Available online at www.sciencedirect.com

SciVerse ScienceDirect

Current Opinion in Plant Biology

Towards understanding how molecular networks evolve in plants

Lee Chae[1], Insuk Lee[2], Junha Shin[2] and Seung Yon Rhee[1]

Residing beneath the phenotypic landscape of a plant are intricate and dynamic networks of genes and proteins. As evolution operates on phenotypes, we expect its forces to shape somehow these underlying molecular networks. In this review, we discuss progress being made to elucidate the nature of these forces and their impact on the composition and structure of molecular networks. We also outline current limitations and open questions facing the broader field of plant network analysis.

Addresses
[1] Department of Plant Biology, Carnegie Institution for Science, 260 Panama St, Stanford, CA 94305, USA
[2] Department of Biotechnology, College of Life Science and Biotechnology, Yonsei University, 262 Seongsanno, Seodaemun-gu, Seoul 120-749, Republic of Korea

Corresponding authors: Chae, Lee (lchae@stanford.edu) and Lee, Insuk (insuklee@yonsei.ac.kr), Rhee, Seung Yon (rhee@acoma.stanford.edu)

Current Opinion in Plant Biology 2012, **15**:177–184

This review comes from a themed issue on
Genome studies and molecular genetics
Edited by Yves Van de Peer and J. Chris Pires

Available online 24th January 2012

1369-5266/$ – see front matter

DOI 10.1016/j.pbi.2012.01.006

Introduction

Complex plant traits, such as the clever mimicry of orchid flowers and exploratory dances of pea tendrils, are true puzzles of nature. In some instances, the molecular pieces to a puzzle sit scattered on the table, but we have little idea how they integrate to enable the final masterpiece. In other cases, we may see the puzzle's final form, but have trouble recognizing the pieces. One way to investigate such biological puzzles and the molecular pieces that conspire to produce them is to model them using network-theory-based approaches [1].

In network analysis, complex systems are modeled as networks whose functional components are represented as nodes and the relationships among them as edges, or links, between nodes. For example, in a protein–protein interaction network, nodes represent proteins and edges depict a physical interaction between two nodes (Figure 1). A variety of biological systems have been investigated on a genome-wide scale using network analysis, including protein–protein interaction, metabolic, gene co-function, co-expression, and regulatory networks. Current genome-wide network reconstructions have typically targeted simple, unicellular organisms such as yeast and bacteria, which benefit from a large amount of available omics-type data. With the exception of gene co-expression networks, only a limited number of genome-scale biological networks have been reconstructed for plants due to a lack of appropriate data (Table 1).

The graph-based network analyses we discuss here provide static portraits of molecular networks. While these snapshots capture an important perspective on a network's fundamental structure and organization, they do not provide a dynamic view of its behavior. A number of modeling frameworks attempt to portray the dynamic properties of networks, including those based on ordinary differential equations and dynamical Bayesian networks, among others [2]. However, the absence of sufficient experimental data is a common obstacle for all network modeling approaches, with genome-wide kinetic data being particularly limited.

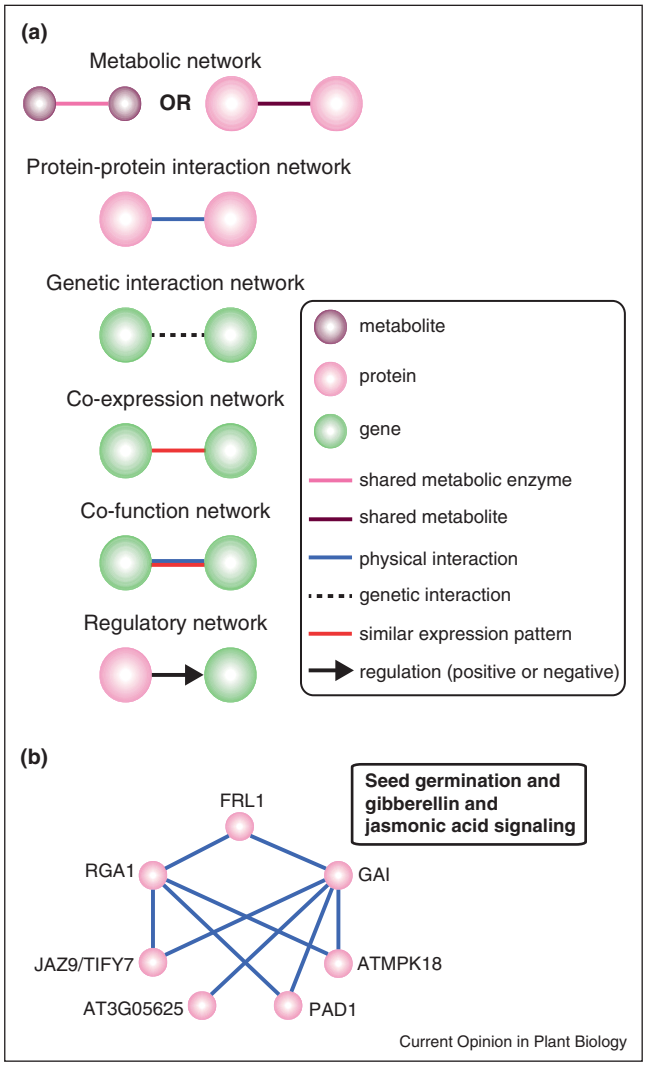
Several reviews describe how to reconstruct genome-wide molecular networks and how to use them in a variety of research problems [1,3•,4,5•,6]. Here we focus on the problem of how biological networks evolve, an emerging pursuit sometimes called 'evolutionary systems biology' [7••,8]. These studies attempt to understand how evolution has shaped the composition, structure, and function of biological networks. Many of the findings we discuss deal primarily with bacterial and yeast networks. As more genome-scale plant networks become available, we expect a greater focus to be placed on investigating network evolution in plants.

