysis of posttranslational modifications of proteins by tandem mass spectrometry. *Biotechniques* 40, 790–798.

Laugesen, S., Bak-Jensen, K. S., Hagglund, P., Henriksen, A., Finnie, C., Svensson, B., et al. (2007). Barley peroxidase isozymes – expression and post-translational modification in mature seeds as identified by two-dimensional gel electrophoresis and mass spectrometry. *Int. J. Mass Spectrom.* 268, 244–253.

Lemeer, S., Hahne, H., Pachl, F., and Kuster, B. (2012). Software tools for MS-based quantitative proteomics: a brief overview. *Methods Mol. Biol.* 893, 489–499.

Lin, Y., Liu, H., Liu, Z. H., Wang, X. C., and Liang, S. P. (2012). Shotgun analysis of membrane proteomes using a novel combinative strategy of solution-based sample preparation coupled with liquid chromatography-tandem mass spectrometry. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 901, 18–24.

Mano, S., Miwa, T., Nishikawa, S., Mimura, T., and Nishimura, M. (2008). The plant organelles database (PODB): a collection of visualized plant organelles and protocols for plant organelle research. *Nucleic Acids Res.* 36, 929–937.

Manza, L. L., Stamer, S. L., Ham, A. J. L., Codreanu, S. G., and Liebler, D. C. (2005). Sample preparation and digestion for proteomic analyses using spin filters. *Proteomics* 5, 1742–1745.

Mayer, K. F., Waugh, R., Langridge, P., Close, T. J., Wise, R. P., Graner, A., et al. (2012). A physical, genetic and functional sequence assembly of the barley genome. *Nature.* 491, 711–716.

Melo-Braga, M. N., Verano-Braga, T., Leon, I. R., Antonacci, D., Nogueira, F. C. S., Thelen, J. J., et al. (2012). Modulation of protein phosphorylation, *N*-glycosylation and Lys-acetylation in grape (*Vitis vinifera*) mesocarp and exocarp owing to *Lobesia botrana* infection. *Mol. Cell. Proteomics* 11, 945–956.

Metodiev, M., and Demirevskakepova, K. (1992). Rbisco quantitation in leaves of different barley varieties by enzyme-linked-immunosorbent-assay. *J. Exp. Bot.* 43, 155–158.

Millar, A. H., Heazlewood, J. L., Kristensen, B. K., Braun, H. P., and Moller, I. M. (2005). The plant mitochondrial proteome. *Trends Plant Sci.* 10, 36–43.

Mithoe, S. C., and Menke, F. L. H. (2011). Phosphoproteomics perspective on plant signal transduction and tyrosine phosphorylation. *Phytochemistry* 72, 997–1006.

Moller, A. L. B., Pedas, P., Andersen, B., Svensson, B., Schjoerring, J. K., and Finnie, C. (2011a). Responses of barley root and shoot proteomes to long-term nitrogen deficiency, short-term nitrogen starvation and ammonium. *Plant Cell Environ.* 34, 2024–2037.

Moller, I. M., Rogowska-Wrzesinska, A., and Rao, R. S. P. (2011b). Protein carbonylation and metal-catalyzed protein oxidation in a cellular perspective. *J. Proteomics* 74, 2228–2242.

Nawrocki, A., Thorup-Kristensen, K., and Jensen, O. N. (2011). Quantitative proteomics by 2DE and MALDI MS/MS uncover the effects of organic and conventional cropping methods on vegetable products. *J. Proteomics* 74, 2810–2825.

Neilson, K. A., Mariani, M., and Haynes, P. A. (2011). Quantitative proteomic analysis of cold-responsive proteins in rice. *Proteomics* 11, 1696–1706.

Nelson, C. J., Huttlin, E. L., Hegeman, A. D., Harms, A. C., and Sussman, M. R. (2007). Implications of N-15-metabolic labeling for automated peptide identification in *Arabidopsis thaliana*. *Proteomics* 7, 1279–1292.

Neuburger, M., Journet, E. P., Bligny, R., Carde, J. P., and Douce, R. (1982). Purification of plant-mitochondria by isopycnic centrifugation in density gradients of percollLL. *Arch. Biochem. Biophys.* 217, 312–323.

Ng, D. W. K., Zhang, C., Miller, M., Shen, Z., Briggs, S. P., and Chen, Z. J. (2012). Proteomic divergence in *Arabidopsis* autopolyploids and allopolyploids and their progenitors. *Heredity (Edinb)* 108, 419–430.

Norrgran, J., Williams, T. L., Woolfitt, A. R., Solano, M. I., Pirkle, J. L., and Barr, J. R. (2009). Optimization of digestion parameters for protein quantification. *Anal. Biochem.* 393, 48–55.

Okada, Y., Limure, T., Takoi, K., Kaneko, T., Kihara, M., Hayashi, K., et al. (2008). The influence of barley malt protein modification on beer foam stability and their relationship to the barley dimeric alpha-amylase inhibitor-1 (BDAI-1) as a possible foam-promoting protein. *J. Agric. Food Chem.* 56, 1458–1464.

Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., Steen, H., Pandey, A., et al. (2002). Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics* 1, 376–386.

Ong, S. E., and Mann, M. (2005). Mass spectrometry-based proteomics turns quantitative. *Nat. Chem. Biol.* 1, 252–262.

Ostergaard, O., Melchior, S., Roepstorff, P., and Svensson, B. (2002). Initial proteome analysis of mature barley seeds and malt. *Proteomics* 2, 733–739.

Patterson, J., Ford, K., Cassin, A., Natera, S., and Bacic, A. (2007). Increased abundance of proteins involved in phytosiderophore production in boron-tolerant barley. *Plant Physiol.* 144, 1612–1631.

Patterson, S. D., and Aebersold, R. H. (2003). Proteomics: the first decade and beyond. *Nat. Genet.* 33, 311–323.

Patton, W. F. (2002). Detection technologies in proteome analysis. *J. Chromatogr B Analyt. Technol. Biomed. Life Sci.* 771, 3–31.

Peltier, J. B., Cai, Y., Sun, Q., Zabrouskov, V., Giacomelli, L., Rudella, A., et al. (2006). The oligomeric stromal proteome of *Arabidopsis thaliana* chloroplasts. *Mol. Cell. Proteomics* 5, 114–133.

Peltier, J. B., Emanuelsson, O., Kalume, D. E., Ytterberg, J., Friso, G., Rudella, A., et al. (2002). Central functions of the lumenal and peripheral thylakoid proteome of *Arabidopsis* determined by experimentation and genome-wide prediction. *Plant Cell* 14, 211–236.

Peltier, J. B., Friso, G., Kalume, D. E., Roepstorff, P., Nilsson, F., Adamska, I., et al. (2000). Proteomics of the chloroplast: systematic identification and targeting analysis of lumenal and peripheral thylakoid proteins. *Plant Cell* 12, 319–341.

Peltier, J. B., Ytterberg, A. J., Sun, Q., and van Wijk, K. J. (2004). New functions of the thylakoid membrane proteome of *Arabidopsis thaliana* revealed by a simple, fast, and versatile fractionation strategy. *J. Biol. Chem.* 279, 49367–49383.

Perrocheau, L., Rogniaux, H., Boivin, P., and Marion, D. (2005). Probing heat-stable water-soluble proteins from barley to malt and beer. *Proteomics* 5, 2849–2858.

Petry-Podgorska, I., Zidkova, J., Flodrova, D., and Bobalova, J. (2010). 2D-HPLC and MALDI-TOF/TOF analysis of barley proteins glycated during brewing. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 878, 3143–3148.

Ploscher, M., Granvogl, B., Zoryan, M., Reisinger, V., and Eichacker, L. A. (2009). Mass spectrometric characterization of membrane integral low molecular weight proteins from photosystem II in barley etioplasts. *Proteomics,* 9 625–635.

Ploscher, M., Reisinger, V., and Eichacker, L. A. (2011). Proteomic comparison of etioplast and chloroplast protein complexes. *J. Proteomics* 74, 1256–1265.

Podwojski, K., Stephan, C., and Eisenacher, M. (2012). Important issues in planning a proteomics experiment: statistical considerations of quantitative proteomic data. *Methods Mol. Biol.* 893, 3–21.

Pottiez, G., Wiederin, J., Fox, H. S., and Ciborowski, P. (2012). Comparison of 4-plex to 8-plex iTRAQ quantitative measurements of proteins in human plasma samples. *J. Proteome Res.* 11, 3774–3781.

Reisinger, V., and Eichacker, L. A. (2007). How to analyze protein complexes by 2D blue native SDS-PAGE. *Proteomics* 7(Suppl. 1), 6–16.

Ross, P. L., Huang, Y. L. N., Marchese, J. N., Williamson, B., Parker, K., Hattan, S., et al. (2004). Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol. Cell. Proteomics* 3, 1154–1169.

Sachon, E., Mohammed, S., Bache, N., and Jensen, O. N. (2006). Phosphopeptide quantitation using amine-reactive isobaric tagging reagents and tandem mass spectrometry: application to proteins isolated by gel electrophoresis. *Rapid Commun. Mass Spectrom.* 20, 1127–1134.

Saeed, M., Dahab, A. H. A., Guo, W. Z., and Zhang, T. Z. (2012). A cascade of recently discovered molecular mechanisms involved in abiotic stress tolerance of plants. *OMICS* 16, 188–199.

Saisho, D., and Takeda, K. (2011). Barley: emergence as a new research material of crop science. *Plant Cell Physiol.* 52, 724–727.

Schmidt, A., Kellermann, J., and Lottspeich, F. (2005). A novel strategy for quantitative proteomics using isotope-coded protein labels. *Proteomics* 5, 4–15.

Schubert, M., Petersson, U. A., Haas, B. J., Funk, C., Schroder, W. P., and Kieselbach, T. (2002). Proteome map of the chloroplast lumen of *Arabidopsis thaliana*. *J. Biol. Chem.* 277, 8354–8365.

Schulte, D., Close, T. J., Graner, A., Langridge, P., Matsumoto, T., Muehlbauer, G., et al. (2009). The international barley sequencing consortium-at the threshold of efficient access to the barley genome. *Plant Physiol.* 149, 142–147.

Shevchenko, A., Tomas, H., Havlis, J., Olsen, J. V., and Mann, M. (2006). In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protoc.* 1, 2856–2860.

Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996). Mass spectrometric sequencing of proteins from silver stained polyacrylamide gels. *Anal. Chem.* 68, 850–858.

Silva, J. C., Gorenstein, M. V., Li, G. Z., Vissers, J. P. C., and Geromanos, S. J. (2006). Absolute quantification of proteins by LCMSE – a virtue of parallel MS acquisition. *Mol. Cell. Proteomics* 5, 144–156.

Small, I., Peeters, N., Legeai, F., and Lurin, C. (2004). Predotar: a tool for rapidly screening proteomes for N-terminal targeting sequences. *Proteomics* 4, 1581–1590.

Speers, A. E., and Wu, C. C. (2007). Proteomics of integral membrane proteins theory and application. *Chem. Rev.* 107, 3687–3714.

Syed, N. H., Kalyna, M., Marquez, Y., Barta, A., and Brown, J. W. S. (2012). Alternative splicing in plants – coming of age. *Trends Plant Sci.* 17, 616–623.

Thelen, J. J., and Peck, S. C. (2007). Quantitative proteomics in plants: choices in abundance. *Plant Cell* 19, 3339–3346.

van Bentem, S. D., Roitinger, E., Anrather, D., Csaszar, E., and Hirt, H. (2006). Phosphoproteomics as a tool to unravel plant regulatory mechanisms. *Physiol. Plant.* 126, 110–119.

van Wijk, K. J. (2004). Plastid proteomics. *Plant Physiol. Biochem.* 42, 963–977.

Washburn, M. P., Wolters, D., and Yates, J. R. (2001). Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat. Biotechnol.* 19, 242–247.

Wendelboe-Nelson, C., and Morris, P. C. (2012). Proteins linked to drought tolerance revealed by DIGE analysis of drought resistant and susceptible barley varieties. *Proteomics* 12, 3374–3385.

Wienkoop, S., Larrainzar, E., Niemann, M., Gonzalez, E. M., Lehmann, U., and Weckwerth, W. (2006). Stable isotope-free quantitative shotgun proteomics combined with sample pattern recognition for rapid diagnostics. *J. Sep. Sci.* 29, 2793–2801.

Wiese, S., Reidegeld, K. A., Meyer, H. E., and Warscheid, B. (2007). Protein labeling by iTRAQ: A new tool for quantitative mass spectrometry in proteome research. *Proteomics* 7, 340–350.

Wisniewski, J. R., Zougman, A., Nagaraj, N., and Mann, M. (2009). Universal sample preparation method for

proteome analysis. *Nat. Methods* 6, 359–362.

Wortman, J. R., Haas, B. J., Hannick, L. I., Smith, R. K., Maiti, R., Ronning, C. M., et al. (2003). Annotation of the *Arabidopsis* genome. *Plant Physiol.* 132, 461–468.

Xiao, Z., Conrads, T. P., Lucas, D. A., Janini, G. M., Schaefer, C. F., Buetow, K. H., et al. (2004). Direct ampholyte-free liquid-phase isoelectric peptide focusing: application to the human serum proteome. *Electrophoresis* 25, 128–133.

Yeung, Y. G., and Stanley, E. R. (2010). Rapid detergent removal from peptide samples with ethyl acetate for mass spectrometry analysis. *Curr. Protoc. Protein Sci.* Chapter 16:Unit 16.12.

Ytterberg, A. J., and Jensen, O. N. (2010). Modification-specific proteomics in plant biology. *J. Proteomics* 73, 2249–2266.

Zybailov, B., Rutschow, H., Friso, G., Rudella, A., Emanuelsson, O., Sun, Q., et al. (2008). Sorting signals, N-terminal modifications and abundance of the chloroplast proteome. *PLoS ONE* 3:e1994. doi: 10.1371/journal.pone.0001994

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 10 January 2013; paper pending published: 30 January 2013; accepted: 26 February 2013; published online: 18 March 2013.

Citation: Petersen J, Rogowska-Wrzesinska A and Jensen ON (2013) Functional proteomics of barley and barley chloroplasts – strategies, methods and perspectives. Front. Plant Sci. ***4****:52. doi: 10.3389/fpls.2013.00052*

This article was submitted to Frontiers in Plant Proteomics, a specialty of Frontiers in Plant Science.

Copyright © 2013 Petersen, Rogowska-Wrzesinska and Jensen. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.

frontiers in
PLANT SCIENCE

REVIEW ARTICLE
published: 11 January 2013
doi: 10.3389/fpls.2012.00307

Dissecting plasmodesmata molecular composition by mass spectrometry-based proteomics

Magali S. Salmon and Emmanuelle M. F. Bayer*

Laboratory of Membrane Biogenesis, CNRS UMR5200, University of Bordeaux, Bordeaux, France

Edited by:
Nicolas L. Taylor, The University of Western Australia, Australia

Reviewed by:
Subhra Chakraborty, National Institute of Plant Genome Research, India
Borjana Arsova, Heinrich-Heine University, Germany

****Correspondence:***
Emmanuelle M. F. Bayer, Laboratory of Membrane Biogenesis, CNRS UMR5200, Campus INRA de Bordeaux, 71 Avenue E. Bourlaux, 33883 Villenave d'Ornon Cedex, France.
e-mail: emmanuelle.bayer@u-bordeaux2.fr

In plants, the intercellular communication through the membranous channels called plasmodesmata (PD; singular plasmodesma) plays pivotal roles in the orchestration of development, defence responses, and viral propagation. PD are dynamic structures embedded in the plant cell wall that are defined by specialized domains of the endoplasmic reticulum (ER) and the plasma membrane (PM). PD structure and unique functions are guaranteed by their particular molecular composition. Yet, up to recent years and despite numerous approaches such as mutant screens, immunolocalization, or screening of random cDNAs, only few PD proteins had been conclusively identified and characterized. A clear breakthrough in the search of PD constituents came from mass-spectrometry-based proteomic approaches coupled with subcellular fractionation strategies. Due to their position, firmly anchored in the extracellular matrix, PD are notoriously difficult to isolate for biochemical analysis. Proteomic-based approaches have therefore first relied on the use of cell wall fractions containing embedded PD then on "free" PD fractions whereby PD membranes were released from the walls by enzymatic degradation. To discriminate between likely contaminants and PD protein candidates, bioinformatics tools have often been used in combination with proteomic approaches. GFP fusion proteins of selected candidates have confirmed the PD association of several protein families. Here we review the accomplishments and limitations of the proteomic-based strategies to unravel the functional and structural complexity of PD. We also discuss the role of the identified PD-associated proteins.

Keywords: plasmodesmata, wall, proteomics, subcellular fractionation, *Arabidopsis* suspension cells

INTRODUCTION

In plants, intercellular communication must overcome the rigid pectocellulosic wall that encompasses all cells. To achieve that plants have developed membranous pores called plasmodesmata (PD) that perforate the extracellular matrix providing symplastic connections between most cell types (Maule, 2008; Xu and Jackson, 2010; Maule et al., 2011). PD are central to a wide range of biological processes that require cell-to-cell communication such as cell fate specification, coordinated growth and development, and transport of carbohydrates. Plant viruses but also fungus can exploit PD transport machinery to establish infection. The emerging view is that PD may well represent a consensus target for pathogens and play a crucial role in defense signaling (Kankanala et al., 2007; Lee and Lu, 2011; Lee et al., 2011). Data regarding PD structure mainly derives from electron microscopy (Helper, 1982; Overall et al., 1982; Tilney et al., 1991; Ding et al., 1992; Botha et al., 1993). PD are lined by the plasma membrane (PM) and contain a central rod, the desmotubule, which is derived from, and continuous with, the endoplasmic reticulum (ER) (**Figure 1**). Both membrane domains are linked by bridging-like elements whose identity remains a matter of speculation. The space between the PM and the desmotubule is called the cytoplasmic sleeve and provides a conduit through which molecules below the size exclusion limit (SEL) can diffuse between cells in either soluble form or laterally within the membrane phases. Although PD guarantee both cytosolic and membrane continuity between plant cells, the exchange of molecules is under tight control. Non-selective trafficking through diffusion hinges on the number and SEL of PD at a given cellular interface. Both parameters vary depending on the cell type and developmental stage of the tissue considered. An additional level of regulation involves the selective trafficking of specific macromolecules whose size is above the SEL. Such targeted movement implies direct interaction between the trafficking cargo and PD components and results in transient opening of the channels. Understanding of how PD dictate cellular connectivity in such circumstances is dependent on comprehensive knowledge of the composition of PD and functional characterization of their constituents.

THE LONG QUEST FOR PLASMODESMAL PROTEIN CONSTITUENTS

For a long time, the sparse information available about PD constituents has hindered progress in our understanding as to how these membranous structures function. Over the last 30 years the search for PD proteins has been a constant topic of research and endeavors to identify them have employed a wide diversity of approaches (Faulkner and Maule, 2011). Genetic-based

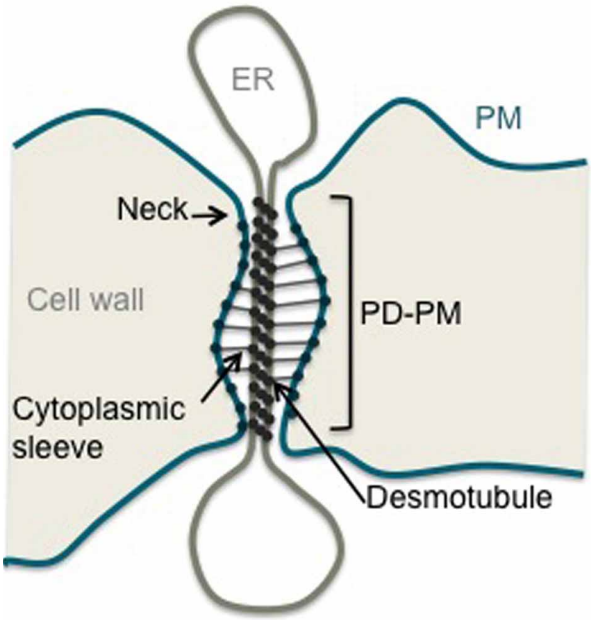


FIGURE 1 | Structure of a plasmodesma. Schematic representation of a simple stranded PD. ER, Endoplasmic Reticulum; PM, Plasma Membrane; PD-PM, Plasma Membrane lining PD.

approaches have failed to divulge PD structural and regulatory components; this is likely due to the critical role that PD play in growth and development. However, they have supplied critical guidance toward PD functional mechanisms by enabling the identification of proteins, such as a m-type thioredoxin or RNA helicases, which impact on PD permeability but are localized in other subcellular compartments (Kobayashi et al., 2007; Benitez-Alfonso et al., 2009; Stonebloom et al., 2009; Guseman et al., 2010). Targeted approaches aimed at identifying PD receptors have taken advantage of viral movement proteins which accumulate at PD and modify their SEL to permit virus transfer (Benitez-Alfonso et al., 2010). Screens were developed using viral proteins as baits but yielded limited success (Citovsky et al., 1993; Kragler et al., 2000; Paape et al., 2006). Unexpectingly, immunolocalization strategies turned out to be relatively successful. The idea was to identify proteins with established functions that associated with PD. Notably, a close association between PD and elements of the cytoskeleton, especially actin and myosin, were revealed (White et al., 1994; Blackman and Overall, 1998; Radford and White, 1998; Reichelt et al., 1999). They have since been shown to have critical roles in the regulation of cell-to-cell movement and control of PD SEL (White et al., 1994; Ding et al., 1996; Su et al., 2010; White and Barton, 2011; Deeks et al., 2012). Immunological approaches were nevertheless limited to known proteins with available antibodies, and did not lead to unambiguous protein identification.

The need to identify novel PD proteins lead to the development of high throughput screens. Plant cDNAs libraries fused to the fluorescent tag GFP were utilized to this end (Cutler et al., 2000; Escobar et al., 2003). While theoretically appealing, these approaches did not succeed in identifying PD proteins. A different approach for the identification of PD components was required, shifting the focus to the potential for biochemical isolation and proteomic analysis of PD-enriched fractions.

PURIFYING PD-ENRICHED SUBCELLULAR FRACTIONS: FIRST STEPS TOWARD THE HOLY GRAIL

Access to PD structures by subcellular fractionation is rendered difficult both by their location, embedded in the extracellular matrix, and by the small physical contribution they make to total plant tissue mass. In fact, PD are not simply inserted into the wall but firmly anchored into it, probably through the action of proteins and/or wall polymers, that would provide stable bridges between the PM and the wall (Brecknock et al., 2011). Even during an intense plasmolysis treatment, PD stay embedded in the wall matrix while the protoplast retracts (Tilney et al., 1991). However, what was first viewed as a hurdle to PD isolation turned out to be a major advantage. Thus, PD-enriched fractions were readily obtained by purifying wall fragments from plant tissues by mechanical disruption of tissues (French Press, N_2 pressure bomb, grinding in liquid nitrogen) followed by successive low speed centrifugations to recover and wash wall fragments.

The first attempts to identify PD-associated proteins from purified cell walls, relied on plant tissues known to be rich in PD (Monzer and Kloth, 1991; Kotlizky et al., 1992; Turner et al., 1994; Epel et al., 1995, 1996). With maize mesocotyls as source material, Epel et al. (1996) identified a 41 kDa protein enriched in wall extracts. Screening an expression library, the authors identified Reversibly Glycosylated Polypeptide 2 (RGP2) whose homolog in *Arabidopsis* was subsequently found to be enriched at PD (Sagi et al., 2005). Similarly, monoclonal antibodies raised against maize root tip cell wall proteins (JIM64 and JIM67) were shown to associate with PD in trichomes and mesophyll cells of *N. clevelandii* (Turner et al., 1994; Waigmann et al., 1997) but the identity of their antigen has not yet been retrieved.

Differentiated plant tissues however are often resistant to disruption making the preparation of pure cell wall fractions difficult. This potential drawback is of some importance as the identification of PD components lies in minimizing the level of contamination from intact cells, trapped subcellular organelles, or adhering membranes. As an alternative, the use of liquid cultured cells was investigated by several groups (Lee et al., 2003, 2005; Bayer et al., 2004, 2006; Fernandez-Calvino et al., 2011; Jo et al., 2011). Suspension cells provided an attractive system, as they comprise a friable population of relatively uniform, large cells that lay down abundant primary PD on division walls enabling the recovery of pure wall fractions, containing intact PD (Bayer et al., 2004; **Figure 2**). Moreover, the amount of plant material that could be processed is not a limiting factor. Using the non-cell-autonomous *Cucurbita maxima* phloem protein (CmPP16) as a bait, the group of Bill Lucas identified a Non-Cell-Autonomous-Protein-Pathway1 (NACPP1; Lee et al., 2003) and recently a Plasmodesmal Germin-like Protein1 (PDGLP1; Ham et al., 2012) from the PD-enriched

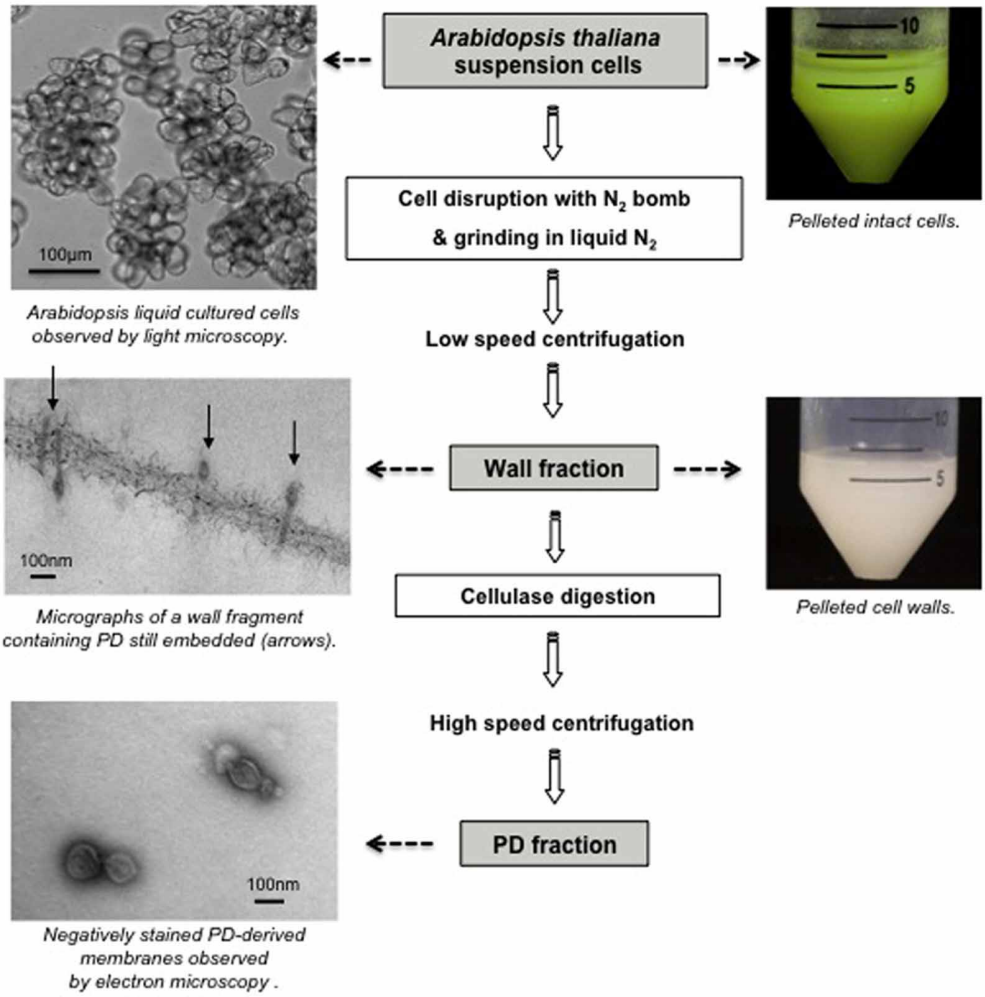


FIGURE 2 | Purification of PD-enriched wall fraction and "free" PD fraction from *Arabidopsis thaliana* suspension cells.

wall fraction of BY-2 cells. NCAPP1 associates to ER-domains close to the channels where it possibly acts as a shuttle for PD translocation. PDGLP proteins are PD-located and affect root growth when over expressed. Kinase activity essays on the same BY-2 subcellular fraction, lead to the identification of a PD-Associated Protein Kinase (PAPK) that was shown to phosphorylate the movement protein of tobacco mosaic virus (Lee et al., 2005).

With the aim of analyzing the proteome of PD-enriched fraction, Bayer et al. (2004) selected *A. thaliana* suspension culture owing to the extensive genomic information available. Although PD-enriched wall fractions have undoubtedly been of great value in the identification of PD constituents (Lee et al., 2003, 2005; Faulkner et al., 2005; Sagi et al., 2005; Thomas et al., 2008; Simpson et al., 2009; Jo et al., 2011), the contribution of PD proteins to the total wall protein extract was still relatively low. Success in isolating "free" PD from purified cell walls was first reported by Epel group (Epel et al., 1995), with the crucial advance being that PD-derived membranes were released from their position embedded in the wall by treatment with cellulase. This technique was used by Fernandez-Calvino et al. (2011) on *Arabidopsis* cell cultures and produced a final fraction with clear enrichment in known PD-proteins. Ultimately, biochemical fractionation of PD has presented the most straightforward and promising strategy for proteomic-based identification of PD components.

