Review

Plant systems biology comes of age

Joshua S. Yuan[1,2], David W. Galbraith[3], Susie Y. Dai[4], Patrick Griffin[4] and C. Neal Stewart Jr.[2]

[1] UTIA Genomics Hub, University of Tennessee, Knoxville, TN 37996, USA
[2] Department of Plant Sciences, University of Tennessee, Knoxville, TN 37996, USA
[3] BIO5 Institute for Collaborative Bioresearch and Department of Plant Sciences, University of Arizona, Tucson, AZ 85721, USA
[4] Department of Molecular Therapeutics, The Scripps Research Institute, Jupiter, FL 33458, USA

'Omics' research approaches have produced copious data for living systems, which have necessitated the development of systems biology to integrate multidimensional biological information into networks and models. Applications of systems biology to plant science have been rapid, and have increased our knowledge about circadian rhythms, multigenic traits, stress responses and plant defenses, and have advanced the virtual plant project.

Biological systems and systems biology in plants

The practice of integrating physiological, morphological, molecular, biochemical and genetic information has long been applied to biological research, and in diverse fields such as plant breeding and ecology [1]. The development of modern systems biology was driven by the need to assimilate the large amounts of data generated by genome-scale studies into biologically meaningful interpretations. Nevertheless, the definition of systems biology is still contentious; some researchers emphasize the role of dynamic modeling, whereas others stress multidimensional data analysis. Considering the infancy of the field, this dichotomy is not surprising [2]. Here, we define systems biology as the study of interactions among biological components using models and/or networks to integrate genes, metabolites, proteins, regulatory elements and other biological components. We focus mainly on networks. Both component integration and dynamic interactions are key features of systems biology. There is an immediate need for systems integration in plant biology, considering the large datasets generated from different omics technologies such as genomics, proteomics, transcriptomics, interactomics and metabolomics (Box 1). Contrasting with the immense recent efforts of generating large datasets, there has been less development of platforms to integrate multidimensional data to derive models for describing biological interactions in plants [1–3]. Some prominent reviews have been published regarding one or two aspects of plant systems biology or its history [1,3–5]. Other reviews have discussed the techniques used in systems biology, such as microarray, next-generation sequencing, synthetic genetic array, pull-down assays, the yeast two-hybrid system, systems biology markup language and other important tools [6–9]. However, most of the previous plant systems biology reviews focused on the application of systems biology in one area of plant biology [5,10–20], and a wide-ranging review covering different types of networks and broad application perspectives is still not available. We therefore aim to provide a comprehensive review on recent advances in the integration of multidimensional data into networks and the application of these networks in addressing questions in plant biology.

Advances in plant systems biology – a network perspective

Network (see Glossary) construction and analysis is one of the most common approaches to describe biological systems. Networks can be either static or dynamic, and their components can include genes, proteins, *cis*-elements, metabolites and other molecules. Here, we focus on four network types, including gene-to-metabolite networks, protein–protein interaction networks, transcriptional regulatory networks and gene regulatory networks (Figure 1). The first three types of networks are often static, whereas the gene regulatory network frequently emphasizes the dynamic changes of processes.

Gene-to-metabolite networks

Gene-to-metabolite networks define the interactions among genes and metabolites and are typically constructed using multivariate analysis or data mining

Glossary

Data mining: the computational process to search for consistent patterns or systemic relationships among the variables in a complex dataset. It includes several popular techniques, such as neural network, decision tree and logistic regression.
Data modeling: the computational approach to determine relationships among concepts or objects. In systems biology, data modeling is used to elucidate relationships among system elements based on the biological information.
Data visualization: techniques to describe relationships among the elements with graphics, images or animations.
Dynamic network: this network integrates interactive and kinetic changes of a system. The network is often built up through the reiterative modeling and validation processes. It can be used to predict the outcome of a stimulus in the network.
Meta-analysis: statistical approaches to integrate data from multiple studies to derive related hypotheses, combining multiple analyses for a grand analysis. In practice, meta-analysis can be used to establish common relationships among components based on different types of data and among various empirical studies that might not be directly comparable with one another.
Multivariate analysis: statistical approaches used to analyze more than one variable at a time. In practice, these methods are used to reduce the dimension of the data in a complex dataset with multiple variables.
Network: the interaction and integration of multiple components in organisms using computational models and visualized by node and connection diagrams.
Static network: in contrast to a dynamic network, a static network is constructed without kinetic information. It is used more often to describe the associations among system components.

Corresponding author: Stewart, C.N. Jr. (nealstewart@utk.edu).

1360-1385/$ – see front matter doi:10.1016/j.tplants.2008.02.003 Available online 7 March 2008

Box 1. Systems biology and omics

Systems biology is related but not synonymous to the postgenomics omics technologies such as microarrays for high-throughput generation of large-scale data. Genome sequencing enables functional or comparative studies of plant genomes [48]. Genome information also leads to the study of the mRNA transcripts and proteins of an organism as a whole, which are referred to as transcriptomics and proteomics, respectively [7]. Research of all or most of the metabolites in an organism is referred to as metabolomics [81]. The study of genome-scale interactions among proteins is referred to as interactomics [46]. The advances in RNA interference and other mutagenesis technologies have enabled high-throughput phenotype screens for genes; this is referred to as phenomics [15]. With the recent advances in analytical techniques, the list of omics is growing to include, for example, glycomics, fluxomics and ionomics [82–84].