Evolution of network composition

Gene duplication as a driving force for node evolution

How do nodes evolve and what evolutionary forces shape their divergence? Nodes in a molecular network typically represent genes and proteins. Genes can duplicate at the single-gene, chromosome, and whole-genome level and these events provide a means for the addition and functional divergence of network nodes [9]. For example, many innovations in metabolic networks come from duplications of existing enzymes [10,11•,12]. More so than single-gene duplications, whole-genome duplications have the potential to create large-scale changes in molecular networks. For example, the evolution of a protein–protein interaction network of transcription factors in several plant species can be mainly attributed to successive rounds of whole-genome duplications, as opposed to small-scale duplication events [13].

Figure 1



Nodes and edges can depict a variety of molecular components and their interactions within molecular networks. **(a)** Nodes of a molecular network typically represent genes, proteins, or chemical compounds. Edges signify a biological relationship between nodes. For example, in a co-expression network, two genes that display similar transcriptional patterns are depicted as nodes with an edge linking them. In some networks, edges may be directional, as seen in a regulatory network, where a node representing a protein is linked to a gene whose expression level it regulates. Note that in this example, a node can represent either a protein or a gene. Co-function networks are unique in that the edges represent a high likelihood that two genes have similar function. Edges are based on the integration of a variety of genome-scale functional data. Table 1 describes the types of nodes and edges found in common molecular networks and contains references to examples of these networks in plants. **(b)** An example molecular network composed of interactions among Arabidopsis proteins involved in seed germination and gibberellin and jasmonic acid signaling, mapped by yeast two-hybrid analysis, elucidates a functional module.Adapted from [17••].

After a duplication event, genes can either be lost or retained. Unraveling the dynamics, mechanisms, and causes of genome architecture reorganization after duplication is an active area of research [9]. One possible reason for retention within metabolic networks is a boost in metabolic flux resulting from the existence of additional copies of the enzyme-encoding gene [14]. While increases in flux may be a source of selective advantage, other possible factors for retention also include: (1) compensation for genetic malfunctions; (2) maintenance of gene balance, according to the gene dosage balance hypothesis [15]; (3) subfunctionalization, in which gene pairs diversify in the timing or location of their activity; and (4) neo-functionalization, where duplicated genes handle separate functional roles [9,16]. However, determining the existence and relative effects of these factors in network evolution is challenging as experimental evidences are limited, and in some cases (sub- and neo-functionalization), innovations in measuring their effects are needed.

Rewiring of edges during network evolution

Molecular networks can diversify not only through the introduction of new nodes, but also as a result of rewiring events in which edges are gained or lost. A gain of an edge between two nodes represents the appearance of new functionality, such as when an enzyme within a metabolic network evolves to catalyze a new substrate. The loss of an edge can result in functional divergence, as when duplicate copies of a protein evolve to bind different interaction partners, losing a subset of their initial interactions in the process. Such rewiring of interactions among sets of duplicated proteins has contributed to the functional divergence of the protein–protein interaction network in Arabidopsis [17••]. Furthermore, the frequency of rewiring in the Arabidopsis protein–protein interaction network has differed depending on the type of duplication. Protein pairs derived from whole genome duplication events retained their interactions to a greater degree than those generated by small-scale gene duplications, suggesting that relative modes of duplication have distinct roles in the addition and divergence of edges during network evolution [18].

How fast can network edges be rewired? Different types of molecular networks undergo edge rewiring at different rates. Across eukaryotic protein interaction networks, these rates are estimated to average 10^{-5} to 10^{-6} interactions per protein pair per million years of divergence [19,20]. By contrast, transcriptional regulatory networks appear to experience higher rewiring rates [21••], while natural selection apparently has not favored extensive rewiring among subunits of protein complexes [22].

As with node evolution, unearthing the mechanisms that drive network rewiring is a formidable challenge. Posited mechanisms, such as gene duplication followed

Table 1

Different types of genome-wide molecular networks

Network	Nodes	Edges	Reconstruction methods
Metabolic network	Metabolites, enzymes, or reactions	An edge can be a reaction that consumes one metabolite to produce another; a linkage between enzymes participating in consecutive reactions; or, a metabolite produced by one reaction and consumed by another reaction.	Computational prediction and curation of experimentally determined enzymes and reactions [62,73].
Protein–protein interaction network	Proteins	Edges between two proteins indicate the pair can engage in a physical interaction.	Experimental detection of physical contact between proteins [17••,63].
Genetic interaction network	Genes	Edges are drawn between two genes if they have non-additive phenotypes.	Experimental detection of non-additive phenotypic effects of mutations in two genes [6,64].
Co-expression network	Genes	An edge between two genes represents similarity in their expression patterns, usually across spatiotemporal contexts.	Statistical detection of correlation of gene expression across multiple conditions [52,65–69].
Co-function network	Genes	An edge represents the probability that two genes function in the same pathway.	Statistical assessment of two genes participating in the same pathway [26••,70•].
Regulatory network	Genes	Edges depict the case where one gene directly regulates a second gene.	Experimental detection of transcription factor-DNA binding [71].

by neo-functionalization or subfunctionalization due to edge gains and losses, are hard to isolate without proper reference networks, though strategies for confronting this challenge are emerging [23].