COMBINING SUBCELLULAR FRACTIONATION AND PROTEOMIC APPROACHES TO DEFINE THE PD PROTEOME

Proteomic analyses have emerged as powerful tools for large-scale analysis of complex protein mixtures. Combined with the development of subcellular fractionation strategies these approaches have permitted the identification of an unprecedented number of PD-associated proteins. These technologies have transformed what in the past could only be the result of laborious sequencing of few selected proteins enriched in wall or PD fractions, into a non-targeted approach whereby most, if not all, proteins present in a given sample could be identified.

A limited number of laboratories have actually explored proteomic technologies. Most research teams have only revealed the identity of "confirmed" PD proteins from their proteomic datasets (Sagi et al., 2005; Levy et al., 2007; Jo et al., 2011) but few groups made available the complete list of proteins identified from their PD-enriched fractions (Faulkner et al., 2005; Bayer et al., 2006; Fernandez-Calvino et al., 2011). These publically available databases certainly provide a rich source that can be exploited by all for further identification of PD proteins.

The most comprehensive proteomic analysis of PD proteins was undertaken by the Maule laboratory. Working with *Arabidopsis* suspension cells, the proteome of the wall fraction was first established (Bayer et al., 2006) and with the further refinement in the purification technique, that of the PD fraction (Fernandez-Calvino et al., 2011). Protein MS is coupled and highly dependent on separation strategies that simplify complex biological samples prior to application to the mass analyzer. Sufficient separation is required for both sensitivity and accuracy. Due to the likely hydrophobic nature of PD constituents, gel separation of wall extracts by means of 2D electrophoresis turned out to be inappropriate as most membrane proteins were not resolved (Bayer, unpublished). Instead, a non-gel approach, the Multidimensional Protein Identification Technology (MudPIT; Washburn et al., 2001), which consists of 2D liquid chromatography (2D-LC) directly coupled to a tandem MS, was used to analyze the total wall extract. The subsequent analysis of the PD fraction employed a nano-LC ion trap MS/MS method using an LTQ-Orbitrap™ analyzer that features high resolution, high mass accuracy, and a wide mass-to-charge range (Fernandez-Calvino et al., 2011). Both studies generated exhaustive lists of 792 and 1341 unique protein sequences for the wall and PD fractions, respectively, among which PD components are represented.

SELECTING PD POTENTIAL CANDIDATES FROM PROTEOMIC DATABASES

Sensitive proteomic detection systems have the potential to generate large datasets. Hundreds of proteins can be identified and even with relatively pure samples, minor contaminants are present and cannot be easily discriminated from the proteins of interest. Considering the methodology, what is gained by subcellular fractionation is partially lost by an increase in sensitivity.

To overcome these drawbacks, an elegant approach was developed by the Overall laboratory, who exploited the anatomy of the green alga *Chara corallina* (Blackman and Overall, 1998; Faulkner et al., 2005). The protein profile of wall extracts containing PD (nodal complexes) with those of walls without PD (external internodal walls) were compared by 2D electrophoresis and proteins unique to nodal complexes were analyzed by LC-MS/MS. Some showed sequence similarity to previously identified PD-associated proteins but the approach suffered from the absence of a sequenced genome. A similar approach would be difficult with land plant tissues as virtually all cells are connected with PD.

An alternative strategy consists on downstream analysis of the proteomic datasets generated using bioinformatic tools, databases, and literature sources. This approach was employed by the Maule laboratory following the establishment of *Arabidopsis* cell wall proteome, where PD components accounted for a small proportion of total proteins (Bayer et al., 2006). The selection of potential candidates had to rely on specific characteristics that would distinguish PD-associated proteins from "classical" wall proteins and cytoplasmic contaminants. Since little was known about the structure and function of PD, this was a largely subjective process of elimination. However, based on the nature of PD, the authors argued that a proportion of their protein components would be transported along the secretory pathway to reach either the desmotubule or the PM. Many PD proteins were also expected to be membrane-associated. Candidates were therefore selected based upon two main criteria. First, the preprotein sequence had to contain a N-terminal signal peptide for secretion via the ER and second, to be membrane-associated *via* either a transmembrane domain (TMD) or a Glycosyl Phosphatidyl Inositol (GPI) anchor. A conspicuous drawback of such selection strategy is that it precludes any PD proteins that would associate with PD by other means. A similar strategy was later on also applied to the *Arabidopsis* PD fraction which despite a major enrichment in PD-derived membranes gave rise a colossal proteomic dataset including likely contaminants (Fernandez-Calvino et al., 2011). Jo et al. (2011), who analyzed the wall proteome of rice callus cultures, also focused on membrane-associated proteins to identify PD constituents. The proteomic databases generated from *Arabidopsis* wall and PD fractions were searched using bioinformatic prediction programmes, databases, and published work. In each case about 10% of the proteins identified were shown to fulfill the criteria for PD association and were therefore elected for further analysis. Ultimate confirmation of the physical association of selected candidates with PD structures was then achieved through transient expression of GFP fusion products in leaves and eventually by immunolocalization with electron microscopy. So far, this approach resulted in the conclusive identification of several PD-associated proteins including Plasmodesmata Located Proteins (PDLP; Thomas et al., 2008), Plasmodesmal Callose Binding proteins (PDCB; Simpson et al., 2009), Receptor-Like Kinases (RLK; Fernandez-Calvino et al., 2011), and Tetraspanin (Fernandez-Calvino et al., 2011). We have compiled in **Table 1** all PD proteins that have been identified through subcellular fractionation and proteomic-based strategies and confirmed through GFP tagging or immunolocalization.

WHAT HAVE WE LEARNT FROM PROTEOMIC ANALYSIS?

These proteomic-based studies, combined with functional analysis of identified PD components, have greatly contributed to elucidate PD organization and regulatory principles. For instance, an interesting finding was that PD house receptor-like activities, such as receptor-like kinases (Fernandez-Calvino et al., 2011; Jo et al., 2011). This implies a role for the channels in signaling events and emphasizes the potential for extracellular stimuli to influence cell-to-cell communication. In the same vein, Thomas et al. (2008) identified from *Arabidopsis* cell wall extracts a new family of receptor-like transmembrane proteins named PDLP which were later on shown to act as receptors for viral movement proteins (Amari et al., 2010). An existing

Table 1 | List of confirmed PD-associated proteins identified through subcellular fractionation and proteomic analysis.

Protein	Gene	Biological material used for subcellular fractionation	Description and putative function	Localization	References
Class1 Reversibly Glycosylated Polypeptide (^{C1}RGP)	Ortholog in *Arabidopsis* At5g15650 (AtRGP2)	Maize mesocotyl *N. tabacum* BY-2 suspension cells	May shuttle UDP-sugar to or from glycosyltransferase	PD, Golgi	Epel et al., 1996 Sagi et al., 2005
β1.3 Glucanase	At5g42100	*A. thaliana cry2* mutant	Degradation of callose	PD, PM	Levy et al., 2007
Plasmodesmata Located Protein (PDLP) family	At5g43980 At3g04370 At2g33330 A1g04520 At3g60720 At1g70690 At5g37660 At2g01660	*A. thaliana* suspension cells	Type I membrane receptor Receptor of viral movement protein	PD	Bayer et al., 2006 Thomas et al., 2008
Plasmodesmal Callose Binding (PDCB) family	At5g61130 At5g08000 At1g18650 At1g69295 At3g58100	*A. thaliana* suspension cells	Callose binding protein through X8 domain	PD	Bayer et al., 2006 Simpson et al., 2009
Leucine Rich Repeat Receptor-Like Kinase (LRR RLK)	At1g56145	*A. thaliana* suspension cells	Signaling	PD, PM	Fernandez-Calvino et al., 2011
Catharanthus roseus Receptor-Like Kinase1-like (crRLK1L)	At5g24010	*A. thaliana* suspension cells	Signaling	PD, PM	Fernandez-Calvino et al., 2011
S-domain Receptor-Like Kinase	At4g21380	*A. thaliana* suspension cells	Signaling	PD, PM	Fernandez-Calvino et al., 2011
Tetraspanin3 (TET3)	At3g45600	*A. thaliana* suspension cells	Formation of specialized membrane microdomains	PD, PM	Fernandez-Calvino et al., 2011
Hypothetical protein	At3g15480	*A. thaliana* suspension cells	–	PD, PM	Fernandez-Calvino et al., 2011
Leucine Rich Repeat Receptor-Like Kinase (LRR RLK)	OsO6g47750 OsO2g05960 OsO9g02250	*Rice callus* suspension cells	Signaling	PD	Jo et al., 2011
Lectin Receptor-Like Kinase	Os04g01874	*Rice callus* suspension cells	Signaling	PD	Jo et al., 2011
Wall-Associated Kinase	OsO3g12470 OsO4g51050	*Rice callus* suspension cells	Signaling	PD	Jo et al., 2011
Plasmodesmal-Associated Protein Kinase1 (PAPK1)	Ortholog in *Arabidopsis* At4g28540 (Casein Kinase Like6)	*N. tabacum* BY-2 suspension cells	Signaling Phosphorylation of viral movement protein	PD	Lee et al., 2005
Nt-Plasmodesmal Germin-Like Protein1	Orthologs in *Arabidopsis* At1g09560 (PDGLP1) At1g02335 (PDGLP2)	*N. tabacum* BY-2 suspension cells	Regulation of primary roots growth	PD	Ham et al., 2012

discovery was that PDLP TMD was sufficient for PD targeting indicating that the sorting signals were recognized within the lipid bilayer (Thomas et al., 2008). This, together with the recent finding that lipid rafts, liquid-ordered sterols, and sphingolipids enriched PM microdomains, may associate with PD, raises questions about the role of lipids in defining PD specialized membranes (Raffaele et al., 2009; Mongrand et al., 2010; Tilsner et al., 2011). It is conceivable that the PM region lining PD may itself be sub-divided into functional domains. Sterol-enriched microdomains could well accumulate at the neck region of PD where GPI-anchored proteins such as PDCB or the β1–3 glucanases accumulate to control callose homeostasis and influence PD permeability (Levy et al., 2007; Simpson et al., 2009; Rinne et al., 2011). Hence, GPI anchors preferentially associate with liquid-ordered membrane domains (Sangiorgio et al., 2004; Borner et al., 2005; Kierszniowska et al., 2008). Through its X8 callose-binding domain, PDCB provides a physical link between PD and the wall and may even participate in stabilizing raft domains at PD (Simpson et al., 2009). The presence of functional subdomains at PD is also supported by the presence of TET3 a member of the tetraspanin family (Fernandez-Calvino et al., 2011). Tetraspanins are hydrophobic proteins that have the ability to associate with one another and to recruit specific proteins to build up tetraspanin-enriched microdomains that in mammalian regulate processes such as cell adhesion, signaling, and intracellular trafficking (Stipp et al., 2003; Yunta and Lazo, 2003; Rubinstein, 2011). Like rafts they enable membrane compartmentalization, a process that is required for PD to ensure their unique function.

We must also consider that PD are physically and functionally connected with the endomembrane system. In addition to the continuity of the ER with the desmotubule, the vast majority of PD components identified to date use the secretory pathway for delivery to the channels. For instance, Golgi disrupting treatments prevent both PDLP1 and RGP2 from reaching PD (Sagi et al., 2005; Thomas et al., 2008). Similarly, many plant viruses, which replicate in association with the endomembrane system, traffic to PD along the ER (Niehl and Heinlein, 2011). A number of PD located proteins also associate with the PM (LRR kinases; Jo et al., 2011), the Golgi (RGP2; Sagi et al., 2005), or the ER (calreticulin, Baluška et al., 1999; Chen et al., 2005) highlighting the potential for functional and dynamic relationships with other membrane compartments.

CONCLUSION AND PERSPECTIVES

The proteomic-based identification of PD components, combined with imaging techniques, pharmacological, and genetic approaches have brought substantial insight into the complexity of PD structure and dynamics. However, our understanding of PD function is still far from comprehensive and much remains to be determined before we fully comprehend the regulatory mechanisms governing symplastic transport. Many of the identified PD proteins still await functional characterization and advances in this area will provide exciting insights. Moreover, current findings concentrate on proteins with a membrane-localized signature, excluding for instance PD-associated soluble proteins or proteins transiently interacting with the channels which are both likely to be lost during PD purification due to extensive washes with salt containing buffer. Finally, many biological processes governed by symplastic transport probably come with a significant remodeling of PD constituents dictating that there are many more analyses to be done before functional PD components are fully described.

ACKNOWLEDGMENTS

We thank Dr. Christine Faulkner for comments on the manuscript before the submission.

REFERENCES

Amari, K., Boutant, E., Hofmann, C., Schmitt-Keichinger, C., Fernandez-Calvino, L., Didier, P., et al. (2010). A family of plasmodesmal proteins with receptor-like properties for plant viral movement proteins. *PLoS Pathog.* 6:e1001119. doi: 10.1371/journal.ppat.1001119

Baluška, F., Samaj, J., Napier, R., and Volkmann, D. (1999). Maize calreticulin localizes preferentially to plasmodesmata in root apex. *Plant J.* 19, 481–488.

Bayer, E., Thomas, C. L., and Maule, A. J. (2004). Plasmodesmata in *Arabidopsis thaliana* suspension cells. *Protoplasma* 223, 93–102.

Bayer, E. M., Bottrill, A. R., Walshaw, J., Vigouroux, M., Naldrett, M. J., Thomas, C. L., et al. (2006). Arabidopsis cell wall proteome defined using multidimensional protein identification technology. *Proteomics* 6, 301–311.

Benitez-Alfonso, Y., Cilia, M., San Roman, A., Thomas, C., Maule, A., Hearn, S., et al. (2009). Control of Arabidopsis meristem development by thioredoxin-dependent regulation of intercellular transport. *Proc. Natl. Acad. Sci. U.S.A.* 106, 3615–3620.

Benitez-Alfonso, Y., Faulkner, C., Ritzenthaler, C., and Maule, A. J. (2010). Plasmodesmata: gateways to local and systemic virus infection. *Mol. Plant Microbe Interact.* 23, 1403–1412.

Blackman, L. M., and Overall, R. L. (1998). Immunolocalisation of the cytoskeleton to plasmodesmata of *Chara corallina*. *Plant J.* 14, 733–741.

Borner, G. H., Sherrier, D. J., Weimar, T., Michaelson, L. V., Hawkins, N. D., Macaskill, A., et al. (2005). Analysis of detergent-resistant membranes in Arabidopsis. Evidence for plasma membrane lipid rafts. *Plant Physiol.* 137, 104–116.

Botha, C. E. J., Hartley, B. J., and Cross, R. H. M. (1993). The ultrastructure and computer-enhanced digital image analysis of plasmodesmata at the Kranz mesophyll-bundle sheath interface of *Themeda triandra* var. imbersis (Rezt) A. camus in conventianally-fixed leaf blades. *Ann. Bot.* 72, 255–261.

Brecknock, S., Dibbayawan, T. P., Vesk, M., Vesk, P. A., Faulkner, C., Barton, D. A., et al. (2011). High resolution scanning electron microscopy of plasmodesmata. *Planta* 234, 749–758.

Chen, M. H., Tian, G. W., Gafni, Y., and Citovsky, V. (2005). Effects of calreticulin on viral cell-to-cell movement. *Plant Physiol.* 138, 1866–1876.

Citovsky, V., McLean, B. G., Zupan, J. R., and Zambryski, P. (1993). Phosphorylation of tobacco mosaic virus cell-to-cell movement protein by a developmentally regulated plant cell wall-associated protein kinase. *Genes Dev.* 7, 904–910.

Cutler, S. R., Ehrhardt, D. W., Griffitts, J. S., and Somerville, C. R. (2000). Random GFP::cDNA fusions enable visualization of subcellular structures in cells of Arabidopsis at a high frequency. *Proc. Natl. Acad. Sci. U.S.A.* 97, 3718–3723.

Deeks, M. J., Calcutt, J. R., Ingle, E. K., Hawkins, T. J., Chapman, S., Richardson, A. C., et al. (2012). A superfamily of actin-binding proteins at the actin-membrane nexus of higher plants. *Curr. Biol.* 22, 1595–1600.

Ding, B., Kwon, M.-O., and Warnberg, L. (1996). Evidence that actin filaments are involved in controlling the permeability of plasmodesmata in tobacco mesophyll. *Plant J.* 10, 157–164.

Ding, B., Turgeon, R., and Parthasarathy, M. V. (1992). Substructure of freeze substituted plasmodesmata. *Protoplasma* 169, 28–41.

Epel, B., Van Lent, J. W. M., Cohen, L., Kotlizky, G., Katz, A., and

Yahalom, A. (1996). A 41 kDa protein isolated from maize mesocotyl cell walls immunolocalizes to plasmodesmata. *Protoplasma* 191, 70–78.

Epel, B. L., Kuchuck, B., Kotlizky, G., Shurtz, S., Erlanger, M., and Yahalom, A. (1995). Isolation and characterization of plasmodesmata. *Methods Cell Biol.* 50, 237–253.

Escobar, N. M., Haupt, S., Thow, G., Boevink, P., Chapman, S., and Oparka, K. (2003). High-throughput viral expression of cDNA-green fluorescent protein fusions reveals novel subcellular addresses and identifies unique proteins that interact with plasmodesmata. *Plant Cell* 15, 1507–1523.

Faulkner, C., and Maule, A. (2011). Opportunities and successes in the search for plasmodesmal proteins. *Protoplasma* 248, 27–38.

Faulkner, C. R., Blackman, L. M., Cordwell, S. J., and Overall, R. L. (2005). Proteomic identification of putative plasmodesmatal proteins from *Chara corallina*. *Proteomics* 5, 2866–2875.

Fernandez-Calvino, L., Faulkner, C., Walshaw, J., Saalbach, G., Bayer, E., Benitez-Alfonso, Y., et al. (2011). Arabidopsis plasmodesmal proteome. *PLoS ONE* 6:e18880. doi: 10.1371/journal.pone.0018880

Guseman, J. M., Lee, J. S., Bogenschutz, N. L., Peterson, K. M., Virata, R. E., Xie, B., et al. (2010). Dysregulation of cell-to-cell connectivity and stomatal patterning by loss-of-function mutation in Arabidopsis chorus (glucan synthase-like 8). *Development* 137, 1731–1741.

Ham, B. K., Li, G., Kang, B. H., Zeng, F., and Lucas, W. J. (2012). Overexpression of Arabidopsis Plasmodesmata germin-like proteins disrupts root growth and development. *Plant Cell* 24, 3630–3648.

Helper, P. K. (1982). Endoplasmic reticulum in the formation of the cell plate and plasmodesmata. *Protoplasma* 111, 121–133.

Jo, Y., Cho, W. K., Rim, Y., Moon, J., Chen, X. Y., Chu, H., et al. (2011). Plasmodesmal receptor-like kinases identified through analysis of rice cell wall extracted proteins. *Protoplasma* 248, 191–203.

Kankanala, P., Czymmek, K., and Valent, B. (2007). Roles for rice membrane dynamics and plasmodesmata during biotrophic invasion by the blast fungus. *Plant Cell* 19, 706–724.

Kierszniowska, S., Seiwert, B., and Schulze, W. X. (2008). Definition of Arabidopsis sterol-rich membrane microdomains by differential treatment with methyl-beta-cyclodextrin and quantitative proteomics. *Mol. Cell. Proteomics* 8, 612–623.

Kobayashi, K., Otegui, M. S., Krishnakumar, S., Mindrinos, M., and Zambryski, P. (2007). INCREASED SIZE EXCLUSION LIMIT 2 encodes a putative DEVH box RNA helicase involved in plasmodesmata function during Arabidopsis embryogenesis. *Plant Cell* 19, 1885–1897.

Kotlizky, G., Shurtz, S., Yahalom, A., Malik, Z., Traub, O., and Epel, B. L. (1992). An improved procedure for the isolation of plasmodesmata embedded in clean maize cell walls. *Plant J.* 2, 623–630.

Kragler, F., Monzer, J., Xoconostle-Cazares, B., and Lucas, W. J. (2000). Peptide antagonists of the plasmodesmal macromolecular trafficking pathway. *EMBO J.* 19, 2856–2868.

Lee, J. Y., and Lu, H. (2011). Plasmodesmata: the battleground against intruders. *Trends Plant Sci.* 16, 201–210.

Lee, J. Y., Taoka, K., Yoo, B. C., Ben-Nissan, G., Kim, D. J., and Lucas, W. J. (2005). Plasmodesmal-associated protein kinase in tobacco and Arabidopsis recognizes a subset of non-cell-autonomous proteins. *Plant Cell* 17, 2817–2831.

Lee, J. Y., Wang, X., Cui, W., Sager, R., Modla, S., Czymmek, K., et al. (2011). A plasmodesmata-localized protein mediates crosstalk between cell-to-cell communication and innate immunity in Arabidopsis. *Plant Cell* 23, 3353–3373.

Lee, J. Y., Yoo, B. C., Rojas, M. R., Gomez-Ospina, N., Staehelin, L. A., and Lucas, W. J. (2003). Selective trafficking of non-cell-autonomous proteins mediated by NtNCAPP1. *Science* 299, 392–396.

Levy, A., Erlanger, M., Rosenthal, M., and Epel, B. L. (2007). A plasmodesmata-associated beta-1, 3-glucanase in Arabidopsis. *Plant J.* 49, 669–682.

Maule, A. J. (2008). Plasmodesmata: structure, function and biogenesis. *Curr. Opin. Plant Biol.* 11, 680–686.

Maule, A. J., Benitez-Alfonso, Y., and Faulkner, C. (2011). Plasmodesmata – membrane tunnels with attitude. *Curr. Opin. Plant Biol.* 14, 683–690.

Mongrand, S., Stanislas, T., Bayer, E. M., Lherminier, J., and Simon-Plas, F. (2010). Membrane rafts in plant cells. *Trends Plant Sci.* 15, 656–663.

Monzer, J., and Kloth, S. (1991). The preparation of plasmodesmata from plant tissue homogenates: access to the biochemical characterisation of plasmodesmata-related polypeptides. *Bot. Acta* 104, 82–84.

Niehl, A., and Heinlein, M. (2011). Cellular pathways for viral transport through plasmodesmata. *Protoplasma* 248, 75–99.

Overall, R. L., Wolfe, J., and Gunning, B. E. S. (1982). Intercellular communication in *Azolla*. I. Ultrastructure of plasmodesmata. *Protoplasma* 111, 134–150.

Paape, M., Solovyev, A. G., Erokhina, T. N., Minina, E. A., Schepetilnikov, M. V., Lesemann, D. E., et al. (2006). At-4/1, an interactor of the Tomato spotted wilt virus movement protein, belongs to a new family of plant proteins capable of directed intra- and intercellular trafficking. *Mol. Plant Microbe Interact.* 19, 874–883.

Radford, J. E., and White, R. G. (1998). Localization of a myosin-like protein to plasmodesmata. *Plant J.* 14, 743–750.

Raffaele, S., Bayer, E., Lafarge, D., Cluzet, S., German-Retana, S., Boubekeur, T., et al. (2009). Remorin, a solanacae protein resident in membrane rafts and plasmodesmata, impairs potato virus X movement. *Plant Cell* 21, 1541–1555.

Reichelt, S., Knight, A. E., Hodge, T. P., Baluska, F., Samaj, J., Volkmann, D., et al. (1999). Characterization of the unconventional myosin VIII in plant cells and its localization at the post-cytokinetic cell wall. *Plant J.* 19, 555–567.

Rinne, P. L., Welling, A., Vahala, J., Ripel, L., Ruonala, R., Kangasjarvi, J., et al. (2011). Chilling of dormant buds hyperinduces FLOWERING LOCUS T and recruits GA-inducible 1,3-beta-glucanases to reopen signal conduits and release dormancy in Populus. *Plant Cell* 23, 130–146.

Rubinstein, E. (2011). The complexity of tetraspanins. *Biochem. Soc. Trans.* 39, 501–505.

Sagi, G., Katz, A., Guenoune-Gelbart, D., and Epel, B. L. (2005). Class 1 reversibly glycosylated polypeptides are plasmodesmal-associated proteins delivered to plasmodesmata via the golgi apparatus. *Plant Cell* 17, 1788–1800.

Sangiorgio, V., Pitto, M., Palestini, P., and Masserini, M. (2004). GPI_Anchored proteins and lipid rafts. *Ital. J. Biochem.* 53, 98–111.

Simpson, C., Thomas, C., Findlay, K., Bayer, E., and Maule, A. J. (2009). An Arabidopsis GPI-anchor plasmodesmal neck protein with callose binding activity and potential to regulate cell-to-cell trafficking. *Plant Cell* 21, 581–594.

Stipp, C. S., Kolesnikova, T. V., and Hemler, M. E. (2003). Functional domains in tetraspanin proteins. *Trends Biochem. Sci.* 28, 106–112.

Stonebloom, S., Burch-Smith, T., Kim, I., Meinke, D., Mindrinos, M., and Zambryski, P. (2009). Loss of the plant DEAD-box protein ISE1 leads to defective mitochondria and increased cell-to-cell transport via plasmodesmata. *Proc. Natl. Acad. Sci. U.S.A.* 106, 17229–17234.

Su, S., Liu, Z., Chen, C., Zhang, Y., Wang, X., Zhu, L., et al. (2010). Cucumber mosaic virus movement protein severs actin filaments to increase the plasmodesmal size exclusion limit in tobacco. *Plant Cell* 22, 1373–1387.

Thomas, C. L., Bayer, E. M., Ritzenthaler, C., Fernandez-Calvino, L., and Maule, A. J. (2008). Specific targeting of a plasmodesmal protein affecting cell-to-cell communication. *PLoS Biol.* 6:e7. doi: 10.1371/journal.pbio.0060007

Tilney, L. G., Cooke, T. J., Connelly, P. S., and Tilney, M. S. (1991). The structure of plasmodesmata as revealed by plasmolysis, detergent extraction, and protease digestion. *J. Cell Biol.* 112, 739–747.

Tilsner, J., Amari, K., and Torrance, L. (2011). Plasmodesmata viewed as specialised membrane adhesion sites. *Protoplasma* 248, 39–60.

Turner, A., Wells, B., and Roberts, K. (1994). Plasmodesmata of maize root tips: structure and composition. *J. Cell Sci.* 107, 3351–3361.

Waigmann, E., Turner, A., Peart, J., Roberts, K., and Zambryski, P. (1997). Ultrastructural analysis of leaf trichome plasmodesmata reveals major differences from mesophyll plasmodesmata. *Planta* 203, 75–84.

Washburn, M. P., Wolters, D., and Yates, J. R. 3rd. (2001). Large-scale analysis of the yeast proteome by multidimensional protein

identification technology. *Nat. Biotechnol.* 19, 242–247.

White, R. G., Badelt, K., Overall, R. L., and Vesk, M. (1994). Actin associated with plasmodesmata. *Protoplasma* 180, 169–184.

White, R. G., and Barton, D. A. (2011). The cytoskeleton in plasmodesmata: a role in intercellular transport? *J. Exp. Bot.* 62, 5249–5266.

Xu, X. M., and Jackson, D. (2010). Lights at the end of the tunnel: new views of plasmodesmal structure and function. *Curr. Opin. Plant Biol.* 13, 684–692.

Yunta, M., and Lazo, P. A. (2003). Tetraspanin proteins as organisers of membrane microdomains and signalling complexes. *Cell. Signal.* 15, 559–564.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 31 October 2012; accepted: 21 December 2012; published online: 11 January 2013.

Citation: Salmon MS and Bayer EMF (2013) Dissecting plasmodesmata molecular composition by mass spectrometry-based proteomics. Front. Plant Sci. ***3****:307. doi: 10.3389/fpls.2012.00307*

This article was submitted to Frontiers in Plant Proteomics, a specialty of Frontiers in Plant Science.

Copyright © 2013 Salmon and Bayer. This is an open-access article distributed under the terms of the ***Creative Commons Attribution License****, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.*

frontiers in
PLANT SCIENCE

MINI REVIEW ARTICLE
published: 25 April 2013
doi: 10.3389/fpls.2013.00114

When proteomics reveals unsuspected roles: the plastoglobule example

Houda Nacir*[1,2] *and Claire Bréhélin*[1,2]

[1] *Laboratoire de Biogenèse Membranaire, CNRS, Villenave d'Ornon, France*
[2] *Laboratoire de Biogenèse Membranaire, Université de Bordeaux, Villenave d'Ornon, France*

Edited by:
Nicolas L. Taylor, The University of Western Australia, Australia

Reviewed by:
Marcel Kuntz, Centre National de la Recherche Scientifique, France
Jotham R. Austin II, The University of Chicago, USA

****Correspondence:***
Claire Bréhélin, Laboratoire de Biogenèse Membranaire, CNRS – Université de Bordeaux, UMR5200, Campus INRA de Bordeaux, 71 Avenue E. Bourlaux, BP 81, F-33883 Villenave d'Ornon Cedex, France.
e-mail: claire.brehelin@u-bordeaux2.fr

Plastoglobules are globular compartments found in plastids. Before initial proteomic studies were published, these particles were often viewed as passive lipid droplets whose unique role was to store lipids coming from the thylakoid turn-over, or to accumulate carotenoids in the chromoplasts. Yet, two proteomic studies, published concomitantly, suggested for the first time that plastoglobules are more than "junk cupboards" for lipids. Indeed, both studies demonstrated that plastoglobules do not only include structural proteins belonging to the plastoglobulin/fibrillin family, but also contain active enzymes. The specific plastoglobule localization of these enzymes has been confirmed by different approaches such as immunogold localization and GFP protein fusions, thus providing evidence that plastoglobules actively participate in diverse pathways of plastid metabolism. These proteomic studies have been the basis for numerous recent works investigating plastoglobule function. However, a lot still needs to be discovered about the molecular composition and the role of plastoglobules. In this chapter, we will describe how the proteomic approaches have launched new perspectives on plastoglobule functions.