Regardless of which omics techniques are used to generate data, the goal for systems biology is to define the structure, dynamics and control of biological systems. Thus, systems biology extends beyond the omics list of genes or proteins to the functional whole plant, which requires bioinformatics to model suites of biological data into networks, which are representations of the actual plant. Bioinformatics, then, is the key for successful integration and presentation of systems biology data. Considering the complexity and diversity of biological systems, the computational models for different applications are expected to be diverse.

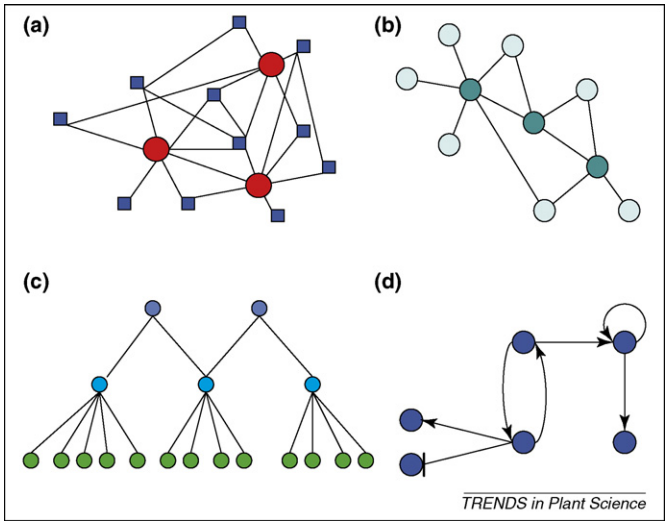

Figure 1. Networks in plant systems biology. The four common networks used in plant systems biology study are shown. **(a)** The gene-to-metabolite network often derives from correlation analysis of gene and metabolite profiling under multiple conditions. In the network, genes are typically symbolized differently to metabolites. Here, genes are represented by circles and the metabolites by squares. The interactions are symbolized with lines. **(b)** Interactomes can be derived from genetic or protein-binding assays such as yeast two-hybrid assay and co-immunoprecipitation. These can be visualized with circles representing genes or proteins and lines representing interactions. The genes in centralized hub locations with many interactions among multiple genes are often symbolized with a different color. **(c)** The transcriptional regulatory network is highly diverse and can be presented as a hierarchical structure. The elements at the top are expected to be general regulatory genes. **(d)** The gene regulatory network can be derived from gene expression profiles, mutant analysis and other data. The gene regulatory network is dynamic, and system dynamics need to be visualized in the graph by the different symbols of the lines. The genes are often represented by circles and the interactions between genes are often represented by lines. Different symbols at the end of the lines can describe different types of interactions, including gene activation and repression.

of gene profiling and metabolite profiling data under different conditions (Figure 1a). The outputs from statistical analyses are often visualized based on the distance calculated among genes and metabolites according to their profiling patterns. If a gene is determined as being 'close' to a metabolite, it might be implicated in the biosynthesis or regulation of the latter. The study of gene-to-metabolite networks is more complex in plants, particularly in comparison with mammals, because of the greater diversity and larger numbers of metabolites produced by plants as an adaptation to their sessile life style.

Early research used gene-to-metabolite networks to dissect the dynamic responses during sulfur and nitrogen starvation in *Arabidopsis* [21]. The work integrated microarray-based gene profiling with liquid chromatography mass spectrometry (MS) and Fourier transform–ion cyclotron MS-based metabolite profiling using multivariate analysis methods including self-organizing map and principal components analysis to derive gene-to-metabolite associations [21,22]. In a follow-up study, analyses of gene-to-metabolite networks led to identification of several previously unknown desulfoglucosinolate sulfotransferases and candidate transcriptional factors regulating anthocyanin biosynthesis [23]. By combining systems biology with reductionism, coordinated changes of genes and metabolites following alteration of expression of key signal transduction pathway genes have been elucidated [24,25]. Recently, gene-to-metabolite networks have been characterized for stress responses, plant defense and hormone-induced responses [26–28]. Moreover, the gene-to-metabolite networks have also been constructed for plant species with limited available genome information, such as Madagascar periwinkle (*Catharanthus roseus*) [29]; the network analysis in this case led to the discovery of novel candidate genes for terpenoid indole alkaloid biosynthesis.

In silico analysis of previously published gene expression datasets can be used to construct metabolite- or gene-to-metabolite networks. This type of analysis is based on the assumption that genes within the same metabolic pathway are frequently coregulated. Such meta-analysis of microarray data identifies coregulated genes with higher correlation efficiency, and these groups of coregulated genes can then be associated with particular metabolic pathways. This approach has recently been used to identify the genes involved in metabolite biosynthesis and transcriptional regulation of 140 different pathways [30].

Another approach to integrate gene expression data with metabolism was to present the expression data in the context of metabolic pathways using MapMan, software enabling statistical treatment of multiple microarray datasets to display the significantly changed genes in the corresponding metabolic pathway [31]. This approach was successfully applied to identify the genes and metabolic pathways involved in the response to nitrogen deficiency and during diurnal cycles [32,33]. Notably, the integration of gene expression data with sugar-related pathways in both wild-type *Arabidopsis* and a starchless phosphoglucomutase (*pgm*) mutant enabled the dissection of the intervening sugar and circadian rhythm signaling and the identification of sugar-responsive genes during the diurnal cycle [32,34].

