

UC Davis

UC Davis Previously Published Works

Title

Plant Networks as Traits and Hypotheses: Moving Beyond Description

Permalink

<https://escholarship.org/uc/item/99z8c7mn>

Journal

Trends in Plant Science, 24(9)

ISSN

1360-1385

Authors

Marshall-Colón, Amy
Kliebenstein, Daniel J

Publication Date

2019-09-01

DOI

10.1016/j.tplants.2019.06.003

Peer reviewed

Review

Plant Networks as Traits and Hypotheses: Moving Beyond Description

Amy Marshall-Colón¹ and Daniel J. Kliebenstein^{2,3,*}

Biology relies on the central thesis that the genes in an organism encode molecular mechanisms that combine with stimuli and raw materials from the environment to create a final phenotypic expression representative of the genomic programming. While conceptually simple, the genotype-to-phenotype linkage in a eukaryotic organism relies on the interactions of thousands of genes and an environment with a potentially unknowable level of complexity. Modern biology has moved to the use of networks in systems biology to try to simplify this complexity to decode how an organism's genome works. Previously, biological networks were basic ways to organize, simplify, and analyze data. However, recent advances are allowing networks to move beyond description and become phenotypes or hypotheses in their own right. This review discusses these efforts, like mapping responses across biological scales, including relationships among cellular entities, and the direct use of networks as traits or hypotheses.

Biological Networks

The modern molecular synthesis proposes that the genes in an organism encode molecular mechanisms that combine with signals and raw materials from the environment to create the final phenotypic expression. This genotype-to-phenotype linkage is conceptually simple, but in reality a eukaryotic organism has thousands of genes and the environment has an unknown and potentially unknowable level of complexity. This complexity has led to current efforts to utilize systems analysis to decode how an organism's genome works to create the optimal phenotype for a given environment. One simplification in the systems biology toolkit is to use biological networks to organize, simplify, and analyze data. These networks can be used to map responses across biological scales, including relationships among cellular entities, such as protein–protein interactions, gene–gene coexpression, and protein:DNA binding. The description of biological processes with networks has uncovered regulatory factors that exert control over many biological processes [1–3], regulatory motifs that shed light on signal transduction [4–7], and putative roles for genes of unknown function [8–10]. The construction of gene regulatory networks (GRNs) provides insight about the regulation – direct, indirect, or hierarchical – in a network by layering DNA-binding data (predicted or observed) onto coexpression networks. However, the predominant use of biological networks has been to describe relationships among molecules at a snapshot in time or to focus on the potential role of hub genes, yielding important insights but with significant limitations [4]. For example, downstream functional analysis of hub genes identified by network analysis can result in no obvious physiological phenotype, potentially due to functional redundancy, or the network may not translate to another experiment [11]. While some of this difficulty is caused by the available data being limiting for complete network inference, there is a key need to improve our ability to derive functional and predictive information from any given network. The role of data limitation is covered extensively in all of the citations in this review. Here, we cover recent advances in deriving novel predictive insight from networks across different scales and the use of networks as direct measures of traits to quantify complex processes.

Highlights

After a number of years, systems biology and networks are being transitioned from generating hypotheses to being used to directly test hypotheses.

Numerous statistical methods are now available to create predictive and testable hypotheses from a wide array of omics and systems data.

Eigengenes allows the use, measuring, and testing of networks directly as any physiological, developmental trait.

Community solutions surrounding sharing and visualizing networks are essential to allow the broader utilization of these approaches.

¹Department of Plant Biology, University of Illinois at Urbana–Champaign, Urbana, IL 61801, USA

²Department of Plant Sciences, University of California, Davis, One Shields Avenue, Davis, CA 95616, USA

³DynaMo Center of Excellence, University of Copenhagen, Thorvaldsensvej 40, DK-1871 Frederiksberg C, Denmark

*Correspondence:
Kliebenstein@ucdavis.edu
(D.J. Kliebenstein).

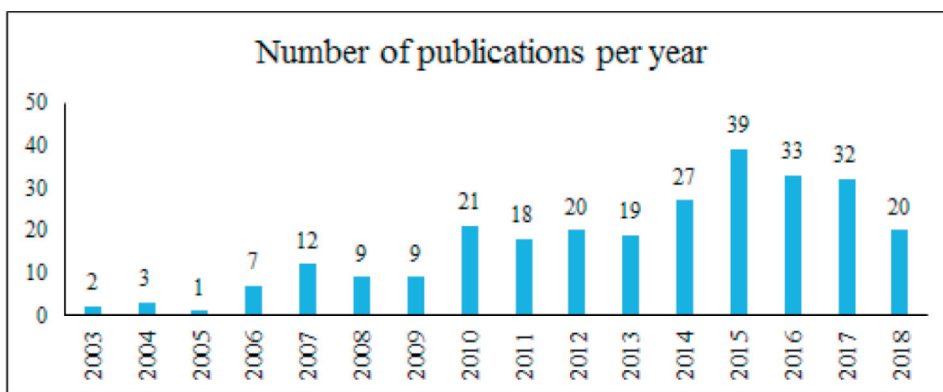
Time Course and Dynamics

The first biological networks were derived from one or a few conditions with little or no temporal analysis. This discord with the temporal nature of any response in an organism led researchers to propose that extensive time-course datasets would provide more informative/predictive networks (Figure 1). The increasing availability and decreasing cost of new experimental technologies are expediting this shift. While the process of generating time-series omics data is straightforward, the construction of time-resolved networks presents new analytical and computational challenges. Statistics, machine learning (ML), and calculus-based computational strategies have all been used to build temporal biological networks. Many of these strategies can successfully be applied to developmental and chronological time-series data [12–15]. In this section, we explore different computational strategies used by the community to generate and interpret time-resolved biological networks.

Statistical Approaches

Temporal regulatory networks are often built on simple correlation methodologies due to simplicity in implementation and interpretation. Correlation methods are used to cluster temporal gene expression patterns and identify network modules of coexpressed and coregulated genes [10, 16, 17]. A GRN of the jasmonic acid (JA) signaling network was built by connecting transcription factors (TFs) to gene clusters based on the temporal phase when they are first differentially expressed (DE) [10]. Network analysis revealed temporal phases of activation and repression, where biological processes associated with methyl jasmonate response were initiated at different times after the signal was perceived. While simple correlation is informative and useful for hypothesis generation, the goal of time-series network analysis is to identify causal relationships. Time-lag Pearson correlation is a statistical approach that utilizes the time lag between cause and effect to identify potential causality. A study of temporal nitrogen (N) deprivation in *Chlamydomonas* used the time-lag Pearson approach to identify potential TFs or other regulators that control the expression of target genes and/or metabolites in response to N deprivation [12].

