Published in final edited form as:

Curr Opin Plant Biol. 2010 April ; 13(2): 126–131. doi:10.1016/j.pbi.2009.11.005.

OMICS MEET NETWORKS - USING SYSTEMS APPROACHES TO INFER REGULATORY NETWORKS IN PLANTS

Miguel A. Moreno-Risueno, Wolfgang Busch, and Philip N. Benfey

Duke University, Department of Biology and Center for Systems Biology, United States

Abstract

Many genomic-scale datasets in plants have been generated over the last few years. This substantial achievement has led to impressive progress, including some of the most detailed molecular maps in any multicellular organism. Networks and pathways have been reconstructed using transcriptome, genome-wide transcription factor binding, proteome and metabolome data, and subsequently used to infer functional interactions among genes, proteins and metabolites. However, more sophisticated systems biology approaches are needed to integrate different omics data sets. Ultimately, the integration of diverse and massive datasets into coherent models will improve our understanding of the molecular networks that underlie biological processes.

Introduction

In the past decades, most plant biology studies focused on identifying the function of single genes within the context of specific responses or pathways. Recent genome-scale analyses in Arabidopsis and several crop species have allowed plant scientists to globally address the function of pathways and regulatory networks. This has significantly changed our concept of how genes work together. In the era of omics, major challenges include interpretation and integration of large datasets to understand the principles underlying the regulation of networks and their relevance. In this review, we focus on how different genome-wide data sets have been and can be used to reconstruct biological networks.

Reconstructing networks from gene expression data

Understanding plant regulatory networks and the biological principles by which they are governed requires knowledge of genome-scale responses during development and to environmental stimuli. A set of studies has examined the transcriptome of different organs and developmental stages of Arabidopsis in response to about 40 conditions [1-3]. These studies constitute a major step towards the identification of gene regulatory networks in plants; however, intrinsic complexity of some of the networks might have been overlooked. Cell-type-specific transcriptomic profiling of the Arabidopsis root and of distinct functional domains in the shoot apical meristem have unveiled gene expression patterns with unprecedented resolution [4-6]. In the root, this cell-type specific data has been combined with transcriptional profiling of 13 longitudinal sections, which represent developmental stages, to create a spatiotemporal map [5]. Similarly in rice, gene expression data from 40 different cell-types at different developmental stages has been analyzed to produce a transcriptome atlas [7].

© 2009 Elsevier Ltd. All rights reserved.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Furthermore, profiling root cell types versus whole roots in response to salt and iron deprivation has revealed additional transcriptional complexity and has generated valuable information about specific networks operating at the tissue level [8].

Mining expression datasets, such as the transcriptomic maps, usually relies on different clustering algorithms to find groups of co-expressed genes. Genes belonging to the same co-expression clusters are hypothesized to be co-regulated genes under the same internal or external cues by similar transcription factors and form a transcriptional module or subnetwork (the so called ‘guilt by association’ principle). Based on this principle, regulatory hierarchies of gene expression can be inferred (Figure 1). Conventional microarray clustering approaches are based on separating a set of elements (genes) into related subsets based on a distance metric, or by principal-component analysis. However, analyzing more complex gene expression data from large-scale microarray datasets requires more sophisticated approaches. In the root map, finding transcriptional modules has been limited to identify a set of unique expression patterns within the more variable genes in the dataset. These clusters form the dominant regulatory modules and have been shown to be biologically significant [5,9]. In another approach, genome-scale data from time-course experiments under different environmental conditions have been processed using predefined hypotheses to identify relevant patterns associated with specific *cis*-regulatory elements bound by light-regulated genes and other transcription factors (TFs) [10].

Transcriptome data can be used to reconstruct regulatory networks. However, in order to build models, methods that describe nodes (genes) and their regulatory interactions (edges) are required. In several single-cell organisms and human B-cells, it is possible to predict regulators for clusters of co-expressed genes that are treated as a single unit, or alternatively, to infer a network of pairwise interactions between each regulator and its putative targets [11-13]. In plants, none of these approaches has been shown to produce meaningful results, due in all likelihood to limited information about TFs and *cis*-regulatory motifs. However, a network predicting about 19,000 interactions (edges) among 6760 genes (nodes) has been reconstructed using about 2000 microarrays and a modified graphical Gaussian model based on partial correlations [14]. This method successfully inferred subnetworks related to metabolic functions and stress responses that are potentially useful to predict the function of novel genes. Furthermore, integration of coexpression data, gene ontology annotations, and *cis*-motifs has been shown to coherently link genes to specific biological functions [15]. These approaches facilitate further dissection of these novel functions by means of gain- or loss-of-function mutants or RNAi/amiRNA strategies [16-18]. However, high-throughput approaches that test the *in vivo* relevance of edges and nodes within a predicted network are needed to reveal the biological significance of such interactions.

