

# Making the right connections: biological networks in the light of evolution

Christopher G. Knight,<sup>1\*</sup> and John W. Pinney<sup>2</sup>

<sup>1</sup>Faculty of Life Sciences, The University of Manchester, Michael Smith Building, Oxford Road, Manchester M13 9PT, UK

<sup>2</sup>Centre for Bioinformatics, Division of Molecular Biosciences, Imperial College London, London SW7 2AZ, UK

**Our understanding of how evolution acts on biological networks remains patchy, as is our knowledge of how that action is best identified, modelled and understood. Starting with network structure and the evolution of protein–protein interaction networks, we briefly survey the ways in which network evolution is being addressed in the fields of systems biology, development and ecology. The approaches highlighted demonstrate a movement away from a focus on network topology towards a more integrated view, placing biological properties centre-stage. We argue that there remains great potential in a closer synergy between evolutionary biology and biological network analysis, although that may require the development of novel approaches and even different analogies for biological networks themselves.**

**Keywords:** correlation network; eQTL network; eurovision song contest; gene regulatory network; protein–protein interaction network

## Introduction

Many diverse and distinct biological systems may be represented as networks (Fig. 1). It is perhaps reasonable to expect evolutionary processes to act on each in a different way. Yet common network representations encourage the use of common analytical tools and suggest the potential for cross-fertilisation of ideas and techniques. Studying evolution at the level of the network representations of biological systems may therefore provide a broad and unified view of evolution itself. Conversely, an evolutionary view may be just what is required to make sense of increasingly numerous, complex and confusing networks. Systems biology (SB), one of the main modern purveyors of network representations of biological systems, sets out to navigate between biological levels, molecular and functional, in a mathematically explicit way.<sup>(1)</sup> It was to help navigate between such levels, molecular

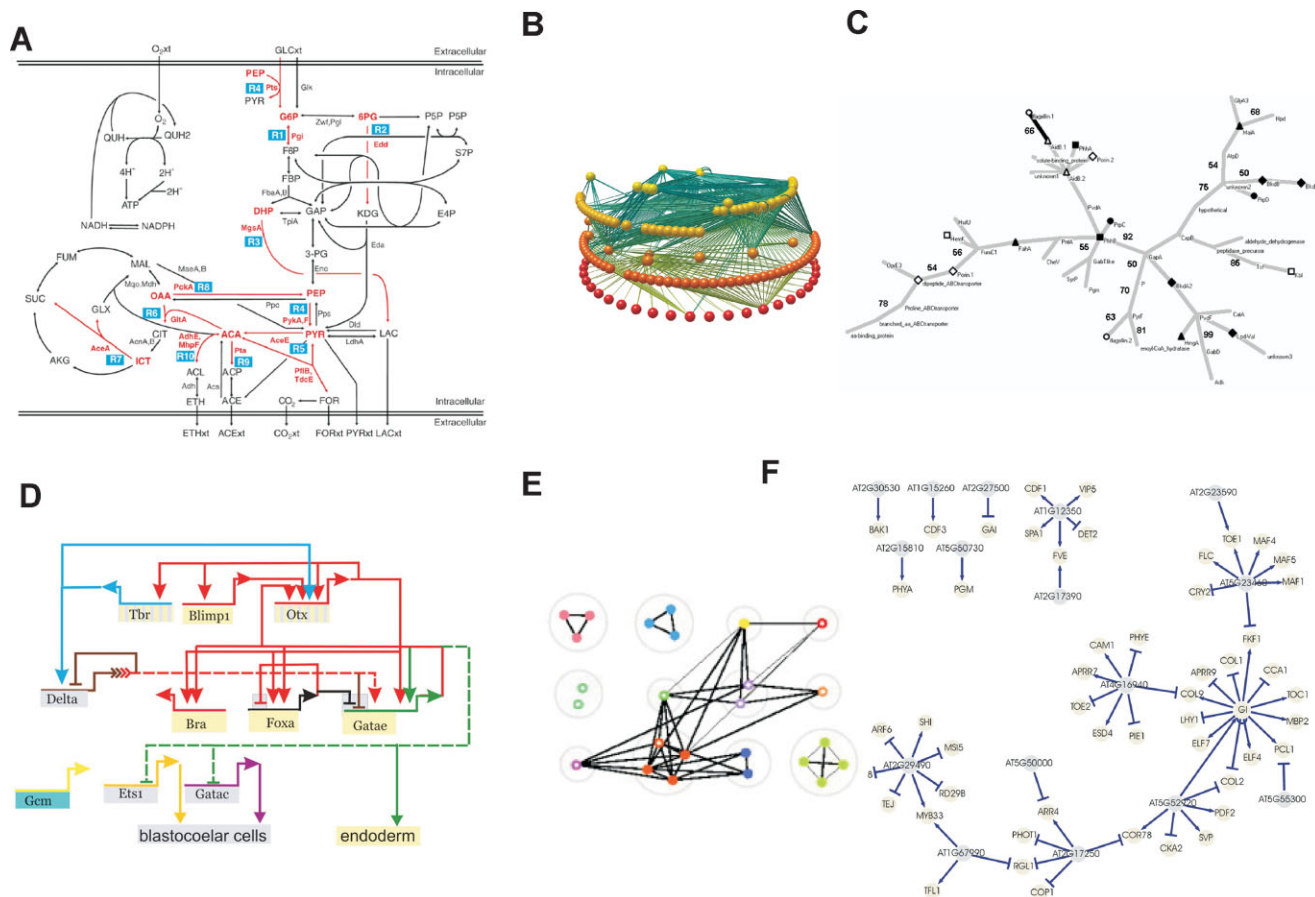
and organismic, that Dobzhansky coined the axiom ‘Nothing makes sense in biology except in the light of evolution’.<sup>(2)</sup> With networks now being invoked in fields from paleobiology<sup>(3)</sup> to human behaviour,<sup>(4)</sup> the issue is shifting from the existence and identity of such networks to their biology and the insights they might provide into evolution at the systems level. It is therefore appropriate to ask how much evolutionary sense is being made of biological networks.