Keywords: plastoglobule, proteomics, *Arabidopsis*, plastids, stress, subcellular fractionation, fibrillin

INTRODUCTION

In addition to the network of thylakoid membranes which are the site of photosynthesis, plastids contain in their soluble phase, the stroma, some enigmatic lipoprotein bodies named the plastoglobules (cf. **Figure 1**). Plastoglobules can be found in diverse types of plastids, from proplastids (for review, see Nagata et al., 2002) to gerontoplasts (Kovacs et al., 2008) or etioplasts (Seyyedi et al., 1999). Although the origin of plastoglobules remains unclear, they may be closely linked to thylakoid development and dismantlement. Indeed, it has been observed that plastoglobule abundance increases when photosynthetic activity of green tissues decreases and thylakoids break down, like for example in senescent chloroplasts (Lichtenthaler, 1968; Guiamét et al., 1999; Ghosh et al., 2001), or during fruit maturation and ripening, when chloroplasts turn into chromoplasts and thylakoids disintegrate (Deruere et al., 1994; Vishnevetsky et al., 1999; Bonora et al., 2000). Reciprocally, plastoglobules are thought to be lipid reservoirs in greening tissue (Kessler et al., 1999), allowing the rapid formation of thylakoids. For example they may be involved in the formation of thylakoid membranes in de-etiolating plastids: etioplasts with poorly developed thylakoids have more plastoglobules than chloroplasts, but the plastoglobule abundance decreases during thylakoid biogenesis (Sprey and Lichtenthaler, 1966; Lichtenthaler and Peveling, 1967; Lichtenthaler, 1968). A tomographic study showing that plastoglobules are physically linked to thylakoid membranes (Austin et al., 2006) reinforced the idea that plastoglobules and thylakoids indeed share a functional relationship. The physical connection between the two compartments would allow channeling of molecules in both directions.

Plastoglobules are composed of an outer polar lipid monolayer containing neutral lipids (mainly prenylquinones, triacylglycerol, and carotenoids), and harbor proteins (for review, see Bréhélin et al., 2007; Bréhélin and Kessler, 2008). The diameter of plastoglobules is around 50–100 nm but they can enlarge to several micrometers (Thomson and Platt, 1973) depending on various factors such as plant species, plastid types, developmental stages, and environmental conditions. Numerous studies have described an increase of plastoglobule size and/or number under various environmental conditions (for review, see Bréhélin et al., 2007; Bréhélin and Kessler, 2008), such as drought (Rey et al., 2000), salt stress (Locy et al., 1996; Ben Khaled et al., 2003), or in the presence of heavy metals (Baszynski et al., 1980). Based on these ultrastructural observations, the involvement of the plastoglobules in plant responses to stress has been suggested, but biochemical or physiological evidence is missing. The exact role of plastoglobules in plant adaptation to stresses remains poorly understood. Yet, advances are being made in understanding some of their functions, mostly thanks to proteomics.

DECIPHERING THE NATURE AND ROLES OF PLASTOGLOBULES: FROM ULTRASTRUCTURAL BASED SPECULATIONS TO PROTEOMIC INDICATIONS

The progress made in plant electron microscopy allowed the first descriptions of plastoglobules: Hodge et al. (1955) observed the presence of "dense spherical bodies" in stroma of maize mesophyll chloroplasts while Falk (1960) reported the existence in *Ficus elastica* chloroplasts of "osmiophilic spheres" and "magnoglobuli" ranging from 0.13 to 2.5 μm in diameter. Menke (1962) stated that the chemical composition of the "spherical inclusions known as

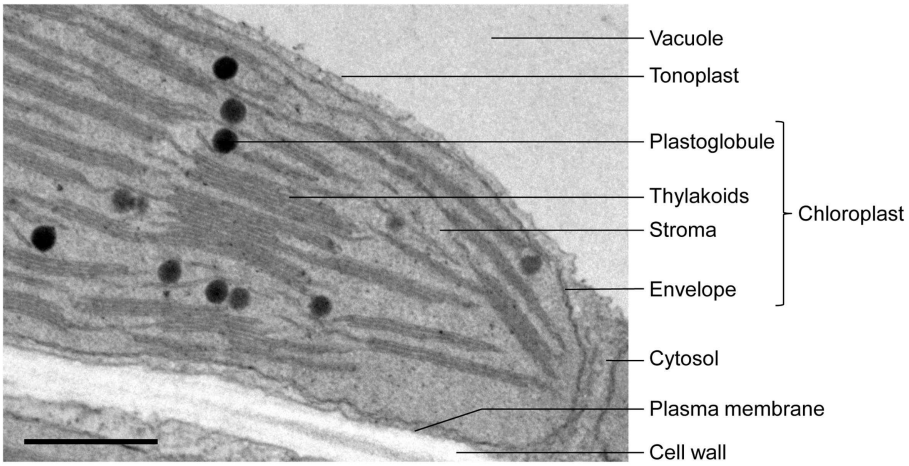


FIGURE 1 | Electron micrograph of an *Arabidopsis* leaf chloroplast showing plastoglobules in proximity with thylakoids. Bar: 500 nm.

osmiophilic granules or globules" was unknown, but that they were made of ether-soluble compounds, thus highlighting our ignorance of the plastoglobule composition, excepted for their lipidic nature.

The first protocols for the isolation of "osmiophilic globules" were then rapidly set up (Park and Pon, 1961; Bailey and Whyborn, 1963; Greenwood et al., 1963). They all followed a similar scheme. First, integral chloroplasts were purified from other cell components by centrifugation. Next, the chloroplasts were disrupted and plastoglobules separated from chloroplast membranes by differential centrifugation, thanks to their relatively low density. The subcellular fractionation of plastoglobules enabled scientists to investigate their chemical nature, especially with regard to their lipid and pigment contents (Bailey and Whyborn, 1963; Greenwood et al., 1963; Lichtenthaler, 1969). These studies reported the presence, in chloroplast plastoglobules, of several prenylquinones (tocopherol, phylloquinone, plastoquinone) while no significant amounts of carotenoids were detected.

While purification protocols were rapidly and easily set up, making purified plastoglobules available, the protein composition of this compartment has only started to be investigated 30 years later. Indeed, plastoglobules were long thought to be passive lipid droplets, accumulating pigments, and lipids originating from thylakoid disintegration (Smith et al., 2000). One of the first evidence for the association of proteins with plastoglobules came with the immunogold labeling of geranylgeranyl pyrophosphate synthase (GGPPS) in *Capsicum* fruits by Cheniclet et al. (1992). The authors described the presence of a pool of GGPPS around the plastoglobules. However, GGPPS is a functionally soluble enzyme and its specific physical association with plastoglobules was never confirmed. Pozueta-Romero et al. (1997) demonstrated that a major protein of bell pepper chromoplasts, the fibrillin, was a genuine component of plastoglobules and was located at their periphery. This protein was previously called fibrillin because of its high abundance in fibrils, a specialized structure of some chromoplasts wherein carotenoids accumulate (Deruere et al., 1994). It was proposed that fibrillin could built a compatible interface between the hydrophobic core of plastoglobule and the surrounding hydrophilic stroma, thereby allowing the maintenance of their structure and preventing them from coalescence (Deruere et al., 1994; Rey et al., 2000; Simkin et al., 2007). Afterward Kessler et al. (1999) showed that plastoglobules contained at least a dozen of different proteins which they named plastoglobulins. They characterized one of these plastoglobulins and showed that it belonged to the fibrillin family. Thus at the end of the twentieth century, plastoglobules were still generally viewed as passive lipid bags delimited by a coat of proteins whose nature and function were unknown. In this respect, proteomics allowed important improvement in the understanding of plastoglobule function, providing the first evidence for the presence of active enzymes in this compartment.

In 2006, two independent laboratories established for the first time the proteome of plastoglobules. While hundreds of proteins are usually listed in subcellular proteomic studies (Wienkoop et al., 2010), only as few as 30 proteins (cf. **Table 1**) were identified in *Arabidopsis thaliana* chloroplast plastoglobules (Vidi et al., 2006; Ytterberg et al., 2006), implying that plastoglobules are highly specialized sites dedicated to a restricted set of tasks in plastids. As expected, a major part of the proteome was constituted by proteins belonging to the plastoglobulin/PAP/fibrillin family, and another part was composed of proteins with unknown function. More astonishing, was the identification of about 10 known or putative metabolic enzymes, suggesting an active role for plastoglobules in some plastid metabolic pathways.

By combining shotgun proteomics with spectral-counting techniques, a quantitative proteomic approach has recently been applied to the plastoglobule proteome in order to detect low abundant proteins and quantify the relative abundance of each protein within plastoglobules (Lundquist et al., 2012b). By using defined selection filters (presence in biological and technical replicates, enrichment in the plastoglobule fraction, previously characterized subcellular localization), the core plastoglobule proteome was then restricted to 30 proteins. A striking observation was that despite the increased sensitivity of current mass spectrometers,

the plastoglobule proteome size did not enlarge. Only seven new low abundant proteins were added to the plastoglobule proteome while others were considered as likely contaminants and therefore were removed from the previous proteome (cf. **Table 1**). This new proteome was established with *Arabidopsis* plants submitted to high light stress, which could explain at least part of the observed variations.

THE PROTEIN COMPOSITION OF *ARABIDOPSIS* CHLOROPLAST PLASTOGLOBULE *PER SE*

A detailed comparison of the three different plastoglobule proteomes has recently been described (Lundquist et al., 2012b). Thus this review will only briefly summarize the function of the main protein components of the plastoglobules.

THE PLASTOGLOBULIN/PAP/FIBRILLIN FAMILY

Plastoglobulins (also called PAP for Plastid-lipid-Associated Protein, or fibrillin) represent more than 50% of the protein mass of the plastoglobule core proteome in *Arabidopsis* chloroplasts (Lundquist et al., 2012b). Proteomic studies have demonstrated that several members of this family are associated with plastoglobules, some being exclusively localized in plastoglobules (Kessler et al., 1999; Austin et al., 2006; Vidi et al., 2006) while others may partition between plastoglobules, thylakoids, and stroma (Rey et al., 2000; Lundquist et al., 2012b). However it still remains to be determined if the plastoglobulin composition varies within the plastoglobule population of a single plastid and if the presence of one specific plastoglobulin defines a kind of specialization of the plastoglobules (Vidi et al., 2006). Indeed, the abundance of each plastoglobulin within the plastoglobule proteome is not uniform, ranging from 1.8 to 16.1% (Lundquist et al., 2012b), and some plastoglobulins accumulate in plastoglobules under high light conditions, while others accumulate after dark treatment (Ytterberg et al., 2006). This suggests that plastoglobulins have diverse functions. In agreement, plastoglobulin mutant phenotypes suggest an implication of some of these proteins in plant growth regulation and development, as well as in stress tolerance and disease resistance (for review, see Singh and McNellis, 2011). Notably, plastoglobulins have recently been demonstrated to be implicated in jasmonate biosynthesis (Youssef et al., 2010) or plastoquinone accumulation (Singh et al., 2010, 2012), which illustrate their role in stress tolerance. However, the exact mechanism of action of plastoglobulins still needs to be clarified.

THE TOCOPHEROL CYCLASE VTE1

The tocopherol cyclase (VTE1) story illustrates how proteomics can sometimes prompt us to reconsider accepted models. VTE1 catalyzes the second to last step of α-tocopherol (vitamin E) synthesis consisting of the cyclization of 2,3-dimethyl-5-phytyl-1,4-hydroquinol (DMPQ) to γ-tocopherol (Porfirova et al., 2002). Before the plastoglobule proteomic studies were published, it was generally thought that the entire pathway for the vitamin E biosynthesis was taking place at the plastid envelope membrane (Soll et al., 1985). However, proteomics, coupled with immunolocalization and GFP-fusion studies, have demonstrated a specific localization of VTE1 in plastoglobules (Austin et al., 2006; Vidi et al., 2006; Ytterberg et al., 2006; Lundquist et al., 2012b). In addition to DMPQ, VTE1 also catalyzes the conversion of plastoquinone (PQH_2-9), another prenyl quinone, into plastochromanol (PC-8) (Szymanska and Kruk, 2010; Zbierzak et al., 2010). Both substrates (DMPQ and PQH_2-9) and products (γ-tocopherol and PC-8) are present at least partially in plastoglobules (Vidi et al., 2006; Zbierzak et al., 2010), providing an additional evidence for the implication of plastoglobules in the synthesis of prenyl quinones. Tocopherols, as well as plastoquinone and plastochromanol have antioxidant activity, exerting a photoprotective role to thylakoid lipids and photosystem II (Eugeni Piller et al., 2012). Thus, plastoglobules may represent an antioxidant reservoir available to protect the thylakoid membranes from oxidative stress.

THE NAD(P)H QUINONE DEHYDROGENASE C1 (NDC1)

When the first plastoglobule proteomes were published, the function of NDC1 was unknown, and its localization was believed to be mitochondrial (Michalecka et al., 2003). The identification of NDC1 in the plastoglobule proteome prompted Kessler and colleagues to investigate its localization and function (Eugeni Piller et al., 2011). Its localization in plastoglobules was confirmed by means of GFP-fusion constructs and its dual localization in plastoglobules and mitochondria demonstrated by western blot analysis. The authors also showed that in the knock-out *ndc1 Arabidopsis* mutant, the plastoquinone pool was more oxidized than in the wild type, demonstrating that NDC1 is involved in the regeneration of reduced plastoquinone. In addition, the *ndc1* mutant was almost totally deprived of phylloquinone (vitamin K_1), suggesting that the enzyme plays a part in the phylloquinone production.

THE CAROTENOID CLEAVAGE DIOXYGENASE 4

Carotenoid cleavage dioxygenase 4 (CCD4) (also named NCED4 for 9-cis epoxy-carotenoid dioxygenase 4) has been reported to occur in the plastoglobule proteomes (Vidi et al., 2006; Ytterberg et al., 2006; Lundquist et al., 2012b). The members of the CCD family cleave different carotenoids and xanthophylls to apocarotenoids such as abscisic acid. *In vitro*, AtCCD4 cleaves preferentially the apocarotenoid 8′-apo-β-caroten-8′ to yield β-ionone (Huang et al., 2009). However, the *in planta* substrate for this enzyme has not been discovered yet, and its current function in plastoglobule is unknown. The stable isotope experiments performed by Ytterberg et al. (2006) showed an accumulation of the enzyme after dark treatment compared to high light treatment, suggesting an implication of AtCCD4 in carotenoid breakdown.

ABC1 KINASES

Six members of the activity of BC1 (ABC1 complex) kinases have been identified in plastoglobules, representing the second most abundant protein family of the plastoglobule proteome (Lundquist et al., 2012b). The ABC1 kinases belong to the atypical protein kinase superfamily. In *Arabidopsis*, this superfamily is composed of 15 members, among which six are most likely mitochondrial and the remaining ones are supposed to be targeted to plastids (Lundquist et al., 2012a). An ABC1 kinase was first described as playing an essential role in electron transfer in the bc1 complex of *Saccharomyces cerevisiae* (Bousquet et al., 1991). Two plastidial ABC1 kinases, AtOSA1, and AtACDO1 which locates to plastoglobules, were proposed to act against photooxidative stress

Table 1 | The chloroplast plastoglobule proteome determined by different proteomic studies.

Accession number	Name/description	Percent mass PG core[a]	Prot. reference[b]	Confirmed plastoglobule location[c]
At4g04020	AtPGL35/FBN1a; Stress tolerance (Youssef et al., 2010)	16.1	Vidi et al. (2006), Ytterberg et al. (2006), Lundquist et al. (2012b)	GFP-fusion (Vidi et al., 2006), immunogold (Ytterberg et al., 2006)
At3g23400	AtPGL30.4/FBN4; Plastoglobule development, stress tolerance (Singh et al., 2010)	11.9	Vidi et al. (2006), Ytterberg et al. (2006), Lundquist et al. (2012b)	GFP-fusion (Vidi et al., 2006)
At4g22240	AtPGL33/FBN1b; Stress tolerance (Youssef et al., 2010)	9.6	Vidi et al. (2006), Ytterberg et al. (2006), Lundquist et al. (2012b)	
At2g35490	AtPGL40/FBN2; Stress tolerance (Youssef et al., 2010)	7.1	Vidi et al. (2006), Ytterberg et al. (2006), Lundquist et al. (2012b)	
At3g58010	AtPGL34/FBN7a	3.5	Vidi et al. (2006), Ytterberg et al. (2006), Lundquist et al. (2012b)	GFP-fusion (Vidi et al., 2006, 2007)
At2g42130	AtPGL30/FBN7b	2.1	Vidi et al. (2006), Ytterberg et al. (2006), Lundquist et al. (2012b)	
At2g46910	AtPGL31/FBN8	1.8	Ytterberg et al. (2006), Lundquist et al. (2012b)	
At5g05200	ABC1 kinase 9; Lundquist et al. (2012a)	5.2	Vidi et al. (2006), Ytterberg et al. (2006), Lundquist et al. (2012b)	
At4g31390	ABC1 kinase 1/AtACDO1; Photooxidative stress tolerance (Yang et al., 2012)	4.5	Ytterberg et al. (2006), Lundquist et al. (2012b)	
At1g79600	ABC1 kinase 3; Lundquist et al. (2012a)	4.3	Vidi et al. (2006), Ytterberg et al. (2006), Lundquist et al. (2012b)	
At3g24190	ABC1 kinase 6; Lundquist et al. (2012a)	2.6	Lundquist et al. (2012b)	
At1g71810	ABC1 kinase 5; Lundquist et al. (2012a)	1.7	Ytterberg et al. (2006), Lundquist et al. (2012b)	
At3g07700	ABC1 kinase 7; Lundquist et al. (2012a)	0.8	Lundquist et al. (2012b)	
At4g19170	Carotenoid cleavage dioxygenase (AtCDD4) Carotenoid catabolism (Huang et al., 2009)	3.3	Vidi et al. (2006), Ytterberg et al. (2006), Lundquist et al. (2012b)	
At4g32770	Tocopherol cyclase (VTE1); Vitamin E biosynthesis (Porfirova et al., 2002)	2.6	Vidi et al. (2006), Ytterberg et al. (2006), Lundquist et al. (2012b)	GFP-fusion (Vidi et al., 2006), immunogold (Ytterberg et al., 2006)
At5g08740	NAD(P)H dehydrogenase C1 (NDC1); Phylloquinone synthesis, plastoquinone reduction	2.5	Vidi et al. (2006), Ytterberg et al. (2006), Lundquist et al. (2012b)	GFP-fusion (Eugeni Piller et al., 2011)
At1g54570	PES1; Phytyl ester synthesis (Lippold et al., 2012)	2.6	Vidi et al. (2006), Ytterberg et al. (2006), Lundquist et al. (2012b)	
At3g26840	PES2; Phytyl ester synthesis (Lippold et al., 2012)	1.4	Vidi et al. (2006), Ytterberg et al. (2006), Lundquist et al. (2012b)	
At1g78140	Protein with Methyltransferase type 11 domain-1	1.5	Vidi et al. (2006), Ytterberg et al. (2006), Lundquist et al. (2012b)	
At2g41040	Protein with Methyltransferase type 11 domain-2	1.5	Vidi et al. (2006), Ytterberg et al. (2006), Lundquist et al. (2012b)	

(Continued)

Table 1 | Continued

Accession number	Name/description	Percent mass PG core[a]	Prot. reference[b]	Confirmed plastoglobule location[c]
At1g32220	Protein with NAD-dependent epimerase/dehydratase domain	2.1	Vidi et al. (2006), Ytterberg et al. (2006), Lundquist et al. (2012b)	
At1g06690	Protein with aldo-keto reductase domain	1.5	Vidi et al. (2006), Ytterberg et al. (2006), Lundquist et al. (2012b)	
At4g39730	Lipase/lipooxygenase, PLAT/LH2 family protein	1.6	Lundquist et al. (2012b)	
At1g73750	Protein with α/β hydrolase domain	0.4	Lundquist et al. (2012b)	
At5g41120	Esterase/lipase/thioesterase family protein	0.3	Lundquist et al. (2012b)	
At5g42650	Allene oxide synthase (AOS) Jasmonic acid biosynthesis (Laudert et al., 1996)		Vidi et al. (2006), Ytterberg et al. (2006)	GFP-fusion (Vidi et al., 2006)
At2g21330	Fructose-bisphosphate aldolase-1		Vidi et al. (2006), Ytterberg et al. (2006)	GFP-fusion, enzymatic activity (Vidi et al., 2006)
At4g38970	Fructose-bisphosphate aldolase-2		Vidi et al. (2006), Ytterberg et al. (2006)	GFP-fusion, enzymatic activity (Vidi et al., 2006)
At2g01140	Putative fructose-bisphosphate aldolase-3		Vidi et al. (2006), Ytterberg et al. (2006)	
At3g10130	Protein with SOUL heme binding domain	1.8	Vidi et al. (2006), Ytterberg et al. (2006), Lundquist et al. (2012b)	
At2g34460	NAD(P)-binding Rossmann-fold superfamily protein	1.5	Vidi et al. (2006), Ytterberg et al. (2006), Lundquist et al. (2012b)	
At4g13200	Unknown protein-1	1.9	Vidi et al. (2006), Ytterberg et al. (2006), Lundquist et al. (2012b)	
At3g43540	Unknown protein-2 (DUF 1350)	1.3	Lundquist et al. (2012b)	
At3g27110	Protein with peptidase M48 domain	0.3	Lundquist et al. (2012b)	

[a] Contribution of each protein to the protein mass of the total PG core proteome (in %) determined in Lundquist et al. (2012b). [b] Plastoglobule proteome from which the protein was identified. [c] Reference confirming the localization of the protein in plastoglobules.

(Jasinski et al., 2008; Yang et al., 2012). However, the function of the plant ABC1 kinases is unknown, and the significance of the localization of the six members of this family in plastoglobules, representing 18% of the protein plastoglobule mass, is still mysterious.

PES1 AND 2

Plastoglobules were demonstrated to be the site of accumulation of fatty acid phytyl esters (FAPEs) under stress conditions or during senescence (Ischebeck et al., 2006; Gaude et al., 2007). FAPEs consist of a phytol molecule, originating from the breakdown of chlorophyll, esterified to an acyl group removed from galactolipids. The accumulation of FAPEs in plastoglobules is believed to prevent the membranes from the detergent-like properties of free phytol and acyl groups (Bréhélin and Kessler, 2008). The enzyme(s) responsible for this synthesis were initially unknown. However, two putative acyltransferases with sequence similarities to esterases/lipases/thioesterases (At1g54570 and At3g26840) were identified in the plastoglobule core proteome (Vidi et al., 2006; Ytterberg et al., 2006; Lundquist et al., 2012b). Lippold et al. (2012) therefore hypothesized that these enzymes could be involved in the FAPE synthesis. Using reverse genetic and heterologous expression approaches, they demonstrated that indeed both proteins catalyzed the formation of FAPEs during stress conditions, and were therefore likely involved in the maintenance of the photosynthetic membrane integrity.

OTHER PROTEINS

Three isoforms of fructose-1,6-biphosphate aldolases (FBPA), and the allene oxide synthase (AOS) were identified in the two first published proteomes (Vidi et al., 2006; Ytterberg et al., 2006) but excluded from the plastoglobule proteome established by Lundquist et al. (2012b). AOS is implicated in jasmonate synthesis (Schaller and Stintzi, 2009), while FBPA participates to the Calvin cycle and glycolysis. The localization in plastoglobules of the four enzymes was confirmed by transient expression of GFP-tagged constructs (Vidi et al., 2006) and an FBPA activity was measured in

the plastoglobule fractions (Vidi et al., 2006). These enzymes were removed from the "core plastoglobule proteome" because they were not enriched in the plastoglobule fraction. They could however partition between plastoglobules and other compartments of the chloroplast. These four enzymes should perhaps not be excluded from the plastoglobule proteome but rather be considered as enzymes with roles in plastoglobules as well as in other plastid compartments.

Some additional low abundant proteins were described in plastoglobules. The majority was represented by proteins with unknown function, such as two proteins with methyltransferase domains, a SOUL heme binding protein, or newly identified proteins (Lundquist et al., 2012b) such as the M48 metalloprotease. There is no doubt that the understanding of their function will reveal another panel of the plastoglobule story.

OTHER PLASTOGLOBULE PROTEOMES

The specialized structures wherein carotenoids accumulate during chromoplastogenesis define the morphology of chromoplasts (reviewed in Egea et al., 2010). For instance, during globular chromoplast formation, carotenoids accumulate in plastoglobules, leading to an increase of plastoglobule size and/or number (Jeffery et al., 2012). Ytterberg et al. (2006) analyzed the proteome of red pepper chromoplast plastoglobules and showed that it contains (i) plastoglobulins, one of the most abundant proteins of pepper chromoplasts (Siddique et al., 2006), which are known to be involved in carotenoid sequestration (for review, see Bréhélin and Kessler, 2008; Egea et al., 2010), (ii) enzymes already characterized in plastoglobules such as VTE1 and FBPA, and (iii) enzymes involved in carotenoid synthesis including ζ-carotene desaturase (ZDS), lycopene β-cyclase (LYCB), and two β-carotene β-hydroxylases. In addition, phytoene synthases were proposed to locate to plastoglobules based on evidence from GFP-fusion localization experiments (Shumskaya et al., 2012). The presence of such enzymes suggests that plastoglobules are not only involved in the sequestration of carotenoids but also in carotenoid biosynthesis. However uncertainty still persists about the localization of ZDS and LYCB since they were characterized in envelope fraction of *Arabidopsis* chloroplasts by spectral-counting proteomics (Joyard et al., 2009). The possibility remains that the envelope fraction was contaminated with plastoglobules in these latter experiments. Yet, in the same study, plastoglobulins were not found in the envelope fraction but in the thylakoids, suggesting that plastoglobules are rather associated with the thylakoid membranes in this preparation. Another explanation could be that the localization of these enzymes differs depending on plastid type (chloroplast or chromoplast), organ (leaf or fruit), or species. This underlines the need for other chromoplast plastoglobule proteomic studies.

Finally, other proteomic data about plastoglobules could be taken from the proteome of *Chlamydomonas reinhardtii* eyespot. The eyespot apparatus is believed to play the role of a directional light sensor (reviewed in Kreimer, 2009). It is constituted by two layers of carotenoid-rich globules associated with thylakoids. Transmission electron microscopy observations suggested similarities between the globules of the eyespot apparatus and plastoglobules. The resemblance of both compartments is confirmed by proteomics. Indeed, the proteomes of *Chlamydomonas reinhardtii* eyespot and plastoglobules contain common homologous proteins, such as proteins with plastoglobulin domain, ABC1 kinases, or FBPA (Schmidt et al., 2006; Kreimer, 2009). Proteins with plastoglobulin domains may prevent the coalescence of the carotenoid-rich globules and maintain interactions with membranes. The actual function of the other proteins found both in eyespot and plastoglobules is not yet understood.

While the plastoglobule proteomes from different plastid types share proteins in common, they also contain specific proteins depending on the plastid type. This variation in the plastoglobule protein composition may indicate a possible specialization of the plastoglobule function depending on the tissues considered.

CONCLUSION

Proteomic approaches have brought substantial insight into our understanding of plastoglobule functions. Our conception of the plastoglobules has dramatically changed from insipid passive lipid droplets inside plastids to particles with active role at the crossroad of diverse metabolic pathways, for example in vitamin biosynthesis. Notably, the common denominator of the vast majority of plastoglobule proteins is a possible involvement in the response to stress. Recently, our understanding of the plastoglobule function has reached an upper level with the combination of proteomics and co-expression analysis (Lundquist et al., 2012b). Assuming that a set of coexpressed genes is involved in the same or related metabolic pathway, a co-expression network of the core plastoglobule genes has been built, with the goal to provide a framework to better decipher plastoglobule roles. Four major co-expression modules were defined, with specific functions. Thus a model was proposed, where plastoglobules are involved in senescence, plastid biogenesis and proteolysis, redox regulation, photoacclimation, and isoprenoid biosynthesis (Lundquist et al., 2012b). Nevertheless, major efforts still need to be accomplished to comprehensively understand the role of plastoglobules in the plant biology, and especially their involvement in the plant responses to stress.

ACKNOWLEDGMENTS

We thank the Plant Imaging Platform at the Bordeaux Imaging Center (http://www.bic.u-bordeaux2.fr/index.php/fr/imagerie-vegetale) for contribution to imagery equipment and J.-J. Bessoule, C. Garcion, and E. Bayer for critical reading of the manuscript. Houda Nacir is recipient of a fellowship from the Conseil Régional d'Aquitaine. Claire Bréhélin acknowledges the Conseil Régional d'Aquitaine for financial support.

REFERENCES

Austin, J. R., Frost, E., Vidi, P. A., Kessler, F., and Staehelin, L. A. (2006). Plastoglobules are lipoprotein subcompartments of the chloroplast that are permanently coupled to thylakoid membranes and contain biosynthetic enzymes. *Plant Cell* 18, 1693–1703.