Over the past four years, different types of gene-to-metabolite networks have added new perspectives and insights in plant science. First, the gene-to-metabolite network clarified how biological processes are interrelated,

enabling substantial improvements in omics data interpretation, better prediction of outcomes for system perturbations and conceptual reconstruction of interactive biological systems with multiple components, including enzyme activities, gene expression and metabolite levels [35]. Second, the systemic view enables the discovery of key regulatory components in a biological system and thus improves gene function annotation [24]. Third, the association of genes with metabolites enables the discovery of new genes involved in metabolite biosynthesis, transport, regulation and modification, in addition to their regulation [23]. Despite progress, the application of the gene-to-metabolite network approach has been hampered by technological and computational limitations. The structure of many plant metabolites cannot be readily resolved by MS-based methods [36]. The gene expression analyses using microarray or serial analysis of gene expression often require extensive genome information of the plant species studied [7]. In addition, linear correlation has been used to establish most of the gene-to-metabolite network, which cannot reflect the dynamics of the system. Technological challenges can be met using improved MS and nuclear magnetic resonance-based metabolite identification, in addition to the application of next-generation sequencing in transcriptome profiling [3,6]. More precise gene-to-metabolite network construction will be possible with the rapid progress in both technology and algorithm development.

Plant interactome and protein interaction networks

Two types of interactome (Box 1) have recently been recognized: genetic and physical (Figure 1b). A genetic interactome is a network of genes characterized on the basis of genetic interactions serving to elucidate gene function within physiological processes [9]. A good example of this approach is systematic genetic analysis (SGA) for yeast [37]. One characteristic feature of the yeast genetic interactome is that genes having important cellular functions typically form interaction hubs, at the center of multiple interactions with many other genes [38]. The SGA-based genetic interactome has also been applied successfully in the nematode (*Caenorhabditis elegans*) [39]. Despite its potential, the application of SGA to plants is often complicated by polyploidy and a comparatively long life cycle of plants that decreases precision and increases the time needed for genetic manipulation. These factors confound the emergence of phenotypes because of redundancy and noisy data. Nonetheless, similar approaches of multiple mutations have been adopted to identify the two-component interaction network for *Arabidopsis* type-A response regulators involved in cytokinin signaling [40,41]. Even though the establishment of a genome-wide genetic interactome map currently might be difficult, the local genetic interaction network of proteins can provide a better understanding of gene functions and their interactions. Besides the SGA interactome, an RNA interference-based method of downregulating target genes can also help to create genetic interactome maps, considering the relatively facile transformation of *Arabidopsis* [40,41].

As compared with genetic interactions, physical interactions among proteins are relatively easier to characterize in plants. Two main techniques are the yeast two-hybrid system, and anti-tag immunoprecipitation coupled with tandem MS characterization [42]. Genome-wide studies of the physical interactome have reached a particularly advanced stage in biological systems of lesser complexity than plants [43,44]. Comprehensive physical interactomes based on the yeast two-hybrid system have been described for yeast, *Drosophila,* and *C. elegans* [45,46]. MS methods have also resulted in genome-scale descriptions of the yeast interactome [47]. To date, neither technique has been used for similar interactome research in plants but initial steps toward this goal are being taken, such as the identification and archiving of recombinant clones comprising full-length cDNAs for the model species *Arabidopsis* [48].

Published physical interactome work in plants has involved either large-scale *in silico* methods or gene family-level protein–protein interaction network studies. The *in silico* approach maps are called 'interologs' (interacting othologs). Interologs are the predicted interactive protein pairs, and they are identified by the sequence similarities to the proteins that are known to interact in the reference species such as yeast and human [49]. The *in silico* predicted interactome based on the interologs in *Arabidopsis* was developed as a potential guide map for a genome-level interactome in plants [50,51].

Despite the absence of an experimentally established global interactome for any plant species, studies focusing on local protein–protein interaction networks have been particularly productive. An interaction network based on the yeast two-hybrid system for essentially all *Arabidopsis* MADS box domain DNA-binding proteins revealed both specific heterodimeric and homodimeric interactions [52]. In a similar study, the interactions between myeloblastosis (MYB) protein and R/B-like basic-helix–loop-helix (BHLH) were characterized, which helped to distinguish the functions of different MYB proteins with similar sequences [53]. The research helped to characterize the functions of MYB DNA-binding proteins and BHLH transcriptional factors, both of which are involved in a variety of processes, including cell cycle, cell proliferation and cell lineage establishment [53]. Recently, a comprehensive study has helped to identify a protein–protein interaction network of over 70 proteins in wheat for abiotic stress response and development [54]. Besides the yeast two-hybrid system and anti-tag co-immunoprecipitation, protein microarrays were also used to probe the interaction between proteins globally in *Arabidopsis* [55].

Overall, local- and gene family-based interaction networks have shown much potential for a plant interactome to provide a global view of changes in biological processes, identify the key regulatory proteins and offer in-depth understanding of signal transduction. The potential of interactomes in understanding the systemic regulation of biological processes is unquestionable; however, the techniques currently used to build the interactome, especially the physical interactome, often identify false-positive interactions. Therefore, biochemical methods such as the fluorescence resonance energy transfer (FRET) assay are used to validate the interactomic methods [56]. However, both the FRET assay and the yeast two-hybrid system are

in vitro methods, and only co-immunoprecipitation can be performed *in vivo*, enabling protein–protein interactions to be characterized directly in plants of interest. Because each platform has limitations, using two or more platforms should yield more reliable interactome models compared with any single platform.