Others have considered the rise of network thinking in relation to evolution<sup>(5)</sup> and the importance of network analyses in bridging the gap between evolutionary biology and other fields.<sup>(6)</sup> Specific issues of modularity,<sup>(7)</sup> robustness and evolvability<sup>(8)</sup> and the related phenomena of pleiotropy and epistasis<sup>(9)</sup> in relation to biological networks have been subjects of much interest. We necessarily encounter these themes, but focus mainly on existing paradigms of how networks evolve. We ask what can be understood about the biological mechanisms involved in network evolution and highlight emerging approaches that may help us find answers to this question.

## Network structure

The theory of complex networks and their properties has been extensively reviewed elsewhere, both generally<sup>(10)</sup> and regarding their biological applications.<sup>(11)</sup> Throughout the short history of this science, research has tended to focus on the description of a network’s structure in terms of global summary statistics such as its observed degree (*i.e.* connectivity) distribution, mean clustering coefficient or characteristic path length. The most famous example is the power law (or ‘scale-free’) degree distribution, where the frequency of nodes having a degree  $k$  has the form  $P(k) \propto k^{-\gamma}$ , where  $\gamma > 1$ . Networks of this form have few nodes with many connections, but many nodes with only one or two connections. Power law behaviour in the structure of networks was first noted by Price in the patterns of citations in scientific publications<sup>(12)</sup> and subsequently popularised by Barabasi and Albert, who observed power law distributions for network data concerning the World-Wide-Web, actor collaborations and a power grid.<sup>(13)</sup> Subsequent work from this and other research groups reported power law distributions for

**Abbreviations:** eQTL, expression quantitative trait locus; FBA, flux balance analysis; GRN, gene regulatory network; MST, minimal spanning tree; PPI, protein–protein interaction; QTL, quantitative trait locus; SB, systems biology.  
**\*Correspondence to:** C. G. Knight, Michael Smith Building, Oxford Road, Manchester, M13 9PT.  
E-mail: chris.knight@manchester.ac.uk



