



Medicago PhosphoProtein Database: a repository for *Medicago truncatula* phosphoprotein data

Christopher M. Rose¹, Muthusubramanian Venkateshwaran², Paul A. Grimsrud³, Michael S. Westphall¹, Michael R. Sussman³, Joshua J. Coon^{1,4} and Jean-Michel Ané^{2*}

¹ Department of Chemistry, University of Wisconsin, Madison, WI, USA

² Department of Agronomy, University of Wisconsin, Madison, WI, USA

³ Department of Biochemistry, University of Wisconsin, Madison, WI, USA

⁴ Department of Biomolecular Chemistry, University of Wisconsin, Madison, WI, USA

Edited by:

Joshua L. Heazlewood, Lawrence Berkeley National Laboratory, USA

Reviewed by:

Dong Xu, University of Missouri, USA

Stefanie Wienkoop, University of Vienna, Austria

*Correspondence:

Jean-Michel Ané, Department of Agronomy, University of Wisconsin, 348 Moore Hall, 1575 Linden Drive, Madison, WI 53706, USA.
e-mail: jane@wisc.edu

The ability of legume crops to fix atmospheric nitrogen via a symbiotic association with soil rhizobia makes them an essential component of many agricultural systems. Initiation of this symbiosis requires protein phosphorylation-mediated signaling in response to rhizobial signals named Nod factors. *Medicago truncatula* (Medicago) is the model system for studying legume biology, making the study of its phosphoproteome essential. Here, we describe the Medicago PhosphoProtein Database (MPPD; <http://phospho.medicago.wisc.edu>), a repository built to house phosphoprotein, phosphopeptide, and phosphosite data specific to Medicago. Currently, the MPPD holds 3,457 unique phosphopeptides that contain 3,404 non-redundant sites of phosphorylation on 829 proteins. Through the web-based interface, users are allowed to browse identified proteins or search for proteins of interest. Furthermore, we allow users to conduct BLAST searches of the database using both peptide sequences and phosphorylation motifs as queries. The data contained within the database are available for download to be investigated at the user's discretion. The MPPD will be updated continually with novel phosphoprotein and phosphopeptide identifications, with the intent of constructing an unparalleled compendium of large-scale Medicago phosphorylation data.

Keywords: *Medicago truncatula*, phosphoproteome, ETD, CAD, proteomics

INTRODUCTION

Legumes, also known as Fabaceae, are a very large and economically important group of plants (Graham and Vance, 2003). They are used as food crops, forages, and green manure throughout the world. Most legumes can develop a nitrogen-fixing association with soil bacteria known as rhizobia, which results in the formation of root nodules (Jones et al., 2007). Rhizobia thrive inside these nodules and fix atmospheric nitrogen in exchange for a carbon source (Venkateshwaran and Ané, 2011). Many legumes also establish symbiotic associations with arbuscular mycorrhizal fungi which facilitate the acquisition of nutrients (phosphorous, nitrogen, etc.) and provide some protection against environmental stresses (Ruiz-Lozano et al., 1995; Schutzendubel and Polle, 2002; Bonfante and Genre, 2010). *Medicago truncatula* (Medicago) is a well-established model for studying legume biology and, in particular, the molecular mechanisms mediating symbiotic associations (Cook, 1999). The Medicago research community has developed many genetic, proteomic, and genomic tools, including the recent release of its genome sequence (Thoquet et al., 2002; Ané et al., 2008; Colditz and Braun, 2010; Young et al., 2011).

Genetic studies in Medicago have allowed a precise dissection of the molecular mechanisms controlling the early steps of the rhizobia-legume symbiosis (Riely et al., 2004, 2006). In response to plant signals, rhizobia produce Nod factors, which

are lipochitooligosaccharides required for both root colonization and nodule organogenesis (Dénarié et al., 1996; Brelles-Marino and Ané, 2008). Nod factors are recognized by LysM-receptor kinases on the plasma membrane, including Nod Factor Perception (NFP) and LYK3 (Ben Amor et al., 2003; Arrighi et al., 2006; Smit et al., 2007). Does not Make Infections 2 (DMI2) is another receptor-like kinase residing on the plasma membrane but with leucine-rich repeats (LRR) and plays a role downstream of NFP (Endre et al., 2002; Limpens et al., 2005). The signals are transduced from the plasma membrane to the nucleus, where they activate oscillations of calcium concentrations (calcium spiking) and a calcium/calmodulin-dependent protein kinase (CCaMK) named DMI3 (Ehrhardt et al., 1996; Lévy et al., 2004; Peiter et al., 2007; Shimoda et al., 2012). Proteins interacting with and phosphorylated by protein kinases, such as LYK3, DMI2, and DMI3, have been identified by targeted studies and mediate symbiotic signaling (Kevei et al., 2007; Messinese et al., 2007; Lefebvre et al., 2010; Mbengue et al., 2010; Horvath et al., 2011; Chen et al., 2012). Hence, protein phosphorylation plays a central role in symbiotic signaling, but this is also true for many cellular signaling events mediating developmental processes and responses to external stimuli (Laugesen et al., 2006; Peck, 2006; Huber, 2007).