Sometimes correlation-based methods are a first step in an analysis pipeline to create causal hypotheses about gene relationships [18, 19]. WIGWAMS is a computational method based on



Trends in Plant Science

Figure 1. Plant Publications Involving Time-Resolved Networks of Omics-Scale Data. Data were derived by searching PubMed using the following search terms: "(((("time series" OR "time course" OR temporal) AND (network OR model) AND (apple OR arabidopsis OR crop OR maize OR soybean OR wheat OR rice OR corn OR poplar OR sorghum OR brassica OR tobacco OR tomato OR petunia OR nicotiana) AND (transcriptome OR "RNA-seq" OR eQTL OR "high throughput" OR "metabolome" OR metabolomics OR transcriptomics OR genomics OR proteomics OR proteome OR genome))) NOT "review"[Publication Type])" on 1 October 2019. Pruning of the results was done by manual curation.

Pearson correlation that takes advantage of time-series data to find gene modules that are dependently coexpressed (Table 1) [19]. WIGWAMS investigates DE genes by searching for time lags to identify genes that regulate entire modules and was used to identify biologically relevant modules across six time-series datasets. Another method, Causal structure identification (CSI), was used to infer a regulatory network of the arabidopsis response to *Botrytis cinerea* infection [18]. First, temporal gene expression clusters were identified using correlation, then a GRN was inferred by correlating cluster means with *B. cinerea* growth. This identified a temporal pattern of transcriptional response to *B. cinerea* infection, where most gene clusters are “turned on” downstream of *B. cinerea* growth while one gene cluster was upstream of the *B. cinerea* growth curve. Exploration of TFs in each of these network modules uncovered their predicted roles in plant response to infection, such as *TGA3* directly regulating the expression of 193 genes.

Integrating the direct measurement of pathogen growth into the plant response network analysis greatly increased the information obtained from the network. Similarly, Greenham *et al.* integrated gene networks with physiological data to investigate temporal changes in the arabidopsis transcriptome in response to drought [17]. Weighted gene correlation network analysis (WGCNA) was used to create gene modules and eigengenes (Box 1 and Figure 2) to link with physiological measurements over a drought time course. Modules linked to perturbed circadian expression had positive or negative correlations with F_v'/F_m' (maximum efficiency of photosystem II), g_s (stomatal conductance), and nonstructural carbohydrate (NSC) measurements. The drought-related eigengenes contained drought-responsive genes involved in the abiotic stress response, photosynthesis, the light response, glucosinolate biosynthesis, amino acid biosynthesis, phosphatase activity, and nitrogen metabolism. In both cases, temporal resolution aided the ability to infer the link between the dynamic transcriptome and the physiological responses.

A dynamic Bayesian network (DBN) can model multivariate time series that are often too complex for other types of gene network modeling and can detect time-series anomalies and hidden patterns. A DBN approach called Metropolis variational Bayesian state space modeling (M-VBSSM) was used to identify hub genes involved in the gradual drought response in arabidopsis [20]. Changes to the transcriptome, metabolome, and physiology were measured in response to the onset of drought, and the GRN identified regulatory genes involved in the perception and signaling of drought stress.

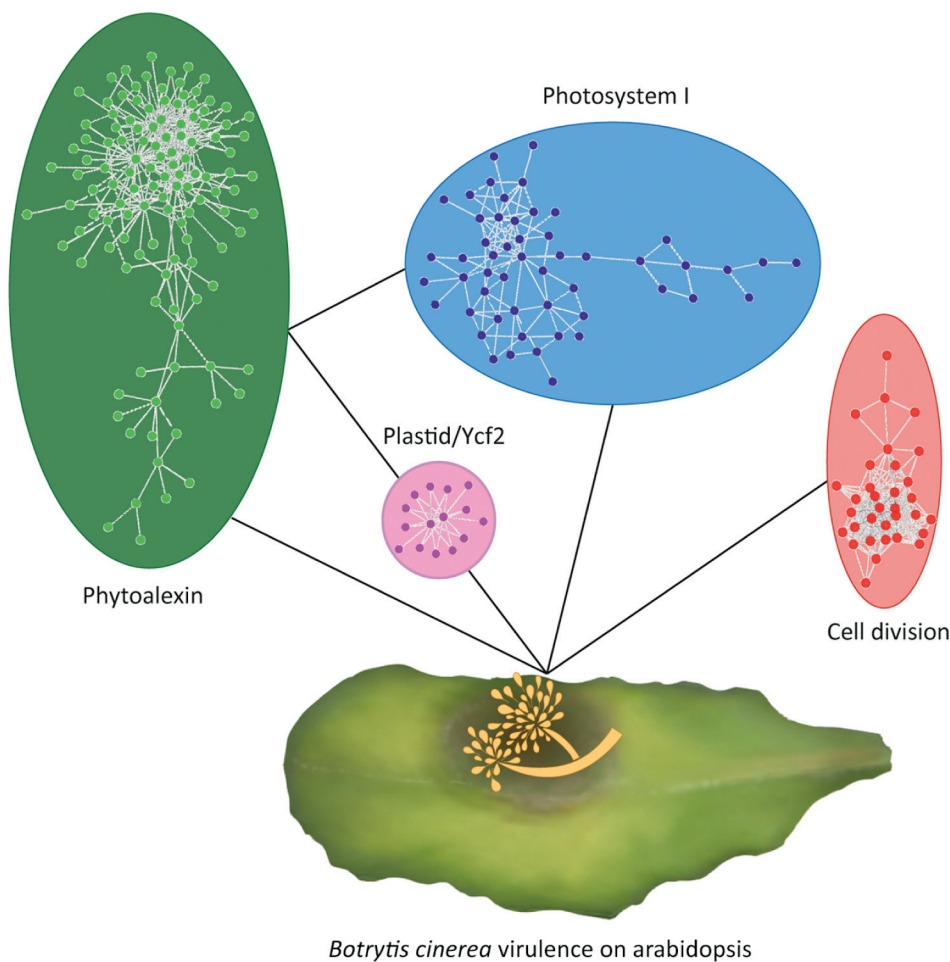
Table 1. Mathematical Approaches for Building Biological Networks

Model type		Description	Refs
Statistics	Correlation	Dependencies between expression patterns are found by correlating variables using methods like Pearson's correlation coefficient, Spearman's rank correlation, and partial correlations; these methods are often used to build coexpression networks based on hard thresholding	[4]
	WGCNA	WGCNA uses unsupervised correlation analysis and soft thresholding to convert coexpression measures in adjacency matrices to a connection weight; key features of WGCNA networks are modules that can be used as module eigengenes (Box 1)	[17,89]
	Bayesian inference	A probabilistic and conditional method that statistically addresses direct and indirect causal effects in a network; conditional dependencies are represented as directed acyclic graphs	[102,103]
ML	Neural networks	Algorithmic approach designed to recognize patterns through unsupervised clustering followed by supervised classification	[28,29,104]
	State-space models (i.e., DFG)	Model the joint probabilities between hidden and observed variables, where the state at a given time depends on the states and observations over many past time steps	[23,24,105]
Algebra and calculus	Boolean	Algebraic method that uses logic operators such as (and, or, not) as rules that govern node states in a network, resulting in a directed graph	[106]
	ODEs	Calculus-based method to model time-varying or continuous interactions in nonlinear systems; this method is powerful by being mechanistic and predictive, but is limited to small networks	[107,108]

Box 1. Eigengenes: Creation and Use

One method to translate a network into a tool to quantitatively measure phenotypes is the eigengene. The eigengene concept has been popularized as an extension of the WGCNA where it aids in the summarization of highly correlated clusters of genes. An eigengene is a vector that comprises a series of transcripts and associated constants by which those transcripts' accumulation values are multiplied. It is possible to obtain a single estimate of how these genes behave by taking the transcripts in an eigengene, multiplying their accumulation by the associated constants, and adding up these values. This creates an eigenvalue for the compilation of genes in an eigengene that can then be used as any other quantitative measurement. For example, eigenvalues associated with module eigengenes can be correlated with non-gene expression data such as phenotypic data to identify links between transcript networks and possible outputs.