**Figure 1.** Recent examples of networks used in evolution-related studies in diverse areas. **A:** Metabolic network of central carbon metabolism in *E. coli*, as used for evaluating flux balance analysis (FBA) objective functions.<sup>(100)</sup> **B:** Food-web network of species in the Burgess Shale.<sup>(3)</sup> **C:** Correlation network of proteins affected in a bacterial experimental evolution.<sup>(72)</sup> **D:** Gene regulatory network (GRN) for endomesodermal specification in sea star.<sup>(90)</sup> **E:** Inferred ancestral chordate protein–protein interaction (PPI) network for bZIP transcription factors.<sup>(62)</sup> **F:** Regulatory network of genes involved in the transition to flowering in *Arabidopsis* inferred from expression quantitative trait locus (eQTLs).<sup>(89)</sup>

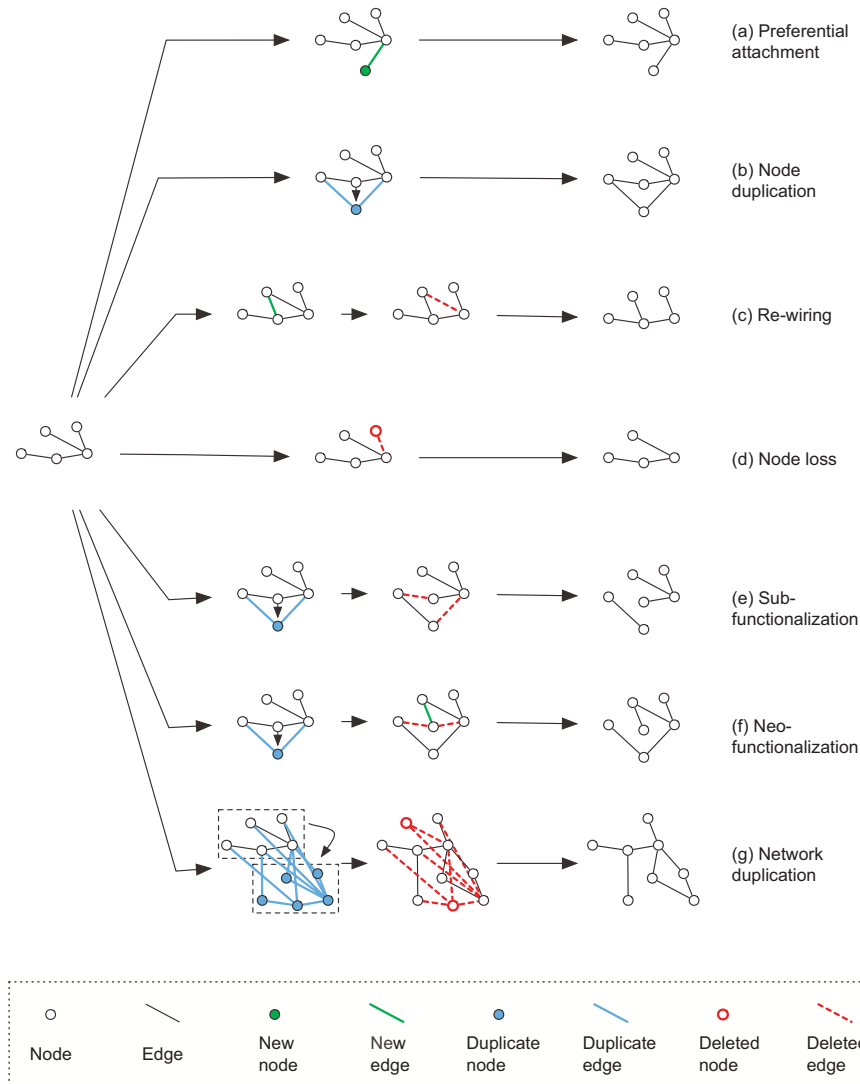
many biological networks, both subcellular [e.g. metabolic<sup>(14)</sup> and protein–protein interaction (PPI)<sup>(15)</sup>] and organismal.<sup>(4)</sup> It is important to note, however, that the degree distribution alone is a poor encapsulation of a network’s topology, since a given degree distribution may be satisfied by many different networks with substantially different architectures.<sup>(16)</sup> Additionally, there are various alternative forms proposed for the degree distributions of biological networks. There is currently much controversy over which of these might in fact fit the observed network data the best,<sup>(17)</sup> whether such distributions have any real biological significance,<sup>(18)</sup> and to what extent the values of network summary statistics might be affected by noise,<sup>(19)</sup> sampling<sup>(20)</sup> or data handling.<sup>(21)</sup>