Except for few studies (Laugesen et al., 2006; Lima et al., 2006; Wienkoop et al., 2008), the Medicago phosphoproteome

has not been studied extensively *in vivo* until we published a large-scale phosphoproteomic study in 2010 (Grimsrud et al., 2010). This work identified phosphorylation sites on proteins isolated from Medicago roots (from whole cell lysate and membrane-enriched fractions) using immobilized metal affinity chromatography (IMAC) and tandem mass spectrometry. The data collected from this large-scale study were used to create an online Medicago PhosphoProtein Database (MPPD; <http://www.phospho.medicago.wisc.edu>). The salient features of MPPD are discussed below.

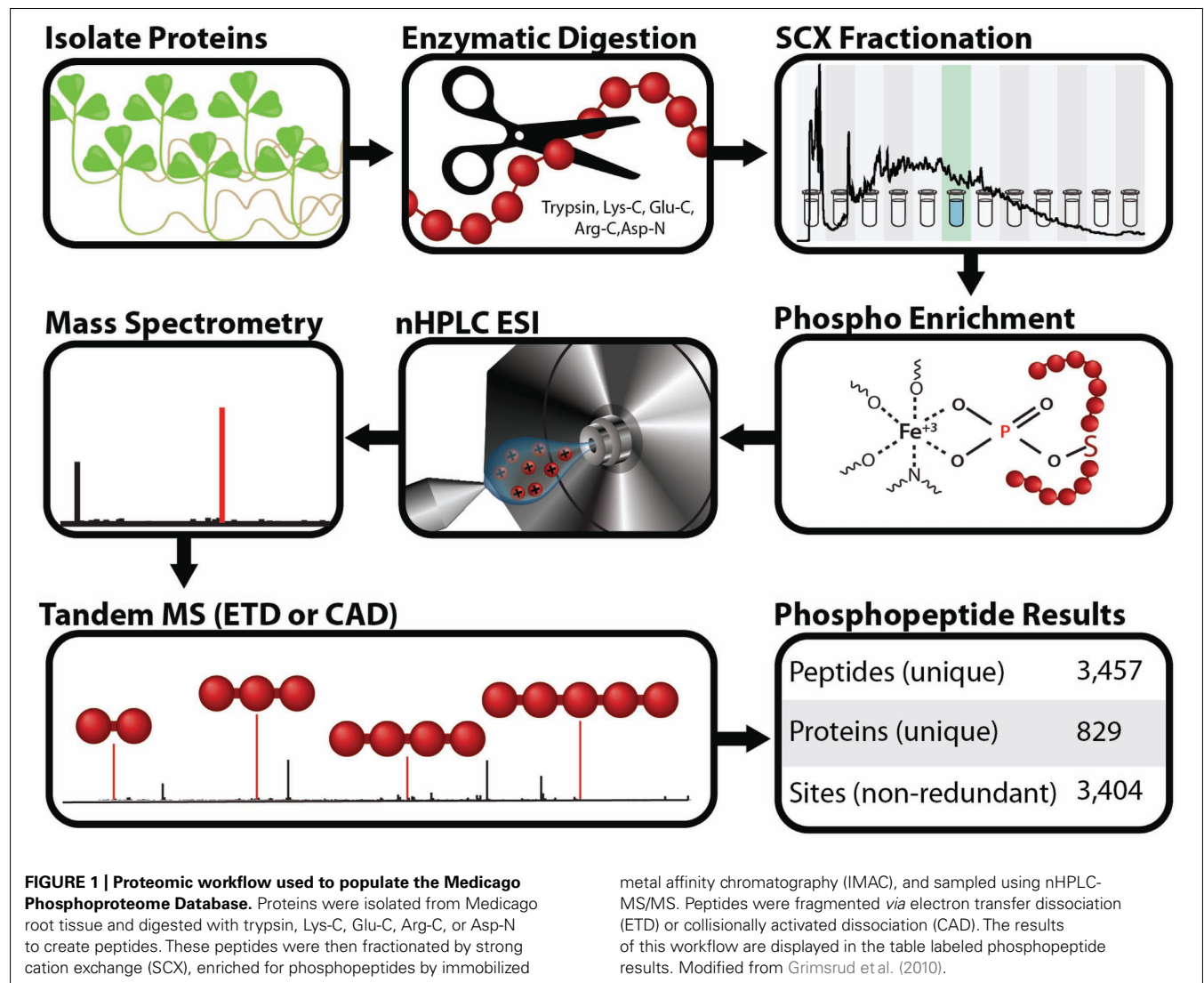
MASS SPECTROMETRY ANALYSIS OF THE MEDICAGO PHOSPHOPROTEOME

The MPPD was populated using the phosphoproteomic workflow presented in **Figure 1**. Proteins were isolated from Medicago plant tissue and digested with trypsin, Lys-C, Glu-C, Arg-C, or Asp-N. To reduce sample complexity, resulting peptides were then fractionated by strong cation exchange chromatography. Phosphopeptides were then enriched by IMAC and analyzed using