Integration of gene expression and transcription factor binding data

For a comprehensive reconstruction of transcriptional networks it is essential to determine if a transcription factor regulates a gene in a direct or an indirect fashion (Figure 1). However, it is not possible to determine such regulatory interactions with high confidence on the basis of transcriptome data alone. Chromatin immunoprecipitation coupled with hybridization to whole genome arrays (ChIP-chip) or deep sequencing (ChIPseq) promise to fill this gap.

Despite the assumption that ChIP-chip data would enable the complete reconstruction of transcriptional networks, new complexities arose with those data sets. Previous studies identified varying numbers of binding sites for different transcription factors, ranging from less than 100 [19] to several hundred [20-22] and even up to several thousand [23,24]. Interestingly, only a very small fraction of the genes that were bound responded transcriptionally to altered levels of the respective transcription factor. In particular, most of

the cases reported that less than 10% of the directly bound genes were among genes with significant transcriptional response [21,23,24].

The integration of data from transcriptome and ChIP-chip experiments was typically attempted by classifying transcriptionally responsive genes into a high-confidence direct target set and another set that included indirect target genes [20,22-24]. A more sophisticated approach was undertaken to study transcriptional networks of trichome development. In this case, transcriptome profiles of mutant plants and inducible versions of the transcription factors GLABRA3 (GL3) and GLABRA1 (GL1) were integrated via a meta-analysis. This enabled the detection of a minimal set of genes that were uniquely associated with the formation of trichomes. This gene set, interpreted with results from individual ChIP-chip experiments for both transcription factors led to a multi-tiered reconstruction of the GL1/GL3 regulated network of trichome formation. [21].

It is startling that almost a decade after the first introduction of the ChIP-chip methodology, only a very limited number of genome wide transcription factor occupancy studies have been successfully conducted in the plant field. Typically each study confirmed known binding motifs and showed that most binding seemed to take place within the 2000 bp proximal upstream region of a gene and identified the enrichment of GO categories that were consistent with prior biological knowledge [20-24]. Because whole genome transcription factor binding data should, in theory, explain a massive part of transcriptional regulation, the low amount of correlation between directly bound genes and genes that are responding to alterations in transcription factor levels is puzzling. However, similar low level agreement between ChIP-chip and expression data was also frequently found in other organisms [25]. This may indicate that many binding events are silent most of the time, but they may play a crucial role in dynamic transient regulatory interactions. Since this often cannot be easily assayed, more sophisticated approaches to monitor transient expression changes of target genes, or elucidation of the function of a transcription factor in different developmental or environmental contexts should increase the agreement of these data sets. A more complete understanding of transcription factor function is obligatory to define transcriptional modules with a high confidence and refine the networks that were originally reconstructed through 'guilt by association' principles.

Interpretation and integration of omics data

Proteomic studies use Mass Spectrometry (MS) based methods to profile the proteome, phosphoproteome or plasma membrane proteome, while metabolomic studies use MS or Nuclear Magnetic Resonance to profile the metabolites of whole plants, organs, cells or organelles [26-29]. Remarkably, analyses of the proteome of six organs plus undifferentiated cultured cells identified around 13,000 distinct proteins to construct the first Arabidopsis organ proteome map [30]. The correlation of these data with transcriptome profiles suggests that protein accumulation in Arabidopsis is primarily regulated at the transcriptional level. Notably, profiling specific cell-types or enriching certain sub-cellular components seems to increase the resolution of these studies and provides deeper biological insight. MS analysis of isolated guard cells has detected 336 novel proteins and has identified a novel glucosinolate-myrosinase system required for ABA signaling [31]. Moreover, isolation of plasma membrane proteins revealed that a small set of kinases, called BSKs, transduce brassinosteroid signaling downstream of the BRI1 receptor that had been overlooked when total proteins were analyzed [32].