In summary, gene duplication combined with linkage rewiring provides a fundamental engine for the growth and diversification of network nodes and edges, though we do not yet know what drives this engine and how the engine works. The next section addresses the question of how evolution can affect the shape of molecular networks.

Evolution of network structure

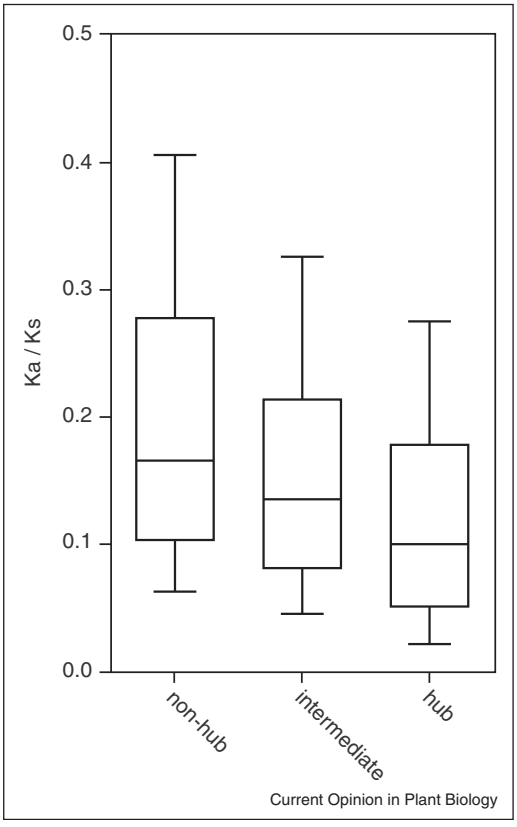
Networks have characteristic features in their shape (also referred to as topology) [24]. Within molecular networks, a few nodes, called hubs, have many connections to other nodes, while most nodes have only a few connections in a distribution sometimes referred to as 'scale-free' [25]. Also, nodes are found in clustered subsets of highly interconnected members, and are believed to act together as functional modules. Finally, networks typically contain recurring motifs of linkage patterns, representing modes of regulation. Why are these features prevalent in molecular networks and how does evolution shape network topology? Here we examine how evolution shapes connectivity of nodes and cohesiveness of functional modules.

Network connectivity and evolutionary rate

Is the evolutionary rate of a gene or protein affected by its physical and functional interactions within a network? For example, in a genomewide co-function network of Arabidopsis [26••], highly connected proteins (hubs) evolve more slowly than proteins with less connectivity (Figure 2). A similar effect was observed in the yeast metabolic network, as well as the yeast protein–protein interaction network [27,28•], although the latter observation has been contested owing to confounding factors, such as gene expression level [29,30]. Unlike in yeast, however, the inverse correlation between gene evolutionary rate and network connectivity in the Arabidopsis co-function network (Figure 2) still exists even after controlling for expression level (−0.18 vs. −0.13, Pearson vs. partial correlation, fixing for expression level, unpublished results).

One explanation for the relationship between high connectivity and slow evolutionary rates is that hub molecules are indispensable in networks. Hence, most mutations in their sequences are not favored by natural selection. This 'centrality-lethality' idea was originally proposed to explain the slow evolutionary rates of hubs in protein–protein interaction networks [31•,32]. But the centrality-lethality relationship is not observable in other types of molecular networks, as highly connected enzymes in metabolic networks—though perhaps evolving more slowly—are no more essential than less connected enzymes [28•,33]. Also, central transcription factors in a yeast gene regulatory network actually have a higher rate of evolution than less connected transcription factors [34]. Furthermore, correlations between a node's connectivity and its duplicability (probability of being retained from gene duplication, which is another form of evolutionary constraint) vary according to the type of biological network and organism, including the Arabidopsis metabolic network [28•,35–37].

Figure 2



Highly connected genes (hubs) evolve more slowly than those with less connectivity (intermediate and non-hubs) in Arabidopsis. We used 8789 genes in the genome-wide co-function network, AraNet [26••], with evolutionary rates represented as a ratio between nonsynonymous substitution and synonymous substitution (Ka/Ks). The evolutionary rates were estimated from comparisons of *Arabidopsis thaliana* and *Arabidopsis lyrata* [38]. The higher Ka/Ks indicates a faster evolutionary rate. Evolutionary rate inversely correlates with connectivity in AraNet (Pearson's correlation coefficient = −0.18), even after having been controlled by fixing expression level (partial correlation coefficient = −0.13). For visual clarity of correlation analysis, we divided these genes into three classes by degree of network connectivity: hub (>200 links), intermediate (between 20 and 200 links), non-hub (<20 links). Distributions of Ka/Ks for each class of genes (1541 hubs, 4043 intermediates, and 3205 non-hubs) were summarized as box-and-whisker plots, showing 90%, 75%, 50%, 25%, and 10% quantiles from the top whisker. Differences among the three classes are statistically significant: non-hub vs. intermediate (p-value = 0), intermediate vs. hub (p-value = 1.38×10^{-256}), and non-hub vs. hub (p-value = 0) by Wilcoxon rank sum test.