### Evolution of topology and beyond

When researchers started to address the question of how these networks evolved, belief in the primacy of the degree

distribution led to a focus on evolutionary mechanisms that would generate power law networks.<sup>(22)</sup> Just as there may be many plausible topological models to fit a particular degree distribution, there are many plausible stochastic models of network evolution that could generate a given topology.<sup>(23)</sup> For example, the preferential attachment model (Fig. 2A)<sup>(13)</sup> is one simple way to generate a power-law network by the progressive addition of nodes, where each new node is attached to an existing node with a probability related to the degree of that node. However, preferential attachment seems a particularly unreasonable mechanism for the evolution of many biological systems. Several biologically motivated schemes incorporating node duplication (Fig. 2B) and subsequent loss and/or gain of interactions (Fig. 2C) have been proposed.<sup>(24,25)</sup>

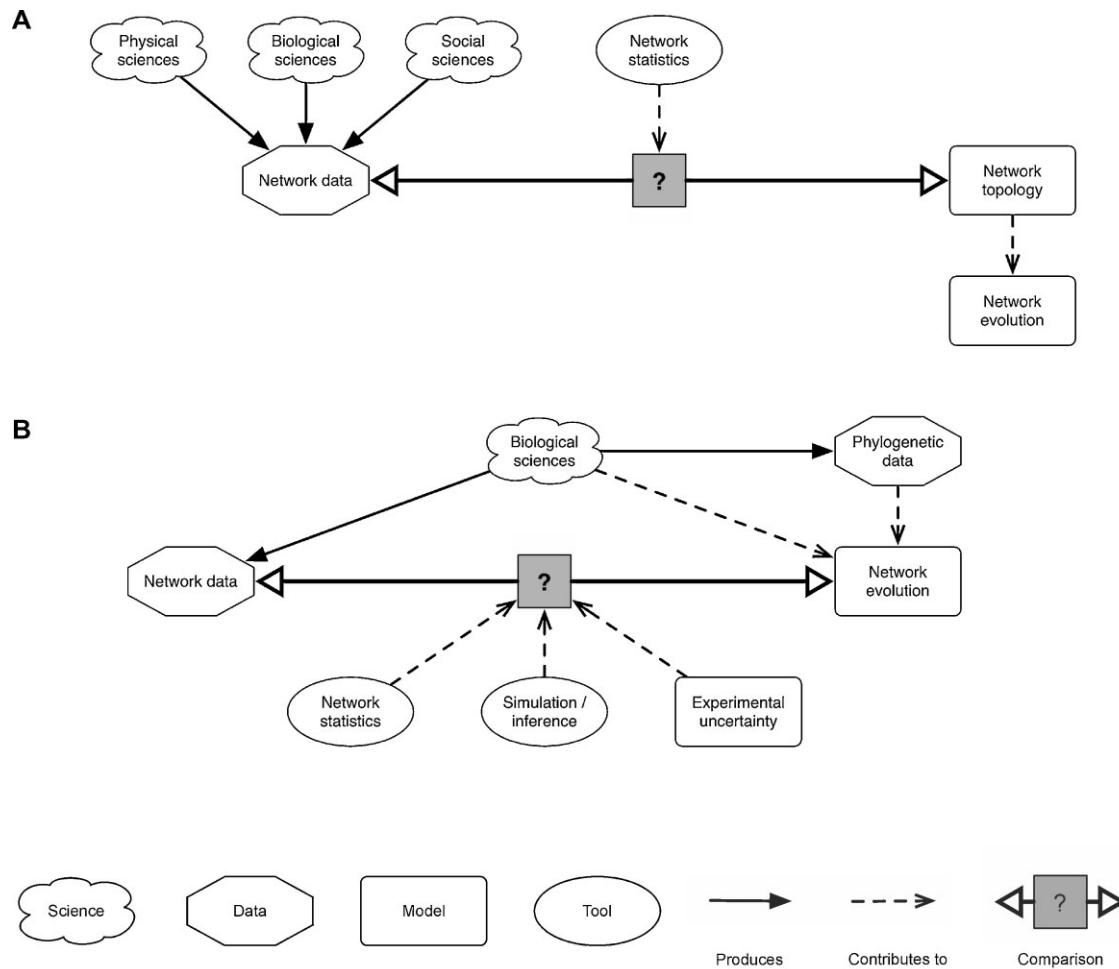
Much of this research into network evolution is typified by a paradigm in which a topological model is described, justified to some degree in its fit to observed network data, then subsequently discussed in terms of its evolutionary implications (Fig. 3A). However, in spite of its popularity, this