Models of subsets of genes were compared and the best model (high marginal likelihood) for interacting genes was chosen for downstream analysis. This led to the identification of *AGL22*, a unique hub TF involved in the regulation of drought perception and signaling [20].



Trends in Plant Science

Figure 2. Using Eigengenes to Link Networks and Traits. The ellipses define specific networks of genes whose transcripts are correlated with each other and share a common function as listed by the ellipse. By combining the expression values across these genes into single eigengene descriptions, it is possible to test how the networks correlate with each other and with a specific output trait, in this case the virulence of a specific pathogen. The lines between networks or traits show whether the eigengenes or traits were positively (blue) or negatively (red) correlated. The phytoalexin network comprises the biosynthetic and regulatory genes to make camalexin, photosystem I is the structural and enzyme genes for photosystem I, cell division is genes regulating foliar cell division, and the plastid/Ycf2 network is plastid-encoded genes that coexpress with the Ycf2/AtCg00860 gene.

Other approaches for modeling causal relationships in a time-resolved GRN rely on probability algorithms that are closer to ML than classic statistical models. The Signaling and Dynamic Regulatory Events Miner (SDREM) uses condition-specific time-series transcriptome data to reconstruct activated regulatory networks in response to a stress using temporal data to provide causal information [21]. The SDREM incorporates information about static protein–protein and protein–DNA interactions with time-series expression data to find condition-specific signaling pathways that activate known interactions. Regulatory bifurcation events in a time series, where two previously coexpressed genes diverge in expression pattern, are annotated with TFs predicted to commonly regulate those genes. This associates temporal information (time splits) with static protein–DNA interaction data, even if the predicted TF is not consistent with known signaling pathways. SDREM was used to model the arabidopsis temporal immune response to Hpa infection and predicted 83 interactions between signaling proteins and TFs with known biological relevance to Hpa infection response [21].

ML Approaches

ML algorithms are an alternative approach to infer causal relationships in temporal GRNs [22]. However, ML algorithms require high-quality training and testing data for effective learning. State-space models (SSMs) are one type of ML algorithm that can use both noisy-observed and noiseless-latent values of gene expression, where the objective is to solve for unobserved gene expression values at future time points based on previously observed gene expression values. A SSM was applied to learn the interactions between TFs and target genes over a nitrogen-response time course [23]. The SSM predicted the correct direction of change in gene expression values for numerous genes not included in the training set. This revealed the TF *SPL9* as a regulatory hub for N-responsive genes that played a role in attenuating the response to nitrate. A dynamic factor graph (DFG), a type of state-space model, was used to prune this GRN of the nitrogen response in whole plants by providing a measure of influence for each TF. This led to a more refined GRN that retained the top 10% of DFG-predicted interactions with high confidence [24]. The final dynamic GRN revealed new hub genes involved in the regulation of the nitrogen response, including CRF4, which was experimentally validated to directly regulate N uptake and assimilation.

ML algorithms can also incorporate multiple layers of information to generate dynamic GRNs [25]. Genome-scale measurements (RNA-seq, ATAC-seq, and TF binding motifs) and physiological data were used to learn GRNs associated with rice responses to environmental perturbations via environmental gene regulatory influence networks (EGRINs) [26]. The EGRIN was used to explore key regulators involved in rice responses to either high temperature or water deficit. Network inference was based on the inferelator algorithm [27] to construct a GRN based on TF activity and not based on coexpression. The network identified novel roles for heat shock factor (HSF) family TFs in regulating the abiotic stress response and circadian clock in rice. Similarly, an artificial neural network (ANN) approach was used to predict relationships among transcripts and metabolites in crops harboring introgressed exotic alleles [28]. The software omeSOM was developed to generate neural models called self-organizing maps (SOMs) used for the unsupervised clustering of metabolite profiles and gene expression in a population of introgression lines (ILs) from a cross between two domesticated tomato species. The SOMs were used to compare patterns of gene and/or metabolite expression across 21 ILs and group them based on molecular similarities. An ANN was recently used to build a GRN of the maize nitrogen response [29]. The stepwise screening approach identified a reduced network in which bZIP108 and WRKY36 were two key TFs influencing the N response.

Calculus and Algebra-Based Approaches

Once GRNs are developed, the goal is to move toward a mechanistic analysis. This often requires a shift to models based in calculus and algebra whose computational intensity limits the size of the

GRN. A dynamic GRN of the floral transition time in maize using ordinary differential equations (ODEs) was built for just four genes [30]. Network topology was derived from a review of the literature on flowering time in maize and the dynamic GRN was parameterized using expression data from different maize genotypes taken during the transition from vegetative to floral stages. The resulting model predicted mutant phenotypes with a high correlation ($R^2 = 0.87$). Another ODE-based dynamic GRN modeled nine genes regulated by the CHE TF in response to *B. cinerea* infection in arabidopsis [31]. To determine the directional influence of CHE on target genes, a dynamic linear 'gray box' model (one that includes *a priori* biological information) was built using a series of ODEs. Treating the GRN as a circuit revealed a number of possible network rewiring strategies that could be implemented using a synthetic biology approach to protect CHE from being downregulated on *B. cinerea* infection [31].

Boolean modeling is a mathematical approach that can somewhat overcome the restricted size of ODE models. A Boolean model of 29 regulatory interactions involved in cell cycle progression in arabidopsis uncovered novel regulatory interactions in the cell cycle that were experimentally validated using gene loss- and gain-of-function lines [32]. Modeling and experimental validation revealed a role for anaphase-promoting complex/cyclosome (APC/C), part of the cell cycle machinery, in mediating hormonal control of cell cycle dynamics. Likewise, Boolean modeling was used to build a 31-interaction hormonal regulatory network controlling cell differentiation in the root apical meristem of arabidopsis [33]. Another common strategy is to combine approaches to build dynamic GRNs, as was done to identify the GRN at the onset of leaf senescence in arabidopsis by combining Boolean and ODE models [34].

Identifying Networks across Molecular Scales

Time-course networks are often built by investigating one molecular class of traits, either transcriptomics or metabolomics. However, there is extensive interest in combining across molecular scales to better develop predictive phenotypic models. Multiscale models can more accurately simulate biological processes by representing the flow of information across biological levels. Network models are ideal for representing predicted relationships between cellular entities like genes, small RNAs, proteins, and metabolites that can be represented as nodes, while the edges mathematically define their relationship. Example relationships include known or predicted protein-protein interactions, gene coexpression, and regulatory binding information [35,36]. Multiscale modeling can be performed using a correlation approach [17,20], by designing a model that uses the output of one biological process as input for the next [37–39], or by using information across scales to constrain model parameters [40,41].