The differences seen in these associations suggest that while the topology of a network may affect evolutionary rates among its components, its effect needs to be evaluated relative to other aspects of gene and protein function. For example, gene evolution rates in yeast appear to be more strongly tied to gene expression levels than to network connectivity [30]. In Arabidopsis, gene evolutionary rates also correlate with gene expression levels and the type of duplication event that generated the gene, when compared to factors such as gene structure, chromosomal positioning, local recombination rates, and gene multi-functionality [38]. Whether and how these evolutionary rates vary within plant networks according to topological features is an open question.

Functional modules as a basic unit of network evolution

Subsets of highly interconnected nodes within biological networks are often implicated to be functional modules [39••,40]. Do selective constraints affect modules, conserving them across evolutionary time scales? About half of the functional modules in molecular networks appear to be comprised of genes whose phylogenetic distributions are more similar than expected by chance [41•]. 'Evolutionary cohesiveness' measures the tendency of members of a module to experience the same evolutionary event such as a gain or loss. In prokaryotes and eukaryotes (including Arabidopsis), approximately 40% and 46% of identified modules, respectively, display evolutionary cohesiveness [42,43].

These findings show that a subset of functional modules have been conserved to some degree during evolution. However, the results also reveal that a large proportion of functional modules can vary in their molecular composition across species. Consequently, it is logical to ask whether selective forces drive the variation seen in functional modules. For instance, plant proteins interacting with pathogen proteins evolve faster than those interacting with other plant proteins, probably owing to an arms-race between pathogen and host proteins [44••]. It would be interesting to see whether pathogen-resistance-based functional modules display greater evolutionary change than other modules.

Finally, despite the different evolutionary paths that networks can take to derive their composition of nodes, edges, and functional modules, their phenotypic output can be remarkably consistent. For example, evolution simulations involving the yeast metabolic network demonstrated similar metabolic capacities for the final network despite being confronted with different selective pressures [45]. Furthermore, the consistent phenotype of a network can be achieved with a large variety of possible genotypes and across a number of different evolutionary scenarios [46,47]. Therefore, understanding the logic and pattern of evolutionary trajectories of molecular networks will be crucial in predicting phenotype from genotype.

In summary, network properties of a protein may affect how it evolves, and reciprocally, the functions of proteins and modules can affect how a network structure evolves. The findings presented here represent just the tip of the iceberg of new knowledge that will be uncovered in this field, provided we overcome some current limitations.

Limitations and challenges

Despite the progress in developing an evolutionary understanding of molecular networks, we face many challenges. First, we still have limited network views of plants. The largest interactome for the reference plant Arabidopsis [17••] covers only 10% of the genome, and the largest inferred co-function network [26••] only 75% of the genome. We do not have a large-scale map of genetic interactions for any plant [48•,49•]. The lack of a genetic interaction map for plants is a seemingly intractable problem, but perhaps molecular networks could guide the selection of gene pairs to test for interaction. For example, most of the genes whose functions were discovered through traditional forward or reverse genetics appear to be highly connected in a genome-wide co-function network of Arabidopsis (Figure 3). Pairwise combinations of a hub gene and each of its immediate neighbors may be good candidates for testing for non-additive genetic interactions. If a hub gene with an identifiable phenotype exists, 'hub gene-neighbor gene' double mutants may show an enhanced or suppressed phenotype.

Second, proteins function in a context-specific manner (e.g. cell type, tissue type, developmental stage, environment), and highly plastic transcriptomes in different cell types suggest variation in biological networks across different cell types [50,51]. But, most experimentally mapped interactions for plants do not account for context specificity. While efforts have begun to account for time and place in reconstructing plant molecular networks [52–54], questions about how evolution shapes context specificity remain to be answered.

Third, different types of biological data need to be integrated for holistic analysis and interpretation of networks. The improved power for gene discovery by combining molecular networks and quantitative trait locus mapping [55] or genome-wide association studies [56•,57•] in plants foreshadows accelerated progress for investigating the evolution of complex traits in plants.

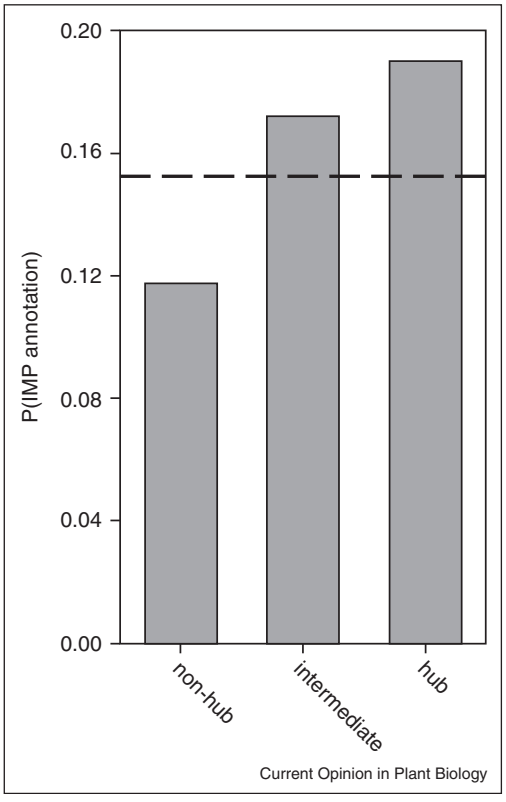
Finally, we need to build genome-wide networks for more plant species to understand how network components and their organization evolved in the plant lineage. For example, genome-wide duplications are rampant in plants [58] and investigating the effects of these large-scale reorganizations of the genome will shed light on our understanding of both micro- and macro-evolutionary processes.

