



The conundrum of discordant protein and mRNA expression. Are plants special?

Isabel Cristina Vélez-Bermúdez and Wolfgang Schmidt*

Integrative Root Development Lab, Institute of Plant and Microbial Biology, Academia Sinica, Taipei, Taiwan
*Correspondence: wosh@gate.sinica.edu.tw

Edited by:

Sabine Lüthje, University of Hamburg, Germany

Reviewed by:

Katja Baerenfaller, Swiss Federal Institute of Technology Zurich, Switzerland
Christine Vogel, New York University, USA

Keywords: omics, systems biology, post-transcriptional regulation, ribosomes, alternative splicing, gene expression

Rapid progress in transcriptional and proteomic profiling methodologies, such as RNA-sequencing and high-resolution tandem mass spectrometry in combination with nanoflow chromatography, allows for a more accurate comparisons of disparate omics data sets at high resolution. Despite the increased fidelity of such surveys, several studies in mammals, yeast and plants have shown that transcript levels are not always a good proxy for protein abundance. While to some extent the modest concordance observed in studies across platforms may result from fundamentally different protocols that are used for the detection of proteins and mRNAs or caused by differences in the sensitivity of detection and data analysis (“false discordance”), it is becoming obvious that transcript/protein discordance is largely of biological origin (“true discordance”) and represents a critical layer of regulatory processes at the post-transcriptional level that is often neglected.

THE MORE THE BETTER? THE CONCEPT OF POTENTIATION

Gene activity is the result of complex dynamics between the transcription rate of the DNA template, the stability of the mRNA, the translation efficiency of the transcript, and degradation of the protein. Interestingly, the concordance between the abundance of orthologous proteins among related species is higher than that between proteins and their cognate mRNAs within a species (Kwon et al., 2013). This suggests that a certain set point of protein abundance is established to ensure optimal function, although this set point is

not necessarily controlled at the transcriptional level. Rather, protein synthesis and feedback inhibition of the synthesis rate appear to dictate protein expression (Kristensen et al., 2013).

Contrary to expectations, direct coupling of transcription and translation, which occurs in prokaryotes, results in lower mRNA/protein concordance when compared to eukaryotes where the two processes are spatially separated (Vogel and Marcotte, 2012). Interestingly, concordance values are highest in single cell eukaryotes and lowest in humans (overview in De Sousa Abreu et al., 2009), indicating that cellular diversity is a large contributor to the difference between mRNA and protein abundance. In particular, decreased mRNA abundance is not conclusive. In a comprehensive survey of changes in protein and transcript profiles in *Arabidopsis* roots in response to phosphate deficiency, we observed a complete lack of correlation between down-regulated transcripts and the amount of their corresponding proteins, whereas for induced genes changes in the levels of mRNAs and proteins were reasonably well correlated (Lan et al., 2012). A similar observation was reported for yeast cells subjected to salt treatment, in which transcript reduction produced only minor changes in the abundance of the corresponding proteins (Lee et al., 2011). Strong induction of transcription is a much better predictor of changes in protein levels than decreased expression. This might be related to the importance of stress-associated proteins that are required to recalibrate cellular metabolism. In fact,

highly and lowly transcribed genes differ in their translational fitness, i.e., their association with polysomes (Preiss et al., 2003). Relatively few studies compare mRNA and protein expression in plants using (mainly *Arabidopsis* and maize), but the general picture that emerges from these studies is a variable correlation of protein–transcript pairs (Baerenfaller et al., 2012; Walley et al., 2013; Ponnala et al., 2014) that is comparable to what has been reported for other organisms. Thus, although technical improvements and the use of techniques measuring *in vivo* translation such as ribosome profiling (Ingolia, 2014) will correct the observed correlations, it appears that at least a substantial proportion of this lack of concordance is of biological origin and thus reflect post-transcriptional regulation rather than technical constrains.

Preferential translation of transcripts derived from highly induced genes, referred to as homo-directional co-regulatory response or potentiation (Preiss et al., 2003), amplifies transcriptional changes at the translational level, leading to a fast and robust change in gene activity. In plants, translation efficiency can change dramatically in response to abiotic stress, leading to a massive bias in the pool of mRNAs that are actively translated (Mustroph et al., 2009; Juntawong et al., 2014). Potentiation of gene expression can be encoded in both DNA and mRNA; translation of *Arabidopsis* mRNAs upon light treatment is dependent on the presence of the sequence motifs TAGGGTTT or AAAACCCCT in their 5' UTR (Liu et al., 2012). These elements are also required

for transcription (Tremousaygue et al., 1999; Tatematsu et al., 2005), suggesting that such co-regulatory responses have evolved to switch the expression of genes into a fast-forward mode when the demand for the encoded proteins is high. One may speculate that, while gene induction is amplified by preferred translation of transcripts derived from these genes to secure fast acclimation, down-regulation of gene expression is a weaker signal and the full scale of regulatory post-translational mechanisms results in pronounced discordance between transcripts and proteins.