One area where integrative models have been widely utilized is in natural variation studies that move across molecular scales to identify loci that are causal for metabolic or other phenotypes. eQuantitative trait loci (eQTLs), mQTLs, and pQTLs are loci where genetic variation influences levels of transcripts, metabolites, and proteins, respectively. Blending them can identify causal loci that control variation across all of these levels [42–46]. GWA of six tocopherol compounds revealed association signals between α -tocopherol and the biosynthetic pathway gene *ZmVTE4* and between the compound tocotrienol and the tocopherol biosynthetic pathway gene *ZmVTE1* [47]. This showed that favorable alleles of the mQTLs may increase levels of vitamin E in maize. While a number of eQTL and mQTL studies have been performed in plants, pQTL analyses are less common but are a target for future studies. A recent pQTL analysis revealed chromosomal hotspots for quality traits for starch content and cold sweetening in potatoes [42].

Network analysis has also been used with genome-wide association studies (GWASs) to enhance the biological interpretation and prioritize candidate GWAS loci [48–57]. This approach is

gaining in popularity because GWAS results on their own do not provide the necessary context to interpret findings, as pathways affected by a genetic locus are often not immediately apparent. To aid in filtering this onslaught of information, metanetworks are generated by considering pairwise relationships between nodes representing cellular entities such as genes, proteins, and metabolites [48,49,53]. These networks are used to test for enrichments in the GWAS loci that may describe the overall mechanistic basis of the trait [55,56]. For example, in an analysis of quantitative resistance to *B. cinerea* in arabidopsis, it was shown that the vast majority of likely causal loci were in downstream physical and metabolic defenses rather than in the upstream signaling events [56]. Moving beyond these broader descriptions requires a rapid and efficient ability to identify true causal loci in GWASs. An efficient way to do this is to take the GWAS candidates and utilize previous network descriptions to identify genes that both associate with the phenotype and have a mutual network association. This can identify true causal loci with a >60% validation rate for plant metabolic, growth, and defense traits [55–57]. This has recently been implemented in a formalized architecture via coanalysis of molecular components (Camoco). This was developed to integrate gene coexpression networks with GWAS results to identify causal genes contributing to phenotypes of interest. Camoco was applied to maize to identify genes underlying the grain ionome [54]. The web application Arabidopsis Genome-Wide Association Boosting (araGWAB) overcomes weak phenotype-association signals from GWASs by integrating cofunctional gene network information [51].

Testing Network Structure Requires Factorial Genetic Perturbation

Time courses and multiscale integration will create theoretical maps of a network that then require validation. However, biological networks are assumed to be buffered against perturbations. This buffering is a characteristic of biological networks compared with random networks that can be generated by functional redundancy of genes and the presence of feedforward and feedback regulatory loops. Random networks often have an evenly distributed topology with an equal number of degrees across nodes. Biological networks display scale-free topology in which there are many nodes of low degree and few nodes of high degree [58]. While this may have a beneficial fitness effect, buffering makes the interior structure of a network difficult to ascertain using single perturbation, chemical, biotic, or mutational approaches. A solution to bypass this buffering is to incorporate combinatorial perturbations.

Combinatorial perturbation has been applied to map plant defense regulatory networks by conducting transcriptomics across combinatorial mutants that abolish one, two, or more disease signaling pathways [59–62]. Phenotyping factorial mutant combinations showed how the plant response to effectors like flagellin-22 is buffered by extensive interactions between all of the underlying regulatory mechanisms of jasmonate (JA), ethylene (ET), phytoalexin-deficient 4 (PAD4), and salicylate (SA) [59]. Buffering by these interactions prevented their identification by traditional single- or even double-mutant approaches [59]. Proper identification of them required all of the single, double-, and triple-mutant combinations. The use of virulent and avirulent *Pseudomonas syringae* identified a similar connectivity pattern between the subnetworks [62]. The use of the stepwise mutants identified a new aspect of the defense signaling network. The coexpression networks' responses to the pathogen were identical in the wild type (WT) and mutants that abolished the JA, ET, PAD4, and SA signaling pathways, but the magnitude was diminished [62–64]. The use of combinatorial perturbations revealed that two types of network are needed to describe the system: one that describes the magnitude of the transcriptome response and one that describes the coexpression network structure. The known genes controlled the magnitude network, while it remains an open question about which genes pattern the defense coexpression structure. Further, as these observations required factorial perturbations to remove

the buffering veil, what would factorial mutations identify in other critical responses that have largely been studied by single-mutant analysis?

Using Networks to Generate and Directly Test Hypotheses

The identification of networks and testing of their potential validity is often the end point for numerous network and systems biology studies. However, the ultimate goal of a network is to create a new view of biology that can be used to measure or test biology. In conjunction with this, there has been a wave of research wherein either the networks are used to make new predictions that are then tested or the network itself has been utilized to create a new way to phenotype dynamic systems to get at their underpinning causality. The following examples highlight the efforts to move networks beyond description and into actual implementation to show and identify causality.

Environmental sensing and signaling by plants has been a major target of predictive network modeling. A number of studies have explored plant physiological and molecular responses to a change in nitrogen supply and have uncovered important transcriptional regulators [4,24,65–67]. Recently, a GRN of root and shoot architectural changes in response to nitrogen identified TFs regulating N-associated genes [65]. Yeast one-hybrid assays were used to screen TFs and refine the network to describe the hierarchical regulation of TFs in N signaling and assimilation. Functional validation using mutant screens confirmed model-predicted relationships and revealed a specific regulatory network that contributes to root and shoot architectural changes in response to external nitrogen availability [65]. Biological network modeling has led to the identification of a number of genes and regulatory factors involved in plant susceptibility and resistance to disease through analysis of hubbiness and other measures of node importance [68–72].

Another area of complexity that is rapidly being addressed via the use of predictive networks is plant specialized or secondary metabolism. These compounds are critical for a plant's ability to defend against and interact with other organisms in its environment [73]. However, these compounds are highly diverse and frequently lineage specific, with individuals and species often having different compounds compared with their closest relatives [74]. The enzyme families for the production of these compounds are relatively easy to predict on homology but because their activity can be dramatically shifted by single amino acid changes there is little ability to predict a specific enzyme's activity. Previous work had shown that it was possible to identify the genes in a pathway one by one using coexpression, but this was a long process to identify and validate all of the steps in a pathway [75–79]. This stepwise functionality has recently been shown to be extendable to identify entire pathways in a single-network approach. Use of a catalog of transcriptomics data and mapping of coexpression networks showed that these networks frequently partition into discrete specialized metabolic pathways [80].