Open questions

Network theory allows us to model the molecular underpinnings of complex biological systems, and evolutionary studies of these systems help us understand their properties, including their organization, dynamics, and robustness. The ultimate goal is to understand how these systems function to produce phenotypes. With that in mind, we conclude with the following open questions as possible avenues towards meeting this goal:

- Can we complete the reconstruction of plant molecular networks? How do we define and assess completeness?
- How do we integrate networks of different types and levels of organization ranging from metabolites, genes, transcripts, proteins, reactions, pathways, functional modules, regulatory motifs, subcellular compartments, cells, tissues, organs, organ systems, organisms, and ecosystems?
- Most functional modules are made up of more than two genes. Therefore, binary genetic interaction studies will not uncover the function of most of these modules. How can we systematically elucidate the functions of these multi-genic modules?

Figure 3



Arabidopsis genes with characterized mutant phenotypes are more connected in a genome-wide co-function network (AraNet). Among 19,647 genes in AraNet [26••], 2888 genes (~14.7%) are annotated with Gene Ontology biological process terms with support from mutant phenotype characterization (evidence code IMP: Inferred from Mutant Phenotype) [72]. Hub genes (as defined in Figure 2) are enriched in IMP annotation (17.4%, 574 out of 3281, p-value = 6.62×10^{-7}, binomial distribution), whereas non-hubs are depleted in IMP annotation (11.5%, 887 out of 7699, p-value = 1.80×10^{-16}, binomial distribution). Dashed line shows the background portion of all IMP-annotated genes in AraNet (2888/19,647 = 14.7%).

- Can traits or biological processes be represented as networks and if so how can we build such networks? How will these trait networks relate to molecular networks?
- Which plant traits are 'keystone' traits in the network that serve as the tipping points of adaptive landscapes? Which traits are versatile or exploratory [e.g. [59]]?
- How can we model agronomically important traits, such as domestication, heterosis, and yield, using network analysis?
- What is the best language and representation of dynamic networks (e.g. Systems Biology Markup Language [60])?
- How predictable or repeatable are evolutionary trajectories of networks in plants [e.g. [61]]?

Acknowledgements

This work was supported by Carnegie Institution for Science (S.Y.R.), grants from the National Science Foundation to S.Y.R. (DBI-0640769 and IOS-1026003), a grant from the National Research Foundation funded by the Korean Government (No. 20100017649) and POSCO TJ Park Science Fellowship to I.L.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Feist AM, Herrgård MJ, Thiele I, Reed JL, Palsson BØ: **Reconstruction of biochemical networks in microorganisms**. *Nat Rev Microbiol* 2009, **7**:129-143.

2. de Jong H: **Modeling and simulation of genetic regulatory systems: a literature review**. *J Comput Biol* 2002, **9**:67-103.

3. • Feist AM, Palsson BO: **The growing scope of applications of genome-scale metabolic reconstructions using *Escherichia coli***. *Nat Biotechnol* 2008, **26**:659-667.
This article reviews current and proposed research applications for genome-scale metabolic networks, including metabolic engineering, model-based discovery, and analysis of large-scale phenotype screens, using the *E. coli* metabolic network as an example.

4. Yamada T, Bork P: **Evolution of biomolecular networks: lessons from metabolic and protein interactions**. *Nat Rev Mol Cell Biol* 2009, **10**:791-803.

5. • Barabasi AL, Gulbahce N, Loscalzo J: **Network medicine: a network-based approach to human disease**. *Nat Rev Genet* 2010, **12**:56-68.
This review article summarizes general network properties of disease genes, and their application for disease gene prediction and disease classification. The same concept can be applied to predict phenotypic genes in plants.

6. Boone C, Bussey H, Andrews BJ: **Exploring genetic interactions and networks with yeast**. *Nat Rev Genet* 2007, **8**:437-449.

7. •• Koonin EV, Wolf YI: **Evolutionary systems biology: links between gene evolution and function**. *Curr Opin Biotechnol* 2006, **17**:481-487.
This review article comprehensively summarizes reported correlations between multiple genome-wide variables that characterize genomic function and evolution, and discusses their biological meanings.

8. Loewe L: **A framework for evolutionary systems biology**. *BMC Syst Biol* 2009, **3**:27.

9. Freeling M: **Bias in plant gene content following different sorts of duplication: tandem, whole-genome, segmental, or by transposition**. *Annu Rev Plant Biol* 2009, **60**:433-453.

10. Alves R, Chaleil RA, Sternberg MJ: **Evolution of enzymes in metabolism: a network perspective**. *J Mol Biol* 2002, **320**:751-770.

11. • Caetano-Anolles G, Yafremava LS, Gee H, Caetano-Anolles D, Kim HS, Mittenthal JE: **The origin and evolution of modern metabolism**. *Int J Biochem Cell Biol* 2009, **41**:285-297.
A good review of the main hypotheses proposed for metabolic enzyme and pathway evolution. Also presents an analysis of metabolic evolution based on protein structure data.

12. Diaz-Mejia JJ, Perez-Rueda E, Segovia L: **A network perspective on the evolution of metabolism by gene duplication**. *Genome Biol* 2007, **8**:R26.