Bailey, J. L., and Whyborn, A. G. (1963). The osmiophilic globules of chloroplasts. II. Globules of the spinach-beet chloroplast. *Biochim. Biophys. Acta* 78, 163–174.

Baszynski, T., Wajda, L., Krol, M., Wolinska, D., Krupa, Z., and Tukendorf, A. (1980). Photosynthetic activities of cadmium-treated tomato plants. *Physiol. Plant* 48, 365–370.

Ben Khaled, L., Gomez, A. M., Honrubia, M., and Oihabi, A. (2003). Effet du stress salin en milieu hydroponique sur le trèfle inoculé par le Rhizobium. *Agronomie* 23, 553–560.

Bonora, A., Pancaldi, S., Gualandri, R., and Fasulo, M. P. (2000). Carotenoid and ultrastructure variations in plastids of *Arum italicum* Miller fruit during maturation and ripening. *J. Exp. Bot.* 51, 873–884.

Bousquet, I., Dujardin, G., and Slonimski, P. P. (1991). ABC1, a novel yeast nuclear gene has a dual function in mitochondria: it suppresses a cytochrome b mRNA translation defect and is essential for the electron transfer in the bc 1 complex. *EMBO J.* 10, 2023–2031.

Bréhélin, C., and Kessler, F. (2008). The plastoglobule: a bag full of lipid biochemistry tricks. *Photochem. Photobiol.* 84, 1388–1394.

Bréhélin, C., Kessler, F., and van Wijk, K. J. (2007). Plastoglobules: versatile lipoprotein particles in plastids. *Trends Plant Sci.* 12, 260–266.

Cheniclet, C., Rafia, F., Saintguily, A., Verna, A., and Carde, J. P. (1992). Localization of the enzyme geranylgeranyl pyrophosphate synthase in capsicum fruits by immunogold cytochemistry after conventional chemical fixation or quick-freezing followed by freeze-substitution – labeling evolution during fruit ripening. *Biol. Cell* 75, 145–154.

Deruere, J., Romer, S., d'Harlingue, A., Backhaus, R. A., Kuntz, M., and Camara, B. (1994). Fibril assembly and carotenoid overaccumulation in chromoplasts: a model for supramolecular lipoprotein structures. *Plant Cell* 6, 119–133.

Egea, I., Barsan, C., Bian, W., Purgatto, E., Latche, A., Chervin, C., et al. (2010). Chromoplast differentiation: current status and perspectives. *Plant Cell Physiol.* 51, 1601–1611.

Eugeni Piller, L., Abraham, M., Dormann, P., Kessler, F., and Besagni, C. (2012). Plastid lipid droplets at the crossroads of prenylquinone metabolism. *J. Exp. Bot.* 63, 1609–1618.

Eugeni Piller, L., Besagni, C., Ksas, B., Rumeau, D., Brehelin, C., Glauser, G., et al. (2011). Chloroplast lipid droplet type II NAD(P)H quinone oxidoreductase is essential for prenylquinone metabolism and vitamin K1 accumulation. *Proc. Natl. Acad. Sci. U.S.A.* 108, 14354–14359.

Falk, H. (1960). Magnoglobuli in chloroplasten von *Ficus elastica* roxb. *Planta* 55, 525–532.

Gaude, N., Bréhélin, C., Tischendorf, G., Kessler, F., and Dormann, P. (2007). Nitrogen deficiency in *Arabidopsis* affects galactolipid composition and gene expression and results in accumulation of fatty acid phytyl esters. *Plant J.* 49, 729–739.

Ghosh, S., Mahoney, S. R., Penterman, J. N., Peirson, D., and Dumbroff, E. B. (2001). Ultrastructural and biochemical changes in chloroplasts during *Brassica napus* senescence. *Plant Physiol. Biochem.* 39, 777–784.

Greenwood, A. D., Leech, R. M., and Williams, J. P. (1963). The osmiophilic globules of chloroplasts. I. Osmiophilic globules as a normal component of chloroplasts and their isolation and composition in *Vicia faba* L. *Biochim. Biophys. Acta* 78, 148–162.

Guiamét, J. J., Pichersky, E., and Noodén, L. D. (1999). Mass exodus from senescing soybean chloroplasts. *Plant Cell Physiol.* 40, 986–992.

Hodge, A. J., McLean, J. D., and Mercer, F. V. (1955). Ultrastructure of the lamellae and grana in the chloroplasts of *Zea mays* L. *J. Biophys. Biochem. Cytol.* 1, 605.

Huang, F. C., Molnar, P., and Schwab, W. (2009). Cloning and functional characterization of carotenoid cleavage dioxygenase 4 genes. *J. Exp. Bot.* 60, 3011–3022.

Ischebeck, T., Zbierzak, A. M., Kanwischer, M., and Dormann, P. (2006). A salvage pathway for phytol metabolism in *Arabidopsis*. *J. Biol. Chem.* 281, 2470–2477.

Jasinski, M., Sudre, D., Schansker, G., Schellenberg, M., Constant, S., Martinoia, E., et al. (2008). AtOSA1, a member of the Abc1-like family, as a new factor in cadmium and oxidative stress response. *Plant Physiol.* 147, 719–731.

Jeffery, J., Holzenburg, A., and King, S. (2012). Physical barriers to carotenoid bioaccessibility. *J. Sci. Food Agric.* 92, 2594–2602.

Joyard, J., Ferro, M., Masselon, C., Seigneurin-Berny, D., Salvi, D., Garin, J., et al. (2009). Chloroplast proteomics and the compartmentation of plastidial isoprenoid biosynthetic pathways. *Mol. Plant.* 2, 1154–1180.

Kessler, F., Schnell, D., and Blobel, G. (1999). Identification of proteins associated with plastoglobules isolated from pea (*Pisum sativum* L.) chloroplasts. *Planta* 208, 107–113.

Kovacs, E., Kristof, Z., Perlaki, R., and Szollosi, D. (2008). Cell wall metabolism during ripening and storage of nonclimacteric sour cherry (*Prunus cerasus* L., cv. Kantorjanosi). *Acta Aliment.* 37, 415–426.

Kreimer, G. (2009). The green algal eyespot apparatus: a primordial visual system and more? *Curr. Genet.* 55, 19–43.

Laudert, D., Pfannschmidt, U., Lottspeich, F., Hollander-Czytko, H., and Weiler, E. W. (1996). Cloning, molecular and functional characterization of *Arabidopsis thaliana* allene oxide synthase (CYP 74), the first enzyme of the octadecanoid pathway to jasmonates. *Plant Mol. Biol.* 31, 323–335.

Lichtenthaler, H. K. (1968). Plastoglobuli and the fine structure of plastids. *Endeavor* 27, 144–149.

Lichtenthaler, H. K. (1969). Die Plastoglobuli. *Protoplasma* 68, 65–77.

Lichtenthaler, H. K., and Peveling, E. (1967). Plastoglobuli in verschiedenen differenzierungsstadien der plastiden bei *Allium cepa* L. *Planta* 72, 1–13.

Lippold, F., vom Dorp, K., Abraham, M., Holzl, G., Wewer, V., Yilmaz, J. L., et al. (2012). Fatty acid phytyl ester synthesis in chloroplasts of *Arabidopsis*. *Plant Cell* 24, 2001–2014.

Locy, R. D., Chang, C. C., Nielsen, B. L., and Singh, N. K. (1996). Photosynthesis in salt-adapted heterotrophic tobacco cells and regenerated plants. *Plant Physiol.* 110, 321–328.

Lundquist, P. K., Davis, J. I., and van Wijk, K. J. (2012a). ABC1K atypical kinases in plants: filling the organellar kinase void. *Trends Plant Sci.* 17, 546–555.

Lundquist, P. K., Poliakov, A., Bhuiyan, N. H., Zybailov, B., Sun, Q., and van Wijk, K. J. (2012b). The functional network of the *Arabidopsis* plastoglobule proteome based on quantitative proteomics and genome-wide coexpression analysis. *Plant Physiol.* 158, 1172–1192.

Menke, W. (1962). Structure and chemistry of plastids. *Annu. Rev. Plant Physiol.* 13, 27.

Michalecka, A. M., Svensson, A. S., Johansson, F. I., Agius, S. C., Johanson, U., Brennicke, A., et al. (2003). Arabidopsis genes encoding mitochondrial type II NAD(P)H dehydrogenases have different evolutionary origin and show distinct responses to light. *Plant Physiol.* 133, 642–652.

Nagata, N., Suzuki, M., Yoshida, S., and Muranaka, T. (2002). Mevalonic acid partially restores chloroplast and etioplast development in *Arabidopsis* lacking the non-mevalonate pathway. *Planta* 216, 345–350.

Park, R. B., and Pon, N. G. (1961). Correlation of structure with function in *Spinacia oleracea* chloroplasts. *J. Mol. Biol.* 3, 1–10.

Porfirova, S., Bergmuller, E., Tropf, S., Lemke, R., and Dormann, P. (2002). Isolation of an *Arabidopsis* mutant lacking vitamin E and identification of a cyclase essential for all tocopherol biosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* 99, 12495–12500.

Pozueta-Romero, J., Rafia, F., Houlne, G., Cheniclet, C., Carde, J. P., Schantz, M. L., et al. (1997). A ubiquitous plant housekeeping gene, PAP, encodes a major protein component of bell pepper chromoplasts. *Plant Physiol.* 115, 1185–1194.

Rey, P., Gillet, B., Romer, S., Eymery, F., Massimino, J., Peltier, G., et al. (2000). Over-expression of a pepper plastid lipid-associated protein in tobacco leads to changes in plastid ultrastructure and plant development upon stress. *Plant J.* 21, 483–494.

Schaller, A., and Stintzi, A. (2009). Enzymes in jasmonate biosynthesis – structure, function, regulation. *Phytochemistry* 70, 1532–1538.

Schmidt, M., Gessner, G., Luff, M., Heiland, I., Wagner, V., Kaminski, M., et al. (2006). Proteomic analysis of the eyespot of *Chlamydomonas reinhardtii* provides novel insights into its components and tactic movements. *Plant Cell* 18, 1908–1930.

Seyyedi, M., Timko, M. P., and Sundqvist, C. (1999). Protochlorophyllide, NADPH-protochlorophyllide oxidoreductase, and chlorophyll formation in the lip1 mutant of pea. *Physiol. Plant* 106, 344–354.

Shumskaya, M., Bradbury, L. M., Monaco, R. R., and Wurtzel, E. T. (2012). Plastid localization of the key carotenoid enzyme phytoene synthase is altered by isozyme, allelic variation, and activity. *Plant Cell* 24, 3725–3741.

Siddique, M. A., Grossmann, J., Gruissem, W., and Baginsky, S. (2006). Proteome analysis of bell pepper (*Capsicum annuum* L.) chromoplasts. *Plant Cell Physiol.* 47, 1663–1673.

Simkin, A. J., Gaffe, J., Alcaraz, J. P., Carde, J. P., Bramley, P. M., Fraser, P. D., et al. (2007). Fibrillin influence on plastid ultrastructure and pigment content in tomato fruit. *Phytochemistry* 68, 1545–1556.

Singh, D. K., Laremore, T. N., Smith, P. B., Maximova, S. N., and

McNellis, T. W. (2012). Knockdown of FIBRILLIN4 gene expression in apple decreases plastoglobule plastoquinone content. *PLoS ONE* 7:e47547. doi:10.1371/journal.pone.0047547

Singh, D. K., Maximova, S. N., Jensen, P. J., Lehman, B. L., Ngugi, H. K., and McNellis, T. W. (2010). FIBRILLIN4 is required for plastoglobule development and stress resistance in apple and *Arabidopsis*. *Plant Physiol.* 154, 1281–1293.

Singh, D. K., and McNellis, T. W. (2011). Fibrillin protein function: the tip of the iceberg? *Trends Plant Sci.* 16, 432–441.

Smith, M. D., Licatalosi, D. D., and Thompson, J. E. (2000). Co-association of cytochrome f catabolites and plastid-lipid-associated protein with chloroplast lipid particles. *Plant Physiol.* 124, 211–221.

Soll, J., Schultz, G., Joyard, J., Douce, R., and Block, M. A. (1985). Localization of synthesis of prenylquinones in isolated outer and inner envelope membranes from spinach chloroplasts. *Arch. Biochem. Biophys.* 238, 290–299.

Sprey, B., and Lichtenthaler, H. (1966). Zur frage der Beziehungen zwischen plastoglobuli und thylakoidgenese in gerstenkeimlingen. *Z. Naturforsch.* 21b, 697–699.

Szymanska, R., and Kruk, J. (2010). Plastoquinol is the main prenyllipid synthesized during acclimation to high light conditions in Arabidopsis and is converted to plastochromanol by tocopherol cyclase. *Plant Cell Physiol.* 51, 537–545.

Thomson, W. W., and Platt, K. A. (1973). Plastid ultrastructure in the barrel cactus, *Echinocactus acanthodes*. *New Phytol.* 72, 791–797.

Vidi, P. A., Kanwischer, M., Baginsky, S., Austin, J. R., Csucs, G., Dormann, P., et al. (2006). Tocopherol cyclase (VTE1) localization and vitamin E accumulation in chloroplast plastoglobule lipoprotein particles. *J. Biol. Chem.* 281, 11225–11234.

Vidi, P. A., Kessler, F., and Bréhélin, C. (2007). Plastoglobules: a new address for targeting recombinant proteins in the chloroplast. *BMC Biotechnol.* 7:4. doi:10.1186/1472-6750-7-4

Vishnevetsky, M., Ovadis, M., and Vainstein, A. (1999). Carotenoid sequestration in plants: the role of carotenoid-associated proteins. *Trends Plant Sci.* 4, 232–235.

Wienkoop, S., Baginsky, S., and Weckwerth, W. (2010). *Arabidopsis thaliana* as a model organism for plant proteome research. *J. Proteomics* 73, 2239–2248.

Yang, S., Zeng, X., Li, T., Liu, M., Zhang, S., Gao, S., et al. (2012). AtACDO1, an ABC1-like kinase gene, is involved in chlorophyll degradation and the response to photooxidative stress in *Arabidopsis*. *J. Exp. Bot.* 63, 3959–3973.

Youssef, A., Laizet, Y., Block, M. A., Marechal, E., Alcaraz, J. P., Larson, T. R., et al. (2010). Plant lipid-associated fibrillin proteins condition jasmonate production under photosynthetic stress. *Plant J.* 61, 436–445.

Ytterberg, A. J., Peltier, J. B., and van Wijk, K. J. (2006). Protein profiling of plastoglobules in chloroplasts and chromoplasts. A surprising site for differential accumulation of metabolic enzymes. *Plant Physiol.* 140, 984–997.

Zbierzak, A. M., Kanwischer, M., Wille, C., Vidi, P. A., Giavalisco, P., Lohmann, A., et al. (2010). Intersection of the tocopherol and plastoquinol metabolic pathways at the plastoglobule. *Biochem. J.* 425, 389–399.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 21 December 2012; paper pending published: 04 February 2013; accepted: 11 April 2013; published online: 25 April 2013.

Citation: Nacir H and Bréhélin C (2013) When proteomics reveals unsuspected roles: the plastoglobule example. Front. Plant Sci. ***4****:114. doi: 10.3389/fpls.2013.00114*

This article was submitted to Frontiers in Plant Proteomics, a specialty of Frontiers in Plant Science.

Copyright © 2013 Nacir and Bréhélin. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.

frontiers in
PLANT SCIENCE

REVIEW ARTICLE
published: 24 June 2013
doi: 10.3389/fpls.2013.00214

Fluorescent protein tagging as a tool to define the subcellular distribution of proteins in plants

Sandra K. Tanz[1]*, Ian Castleden[2], Ian D. Small[1,2] and A. Harvey Millar[1,2,3]

[1] ARC Centre of Excellence in Plant Energy Biology, The University of Western Australia, Perth, WA, Australia
[2] Centre of Excellence in Computational Systems Biology, The University of Western Australia, Perth, WA, Australia
[3] Centre for Comparative Analysis on Biomolecular Networks (CABiN), The University of Western Australia, Perth, WA, Australia

Edited by:
Katja Baerenfaller, Swiss Federal Institute of Technology Zurich, Switzerland

Reviewed by:
Martin Hajduch, Slovak Academy of Sciences, Slovakia
Borjana Arsova, Heinrich-Heine University, Germany

***Correspondence:**
Sandra K. Tanz, ARC Centre of Excellence in Plant Energy Biology, The University of Western Australia, 35 Stirling Highway, Crawley, Perth, WA 6009, Australia
e-mail: sandra.tanz@uwa.edu.au

Fluorescent protein (FP) tagging approaches are widely used to determine the subcellular location of plant proteins. Here we give a brief overview of FP approaches, highlight potential technical problems, and discuss what to consider when designing FP/protein fusion constructs and performing transformation assays. We analyze published FP tagging data sets along with data from proteomics studies collated in SUBA3, a subcellular location database for Arabidopsis proteins, and assess the reliability of these data sets by comparing them. We also outline the limitations of the FP tagging approach for defining protein location and investigate multiple localization claims by FP tagging. We conclude that the collation of localization datasets in databases like SUBA3 is helpful for revealing discrepancies in location attributions by different techniques and/or by different research groups.

Keywords: FP tagging, subcellular localization, database, Arabidopsis, subcellular proteomics

INTRODUCTION

Plant systems are comprised of a complex network where organs, tissues, and cell types interact with each other. Each cell, in turn, is characterized by a comparably complex network of subcellular compartments that are morphologically and functionally different. Proteins located in these subcellular compartments often share similar attributes and play roles in defining the function of these distinct cellular environments. To understand how plant cells are functionally structured, we need to know where enzymes and regulatory proteins are located within the cell at certain points in development and under particular environmental conditions (Millar et al., 2009).

Different methods can be employed to help to determine a protein's intracellular location. Computational programs that can predict the subcellular location from the protein's nucleic acid sequence are useful but not conclusive (Richly and Leister, 2004; Heazlewood et al., 2005; Reumann, 2011). In addition, some proteins exist in multiple locations (Small et al., 1998; Carrie and Small, 2012) but only a few prediction programs deal with multiple locations effectively, such as ATP (Mitschke et al., 2009), Plant-mPLoc (Chou and Shen, 2010), WOLF PSORT (Horton et al., 2007), and YLoc (Briesemeister et al., 2010; for an overview of protein localization predictors see also Tanz and Small, 2011). *In vitro* uptake studies of an exogenously added protein into an isolated organelle has been a powerful tool for detailed studies of the import process but does not reproduce the complex intracellular environment and might not always reveal targeting preference between organelles (Rudhe et al., 2002; Chew et al., 2003). Immunolabeling of proteins in tissue sections, where specific antibodies recognize the native conformation of the protein, can be laborious and time-consuming and may not always be successful. This approach is also problematic when dealing with proteins with closely related sequences. Proteomic studies employing cell fractionation and mass spectrometry (MS) to identify peptides in the purified subcellular compartments result in large, information-rich datasets (Jaquinod et al., 2007; Reumann et al., 2007, 2009; Eubel et al., 2008; Mitra et al., 2009; Ferro et al., 2010; Olinares et al., 2010; Ito et al., 2011; Klodmann et al., 2011; Lee et al., 2011; Taylor et al., 2011; Zhang and Peck, 2011; Lundquist et al., 2012). However, MS can be technically challenging as contamination of the subcellular preparation with proteins from other parts of the cell is a frequent problem and low abundance, small and hydrophobic proteins can be missed employing this approach. Fusion of fluorescent protein (FP) coding sequences to the coding regions of genes of unknown location is relatively simple and fast and can be directed to specific proteins of interest, and as a result FP tagging has become the method of choice for many plant biologists.

FP tagging and subcellular proteomic studies have become the dominant tools for determining the location of a protein within the plant cell and provide complementary and independent information. However, these high-throughput approaches are prone to both false-negative and false-positive claims of protein location. In addition, the FP tagging approach defines a protein's targeting ability and defines a final location by accumulated fluorescent signal, while the subcellular proteomics approach determines, in steady-state, where the native protein accumulates in the cell. While it is expected that these two approaches should reveal matching results in most cases, they will not always agree even when the data from both methods is sound (Millar et al., 2009). Collating location data sets of different approaches in databases like SUBA (Heazlewood et al., 2007; Tanz et al., 2013) allows users to assess these data collectively and can expose discrepancies and conflicts in location attributions by different methods and/or by

different research groups. In this report we review the current location data sets in SUBA3 (Tanz et al., 2013). Specifically, we focus on the subcellular location data by FP tagging and examine the broader reliability of these data compared to other experimental claims, discuss the limitations of the approach, and analyze localization claims by FP for the same protein in multiple locations.

THE FP TAGGING APPROACH

FP TAGGING IN PLANTS

Expression of the green fluorescent protein (GFP) from the jellyfish *Aequorea victoria* and its spectral variants within cells (Chalfie et al., 1994; Zacharias and Tsien, 2006) has stimulated many experiments to gain new insights into the organization of cellular metabolism and to better understand compartmentation of cells. FP tagging can now provide answers to the following questions: Where do proteins localize within the cell? Where do dynamic proteins move within the cell? How do individual proteins behave in response to developmental and environmental changes? However, heterologous expression of GFP in plant cells has not always been straightforward. Initially, GFP tagging was only successful in animal and fungal cells, whereas only poor GFP expression levels were observed in plant cells. This was due to the presence of a cryptic intron in the original jellyfish GFP sequence, which was incorrectly removed in plant systems. Modifications to the GFP codon sequence abolished the erroneous removal of part of the sequence and restored the expression of GFP in plant systems (Haseloff et al., 1997; Rouwendal et al., 1997).

Today, GFP and its derivatives and homologs (here collectively referred to as fluorescent proteins or FPs) are the most important fluorophores for plant cell biology and their use has been reported extensively in the literature (reviews include Hanson and Kohler, 2001; Ehrhardt, 2003; Dixit et al., 2006; Fricker et al., 2006; Berg and Beachy, 2008). Untargeted or "free" FPs are localized to the cytoplasm in plant cells but also go into the nucleus due to their small size. In addition, FPs have been targeted to all plant organelles using FP fusions incorporating location-specific signal sequences (Tian et al., 2004). In fact, a set of fluorescent organelle markers has been generated based on well-established targeting sequences (Nelson et al., 2007). All markers were generated with four different FPs in two different binary plasmids to allow for flexible combinations during co-localization studies (Nelson et al., 2007). The use of FPs to localize individual proteins is based on the ability to engineer FP fusions, with FP tagged onto the protein of interest, allowing it to be observed within intact tissue. FPs have even been used to tag viral proteins to investigate the interaction of such proteins with plant organelles (Lazarowitz and Beachy, 1999; Ueki and Citovsky, 2011). FP imaging does not require staining and allows analysis of cells in a relatively undisturbed, living state. This non-invasive way of monitoring localization and dynamics of proteins as well as there being no need for exogenous substrates or co-factors (Chalfie et al., 1994) are the main advantages of FP tagging.

A disadvantage with FP imaging, particularly in plants, has been the autofluorescence of cellular components such as cell walls and plastids, which may overlap with FP spectral signals (Deblasio et al., 2010). For example, interference by autofluorescence from the cell wall could be a problem for the localization of low abundant plasma membrane proteins. However, most modern confocal microscopes are now able to account for background autofluorescence and subtract it from FP signals based on the unique spectral profile of non-FP expressing reference images.

As increasing numbers of plant genomes are fully sequenced, high-throughput FP screens are being employed to identify gene function and regulatory networks (Cutler et al., 2000; Escobar et al., 2003; Tian et al., 2004; Koroleva et al., 2005; Marion et al., 2008). For example, a library of Arabidopsis cDNAs was generated and fused to the 3′ end of GFP. The library was then transformed into Arabidopsis *en masse* and the progeny screened for transgenic plants showing different subcellular localization patterns (Cutler et al., 2000). In a complementary study, open reading frame cDNA clones were GFP-tagged at their 3′ end and transformed cell cultures were screened for localization patterns (Koroleva et al., 2005). The Arabidopsis localizome project uses a recombineering-based gene tagging approach to generate FP fusion proteins in their chromosomal context (Zhou et al., 2011). A bacterial homologous recombination system is used to insert FP tags into genes of interest that are harbored by transformation-competent bacterial artificial chromosomes (TAC; Zhou et al., 2011). This ensures that all *cis*-regulatory sequences of a gene are included and because the genes are not amplified by PCR there is no limit to the size of a gene that can be tagged. Thus, this is a promising approach for the future that will eliminate many of the current problems encountered during FP tagging studies (see section Considerations with FP/Protein Fusions).

CONSIDERATIONS WITH FP/PROTEIN FUSIONS

The fusion of FP to enzymes often does not inhibit their catalytic activity and FP tagging is generally thought to be a "safe method" to determine the subcellular location of a protein. Indeed expressions of FP fusions of proteins have been reported to functionally complement knockout mutants (Sedbrook et al., 2002; Benkova et al., 2003; Kim et al., 2003). However, it is possible that in some cases the FP/protein fusion and the wild-type protein will differ in their subcellular locations leading to false positive results. Careful consideration is required where a protein is tagged, as the presence of the FP could hinder proper localization encoded by a transit sequence on the attached protein.

FP coding sequences are typically fused to either the 5′ or 3′ end of the coding region of a DNA sequence in question, generating N- or C-terminal FP fusions (Cutler et al., 2000; Huh et al., 2003). Alternatively, proteins can be tagged at a selected internal site, which has the advantage that targeting signals present at the 5′ or 3′ end of the coding region are not masked by the FP. For example, N-terminal fusions (FP is fused to the N terminus of the protein of interest) interfere with plastid and mitochondrial localization signals and are also likely to abrogate endoplasmic reticulum (ER) signal peptides. C-terminal fusions (FP is fused to the C terminus of the protein of interest) may also cause many proteins to mislocalize, particularly peroxisomal proteins. In addition, C-terminal fusions could mask stem-loop structures in the 3′ part of the coding sequence and the 3′ untranslated region, which are necessary for the accurate localization

of certain mRNAs (Chartrand et al., 1999). N- or C-terminal fusions may also interfere with posttranslational modification sites, such as myristylation or farnesylation sites important for membrane targeting. Indeed, some plasma membrane proteins failed to localize to the plasma membrane using N- or C-terminal tags but internally tagged proteins localized correctly (Sedbrook et al., 2002; Gardiner et al., 2003; Tian et al., 2004). In addition, more and more multi-targeted proteins are being identified. For example, proteins with peroxisomal targeting signals and chloroplast or mitochondrial transit peptides have only been identified when analyzed with separate N- and C-terminal fusion constructs (Carrie et al., 2008; Hooks et al., 2012). Thus, for correct localization it is crucial to examine N- and C-terminal FP fusion constructs and/or internally tagged proteins.

Similarly, the length of a protein sequence for fusion with an FP needs to be considered. Using the full-length sequence of a protein is desirable; however, some genes might be too long to be easily cloned into an expression vector and thus partial sequences are frequently used for localization by FP tagging. Most plastid or mitochondrial targeting sequences are located at the N-terminus and the N-terminal ~100 amino acids are generally sufficient for correct subcellular localization. However, in this case a possible second C-terminal or internally located targeting sequence might be missed, as in the case of multi-targeted proteins (Carrie et al., 2009; Hooks et al., 2012).

The promoter used in front of an FP fusion construct also needs to be considered. Often the CMV 35S promoter is used instead of the native gene promoter, which could lead to higher expression levels of the fusion construct than for the endogenous protein, and subsequently could lead to mistargeting. This could particularly affect nuclear-encoded proteins targeted to organelles, where high protein abundance could result in incomplete import. Theoretically this might also account for some false claims of dual targeting of proteins between the cytoplasm and various organelles.

In addition, the fused FP could be the reason for a conformational modification in the attached protein and a localization signal could become active, which is normally isolated in the absence of FP or when it is lacking some endogenous ligand. Also, the abundance of the fused FP may be very different from the native protein, leading to mislocation, aggregation, metabolic disturbance or the like.

CONSIDERATIONS WITH TRANSFORMATION ASSAYS DURING FP TAGGING

FP fusion constructs can be introduced into plant cells for transient assays or stably expressed in transgenic plants. With the latter, many different cell types can be investigated in which the FP/protein fusion is expressed, while not all cell types are suitable for transient expression. In addition, cell damage often occurs during DNA uptake in transient assays and inconsistent amounts of FP fusion constructs can be delivered into the cells. Thus, it is more reliable overall to analyse healthy stable transformants to define protein location by FP. However, the simplicity and speed of transient assays makes them a very valuable tool, especially when considering the extra labor and analysis it takes to generate and test stable transgenic plants. Onion epidermis is a favorite material for biolistic transient assays, because of its clear cytoplasm and single layer of living cells. Similarly, Arabidopsis cell culture, Arabidopsis seedlings and young detached leaves have also been successfully used in transient assays. Following particle bombardment with various constructs, cellular compartments such as ER, Golgi, vacuole, mitochondria, plastids and plasma membrane can all be labeled by different transiently expressed FP fusions in Arabidopsis (Nelson et al., 2007). Other popular transient expression methods include the protein expression in isolated protoplasts by electroporation or using polyethylene glycol (Miao and Jiang, 2007; Yoo et al., 2007) and the Agrobacterium-mediated infiltration in *Nicotiana benthamiana* (Yang et al., 2000) or Arabidopsis leaves (Tsuda et al., 2012).