Recently, this concept that plants' specialized metabolic pathways are coordinate transcriptional units has been used to predict entire biosynthetic pathways for new metabolites [81–86]. By using coexpression to find enzyme genes that coexpress in response to pathogen attack in *Arabidopsis* and *Brassica*, it was possible to identify potential complete theoretical pathways for new phytoalexins that could be preliminarily validated in tobacco [83,84]. Interestingly, the transcriptome does not have to be structured by pathogen treatment, and simple coexpression networks across any dataset can often identify pathways for specialized metabolites that can play other biotic or abiotic roles [87]. Coexpression can be rapidly applied to the identification and cloning of nearly completely unstudied pathways in unstudied plant species. *De novo* transcriptomics in mayapple followed by factorial testing in tobacco identified a six-enzyme pathway for the therapeutic etoposide aglycone [85]. However, there is a critical need to accelerate the rate of

testing of these predictions in the native system to assess the biological function of these rapidly accumulating new pathways.

It is also possible to directly apply networks as a trait. One of the most common ways is to utilize an eigengene whose value can be used as any other phenotype (Box 1) [88–90]. For example, eigengenes have been used to identify quantitative trait loci that control the expression of biosynthetic networks [90]. This approach enabled the direct comparison of how different networks behave in response to the same stress, such as comparing abiotic stress networks across a range of common stresses [91] and further allowed the direct comparison of phenotypic trait variation to specific transcriptomic networks in cotton, maize and *Medicago* [92–94]. Using a network as a direct measurement of the transcriptome has the benefit of decreasing the number of tests and hence decreases the multiple testing difficulty. The network value may be a better measurement of how the system is behaving versus individual genes and thus may be a more direct comparison.

Using a network to create a quantitative estimate also enables novel analysis of complex systems. Time courses are difficult experimental systems because, while time courses are possible in one or a few genotypes, larger experimental perturbations such as time × environment or time × genotype rapidly grow too large to be feasible. A solution is to take data from a fine-scale time course to partition genes into networks that respond at specific times with specific patterns [7, 17, 18, 69, 95]. These networks then represent discrete time-dependent outputs that reflect the system. Eigengenes can then be derived to test whether the temporal system has been altered in a single time point. For example, using this approach with the circadian clock showed that it was possible to map and validate clock QTLs that alter the circadian clock using single-timepoint transcriptomics [96]. This direct use of networks to measure time-dependent processes at a single time could be applied to any system and would greatly expand the potential experimental space that could be tested.

Limitations on Network Implementation

Technological advances in sequencing and phenotyping suggest that the use of networks will only increase in biology. These advances also mean that each network is summarizing more and more data and investment. To maximize the ability to use networks to accelerate discovery, we must maximize the ability to use and reuse networks. However, this has several limitations.

A key difficulty is the lack of standardized methods for network construction. This is an advantage in that it allows great creativity in the generation of new networks, but it complicates their use by other groups. As demonstrated in this review, there is no one-size-fits-all method for building or analyzing biological networks, as the methodology must match the question being asked and the structure of the data. For some studies, pre-network clustering will have an advantage, while for others post-network clustering will be more informative. Likewise, a simple coexpression network based on correlations will be adequate to address some biological questions surrounding omics data while more mathematically complex network edges will be preferable for others. It is most important that the networks are built using sound mathematical approaches and are thoroughly tested experimentally. Strong, validated core modules can then be translated between approaches. Likewise, there is room for improvement in methods for comparing networks with more readily identify commonalities and divergences. While some methods exist in defined workflows and packages, like the consensus network analysis in WGCNA [89], it is a computationally challenging problem due to difficulties in alignment and semantic reconciliation. Solving

Table 2. Online Tools and Software for Network Construction, Analysis, and Visualization

Tool or software	Description	Refs
Cytoscape	Free tool for network construction and analysis	https://cytoscape.org/ [98]
Gephi	Free tool for network construction and analysis	https://gephi.org/ [100]
BioTapestry	Tool for building and annotating developmental models and/or models of increasing complexity over time	http://www.biotapestry.org/ [99]
Stanford Network Analysis Platform (SNAP)	For the analysis and manipulation of large networks and dynamic graphs	https://snap.stanford.edu/snap/ [109]
Network science tutorials	Network tutorials in R, Gephi, and igraph by Dr Katherine Ognyanova	http://kateto.net/tutorials/
VisANT	Network and pathway construction for multiple species	http://visant.bu.edu/ [110]
VirtualPlant	One-stop shop for gene expression analysis and network construction with layered information; easy-to-build GRNs in a number of plant species	http://virtualplant.bio.nyu.edu/cgi-bin/vpweb/ [36]
FunctionalNet	Network building for many different taxa including plants	http://www.functionalnet.org/

Outstanding Questions

How often do factorial mutants illuminate new network observations not found in single-mutant analysis?

How does evolution change/destroy/create networks between or within species?

How does causality permeate through a molecular network?

What are the best approaches to create conditional networks?

Are networks better measures of molecular traits than individual transcripts, metabolites, or proteins?

this problem will allow us to leverage across diversity networks and attain new insights about common regulatory drivers and system-wide responses to perturbations.

Another significant difficulty in biological networks is data visualization that describes information in a way that is rapidly perceived by the human brain [97]. Scientific visualization has the potential to make multivariate data easier to understand and enable networks to be used by others. Online tools and software have been created to facilitate network construction and visualization (Table 2). Cytoscape is a commonly used tool with built-in network analysis algorithms and dozens of plugins available to analyze and visualize networks [98]. Network graphs, in theory, organize omics data and ease the interpretation of significant relationships defined in the network. However, large coexpression networks are often incomprehensible as static images. Interactive visualizations that can be created using applications like BioTapestry [99] and Gephi [100] make large networks tangible and more intelligible. A possibility in the future is to visualize network graphs using immersive visualization like virtual reality (VR) [97]. It was shown that network graph exploration using VR improved user interpretation of network relationships and helped in identifying meaningful patterns in the data [101]. Broader development of new visualization tools for factorial networks is essential to allow their wider use by the community.

Concluding Remarks and Future Perspectives

Networks have been a core of the genomic revolution but we are only just beginning to understand their true strength in creating new testable hypotheses, identifying causal loci in polygenic traits, and being direct phenotypes in their own right. This shift in the community from a focus on creating networks to using them for new biology can potentially free us to develop new thoughts and observations on biology that move dramatically beyond the descriptive. To achieve this goal will require the development of new datasets from which to generate and test networks in addition to new methods for creating, sharing, and visualizing these networks. Given that the smallest plant genomes have over 20 000 genes that all function and coordinate to create any given phenotype, it will only be via these large-scale integrative and factorial approaches that we can truly generate any predictive nature in translating from genotype to phenotype (see Outstanding Questions).

Acknowledgments

Funding for this work was provided by the NSF awards IOS 1339125, IOS 1655810, and IOS 1547796 to D.J.K. and NSF-IOS 1645875 to A.M.-C., the USDA National Institute of Food and Agriculture, Hatch project number CA-D-PLS-7033-H to D.J.K., and by the Danish National Research Foundation (DNRF99) grant to D.J.K. Research reported in the publication was supported by the Foundation for Food and Agriculture Research under award number – Grant ID: 602757 to A.M.-C. The content of this publication is solely the responsibility of the authors and does not necessarily represent the official views of the Foundation for Food and Agriculture Research.