Transcriptional regulatory networks

Whereas the interactome describes protein–protein interactions, the transcription regulatory network describes the interaction between transcriptional regulatory genes and downstream genes, as shown in Figure 1c [17,57]. The transcriptional regulatory network was one of the first types to be constructed in molecular and cellular biology [17]. The data are often based on interactions between regulatory genes and downstream genes, as defined by mutant studies, global gene expression profiling, computational prediction of *cis*-elements and protein–DNA interaction studies using gel-shift assays. Examples of such networks, which also can incorporate information from the existing literature, have been constructed to infer signaling events involved in plant development, defense and physiological changes [58–60]. Recently, several computational tools and databases have been developed to identify the interaction between transcription factors and *cis*-elements, an important component in the construction of transcriptional regulatory networks [25,61,62].

This approach has recently been used to study transcription factors involved in glucosinolate biosynthesis [25]. Glucosinolate biosynthesis genes were first identified through meta-analysis of more than 1000 microarray experiments in public databases by building a gene-to-gene interaction network [25]. The mutants for candidate *myb* genes were then analyzed by DNA microarray experiments to identify the downstream genes controlled by the *myb* genes. A transcriptional regulatory network for glucosinolate biosynthesis genes was made from gene profiling data, mutant analysis and previous publications [25]. *Myb28*, *Myb29* and *Myb34* were identified to be the upstream transcription factors regulating glucosinolate biosynthesis, and all three genes were found to be downregulated in response to sulfur deficiency, indicating that plants mounted an effort to relocate the limited sulfur resources to produce sulfur-containing amino acids and glutathione biosynthesis [25]. The research highlights the potential of using transcriptional networks to describe systems-level regulation of a biological process in addition to discovering key transcription factors [25].

Gene regulatory networks

A gene regulatory network describes how genes interact with one another during the biological process to perform a function [16,20]. The relationships among the genes in the network can be defined in terms of activation, repression and other types of functional interactions. This type of network is more general than that described earlier because it incorporates post-transcriptional events such as protein targeting and covalent protein modification. For the purposes of mathematical modeling, the relationship between interacting genes is almost always simplified to activation and repression. Gene regulatory networks are often derived from a computational model based on previously published gene expression and mutant study data, and the model is subsequently validated and refined by perturbing the system, either using gene knockouts or specific biotic or abiotic treatments [16,20]. Network visualizations typically represent genes as 'nodes', and their relationships, simplified to activation or repression, as lines, as shown in Figure 1d.

One important benefit of the gene regulatory network is its capacity to describe the dynamics of a biological system. Discrete or continuous models are often employed to represent the changes during a process or a period of time. Unlike the other networks, which are static, gene regulatory network models are often dynamic [63].

Gene regulatory networks have been constructed to study several different developmental and physiological processes in plants [20,63–66]. For example, a discrete dynamic model has been developed that integrates previously published gene expression and genetic data during flower development in *Arabidopsis* [64]. The results implied that cell fate in flower development is not controlled by a precise signaling pathway, but rather by the dynamics of the gene network, reflected by a balance of effects provided by different genes [64]. In another example, gene regulation network analysis was also used to model the essential components controlling guard cell size [63]. In this study, the network organization data were first extracted from previous publications to construct the network structure. The probability of guard cell closure could be predicted through a dynamic model with Boolean (logic) functions when the network system is disturbed using external stimulations such as abscisic acid treatments. The final model provided a comprehensive overview of genes interactively involved in controlling stomata, and was validated by accurate prediction of the probability of guard cell closure when the network system was disturbed by external stimulation [63]. Largely, a gene regulatory network represents a viable approach to explain gene function in biological processes in a dynamic fashion.

Overall, biological network construction has been the most popular approach in systems biology to describe the physiology of an organism or a biological process. The advantages of networks lie in their capacity to understand the inherent biological complexity by identifying key components and interactions for system regulation. However, it is important to recognize that a network often has assumptions for data modeling, and thus cannot represent the exact biological system, no matter how perfectly it was built up. Despite the limitations, networks have provided new perspectives for interpreting omics data and have been applied to study a variety of plant biological questions.

Advances in plant systems biology – a biological perspective

There are several practical problems plaguing agriculture that are being addressed using systems biology. Examples of areas in plant science that have been addressed using systems biology are quantitative traits and plant stress and defense.

Systemic responses to abiotic stress

All four types of networks have been used to study plant responses to abiotic stress. The gene-to-metabolite network has been used to study *Arabidopsis thaliana* responses to nutritional deficiency [21,22]. A protein–protein interaction network was used to identify the key groups of genes involved in abiotic stress responses and flowering control in wheat [54]. A transcriptional regulatory network was built to elucidate the molecular mechanisms of dehydration and cold stress responses in *Arabidopsis* [58]. A gene regulatory network was successfully applied to explain the dynamics of aperture opening and closure in *Arabidopsis* in response to environmental stimuli [63]. In addition, a combined study of transcriptomics, quantitative trait loci (QTLs), mutant studies and yeast two-hybrid assays has helped to characterize genes involved in stress response and seed germination in rice [54,67]. Using systems biology platforms, plant stress response networks and dynamics were generated, which enabled the discovery of several important plant stress response genes [67]. These examples illustrate the power of using different types of networks to understand complex responses to stress, especially indicating how stresses induce the same and related genes.