ANALYSIS OF FP TAGGING DATA IN SUBA3

THE RELIABILITY OF FP LOCALIZATION DATA

Given that various approaches have been used to define the location of proteins, and each has its own drawbacks, it is important to ask: What is the reliability of the FP tagging approach? In an attempt to answer this question we have analyzed subcellular localization data in SUBA (Heazlewood et al., 2007; Tanz et al., 2013). At the time of writing, SUBA3 contains a total of 3788 entries based on FP tagging studies from 1074 different publications, representing 2477 unique proteins. Of these, 443 proteins have been localized at least twice independently by FP, and for 375 proteins the independent FP localizations agree. Thus, for 85% of cases, the FP data are internally consistent, whereas they disagree in the cases of 123 proteins (28%). For 13% of proteins, the FP localization of one publication has been shown to agree with a second publication, and shown to disagree with a third publication; these proteins count toward both groups. Additional data based on subcellular MS-based proteomics from 122 different publications add 22,191 entries on 7685 distinct proteins. Calculating the percentage of FP tagging and MS agreements/disagreements for proteins for which both FP tagging and proteomics data are available shows that 61% of the data agree and 39% disagree. The remaining 1593 FP entries are not confirmed nor do they disagree with MS data because no independent subcellular proteomics data relating to these proteins have been published to our knowledge. Analyzing the FP data set further and comparing it to data from subcellular MS-based proteomics reveals that 849 out of 2996 FP protein claims agree with proteomics data (**Table 1**). The number of protein claims (2996) is different to the number of unique proteins (2477) because it includes cases where the same protein has been found in multiple compartments and thus accounts for multiple entries, and it is also different to the total FP entries (3788) as a protein is only counted once per location regardless how many researchers have found it in the same location. In these 849 cases, the protein's targeting ability tested by FP tagging agrees with the protein's accumulation tested by subcellular MS and we can be confident of the location claim and how the protein got there. On the contrary, for 554 FP claims a different location has been reported by MS studies. Thus, published disagreement of subcellular location exist for these FP claims and the protein's targeting ability appears to disagree with the claimed location of the protein's accumulation.

Table 1 | Number of localizations by FP tagging for each of the 11 subcellular compartments in SUBA3.

Intracellular location	FP localizations	FP localizations confirmed by MS	FP localizations contradicted by MS	FP and MS localizations agree/disagree (% of proteins with both FP and MS data)	FP localizations neither confirmed nor contradicted
Cytoskeleton	63	0	24	0/100	39
Cytosol	514	119	101	54/46	294
ER	189	31	56	36/64	102
Extracellular	40	9	10	47/53	21
Golgi	145	42	57	42/58	46
Mitochondrion	312	90	54	63/37	168
Nucleus	779	95	149	39/61	535
Peroxisome	130	59	33	64/36	38
PM	245	122	17	88/12	106
Plastid	486	248	34	88/12	203
Vacuole	93	34	18	65/34	41
Total	2996	849	554	61/39	1593

Also shown are the numbers of FP localizations that overlap with MS localization and are thus confirmed by this approach, the numbers of FP localizations that disagree with MS localizations, the percentage of agreements and disagreements of FP with MS localizations for proteins for which FP tagging and proteomics data are both available, and the numbers of FP localizations that are neither confirmed nor contradicted by MS data. Data sets were extracted from the SUBA3 database (http://suba.plantenergy.uwa.edu.au).

Abbreviations: FP, fluorescent protein; MS, mass spectrometry; ER, endoplasmic reticulum; PM, plasma membrane.

A detailed list of the existing FP data for each of the 11 compartments in SUBA3 is shown in **Table 1**, along with the independent confirmations and disagreements by published subcellular proteomics data. For most of the compartments, the agreements between the claims for localization by FP tagging and subcellular MS lie between 36% and 65% for proteins with both FP and MS data available (**Table 1**). However, for two compartments, namely plastid and plasma membrane, 88% of proteins for which FP and MS data are available show an agreement and only 12% of FP data do not agree with the MS localization data (**Table 1**). The relatively high discrepancy between FP and MS data for most of the other compartments (35–64%, **Table 1**), likely highlights technical problems in false positive rates with both the MS and FP tagging approaches but further analysis will be required to confirm this.

The three organelles plastid, mitochondrion and peroxisome were chosen as examples to closer investigate the proteins for which a disagreement between FP and MS data has been observed.

Plastid

A total of 486 proteins have been localized to the plastid by FP tagging (**Table 1**). From these, the published plastid FP localizations of 34 proteins appear to disagree with the locations claimed by proteomics studies (Supplementary Table 1). For eight of these proteins, additional FP location data for the same proteins agree with MS location claims and thus the whole FP data set does not strictly disagree with the proteomics (Supplementary Table 1, AGIs with asterisk). Investigating the 34 proteins more closely reveals that seven proteins are known to be dual-targeted or dynamic so here the two data sets may both be correct (Supplementary Table 1, yellow). Another eight proteins clearly have a function in the plastid with two of these located in a second compartment other than the one determined by MS (Gao et al., 2003; Lurin et al., 2004; Murcha et al., 2007; Yu et al., 2008; Sun et al., 2010; Skalitzky et al., 2011). Thus, the disagreements are due to technical issues with the MS approach and could result from contamination of these proteins in sample preparations of other subcellular structures (Supplementary Table 1, blue). One of these proteins is OEP16 (At4g16160), localized by FP tagging to the plastid and by MS to the cytosol, but it has been confirmed by *in vitro* imports to be targeted to plastids and not to mitochondria, unlike the mitochondrial isoforms of this protein family (Murcha et al., 2007). The disagreement is likely due to be an error or contamination in the MS approach (Supplementary Table 1, blue). One protein (Complex I subunit At2g02510) clearly functions in the mitochondrion (Brugiere et al., 2004; Meyer et al., 2008; Klodmann et al., 2011), and the disagreement in localization is due to technical issues with the FP tagging approach (Supplementary Table 1, green). These include artifacts that may result from the foreign passenger protein affecting the targeting ability of the protein of interest, such as difference in abundance of the fusion protein, conformational changes or activation of a localization signal in the attached protein (see section Considerations with FP/Protein Fusions). The remaining 18 proteins are either unknown multi-targeted proteins located to the plastid and other compartments in the cell or the disagreement between FP and MS data is due to limitations of one or both approaches.

An interesting example for when experimental data appear to disagree but when in fact they actually complement each other is alanyl-tRNA synthetase (At1g50200). FP tagging studies found this protein to be targeted to plastids and mitochondria, whereas proteomics studies found it in the cytosol (Supplementary Table 1). Analysis of the transcription of the gene showed the presence of two translation initiation codons

(Mireau et al., 1996). Translation from the upstream AUG generates an N-terminal extension with features that target the protein to the mitochondrion and plastid, whereas most ribosomes initiate on the downstream AUG to give the shorter polypeptide corresponding in size to the cytosolic enzyme (Mireau et al., 1996). Examining the peptides identified in the cytosolic MS study (Ito et al., 2011) showed that all the cytosolic peptides significantly matching to At1g50200 (see Ito et al., 2011; Supplementary Table 1, protein hit number 68) are downstream of the second start methionine. Thus, alanyl-tRNA synthetase is only expressed at low levels in mitochondria and plastids, which explains why MS studies have not found it in these organelles but only in the cytosol and why FP studies, using the full-length sequence, have only found it in plastids and mitochondria but not in the cytosol.

Mitochondrion

Examining the 54 proteins that have been localized to the mitochondrion by FP tagging but elsewhere by subcellular MS studies shows that as many as 37 of these have additional FP data that agree with MS locations (Supplementary Table 1, AGIs with asterisk). Twenty six of these 54 proteins are known dual-targeted or dynamic proteins (Supplementary Table 1, yellow). In both cases no strict disagreement exists. Eight proteins are clearly localized to and have a function in the mitochondrion as defined by FP tagging (six of these are additionally targeted to a second compartment different to the one defined by MS) and the location disagreements are due to technical issues with the MS approach (Supplementary Table 1, blue) (Souciet et al., 1999; Escobar et al., 2003; Michalecka et al., 2003; Duchene et al., 2005; Murcha et al., 2007; Carrie et al., 2008, 2009; Palmieri et al., 2009). Another seven proteins are clearly not located in the mitochondrion but function in the plastid (Hjelmstad and Bell, 1990; Froehlich et al., 2003; Asano et al., 2004; Chew et al., 2004; Friso et al., 2004; Kleffmann et al., 2004; Peltier et al., 2004; Giacomelli et al., 2006; Peltier et al., 2006; Rutschow et al., 2008; Zybailov et al., 2008; Ferro et al., 2010; Olinares et al., 2010; Granlund et al., 2011), and here the disagreement in location is due to technical issues with the FP tagging approach (Supplementary Table 1, green). The remaining 13 proteins are either unknown multi-targeted proteins or the disagreement is due to limitations of the FP tagging or the subcellular MS approach.

Peroxisome

One hundred and thirty proteins are localized to the peroxisome by FP tagging, of which 33 are localized elsewhere by proteomic studies (**Table 1**). Eight of these have additional FP data that agree with MS locations (Supplementary Table 1, AGIs with asterisk). Eight of the 33 proteins are known to be dual-targeted or dynamic proteins and the two data sets do not necessarily disagree (Supplementary Table 1, yellow). Three proteins are clearly localized to the peroxisome and have a function in the peroxisome (Cutler et al., 2000; Carrie et al., 2008, 2009) as defined by FP tagging [with two of them, a substrate carrier (At3g55640) and a NAD(P)H dehydrogenase (At4g28220), also localized to another compartment different to the one determined by MS], and the location disagreement is due to technical issues with the MS approach (Supplementary Table 1, blue). Four proteins are either unknown multi-targeted proteins or the location difference is due to limitations of one or both approaches (Supplementary Table 1, no color). However, about half of the location discrepancies between the two methods are due to technical issues with the FP tagging approach as most proteins are most likely not localized to the peroxisome and have functions elsewhere in the cell (Supplementary Table 1, green).

MULTIPLE LOCALIZATION CLAIMS BY FP TAGGING

The redundancy that is apparent between 2996 FP localizations in **Table 1**, but 2477 unique proteins localized by FP tagging, is either due to multiple locations claimed by single literature reports or independent reports claim different locations for a single protein. Examples for the former include dual-targeted proteins to chloroplasts and mitochondria (Peeters and Small, 2001; Carrie and Small, 2012), to mitochondria and peroxisomes (Carrie et al., 2009), and to mitochondria and nucleus (Carrie et al., 2009; Hammani et al., 2011).

Analyzing only the FP tagging data in SUBA3 generated a total of 739 claims where proteins are localized to two different locations (**Table 2**). The 739 claims comprise 545 distinct proteins that have been localized to at least two different cellular compartments by FP tagging. A paired matrix of these data displays these dual localization claims for each possible subcellular compartment combination (**Table 2**). There is typically 1–20% overlap between any two subcellular proteomes. However, a 31% and 46% overlap exists between nucleus and cytosol and a 20% and 32% overlap between plastid and mitochondrion (**Table 2**). This can be partially explained by dynamic proteins that can move between nucleus and cytosol and proteins that are dual-targeted to these compartments. No doubt, the FP tagging approach has its limitations and some false positive results must also be contributing to these overlaps. Furthermore, a dual localization to the nucleus and cytosol can be due to FP artifacts, including GFP localizing by itself to the cytosol and the nucleus, which can generate false positive results to these two compartments.

Of the 739 claims where proteins are localized to two different locations, 80% (595 dual claims) are by the same literature reports. These comprise 491 proteins and because the dual location is reported by the same publication these are presumably dual- or multi-targeted proteins. 20% of these claims (representing 105 proteins) demonstrate a conflict in the literature (as they appear as different publications that contradict each other) and may highlight problems associated with the use of different FP tagging approaches. However, this set could also include biological discoveries such as identification of an unknown dual-targeted protein or showing dynamic proteins that move around in the cell in different cell types or treatments.

As examples for further investigation, the dual FP localization claims for mitochondrion/plastid, mitochondrion/peroxisome, and plastid/peroxisome were chosen.

Mitochondrion and plastid

Examining the literature references of the 100 proteins that have been located by FP tagging to the plastid and mitochondrion (**Table 2**) reveals that the dual localizations of 92 proteins are

Table 2 | A paired matrix showing dual FP localization claims for each possible subcellular compartment combination.

FP localization	Cytoskeleton	Cytosol	ER	Extracellular	Golgi	Mitochondrion	Nucleus	Peroxisome	PM	Plastid	Vacuole
Cytoskeleton	63										
Cytosol	11	514									
ER	3	15	189								
Extracellular	1	3	1	40							
Golgi	3	21	28	2	145						
Mitochondrion	2	12	4	0	2	312					
Nucleus	12	238	13	4	1	8	779				
Peroxisome	2	11	2	0	1	10	5	130			
PM	8	45	21	14	22	3	23	0	145		
Plastid	2	16	10	0	2	100	7	8	3	486	
Vacuole	1	5	10	1	10	0	2	2	7	2	93

In total, 739 claims are listed, comprising 545 distinct proteins that have been localized to at least two different cellular compartments by FP tagging. The matrix diagonal shows the set of proteins claimed in each compartment. In the matrix below the diagonal, the two-way comparisons of claims for proteins to be present in different compartments are shown. Data sets were extracted from the SUBA3 database (http://suba.plantenergy.uwa.edu.au).
Abbreviations: FP, fluorescent protein; ER, endoplasmic reticulum; PM, plasma membrane.

described in the same literature reports and these proteins are presumably dual-targeted (Supplementary Table 2, "Y"). Indeed when investigating the function of these proteins, many are known dual-targeted proteins (Supplementary Table 2, yellow). Nevertheless, four proteins are likely to be only located to the mitochondrion (Supplementary Table 2, orange) and another eight only located in plastids (Supplementary Table 2, green). Thus, here the apparent dual location is due to technical issues with the FP tagging approach that could involve a difference in abundance of the fusion protein or conformational changes leading to activation of a localization signal in the attached protein (see section Considerations with FP/Protein Fusions). For eight proteins a literature conflict exists and independent reports claim mitochondrial and plastid locations for a single protein. These proteins are either dual-located proteins, or the dual localizations are false positives due to technical problems with the FP tagging approach. In fact, based on their function and from independent literature reports, two of these eight proteins are already known dual-located proteins [dynamin 3A (At4g33650) and lon1 protease (At5g26860); Supplementary Table 2, yellow] and four are known to be located in the plastid only (Supplementary Table 2, green) indicating an issue with the FP approach.

Mitochondrion and peroxisome

Ten proteins have been localized to mitochondria and peroxisomes by FP tagging (**Table 2**) and the dual-locations of all ten proteins are each reported by the same publication, indicating all ten proteins are probably truly dual-targeted (Supplementary Table 2, "Y"). In fact, more than half of the proteins are known dual-targeted proteins from other literature (Supplementary Table 2, yellow).

Peroxisome and plastid

Of the eight distinct proteins that have been localized to the peroxisome and plastid by FP tagging, five proteins are presumably dual-targeted (same publication; Supplementary Table 2, "Y"), of which two are known dual-targeted proteins based on the function (Supplementary Table 2, yellow). The remaining three proteins demonstrate a conflict in the literature (Supplementary Table 2, "N"), of which two are clearly only located in the plastid [Rubisco small chain 1A (At1g67090; Parry et al., 2003) and chaperonin 20 (At5g20720; Carrie et al., 2009)] and the multiple localizations of these proteins likely represent technical problems with the FP tagging method (Supplementary Table 2, green). The third is the same dynamin 3A (At4g33650) noted above; the plastid claim for this protein by FP pre-dated the dual-targeting claim in mitochondria and peroxisomes by 6 years. While an explanation of why a plastid FP location was found has not been provided, the weight of genetic and other evidence appears to suggest this is a technical problem with the FP claim of the plastid location (Mano et al., 2004).

CONCLUSIONS

FP tagging with its rapidity and simplicity has become a very important tool for plant biologists to localize proteins at a subcellular level. The analysis of the FP-tagging localization dataset along with the subcellular proteomics data, both available in SUBA3, has revealed subcellular compartments where up to 88% the FP localizations have been confirmed by subcellular proteomics for proteins for which both data are available. Thus, here the protein's targeting ability agrees with its observed protein's accumulation. The more data become available in the future, the better the coverage of each subcellular proteome and the higher the agreement between different methods is likely to be. However, with more data the number of disagreements between methods will also increase. Examining the number of existing disagreements between FP tagging and MS for the individual subcellular compartments has already exposed discrepancies in location attributions between the two methods as high as 39% of the total FP datasets for proteins for which both FP and MS data are available. Such a high discrepancy highlights problems with both the MS and FP tagging approaches, which are evident

when looking closely at the organelle examples of the plastid, mitochondrion and peroxisome. Apart from the technical issues and limitations of both approaches, the disagreements can also be due to unknown biology (dual-targeted proteins or dynamic proteins). Similarly, investigating the localization disagreements within the FP tagging method showed that the majority of multiple localization claims (80%) are due to multi-targeted proteins. The remaining 20% demonstrate a conflict in location attributions by different research groups and are possibly due to problems with the FP tagging approach, but may in some cases include dynamic proteins or unknown dual-targeted proteins. To be able to assess such localization data and draw conclusions about the reliability of localization methods and expose their limitations, collation of published results in databases like SUBA3 is extremely helpful. The intersections where existing data disagree could be avenues for new biological discoveries to be made.

ACKNOWLEDGMENTS

This work was supported by the Australian Research Council [CE0561495 to A. Harvey Millar and Ian D. Small, FT110100242 to A. Harvey Millar, DE120100307 to Sandra K. Tanz]; and the Government of Western Australia through funding for the WA Centre of Excellence for Computational Systems Biology [DIR WA CoE].

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://www.frontiersin.org/Plant_Proteomics/10.3389/fpls.2013.00214/abstract

REFERENCES

Asano, T., Yoshioka, Y., Kurei, S., Sakamoto, W., and Machida, Y. (2004). A mutation of the CRUMPLED LEAF gene that encodes a protein localized in the outer envelope membrane of plastids affects the pattern of cell division, cell differentiation, and plastid division in Arabidopsis. *Plant J.* 38, 448–459. doi: 10.1111/j.1365-313X.2004.02057.x

Benkova, E., Michniewicz, M., Sauer, M., Teichmann, T., Seifertova, D., Jurgens, G., et al. (2003). Local, efflux-dependent auxin gradients as a common module for plant organ formation. *Cell* 115, 591–602. doi: 10.1016/S0092-8674(03)00924-3

Berg, R. H., and Beachy, R. N. (2008). Fluorescent protein applications in plants. *Method Cell Biol.* 85, 153–177. doi: 10.1016/S0091-679X(08)85008-X

Briesemeister, S., Rahnenfuhrer, J., and Kohlbacher, O. (2010). YLoc–an interpretable web server for predicting subcellular localization. *Nucleic Acids Res.* 38, W497–W502. doi: 10.1093/nar/gkq477

Brugiere, S., Kowalski, S., Ferro, M., Seigneurin-Berny, D., Miras, S., Salvi, D., et al. (2004). The hydrophobic proteome of mitochondrial membranes from Arabidopsis cell suspensions. *Phytochemistry* 65, 1693–1707. doi: 10.1016/j.phytochem.2004.03.028

Carrie, C., Kuhn, K., Murcha, M. W., Duncan, O., Small, I. D., O'Toole, N., et al. (2009). Approaches to defining dual-targeted proteins in Arabidopsis. *Plant J.* 57, 1128–1139. doi: 10.1111/j.1365-313X.2008.03745.x

Carrie, C., Murcha, M. W., Kuehn, K., Duncan, O., Barthet, M., Smith, P. M., et al. (2008). Type II NAD(P)H dehydrogenases are targeted to mitochondria and chloroplasts or peroxisomes in *Arabidopsis thaliana*. *FEBS Lett.* 582, 3073–3079. doi: 10.1016/j.febslet.2008.07.061

Carrie, C., and Small, I. (2012). A reevaluation of dual-targeting of proteins to mitochondria and chloroplasts. *Biochim. Biophys. Acta.* 1833, 253–259. doi: 10.1016/j.bbamcr.2012.05.029

Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W., and Prasher, D. C. (1994). Green fluorescent protein as a marker for gene expression. *Science* 263, 802–805. doi: 10.1126/science.8303295

Chartrand, P., Meng, X. H., Singer, R. H., and Long, R. M. (1999). Structural elements required for the localization of ASH1 mRNA and of a green fluorescent protein reporter particle *in vivo*. *Curr. Biol.* 9, 333–336. doi: 10.1016/S0960-9822(99)80144-4

Chew, O., Lister, R., Qbadou, S., Heazlewood, J. L., Soll, J., Schleiff, E., et al. (2004). A plant outer mitochondrial membrane protein with high amino acid sequence identity to a chloroplast protein import receptor. *FEBS Lett.* 557, 109–114. doi: 10.1016/S0014-5793(03)01457-1

Chew, O., Rudhe, C., Glaser, E., and Whelan, J. (2003). Characterization of the targeting signal of dual-targeted pea glutathione reductase. *Plant Mol. Biol.* 53, 341–356. doi: 10.1023/B:PLAN.0000006939.87660.4f

Chou, K. C., and Shen, H. B. (2010). Plant-mPLoc: a top-down strategy to augment the power for predicting plant protein subcellular localization. *PLoS ONE* 5:e11335. doi: 10.1371/journal.pone.0011335

Cutler, S. R., Ehrhardt, D. W., Griffitts, J. S., and Somerville, C. R. (2000). Random GFP: cDNA fusions enable visualization of subcellular structures in cells of Arabidopsis at a high frequency. *Proc. Natl. Acad. Sci. U.S.A.* 97, 3718–3723. doi: 10.1073/pnas.97.7.3718

Deblasio, S. L., Sylvester, A. W., and Jackson, D. (2010). Illuminating plant biology: using fluorescent proteins for high-throughput analysis of protein localization and function in plants. *Brief. Funct. Genom.* 9, 129–138. doi: 10.1093/bfgp/elp060

Dixit, R., Cyr, R., and Gilroy, S. (2006). Using intrinsically fluorescent proteins for plant cell imaging. *Plant J.* 45, 599–615. doi: 10.1111/j.1365-313X.2006.02658.x

Duchene, A. M., Giritch, A., Hoffmann, B., Cognat, V., Lancelin, D., Peeters, N. M., et al. (2005). Dual targeting is the rule for organellar aminoacyl-tRNA synthetases in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* 102, 16484–16489. doi: 10.1073/pnas.0504682102

Ehrhardt, D. (2003). GFP technology for live cell imaging. *Curr. Opin. Plant Biol.* 6, 622–628. doi: 10.1016/j.pbi.2003.09.014

Escobar, N. M., Haupt, S., Thow, G., Boevink, P., Chapman, S., and Oparka, K. (2003). High-throughput viral expression of cDNA-green fluorescent protein fusions reveals novel subcellular addresses and identifies unique proteins that interact with plasmodesmata. *Plant Cell.* 15, 1507–1523. doi: 10.1105/tpc.013284

Eubel, H., Meyer, E. H., Taylor, N. L., Bussell, J. D., O'Toole, N., Heazlewood, J. L., et al. (2008). Novel proteins, putative membrane transporters, and an integrated metabolic network are revealed by quantitative proteomic analysis of Arabidopsis cell culture peroxisomes. *Plant Physiol.* 148, 1809–1829. doi: 10.1104/pp.108.129999

Ferro, M., Brugiere, S., Salvi, D., Seigneurin-Berny, D., Court, M., Moyet, L., et al. (2010). AT_CHLORO, a comprehensive chloroplast proteome database with subplastidial localization and curated information on envelope proteins. *Mol. Cell Proteome.* 9, 1063–1084. doi: 10.1074/mcp.M900325-MCP200

Fricker, M., Runions, J., and Moore, I. (2006). Quantitative fluorescence microscopy: from art to science. *Annu. Rev. Plant Biol.* 57, 79–107. doi: 10.1146/annurev.arplant.57.032905.105239

Friso, G., Giacomelli, L., Ytterberg, A. J., Peltier, J. B., Rudella, A., Sun, Q., et al. (2004). In-depth analysis of the thylakoid membrane proteome of *Arabidopsis thaliana* chloroplasts: new proteins, new functions, and a plastid proteome database. *Plant Cell* 16, 478–499. doi: 10.1105/tpc.017814

Froehlich, J. E., Wilkerson, C. G., Ray, W. K., McAndrew, R. S., Osteryoung, K. W., Gage, D. A., et al. (2003). Proteomic study of the *Arabidopsis thaliana* chloroplastic envelope membrane utilizing alternatives to traditional two-dimensional electrophoresis. *J. Proteome. Res.* 2, 413–425. doi: 10.1021/pr034025j

Gao, H., Kadirjan-Kalbach, D., Froehlich, J. E., and Osteryoung, K. W. (2003). ARC5, a cytosolic dynamin-like protein from plants, is part of the chloroplast division machinery. *Proc. Natl. Acad. Sci. U.S.A.* 100, 4328–4333. doi: 10.1073/pnas.0530206100

Gardiner, J. C., Taylor, N. G., and Turner, S. R. (2003). Control of cellulose synthase complex localization in developing xylem. *Plant*

Cell 15, 1740–1748. doi: 10.1105/tpc.012815

Giacomelli, L., Rudella, A., and Van Wijk, K. J. (2006). High light response of the thylakoid proteome in arabidopsis wild type and the ascorbate-deficient mutant vtc2-2. A comparative proteomics study. *Plant Physiol.* 141, 685–701. doi: 10.1104/pp.106.080150

Granlund, I., Kieselbach, T., Alm, R., Schroder, W. P., and Emanuelsson, C. (2011). Clustering of MS spectra for improved protein identification rate and screening for protein variants and modifications by MALDI-MS/MS. *J. Proteom.* 74, 1190–1200. doi: 10.1016/j.jprot.2011.04.008

Hammani, K., Gobert, A., Hleibieh, K., Choulier, L., Small, I., and Giege, P. (2011). An Arabidopsis dual-localized pentatricopeptide repeat protein interacts with nuclear proteins involved in gene expression regulation. *Plant Cell* 23, 730–740. doi: 10.1105/tpc.110.081638

Hanson, M. R., and Kohler, R. H. (2001). GFP imaging: methodology and application to investigate cellular compartmentation in plants. *J. Exp. Bot.* 52, 529–539. doi: 10.1093/jexbot/52.356.529

Haseloff, J., Siemering, K. R., Prasher, D. C., and Hodge, S. (1997). Removal of a cryptic intron and subcellular localization of green fluorescent protein are required to mark transgenic Arabidopsis plants brightly. *Proc. Natl. Acad. Sci. U.S.A.* 94, 2122–2127. doi: 10.1073/pnas.94.6.2122

Heazlewood, J. L., Tonti-Filippini, J., Verboom, R. E., and Millar, A. H. (2005). Combining experimental and predicted datasets for determination of the subcellular location of proteins in Arabidopsis. *Plant Physiol.* 139, 598–609. doi: 10.1104/pp.105.065532

Heazlewood, J. L., Verboom, R. E., Tonti-Filippini, J., Small, I., and Millar, A. H. (2007). SUBA: the Arabidopsis Subcellular Database. *Nucleic Acids Res.* 35, D213–D218. doi: 10.1093/nar/gkl863

Hjelmstad, R. H., and Bell, R. M. (1990). The sn-1,2-diacylglycerol cholinephosphotransferase of Saccharomyces cerevisiae. Nucleotide sequence, transcriptional mapping, and gene product analysis of the CPT1 gene. *J. Biol. Chem.* 265, 1755–1764.

Hooks, K. B., Turner, J. E., Graham, I. A., Runions, J., and Hooks, M. A. (2012). GFP-tagging of Arabidopsis acyl-activating enzymes raises the issue of peroxisome-chloroplast import competition versus dual localization. *J. Plant Physiol.* 169, 1631–1638. doi: 10.1016/j.jplph.2012.05.026

Horton, P., Park, K. J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C. J., et al. (2007). WoLF PSORT: protein localization predictor. *Nucleic Acids Res.* 35, W585–W587. doi: 10.1093/nar/gkm259

Huh, W. K., Falvo, J. V., Gerke, L. C., Carroll, A. S., Howson, R. W., Weissman, J. S., et al. (2003). Global analysis of protein localization in budding yeast. *Nature* 425, 686–691. doi: 10.1038/nature02026

Ito, J., Batth, T. S., Petzold, C. J., Redding-Johanson, A. M., Mukhopadhyay, A., Verboom, R., et al. (2011). Analysis of the *Arabidopsis cytosolic* proteome highlights subcellular partitioning of central plant metabolism. *J. Proteome Res.* 10, 1571–1582. doi: 10.1021/pr1009433

Jaquinod, M., Villiers, F., Kieffer-Jaquinod, S., Hugouvieux, V., Bruley, C., Garin, J., et al. (2007). A proteomics dissection of *Arabidopsis thaliana* vacuoles isolated from cell culture. *Mol. Cell Proteom.* 6, 394–412. doi: 10.1074/mcp.M600250-MCP200

Kim, J. Y., Yuan, Z., and Jackson, D. (2003). Developmental regulation and significance of KNOX protein trafficking in Arabidopsis. *Development* 130, 4351–4362. doi: 10.1242/dev.00618

Kleffmann, T., Russenberger, D., Von Zychlinski, A., Christopher, W., Sjolander, K., Gruissem, W., et al. (2004). The *Arabidopsis thaliana* chloroplast proteome reveals pathway abundance and novel protein functions. *Curr. Biol.* 14, 354–362. doi: 10.1016/j.cub.2004.02.039

Klodmann, J., Senkler, M., Rode, C., and Braun, H. P. (2011). Defining the protein complex proteome of plant mitochondria. *Plant Physiol.* 157, 587–598. doi: 10.1104/pp.111.182352

Koroleva, O. A., Tomlinson, M. L., Leader, D., Shaw, P., and Doonan, J. H. (2005). High-throughput protein localization in Arabidopsis using Agrobacterium-mediated transient expression of GFP-ORF fusions. *Plant J.* 41, 162–174. doi: 10.1111/j.1365-313X.2004.02281.x

Lazarowitz, S. G., and Beachy, R. N. (1999). Viral movement proteins as probes for intracellular and intercellular trafficking in plants. *Plant Cell* 11, 535–548.