References

- Ikeuchi, M. *et al.* (2018) A gene regulatory network for cellular reprogramming in plant regeneration. *Plant Cell Physiol.* 59, 770–782
- Xiong, W. *et al.* (2017) Highly interwoven communities of a gene regulatory network unveil topologically important genes for maize seed development. *Plant J.* 92, 1143–1156
- Taylor-Teeple, M. *et al.* (2015) An *Arabidopsis* gene regulatory network for secondary cell wall synthesis. *Nature* 517, 558–559
- Gutiérrez, R.A. *et al.* (2008) Systems approach identifies an organic nitrogen-responsive gene network that is regulated by the master clock control gene *CCA1*. *Proc. Natl. Acad. Sci. U. S. A.* 105, 4939–4944
- Jiang, W. *et al.* (2017) Transcriptional regulatory network of WOX11 is involved in the control of crown root development, cytokinin signals, and redox in rice. *J. Exp. Bot.* 68, 2787–2798
- Kim, H.J. *et al.* (2014) Gene regulatory cascade of senescence-associated NAC transcription factors activated by ETHYLENE-INSENSITIVE2-mediated leaf senescence signalling in *Arabidopsis*. *J. Exp. Bot.* 65, 4023–4036
- Hammer, S.L. *et al.* (2000) Orchestrated transcription of key pathways in *Arabidopsis* by the circadian clock. *Science* 290, 2110–2113
- Hwang, S. *et al.* (2011) Systematic prediction of gene function in *Arabidopsis thaliana* using a probabilistic functional gene network. *Nat. Protoc.* 6, 1429–1442
- Wang, W. *et al.* (2017) Genome-wide characterization of the aldehyde dehydrogenase gene superfamily in soybean and its potential role in drought stress response. *BMC Genomics* 18, 518
- Hickman, R. *et al.* (2017) Architecture and dynamics of the jasmonic acid gene regulatory network. *Plant Cell* 29, 2086–2105
- Simon, M. *et al.* (2013) Tissue-specific profiling reveals transcriptome alterations in *Arabidopsis* mutants lacking morphological phenotypes. *Plant Cell* 25, 3175–3185
- Haque, S. *et al.* (2019) Computational prediction of gene regulatory networks in plant growth and development. *Curr. Opin. Plant Biol.* 47, 96–105
- Alvarez-Buylla, E.R. *et al.* (2007) Gene regulatory network models for plant development. *Curr. Opin. Plant Biol.* 10, 83–91
- Wils, C.R. and Kaufmann, K. (2017) Gene-regulatory networks controlling inflorescence and flower development in *Arabidopsis thaliana*. *Biochim. Biophys. Acta* 1860, 95–105
- Shibata, M. and Sugimoto, K. (2019) A gene regulatory network for root hair development. *J. Plant Res.* 132, 301–309
- Gargouri, M. *et al.* (2015) Identification of regulatory network hubs that control lipid metabolism in *Chlamydomonas reinhardtii*. *J. Exp. Bot.* 66, 4551–4566
- Greenham, K. *et al.* (2017) Temporal network analysis identifies early physiological and transcriptomic indicators of mild drought in *Brassica rapa*. *Elife* 6, e29655
- Windram, O. *et al.* (2012) *Arabidopsis* defense against *Botrytis cinerea*: chronology and regulation deciphered by high-resolution temporal transcriptomic analysis. *Plant Cell* 24, 3530–3557
- Polanski, K. *et al.* (2014) Wigwags: identifying gene modules co-regulated across multiple biological conditions. *Bioinformatics* 30, 962–970
- Bechtold, U. *et al.* (2016) Time-series transcriptomics reveals that *AGAMOUS-LIKE22* affects primary metabolism and developmental processes in drought-stressed *Arabidopsis*. *Plant Cell* 28, 345–366
- Gitter, A. *et al.* (2013) Linking the signaling cascades and dynamic regulatory networks controlling stress responses. *Genome Res.* 23, 365–376
- Mochida, K. *et al.* (2018) Statistical and machine learning approaches to predict gene regulatory networks from transcriptome datasets. *Front. Plant Sci.* 9, 1770
- Krouk, G. *et al.* (2010) Predictive network modeling of the high-resolution dynamic plant transcriptome in response to nitrate. *Genome Biol.* 11, R123
- Varala, K. *et al.* (2018) Temporal transcriptional logic of dynamic regulatory networks underlying nitrogen signaling and use in plants. *Proc. Natl. Acad. Sci. U. S. A.* 115, 6494–6499
- Bonneau, R. *et al.* (2006) The Inferelator: an algorithm for learning parsimonious regulatory networks from systems-biology data sets *de novo*. *Genome Biol.* 7, 16
- Wilkins, O. *et al.* (2016) EGRINs (environmental gene regulatory influence networks) in rice that function in the response to water deficit, high temperature, and agricultural environments. *Plant Cell* 28, 2365–2384
- Arrieta-Ortiz, M.L. *et al.* (2015) An experimentally supported model of the *Bacillus subtilis* global transcriptional regulatory network. *Mol. Syst. Biol.* 11, 839
- Milone, D.H. *et al.* (2010) omeSOM: a software for clustering and visualization of transcriptional and metabolite data mined from interspecific crosses of crop plants. *BMC Bioinformatics* 11, 438
- Jiang, L. *et al.* (2018) Analysis of gene regulatory networks of maize in response to nitrogen. *Genes (Basel)* 9, E151
- Dong, Z. *et al.* (2012) A gene regulatory network model for floral transition of the shoot apex in maize and its dynamic modeling. *PLoS One* 7, e43450
- Foo, M. *et al.* (2018) A framework for engineering stress resilient plants using genetic feedback control and regulatory network rewiring. *ACS Synth. Biol.* 7, 1553–1564
- Ortiz-Gutiérrez, E. *et al.* (2015) A dynamic gene regulatory network model that recovers the cyclic behavior of *Arabidopsis thaliana* cell cycle. *PLoS Comput. Biol.* 11, e1004486
- García-Gómez, M.L. *et al.* (2017) A dynamic genetic-hormonal regulatory network model explains multiple cellular behaviors of the root apical meristem of *Arabidopsis thaliana*. *PLoS Comput. Biol.* 13, e1005488
- Mishra, B. *et al.* (2018) Dynamic modeling of transcriptional gene regulatory network uncovers distinct pathways during the onset of *Arabidopsis* leaf senescence. *NPJ Syst. Biol. Appl.* 4, 4
- Di Silvestre, D. *et al.* (2018) Large scale proteomic data and network-based systems biology approaches to explore the plant world. *Proteomes* 6, 27
- Katari, M.S. *et al.* (2010) VirtualPlant: a software platform to support systems biology research. *Plant Physiol.* 152, 500–515
- Srinivasan, V. *et al.* (2017) Decreasing, not increasing, leaf area will raise crop yields under global atmospheric change. *Glob. Chang. Biol.* 23, 1626–1635
- Wang, Y. *et al.* (2014) Three distinct biochemical subtypes of C4 photosynthesis? A modelling analysis. *J. Exp. Bot.* 65, 3567–3578
- Kannan, K. *et al.* (2019) Combining gene network, metabolic, and leaf-level models show means to future-proof soybean