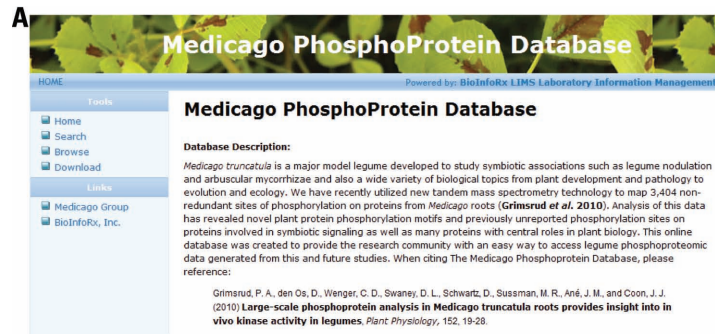
an electron transfer dissociation (ETD)-enabled LTQ Orbitrap mass spectrometer (Thermo-Fisher). To increase proteome coverage, both collisionally activated dissociation (CAD) and ETD (Syka et al., 2004) were used for peptide fragmentation. Spectra were searched against a Medicago protein database using the Open Mass Spectrometry Search Algorithm (OMSSA; Geer et al., 2004). Identifications were then filtered to 1% FDR at both the peptide and protein level. This analysis produced 3,457 unique phosphopeptides, 829 unique proteins, and 3,404 non-redundant sites of phosphorylation (Grimsrud et al., 2010). The entirety of this data is contained within the MPPD and is freely available for download.

OVERVIEW OF THE MEDICAGO PHOSPHOPROTEIN DATABASE

The MPPD (<http://phospho.medicago.wisc.edu>) is a web-based resource that allows users to search for a particular protein of interest, BLAST (Basic Local Alignment Search Tool; Altschul et al., 1990) a protein sequence, browse the entire phosphoproteomic



A



B

Search Medicago PhosphoProtein Database

Search Terms:

Description (e.g. *ABD32373*, *kinase*, *binding*)
 BLAST Protein Sequence (E-value threshold:
 Amino Acid Sequence / Phosphorylation Motif (e.g. *SLxL*, *cys sensitive*)

Tip: Use x (single) or * (multiple) as wild card to match amino acids. Lower-case letters indicate sites of phosphorylation.

Note: You can enter multiple terms by putting one term per row

Group: 1 Proteins: 2 Peptides: 6 P Score: 2.94E-33

Description: IMG1Medtr5g027010.1 interacting protein of DMI3, putative chr05_pseudomolecule_IMGAG_V3 10960357-10964849 E EGN_Med90450 20090702 (MS3.0_proteins_20090702_NAMED)

Coverage 4.6200
Redundancy 0.0800

Sequence:	1	11	21	31	41	50
MEGRGFSGLY	KNSSEELFLK	TVNESPIMP	VPTMMLGFK	EVSQHFRLE		50
EELFKRWLTN	DQEARPSILG	YNSSSMGLNS	RLSKRISTEI	ANMSNQOHIG		100
VASEGRNNDK	SCLQNNFLAN	DVSSDFNFFI	RDPVDRELQS	SNLFLAKAWF		150
ITDQRMTSR	SEELARRYTE	MQNSGAPGGL	DSMFVPEID	TNTRKELAN		200
FNQFDLQMC	ELPQRQTFM	SPNSGSETF	NHQLVQVQR	VSSCVSLKGS		250
TLQRKLECG	VEREAEDGL	NEIFCRPEL	FQSAFNEES	WNCQKLVNVO		300
GDFTDQVNDP	GVMTLEGT	NFVLDGFANQ	TNQLQGRAS	GEPSQSESA		350
AAPVISSGLD	ACEGPNNSQ	TLGDSWQVQ	GESTQNKVRG	VREQIMDNLK		400
DDRRKSLER	YGSVTSVSD	GRMDNPKRR	VERSRMAEA	KERNLTPITP		450
SDMQAILKRC	ENLEKEVRSI	KLNSLPMNRK	DSEQTKQIED	LQKQEDLAD		500
ERELEETE	RILETGKI					

Green letters = localized sites of phosphorylation.
Red letters = non-localized sites from ambiguous positional isomers.

C

Search Results

No.	E Value	P Value	Charge State	Theoretical Neutral Mass (Da)	Experimental Neutral Mass (Da)	Precursor Mass Error (ppm)	Matching Fragments	Total Fragments	All Phosphorylations Localized?
1	0.00760387	9.08E-08	3	1097.5343	1097.5393	4.5519	13	14	YES
2	0.0714818	8.54E-07	3	1097.5343	1097.5397	4.9689	13	14	YES

Dissociation Method	Phosphopeptide Sequence	Reference
ETD	SRSELRR	Grimsrud et al., <i>Plant Physiology</i> . 152: 19-28, 2010.
ETD	SRSELRR	Grimsrud et al., <i>Plant Physiology</i> . 152: 19-28, 2010.

FIGURE 2 | Homepage and search page of the Medicago PhosphoProtein Database. (A) The MPPD homepage contains a description of the database as well as instructions on how to search the database. **(B)** Search result for the protein, Interacting Protein of DMI3 (IPD3). A protein description search returned results listing the protein description, protein coverage, sequence redundancy, number of protein groups, number of proteins and peptides within a protein group, and the protein group *p*-score. Modified amino acids are displayed in lowercase with green and red letters representing localized

and unlocalized phosphosites, respectively. **(C)** Selecting amino acid s161 displays information related to the mass spectrometric identification of the peptide used to identify the given phosphosite. This page displays the *e*-value and *p*-value assigned by OMSSA, the precursor charge state, precursor theoretical mass, precursor experimental mass, precursor mass error, number of matching fragments, number of possible fragments, whether all phosphosites were localized, the dissociation method, the phosphopeptide sequence, and the reference in which the phosphosite was described.

database, and download the mass spectrometric data derived from a phosphoproteomic analysis of Medicago. The MPPD home page displays a description of the database and provides a reference to the publication from which the data was derived (Figure 2A). The homepage also provides detailed instructions for searching the database as well as an explanation of the various options for handling the data (i.e., search, browse, download). Here, we discuss each of these features and provide a guide to navigating the MPPD.