Lee, C. P., Eubel, H., O'Toole, N., and Millar, A. H. (2011). Combining proteomics of root and shoot mitochondria and transcript analysis to define constitutive and variable components in plant mitochondria. *Phytochemistry* 72, 1092–1108. doi: 10.1016/j.phytochem.2010.12.004

Lundquist, P. K., Poliakov, A., Bhuiyan, N. H., Zybailov, B., Sun, Q., and Van Wijk, K. J. (2012). The functional network of the Arabidopsis plastoglobule proteome based on quantitative proteomics and genome-wide coexpression analysis. *Plant Physiol.* 158, 1172–1192. doi: 10.1104/pp.111.193144

Lurin, C., Andres, C., Aubourg, S., Bellaoui, M., Bitton, F., Bruyere, C., et al. (2004). Genome-wide analysis of Arabidopsis pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. *Plant Cell* 16, 2089–2103. doi: 10.1105/tpc.104.022236

Mano, S., Nakamori, C., Kondo, M., Hayashi, M., and Nishimura, M. (2004). An Arabidopsis dynamin-related protein, DRP3A, controls both peroxisomal and mitochondrial division. *Plant J.* 38, 487–498. doi: 10.1111/j.1365-313X.2004.02063.x

Marion, J., Bach, L., Bellec, Y., Meyer, C., Gissot, L., and Faure, J. D. (2008). Systematic analysis of protein subcellular localization and interaction using high-throughput transient transformation of Arabidopsis seedlings. *Plant J.* 56, 169–179. doi: 10.1111/j.1365-313X.2008.03596.x

Meyer, E. H., Taylor, N. L., and Millar, A. H. (2008). Resolving and identifying protein components of plant mitochondrial respiratory complexes using three dimensions of gel electrophoresis. *J. Proteome Res.* 7, 786–794. doi: 10.1021/pr700595p

Miao, Y., and Jiang, L. (2007). Transient expression of fluorescent fusion proteins in protoplasts of suspension cultured cells. *Nat. Protoc.* 2, 2348–2353. doi: 10.1038/nprot.2007.360

Michalecka, A. M., Svensson, A. S., Johansson, F. I., Agius, S. C., Johanson, U., Brennicke, A., et al. (2003). Arabidopsis genes encoding mitochondrial type II NAD(P)H dehydrogenases have different evolutionary origin and show distinct responses to light. *Plant Physiol.* 133, 642–652. doi: 10.1104/pp.103.024208

Millar, A. H., Carrie, C., Pogson, B., and Whelan, J. (2009). Exploring the function-location nexus: using multiple lines of evidence in defining the subcellular location of plant proteins. *Plant Cell* 21, 1625–1631. doi: 10.1105/tpc.109.066019

Mireau, H., Lancelin, D., and Small, I. D. (1996). The same Arabidopsis gene encodes both cytosolic and mitochondrial alanyl-tRNA synthetases. *Plant Cell* 8, 1027–1039.

Mitra, S. K., Walters, B. T., Clouse, S. D., and Goshe, M. B. (2009). An efficient organic solvent based extraction method for the proteomic analysis of *Arabidopsis plasma* membranes. *J. Proteome Res.* 8, 2752–2767. doi: 10.1021/pr801044y

Mitschke, J., Fuss, J., Blum, T., Hoglund, A., Reski, R., Kohlbacher, O., et al. (2009). Prediction of dual protein targeting to plant organelles. *New Phytol.* 183, 224–235. doi: 10.1111/j.1469-8137.2009.02832.x

Murcha, M. W., Elhafez, D., Lister, R., Tonti-Filippini, J., Baumgartner, M., Philippar, K., et al. (2007). Characterization of the preprotein and amino acid transporter gene family in Arabidopsis. *Plant Physiol* 143, 199–212. doi: 10.1104/pp.106.090688

Nelson, B. K., Cai, X., and Nebenfuhr, A. (2007). A multicolored set of *in vivo* organelle markers for co-localization studies in Arabidopsis and other plants. *Plant J.* 51, 1126–1136. doi: 10.1111/j.1365-313X.2007.03212.x

Olinares, P. D., Ponnala, L., and Van Wijk, K. J. (2010). Megadalton complexes in the chloroplast stroma of *Arabidopsis thaliana* characterized by size exclusion chromatography, mass spectrometry, and hierarchical clustering. *Mol. Cell Proteom.* 9, 1594–1615. doi: 10.1074/mcp.M000038-MCP201

Palmieri, F., Rieder, B., Ventrella, A., Blanco, E., Do, P. T., Nunes-Nesi, A., et al. (2009). Molecular identification and functional characterization of *Arabidopsis thaliana* mitochondrial and chloroplastic NAD+ carrier proteins. *J. Biol. Chem.* 284, 31249–31259. doi: 10.1074/jbc.M109.041830

Parry, M. A., Andralojc, P. J., Mitchell, R. A., Madgwick, P. J., and Keys, A. J. (2003). Manipulation of Rubisco: the amount, activity, function and regulation. *J. Exp. Bot.* 54, 1321–1333. doi: 10.1093/jxb/erg141

Peeters, N., and Small, I. (2001). Dual targeting to mitochondria and chloroplasts. *Biochim. Biophys. Acta* 1541, 54–63. doi: 10.1016/S0167-4889(01)00146-X

Peltier, J. B., Cai, Y., Sun, Q., Zabrouskov, V., Giacomelli, L., Rudella, A., et al. (2006). The oligomeric stromal proteome of

Arabidopsis thaliana chloroplasts. *Mol. Cell Proteom.* 5, 114–133. doi: 10.1074/mcp.M500180-MCP200

Peltier, J. B., Ytterberg, A. J., Sun, Q., and Van Wijk, K. J. (2004). New functions of the thylakoid membrane proteome of *Arabidopsis thaliana* revealed by a simple, fast, and versatile fractionation strategy. *J. Biol. Chem.* 279, 49367–49383. doi: 10.1074/jbc.M406763200

Reumann, S. (2011). Toward a definition of the complete proteome of plant peroxisomes: where experimental proteomics must be complemented by bioinformatics. *Proteomics* 11, 1764–1779. doi: 10.1002/pmic.201000681

Reumann, S., Babujee, L., Ma, C., Wienkoop, S., Siemsen, T., Antonicelli, G. E., et al. (2007). Proteome analysis of Arabidopsis leaf peroxisomes reveals novel targeting peptides, metabolic pathways, and defense mechanisms. *Plant Cell* 19, 3170–3193. doi: 10.1105/tpc.107.050989

Reumann, S., Quan, S., Aung, K., Yang, P., Manandhar-Shrestha, K., Holbrook, D., et al. (2009). In-depth proteome analysis of Arabidopsis leaf peroxisomes combined with *in vivo* subcellular targeting verification indicates novel metabolic and regulatory functions of peroxisomes. *Plant Physiol.* 150, 125–143. doi: 10.1104/pp.109.137703

Richly, E., and Leister, D. (2004). An improved prediction of chloroplast proteins reveals diversities and commonalities in the chloroplast proteomes of Arabidopsis and rice. *Gene* 329, 11–16. doi: 10.1016/j.gene.2004.01.008

Rouwendal, G. J., Mendes, O., Wolbert, E. J., and Douwe De Boer, A. (1997). Enhanced expression in tobacco of the gene encoding green fluorescent protein by modification of its codon usage. *Plant Mol. Biol.* 33, 989–999. doi: 10.1023/A:1005740823703

Rudhe, C., Chew, O., Whelan, J., and Glaser, E. (2002). A novel *in vitro* system for simultaneous import of precursor proteins into mitochondria and chloroplasts. *Plant J.* 30, 213–220. doi: 10.1046/j.1365-313X.2002.01280.x

Rutschow, H., Ytterberg, A. J., Friso, G., Nilsson, R., and Van Wijk, K. J. (2008). Quantitative proteomics of a chloroplast SRP54 sorting mutant and its genetic interactions with CLPC1 in Arabidopsis. *Plant Physiol.* 148, 156–175. doi: 10.1104/pp.108.124545

Sedbrook, J. C., Carroll, K. L., Hung, K. F., Masson, P. H., and Somerville, C. R. (2002). The Arabidopsis SKU5 gene encodes an extracellular glycosyl phosphatidylinositol-anchored glycoprotein involved in directional root growth. *Plant Cell* 14, 1635–1648. doi: 10.1105/tpc.002360

Skalitzky, C. A., Martin, J. R., Harwood, J. H., Beirne, J. J., Adamczyk, B. J., Heck, G. R., et al. (2011). Plastids contain a second sec translocase system with essential functions. *Plant Physiol.* 155, 354–369. doi: 10.1104/pp.110.166546

Small, I., Wintz, H., Akashi, K., and Mireau, H. (1998). Two birds with one stone: genes that encode products targeted to two or more compartments. *Plant Mol. Biol.* 38, 265–277. doi: 10.1023/A:1006081903354

Souciet, G., Menand, B., Ovesna, J., Cosset, A., Dietrich, A., and Wintz, H. (1999). Characterization of two bifunctional Arabdopsis thaliana genes coding for mitochondrial and cytosolic forms of valyl-tRNA synthetase and threonyl-tRNA synthetase by alternative use of two in-frame AUGs. *Eur. J. Biochem.* 266, 848–854. doi: 10.1046/j.1432-1327.1999.00922.x

Sun, X., Fu, T., Chen, N., Guo, J., Ma, J., Zou, M., et al. (2010). The stromal chloroplast Deg7 protease participates in the repair of photosystem II after photoinhibition in Arabidopsis. *Plant Physiol.* 152, 1263–1273. doi: 10.1104/pp.109.150722

Tanz, S. K., Castleden, I., Hooper, C. M., Vacher, M., Small, I., and Millar, H. A. (2013). SUBA3: a database for integrating experimentation and prediction to define the SUBcellular location of proteins in Arabidopsis. *Nucleic Acids Res.* 41, D1185–D1191. doi: 10.1093/nar/gks1151

Tanz, S. K., and Small, I. (2011). In silico methods for identifying organellar and suborganellar targeting peptides in *Arabidopsis chloroplast* proteins and for predicting the topology of membrane proteins. *Method Mol. Biol.* 774, 243–280. doi: 10.1007/978-1-61779-234-2_16

Taylor, N. L., Heazlewood, J. L., and Millar, A. H. (2011). The *Arabidopsis thaliana* 2-D gel mitochondrial proteome: refining the value of reference maps for assessing protein abundance, contaminants and post-translational modifications. *Proteomics* 11, 1720–1733. doi: 10.1002/pmic.201000620

Tian, G. W., Mohanty, A., Chary, S. N., Li, S., Paap, B., Drakakaki, G., et al. (2004). High-throughput fluorescent tagging of full-length Arabidopsis gene products in planta. *Plant Physiol.* 135, 25–38. doi: 10.1104/pp.104.040139

Tsuda, K., Qi, Y., Nguyen Le, V., Bethke, G., Tsuda, Y., Glazebrook, J., and Katagiri, F. (2012). An efficient Agrobacterium-mediated transient transformation of Arabidopsis. *Plant J.* 69, 713–719. doi: 10.1111/j.1365-313X.2011.04819.x

Ueki, S., and Citovsky, V. (2011). To gate, or not to gate: regulatory mechanisms for intercellular protein transport and virus movement in plants. *Mol. Plant* 4, 782–793. doi: 10.1093/mp/ssr060

Yang, Y., Li, R., and Qi, M. (2000). *In vivo* analysis of plant promoters and transcription factors by agroinfiltration of tobacco leaves. *Plant J.* 22, 543–551. doi: 10.1046/j.1365-313x.2000.00760.x

Yoo, S. D., Cho, Y. H., and Sheen, J. (2007). Arabidopsis mesophyll protoplasts: a versatile cell system for transient gene expression analysis. *Nat. Protoc.* 2, 1565–1572. doi: 10.1038/nprot.2007.199

Yu, F., Liu, X., Alsheikh, M., Park, S., and Rodermel, S. (2008). Mutations in SUPPRESSOR OF VARIEGATION1, a factor required for normal chloroplast translation, suppress var2-mediated leaf variegation in Arabidopsis. *Plant Cell* 20, 1786–1804. doi: 10.1105/tpc.107.054965

Zacharias, D. A., and Tsien, R. Y. (2006). Molecular biology and mutation of green fluorescent protein. *Method Biochem. Anal.* 47, 83–120.

Zhang, Z. J., and Peck, S. C. (2011). Simplified enrichment of plasma membrane proteins for proteomic analyses in *Arabidopsis thaliana*. *Proteomics* 11, 1780–1788. doi: 10.1002/pmic.201000648

Zhou, R., Benavente, L. M., Stepanova, A. N., and Alonso, J. M. (2011). A recombineering-based gene tagging system for Arabidopsis. *Plant J.* 66, 712–723. doi: 10.1111/j.1365-313X.2011.04524.x

Zybailov, B., Rutschow, H., Friso, G., Rudella, A., Emanuelsson, O., Sun, Q., et al. (2008). Sorting signals, N-terminal modifications and abundance of the chloroplast proteome. *PLoS ONE* 3:e1994. doi: 10.1371/journal.pone.0001994

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 08 April 2013; paper pending published: 27 April 2013; accepted: 05 June 2013; published online: 24 June 2013.

*Citation: Tanz SK, Castleden I, Small ID and Millar AH (2013) Fluorescent protein tagging as a tool to define the subcellular distribution of proteins in plants. Front. Plant Sci. **4**:214. doi: 10.3389/fpls.2013.00214*

This article was submitted to Frontiers in Plant Proteomics, a specialty of Frontiers in Plant Science.

Copyright © 2013 Tanz, Castleden, Small and Millar. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.

frontiers in
PLANT SCIENCE

REVIEW ARTICLE
published: 30 April 2013
doi: 10.3389/fpls.2013.00112

Sub-cellular proteomics of *Medicago truncatula*

Jeonghoon Lee, Zhentian Lei, Bonnie S. Watson and Lloyd W. Sumner*

Plant Biology Division, The Samuel Roberts Noble Foundation, Ardmore, OK, USA

Edited by:
Nicolas L. Taylor, The University of Western Australia, Australia

Reviewed by:
Holger Eubel, Leibniz Universität Hannover, Germany
Frank Colditz, Leibniz University of Hannover, Germany
Ulrike Mathesius, Australian National University, Australia

****Correspondence:***
Lloyd W. Sumner, Plant Biology Division, The Samuel Roberts Noble Foundation, 2510 Sam Noble Parkway, Ardmore, OK 73401, USA.
e-mail: lwsumner@noble.org

Medicago truncatula is a leading model species and substantial molecular, genetic, genomics, proteomics, and metabolomics resources have been developed for this species to facilitate the study of legume biology. Currently, over 60 proteomics studies of *M. truncatula* have been published. Many of these have focused upon the unique symbiosis formed between legumes and nitrogen fixing rhizobia bacteria, while others have focused on seed development and the specialized proteomes of distinct tissues/organs. These include the characterization of sub-cellular organelle proteomes such as nuclei and mitochondria, as well as proteins distributed in plasma or microsomal membranes from various tissues. The isolation of sub-cellular proteins typically requires a series of steps that are labor-intensive. Thus, efficient protocols for sub-cellular fractionation, purification, and enrichment are necessary for each cellular compartment. In addition, protein extraction, solubilization, separation, and digestion prior to mass spectral identification are important to enhance the detection of low abundance proteins and to increase the overall detectable proportion of the sub-cellular proteome. This review summarizes the sub-cellular proteomics studies in *M. truncatula*.

Keywords: ***Medicago truncatula*, subcellular proteomics, mass spectrometry based proteomics, legumes, nodulation, arbuscular mycorrhizal symbiosis**

INTRODUCTION

Proteomics has become an important research tool to study complex biological systems in the post-genomics era, and the large-scale, systematic analysis of tissue and organelle specific proteins provides a more direct view of cellular processes not available through the measurement of DNA. Proteomics can provide insight on the specialized biochemistry of distinct tissues, protein localization, protein–protein interactions, enzymatic complexes, protein-metabolite complexes, post-translational modifications, and cellular signaling (Kersten et al., 2002; Baginsky, 2009). There are two general strategies for large-scale proteome analysis: bottom-up and top-down. With the bottom-up approach, complex protein mixtures are digested and resultant peptides analyzed by mass spectrometry (MS) for protein identification and quantification. Often the complex mixtures are purified using electrophoretic or chromatographic separations to render enriched or purified proteins which are then subjected to proteolytic digestion and mass spectral protein identification. One example would be two-dimensional polyacrylamide gel electrophoresis used for separation of complex protein mixtures followed by proteolytic digestion of isolated spots and mass spectral fingerprinting (Egelhofer et al., 2002). An alternate, "shotgun" proteomics approach to bottom-up sequencing uses enzymatic digestion of complex mixtures and multi-dimensional separations, such as ion exchange chromatography and high performance liquid chromatography, of the proteolytic fragments (Motoyama et al., 2006). Mass spectrometric peptide mapping and database searching is then performed.

Progress in the area of proteomics has relied heavily on the development of analytical tools for the sensitive, selective, and high-throughput studies of protein analytes (Aebersold and Goodlett, 2001). MS has evolved into a primary analytical tool for proteomics research, especially when coupled with high resolution separation techniques, due to the high information content that can be derived from these coupled techniques (Aebersold and Mann, 2003). Advances in MS have been substantially facilitated by two ionizations techniques; electrospray ionization (ESI) and matrix-assisted laser desorption/ionization (MALDI). Over the course of the past two decades, these ionization methods have become indispensable for the analysis of biological molecules, especially proteins and peptides. ESI MS produces highly charged ions directly from liquids and is therefore useful for coupling to liquid separations (Fenn et al., 1989; Griffiths et al., 2001; Domon and Aebersold, 2006). MALDI is fast and efficient and has a high tolerance to non-volatile buffers and impurities (Hillenkamp et al., 1991; Hardouin, 2007). The samples for MALDI are typically applied to solid supports and used off-line from liquid or gel separations (Walker et al., 1995; Rejtar et al., 2002).

Separation of the proteome is challenging due to the complexity of the proteome. Through post-translational modifications and differential splicing, 5–10 different protein variants can often be produced from each gene (Collins et al., 2004; Fröhlich and Arnold, 2006). A further complication is the dynamic range of protein expression at the cellular level, which can range from one up to 10^6–10^9 copies per cell (Corthals et al., 2000). Many important proteins are present at low abundance and are difficult to isolate from complex mixtures containing more highly abundant proteins with current separation methods (Hancock et al., 2002; Hunter et al., 2002). As an example, two-dimensional gel electrophoresis (2-DE) can separate up to 11,000 proteins from a whole cell lysate, but is restricted to the most highly abundant proteins in

the sample (Gygi et al., 2000). High peak capacity separations of proteins with better resolution and faster analysis times are required for continued improvement of proteomic analyses. This can be accomplished in a bottom-up fashion, in which the proteins are proteolytically digested into peptide fragments and separated before MS analysis.

Legumes are economically valuable crops in the United States, Asia, South America, and throughout the world. Legumes form unique symbiotic relationships with nitrogen-fixing soil bacteria such as *Rhizobia* which provide nitrogen to the plant (Sprent, 2007). The accumulation of high protein levels in legumes provides an economical dietary source of proteins for humans and animals. *Medicago truncatula* (**Figure 1**) has been utilized as a model legume because of its small diploid genome compared to other legumes which have large, complex polyploidy genomes (Bell et al., 2001). It also has a short generation time, prolific seed production, and good transformation efficiency (Cook, 1999; Trieu et al., 2000). Many molecular, genetic, and biochemical resources are now available for *M. truncatula,* and it is a mature model for the study of legumes and legume biology. Legumes are also a rich source of a wide variety of natural products such as flavonoids, isoflavonoids, saponins, and alkaloids which have potential in pharmaceutical and biotechnological applications (Dixon and Sumner, 2003). Many natural products are produced in specific tissues such as glandular trichomes and/or stored in sub-cellular vacuoles.

Cellular proteins are actively and passively transported across cellular and organelle membranes during normal homeostasis and in response to stress and external stimuli (Agrawal et al., 2005a,b). Accurate identification and quantification of a sub-cellular proteome is very useful and can provide insight into cellular/organelle function and dynamics (Lilley and Dupree, 2007). In addition to the field of functional proteomics, sub-cellular proteomics can provide insight into the molecular mechanisms of plant cell modulation of protein accumulation in intracellular compartments in response to various perturbations, and thus provides refined knowledge about signal transduction in organelles (Hossain et al., 2012). In order to achieve an accurate description concerning the sub-cellular localization of specific proteins, rigorous isolation procedures are important to prevent cross contamination from other cellular and organelle proteins. Some of the major sub-cellular organelles are nucleus, mitochondria, chloroplasts, plastids, peroxisomes, plasma membrane, cytosolic ribosome, and extra cellular structures. The proteomes of multiple tissues, sub-cellular organelles, and membrane systems have been describe from *M. truncatula* (**Figure 1**). Several of the global proteomics approaches for protein characterization from different tissues/organs of *M. truncatula* were based on 2-DE and MS (Mathesius et al., 2001; Watson et al., 2003). These studies focused upon total protein characterization but the analytical tools are still useful for the analysis of proteins from individual cellular compartments (Valot et al., 2006; Soares et al., 2007). The combination of protein separation by one-dimensional electrophoresis with LC-MS/MS has also been utilized for identification of differentially expressed sub-cellular proteins (Lefebvre et al., 2007; Daher et al., 2010). Due to hydrophobic nature of most membrane proteins, gel-free LC-MS/MS or multi-dimensional protein identification technology (MudPIT) systems have also been used to investigate sub-cellular membrane components (Zhang et al., 2006).

In this article, the various proteomic studies focused on sub-cellular organelles of *M. trucatula* are briefly reviewed. Several analytical technologies employed to study specific *M. trucatula* sub-cellular compartments are also presented.

M. truncatula SUB-CELLULAR PROTEOME

The plant nucleus is of great interest because it contains the genomic content and is the critical site of transcription and replication in eukaryote cells. Repetto et al. (2008) used proteome analyses to identify nuclear regulators of developing *M. truncatula* seeds and to understand the molecular mechanisms of seed development. The purity of the nuclear preparations were assessed by Western blot analyses using antibodies directed against histone H1, uridine diphosphate (UDP)-glucose pyrophosphorylase, vacuolar ATPase, photosystem II reaction center protein D1, and mitochondrial porin. A total of 179 peptides from 143 different proteins were identified using nano LC-MS/MS analysis of one-dimensional SDS-PAGE in-gel digests obtained from the seed nuclear proteome at 12 days after pollination (dap) which marks the switch to seed filling. Identified proteins were associated with roles in biogenesis of ribosomal subunits or nucleocytoplasmic trafficking. The results showed that 12-dap seeds accumulated ribosomal proteins in preparation for protein synthesis activity prior to seed filling. Other identified proteins were related to chromatin structure, transcription, RNA maturation, silencing, and transport to regulate gene expression and seed development.

Mitochondria are semiautonomous organelles involved in energy metabolism. They are also involved in many other cellular functions including amino acid and nucleotide metabolism, synthesis of cofactors, and photosynthesis. The majority of mitochondrial proteins involved in different metabolic processes are encoded by nuclear DNA, translated in the cytosol, and imported into mitochondria. Dubinin et al. (2011) isolated *M. truncatula* mitochondrial fractions from root cell suspension cultures using Percoll gradient ultracentrifugation. The mitochondrial proteome of *M. truncatula* was characterized using 2-DE IEF/Tricine SDS-PAGE and 2-DE blue-native (BN)/Tricine SDS-PAGE. A total of 144 proteins were identified by 2-DE IEF/SDS-PAGE and 51 proteins were identified on the 2-DE BN/SDS-PAGE. The identified proteins were related to oxidative phosphorylation (OXPHOS), pyruvate decarboxylation and citric acid cycle, amino acid degradation, and ATP synthesis. The *M. truncatula* 2-DE mitochondrial proteome maps revealed similarities to the one from *Arabidopsis*. However, the abundance of complex II was increased in *M. truncatula* compared to *Arabidopsis*, which indicates increased citric acid cycle activity in *M. truncatula* mitochondria. Highly abundant prohibitin complexes were also present in the mitochondrial proteome of *M. truncatula*.

Intercellular fluid (IF) and ionically bound (IB) proteins of *M. truncatula* leaf apoplast were analyzed using 2-DE separation coupled with MALDI-TOF/TOF MS identifications (Soares et al., 2007). The apoplast contains proteins of various functions including cell expansion, growth cessation, signaling, and response to biotic and abiotic stresses. Compared to other sub-cellular

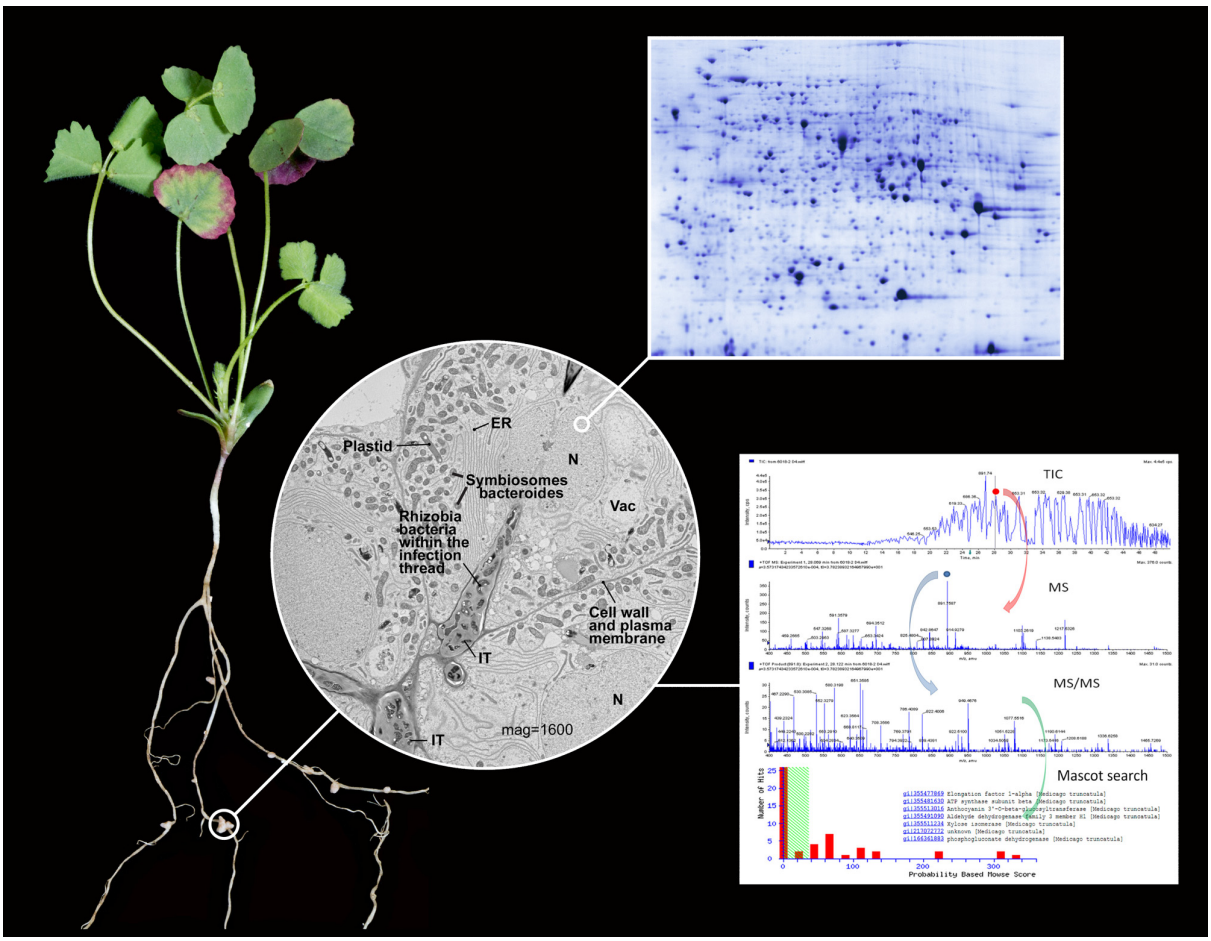


FIGURE 1 | Illustration of the sub-cellular proteomics work flow. A 4-week old *Medicago truncatula* R108 plant is shown with root nodules 18 days post inoculation with *S. meliloti*. The root nodules are specialized root organs resulting from the symbiotic infection of legumes with soil rhizobia bacteria that enable nitrogen fixation. Symbiotic nitrogen fixation provides a ready supply of nitrogen to the plant and carbon to the bacteria, and results in plants with high protein content. Thus, legumes serve as important nutritional resources throughout the world for humans and animals. The central micrograph shows the cellular components of *M. truncatula* root nodule during an early stage of infection and at a magnification of 1600. Clearly visible and labeled are traditional organelles such as the nucleus (N), vacuole (Vac), endoplasmic reticulum (ER), plastid (P) containing starch grains, and cell wall/plasma membrane. The micrograph provides an image of the rhizobia infection thread (IT) which results from the invagination of the cell wall/plasma membrane and contain rhizobium bacteria. Bacteroides are formed as the bacteria segregate from the infection thread and enter the cell and are encapsulation by a symbiosome plant membrane. The number of bacteroides generally increases with the maturity of the nodule cell. Many of the sub-cellular organelles and especially those related to symbiosis have been isolated and analyzed using various proteomics methods such as 2-DE or LC-MS/MS to better understand the protein composition and function of these specialized organelles. An example of a nano LC-MS/MS shotgun proteomics experiment is provided that includes the total ion chromatogram (TIC), full scan mass spectrum (MS), and tandem mass spectrum (MS/MS). These data are used to search predicted proteins from DNA or RNA or sequenced protein databases to identify proteins observed in the LC-MS/MS experiment. The *M. truncatula* micrograph was provided by Dr. Jin Nakashima, Manager of the Noble Foundation Cellular Imaging Facility, and Jiejan Xi. The plant photo was provided by Mr. Broderick Sterns. The representative 2-DE and LC-MS/MS images were provided by the authors.

proteomics reports, a lower number of proteins from apoplast were characterized because of the difficulty in isolating apoplastic proteins free from intracellular contamination. Proteins were extracted with a sodium acetate buffer and were isolated by centrifugation followed by pressure-assisted filtration using molecular weight cut-off membrane filters. Soares et al. (2007) separated 220 IF and 84 IB proteins in the *M. truncatula* leaf apoplast. Malate dehydrogenase activity was measured as a marker of cytosolic contamination from leaf extracts. An analysis of 2-DE of apoplastic proteins revealed the IF and IB proteins consisted largely of different protein populations representing distinct functional components of the apoplast. The authors found a high number of chitinases and β-1,3-glucanases which are glycine-rich proteins associated with defense and predominate the IF. Identified IB proteins were related to energy production/conversion, oxidoreductases, transport/binding of solutes, and cell wall structure.