- photosynthesis under rising CO₂. *bioRxiv* Published online March 20, 2019. <https://doi.org/10.1101/582981>
40. Dal'Molin, C.G. *et al.* (2010) C4GEM, a genome-scale metabolic model to study C4 plant metabolism. *Plant Physiol.* 154, 1871–1885
 41. Robaina-Estévez, S. and Nikoloski, Z. (2016) Metabolic network constrains gene regulation of C4 photosynthesis: the case of maize. *Plant Cell Physiol.* 57, 933–943
 42. Acharjee, A. *et al.* (2018) Genetical genomics of quality related traits in potato tubers using proteomics. *BMC Plant Biol.* 18, 20
 43. Kliebenstein, D.J. (2012) Model misinterpretation within biology: phenotypes, statistics, networks, and inference. *Front. Plant Sci.* 3, 13
 44. Knoch, D. *et al.* (2017) Genetic dissection of metabolite variation in *Arabidopsis* seeds: evidence for mQTL hotspots and a master regulatory locus of seed metabolism. *J. Exp. Bot.* 68, 1655–1667
 45. Wentzell, A.M. *et al.* (2008) Genetic networks controlling structural outcome of glucosinolate activation across development. *PLoS Genet.* 4, e1000234
 46. Wentzell, A.M. *et al.* (2007) Linking metabolic QTLs with network and cis-eQTLs controlling biosynthetic pathways. *PLoS Genet.* 3, 1687–1701
 47. Lipka, A.E. *et al.* (2013) Genome-wide association study and pathway-level analysis of tocopherol levels in maize grain. *G3 (Bethesda)* 3, 1287–1299
 48. Bunyavanich, S. *et al.* (2014) Integrated genome-wide association, coexpression network, and expression single nucleotide polymorphism analysis identifies novel pathway in allergic rhinitis. *BMC Med. Genet.* 7, 48
 49. Jia, P. and Zhao, Z. (2014) Network-assisted analysis to prioritize GWAS results: principles, methods and perspectives. *Hum. Genet.* 133, 125–138
 50. Lee, T. and Kim, H. (2015) Network-assisted crop systems genetics: network inference and integrative analysis. *Curr. Opin. Plant Biol.* 24, 61–70
 51. Lee, T. and Lee, I. (2018) araGWAB: network-based boosting of genome-wide association studies in *Arabidopsis thaliana*. *Sci. Rep.* 8, 2925
 52. Shen, M. *et al.* (2013) Leveraging non-targeted metabolite profiling via statistical genomics. *PLoS One* 8, e57667
 53. Zhu, J. *et al.* (2012) Stitching together multiple data dimensions reveals interacting metabolomic and transcriptomic networks that modulate cell regulation. *PLoS Biol.* 10, e1001301
 54. Schaefer, R.J. *et al.* (2018) Integrating coexpression networks with GWAS to prioritize causal genes in maize. *Plant Cell* 30, 2922–2942
 55. Chan, E.K. *et al.* (2011) Combining genome-wide association mapping and transcriptional networks to identify novel genes controlling glucosinolates in *Arabidopsis thaliana*. *PLoS Biol.* 9, e1001125
 56. Corwin, J.A. *et al.* (2016) The quantitative basis of the *Arabidopsis* innate immune system to endemic pathogens depends on pathogen genetics. *PLoS Genet.* 12, e1005789
 57. Fordyce, R.F. *et al.* (2018) Digital imaging combined with genome-wide association mapping links loci to plant–pathogen interaction traits. *Plant Physiol.* 178, 1406–1422
 58. Barabasi, A.L. (2009) Scale-free networks: a decade and beyond. *Science* 325, 412–413
 59. Hillmer, R.A. *et al.* (2017) The highly buffered *Arabidopsis* immune signaling network conceals the functions of its components. *PLoS Genet.* 13, e1006639
 60. Hatsugai, N. *et al.* (2017) A plant effector-triggered immunity signaling sector is inhibited by pattern-triggered immunity. *EMBO J.* 36, 2758–2769
 61. Kim, Y. *et al.* (2014) Mechanisms underlying robustness and tunability in a plant immune signaling network. *Cell Host Microbe* 15, 84–94
 62. Mine, A. *et al.* (2018) The defense phytohormone signaling network enables rapid, high-amplitude transcriptional reprogramming during effector-triggered immunity. *Plant Cell* 30, 1199–1219
 63. Li, B.H. *et al.* (2018) Network-guided discovery of extensive epistasis between transcription factors involved in aliphatic glucosinolate biosynthesis. *Plant Cell* 30, 178–195
 64. Li, B. *et al.* (2014) Promoter based integration in plant defense regulation. *Plant Physiol.* 166, 1803–1820
 65. Gaudinier, A. *et al.* (2018) Transcriptional regulation of nitrogen-associated metabolism and growth. *Nature* 563, 259–264
 66. Medici, A. *et al.* (2015) ATNIGT1/HRS1 integrates nitrate and phosphate signals at the *Arabidopsis* root tip. *Nat. Commun.* 6, 6274
 67. Ristova, D. *et al.* (2016) Combinatorial interaction network of transcriptomic and phenotypic responses to nitrogen and hormones in the *Arabidopsis thaliana* root. *Sci. Signal.* 9, rs13
 68. Girard, I.J. *et al.* (2017) RNA sequencing of *Brassica napus* reveals cellular redox control of *Sclerotinia* infection. *J. Exp. Bot.* 68, 5079–5091
 69. Lewis, L.A. *et al.* (2015) Transcriptional dynamics driving MAMP-triggered immunity and pathogen effector-mediated immunosuppression in *Arabidopsis* leaves following infection with *Pseudomonas syringae* pv tomato DC3000. *Plant Cell* 27, 3038–3064
 70. Zhang, W. *et al.* (2017) Plastic transcriptomes stabilize immunity to pathogen diversity: the jasmonic acid and salicylic acid networks within the *Arabidopsis/Botrytis* pathosystem. *Plant Cell* 29, 2727–2752
 71. Crocco, C.D. *et al.* (2015) The transcriptional regulator BBX24 impairs DELLA activity to promote shade avoidance in *Arabidopsis thaliana*. *Nat. Commun.* 6, 6202
 72. Cui, J. *et al.* (2017) Comparative transcriptome analysis between resistant and susceptible tomato allows the identification of lncRNA16397 conferring resistance to *Phytophthora infestans* by co-expressing glutaredoxin. *Plant J.* 89, 577–589
 73. Kliebenstein, D.J. (2004) Secondary metabolites and plant/environment interactions: a view through *Arabidopsis thaliana* tinted glasses. *Plant Cell Environ.* 27, 675–684
 74. Osbourn, A.E. and Kliebenstein, D.J. (2012) Making new molecules – evolution of pathways for novel metabolites in plants. *Curr. Opin. Plant Biol.* 15, 415–423
 75. Li, J. *et al.* (2008) Subclade of flavin-monooxygenases involved in aliphatic glucosinolate biosynthesis. *Plant Physiol.* 148, 1721–1733
 76. Hansen, B.G. *et al.* (2008) A novel 2-oxoacid-dependent dioxygenase involved in the formation of the goiterogenic 2-hydroxybut-3-enyl glucosinolate and generalist insect resistance in *Arabidopsis*. *Plant Physiol.* 148, 2096–2108
 77. Hansen, B.G. *et al.* (2007) Identification of a flavin-monooxygenase as the S-oxygenating enzyme in aliphatic glucosinolate biosynthesis in *Arabidopsis*. *Plant J.* 50, 902–910
 78. Saito, K. *et al.* (2008) Decoding genes with coexpression networks and metabolomics – ‘majority report by precogs’. *Trends Plant Sci.* 13, 36–43
 79. Yonekura-Sakakibara, K. *et al.* (2007) Identification of a flavonol 7-O-rhamnosyltransferase gene determining flavonoid pattern in *Arabidopsis* by transcriptome coexpression analysis and reverse genetics. *J. Biol. Chem.* 282, 14932–14941
 80. Wisecaver, J.H. *et al.* (2017) A global coexpression network approach for connecting genes to specialized metabolic pathways in plants. *Plant Cell* 29, 944–959
 81. Mafu, S. *et al.* (2018) Discovery, biosynthesis and stress-related accumulation of dolabradiene-derived defenses in maize. *Plant Physiol.* 176, 2677–2690
 82. Irmisch, S. *et al.* (2018) Discovery of UDP-glycosyltransferases and BAHd-acyltransferases involved in the biosynthesis of the antidiabetic plant metabolite montbretin A. *Plant Cell* 30, 1864–1886
 83. Klein, A.P. and Sattely, E.S. (2017) Biosynthesis of cabbage phytoalexins from indole glucosinolate. *Proc. Natl. Acad. Sci. U. S. A.* 114, 1910–1915
 84. Rajniak, J. *et al.* (2015) A new cyanogenic metabolite in *Arabidopsis* required for inducible pathogen defence. *Nature* 525, 376–379
 85. Lau, W. and Sattely, E.S. (2015) Six enzymes from mayapple that complete the biosynthetic pathway to the etoposide aglycone. *Science* 349, 1224–1228
 86. Klein, N.P. and Sattely, E.S. (2015) Two cytochromes P450 catalyze S-heterocyclizations in cabbage phytoalexin biosynthesis. *Nat. Chem. Biol.* 11, 837–839