**Figure 2.** Illustration of some processes of network evolution. These processes range from **A**: the purely graph-theoretical concept of preferential attachment,<sup>(13)</sup> via increasingly biologically motivated concepts of **B**: node duplication, **C**: re-wiring, **D**: node loss, **E**: sub-functionalization and **F**: neo-functionalization, to **G**: network duplication, analogous to a whole-genome duplication event.

approach appears to have produced few insights into the true evolutionary mechanisms that resulted in present-day biological networks. It also fails to incorporate existing understanding of evolutionary processes. For example, models of evolving PPI networks or gene regulatory networks (GRNs) may need to incorporate gene deletion (Fig. 2D),<sup>(26)</sup> subfunctionalisation and/or neofunctionalisation of duplicated genes (Fig. 2E, F)<sup>(27)</sup> and whole-genome duplication (Fig. 2G).<sup>(28)</sup> Models incorporating such processes can use computational simulation or inference methods to compare the model to the observed data, to fit parameter values and to compare alternative models in terms of their relative likelihoods.<sup>(29)</sup> Choosing a biology-centric rather than topology-centric approach also allows incorporation of known

information concerning evolution of the specific system, *e.g.* phylogenetic trees for the genes involved and the gene duplications and losses that may be inferred from them by cross-species analysis.<sup>(30)</sup> As additional biological factors are implicated in the evolution of networks, it is possible to extend such models to include them. For example, the importance of population genetics in systems evolution has largely been ignored until now, but may prove to have a profound influence on network structures.<sup>(31)</sup> A suggestion for how this alternative research paradigm might be structured is outlined in Fig. 3B. Instead of focussing on a particular theoretical network topology, this approach uses existing biological knowledge to build a realistic model for the evolution of the specific network being studied.



**Figure 3.** Changing research paradigms in the study of biological network evolution. **A:** Throughout the development of network theory, biological networks have been of great interest as data-sets to be analysed alongside examples of technological (e.g. internet, world-wide-web, power grid) and social (e.g. friendship, collaboration) networks. Early work tended to focus on the development of simple models of archetypal network topologies. Although many authors were keen to address the evolution of biological networks, the evolutionary models developed were primarily designed to reproduce the simple topologies under consideration, and as such were rarely tested directly against the data. **B:** A more sophisticated research paradigm for studying the evolution of biological networks starts from the viewpoint that any evolutionary model should relate directly to the biological system under study, with reference to population genetics and genomics where appropriate. Using simulation and probabilistic inference methods, models of network evolution can be tested directly against the biological data, taking factors such as experimental uncertainties and biases into account.

How then can such biology-centric approaches be applied to network evolution? We consider briefly three areas where answers to this question are being worked out: SB, evolution of development ('evo-devo') and ecology.