Extracellular and secreted proteins (cumulatively referred to as the secretome) serve important roles in intercellular communication, development, and defense. Many secreted proteins are found in the apoplast. However, isolating proteins from apoplast is challenging and often contaminated with

intracellular proteins as noted above. To circumvent intracellular contamination and to better understand apoplast protein composition, Kusumawati et al. (2008) proposed cell cultures as a model system to study apoplast proteins and characterized the secreted proteins from the extracellular medium of three *Medicago* cell suspension cultures. These included *M. truncatula* 2HA, *M. truncatula sickle*, and *M. sativa* cell lines. The authors suggest that cell suspension cultures provide an effective method to obtain representative apoplastic proteins without damaging the plant and potentially little or no sample contamination with cytoplasmic proteins. The *M. truncatula* cell culture secretome was isolated using a two-step centrifugation process. A total of 26 proteins were identified in the cultures derived from *M. truncatula* using SDS-PAGE and MALDI-TOF/TOF MS. Among the identified proteins, three secreted proteases including a subtilisin-like serine protease, aspartyl protease, and a serine carboxypeptidase were identified and detected only in *M. truncatula* cell cultures and not in the other Medicago cell cultures. Twelve putative defense response proteins including four chitinases, a peroxidase, three thaumatin-like PR5 proteins, PR1a, PR4a, and a prehevein-like protein were also identified from *M. truncatula* secretome. The authors conclude that all the identified cell culture secreted proteins are part of the classical or non-classical secretion pathways and have not been reported in other *M. truncatula* tissues. Therefore, they further conclude that cell culture secreted proteins are a good approximation of the appoplastic proteins of *Medicago* spp.

Plant membranes are key in maintaining homeostasis and in the transport/storage of cellular materials. They are also important in the perception and transduction of signals that enable effective communication with the cell's surroundings. Valot et al. (2004) reported on the total root microsomal proteome from *M. truncatula* due to the importance of this model legume in studying plant–microbe interactions; especially symbiosis. Microsomal proteins were separated using 2-DE, and three different extraction methods were compared for preparing microsomal proteins. The extraction methods were based on phenol, acetone, and a chloroform/methanol mixture. A total of 440 microsomal proteins were visualized using analytical, silver stained 2-DE. Ninety-six proteins from 115 Coomassie stained micropreparative 2-DE protein spots were identified following in-gel digestion, MALDI-TOF MS peptide mass fingerprinting, and *M. truncatula* clustered EST database searches. This search method led to an increase in the percentage of successfully identified proteins to 83%. The Valot et al. (2004) report provided early methods for *M. truncatula* membrane proteomics and a catalog of membrane proteins useful to future studies of these important cellular components.

Lefebvre et al. (2007) investigated the role of lipid rafts from plasma membranes of *M. truncatula* and provided an extensive analysis of their structure, lipid composition, and protein content. In the process, the authors identified 270 proteins associated with *M. truncatula* lipid rafts. The contamination of the plasma membranes was evaluated through the measurement of azide-sensitive ATPase activity for mitochondria and nitrate-sensitive ATPase for tonoplasts. The major proteins found were related to signaling, transport proteins, redox proteins, cytoskeleton, trafficking, and protein stability. The authors conclude that the lipid rafts contain a complete plasma membrane redox system important for stress mediation and defense, and contain a large number of receptor-like protein kinases (RLKs) that re-enforce the hypothesis that plant lipid rafts are part of the signaling network similar to animals.

The sub-cellular proteomes of membrane-associated protein modifications in response to arbuscular mycorrhizal (AM) symbiosis have also been studied (Valot et al., 2005). This early study focused on the identification of AM fungal proteins in planta. Total membrane fractions were obtained by differential centrifugation. Membrane proteins were then extracted with a cold mixture of chloroform/methanol (6/3, v/v). Membrane proteins were extracted from AM fungal (*Glomus intraradices*) inoculated or non-inoculated *M. truncatula* roots. Thirty-six 2-DE protein spots were differently accumulated where 15 proteins were induced, three were up-regulated, and 18 were down-regulated compared to 2-DE protein spots from non-inoculated roots. Among them, 25 spots were identified with MALDI-TOF MS peptide mass fingerprinting. Most of the identified proteins had not been associated with AM symbiosis previously via proteomics, transcriptomics, nor or suppressive subtractive hybridization which illustrated the potential of direct proteomics. Nineteen of the 25 mycorrhiza-responsive proteins were found to be not regulated by phosphate supply further suggesting that these proteins could serve as markers for AM symbiosis.

Valot et al. (2006) followed with another report focused upon the study of periarbuscular membrane composition which identified 78 proteins from enriched membrane fractions obtained using a discontinuous sucrose gradient method. Marker enzymes for the plant cell membranes were assayed using K^+, Mg^{2+}-ATPase sensitivity assays. Pyrophosphatase, inosine diphosphatase, nicotinamide adenine dinucleotide (NADH)-cytochrome c reductase insensitive to antimycin A and cytochrome c oxidase were used for markers for tonoplast, golgi, endoplasmic reticulum, and mitochondria respectively. In this study, the authors used two-dimensional LC-MS/MS or gelC-LC-MS/MS for protein separation and identification as opposed to the 2-DE method used in their earlier report which allowed for the identification of a larger number of proteins (Valot et al., 2005). Comparison between *G. intraradices* inoculated and uninoculated membrane fractions revealed two differentially accumulated proteins; i.e., H^+-ATPase (Mtha1) and a predicted glycosylphosphatidylinositol-anchored blue copper-binding protein (MtBcp1). The role of these proteins in AM symbiosis remains to be investigated.

Symbiosome membrane proteins from *M. truncatula* and the corresponding *Rhizobia* bacterium *Sinorhizobium meliloti* were investigated using 2-DE and LC-MS/MS (Catalano et al., 2004). Symbiosome membranes derived from the plant plasma membrane encircle *Rhizobia* in infected root nodule cells. See **Figure 1**. This membrane is important for biological nitrogen fixation. Proteins in the symbiosome membrane are critical for transport, energy, metabolic processes, nodule formation, signaling, pathogen response, and protein destination. Symbiosomes were isolated from *M. truncatula* root nodules using differential centrifugation. Western blot analyses were used to evaluate the fraction purity by comparison of protein distributions between 2-DE of symbiosome membrane and other symbiosome fractions.

Fifty-one proteins, mostly associated with protein destination and storage, were identified from the symbiosome membrane. Twenty-eight plant symbiosome proteins were functionally classified into energy and transport, protein destination, nodule-specific, and unclassified categories. The proteomics results provide a better defined biochemical composition of the symbiosome membrane and serve as a hypothesis generating dataset to better understand the mechanism of plant–rhizobia symbiosis.

Larrainzar et al. (2007) characterized the proteome of *M. truncatula* root nodules in response to drought stress. Proteins from nodules of *M. truncatula* in symbiosis with *S. meliloti* were profiled using fast protein liquid chromatography (FPLC) and two-dimensional LC-MS/MS. Western blot analyses were performed using antibodies against NifDK as a marker for bacteroid contamination. A quantitative analysis of plant and bacteroid responses to drought stress was also performed using one-dimensional-LC-MS/MS. The larger identified functional protein classes were involved in protein synthesis and degradation, amino acid metabolism, and glycolytic pathway and TCA cycle. The other functional groups comprised proteins involved in redox state control, defense against biotic and abiotic stress, and signaling processes in nitrogen-fixing nodules. In the quantitative study, the authors found the relative content of five nodule proteins such as Met synthase, SuSy, Asn synthetase (AS), Lb, and the transcriptional eukaryotic elongation factor-2 (eEF-2) decreased following water deficit compared to those from control nodules. The data confirm the role of SuSy and identified four new enzymes involved in drought stress.

Daher et al. (2010) studied the root plastid proteome using nano LC-MS/MS and identified 266 protein candidates which have a role in nucleic acid-related processes, carbohydrate, and nitrogen/sulfur metabolisms, and stress response mechanisms. A major challenge to root plastid proteomics was the isolation of the colorless, low abundant, heterogeneous, and fragile organelles which was ultimately achieved using a differential centrifugation method originally developed for pea (*Pisum sativum*). Another challenge was the identification of a whole set of proteins with diverse chemical properties. The structures of non-photosynthetic plastids are important because they are associated with nitrogen-assimilation, starch, and lipid synthesis. Plastids are also involved in connecting individual organelles during symbiotic nitrogen fixation. In this article, the authors provide an impressive comparison of the *M. truncatula* root plastid proteome relative to the proteomes of amyloplasts, chloroplasts, and proplastids. The authors conclude that the functional distribution of the *M. truncatula* root plastids proteome mainly resemble that of wheat endosperm amyloplasts and tobacco proplastids, but are markedly different from chloroplasts.

CONCLUSION

Sub-cellular proteomics, the analysis of proteins purified from a cell compartment, has emerged as a promising method to better understand the spatial segregation of cellular processes and organelle function. A growing number of studies of plant proteomics have been reported on elucidating the protein functions and dynamics of plant sub-cellular compartments. Although more extensive sub-cellular proteomics literature exists for *Arabidopsis* and rice, research using *M. truncatula* as a model system is increasing with the bulk of these efforts focused upon plant–microbe interactions and symbiosis. The current *M. truncatula* sub-cellular proteomics literature reveals nuclear proteome contains specific proteins involved in signaling and gene regulation; whereas the mitochondrial and chloroplast proteomes revealed protein fractions implicated in energy production, either in electron transport or in ATP production. Additional studies on the apoplast or secreted proteomes revealed predominantly defense related and cell wall modifying proteins. Substantial effort has also been focused on *M. truncatula* membranes during mutualism with AM symbiosis and many of the identified proteins had not been associated with AM symbiosis previously thereby providing additional compositional information relative to this important plant–fungal interaction. Legumes are unique in their ability to form unique symbiotic relationships with soil rhizobia in the fixation of nitrogen. Identified proteins in the *M. truncatula* rhizobia symbiosome membranes are critical for transport, energy, metabolic processes, nodule formation, signaling, pathogen response, and protein destination. These results provide novel insight into the exchange of nutrients and plant–rhizobia communication.

Overall the number of *M. truncatula* sub-cellular identified proteins is still quite modest relative to the number of proteins expected to be present and relative to other plant species such as *Arabidopsis* and rice. More comprehensive sub-cellular proteomics analyses; especially related to symbiosis, are expected to be even more informative. The accurate characterization of the proteome in different sub-cellular locations also remains a challenging task dependent on the successful isolation of purified organelles. The extent of error associated with the interpretation of protein sub-cellular localization is often due to inevitable contamination from other sub-cellular components. Accurate methods to detect and quantify proteins in mixtures are necessary to assess the purity of sub-cellular purification. Although the analysis of marker enzymes based on organelle specific antibodies and the microscopic detection of sub-organelle fractions are still useful to determine the purity, the accuracy is low and varies considerably. However, the combination of 2-DE profiling, MS-based measurements, and the use of genomic data can be useful for comparing protein identities and abundances among samples. In addition to developments in sub-cellular isolation methods, advances in bioinformatics are still needed to enable large-scale comparison and correlation of large datasets from various sub-cellular proteomics studies to better describe complex biological system.

REFERENCES

Aebersold, R., and Goodlett, D. R. (2001). Mass spectrometry in proteomics. *Chem. Rev.* 101, 269–295.

Aebersold, R., and Mann, M. (2003). Mass spectrometry-based proteomics. *Nature* 422, 198–207.

Agrawal, G. K., Yonekura, M., Iwahashi, Y., Iwahashi, H., and Rakwal, R. (2005a). System, trends and perspectives of proteomics in dicot plants Part II: proteomes of the complex developmental stages. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 815, 125–136.

Agrawal, G. K., Yonekura, M., Iwahashi, Y., Iwahashi, H., and Rakwal, R. (2005b). System, trends and

perspectives of proteomics in dicot plants. Part III: unraveling the proteomes influenced by the environment, and at the levels of function and genetic relationships. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 815, 137–145.

Baginsky, S. (2009). Plant proteomics: concepts, applications, and novel strategies for data interpretation. *Mass Spectrom. Rev.* 28, 93–120.

Bell, C. J., Dixon, R. A., Farmer, A. D., Flores, R., Inman, J., Gonzales, R. A., et al. (2001). The *Medicago* genome initiative: a model legume database. *Nucleic Acids Res.* 29, 114–117.

Catalano, C. M., Lane, W. S., and Sherrier, D. J. (2004). Biochemical characterization of symbiosome membrane proteins from *Medicago truncatula* root nodules. *Electrophoresis* 25, 519–531.

Collins, F. S., Lander, E. S., Rogers, J., and Waterson, R. H. (2004). Finishing the euchromatic sequence of the human genome. *Nature* 431, 931–945.

Cook, D. R. (1999). *Medicago truncatula* – a model in the making! *Curr. Opin. Plant Biol.* 2, 301–304.

Corthals, G. L., Wasinger, V. C., Hochstrasser, D. F., and Sanchez, J.-C. (2000). The dynamic range of protein expression: a challenge for proteomic research. *Electrophoresis* 21, 1104–1115.

Daher, Z., Recorbet, G., Valot, B., Robert, F., Balliau, T., Potin, S., et al. (2010). Proteomic analysis of *Medicago truncatula* root plastids. *Proteomics* 10, 2123–2137.

Dixon, R. A., and Sumner, L. W. (2003). Legume natural products: understanding and manipulating complex pathways for human and animal health. *Plant Physiol.* 131, 878–885.

Domon, B., and Aebersold, R. (2006). Mass spectrometry and protein analysis. *Science* 312, 212–217.

Dubinin, J., Braun, H.-P., Schmitz, U., and Colditz, F. (2011). The mitochondrial proteome of the model legume *Medicago truncatula. Biochim. Biophys. Acta* 1814, 1658–1668.

Egelhofer, V., Gobom, J., Seitz, H., Giavalisco, P., Lehrach, H., and Nordhoff, E. (2002). Protein identification by MALDI-TOF-MS peptide mapping: a new strategy. *Anal. Chem.* 74, 1760–1771.

Fenn, J., Mann, M., Meng, C., Wong, S., and Whitehouse, C. (1989). Electrospray ionization for mass spectrometry of large biomolecules. *Science* 246, 64–71.

Fröhlich, T., and Arnold, G. J. (2006). Proteome research based on modern liquid chromatography – tandem mass spectrometry: separation, identification and quantification. *J. Neural Transm.* 113, 973–994.

Griffiths, W. J., Jonsson, A. P., Liu, S., Rai, D. K., and Wang, Y. (2001). Electrospray and tandem mass spectrometry in biochemistry. *Biochem. J.* 355, 545–561.

Gygi, S. P., Corthals, G. L., Zhang, Y., Rochon, Y., and Aebersold, R. (2000). Evaluation of two-dimensional gel electrophoresis-based proteome analysis technology. *Proc. Natl. Acad. Sci. U.S.A.* 97, 9390–9395.

Hancock, W. S., Wu, S. L., and Shieh, P. (2002). The challenges of developing a sound proteomics strategy. *Proteomics* 2, 352–359.

Hardouin, J. (2007). Protein sequence information by matrix-assisted laser desorption/ionization in-source decay mass spectrometry. *Mass Spectrom. Rev.* 26, 672–682.

Hillenkamp, F., Karas, M., Beavis, R. C., and Chait, B. T. (1991). Matrix-assisted laser desorption/ionization mass spectrometry of biopolymers. *Anal. Chem.* 63, 1193–1203.

Hossain, Z., Nouri, M.-Z., and Komatsu, S. (2012). Plant cell organelle proteomics in response to abiotic stress. *J. Proteome Res.* 11, 37–48.

Hunter, T. C., Andon, N. L., Koller, A., Yates, J. R., and Haynes, P. A. (2002). The functional proteomics toolbox: methods and applications. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.*782, 165–181.

Kersten, B., Bürkle, L., Kuhn, E. J., Giavalisco, P., Konthur, Z., Lueking, A., et al. (2002). Large-scale plant proteomics. *Plant Mol. Biol.* 48, 133–141.

Kusumawati, L., Imin, N., and Djordjevic, M. A. (2008). Characterization of the secretome of suspension cultures of *Medicago* species reveals proteins important for defense and development. *J. Proteome Res.* 7, 4508–4520.

Larrainzar, E., Wienkoop, S., Weckwerth, W., Ladrera, R., Arrese-Igor, C., and Gonzalez, E. M. (2007). *Medicago truncatula* root nodule proteome analysis reveals differential plant and bacteroid responses to drought stress. *Plant Physiol.* 144, 1495–1507.

Lefebvre, B., Furt, F., Hartmann, M. A., Michaelson, L. V., Carde, J. P., Sargueil-Boiron, F., et al. (2007). Characterization of lipid rafts from *Medicago truncatula* root plasma membranes: a proteomic study reveals the presence of a raft-associated redox system. *Plant Physiol.* 144, 402–418.

Lilley, K. S., and Dupree, P. (2007). Plant organelle proteomics. *Curr. Opin. Plant Biol.* 10, 594–599.

Mathesius, U., Keijzers, G., Natera, S. H., Weinman, J. J., Djordjevic, M. A., and Rolfe, B. G. (2001). Establishment of a root proteome reference map for the model legume *Medicago truncatula* using the expressed sequence tag database for peptide mass fingerprinting. *Proteomics* 1, 1424–1440.

Motoyama, A., Venable, J. D., Ruse, C. I., and Yates, J. R. (2006). Automated ultra-high-pressure multidimensional protein identification technology (UHP-MudPIT) for improved peptide identification of proteomic samples. *Anal. Chem.* 78, 5109–5118.

Rejtar, T., Hu, P., Juhasz, P., Campbell, J. M., Vestal, M. L., Preisler, J., et al. (2002). Off-line coupling of high-resolution capillary electrophoresis to MALDI-TOF and TOF/TOF MS. *J. Proteome Res.* 1, 171–179.

Repetto, O., Rogniaux, H., Firnhaber, C., Zuber, H., Küster, H., Larré, C., et al. (2008). Exploring the nuclear proteome of *Medicago truncatula* at the switch towards seed filling. *Plant J.* 56, 398–410.

Soares, N. C., Francisco, R., Ricardo, C. P., and Jackson, P. A. (2007). Proteomics of ionically bound and soluble extracellular proteins in *Medicago truncatula* leaves. *Proteomics* 7, 2070–2082.

Sprent, J. I. (2007). Evolving ideas of legume evolution and diversity: a taxonomic perspective on the occurrence of nodulation. *New Phytol.* 174, 11–25.

Trieu, A. T., Burleigh, S. H., Kardailsky, I. V., Maldonado-Mendoza, I. E., Versaw, W. K., Blaylock, L. A., et al. (2000). Transformation of *Medicago truncatula* via infiltration of seedlings or flowering plants with *Agrobacterium. Plant J.* 22, 531–541.

Valot, B., Dieu, M., Recorbet, G., Raes, M., Gianinazzi, S., and Dumas-Gaudot, E. (2005). Identification of membrane-associated proteins regulated by the arbuscular mycorrhizal symbiosis. *Plant Mol. Biol.* 59, 565–580.

Valot, B., Gianinazzi, S., and Eliane, D. G. (2004). Sub-cellular proteomic analysis of a *Medicago truncatula* root microsomal fraction. *Phytochemistry* 65, 1721–1732.

Valot, B., Negroni, L., Zivy, M., Gianinazzi, S., and Dumas-Gaudot, E. (2006). A mass spectrometric approach to identify arbuscular mycorrhiza-related proteins in root plasma membrane fractions. *Proteomics* 6, S145–S155.

Walker, K. L., Chiu, R. W., Monnig, C. A., and Wilkins, C. A. (1995). Off-line coupling of capillary electrophoresis and matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Anal. Chem.* 67, 4197–4204.

Watson, B. S., Asirvatham, V. S., Wang, L., and Sumner, L. W. (2003). Mapping the proteome of barrel medic (*Medicago truncatula*). *Plant Physiol.* 131, 1104–1123.

Zhang, K., McKinlay, C., Hocart, C. H., and Djordjevic, M. A. (2006). The *Medicago truncatula* small protein proteome and peptidome. *J. Proteome Res.* 5, 3355–3367.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 25 January 2013; paper pending published: 15 February 2013; accepted: 26 March 2013; published online: 30 April 2013.

*Citation: Lee J, Lei Z, Watson BS and Sumner LW (2013) Sub-cellular proteomics of Medicago truncatula. Front. Plant Sci. **4**:112. doi: 10.3389/fpls.2013.00112*

This article was submitted to Frontiers in Plant Proteomics, a specialty of Frontiers in Plant Science.

Copyright © 2013 Lee, Lei, Watson and Sumner. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.

frontiers in
PLANT SCIENCE

MINI REVIEW ARTICLE
published: 01 February 2013
doi: 10.3389/fpls.2013.00009

Plant secretome proteomics

Erik Alexandersson*, Ashfaq Ali, Svante Resjö and Erik Andreasson

Department of Plant Protection Biology, Swedish University of Agricultural Sciences, Alnarp, Sweden

Edited by:
Nicolas L. Taylor, The University of Western Australia, Australia

Reviewed by:
Ning Li, The Hong Kong University of Science and Technology, China
Liwen Jiang, The Chinese University of Hong Kong, Hong Kong
Ganesh K. Agrawal, Research Laboratory for Biotechnology and Biochemistry, Nepal

****Correspondence:***
Erik Alexandersson, Department of Plant Protection Biology, Swedish University of Agricultural Sciences, P.O. Box 102, SE-230 53 Alnarp, Sweden.
e-mail: erik.alexandersson@slu.se

The plant secretome refers to the set of proteins secreted out of the plant cell into the surrounding extracellular space commonly referred to as the apoplast. Secreted proteins maintain cell structure and acts in signaling and are crucial for stress responses where they can interact with pathogen effectors and control the extracellular environment. Typically, secreted proteins contain an N-terminal signal peptide and are directed through the endoplasmic reticulum/Golgi pathway. However, in plants many proteins found in the secretome lack such a signature and might follow alternative ways of secretion. This review covers techniques to isolate plant secretomes and how to identify and quantify their constituent proteins. Furthermore, bioinformatical tools to predict secretion signals and define the putative secretome are presented. Findings from proteomic studies and important protein families of plant secretomes, such as proteases and hydrolases, are highlighted.

Keywords: apoplast, mass spectrometry, plant, proteomics, secretome

INTRODUCTION

In plant cells, many proteins undergo secretion or exocytosis to the extracellular space (ECS) in order to maintain cell structure, regulate the external environment and as a part of signaling and defense mechanisms. The ECS is composed by the cell wall and space between these and contains what is often referred to as the apoplastic fluid (APF). The protein composition of the phloem and xylem are not considered in this review.

The word secretome was first used in association with proteins secreted from bacteria (Tjalsma et al., 2000). In plants, there has been an ongoing discussion on how to define secretomics. Agrawal et al. (2010) described it as "*the global study of secreted proteins into the ECS by a cell, tissue, organ or organism at any given time and conditions through known and unknown secretory mechanisms involving constitutive and regulated secretory organelles.*" This is a useful definition even though it should be noted that until now few studies have identified more than 100 proteins and thus cannot be considered to give a "global" view of the secretome.

In "classical" or "conventional" protein secretion, a mechanism highly conserved in eukaryotes, proteins containing a signal peptide are transported via the Golgi apparatus. In the *Arabidopsis* genome, 18% of proteins are predicted to be secreted (the Arabidopsis Genome Initiative, 2000), but recurrently between 40 and 70% of the proteins identified in secretome studies lack a signal peptide, and thus putatively belong to the class of leaderless secreted proteins (LSPs), even though possible contamination of proteins from other cell compartments is a concern (discussed below).

In spite of the fact that commonly the majority of identified secretome proteins lack signal peptides, unconventional protein secretion (UPS) has been little studied in plants, as pointed out in a recent review by Ding et al. (2012). UPS can be divided into two major classes: proteins are either transported in a non-vesicular mode where they pass directly from the cytosol through the plasma membrane or by various vesicular modes with membrane-bounded structures fusing with the plasma membrane before release in the ECS. Recently, a plant-specific compartment named EXPO (exocyst-positive organelle), which appears to mediate UPS without proteins passing the Golgi apparatus, trans-Golgi network, or multi-vascular body, was discovered (Wang et al., 2010).

ISOLATION AND IDENTIFICATION OF SECRETOME PROTEINS

Until the last 3–4 years suspension-cultured cells (SSCs) were the preferred choice for preparation of secretome samples and so far around half of the published studies have been conducted with SSC. Advantages over material derived *in planta* are that cell leakage can be readily estimated by determining the number of dead cells and that the separation of the secretome from intact cells by filtration and/or centrifugation is easier. Still, the general trend is that recent studies are carried out *in planta*, and there are good reasons for this switch since SSC does not provide a natural environment for the cells and physiologically relevant treatments are difficult to apply. Furthermore, it is possible to derive organ- and developmental-specific secretomes using plant material. Jung et al. (2008) reported on a striking difference in rice secretome proteins depending on whether *in planta* or SSC material were used with an overlap of only 6 spots out of 222 after resolution by two-dimensional gel electrophoresis (2D-PAGE) and a difference in the levels of identified proteins with predicted signal peptides of 27 and 76%, respectively, between the two systems. Both secretomes showed low level of contamination as determined by a malate dehydrogenase activity. This indicates that there are large differences in the protein populations derived *in planta* and from SSC and that the secretion mechanisms might be fundamentally different.

When preparing secretomes from intact plant organs caution must be taken to minimize cell breakage and leakage. Non-destructive methods are less likely to give rise to cytosolic

contamination. The most common method, vacuum infiltration-centrifugation, has been practiced for about 50 years (Klement, 1965). Whereas pH of the infiltration buffer affects the metabolic composition, osmolarity and incubation time have little effect on the eluate (Lohaus et al., 2001). Due to differences in sample infiltrability adjustment depending on species may be necessary, e.g., Nouchi et al. (2012) recently suggested a rice-specific method. For potato leaves we first thoroughly rinse the leaves with a buffer to reduce leaf surface tension and facilitate the vacuum infiltration which is repeated once. After infiltration the leaf surfaces should quickly be dried not to dilute samples. Thereafter carefully rolled leaves are transferred to 15 mL Falcon tubes with a washer at the bottom to avoid immersion of leaves into the collected APF (Ali et al., 2012; **Figure 1**). The centrifugation force should not exceed 1000*g* in order to avoid cell breakage (Terry and Bonner, 1980). In general, there is a noticeable lack of studies comparing the effect of different procedures, e.g., buffer concentrations, for secretome isolation.

An alternative, the gravity extraction method (GEM), has been proposed, but it has so far only been used in one study (Jung et al., 2008). In this method, the vacuum infiltration step is omitted and the APF is collected directly to decrease cell damage and solubilization of membrane proteins. After sap isolation it is necessary to add a cetyltrimethylammonium bromide (CTAB) precipitation step to remove interfering compounds such as carbohydrates. Various methodological adaptations might be necessary for efficient secretome preparation, e.g., a water-displacement method was developed to obtain apoplast fluid from stem tissue in poplar (Pechanova et al., 2010).