Defining defense systems

Relatively few reports have been published about systemic modeling of network dynamics involved in plant defense. The capacity of systems biology for dissecting plant defense responses was manifested by a recent study that defined the 'regulatory node' in the transcriptional regulatory network controlling systemic acquired resistance (SAR) [60]. SAR is a mechanism that plants use to defend themselves from disease by mounting a systemic response after an initial pathogen infection. The research combined mutant studies and global gene expression profiling to identify several key *WRKY* genes as the regulatory node downstream of the nonexpresser of pathogen responsiveness gene 1 (*NPR1*), the key salicylic acid pathway regulator during SAR [60]. The research helped to define the roles of different *WRKY* genes in SAR, including establishing *WRKY18* as a crucial positive regulator [60]. Systems biology should be valuable for gaining a better understanding of plant defense, which often involves coordinative changes in secondary metabolite abundance and gene expression [19,25,27]. In particular, we expect that gene-to-metabolite networks will enable a better understanding of secondary metabolite synthesis, transport, dynamics and regulation during the defense process.

Modeling multigenic traits

The need to use systems biology to model multigenic traits is obvious and has been reviewed and discussed previously [68,69]. QTLs can be modeled using phenotypes under different environmental conditions to derive genotype-to-environmental interactions, in which one or more portions of the genome is associated with a phenotype [70–72]. The modeling of QTLs using the systems biology approach has helped to predict multigenic traits such as leaf growth and nitrogen accumulation in maize grain [73]. This approach is not as facile when gene, protein and metabolite information is limited. Comprehensive models using QTL, gene, protein, metabolite and phenotype information integrated will enable more-efficient breeding programs to improve traits in crop species such as maize, soybean and tomato. The so-called expression QTL and metabolite QTL can be used to integrate QTL information with gene expression and metabolite profiling, respectively, and they have already been used for tomato improvement [74] and the exploration of the genetic basis for metabolite diversity across different *Arabidopsis* lines [74,75]. Multigenic trait dissection might be the ultimate practical use of systems biology in crop science. Systems biology is ideally suited to detangle complex interactions that are defined by multigenic traits.

Data integration and the virtual plant project

Beyond all these applications, a more ambitious 'virtual plant project' (http://www.virtualplant.org) aims to use systems biology ultimately to generate dynamic models of the plant itself to describe the biological processes at molecular, cellular, physiological, organismal and ecological levels [76]. *Arabidopsis* has become the default choice for the virtual plant effort, from which an influx of data from genome, transcriptome, proteome, metabolome, interactome and phenome (Box 1) levels will be integrated and modeled to build a multi-network [76]. Towards that end, a virtual peach fruit model was built to identify the optimal conditions for fruit maturation, and many environmental factors and plant features, including metabolite information, were integrated in the model, which led to a reasonable prediction of fruit quality [77]. The integration of the virtual fruit model with transcriptome and metabolome information is expected to accelerate breeding for tomato fruit improvement [28].

The promises and challenges of plant systems biology

There are three domains that must be addressed to take full advantage of plant systems biology: (i) omics technology development; (ii) data integration into usable formats and (iii) data analysis within the domain of bioinformatics. Among these, bioinformatics probably needs the most attention because it is essential that biological data be normalized, standardized and visualized to build integrated models [78–80] (Figure 2). The limitation of systems biology is greatly tied to data modeling, in which analysis always involves generalization, simplification and assumptions. Therefore, networks in systems biology might never completely represent the actual biological system. The interpretation of systems biology data needs to be ever cautious and systems biology approaches can be complemented by reductionist approaches. Nevertheless, plant systems biology has come of age, given its potential to provide crucial understanding of the regulatory networks controlling developmental, physiological and pathological processes in plants. Advances in plant systems biology are needed to take full advantage of the ever-increasing numbers of omics technologies. Advances are needed for both basic plant biology and also the discovery of the key regulatory genes for agricultural purposes. We currently ask crops to serve more purposes than ever before and we