13. Veron AS, Kaufmann K, Bornberg-Bauer E: **Evidence of interaction network evolution by whole-genome duplications: a case study in MADS-box proteins**. *Mol Biol Evol* 2007, **24**:670-678.

14. Papp B, Pal C, Hurst LD: **Metabolic network analysis of the causes and evolution of enzyme dispensability in yeast**. *Nature* 2004, **429**:661-664.

15. Birchler JA, Veitia RA: **The gene balance hypothesis: implications for gene regulation, quantitative traits and evolution**. *New Phytol* 2010, **186**:54-62.

16. Kuepfer L, Sauer U, Blank LM: **Metabolic functions of duplicate genes in *Saccharomyces cerevisiae***. *Genome Res* 2005, **15**:1421-1430.

17. •• Arabidopsis-Interactome-Mapping-Consortium: **Evidence for network evolution in an Arabidopsis interactome map**. *Science* 2011, **333**:601-607.
This first large-scale experiment-based protein–protein interaction map for a plant species covers about 10% of Arabidopsis genome with high confidence. Analysis of the mapped Arabidopsis interactome supports functional diversification of duplicated genes.

18. Casneuf T, De Bodt S, Raes J, Maere S, Van de Peer Y: **Nonrandom divergence of gene expression following gene and genome duplications in the flowering plant *Arabidopsis thaliana***. *Genome Biol* 2006, **7**:R13.

19. Beltrao P, Serrano L: **Specificity and evolvability in eukaryotic protein interaction networks**. *PLoS Comput Biol* 2007, **3**:e25.

20. Wagner A: **The yeast protein interaction network evolves rapidly and contains few redundant duplicate genes**. *Mol Biol Evol* 2001, **18**:1283-1292.

21. •• Shou C, Bhardwaj N, Lam HY, Yan KK, Kim PM, Snyder M, Gerstein MB: **Measuring the evolutionary rewiring of biological networks**. *PLoS Comput Biol* 2011, **7**:e1001050.
This article provides a general computational framework for analyzing rewiring rates between orthologous networks. Using the methods, the authors demonstrated that regulatory networks generally evolve faster than other types of molecular networks.

22. van Dam TJ, Snel B: **Protein complex evolution does not involve extensive network rewiring**. *PLoS Comput Biol* 2008, **4**:e1000132.

23. De Smet R, Van de Peer Y: **Redundancy and rewiring of genetic networks following genome-wide duplication events**. *Curr Opin Plant Biol* 2012, **15**:168-176.

24. Barabasi AL, Oltvai ZN: **Network biology: understanding the cell's functional organization**. *Nat Rev Genet* 2004, **5**:101-113.

25. Réka Albert:: **Scale-free networks in cell biology**. *J Cell Sci* 2005, **118**:4947-4957.

26. •• Lee I, Ambaru B, Thakkar P, Marcotte EM, Rhee SY: **Rational association of genes with traits using a genome-scale gene network for *Arabidopsis thaliana***. *Nat Biotechnol* 2010, **28**:149-156.
This paper presents a co-function gene network, dubbed AraNet, which covers about 75% of the Arabidopsis genome by integration of 24 different types of large-scale genomics data derived from five different species including yeast and animals. Using the AraNet, the authors demonstrated feasibility of network-based rational identification of novel genes for plant traits.

27. Fraser HB, Hirsh AE, Steinmetz LM, Scharfe C, Feldman MW: **Evolutionary rate in the protein interaction network**. *Science* 2002, **296**:750-752.

28. Vitkup D, Kharchenko P, Wagner A: **Influence of metabolic**
• **network structure and function on enzyme evolution**. *Genome Biol* 2006, **7**:R39.
The authors analyze how the evolutionary rates of enzymes in the yeast metabolic network differ depending on their position within the network and their presumed biological importance. They then contrast these findings with trends associated with protein interaction networks to propose basic differences in the two types of networks.

29. Hahn MW, Conant GC, Wagner A: **Molecular evolution in large genetic networks: does connectivity equal constraint?** *J Mol Evol* 2004, **58**:203-211.

30. Drummond DA, Raval A, Wilke CO: **A single determinant dominates the rate of yeast protein evolution**. *Mol Biol Evol* 2006, **23**:327-337.

31. Jeong H, Mason SP, Barabasi AL, Oltvai ZN: **Lethality and**
• **centrality in protein networks**. *Nature* 2001, **411**:41-42.
A report demonstrating that highly connected proteins in a yeast protein interaction network are more likely to be essential for organismal survival.

32. Zotenko E, Mestre J, O'Leary DP, Przytycka TM: **Why do hubs in the yeast protein interaction network tend to be essential: reexamining the connection between the network topology and essentiality**. *PLoS Comput Biol* 2008, **4**:e1000140.

33. del Rio G, Koschutzki D, Coello G: **How to identify essential genes from molecular networks?** *BMC Syst Biol* 2009, **3**:102.

34. Jovelin R, Phillips PC: **Evolutionary rates and centrality in the yeast gene regulatory network**. *Genome Biol* 2009, **10**:R35.

35. Kim WK, Marcotte EM: **Age-dependent evolution of the yeast protein interaction network suggests a limited role of gene duplication and divergence**. *PLoS Comput Biol* 2008, **4**:e1000232.

36. Liang H, Li WH: **Gene essentiality, gene duplicability and protein connectivity in human and mouse**. *Trends Genet* 2007, **23**:375-378.