PURITY ASSESSMENT

Enzyme activity, immunoblotting and microscopy have been used to assess the purity of secretome fractions, e.g., by comparison to the microsomal fraction. Especially *in planta* studies require more stringent assessment of purity to ensure secretome fractions with little intracellular contamination. To estimate cytosolic contamination, enzyme activities of glucose-6-phosphate dehydrogenase, catalase and malate dehydrogenase are commonly measured. Based on malate dehydrogenase activity 1–3% of contamination is usually seen, but up to 10% has been reported (Song et al., 2011). In destructive methods, where the cell-wall fraction is isolated, determining the putative contamination by the plasma membrane is also necessary, e.g., by measuring H-ATPase activity (Pandey et al., 2010; Bhushan et al., 2011). Antibodies against malate dehydrogenase and RuBisCo are also frequently used to determine the contamination level (Pechanova et al., 2010; Gupta et al., 2011). It can be necessary to estimate levels of membrane damage caused by the test condition itself, and for this purpose electrolyte leakage and concentration of malondialdehyde, which is a breakdown product of membrane lipid peroxidation, have been measured (Zhou et al., 2011). In plant–pathogen interaction studies, it should be remembered that cell leakage can be caused by direct damage of hyphae or cell wall maceration, and thus be a result of the biological system itself rather than the isolation procedure. Likewise, plant developmental processes, such as programmed cell death during xylem formation, can release non-secretary cytosolic proteins into the apoplast.

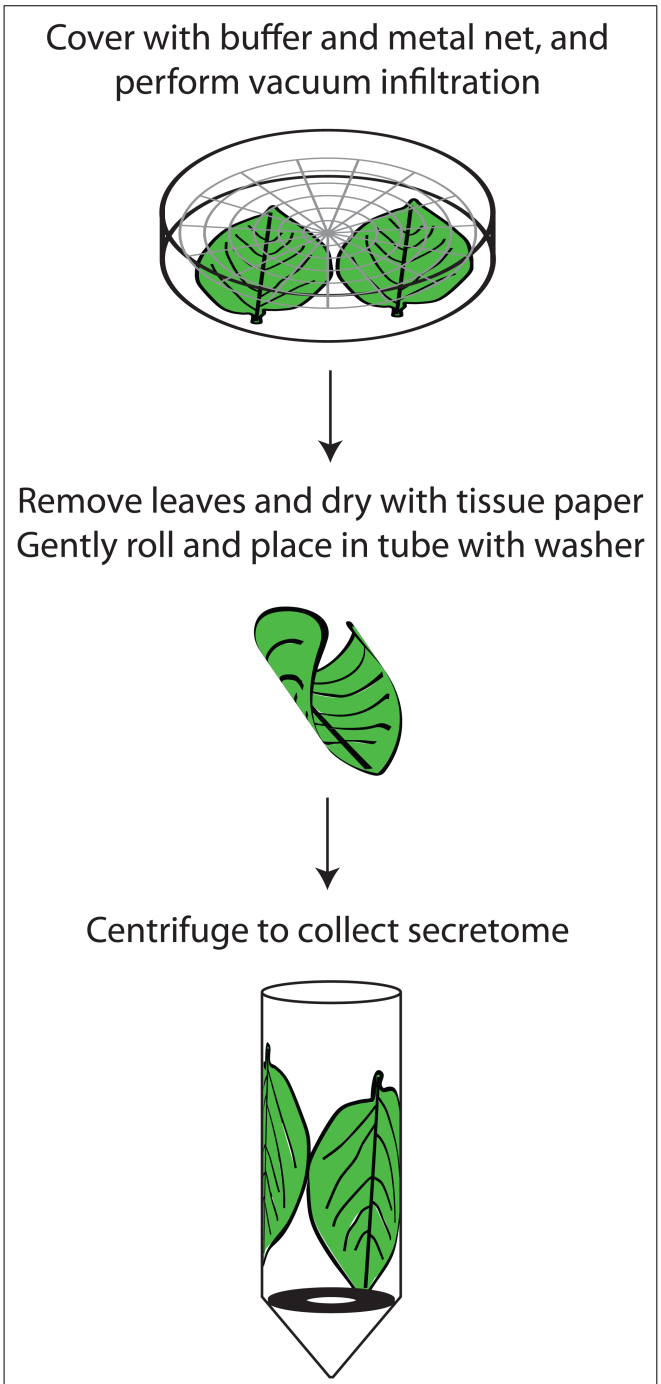


FIGURE 1 | Outline of secretome preparation in potato leaves. The leaves are washed in a beaker by gentle swirling in a solution of 1% Tween-20 and thereafter dried quickly. Up to five potato leaves can be processed at the same time. They are then placed in petri dishes containing 150 mM sodium phosphate pH 7 and 50 mM sodium chloride and covered with a metal grid to ensure that they are completely immersed. Vacuum is applied for two rounds of 5 min each, with the pressure being allowed to return to atmospheric in between. The leaves are then removed, gently dried with tissue paper, placed in tubes containing protease inhibitor and centrifuged for 1000*g* at 4°C. It is vital to roll the leaves gently, preferably several together, when placing them in the tubes to minimize cell breakage and subsequent contamination. A metal washer in the tube separates the extracted secretome from the leaves.

By estimating enrichment or depletion of peptides of a set of marker proteins, quantitative proteomics can be a good method to determine the level of contamination. Due to its sensitivity, selected reaction monitoring mass spectrometry (SRM-MS), discussed below, is particularly promising.

PROTEOMIC ANALYSES

In plant secretome analysis mainly 2D-polyacrylamide gel electrophoresis (PAGE), but also 1D sodium dodecyl sulfate (SDS)-PAGE, have been used (e.g., Gupta et al., 2011; Ali et al., 2012). 2D-PAGE is a well-established and relatively inexpensive method. While transmembrane, highly hydrophobic and very large proteins can be difficult to analyze using 2D-PAGE, apoplast proteins do not generally belong to these categories. Furthermore, 2D-PAGE separation is based on the properties of the intact proteins and splicing and post-translational modifications (PTMs) will affect migration, something that is not always desirable. Finally, multiple proteins are often identified from the same 2D-gel spot, making unambiguous identification difficult.

More recently, high-performance liquid chromatography (HPLC)-based methods, where tryptic peptide digests rather than proteins are analyzed, permit simultaneous identification of larger number of proteins. The proteins can be digested directly or prefractionated, e.g., on 1D-SDS-PAGE. Consequently, HPLC-based methods are useful for the analysis of hydrophobic proteins difficult to retain in solution during 2D-PAGE and for small proteins such as the proteolytic fragments commonly found in the protease-rich apoplast. However, these methods sometimes require more complicated sample preparation and data processing, and since analysis is done for individual peptides splice variants and PTMs may remain unrecognized.

In the HPLC-based methods, peptides can be quantified either by isotopic labeling or by label-free methods (Schulze and Usadel, 2010; Neilson et al., 2011; Yao, 2011). Isobaric tags for relative and absolute quantitation (iTRAQ) was used by Kaffarnik et al. (2009) to analyze secretome of *Arabidopsis* SSC challenged with *Pseudomonas*. In the label-free methods, quantitative data is obtained by analyzing the intensity of the mass spectrometrical signal from a peptide, or by counting the number of times it is identified (spectral counting). The label-free methods are simpler in terms of sample preparation and the number of comparisons you can make is not limited, but analysis can be more computationally challenging.

In SRM-MS, the eluate from an HPLC column is monitored to detect selected peptides enabling high dynamic range (Anderson and Hunter, 2006; Kitteringham et al., 2009). When combined with isotopically labeled internal standard peptides this method allows for sensitive absolute quantification. However, it will only measure peptides from a pre-selected set.

In plant–pathogen interactions identification of secreted proteins from more than one organism is expected. Still, very few pathogen proteins have been identified in interaction studies. Nevertheless, since proteins from the interacting organism can be expected to be a minority, precautions should be taken to avoid false positive hits from host peptides when matching pathogen peptides. In our experience, the use of a combined plant–pathogen protein database extended with a random sequence database for false discovery rate determination is a good approach.

Pathogens and other sources of stress result in a powerful oxidative burst in the secretome. Protein oxidation is known to affect both stability and enzymatical activity (Sweetlove and Moller, 2009) and oxidation proteomics is a field in rapid development. Oxidation products can be identified in global analyses by searching for peptides with oxidized amino acids, or by enrichment-based approaches, e.g., enrichment of carbonylated (Madian and Regnier, 2010) or nitrosylated peptides (Lindermayr et al., 2005). Since oxidative modifications can be quite labile, sample preparation should be optimized to minimize changes in oxidation state (Hawkins et al., 2009). Depending on the possible enrichment strategies, buffer composition should be considered during experimental design, e.g., when isolating modified cysteine residues (Lindermayr et al., 2005). A proteomic investigation of oxidation in the secretome is still lacking but could yield interesting knowledge regarding targets and extent of the oxidative burst.

BIOINFORMATICAL TOOLS AND DATABASES

For *in vitro* prediction of signal peptides SignalP (Petersen et al., 2011) has been widely used. SecretomeP is a prediction method trained on sequence features outside of the signaling peptide of secreted proteins (Bendtsen et al., 2004). It is based on mammalian and bacterial proteins, but interestingly, 60% of the LSPs identified in *Arabidopsis* SSC were predicted to be secreted by SecretomeP (Cheng et al., 2009). New tools are emerging and, e.g., LocTree2 has high prediction success especially for secreted proteins (Goldberg et al., 2012).

In the SUBA3 database 471 *Arabidopsis* proteins are registered as "extracellular" based on MS/MS identification (Heazlewood et al., 2007). Little less than half of these have been reported exclusively in the "extracellular" compartment and less than 10% are not predicted by TargetP to be extracellular proteins. Lum and Min (2011b) identified 1704 plant proteins annotated as secreted in the manually curated UniProt database. Using three prediction tools 97.5% were identified to carry a signal peptide. A database for secreted plant proteins, PlantSecKB, is currently being established (Lum and Min, 2011a).

BIOLOGICAL FINDINGS

For plant secretome studies published before 2010 we mainly refer to Agrawal et al.'s comprehensive review (Agrawal et al., 2010). Since then more than a dozen studies have appeared (highlighted below). Protein families commonly found in the secretome are listed in **Table 1**. For a review on secretomes of oomycetes and fungi we refer to Kamoun (2009).

Plant secretomes have been studied in natural conditions (e.g., Soares et al., 2007), in different cultivars (e.g., Konozy et al., 2012), during nutritional deficiency (Tran and Plaxton, 2008), after hormone treatment (e.g., Cheng et al., 2009), temperature change (Gupta and Deswal, 2012), salt stress (Song et al., 2011), and presence of pathogens and elicitors (e.g., Kim et al., 2009).

Martinez-Esteso et al. (2009) studied the grape secretome of SSC in response to methylated cyclodextrins and methyl jasmonate (MeJA) and could show that the expression levels of peroxidases, pathogenesis-related (PR) proteins, SGNH plant

Table 1 | Proteins families commonly found in the secretome. Protein family name is given together with PLAZA2.5 gene family identifiers and number of members in *Arabidopsis* and rice (spp. *japonica*) according to PLAZA2.5.

Protein family	PLAZA2.5 gene family identifier	Number identified in *A. thaliana*	Number in rice (spp. *japonica*)
Subtilase-related (e.g., P69 protein, proteinase inhibitor I9, serine proteases)	HOM000020	69	77
Eukaryotic aspartyl protease	HOM000037	58	102
Serine carboxypeptidase	HOM000054	54	60
Cysteine protease inhibitor (serpins)	HOM000290	13	20
Pectin methylesterase inhibitor/ Pectin lyase-like	HOM000034	70	47
Kunitz-type trypsin and protease inhibitor	HOM000489	4	3
Glycosyl hydrolases family 17 (e.g., PR2)	HOM000021	95	86
Glycosyl hydrolases family 31	HOM000508	5	6
Chitinase family (Glycoside hydrolase, family 19, PR3)	HOM000272	14	18
Cysteine-rich secretory protein family (e.g., PR1)	HOM000169	23	40
Thaumatin family (e.g., PR5)	HOM000107	22	36
Pectinacetylesterase	HOM000436	12	10
RmlC- and germin-like cupins	HOM000084	33	42
Peroxidases class III	HOM000453	7	10
xyloglucan endotransglucosylase/hydrolase	HOM000089	33	30

lipase-like, xyloglucan endotransglycosylase and subtilisin-like protease were affected. In a similar study, application of elicitors MeJA and cyclodextrins also led to the identification of chitinases and other PR proteins in tomato SSC (Briceno et al., 2012).

Gupta et al. (2011) characterized the secretome from SSC of the legume chickpea and identified over 700 proteins by combining 1D SDS-PAGE and HPLC-MS/MS. By comparing the secretome based on sequence homology to previously published *Arabidopsis*, *Medicago*, and rice data the authors could show a large degree of species-specificity in secreted proteins hinting at differences in the apoplast composition between species and monocots and dicots, something that needs further investigation. Cultivar-specific secretome composition also exists and in the fruit pericarp of three tomato cultivars the percentage of proteins with signal peptides varied with 50–70% (Konozy et al., 2012).

Even if only a few proteins were identified, differences in the effects on exocytosis and protein transport were observed in an elegant experiment using transient over-expression of different SNAREs in tobacco protoplasts (Ul-Rehman et al., 2011)

Several studies have targeted the rhizosphere. Over 100 secreted proteins were identified from rice roots grown in an aseptic hydroculture (Shinano et al., 2011). These proteins are believed to play an important role in the rhizosphere and a relatively high number (54%) had predicted signal peptides. Ma et al. (2010) collected proteins secreted in the mucilage of primary maize roots. Using a combination of 1D SDS-PAGE and HPLC-MS/MS, the presence of 2848 proteins were reported, which is over 50 times more compared to earlier quantitative studies of root mucilage based on 2D-PAGE or MudPIT (Basu et al., 2006; Wen et al., 2007).

The effects in the secretome of rice seedlings were studied under oxidative stress caused by 0.3 and 0.6 mM of hydrogen peroxide (H_2O_2). Of the 54 proteins identified, around half of the responsive proteins were involved in carbohydrate metabolism, with redox homeostasis as the second largest group (Zhou et al., 2011). The typical stress response marker PR1a was also up-regulated. In rice leaves more than 100 identified proteins were shown to be affected by drought stress in a time series spanning over 8 days (Pandey et al., 2010). Similarly to this study, Song et al. (2011) studied the effect of salt stress in rice during a 12 h time course and found 64 proteins with changed abundance. In both studies, proteins related to carbohydrate metabolism were the largest group of proteins with changed abundance.

Gupta and Deswal (2012) explored the secretome of seabuckthorn after low-temperature treatment and identified thaumatin-like protein and chitinase as putative antifreeze proteins. Pechanova et al. (2010) collected secretome samples and measured gene expression by microarrays from poplar growing in a riverine ecosystem exposed to multiple stresses. The composition of the secretome showed clear specificity depending on the tissue and type of stress response.

In plant interaction studies, Goulet et al. (2010) found around 90 proteins, two of which were bacterial, in the leaf apoplast of *Nicotiana benthamiana* infected by the bacterial gene vector *Agrobacterium tumefaciens*. PR proteins were found to be the most abundant proteins in the isolated fraction, and several increased greatly upon infection. Floerl et al. (2012) identified seven proteins, several of which were peroxidases, with changed abundance in the leaf secretome of *Arabidopsis* infected by the soil-borne fungal pathogen *Verticillium longisporum*. Shenton et al. (2012) used virulent and avirulent *Magnaporthe oryzae* strains to compare compatible and incompatible interactions in rice in early and late infection. A number of DUF26 domain-containing proteins increased in the compatible interaction already at 12 h. In the incompatible reaction several PR proteins were accumulated.

Interestingly, one *M. oryzae* protein, a cyclophilin, was identified and that only in the compatible interaction. The authors reduced the detergent concentration in the vacuum infiltration buffer to compensate for Tween-20 used for the *Magnaporthe* inoculation.

CONCLUSIONS AND FUTURE PERSPECTIVES

To date, over 30 secretome studies in more than 10 plant species have shown that hundreds of proteins are secreted into the apoplast. The relatively simple procedure to isolate secretome samples together with the fact that it constitutes the interface between the plant cell and its environment makes it an excellent fraction for identification of biomarkers for signal and stress cues, and highly suitable for monitoring biotic interactions. Secretome studies have firmly established the presence of a substantial level of secreted proteins lacking signal peptides and indicated a large degree of plant species specificity in the composition of secreted proteins. A transition from SSC to *in planta* systems have taken place, but comparative organ-specific studies are still lacking and little is known about the changes in the secretome during plant developmental stages, which are known to affect both metabolism, signaling pathways and resistance levels. Finally, no global study has been done of glycosylation of secreted proteins and little is known of PTMs, such as oxidation, in this fraction. To identify putative effector targets in the secretome, reliable quantitative proteomics will be crucial, since a down-regulation of a protein upon pathogen attack might indicate regulation by pathogen effectors.

Recent technical advances such as improved databases, e.g., based on RNA-seq data, and increased sensitivity of mass spectrometers will aid in the identification of specific isoforms. The regulation of single gene family members important in ecological and agricultural systems can now be dissected even in non-model species. Furthermore, the high through-put of SRM-MS will enable processing of large sample numbers, e.g., by so called eco-proteomics in field-grown material exposed to complex, natural environments, and influenced by multiple organisms. Overall, we are closer than ever to global analyses of plant secretomes similar to what we have seen for some prokaryotes.

ACKNOWLEDGMENT

Swedish Foundation for Strategic Research is thanked for financial support.

REFERENCES

Agrawal, G. K., Jwa, N. S., Lebrun, M. H., Job, D., and Rakwal, R. (2010). Plant secretome: unlocking secrets of the secreted proteins. *Proteomics* 10, 799–827.

Ali, A., Moushib, L. I., Lenman, M., Levander, F., Olsson, K., Carlson-Nilson, U., et al. (2012). Paranoid potato: phytophthora-resistant genotype shows constitutively activated defense. *Plant Signal. Behav.* 7, 400–408.

Anderson, L., and Hunter, C. L. (2006). Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins. *Mol. Cell. Proteomics* 5, 573–588.

Arabidopsis Genome Initiative. (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408, 796–815.

Basu, U., Francis, J., Whittal, R., Stephens, J., Wang, Y., Zaiane, O., et al. (2006). Extracellular proteomes of *Arabidopsis thaliana* and *Brassica napus* roots: analysis and Comparison by MudPIT and LC-MS/MS. *Plant Soil* 286, 357–376.

Bendtsen, J. D., Jensen, L. J., Blom, N., Von Heijne, G., and Brunak, S. (2004). Feature-based prediction of non-classical and leaderless protein secretion. *Protein Eng. Des. Sel.* 17, 349–356.

Bhushan, D., Jaiswal, D. K., Ray, D., Basu, D., Datta, A., Chakraborty, S., et al. (2011). Dehydration-responsive reversible and irreversible changes in the extracellular matrix: comparative proteomics of chickpea genotypes with contrasting tolerance. *J. Proteome Res.* 10, 2027–2046.

Briceno, Z., Almagro, L., Sabater-Jara, A. B., Calderon, A. A., Pedreno, M. A., and Ferrer, M. A. (2012). Enhancement of phytosterols, taraxasterol and induction of extracellular pathogenesis-related proteins in cell cultures of *Solanum lycopersicum* cv Micro-Tom elicited with cyclodextrins and methyl jasmonate. *J. Plant Physiol.* 169, 1050–1058.

Cheng, F. Y., Blackburn, K., Lin, Y. M., Goshe, M. B., and Williamson, J. D. (2009). Absolute protein quantification by LC/MS(E) for global analysis of salicylic acid-induced plant protein secretion responses. *J. Proteome Res.* 8, 82–93.

Ding, Y., Wang, J., Stierhof, Y. D., Robinson, D. G., and Jiang, L. (2012). Unconventional protein secretion. *Trends Plant Sci.* 17, 606–615.

Floerl, S., Majcherczyk, A., Possienke, M., Feussner, K., Tappe, H., Gatz, C., et al. (2012). *Verticillium longisporum* infection affects the leaf apoplastic proteome, metabolome, and cell wall properties in *Arabidopsis thaliana*. *PLoS ONE* 7:e31435. doi: 10.1371/journal.pone.0031435

Goldberg, T., Hamp, T., and Rost, B. (2012). LocTree2 predicts localization for all domains of life. *Bioinformatics* 28, i458–i465.

Goulet, C., Goulet, M. C., and Michaud, D. (2010). 2-DE proteome maps for the leaf apoplast of *Nicotiana benthamiana*. *Proteomics* 10, 2536–2544.

Gupta, R., and Deswal, R. (2012). Low temperature stress modulated secretome analysis and purification of antifreeze protein from *Hippophae rhamnoides*, a Himalayan wonder plant. *J. Proteome Res.* 11, 2684–2696.

Gupta, S., Wardhan, V., Verma, S., Gayali, S., Rajamani, U., Datta, A., et al. (2011). Characterization of the secretome of chickpea suspension culture reveals pathway abundance and the expected and unexpected secreted proteins. *J. Proteome Res.* 10, 5006–5015.

Hawkins, C. L., Morgan, P. E., and Davies, M. J. (2009). Quantification of protein modification by oxidants. *Free Radic. Biol. Med.* 46, 965–988.

Heazlewood, J. L., Verboom, R. E., Tonti-Filippini, J., Small, I., and Millar, A. H. (2007). SUBA: the *Arabidopsis* subcellular database. *Nucleic Acids Res.* 35, D213-D218.

Jung, Y. H., Jeong, S. H., Kim, S. H., Singh, R., Lee, J. E., Cho, Y. S., et al. (2008). Systematic secretome analyses of rice leaf and seed callus suspension-cultured cells: workflow development and establishment of high-density two-dimensional gel reference maps. *J. Proteome Res.* 7, 5187–5210.

Kaffarnik, F. A., Jones, A. M., Rathjen, J. P., and Peck, S. C. (2009). Effector proteins of the bacterial pathogen *Pseudomonas syringae* alter the extracellular proteome of the host plant, *Arabidopsis thaliana*. *Mol. Cell. Proteomics* 8, 145–156.

Kamoun, S. (2009). *The Secretome of Plant-Associated Fungi and Oomycetes.* Berlin Heidelberg: Springer-Verlag, 173–180.

Kim, S. T., Kang, Y. H., Wang, Y., Wu, J., Park, Z. Y., Rakwal, R., et al. (2009). Secretome analysis of differentially induced proteins in rice suspension-cultured cells triggered by rice blast fungus and elicitor. *Proteomics* 9, 1302–1313.

Kitteringham, N. R., Jenkins, R. E., Lane, C. S., Elliott, V. L., and Park, B. K. (2009). Multiple reaction monitoring for quantitative biomarker analysis in proteomics and metabolomics. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* 877, 1229–1239.

Klement, Z. (1965). Method of obtaining fluid from the intercellular spaces of foliage and the fluid's merit as substrate for phytobacterial pathogens. *Phytopathology* 55, 1033–1034.

Konozy, E. H., Rogniaux, H., Causse, M., and Faurobert, M. (2012). Proteomic analysis of tomato (*Solanum lycopersicum*) secretome. *J Plant Res.* doi: 10.1007/s10265-012-0516-4 [Epub ahead of print]

Lindermayr, C., Saalbach, G., and Durner, J. (2005). Proteomic identification of S-nitrosylated proteins in *Arabidopsis*. *Plant Physiol.* 137, 921–930.

Lohaus, G., Pennewiss, K., Sattelmacher, B., Hussmann, M., and Hermann Muehling, K. (2001). Is the infiltration-centrifugation technique appropriate for the isolation of apoplastic fluid? A critical evaluation with different plant species. *Physiol. Plant.* 111, 457–465.

Lum, G., and Min, X. J. (2011a). FunSecKB: the Fungal Secretome KnowledgeBase. *Database (Oxford)* 2011, bar001.

Lum, G., and Min, X. J. (2011b). Plant secretomes: current status and future perspectives. *Plant Omics J.* 4, 114–119.

Ma, W., Muthreich, N., Liao, C., Franz-Wachtel, M., Schutz, W., Zhang, F., et al. (2010). The mucilage proteome of maize (*Zea mays* L.) primary roots. *J. Proteome Res.* 9, 2968–2976.

Madian, A. G., and Regnier, F. E. (2010). Proteomic identification of carbonylated proteins and their oxidation sites. *J. Proteome Res.* 9, 3766–3780.

Martinez-Esteso, M. J., Selles-Marchart, S., Vera-Urbina, J. C., Pedreno, M. A., and Bru-Martinez, R. (2009). Changes of defense proteins in the extracellular proteome of grapevine (*Vitis vinifera* cv. Gamay) cell cultures in response to elicitors. *J. Proteomics* 73, 331–341.

Neilson, K. A., Ali, N. A., Muralidharan, S., Mirzaei, M., Mariani, M., Assadourian, G., et al. (2011). Less label, more free: approaches in label-free quantitative mass spectrometry. *Proteomics* 11, 535–553.

Nouchi, I., Hayashi, K., Hiradate, S., Ishikawa, S., Fukuoka, M., Chen, C. P., et al. (2012). Overcoming the difficulties in collecting apoplastic fluid from rice leaves by the infiltration-centrifugation method. *Plant Cell Physiol.* 53, 1659–1668.

Pandey, A., Rajamani, U., Verma, J., Subba, P., Chakraborty, N., Datta, A., et al. (2010). Identification of extracellular matrix proteins of rice (*Oryza sativa* L.) involved in dehydration-responsive network: a proteomic approach. *J. Proteome Res.* 9, 3443–3464.

Pechanova, O., Hsu, C. Y., Adams, J. P., Pechan, T., Vandervelde, L., Drnevich, J., et al. (2010). Apoplast proteome reveals that extracellular matrix contributes to multistress response in poplar. *BMC Genomics* 11:674. doi: 10.1186/1471-2164-11-674

Petersen, T. N., Brunak, S., Von Heijne, G., and Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* 8, 785–786.

Schulze, W. X., and Usadel, B. (2010). Quantitation in mass-spectrometry-based proteomics. *Annu. Rev. Plant Biol.* 61, 491–516.

Shenton, M., Berberich, T., Kamo, M., Yamashita, T., Taira, H., and Terauchi, R. (2012). Use of intercellular washing fluid to investigate the secreted proteome of the rice–*Magnaporthe* interaction. *J. Plant Res.* 125, 311–316.

Shinano, T., Komatsu, S., Yoshimura, T., Tokutake, S., Kong, F. J., Watanabe, T., et al. (2011). Proteomic analysis of secreted proteins from aseptically grown rice. *Phytochemistry* 72, 312–320.

Soares, N. C., Francisco, R., Ricardo, C. P., and Jackson, P. A. (2007). Proteomics of ionically bound and soluble extracellular proteins in *Medicago truncatula* leaves. *Proteomics* 7, 2070–2082.

Song, Y., Zhang, C., Ge, W., Zhang, Y., Burlingame, A. L., and Guo, Y. (2011). Identification of NaCl stress-responsive apoplastic proteins in rice shoot stems by 2D-DIGE. *J. Proteomics* 74, 1045–1067.

Sweetlove, L. J., and Moller, I. M. (2009). "Chapter 1 Oxidation of proteins in plants – mechanisms and consequences," in *Advances in Botanical Research*, ed. J. Jean-Pierre (Waltham: Academic Press), 1–23.

Terry, M. E., and Bonner, B. A. (1980). An examination of centrifugation as a method of extracting an extracellular solution from peas, and its use for the study of indoleacetic acid-induced growth. *Plant Physiol.* 66, 321–325.

Tjalsma, H., Bolhuis, A., Jongbloed, J. D., Bron, S., and Van Dijl, J. M. (2000). Signal peptide-dependent protein transport in *Bacillus subtilis*: a genome-based survey of the secretome. *Microbiol. Mol. Biol. Rev.* 64, 515–547.

Tran, H. T., and Plaxton, W. C. (2008). Proteomic analysis of alterations in the secretome of *Arabidopsis thaliana* suspension cells subjected to nutritional phosphate deficiency. *Proteomics* 8, 4317–4326.

Ul-Rehman, R., Rinalducci, S., Zolla, L., Dalessandro, G., and Di Sansebastiano, G. P. (2011). *Nicotiana tabacum* protoplasts secretome can evidence relations among regulatory elements of exocytosis mechanisms. *Plant Signal. Behav.* 6, 1140–1145.

Wang, J., Ding, Y., Hillmer, S., Miao, Y., Lo, S. W., Wang, X., et al. (2010). EXPO, an exocyst-positive organelle distinct from multivesicular endosomes and autophagosomes, mediates cytosol to cell wall exocytosis in *Arabidopsis* and tobacco cells. *Plant Cell* 22, 4009–4030.

Wen, F., Vanetten, H. D., Tsaprailis, G., and Hawes, M. C. (2007). Extracellular proteins in pea root tip and border cell exudates. *Plant Physiol.* 143, 773–783.

Yao, X. (2011). Derivatization or not: a choice in quantitative proteomics. *Anal. Chem.* 83, 4427–4439.

Zhou, L., Bokhari, S. A., Dong, C. J., and Liu, J. Y. (2011). Comparative proteomics analysis of the root apoplasts of rice seedlings in response to hydrogen peroxide. *PLoS ONE* 6:e16723. doi: 10.1371/journal.pone.0016723

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 30 October 2012; paper pending published: 03 December 2012; accepted: 11 January 2013; published online: 01 February 2013.

Citation: Alexandersson E, Ali A, Resjö S and Andreasson E (2013) Plant secretome proteomics. Front. Plant Sci. ***4****:9. doi: 10.3389/fpls.2013.00009*

This article was submitted to Frontiers in Plant Proteomics, a specialty of Frontiers in Plant Science.

Copyright © 2013 Alexandersson, Ali, Resjö and Andreasson. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.