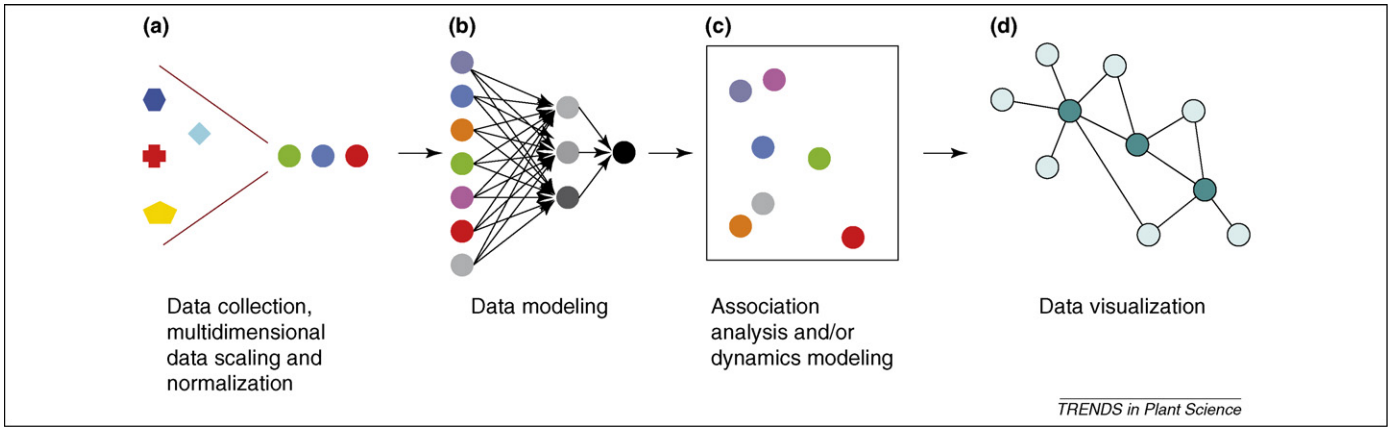

Figure 2. Bioinformatics applications in systems biology. **(a)** The normalization and scaling of the multidimensional omics data. Data from the different omics technologies are expected to have diverse formats and numerical scales of data, and thus should be considered to be multidimensional – that is, beyond two or three dimensions that can be simply visualized. To perform informative analyses, the multidimensional data from different omics techniques should first be integrated and normalized. **(b)** Data modeling for normalized data. **(c)** Constructing interactions or dynamics based on the data modeling. **(d)**. Data visualization. From steps (b) to (d), the normalized data are modeled to estimate the level of interaction and association, and are then visualized. All steps might not be necessary for a given systems biology study. In (a), the different shapes indicate the different types of biological data; in panels (a), (b) and (c), the different-colored circles symbolize the variables representing the data; and (d), the circles represent the elements in the model. Lines represent relationships among the elements.

also expect rapid genetic improvement. By understanding systems integration, we should be able to accelerate crop adaptation for food, feed, biofuels, and industrial and pharmaceutical production.

References

1 Trewavas, A. (2006) A brief history of systems biology. 'Every object that biology studies is a system of systems.' Francois Jacob (1974). *Plant Cell* 18, 2420–2430
2 Kirschner, M.W. (2005) The meaning of systems biology. *Cell* 121, 503–504
3 Schauer, N. and Fernie, A.R. (2006) Plant metabolomics: towards biological function and mechanism. *Trends Plant Sci.* 11, 508–516
4 Uhrig, J.F. (2006) Protein interaction networks in plants. *Planta* 224, 771–781
5 Glinski, M. and Weckwerth, W. (2006) The role of mass spectrometry in plant systems biology. *Mass Spectrom. Rev.* 25, 173–214
6 Ahmadian, A. *et al.* (2006) Pyrosequencing: history, biochemistry and future. *Clin. Chim. Acta* 363, 83–94
7 Galbraith, D.W. (2006) DNA microarray analyses in higher plants. *OMICS* 10, 455–473
8 Kris, R.M. *et al.* (2007) High-throughput, high-sensitivity analysis of gene expression in *Arabidopsis*. *Plant Physiol.* 144, 1256–1266
9 Boone, C. *et al.* (2007) Exploring genetic interactions and networks with yeast. *Nat. Rev. Genet.* 8, 437–449
10 Katagiri, F. (2003) Attacking complex problems with the power of systems biology. *Plant Physiol.* 132, 417–419
11 Sweetlove, L.J. *et al.* (2003) Predictive metabolic engineering: a goal for systems biology. *Plant Physiol.* 132, 420–425
12 Girke, T. *et al.* (2003) Towards a modeling infrastructure for studying plant cells. *Plant Physiol.* 132, 410–414
13 Nemhauser, J.L. *et al.* (2003) Building integrated models of plant growth and development. *Plant Physiol.* 132, 436–439
14 Thum, K.E. *et al.* (2003) Light- and carbon-signaling pathways. Modeling circuits of interactions. *Plant Physiol.* 132, 440–452
15 Alonso, J.M. and Ecker, J.R. (2006) Moving forward in reverse: genetic technologies to enable genome-wide phenomic screens in *Arabidopsis*. *Nat. Rev. Genet.* 7, 524–536