SEARCHING THE MEDICAGO PHOSPHOPROTEIN DATABASE

The MPPD allows users to search the database using three different criteria: (1) protein description, (2) BLAST protein sequence, and (3) peptide sequence/phosphorylation motif. For all search types multiple queries can be executed at one time. The search for protein description allows users to query the database using gene accession number, protein name, or any word contained in the protein description. Figure 2B displays the search results resulting from a search for Interacting Protein of DMI3 (IPD3). IPD3 interacts with and is phosphorylated by DMI3 and is a key regulator of legume–rhizobia symbiotic signaling (Messinese et al., 2007; Horvath et al., 2011). MPPD serves as a repository which contains key information about the phosphorylation sites of several proteins involved in various developmental processes in Medicago. Search results for each query will display the protein description, protein coverage, and sequence redundancy (i.e., the percent of the protein sequenced by multiple peptides). When analyzing shotgun proteomic data, short peptides may match to many proteins; this has been dubbed as the protein inference problem (Nesvizhskii and Aebersold, 2005). To ensure accurate

reporting of the number of protein identifications, proteins with shared peptides are grouped together in protein groups. The light blue bar in Figure 2A contains information regarding the protein group for IPD3. In this case, there was one protein group that contains two separate protein entries from the Medicago database. This group contained six peptides and received a *p*-score of 2.94e-33. Briefly, the *p*-score is calculated by multiplying the *p*-values of all peptides contained within the protein group and is used to conduct false discovery rate analysis at the protein level. In addition, the complete protein sequence is displayed with identified sites of phosphorylation appearing as lower case letters colored green or red with the colors representing localized and unlocalized sites of phosphorylation, respectively. Amino acids highlighted in green (localized sites of phosphorylation) indicate that the mass spectrometric data provided clear evidence for phosphorylation of the given residue. Amino acids highlighted in red indicate that mass spectrometric analysis had determined the presence of phosphorylation, but the site of phosphorylation could not be assigned to a specific residue with high confidence. In the case of unlocalized phosphorylation sites, all possible sites are highlighted in red.

Selecting an amino acid, such as s161, will display the characteristics of the spectra used to identify the given phosphorylation site (Figure 2C). This information includes the *e*-value and *p*-value assigned by OMSSA, precursor charge state, precursor theoretical neutral mass, precursor experimental neutral mass, precursor mass error, number of matching fragments, number of total fragments, localization state of all phosphosites, the MS/MS dissociation method, phosphopeptide sequence, and reference to the paper which describes this identification. The *e*-value provides a

rough characterization of the spectral quality as a lower e -value indicates the given MS/MS spectrum is likely associated with the phosphopeptide in question. The precursor mass error represents the deviation of the experimental mass from the theoretical mass. High accuracy mass analyzers (e.g., Orbitrap, FT-ICR, and TOF) enable greater specificity when searching protein databases and generally a precursor mass error less than 10 ppm will increase the number of peptide identifications (McAlister et al., 2010). The dissociation method is important to note, as certain peptides are more amenable to either collisional or electron-based fragmentation techniques (Swaney et al., 2008). To repeat the identification of phosphosites contained within the MPPD, it is important that the same fragmentation conditions be used. All possible phosphopeptide sequences are also displayed on this page. If the phosphosite was localized, only one sequence will appear per line with the modified residue appearing in lowercase. In the case of unlocalized phosphosites, all possible phosphopeptides will be displayed with each possible phosphosite appearing in lower case.

Assessing sequence homology across species enables researchers to infer function of unknown proteins and make connections to research published on similar proteins in different systems. The MPPD allows users to conduct a BLAST search to query homology of proteins within the Medicago protein database with a user-provided protein sequence. This search returns all of the same protein attributes discussed above, but it adds a link to a flat file containing all alignments above the user-provided threshold. This flat file contains a list of the proteins matched, e -value score of each match, and a visual representation of the alignment.

Phosphorylation motifs can help to elucidate potential kinases which alter the phosphorylation state of a particular amino acid (Schwartz and Gygi, 2005). The MPPD allows users to query the database by entering a short amino acid sequence or phosphorylation motif (e.g., SsLxT). Here, “x” is used as a single wild card amino acid while “*” is used for multiple wild card amino acids. This search is also case-sensitive, as lowercase letters signify a phosphorylated residue. The result of this search is the same for a protein description search (Figure 2B), but the results contain proteins that have the specified amino acid sequence or phosphorylation motif. This feature enables users who are interested in a particular kinase to determine if this protein kinase contains the motif of interest and if a phosphosite at this location was identified in our mass spectrometry analysis.