## Systems biology

As popularised by Kitano,<sup>(1)</sup> SB concerns feedback of wet-lab and dry-lab experiments, putting together a mathematically explicit understanding of the structure and function of sub-cellular networks from their component parts. Since the challenge of constructing any such single network is great

(e.g. those encapsulated by the KEGG databases<sup>(32)</sup>), evolutionary change has typically received only peripheral attention. Furthermore, SB has traditionally advocated applying engineering approaches to biological systems,<sup>(33)</sup> which, while potentially useful for functional network analysis and synthetic biology,<sup>(34)</sup> does not easily accommodate thinking about the evolutionary (as opposed to engineering design) processes that sculpt biological systems. Nonetheless, once constructed, SB network models are adaptable to different tasks, including providing insight into the processes of evolution which constructed them.<sup>(35)</sup> Such an approach has proved useful in studying horizontal gene transfer,<sup>(36)</sup> enzyme dispensability<sup>(37)</sup> and minimised genomes.<sup>(38)</sup>

Evolutionary analysis of SB networks is a relatively new phenomenon; a longer-standing relationship of evolution to SB networks is as an experimental tool. An example is the use of evolution experiments in *Escherichia coli* to test metabolic network model predictions.<sup>(39)</sup> This was successful in as much as bacteria evolved to the model-predicted phenotype. However, the actual evolutionary process whereby the cells evolve to the predicted optimum involves mutations to genes outside the network model.<sup>(40)</sup> This implies that what natural selection identifies as 'the system' is not necessarily limited to specific networks (be they metabolic, transcriptional or at any other level). This finding – that evolution may define biological networks differently to biologists – is an important caveat to which we return below in the context of transcriptional networks.

Another insightful SB example in which evolution has been used as a tool involved experimentally 're-wiring' a transcription network by re-pairing transcription factors and their promoters.<sup>(41)</sup> This study shows the ability of transcriptional networks to tolerate substantial modification, but more importantly demonstrates an interaction (epistasis) between such introduced links and subsequent evolutionary changes: some added links seem consistently to enable more successful adaptation to stressful conditions than the wild-type network. The molecular basis of this evolution remains to be determined and could yield important insights into the relationship between transcriptional control networks and their targets. Such innovative SB approaches offer new ways to tackle old questions surrounding the nature of the evolutionary genetic phenomena of epistasis and pleiotropy.<sup>(42)</sup>

## Evo-devo

Transcription factor networks have also been a key area in evolutionary developmental biology, evo-devo.<sup>(43)</sup> However, unlike in most SB studies, evolution of network structure has been a focus in understanding network models. Similarly, the nature of development imparts different emphases from SB in the network models developed, notably on spatial aspects.<sup>(44)</sup> Evo-devo ideas have been developed in relation to biological network evolution, e.g. concerning homology,<sup>(45)</sup> that may be relevant in other spheres of network evolution. For instance, in the paradigm of eye development there is homology across vast evolutionary distances at the levels of high-level function (photo-reception) and key genes (notably *Pax6*), yet very different patterns at the intermediate levels of morphology (insect vs. vertebrate eyes) and the GRN. To make sense of such patterns, the focus has to be on the historical continuity of networks through evolution<sup>(45)</sup> and 'developmental system drift'.<sup>(46)</sup> The latter phenomenon refers to network structure changes with ultimate function remaining unaltered. It is identifiable in evolutionary network simulations<sup>(18,47)</sup> and

analyses<sup>(48)</sup> and is reminiscent of 'neutral network' evolution seen, for instance, in *in silico* systems of nucleic acid evolution.<sup>(49)</sup>