STOP MAKING SENSE: PRODUCTION OF NON-FUNCTIONAL TRANSCRIPTS TO TUNE PROTEIN ABUNDANCE

In principle, the mechanisms that control RNA turnover, translation, and protein stability are similar among eukaryotes. However, lacking behavioral recourses in coping with unfavorable conditions, plants can adjust their developmental, metabolic and physiological programs in a much broader way than, for example, mammals. The phenotypic plasticity of plants results from sophisticated sensing and signaling circuits that integrate disparate environmental cues and modulate gene activity to adjust the phenotypic readout to the prevailing conditions.

The ability of plants to adapt rapidly to a wide range of conditions is reflected by numerous plant-specific peculiarities in the control of gene activity. This high level of transcriptional regulation may correspond to an equally pronounced abundance of post-transcriptional regulatory processes. An interesting example of a process that is seemingly similar in plants and animals, but that has significantly distinct consequences, is the splicing of pre-mRNAs during transcription. In human cells, the vast majority of intron-containing genes (~95%) are alternatively spliced, leading to the generation of multiple, distinct mRNAs from a single gene. The predominant form of alternative splicing in mammals is the exclusion of cassette exons together with its flanking introns, a process referred to as exon skipping, which leaves the open reading frame uninterrupted (Kornblihtt et al., 2013). Exon skipping is thought to contribute to proteome diversity by producing

protein isoforms that are structurally and functionally distinct. Other forms of alternative splicing, i.e., introns retention or alternative 5' and 3' splice sites, produce transcripts that mostly contain premature stop codons (PTCs), targeting these transcripts for degradation via the nonsense-mediated decay RNA surveillance pathway, a mechanism that prevents the production of truncated proteins by eliminating PTC-containing mRNAs after the pioneer round of translation.

In plants, intron retention and alternative 5' and 3' splice sites comprise the majority of alternative splicing events (Filichkin et al., 2010; Marquez et al., 2012; Li et al., 2013; Wu et al., 2014). Splicing patterns appear to be more complex during acclimation to nutrient deficiencies or light exposure than under constant conditions, resulting in significantly induced intron retention features (Li et al., 2013; Wu et al., 2014). The reasons for the different forms and consequences of alternative splicing in animals and plants are currently unknown. Two plausible scenarios would explain the high number of intron- (and, in most cases, PTC-) containing transcripts in plants. Firstly, non-functional mRNAs could be stored as ribonucleoproteins and processed when needed. This would allow for a fast increase in populations of mature mRNAs that code for proteins that are required upon stress exposure or during development. In an alternative, but not mutually exclusive scenario, non-functional transcripts are produced to tune the abundance of proteins with critical functions in response to a fluctuating environment. A quick shift between the production of functional and non-functional transcripts also aids in rapid re-adjustment of protein levels after the stress is relieved. Production of non-functional transcripts would provide an alternative mechanism for such an adjustment, circumventing changes in transcription rates, which involve recruitment/dismissal of transcription factors and changes in chromatin structure.

Changes in splicing patterns in plants as a means to calibrate protein abundance would imply a feedback mechanism that communicates the demand for a given protein to the splicing machinery, and a switch to adjust the ratio of functional to non-functional transcripts.

Shifting between the production of functional and PTC-containing transcripts could be achieved by post-translational modifications of proteins from the splicing machinery such as serine/arginine (SR)-rich splicing factors to facilitate or repress interaction of the spliceosome with a subset of mRNAs.

Another possible mechanism relies on the presence of information in the differentially retained introns or in the flanking exons, such as *cis*-acting intronic splicing silencers that are differentially recognized by *trans*-acting mRNA-binding proteins. In this scenario, post-translational modifications of mRNA-binding proteins could also modulate the probability of splicing events near a given site. In any case, the production of non-functional mRNA isoforms contributes massively to the apparent transcriptome/proteome discordance in plants, but to a much lesser extent in animals where the vast majority of alternative splicing events yield functional products. Quantitative distinction between functional and non-functional transcripts, presently still a challenging task for genome-wide transcriptional surveys, would most probably change the correlation between transcript and protein levels toward a higher concordance correlation coefficients in plants.