BROWSING AND DOWNLOADING DATA

The MPPD allows users to browse all protein groups contained within the database by selecting the “Browse” option in the tools menu. As described above, protein grouping occurs when multiple entries within the protein database share an identified peptide. To explain peptide identifications with the fewest number of proteins, these protein identifications are placed into one group. When the user selects the browse task, an online table appears, listing each identified protein group, including the associated protein number, number of proteins in the group, number of peptide identifications in the protein groups, the p -score for each protein, and the protein description of the longest protein in the protein group.

As discussed above, the p -score is calculated by multiplying the OMSSA p -value for each peptide in the group and is used to calculate the false discovery rate for the given protein. Users can browse protein entries by changing pages, but to access a protein entry, users must copy the protein description and use the search tool.

To download text files containing all protein group, protein, and peptide identification data, users can select the download option on the tools menu. To download data right click on “Group Information,” “Protein Information,” or “Peptide Information” and select “Save As.” Headers associated with each file are listed on the download page, as the text files do not contain headers.

CONCLUSION

As large-scale phosphoproteomic analysis of plant tissues continues to become more prevalent, tools are needed to enable facile access to data (Jayaraman et al., 2012). Numerous databases for proteomic information in plants exist, including databases focused on sub-cellular fractions (AMPP, Kruff et al., 2001; AraPerox, Reumann et al., 2004; AtNoPDB, Brown et al., 2005; SUBA, Heazlewood et al., 2007; PIProt, Kleffmann et al., 2006; AT_Chloro, Ferro et al., 2010), single species (SpruceDB, Lippert et al., 2009; Soybean Proteome Database, Sakata et al., 2009; PhosPhAt, Durek et al., 2010), multiple species (P3DB, Gao et al., 2009; PPDB, Sun et al., 2009), 2-D gel mapping (GelMap, Rode et al., 2011), and spectral data from mass spectrometry experiments (ProMex, Hummel et al., 2007). Here we highlight a publicly available web portal for Medicago phosphoproteins and phosphopeptides. This online database enables researchers to search for proteins of interest, BLAST for homologous proteins, search for phosphorylation motifs, browse, and download the data. This central repository contains all of the information for researchers to connect phosphosites identified in Medicago to other legumes, providing an invaluable resource for future studies pertaining to legume biology and, particularly, to the legume–rhizobia symbiosis. In addition to the data presented here, the Wisconsin Medicago Group (University of Wisconsin, Madison) has been very active in continuing phosphoproteomic characterization of Medicago in response to symbiotic signals. In particular, we are pursuing quantitative measurements of phosphorylation dynamics within Medicago in response to various symbiotic stimuli. As we report our results, we will continue to build database tools and enable researchers to connect their studies with our results. These tools will include the ability to query quantitative information of phosphorylation state alterations, allowing researchers to determine if proteins of interest are involved in a symbiotic cascade. In addition, future database applications will offer the opportunity for other researchers to upload their own data, creating a large, centralized source for all phosphoproteomic data relating to legumes.

ACKNOWLEDGMENTS

This work was supported by a grant from the National Science Foundation (NSF#0701846) to Michael R. Sussman, Joshua J. Coon, and Jean-Michel Ané. Christopher M. Rose was funded by an NSF Graduate Research Fellowship and NIH Traineeship (T32GM008505).