The mechanics of such network evolution are, however, very difficult to test, given the large timescales of interspecific evolution typical of evo-devo, where we only have developmental details for a few tips in a large phylogenetic tree. Nonetheless, 'micro evo-devo' is progressing in this direction.<sup>(50)</sup> Within-genus work in *Caenorhabditis* nematodes reveals quantitative changes within a signalling network, the overall phenotype remaining constant (changes are described as 'cryptic' since they are only apparent following experimental manipulation).<sup>(51)</sup> At shorter evolutionary timescales again, cryptic quantitative evolution has been identified in *Caenorhabditis elegans* over the course of laboratory culture,<sup>(52)</sup> and experimental evolution has been used.<sup>(53)</sup> Thus, it may be possible to use the *C. elegans* model system to look at steps of developmental network evolution individually and experimentally, which so far has been done only for microbial networks (e.g.<sup>(40)</sup>). Are these concepts and approaches applicable outside development? Quite probably – in the simulations showing developmental system drift the phenotype used is one or more gene expression levels,<sup>(18,47,48)</sup> which is as applicable to unicellular systems as development, and may be more tractable in simpler, experimental evolution systems. More generally, while apparent disconnections between evolutionary behaviour at genotypic, phenotypic and network levels may be particularly acute when the phenotype is as complex as an eye or vertebrate limb, there is no reason to believe that such relationships are any less subtle in simpler and potentially more experimentally tractable systems such as microbial metabolism. Indeed, similarly complex genetic relationships of orthology, paralogy and functional divergence undoubtedly exist, as evidenced in complete microbial genome sequences.<sup>(54)</sup>

## Ecology

Ecology has a long history of using networks – food webs (Fig. 1B) are some of the longest standing networks in any field (e.g. Briand collated 40 published webs over 25 years ago<sup>(55)</sup>). Like evo-devo, network studies in ecology have focused on evolution and the underlying biological (as opposed to purely graph-theoretical) processes. Similarly, network 'dynamics' refers to the evolution of structure over time (a sense in which it has also been used for PPI<sup>(28)</sup> and transcriptional networks<sup>(56)</sup> in contrast to the SB sense of the temporal kinetics of variables within a given network). However, the sorts of processes occurring among networks of biological taxa in ecological networks are scientifically rather distant from the subcellular networks considered so far, although arguably not as distant as the influential social,

citation and world-wide-web network paradigms.<sup>(57)</sup> For instance, an important distinguishing feature of ecological networks from either social or world-wide-web networks is variation in the abundance of the entities represented by network nodes.<sup>(57)</sup> This feature has important implications for designing appropriate null models for network structure<sup>(58)</sup> and is shared with, for instance, PPI networks. Taken to its limit, ecological network evolution is appropriately modelled *via* individual-based models.<sup>(59)</sup> However, even when topology alone has been used, ways have been found to frame meaningful tests of underlying evolutionary processes.<sup>(60)</sup>

Ecology's subject matter means that evolutionary processes acting on the individual or population, notably natural selection, will typically act at or below the level of individual network nodes rather than at the level of the complete network or above, as with subcellular networks. In ecology this has led to clear distinctions being drawn between network structure evolution (especially *via* extinctions) and evolution of the components in the network (phylogeny), highlighting the relationship between the two.<sup>(61)</sup> Such distinctions between levels of evolution are undoubtedly important in other systems, where, *e.g.* the distinction between the evolutionary history of a protein and the evolutionary history of its interactions may not be so obvious, but may be necessary to an understanding of what is going on in evolution (*e.g.*<sup>(62)</sup>). Ecological network evolution is also tackling the move from discrete (binary) networks to quantitative networks (weighted graphs showing the strength of interactions).<sup>(63)</sup> In addition to representing biological aspects of the system, quantitative networks can be used to analyse the effect of sampling effort,<sup>(64)</sup> something that is coming to prominence in PPI networks.<sup>(21)</sup>

## Discrete *versus* quantitative networks

The relationship of discrete 'wiring diagram' networks to quantitative (weighted) representations of the same biological systems is a current challenge for evolutionary network analysis across diverse fields. It is important firstly because the representation used affects the capture of evolutionarily important characteristics.<sup>(65)</sup> Secondly, as demonstrated in evo-devo analyses of nematode vulva development highlighted above, what ends up as discrete network evolution over long (inter-generic) evolutionary timescales (*e.g.*<sup>(66)</sup>) may start as quantitative changes over shorter (intra-generic) timescales.<sup>(51)</sup> Therefore, understanding the mechanisms of network change that occur in the individual steps of evolution will entail obtaining quantitative descriptions of the system, such as the proportion of different cell fate outcomes from signalling networks.<sup>(51)</sup>