BUILD TO ORDER: DO PLANTS PRODUCE SPECIALIZED RIBOSOMES?

Translation is mediated by ribosomes, intricate molecular machines composed of ribosomal RNA and ribosomal proteins (r-proteins) that translate the genetic code encrypted in the DNA into proteins. Because the function of ribosomes is highly conserved in both prokaryotes and eukaryotes, r-proteins are traditionally classified as housekeeping. However, eukaryotic ribosomes contain more proteins than their bacterial counterparts and possess diverse r-RNA modifications that are not found in prokaryotic ribosomes, indicating more sophisticated molecular functions of eukaryotic ribosomes.

In contrast to animals, in which r-proteins are mostly encoded by a single gene, plant r-proteins are encoded by paralogous families comprising several members that generate diverse, functional proteins. For example, the 81 r-proteins of *Arabidopsis* are encoded by more than

200 genes, with each r-protein family consisting of 2–7 members. This does not only complicate coordinate expression of equiamounts of r-proteins to secure ribosomal function, but also allows for a nearly infinite number of differently composed ribosomes, the heterogeneity of which can be further increased by numerous post-translational modifications. In humans and *Drosophila*, defects in r-protein expression have been associated with diseases (Kongsuwan et al., 1985; Uechi et al., 2001), suggesting functions beyond translation. Similarly, in plants several r-protein mutants are affected in cell division and/or cell expansion resulting in deformed leaves (Rosado et al., 2012), indicating specific, extra-ribosomal functions of some r-proteins in developmental processes. It should be noted that in plants accurate detection of protein concentrations of paralogous proteins by mass spectrometry is rendered difficult as these proteins often have identical sequence parts that cannot be distinguished.

A PLANT-SPECIFIC RIBOSOME CODE?

The large number of r-protein paralogs in plants invites speculation as to whether populations of structurally diverse ribosomes produced during development or in response to environmental signals can capture mRNAs differentially and prioritize the translation of specific subsets of mRNAs. The relative incorporation of r-protein paralogs is altered by growth conditions at the protein level (Hummel et al., 2012), and transcripts encoding r-protein accumulate differentially upon iron and phosphate deficiency (Rodríguez-Celma et al., 2013; Wang et al., 2013), suggesting that the translational machinery is remodeled in response to environmental signals. Heterogeneous ribosomal populations would contribute markedly to discordant changes in transcript and protein profiles. Translatome profiling studies support a regulatory intervention of protein abundance at the translational level (Mustroph et al., 2009; Juntawong et al., 2014). A transcriptomic comparison of steady-state and polysome-bound mRNAs revealed that translational control is independent of mRNA abundance (Liu et al., 2012). The authors of this study concluded that

translational control has a greater effect on gene activity than the high steady state mRNA levels. Dynamic changes in r-protein composition would offer a plausible explanation for the observed differential translational efficiency of mRNAs in response to changing environmental conditions.

CONCLUSIONS

Based on the above considerations, we propose that the uncoupling of transcript and protein abundances in plants is driven by additional mechanisms that are not prominent or not present at all in other organisms. This proposal is supported by the predominant forms of alternative splicing in plants (i.e., intron retention and alternative donor or acceptor splice sites) and a highly dynamic ribosome composition that aids in tuning protein profiles to cellular demands during development or stress. Although ribosomal specificity might also exist in mammals (O’Leary et al., 2013), the large number of r-protein genes suggests that heterogeneity of ribosomes in plants is much more pronounced than in other eukaryotes. Also, recruitment of ribosomes to mRNA may have plant-specific dynamics that affect translation efficiency (Lan and Schmidt, 2011). An underappreciated factor is the cell type-specific variation in splicing patterns (Lan et al., 2013), r-protein composition (Whittle and Krochko, 2009), and other, not yet explored processes that may differ among cell types such as mRNA export, and protein stability, which may contribute to the high mRNA/protein discordance in multicellular organisms. Another factor that has not yet been explored at the whole-genome scale is the impact of microRNAs on translation and thus on mRNA/protein abundance correlation in plants. The points brought up here are not only of academic interest. “True” discordance mirrors an underappreciated regulatory layer for determining the final concentrations of proteins. Attempts to engineer the genetics of crop plants to improve plant performance under stress conditions traditionally aim at the control of gene expression. Knowledge regarding post-transcriptional regulation that contributes to transcript/protein discordance may aid in generating stress-resistant germplasm.