REFERENCES

- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410.
- Ané, J. M., Zhu, H., and Frugoli, J. (2008). Recent advances in *Medicago truncatula* genomics. *Int. J. Plant Genomics* 2008, 256597.
- Arrighi, J. F., Barre, A., Ben Amor, B., Bersoult, A., Soriano, L. C., Mirabella, R., de Carvalho-Niebel, E., Journet, E. P., Gherardi, M., Hugué, T., Geurts, R., Dénarié, J., Rouge, P., and Gough, C. (2006). The *Medicago truncatula* lysin [corrected] motif-receptor-like kinase gene family includes *NFP* and new nodule-expressed genes. *Plant Physiol.* 142, 265–279.
- Ben Amor, B., Shaw, S. L., Oldroyd, G. E. D., Maillet, F., Penmetsa, R. V., Cook, D., Long, S. R., Dénarié, J., and Gough, C. (2003). The *NFP* locus of *Medicago truncatula* controls an early step of nod factor signal transduction upstream of a rapid calcium flux and root hair deformation. *Plant J.* 34, 495–506.
- Bonfante, P., and Genre, A. (2010). Mechanisms underlying beneficial plant–fungus interactions in mycorrhizal symbiosis. *Nat. Commun.* 1, 48.
- Brelles-Marino, G. Q., and Ané, J. M. (2008). “Nod factors and the molecular dialogue in the rhizobia–legume interaction,” in *Nitrogen Fixation Research Progress*, ed. G. N. Couto (New York: Nova Science Publishers, Inc.), 173–227.
- Brown, J. W., Shaw, P. J., Shaw, P., and Marshall, D. F. (2005). *Arabidopsis* nucleolar protein database (AtNoPDB). *Nucleic Acids Res.* 33, D633–D636.
- Chen, T., Zhu, H., Ke, D., Cai, K., Wang, C., Gou, H., Hong, Z., and Zhang, Z. (2012). A MAP kinase interacts with SymRK and regulates nodule organogenesis in *Lotus japonicus*. *Plant Cell* 24, 823–838.
- Colditz, F., and Braun, H. P. (2010). *Medicago truncatula* proteomics. *J. Proteomics* 73, 1974–1985.
- Cook, D. R. (1999). *Medicago truncatula* – a model in the making! *Curr. Opin. Plant Biol.* 2, 301–304.
- Dénarié, J., Debelle, F., and Promé, J. C. (1996). Rhizobium lipochitooligosaccharide nodulation factors: signaling molecules mediating recognition and morphogenesis. *Annu. Rev. Biochem.* 65, 503–535.
- Durek, P., Schmidt, R., Heazlewood, J. L., Jones, A., MacLean, D., Nagel, A., Kersten, B., and Schulze, W. X. (2010). PhosPhAt: the *Arabidopsis thaliana* phosphorylation site database. An update. *Nucleic Acids Res.* 38, D828–D834.
- Ehrhardt, D. W., Wais, R., and Long, S. R. (1996). Calcium spiking in plant root hairs responding to rhizobium nodulation signals. *Cell* 85, 673–681.
- Endre, G., Kereszt, A., Kevei, Z., Mihacea, S., Kaló, P., and Kiss, G. B. (2002). A receptor kinase gene regulating symbiotic nodule development. *Nature* 417, 962–966.
- Ferro, M., Brugiére, S., Salvi, D., Seigneurin-Berny, D., Court, M., Moyet, L., Ramus, C., Miras, S., Mellal, M., Le Gall, S., Kieffer-Jaquinod, S., Bruley, C., Garin, J., Joyard, J., Masselon, C., and Rolland, N. (2010). AT_CHLORO, a comprehensive chloroplast proteome database with subplastidial localization and curated information on envelope proteins. *Mol. Cell. Proteomics* 9, 1063–1084.
- Gao, J., Agrawal, G. K., Thelen, J. J., and Xu, D. (2009). P3DB: a plant protein phosphorylation database. *Nucleic Acids Res.* 37, D960–D962.
- Geer, L. Y., Markey, S. P., Kowalak, J. A., Wagner, L., Xu, M., Maynard, D. M., Yang, X., Shi, W., and Bryant, S. H. (2004). Open mass spectrometry search algorithm. *J. Proteome Res.* 3, 958–964.
- Graham, P. H., and Vance, C. P. (2003). Legumes: importance and constraints to greater use. *Plant Physiol.* 131, 872–877.
- Grimsrud, P. A., den Os, D., Wenger, C. D., Swaney, D. L., Schwartz, D., Sussman, M. R., Ane, J. M., and Coon, J. J. (2010). Large-scale phosphoprotein analysis in *Medicago truncatula* roots provides insight into in vivo kinase activity in legumes. *Plant Physiol.* 152, 19–28.
- 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.
- Horvath, B., Yeun, L. H., Domonkos, A., Halasz, G., Gobato, E., Ayaydin, F., Miro, K., Hirsch, S., Sun, J., Tadege, M., Ratet, P., Mysore, K. S., Ane, J. M., Oldroyd, G. E., and Kalo, P. (2011). *Medicago truncatula* IPD3 is a member of the common symbiotic signaling pathway required for rhizobial and mycorrhizal symbioses. *Mol. Plant Microbe Interact.* 24, 1345–1358.
- Huber, S. C. (2007). Exploring the role of protein phosphorylation in plants: from signalling to metabolism. *Biochem. Soc. Trans.* 35, 28–32.