87. Moghe, G.D. *et al.* (2017) Evolutionary routes to biochemical innovation revealed by integrative analysis of a plant-defense related specialized metabolic pathway. *Elife* 6, e28468
88. Zhang, B. and Horvath, S. (2005) A general framework for weighted gene co-expression network analysis. *Stat. Appl. Genet. Mol. Biol.* 4, 17
89. Langfelder, P. and Horvath, S. (2008) WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9, 559
90. Kliebenstein, D.J. *et al.* (2006) Identification of QTLs controlling gene expression networks defined *a priori*. *BMC Bioinformatics* 7, 308
91. Rasmussen, S. *et al.* (2013) Transcriptome responses to combinations of stresses in *Arabidopsis*. *Plant Physiol.* 161, 1783–1794
92. Hu, G.J. *et al.* (2016) Evolutionary conservation and divergence of gene coexpression networks in *Gossypium* (cotton) seeds. *Genome Biol. Evol.* 8, 3765–3783
93. Palakurty, S.X. *et al.* (2018) Cooperation and coexpression: how coexpression networks shift in response to multiple mutualists. *Mol. Ecol.* 27, 1860–1873
94. Schluter, U. *et al.* (2012) Maize source leaf adaptation to nitrogen deficiency affects not only nitrogen and carbon metabolism but also control of phosphate homeostasis. *Plant Physiol.* 160, 1384–1406
95. Harmer, S.L. and Kay, S.A. (2005) Positive and negative factors confer phase-specific circadian regulation of transcription in *Arabidopsis*. *Plant Cell* 17, 1926–1940
96. Kerwin, R.E. *et al.* (2011) Network quantitative trait loci mapping of circadian clock outputs identifies metabolic pathway-to-clock linkages in *Arabidopsis*. *Plant Cell* 23, 471–485
97. Christensen, A.J. *et al.* (2018) Use of computational modeling combined with advanced visualization to develop strategies for the design of crop ideotypes to address food security. *Nutr. Rev.* 76, 332–347
98. Shannon, P. *et al.* (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504
99. Longabaugh, W.J.R. *et al.* (2009) Visualization, documentation, analysis, and communication of large-scale gene regulatory networks. *Biochim. Biophys. Acta* 1789, 363–374
100. Bastian, M. *et al.* (2009) Gephi: an open source software for exploring and manipulating networks. In *Proceedings of the Third International Conference on Weblogs and Social Media, ICWSM 2009*, Association for the Advancement of Artificial Intelligence
101. Kwon, O.-H. *et al.* (2016) A study of layout, rendering, and interaction methods for immersive graph visualization. *IEEE Trans. Vis. Comput. Graph.* 22, 1802–1815
102. Morrissey, E.R. *et al.* (2011) Inferring the time-invariant topology of a nonlinear sparse gene regulatory network using fully Bayesian spline autoregression. *Biostatistics* 12, 682–694
103. Trejo Banos, D. *et al.* (2015) A Bayesian approach for structure learning in oscillating regulatory networks. *Bioinformatics* 31, btv414
104. Niel, C. *et al.* (2015) A survey about methods dedicated to epistasis detection. *Front. Genet.* 6, 285
105. Mirowski, P. and LeCun, Y. (2009) Dynamic factor graphs for time series modeling. In *Machine Learning and Knowledge Discovery in Databases, Part II* (Buntine, W., *et al.*, eds), pp. 128–143, Springer
106. Azpeitia, E. *et al.* (2014) *Gene Regulatory Network Models for Floral Organ Determination*. pp. 441–469, Humana Press
107. Penfold, C.A. and Wild, D.L. (2011) How to infer gene networks from expression profiles, revisited. *Interface Focus* 1, 857–870
108. Balaguer, M.A.D. *et al.* (2017) Predicting gene regulatory networks by combining spatial and temporal gene expression data in *Arabidopsis* root stem cells. *Proc. Natl. Acad. Sci. U. S. A.* 114, E7632–E7640
109. Leskovec, J. and Sosič, R. (2016) SNAP: a general-purpose network analysis and graph-mining library. *ACM Trans. Intell. Syst. Technol.* 8, 1
110. Hu, Z. *et al.* (2013) VisANT 4.0: integrative network platform to connect genes, drugs, diseases and therapies. *Nucleic Acids Res.* 41, W225–W231