Similar to transcriptional networks, protein and metabolic networks can be reconstructed using edges and nodes. In protein networks, nodes are the proteins and edges normally represent protein-protein interactions (interactome) or a functional modification, whereas in metabolic networks, nodes are considered to be metabolites and edges enzymatic reactions or biochemical

modifications (Figure 1). The aforementioned studies can identify nodes of protein networks and be potentially used to infer functional relationships; however, edges are usually identified through different approaches. In several model organisms protein-protein interactions have been mined to produce proteome-scale interactome networks, while in plants this information is still missing. Recently, on a smaller scale, protein microarrays have revealed the target network of several *Arabidopsis* calmodulins [33]; and 10 mitogen-activated protein kinases (MPK) that are in turn phosphorylated by 9 MPK kinases (MKK) [34]. The MKK/MPK network contains 589 nodes and 1331 edges representing predicted phosphorylation events. Interestingly, *in vivo* assays link this network to developmental and pathogen responses [34].

Metabolome studies have the potential to reveal biochemical networks, but integration of unknown metabolites into networks cannot be achieved by correlation analyses as in transcriptional networks. One alternative is combining quantitative trait locus analysis with quantitative metabolite profiling to generate *de novo* biochemical network models [35]. However, most of the work inferring biochemical pathways is based on the principle that under conditions that modulate the accumulation of a certain metabolite, gene expression patterns can then be correlated to decode the function of genes involved in that metabolic pathway. By using this strategy, R2R3-MYB TFs were correctly inferred as involved in glucosinolate biosynthesis [36] and enzymes responsible for variation of chain length of methionine-derived glucosinolates were identified [37]. Moreover, combined transcriptomic and metabolomic profiling of plants over-expressing glucosinolate regulators has revealed connections to other metabolic pathways and primary metabolism [38]. Using similar methods, other studies reported the decoding of novel gene functions in the flavonoid [39] and other biochemical [40] pathways, as well as the possible coupling of the circadian system in *Arabidopsis* with chloroplast and mitochondrion functions [41].

The combination of different omics data will facilitate the functional identification of networks controlling development, growth and response to biotic and abiotic stresses. For instance, it was recently reported that phosphorylation of enzymes involved in metabolic processes correlates with plant growth [42]. Databases have the potential to facilitate hypothesis generation for plant biology by integrating omic datasets. Currently, there are different data repositories, such as ArrayExpress, GEO, CYBEX, PRIDE, Golm Metabolome DataBase and others [43]; while GabiPD [44] intends to be an integrative omics database. However, new computational approaches and interfaces that allow a joint interpretation of massive datasets are needed to form coherent models of biological networks and their function.

Conclusions and future perspectives

Our comprehension of complex molecular networks that underlie biological processes has grown dramatically in the last few years. In addition to the progress discussed above, development in other areas of research has contributed to our understanding of molecular networks, like protein-DNA binding microarrays [45] and genome-wide profiling of histone modifications and DNA methylation [46,47]. However, for future understanding and reconstruction of complex molecular networks, experimental design and technical sophistication of the experiments should be brought to a level that suits the complexity of the networks that are being studied (Figure 1). Whereas chromatin binding maps of transcription factors and whole-genome interactome networks explore a general parameter space for biological networks, cell-type specific data such as transcriptional data, nucleosome occupancy, epigenetic chromatin-landscapes, proteomics and metabolomics are needed to address specific topologies of regulatory networks. In addition, new technologies, such as medium to high throughput microscopy, are emerging and promise to obtain sufficient spatiotemporal resolution for observing the dynamics of aspects of complex systems.

One major challenge of performing system approaches in Arabidopsis is transferring such knowledge to crop species. Translational research seems feasible and approaches similar to those used in Arabidopsis have been applied in rice and other crops, e.g. maize and barley [7, 48-50]. These achievements potentially enable the reconstruction of regulatory pathways and networks in crops, which will lead to new biotechnological applications in the future.