- Hummel, J., Niemann, M., Wienkoop, S., Schulze, W., Steinhauser, D., Selbig, J., Walther, D., and Weckwerth, W. (2007). ProMEX: a mass spectral reference database for proteins and protein phosphorylation sites. *BMC Bioinformatics* 8, 216. doi: 10.1186/1471-2105-8-216
- Jayaraman, D., Forshey, K. L., Grimsrud, P. A., and Ané, J. M. (2012). Leveraging proteomics to understand plant–microbe interactions. *Front. Plant Sci.* 3:44. doi: 10.3389/fpls.2012.00044
- Jones, K. M., Kobayashi, H., Davies, B. W., Taga, M. E., and Walker, G. C. (2007). How rhizobial symbionts invade plants: the Sinorhizobium–*Medicago* model. *Nat. Rev. Microbiol.* 5, 619–633.
- Kevei, Z., Lougnon, G., Mergaert, P., Horvath, G. V., Kereszt, A., Jayaraman, D., Zaman, N., Marcel, F., Regulski, K., Kiss, G. B., Kondorosi, A., Endre, G., Kondorosi, E., and Ané, J. M. (2007). 3-hydroxy-3-methylglutaryl coenzyme A reductase 1 interacts with NORK and is crucial for nodulation in *Medicago truncatula*. *Plant Cell* 19, 3974–3989.
- 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.
- 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, 1694–1710.
- Laugesen, S., Messinese, E., Hem, S., Pichereaux, C., Grat, S., Ranjeva, R., Rossignol, M., and Bono, J. J. (2006). Phosphoproteins analysis in plants: a proteomic approach. *Phytochemistry* 67, 2208–2214.
- Lefebvre, B., Timmers, T., Mbengue, M., Moreau, S., Herve, C., Toth, K., Bittencourt-Silvestre, J., Klaus, D., Deslandes, L., Godiard, L., Murray, J. D., Udvardi, M. K., Raffaele, S., Mongrand, S., Cullimore, J., Gamas, P., Niebel, A., and Ott, T. (2010). A remorin protein interacts with symbiotic receptors and regulates bacterial infection. *Proc. Natl. Acad. Sci. U.S.A.* 107, 2343–2348.
- Lévy, J., Bres, C., Geurts, R., Chalhoub, B., Kulikova, O., Duc, G., Journet, E. P., Ané, J. M., Lauber, E., Bisseling, T., Dénarié, J., Rosenberg, C., and Debelle, F. (2004). A putative Ca²⁺ and calmodulin-dependent protein kinase required for bacterial and fungal symbioses. *Science* 303, 1361–1364.
- Lima, L., Seabra, A., Melo, P., Cullimore, J., and Carvalho, H. (2006). Phosphorylation and subsequent interaction with 14-3-3 proteins regulate plastid glutamine synthetase in *Medicago truncatula*. *Planta* 223, 558–567.
- Limpens, E., Mirabella, R., Fedorova, E., Franken, C., Franssen, H., Bisseling, T., and Geurts, R. (2005). Formation of organelle-like N₂-fixing symbiosomes in legume root nodules is controlled by DMI2. *Proc. Natl. Acad. Sci. U.S.A.* 102, 10375–10380.
- Lippert, D., Yuen, M., and Bohlmann, J. (2009). Spruce proteome DB: a resource for conifer proteomics research. *Tree Genet. Genomes* 5, 723–727.
- Mbengue, M., Camut, S., de Carvalho-Niebel, F., Deslandes, L., Froidure, S., Klaus-Heisen, D., Moreau, S., Rivas, S., Timmers, T., Herve, C., Cullimore, J., and Lefebvre, B. (2010). The *Medicago truncatula* E3 ubiquitin ligase PUB1 interacts with the LYK3 symbiotic receptor and negatively regulates infection and nodulation. *Plant Cell* 22, 3474–3488.
- McAlister, G. C., Phanstiel, D., Wenger, C. D., Lee, M. V., and Coon, J. J. (2010). Analysis of tandem mass spectra by FTMS for improved large-scale proteomics with superior protein quantification. *Anal. Chem.* 82, 316–322.
- Messinese, E., Mun, J. H., Yeun, L. H., Jayaraman, D., Rouge, P., Barre, A., Lougnon, G., Schornack, S., Bono, J. J., Cook, D. R., and Ané, J. M. (2007). A novel nuclear protein interacts with the symbiotic DMI3 calcium- and calmodulin-dependent protein kinase of *Medicago truncatula*. *Mol. Plant Microbe Interact.* 20, 912–921.
- Nesvizhskii, A. I., and Aebersold, R. (2005). Interpretation of shotgun proteomic data: the protein inference problem. *Mol. Cell. Proteomics* 4, 1419–1440.
- Peck, S. C. (2006). Phosphoproteomics in *Arabidopsis*: moving from empirical to predictive science. *J. Exp. Bot.* 57, 1523–1527.
- Peiter, E., Sun, J., Heckmann, A. B., Venkateshwaran, M., Riely, B. K., Otegui, M. S., Edwards, A., Freshour, G., Hahn, M. G., Cook, D. R., Sanders, D., Oldroyd, G. E., Downie, J. A., and Ané, J. M. (2007). The *Medicago truncatula* DMI1 protein modulates cytosolic calcium signaling. *Plant Physiol.* 145, 192–203.
- 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.
- Riely, B. K., Ane, J. M., Penmetsa, R. V., and Cook, D. R. (2004). Genetic and genomic analysis in model legumes bring nod-factor signaling to center stage. *Curr. Opin. Plant Biol.* 7, 408–413.