ACKNOWLEDGMENTS

We thank Marjori Matzke (IPMB, Taiwan) for critical comments on the manuscript. Work in the Schmidt lab is supported by MoST and Academia Sinica.

REFERENCES

- Baerenfaller, K., Massonnet, C., Walsh, S., Baginsky, S., Bühlmann, P., Hennig, L., et al. (2012). Systems-based analysis of *Arabidopsis* leaf growth reveals adaptation to water deficit. *Mol. Syst. Biol.* 8, 606. doi: 10.1038/msb.2012.39
- De Sousa Abreu, R., Penalva, L. O., Marcotte, E. M., and Vogel, C. (2009). Global signatures of protein and mRNA expression levels. *Mol. Biosyst.* 5, 1512–1526. doi: 10.1039/b908315d
- Filichkin, S. A., Priest, H. D., Givan, S. A., Shen, R., Bryant, D. W., Fox, S. E., et al. (2010). Genome-wide mapping of alternative splicing in *Arabidopsis thaliana*. *Genome Res.* 20, 45–58. doi: 10.1101/gr.093302.109
- 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
- Ingolia, N. T. (2014). Ribosome profiling: new views of translation, from single codons to genome scale. *Nat. Rev. Genet.* 15, 205–213. doi: 10.1038/nrg3645
- Juntawong, P., Girke, T., Bazin, J., and Bailey-Serres, J. (2014). Translational dynamics revealed by genome-wide profiling of ribosome footprints in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* 111, E203–E212. doi: 10.1073/pnas.1317811111
- Kongsuwan, K., Yu, Q., Vincent, A., Frisardi, M. C., Rosbash, M., Lengyel, J. A., et al. (1985). A *Drosophila* minute gene encodes a ribosomal protein. *Nature* 317, 555–558. doi: 10.1038/317555a0
- Kornblihtt, A. R., Schor, I. E., Alló, M., Dujardin, G., Petrillo, E., and Muñoz, M. J. (2013). Alternative splicing: a pivotal step between eukaryotic transcription and translation. *Nat. Rev. Mol. Cell Biol.* 14, 153–165. doi: 10.1038/nrm3525
- Kristensen, A. R., Gsponer, J., and Foster, L. J. (2013). Protein synthesis rate is the predominant regulator of protein expression during differentiation. *Mol. Syst. Biol.* 9, 689. doi: 10.1038/msb.2013.47
- Kwon, S. C., Yi, H., Eichelbaum, K., Fohr, S., Fischer, B., You, K. T., et al. (2013). The RNA-binding protein repertoire of embryonic stem cells. *Nat. Struct. Mol. Biol.* 20, 1122–1130. doi: 10.1038/nsmb.2638
- Lan, P., Li, W., Lin, W. D., Santi, S., and Schmidt, W. (2013). Mapping gene activity of *Arabidopsis* root hairs. *Genome Biol.* 14:R67. doi: 10.1186/gb-2013-14-6-r67
- Lan, P., Li, W., and Schmidt, W. (2012). Complementary proteome and transcriptome profiling in phosphate-deficient *Arabidopsis* roots reveals multiple levels of gene regulation. *Mol. Cell. Proteomics* 11, 1156–1166. doi: 10.1074/mcp.M112.020461
- Lan, P., and Schmidt, W. (2011). The enigma of eIF5A in the iron deficiency response of *Arabidopsis*. *Plant Signal. Behav.* 6, 528–530. doi: 10.4161/psb.6.4.14747