Acknowledgments

We thank members of the Benfey lab for critical reading of the manuscript and helpful suggestions; and apologize to those, whose work we could not cover due to space limitations. Work in the Benfey lab on regulatory networks in plants is funded by grants from the NIH, NSF and DARPA.

References

1. Goda H, Sasaki E, Akiyama K, Maruyama-Nakashita A, Nakabayashi K, Li W, Ogawa M, Yamauchi Y, Preston J, Aoki K, et al. The AtGenExpress hormone and chemical treatment data set: experimental design, data evaluation, model data analysis and data access. *Plant J* 2008;55:526–542. [PubMed: 18419781]
2. Kilian J, Whitehead D, Horak J, Wanke D, Weinl S, Batistic O, D'Angelo C, Bornberg-Bauer E, Kudla J, Harter K. The AtGenExpress global stress expression data set: protocols, evaluation and model data analysis of UV-B light, drought and cold stress responses. *Plant J* 2007;50:347–363. [PubMed: 17376166]
3. Schmid M, Davison TS, Henz SR, Pape UJ, Demar M, Vingron M, Scholkopf B, Weigel D, Lohmann JU. A gene expression map of Arabidopsis thaliana development. *Nat Genet* 2005;37:501–506. [PubMed: 15806101]
4. Birnbaum K, Shasha DE, Wang JY, Jung JW, Lambert GM, Galbraith DW, Benfey PN. A gene expression map of the Arabidopsis root. *Science* 2003;302:1956–1960. [PubMed: 14671301]
5. Brady SM, Orlando DA, Lee JY, Wang JY, Koch J, Dinneny JR, Mace D, Ohler U, Benfey PN. A high-resolution root spatiotemporal map reveals dominant expression patterns. *Science* 2007;318:801–806. [PubMed: 17975066] ** In this work, the authors profiled the transcriptome of nearly all the root cell types and 12 developmental stages to create a comprehensive gene expression map and identify significant transcriptional modules.
6. Yadav RK, Girke T, Pasala S, Xie M, Reddy GV. Gene expression map of the Arabidopsis shoot apical meristem stem cell niche. *Proc Natl Acad Sci U S A* 2009;106:4941–4946. [PubMed: 19258454]
7. Jiao Y, Tausta SL, Gandotra N, Sun N, Liu T, Clay NK, Ceserani T, Chen M, Ma L, Holford M, et al. A transcriptome atlas of rice cell types uncovers cellular, functional and developmental hierarchies. *Nat Genet* 2009;41:258–263. [PubMed: 19122662] ** The authors profiled 40 different rice cell types at different developmental stages by laser capture microdissection and microarray analysis to create the first gene expression atlas of a crop plant.
8. Dinneny JR, Long TA, Wang JY, Jung JW, Mace D, Pointer S, Barron C, Brady SM, Schiefelbein J, Benfey PN. Cell identity mediates the response of Arabidopsis roots to abiotic stress. *Science* 2008;320:942–945. [PubMed: 18436742] * The effect of two environmental stresses on the transcriptional responses of the Arabidopsis root was elegantly examined by means of fluorescent activated cell sorting and dissection of four developmental zones
9. Orlando DA, Brady SM, Koch JD, Dinneny JR, Benfey PN. Manipulating large-scale Arabidopsis microarray expression data: identifying dominant expression patterns and biological process enrichment. *Methods Mol Biol* 2009;553:57–77. [PubMed: 19588101]
10. Michael TP, Mockler TC, Breton G, McEntee C, Byer A, Trout JD, Hazen SP, Shen R, Priest HD, Sullivan CM, et al. Network discovery pipeline elucidates conserved time-of-day-specific cis-regulatory modules. *PLoS Genet* 2008;4:e14. [PubMed: 18248097]
11. Basso K, Margolin AA, Stolovitzky G, Klein U, Dalla-Favera R, Califano A. Reverse engineering of regulatory networks in human B cells. *Nat Genet* 2005;37:382–390. [PubMed: 15778709]
12. Joshi A, De Smet R, Marchal K, Van de Peer Y, Michoel T. Module networks revisited: computational assessment and prioritization of model predictions. *Bioinformatics* 2009;25:490–496. [PubMed: 19136553]