16 Alvarez-Buylla, E.R. *et al.* (2007) Gene regulatory network models for plant development. *Curr. Opin. Plant Biol.* 10, 83–91
17 Belostotsky, D.A. and Rose, A.B. (2005) Plant gene expression in the age of systems biology: integrating transcriptional and post-transcriptional events. *Trends Plant Sci.* 10, 347–353
18 Palme, K. (2006) Towards plant systems biology – novel mathematical approaches to enable quantitative analysis of growth processes. *New Phytol.* 171, 443–444
19 Sweetlove, L.J. and Fernie, A.R. (2005) Regulation of metabolic networks: understanding metabolic complexity in the systems biology era. *New Phytol.* 168, 9–23
20 Ueda, H.R. (2006) Systems biology flowering in the plant clock field. *Mol Syst Biol* 2, 60
21 Hirai, M.Y. *et al.* (2004) Integration of transcriptomics and metabolomics for understanding of global responses to nutritional stresses in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U. S. A.* 101, 10205–10210
22 Nikiforova, V.J. *et al.* (2005) Integrative gene-metabolite network with implemented causality deciphers informational fluxes of sulphur stress response. *J. Exp. Bot.* 56, 1887–1896
23 Hirai, M.Y. *et al.* (2005) Elucidation of gene-to-gene and metabolite-to-gene networks in *Arabidopsis* by integration of metabolomics and transcriptomics. *J. Biol. Chem.* 280, 25590–25595
24 Tohge, T. *et al.* (2005) Functional genomics by integrated analysis of metabolome and transcriptome of *Arabidopsis* plants over-expressing an MYB transcription factor. *Plant J.* 42, 218–235
25 Hirai, M.Y. *et al.* (2007) Omics-based identification of *Arabidopsis* Myb transcription factors regulating aliphatic glucosinolate biosynthesis. *Proc. Natl. Acad. Sci. U. S. A.* 104, 6478–6483
26 Goossens, A. *et al.* (2003) A functional genomics approach toward the understanding of secondary metabolism in plant cells. *Proc. Natl. Acad. Sci. U. S. A.* 100, 8595–8600
27 Zulak, K.G. *et al.* (2007) Gene transcript and metabolite profiling of elicitor-induced opium poppy cell cultures reveals the coordinate regulation of primary and secondary metabolism. *Planta* 225, 1085–1106
28 Carrari, F. *et al.* (2006) Integrated analysis of metabolite and transcript levels reveals the metabolic shifts that underlie tomato fruit development and highlight regulatory aspects of metabolic network behavior. *Plant Physiol.* 142, 1380–1396
29 Rischer, H. *et al.* (2006) Gene-to-metabolite networks for terpenoid indole alkaloid biosynthesis in *Catharanthus roseus* cells. *Proc. Natl. Acad. Sci. U. S. A.* 103, 5614–5619
30 Wei, H. *et al.* (2006) Transcriptional coordination of the metabolic network in *Arabidopsis*. *Plant Physiol.* 142, 762–774
31 Usadel, B. *et al.* (2005) Extension of the visualization tool MapMan to allow statistical analysis of arrays, display of corresponding genes, and comparison with known responses. *Plant Physiol.* 138, 1195–1204
32 Blasing, O.E. *et al.* (2005) Sugars and circadian regulation make major contributions to the global regulation of diurnal gene expression in *Arabidopsis*. *Plant Cell* 17, 3257–3281
33 Scheible, W.R. *et al.* (2004) Genome-wide reprogramming of primary and secondary metabolism, protein synthesis, cellular growth processes, and the regulatory infrastructure of *Arabidopsis* in response to nitrogen. *Plant Physiol.* 136, 2483–2499