- Lee, M. V., Topper, S. E., Hubler, S. L., Hose, J., Wenger, C. D., Coon, J. J., et al. (2011). A dynamic model of proteome changes reveals new roles for transcript alteration in yeast. *Mol. Syst. Biol.* 7:514. doi: 10.1038/msb.2011.48
- Li, W., Lin, W. D., Ray, P., Lan, P., and Schmidt, W. (2013). Genome-wide detection of condition-sensitive alternative splicing in Arabidopsis roots. *Plant Physiol.* 162, 1750–1763. doi: 10.1104/pp.113.217778
- 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. doi: 10.1038/msb.2011.97
- Marquez, Y., Brown, J. W., Simpson, C., Barta, A., and Kalyna, M. (2012). Transcriptome survey reveals increased complexity of the alternative splicing landscape in Arabidopsis. *Genome Res.* 22, 1184–1195. doi: 10.1101/gr.134106.111
- Mustroph, A., Zanettia, M. E., Jang, C. J. H., Holtan, H. E., and Repetti, P. P., Galbraith, D. W., et al. (2009). Profiling transcriptomes of discrete cell populations resolves altered cellular priorities during hypoxia in Arabidopsis. *Proc. Natl. Acad. Sci. U.S.A.* 106, 18843–18848. doi: 10.1073/pnas.0906131106
- O'Leary, M. N., Schreiber, K. H., Zhang, Y., Duc, A. C., Rao, S., Hale, J. S., et al. (2013). The ribosomal protein Rpl22 controls ribosome composition by directly repressing expression of its own paralog, Rpl22l1. *PLoS Genet.* 9:e1003708. doi: 10.1371/journal.pgen.1003708
- Ponnala, L., Wang, Y., Sun, Q., and van Wijk, K. J. (2014). Correlation of mRNA and protein abundance in the developing maize leaf. *Plant J.* 78, 424–440. doi: 10.1111/tpj.12482
- Preiss, T., Baron-Benhamou, J., Ansoorge, W., and Hentze, M. W. (2003). Homodirectional changes in transcriptome composition and mRNA translation induced by rapamycin and heat shock. *Nat. Struct. Biol.* 10, 1039–1047. doi: 10.1038/nsb1015
- Rodríguez-Celma, J., Pan, I. C., Li, W. F., Lan, P., Buckhout, T. J., and Schmidt, W. (2013). The transcriptional response of Arabidopsis leaves to Fe deficiency. *Front. Plant Sci.* 4:276. doi: 10.3389/fpls.2013.00276
- Rosado, A., Li, R., van de Ven, W., Hsu, E., and Raikhel, N. V. (2012). Arabidopsis ribosomal proteins control developmental programs through translational regulation of auxin response factors. *Proc. Natl. Acad. Sci. U.S.A.* 109, 19537–19544. doi: 10.1073/pnas.1214774109
- Tatematsu, K., Ward, S., Leyser, O., Kamiya, Y., and Nambara, E. (2005). Identification of cis-elements that regulate gene expression during initiation of axillary bud outgrowth in Arabidopsis. *Plant Physiol.* 138, 757–766. doi: 10.1104/pp.104.057984
- Tremousaygue, D., Manevski, A., Bardet, C., Lescure, N., and Lescure, B. (1999). Plant interstitial telomere motifs participate in the control of gene expression in root meristems. *Plant J.* 20, 553–561. doi: 10.1046/j.1365-313X.1999.00627.x
- Uechi, T., Tanaka, T., and Kenmochi, N. (2001). A complete map of the human ribosomal protein genes: assignment of 80 genes to the cytogenetic map and implications for human disorders. *Genomics* 72, 223–230. doi: 10.1006/geno.2000.6470
- Vogel, C., and Marcotte, E. M. (2012). Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nat. Rev. Genet.* 13, 227–232. doi: 10.1038/nrg3185
- Walley, J. W., Shen, Z., Sartor, R., Wu, K. J., Osborn, J., Smith, L. G., et al. (2013). Reconstruction of protein networks from an atlas of maize seed proteotypes. *Proc. Natl. Acad. Sci. U.S.A.* 110, E4808–E4817. doi: 10.1073/pnas.1319113110
- Wang, J. Y., Lan, P., Gao, H. M., Zheng, L., Li, W. F., and Schmidt, W. (2013). Expression changes of ribosomal proteins in phosphate- and iron-deficient Arabidopsis roots predict stress-specific alterations in ribosome composition. *BMC Genomics* 14:783. doi: 10.1186/1471-2164-14-783
- Whittle, C. A., and Krochko, J. E. (2009). Transcript profiling provides evidence of functional divergence and expression networks among ribosomal protein gene paralogs in *Brassica napus*. *Plant Cell* 21, 2203–2219. doi: 10.1105/tpc.109.068411
- Wu, H. P., Su, Y. S., Chen, H. C., Chen, Y. R., Wu, C. C., Lin, W. D., et al. (2014). Genome-wide analysis of light-regulated alternative splicing mediated by photoreceptors in *Physcomitrella patens*. *Genome Biol.* 15:R10. doi: 10.1186/gb-2014-15-1-r10

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: 07 August 2014; accepted: 21 October 2014; published online: 07 November 2014.

Citation: Vélez-Bermúdez IC and Schmidt W (2014) The conundrum of discordant protein and mRNA expression. *Are plants special? Front. Plant Sci.* 5:619. doi: 10.3389/fpls.2014.00619

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

Copyright © 2014 Vélez-Bermúdez and Schmidt. 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.