13. Segal E, Shapira M, Regev A, Pe'er D, Botstein D, Koller D, Friedman N. Module networks: identifying regulatory modules and their condition-specific regulators from gene expression data. *Nat Genet* 2003;34:166–176. [PubMed: 12740579]
14. Ma S, Gong Q, Bohnert HJ. An Arabidopsis gene network based on the graphical Gaussian model. *Genome Res* 2007;17:1614–1625. [PubMed: 17921353]
15. Vandepoele K, Quimbaya M, Casneuf T, De Veylder L, Van de Peer Y. Unraveling transcriptional control in Arabidopsis using cis-regulatory elements and coexpression networks. *Plant Physiol* 2009;150:535–546. [PubMed: 19357200]
16. Iyer-Pascuzzi A, Simpson J, Herrera-Estrella L, Benfey PN. Functional genomics of root growth and development in Arabidopsis. *Curr Opin Plant Biol* 2009;12:165–171. [PubMed: 19117793]
17. Alonso JM, Stepanova AN, Leisse TJ, Kim CJ, Chen H, Shinn P, Stevenson DK, Zimmerman J, Barajas P, Cheuk R, et al. Genome-wide insertional mutagenesis of Arabidopsis thaliana. *Science* 2003;301:653–657. [PubMed: 12893945]
18. Schwab R, Ossowski S, Riester M, Warthmann N, Weigel D. Highly specific gene silencing by artificial microRNAs in Arabidopsis. *Plant Cell* 2006;18:1121–1133. [PubMed: 16531494]
19. Thibaud-Nissen F, Wu H, Richmond T, Redman JC, Johnson C, Green R, Arias J, Town CD. Development of Arabidopsis whole-genome microarrays and their application to the discovery of binding sites for the TGA2 transcription factor in salicylic acid-treated plants. *Plant J* 2006;47:152–162. [PubMed: 16824183]
20. Mathieu J, Yant LJ, Murdter F, Kuttner F, Schmid M. Repression of flowering by the miR172 target SMZ. *PLoS Biol* 2009;7:e1000148. [PubMed: 19582143]
21. Morohashi K, Grotewold E. A systems approach reveals regulatory circuitry for Arabidopsis trichome initiation by the GL3 and GL1 selectors. *PLoS Genet* 2009;5:e1000396. [PubMed: 19247443] **
The authors used transcriptome and genome wide transcription factor binding (ChIP-chip) data to reconstruct several tiers of the *GL1/GL3*-controlled network of trichome formation.
22. Oh E, Kang H, Yamaguchi S, Park J, Lee D, Kamiya Y, Choi G. Genome-wide analysis of genes targeted by PHYTOCHROME INTERACTING FACTOR 3-LIKE5 during seed germination in Arabidopsis. *Plant Cell* 2009;21:403–419. [PubMed: 19244139]
23. Lee J, He K, Stole V, Lee H, Figueroa P, Gao Y, Tongprasit W, Zhao H, Lee I, Deng XW. Analysis of transcription factor HY5 genomic binding sites revealed its hierarchical role in light regulation of development. *Plant Cell* 2007;19:731–749. [PubMed: 17337630]
24. Zheng Y, Ren N, Wang H, Stromberg AJ, Perry SE. Global Identification of Targets of the Arabidopsis MADS Domain Protein AGAMOUS-Like15. *Plant Cell*. 2009
25. Farnham PJ. Insights from genomic profiling of transcription factors. *Nat Rev Genet* 2009;10:605–616. [PubMed: 19668247]
26. Agrawal GK, Jwa NS, Rakwal R. Rice proteomics: ending phase I and the beginning of phase II. *Proteomics* 2009;9:935–963. [PubMed: 19212951]
27. Jorin-Novo JV, Maldonado AM, Echevarria-Zomeno S, Valledor L, Castillejo MA, Curto M, Valero J, Sghaier B, Donoso G, Redondo I. Plant proteomics update (2007-2008): Second-generation proteomic techniques, an appropriate experimental design, and data analysis to fulfill MIAPE standards, increase plant proteome coverage and expand biological knowledge. *J Proteomics* 2009;72:285–314. [PubMed: 19367730]
28. Kersten B, Agrawal GK, Durek P, Neigenfind J, Schulze W, Walther D, Rakwal R. Plant phosphoproteomics: an update. *Proteomics* 2009;9:964–988. [PubMed: 19212952]
29. Moco S, Schneider B, Vervoort J. Plant Micrometabolomics: The Analysis of Endogenous Metabolites Present in a Plant Cell or Tissue. *J Proteome Res*. 2009
30. Baerenfaller K, Grossmann J, Grobei MA, Hull R, Hirsch-Hoffmann M, Yalovsky S, Zimmermann P, Grossniklaus U, Gruissem W, Baginsky S. Genome-scale proteomics reveals Arabidopsis thaliana gene models and proteome dynamics. *Science* 2008;320:938–941. [PubMed: 18436743] **
The authors created the first Arabidopsis organ proteome map. This large scale-view of the proteome was also used to examine differences between transcriptomic and proteomic data and to identify proteins specifically expressed in different organs.