37. Wu X, Qi X: **Genes encoding hub and bottleneck enzymes of the Arabidopsis metabolic network preferentially retain homeologs through whole genome duplication**. *BMC Evol Biol* 2010, **10**:145.

38. Yang L, Gaut BS: **Factors that contribute to variation in evolutionary rate among Arabidopsis genes**. *Mol Biol Evol* 2011, **28**:2359-2369.

39. Spirin V, Gelfand MS, Mironov AA, Mirny LA: **A metabolic network**
•• **in the evolutionary context: multiscale structure and modularity**. *Proc Natl Acad Sci USA* 2006, **103**:8774-8779.
The authors used a genomic association-based approach to identify evolutionary modules within the *E. coli* metabolic network and compared these modules to known metabolic pathways. They found that in some cases the identified evolutionary modules correspond to whole pathways. In other cases, the evolutionary histories were more complicated, as known pathways could be split into separate evolutionary modules, and some modules could contain portions of separate pathways.

40. von Mering C, Zdobnov EM, Tsoka S, Ciccarelli FD, Pereira-Leal JB, Ouzounis CA, Bork P: **Genome evolution reveals biochemical networks and functional modules**. *Proc Natl Acad Sci USA* 2003, **100**:15428-15433.

41. Snel B, Huynen MA: **Quantifying modularity in the evolution of**
• **biomolecular systems**. *Genome Res* 2004, **14**:391-397.
A systematic study of the evolutionary modularity of nine different sets of functional modules representing different biological networks and different data sources (e.g. manually curated vs. high-throughput data).

42. Campillos M, von Mering C, Jensen LJ, Bork P: **Identification and analysis of evolutionarily cohesive functional modules in protein networks**. *Genome Res* 2006, **16**:374-382.

43. Fokkens L, Snel B: **Cohesive versus flexible evolution of functional modules in eukaryotes**. *PLoS Comput Biol* 2009, **5**:e1000276.

44. Mukhtar MS, Carvunis AR, Dreze M, Epple P, Steinbrenner J,
•• Moore J, Tasan M, Galli M, Hao T, Nishimura MT *et al.*: **Independently evolved virulence effectors converge onto hubs in a plant immune system network**. *Science* 2011, **333**:596-601.
This paper presents a first map of host-pathogen protein–protein interactions between Arabidopsis and bacterial pathogens, based on high confidence large-scale yeast two hybrid analysis. Using the map, authors illustrated independently evolved virulence proteins in different pathogenic species interact with a limited set of hub proteins in the host plant Arabidopsis.

45. van Hoek MJ, Hogeweg P: **Metabolic adaptation after whole genome duplication**. *Mol Biol Evol* 2009, **26**:2441-2453.

46. Matias Rodrigues JF, Wagner A: **Evolutionary plasticity and innovations in complex metabolic reaction networks**. *PLoS Comput Biol* 2009, **5**:e1000613.

47. Samal A, Matias Rodrigues JF, Jost J, Martin OC, Wagner A: **Genotype networks in metabolic reaction spaces**. *BMC Syst Biol* 2010, **4**:30.

48. Dixon SJ, Fedyshyn Y, Koh JL, Prasad TS, Chahwan C, Chua G,
• Toufighi K, Baryshnikova A, Hayles J, Hoe KL *et al.*: **Significant conservation of synthetic lethal genetic interaction networks between distantly related eukaryotes**. *Proc Natl Acad Sci USA* 2008, **105**:16653-16658.
The authors measured the degree of evolutionary conservation of synthetic lethal genetic interaction between *Saccharomyces cerevisiae* (budding yeast) and *Schizosaccharomyces pombe* (fission yeast) and estimated that less than 30% of genetic interactions are common to both species.

49. Roguev A, Bandyopadhyay S, Zofall M, Zhang K, Fischer T,
• Collins SR, Qu H, Shales M, Park HO, Hayles J *et al.*: **Conservation and rewiring of functional modules revealed by an epistasis map in fission yeast**. *Science* 2008, **322**:405-410.
In this work, the authors compare not only negative genetic interactions (e.g. synthetic lethal interaction) but also positive genetic interactions between budding yeast and fission yeast, and found that positive genetic interaction that is enriched for within-pathway gene pairs is more conserved than negative genetic interaction that is enriched for between-pathway gene pairs ($>$50% and $<$18%, respectively). Thus, the low conservation of genetic interaction largely arose from evolutionary difference of cross-talk between functional modules.

50. Brady SM, Orlando DA, Lee JY, Wang JY, Koch J, Dinneny JR, Mace D, Ohler U, Benfey PN: **A high-resolution root spatiotemporal map reveals dominant expression patterns**. *Science* 2007, **318**:801-806.

51. Dinneny JR, Long TA, Wang JY, Jung JW, Mace D, Pointer S, Barron C, Brady SM, Schiefelbein J, Benfey PN: **Cell identity mediates the response of Arabidopsis roots to abiotic stress**. *Science* 2008, **320**:942-945.

52. Bassel GW, Lan H, Glaab E, Gibbs DJ, Gerjets T, Krasnogor N, Bonner AJ, Holdsworth MJ, Provart NJ: **Genome-wide network model capturing seed germination reveals coordinated regulation of plant cellular phase transitions**. *Proc Natl Acad Sci USA* 2011, **108**:9709-9714.

53. Brady SM, Zhang L, Megraw M, Martinez NJ, Jiang E, Yi CS, Liu W, Zeng A, Taylor-Teeples M, Kim D *et al.*: **A stele-enriched gene regulatory network in the Arabidopsis root**. *Mol Syst Biol* 2011, **7**:459.