- Riely, B. K., Mun, J. H., and Ane, J. M. (2006). Unravelling the molecular basis for symbiotic signal transduction in legumes. *Mol. Plant Pathol.* 7, 197–207.
- Rode, C., Senkler, M., Klodmann, J., Winkelmann, T., and Braun, H. P. (2011). GelMap – a novel software tool for building and presenting proteome reference maps. *J. Proteomics* 74, 2214–2219.
- Ruiz-Lozano, J. M., Azcon, R., and Gomez, M. (1995). Effects of arbuscular-mycorrhizal glomus species on drought tolerance: physiological and nutritional plant responses. *Appl. Environ. Microbiol.* 61, 456–460.
- Sakata, K., Ohyanagi, H., Nobori, H., Nakamura, T., Hashiguchi, A., Nanjo, Y., Mikami, Y., Yunokawa, H., and Komatsu, S. (2009). Soybean proteome database: a data resource for plant differential omics. *J. Proteome Res.* 8, 3539–3548.
- Schutzenhubel, A., and Polle, A. (2002). Plant responses to abiotic stresses: heavy metal-induced oxidative stress and protection by mycorrhization. *J. Exp. Bot.* 53, 1351–1365.
- Schwartz, D., and Gygi, S. P. (2005). An iterative statistical approach to the identification of protein phosphorylation motifs from large-scale data sets. *Nat. Biotechnol.* 23, 1391–1398.
- Shimoda, Y., Han, L., Yamazaki, T., Suzuki, R., Hayashi, M., and Imaizumi-Anraku, H. (2012). Rhizobial and fungal symbioses show different requirements for calmodulin binding to calcium calmodulin-dependent protein kinase in *Lotus japonicus*. *Plant Cell* 24, 304–321.
- Smit, P., Limpens, E., Geurts, R., Fedorova, E., Dolgikh, E., Gough, C., and Bisseling, T. (2007). *Medicago* LYK3, an entry receptor in rhizobial nodulation factor signaling. *Plant Physiol.* 145, 183–191.
- Sun, Q., Zybailov, B., Majeran, W., Friso, G., Olinares, P. D., and van Wijk, K. J. (2009). PPDB, the plant proteomics database at cornell. *Nucleic Acids Res.* 37, D969–D974.
- Swaney, D. L., McAlister, G. C., and Coon, J. J. (2008). Decision tree-driven tandem mass spectrometry for shotgun proteomics. *Nat. Methods* 5, 959–964.
- Syka, J. E., Coon, J. J., Schroeder, M. J., Shabanowitz, J., and Hunt, D. F. (2004). Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry. *Proc. Natl. Acad. Sci. U.S.A.* 101, 9528–9533.
- Thoquet, P., Gherardi, M., Journet, E. P., Kereszt, A., Ane, J. M., Prospero, J. M., and Hugué, T. (2002). The molecular genetic linkage map of the model legume *Medicago truncatula*: an essential tool for comparative legume genomics and the isolation of agronomically important genes. *BMC Plant Biol.* 2, 1. doi: 10.1186/1471-2229-2-1
- Venkateshwaran, M., and Ané, J. M. (2011). “Legumes and nitrogen fixation: physiological, molecular, evolutionary perspective and applications,” in *The Molecular Basis of Nutrient Use Efficiency in Crops*, eds M. Hawkesford and P. Barraclough (Hoboken, NJ: John Wiley & Sons), 457–489.
- Wienkoop, S., Larrainzar, E., Glinzski, M., Gonzalez, E. M., Arrese-Igor, C., and Weckwerth, W. (2008). Absolute quantification of *Medicago truncatula* sucrose synthase isoforms and N-metabolism enzymes in symbiotic root nodules and the detection of novel nodule phosphoproteins by mass spectrometry. *J. Exp. Bot.* 59, 3307–3315.
- Young, N. D., Debelle, F., Oldroyd, G. E., Geurts, R., Cannon, S. B., Udvardi, M. K., Benedito, V. A., Mayer, K. F., Gouzy, J., Schoof, H., Van de Peer, Y., Proost, S., Cook, D. R., Meyers, B. C., Spannagl, M., Cheung, F., De Mita, S., Krishnakumar, V., Gundlach, H., Zhou, S., Mudge, J., Bharti, A. K., Murray, J. D., Naoumkina, M. A., Rosen, B., Silverstein, K. A., Tang, H., Rombauts, S., Zhao, P. X., Zhou, P., Barbe, V., Bardou, P., Bechner, M., Bellec, A., Berger, A., Berges, H., Bidwell, S., Bisseling, T., Choise, N., Couloux, A., Denny, R., Deshpande, S., Dai, X., Doyle, J. J., Dudez, A. M., Farmer, A. D., Fouteau, S., Franken, C., Gibelin, C., Gish, J., Goldstein, S., Gonzalez, A. J., Green, P. J., Hallab, A., Hartog, M., Hua, A., Humphray, S. J., Jeong, D. H., Jing, Y., Jocker, A., Kenton, S. M., Kim, D. J., Klee, K., Lai, H., Lang, C., Lin, S., Macmill, S. L., Magdelenat, G., Matthews, L., McCarrison, J., Monaghan, E. L., Mun, J. H., Najjar, F. Z., Nicholson, C., Noiro, C., O’Bleness, M., Paule, C. R., Poulain, J., Prion, F., Qin, B., Qu, C., Retzel, E. F., Riddle, C., Sallet, E., Samain, S., Samson, N., Sanders, I., Saurat, O., Scarpelli, C., Schiex, T., Segurens, B., Severin, A. J., Sherrier, D. J., Shi, R., Sims, S.,
- Singer, S. R., Sinharoy, S., Sterck, L., Viollet, A., Wang, B. B., Wang, K., Wang, M., Wang, X., Warfsmann, J., Weissenbach, J., White, D. D., White, J. D., Wiley, G. B., Wincker, P., Xing, Y., Yang, L., Yao, Z., Ying, F., Zhai, J., Zhou, L., Zuber, A., Denarie, J., Dixon, R. A., May, G. D., Schwartz, D. C., Rogers, J., Quétier, F., Town, C. D., and Roe, B. A. (2011). The *Medicago* genome provides insight into the evolution of rhizobial symbioses. *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: 23 March 2012; accepted: 21 May 2012; published online: 11 June 2012.

Citation: Rose CM, Venkateshwaran M, Grimsrud PA, Westphall MS, Sussman MR, Coon JJ and Ané J-M (2012) *Medicago PhosphoProtein Database: a repository for Medicago truncatula phosphoprotein data*. *Front. Plant Sci.* 3:122. doi: 10.3389/fpls.2012.00122

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

Copyright © 2012 Rose, Venkateshwaran, Grimsrud, Westphall, Sussman, Coon and Ané. This is an open-access article distributed under the terms of the Creative Commons Attribution Non Commercial License, which permits non-commercial use, distribution, and reproduction in other forums, provided the original authors and source are credited.