31. Zhao Z, Zhang W, Stanley BA, Assmann SM. Functional proteomics of *Arabidopsis thaliana* guard cells uncovers new stomatal signaling pathways. *Plant Cell* 2008;20:3210–3226. [PubMed: 19114538]
32. Tang W, Kim TW, Osés-Prieto JA, Sun Y, Deng Z, Zhu S, Wang R, Burlingame AL, Wang ZY. BSKs mediate signal transduction from the receptor kinase BRI1 in *Arabidopsis*. *Science* 2008;321:557–560. [PubMed: 18653891]
33. Popescu SC, Popescu GV, Bachan S, Zhang Z, Seay M, Gerstein M, Snyder M, Dinesh-Kumar SP. Differential binding of calmodulin-related proteins to their targets revealed through high-density *Arabidopsis* protein microarrays. *Proc Natl Acad Sci U S A* 2007;104:4730–4735. [PubMed: 17360592]
34. Popescu SC, Popescu GV, Bachan S, Zhang Z, Gerstein M, Snyder M, Dinesh-Kumar SP. MAPK target networks in *Arabidopsis thaliana* revealed using functional protein microarrays. *Genes Dev* 2009;23:80–92. [PubMed: 19095804] ** The authors used protein microarrays to assay phosphorylation events of 10 *Arabidopsis* mitogen-activated protein kinases (MPK) and 9 MPK kinases over 2158 proteins synthesized *in planta*. This data was used to infer the target network of these MPK/MPKK modules.
35. Rowe HC, Hansen BG, Halkier BA, Kliebenstein DJ. Biochemical networks and epistasis shape the *Arabidopsis thaliana* metabolome. *Plant Cell* 2008;20:1199–1216. [PubMed: 18515501]
36. Hirai MY, Sugiyama K, Sawada Y, Tohge T, Obayashi T, Suzuki A, Araki R, Sakurai N, Suzuki H, Aoki K, et al. Omics-based identification of *Arabidopsis* Myb transcription factors regulating aliphatic glucosinolate biosynthesis. *Proc Natl Acad Sci U S A* 2007;104:6478–6483. [PubMed: 17420480] * The authors combined transcriptomic and metabolomic data to identify two transcription factors involved in the regulation of glucosinolate biosynthesis. This study exemplifies how data integration can be used to predict novel gene functions and further test these hypotheses by reverse genetics.
37. Sawada Y, Kuwahara A, Nagano M, Narisawa T, Sakata A, Saito K, Hirai MY. Omics-based approaches to methionine side chain elongation in *Arabidopsis*: characterization of the genes encoding methylthioalkylmalate isomerase and methylthioalkylmalate dehydrogenase. *Plant Cell Physiol* 2009;50:1181–1190. [PubMed: 19493961]
38. Malitsky S, Blum E, Less H, Venger I, Elbaz M, Morin S, Eshed Y, Aharoni A. The transcript and metabolite networks affected by the two clades of *Arabidopsis* glucosinolate biosynthesis regulators. *Plant Physiol* 2008;148:2021–2049. [PubMed: 18829985]
39. Yonekura-Sakakibara K, Tohge T, Matsuda F, Nakabayashi R, Takayama H, Niida R, Watanabe-Takahashi A, Inoue E, Saito K. Comprehensive flavonol profiling and transcriptome coexpression analysis leading to decoding gene-metabolite correlations in *Arabidopsis*. *Plant Cell* 2008;20:2160–2176. [PubMed: 18757557]
40. Saito K, Hirai MY, Yonekura-Sakakibara K. Decoding genes with coexpression networks and metabolomics - ‘majority report by precogs’. *Trends Plant Sci* 2008;13:36–43. [PubMed: 18160330]
41. Fukushima A, Kusano M, Nakamichi N, Kobayashi M, Hayashi N, Sakakibara H, Mizuno T, Saito K. Impact of clock-associated *Arabidopsis* pseudo-response regulators in metabolic coordination. *Proc Natl Acad Sci U S A* 2009;106:7251–7256. [PubMed: 19359492] * A Combination of transcriptomic and metabolomic profiling was used to infer novel functions for the *Arabidopsis* PSEUDO-RESPONSE REGULATORS 5, 7 and 9 and establish a connection between circadian rhythm and functions of the chloroplast and mitochondrion.
42. Shin R, Alvarez S, Burch AY, Jez JM, Schachtman DP. Phosphoproteomic identification of targets of the *Arabidopsis* sucrose nonfermenting-like kinase SnRK2.8 reveals a connection to metabolic processes. *Proc Natl Acad Sci U S A* 2007;104:6460–6465. [PubMed: 17404219]
43. Brady SM, Provart NJ. Web-queryable large-scale data sets for hypothesis generation in plant biology. *Plant Cell* 2009;21:1034–1051. [PubMed: 19401381]
44. Riano-Pachon DM, Nagel A, Neigenfind J, Wagner R, Basekow R, Weber E, Mueller-Roeber B, Diehl S, Kersten B. GABI PD: the GABI primary database--a plant integrative ‘omics’ database. *Nucleic Acids Res* 2009;37:D954–959. [PubMed: 18812395]
45. Badis G, Berger MF, Philippakis AA, Talukder S, Gehrke AR, Jaeger SA, Chan ET, Metzler G, Vedenko A, Chen X, et al. Diversity and complexity in DNA recognition by transcription factors. *Science* 2009;324:1720–1723. [PubMed: 19443739]