54. Pop A, Huttenhower C, Iyer-Pascuzzi A, Benfey PN, Troyanskaya OG: **Integrated functional networks of process, tissue, and developmental stage specific interactions in *Arabidopsis thaliana***. *BMC Syst Biol* 2010, **4**:180.

55. Jimenez-Gomez JM, Wallace AD, Maloof JN: **Network analysis identifies ELF3 as a QTL for the shade avoidance response in Arabidopsis**. *PLoS Genet* 2010, **6**:.

56. Chan EK, Rowe HC, Corwin JA, Joseph B, Kliebenstein DJ:
• **Combining genome-wide association mapping and transcriptional networks to identify novel genes controlling glucosinolates in *Arabidopsis thaliana***. *PLoS Biol* 2011, **9**:e1001125.
The authors demonstrated that an Arabidopsis co-expression network can successfully prioritize the candidate loci from genome-wide association studies for glucosinolate variation.

57. Zhu C, Li X, Yu J: **Integrating rare-variant testing, function**
• **prediction, and gene network in composite resequencing-based genome-wide association studies (CR-GWAS)**. *Genes Genomes Genet* 2011, **1**:233-243.

This work proposes a novel integrative approach to prioritizing phenotype-associated genes from a genome-wide association study using whole-genome re-sequencing data and a genome-scale co-function gene network.

58. Jiao Y, Wickett NJ, Ayyampalayam S, Chanderbali AS, Landherr L, Ralph PE, Tomsho LP, Hu Y, Liang H, Soltis PS *et al.*: **Ancestral polyploidy in seed plants and angiosperms**. *Nature* 2011, **473**:97-100.

59. Kirschner M, Gerhart J: **Evolvability**. *Proc Natl Acad Sci USA* 1998, **95**:8420-8427.

60. Hucka M, Finney A: **Escalating model sizes and complexities call for standardized forms of representation**. *Mol Syst Biol* 2005, **1**: 2005.0011.

61. Poelwijk FJ, Kiviet DJ, Weinreich DM, Tans SJ: **Empirical fitness landscapes reveal accessible evolutionary paths**. *Nature* 2007, **445**:383-386.

62. Sweetlove LJ, Fell D, Fernie AR: **Getting to grips with the plant metabolic network**. *Biochem J* 2008, **409**:27-41.

63. Lalonde S, Ehrhardt DW, Loque D, Chen J, Rhee SY, Frommer WB: **Molecular and cellular approaches for the detection of protein–protein interactions: latest techniques and current limitations**. *Plant J* 2008, **53**:610-635.

64. Roth FP, Lipshitz HD, Andrews BJ: **Q&A: epistasis**. *J Biol* 2009, **8**:35.

65. Ma S, Gong Q, Bohnert HJ: **An Arabidopsis gene network based on the graphical Gaussian model**. *Genome Res* 2007, **17**:1614-1625.

66. Mutwil M, Klie S, Tohge T, Giorgi FM, Wilkins O, Campbell MM, Fernie AR, Usadel B, Nikoloski Z, Persson S: **PlaNet: combined sequence and expression comparisons across plant networks derived from seven species**. *Plant Cell* 2011, **23**:895-910.

67. Mochida K, Uehara-Yamaguchi Y, Yoshida T, Sakurai T, Shinozaki K: **Global landscape of a co-expressed gene network in barley and its application to gene discovery in Triticeae crops**. *Plant Cell Physiol* 2011, **52**:785-803.

68. Ficklin SP, Luo F, Feltus FA: **The association of multiple interacting genes with specific phenotypes in rice using gene coexpression networks**. *Plant Physiol* 2010, **154**:13-24.

69. Childs KL, Davidson RM, Buell CR: **Gene coexpression network analysis as a source of functional annotation for rice genes**. *PLoS ONE* 2011, **6**:e22196.

70. • Lee I, Seo YS, Coltrane D, Hwang S, Oh T, Marcotte EM, Ronald PC: **Genetic dissection of the biotic stress response using a genome-scale gene network for rice**. *Proc Natl Acad Sci USA* 2011, **108**:18548-18553.

The authors report a genome-scale gene co-function network for rice (*Oryza sativa*), which is the first such kind for crop species. This network covering 50% of the rice genome was used to predict new candidate genes for the rice innate immunity pathway. Three out of five candidate genes in transgenic rice have been validated to be involved in XA21-mediated innate immunity.

71. Yilmaz A, Mejia-Guerra MK, Kurz K, Liang X, Welch L, Grotewold E: **AGRIS: the Arabidopsis gene regulatory information server, an update**. *Nucleic Acids Res* 2011, **39**:D1118-D1122.

72. Berardini TZ, Mundodi S, Reiser L, Huala E, Garcia-Hernandez M, Zhang P, Mueller LA, Yoon J, Doyle A, Lander G *et al.*: **Functional annotation of the Arabidopsis genome using controlled vocabularies**. *Plant Physiol* 2004, **135**:745-755.

73. Zhang P, Dreher K, Karthikeyan A, Chi A, Pujar A, Caspi R, Karp P, Kirkup V, Latendresse M, Lee C *et al.*: **Creation of a genome-wide metabolic pathway database for Populus trichocarpa using a new approach for reconstruction and curation of metabolic pathways for plants**. *Plant Physiol* 2010, **153**: 1479-1491.