34 Gibon, Y. *et al.* (2006) Integration of metabolite with transcript and enzyme activity profiling during diurnal cycles in *Arabidopsis* rosettes. *Genome Biol.* 7, R76
35 Hesse, H. and Hoefgen, R. (2006) On the way to understand biological complexity in plants: S-nutrition as a case study for systems biology. *Cell. Mol. Biol. Lett.* 11, 37–56
36 Hesse, H. and Hoefgen, R. (2006) On the way to understand biological complexity in plants: S-nutrition as a case study for systems biology. *Cell. Mol. Biol. Lett.* 11, 37–55
37 Tong, A.H.Y. *et al.* (2001) Systematic genetic analysis with ordered arrays of yeast deletion mutants. *Science* 294, 2364–2368
38 Tong, A.H.Y. *et al.* (2004) Global mapping of the yeast genetic interaction network. *Science* 303, 808–813
39 Lehner, B. *et al.* (2006) Systematic mapping of genetic interactions in *Caenorhabditis elegans* identifies common modifiers of diverse signaling pathways. *Nat. Genet.* 38, 896–903
40 To, J.P.C. *et al.* (2004) Type-A *Arabidopsis* response regulators are partially redundant negative regulators of cytokinin signaling. *Plant Cell* 16, 658–671
41 Riefler, M. *et al.* (2006) *Arabidopsis* cytokinin receptor mutants reveal functions in shoot growth, leaf senescence, seed size, germination, root development, and cytokinin metabolism. *Plant Cell* 18, 40–54
42 Fields, S. and Song, O.K. (1989) A novel genetic system to detect protein-protein interactions. *Nature* 340, 245–246
43 Ito, T. *et al.* (2001) A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proc. Natl. Acad. Sci. U. S. A.* 98, 4569–4574
44 Uetz, P. *et al.* (2000) A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*. *Nature* 403, 623–627
45 Giot, L. *et al.* (2003) A protein interaction map of *Drosophila melanogaster*. *Science* 302, 1727–1736
46 Li, S. *et al.* (2004) A map of the interactome network of the metazoan *C. elegans*. *Science* 303, 540–543
47 Ho, Y. *et al.* (2002) Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature* 415, 180–183
48 Bevan, M. and Walsh, S. (2005) The *Arabidopsis* genome: a foundation for plant research. *Genome Res.* 15, 1632–1642
49 Matthews, L.R. *et al.* (2001) Identification of potential interaction networks using sequence-based searches for conserved protein-protein interactions or 'interologs'. *Genome Res.* 11, 2120–2126
50 Geisler-Lee, J. *et al.* (2007) A predicted interactome for *Arabidopsis*. *Plant Physiol.* 145, 317–329
51 Yu, H. *et al.* (2004) Annotation transfer between genomes: protein-protein interologs and protein-DNA regulogs. *Genome Res.* 14, 1107–1118
52 de Folter, S. *et al.* (2005) Comprehensive interaction map of the *Arabidopsis* MADS box transcription factors. *Plant Cell* 17, 1424–1433
53 Zimmermann, I.M. *et al.* (2004) Comprehensive identification of *Arabidopsis thaliana* MYB transcription factors interacting with R/B-like BHLH proteins. *Plant J.* 40, 22–34
54 Tardif, G. *et al.* (2007) Interaction network of proteins associated with abiotic stress response and development in wheat. *Plant Mol. Biol.* 63, 703–718
55 Popescu, S.C. *et al.* (2007) Differential binding of calmodulin-related proteins to their targets revealed through high-density *Arabidopsis* protein microarrays. *Proc. Natl. Acad. Sci. U. S. A.* 104, 4730–4735
56 You, X. *et al.* (2006) Intracellular protein interaction mapping with FRET hybrids. *Proc. Natl. Acad. Sci. U. S. A.* 103, 18458–18463
57 Gilchrist, M. *et al.* (2006) Systems biology approaches identify ATF3 as a negative regulator of Toll-like receptor 4. *Nature* 441, 173–178
58 Yamaguchi-Shinozaki, K. and Shinozaki, K. (2006) Transcriptional regulatory networks in cellular responses and tolerance to dehydration and cold stresses. *Annu. Rev. Plant Biol.* 57, 781–803
59 Serna, L. and Martin, C. (2006) Trichomes: different regulatory networks lead to convergent structures. *Trends Plant Sci.* 11, 274–280
60 Wang, D. *et al.* (2006) A genomic approach to identify regulatory nodes in the transcriptional network of systemic acquired resistance in plants. *PLoS Pathogens* 2, 1042–1050
61 Davuluri, R.V. *et al.* (2003) AGRIS: *Arabidopsis* Gene Regulatory Information Server, an information resource of *Arabidopsis* cis-regulatory elements and transcription factors. *BMC Bioinformatics* 4, 25
62 Palaniswamy, S.K. *et al.* (2006) AGRIS and AtRegNet. A platform to link cis-regulatory elements and transcription factors into regulatory networks. *Plant Physiol.* 140, 818–829
63 Li, S. *et al.* (2006) Predicting essential components of signal transduction networks: a dynamic model of guard cell abscisic acid signaling. *PLoS Biol.* 4, 1732–1748
64 Espinosa-soto, C. *et al.* (2004) A gene regulatory network model for cell-fate determination during *Arabidopsis thaliana* flower development that is robust and recovers experimental gene expression profiles. *Plant Cell* 16, 2923–2939
65 Zeilinger, M.N. *et al.* (2006) A novel computational model of the circadian clock in *Arabidopsis* that incorporates PRR7 and PRR9. *Mol. Sys. Biol.* 2, 58
66 Locke, J.C.W. *et al.* (2006) Experimental validation of a predicted feedback loop in the multi-oscillator clock of *Arabidopsis thaliana*. *Mol. Sys. Biol.* 2, 59
67 Cooper, B. *et al.* (2003) A network of rice genes associated with stress response and seed development. *Proc. Natl. Acad. Sci. U. S. A.* 100, 4945–4950
68 Hammer, G. *et al.* (2006) Models for navigating biological complexity in breeding improved crop plants. *Trends Plant Sci.* 11, 587–593
69 Struik, P.C. *et al.* (2005) Complex quality traits: now time to model. *Trends Plant Sci.* 10, 513–516
70 Yin, X. *et al.* (2004) Role of crop physiology in predicting gene-to-phenotype relationships. *Trends Plant Sci.* 9, 426–432
71 Zhao, W. *et al.* (2005) A non-stationary model for functional mapping of complex traits. *Bioinformatics* 21, 2469–2477
72 Wu, R. and Lin, M. (2006) Functional mapping – how to map and study the genetic architecture of dynamic complex traits. *Nat. Rev. Genet.* 7, 229–237
73 Reymond, M. *et al.* (2003) Combining quantitative trait loci analysis and an ecophysiological model to analyze the genetic variability of the responses of maize leaf growth to temperature and water deficit. *Plant Physiol.* 131, 664–675
74 Schauer, N. *et al.* (2006) Comprehensive metabolic profiling and phenotyping of interspecific introgression lines for tomato improvement. *Nat. Biotechnol.* 24, 447–454
75 Keurentjes, J.J.B. *et al.* (2006) The genetics of plant metabolism. *Nat. Genet.* 38, 842–849
76 Gutierrez, R.A. *et al.* (2005) Systems biology for the virtual plant. *Plant Physiol.* 138, 550–554
77 Lescourret, F. and Genard, M. (2005) A virtual peach fruit model simulating changes in fruit quality during the final stage of fruit growth. *Tree Physiol.* 25, 1303–1315
78 Tokimatsu, T. *et al.* (2005) KaPPA-view: a web-based analysis tool for integration of transcript and metabolite data on plant metabolic pathway maps. *Plant Physiol.* 138, 1289–1300
79 Bornholdt, S. (2005) Less is more in modeling large genetic networks. *Science* 310, 449–451
80 Di Ventura, B. *et al.* (2006) From *in vivo* to *in silico* biology and back. *Nature* 443, 527–533
81 Hall, R.D. (2006) Plant metabolomics: from holistic hope, to hype, to hot topic. *New Phytol.* 169, 453–468
82 Salt, D.E. (2004) Update on plant ionomics. *Plant Physiol.* 136, 2451–2456
83 Sekiyama, Y. and Kikuchi, J. (2007) Towards dynamic metabolic network measurements by multi-dimensional NMR-based fluxomics. *Phytochemistry* 68, 2320–2329
84 Turnbull, J.E. and Field, R.A. (2007) Emerging glycomics technologies. *Nat. Chem. Biol.* 3, 74–77