46. Lister R, O'Malley RC, Tonti-Filippini J, Gregory BD, Berry CC, Millar AH, Ecker JR. Highly integrated single-base resolution maps of the epigenome in Arabidopsis. *Cell* 2008;133:523–536. [PubMed: 18423832]
47. Zhang X, Bernatavichute YV, Cokus S, Pellegrini M, Jacobsen SE. Genome-wide analysis of mono-, di- and trimethylation of histone H3 lysine 4 in Arabidopsis thaliana. *Genome Biol* 2009;10:R62. [PubMed: 19508735]
48. Dembinsky D, Woll K, Saleem M, Liu Y, Fu Y, Borsuk LA, Lamkemeyer T, Fladerer C, Madlung J, Barbazuk B, et al. Transcriptomic and proteomic analyses of pericycle cells of the maize primary root. *Plant Physiol* 2007;145:575–588. [PubMed: 17766395] * One of the few studies that combines transcriptomic and proteomic analyses in crops. The authors performed microarrays, high-throughput EST sequencing and two-dimensional electrophoresis on pericycle cells isolated by laser capture microdissection. These data were used to identify genes and proteins specifically expressed in this cell type.
49. Sreenivasulu N, Usadel B, Winter A, Radchuk V, Scholz U, Stein N, Weschke W, Strickert M, Close TJ, Stitt M, et al. Barley grain maturation and germination: metabolic pathway and regulatory network commonalities and differences highlighted by new MapMan/PageMan profiling tools. *Plant Physiol* 2008;146:1738–1758. [PubMed: 18281415]
50. Bonsager BC, Finnie C, Roepstorff P, Svensson B. Spatio-temporal changes in germination and radical elongation of barley seeds tracked by proteome analysis of dissected embryo, aleurone layer, and endosperm tissues. *Proteomics* 2007;7:4528–4540. [PubMed: 18022942]

Figure 1.

Approaches to network inference and reconstruction. Profiling of metabolites, transcripts and proteins as well as chromatin immunoprecipitation coupled with hybridization to microarrays or sequencing can be used to infer functional interactions among genes, proteins and metabolites. Interpretation and integration of omic datasets facilitates hypothesis generation and prediction of novel gene functions in plant biology.