SB has perhaps moved furthest down the road of exploring alternative discrete and quantitative models of the same or

similar systems. Thus, constraint-based analyses of metabolic networks, notably flux balance analysis (FBA), may require little more than discrete wiring diagrams<sup>(67)</sup> (Fig. 1A), but produce only limited steady-state predictions. In contrast, fully quantitative kinetic modelling provides a clearer view of quantities that are more easily measured in real systems (*e.g.* metabolite levels), but requires much more prior information,<sup>(68)</sup> the acquisition of which is a substantial bottleneck (for instance the enzyme kinetic parameters stored in the Sabio-RK database<sup>(69)</sup> are primarily culled from old literature and collected in highly varied conditions; high-throughput approaches using standardised conditions are only slowly being developed). This amounts to a trade-off between the requirement of discrete models for less detailed experimental knowledge and applicability to larger systems, and the greater, subtler and probably more realistic insight of quantitative models. This trade-off will become increasingly apparent across biological network analyses, where the network required to answer specific evolutionary questions will not necessarily be the most quantitative network available. Thus, so far in SB, primarily discrete FBA network analyses have been used effectively to tackle evolutionary questions (*e.g.*<sup>(38)</sup>).

## Phenomenology, null models and mechanism

We currently know so little about the phenomenology of network evolution that charting the 'natural history' of a network evolution will frequently be valuable in itself, without comparison to a null or expected change.<sup>(70)</sup> The natural history task is non-trivial, firstly in terms of separating true evolutionary change from apparent changes due to technical error.<sup>(21)</sup> Secondly, in our ignorance of ancestral network states, evolutionary reconstruction is a delicate exercise.<sup>(62)</sup> Perhaps the neatest answer to the first problem is to concentrate on approaches to network evolution that compare networks with equivalent technical errors. Such a focus argues against non-comparative analyses, such as attempting to draw evolutionary inferences from topological analysis of any single PPI network (see above) and against comparisons of networks constructed from different studies carried out in different laboratories at different times. Conversely, this is an argument for focusing on networks identified and compared in a single study, as in nematode vulva specification studies,<sup>(51)</sup> and the expression quantitative trait locus (eQTL) and correlation network examples discussed below. In a similar vein, the neatest solution to difficulties of phylogenetic reconstruction may be to focus on experimental systems where the ancestral states are observed directly, as in microbial experimental evolution systems<sup>(71)</sup> (see below). Even when all these approaches are

taken, the phenomenology of network evolution may remain challenging to unpick (e.g.<sup>(72)</sup>).

Focusing on single-study comparative and experimental approaches will help make valid evolutionary comparisons. However, the questions of the mechanisms and evolutionary forces underlying network evolution are clearly extremely important. Whether or not there is any empirical basis for doing so, evolutionary, and in particular adaptive, hypotheses of network change are the subject of widespread speculation.<sup>(73)</sup> Full population genetic null models of network evolution would bring clarity to these issues, but are in their infancy and a bottleneck to progress in this direction.<sup>(31)</sup> Nonetheless, simulations are yielding results, notably that non-adaptive evolution can result in pathway architectures more complex than strictly required for a selected function.<sup>(31,74)</sup> Simulation also indicates the importance of considering the network level of evolution, rather than concentrating only on genotype and high-level phenotype.<sup>(6)</sup> For instance, low within-population polymorphism at a single locus coupled with high between-population divergence might be interpreted as a signature of adaptive evolution. However, simulating the network context of a locus's evolution, that signature can be shown to arise in some cases simply from stabilising selection around a given phenotype.<sup>(18)</sup> Such simulations are currently rather abstracted relative to experiment, but, with care, adaptive hypotheses may be testable in real networks.<sup>(75)</sup>

Crucial to unpicking the reality of the roles of adaptive and non-adaptive processes in network evolution will be analysis and comparison of experimentally evolved strains undergoing adaptive, or specifically non-adaptive, evolution. Mutation accumulation experiments consider non-adaptive evolution. The focus is typically on the nature and rates of mutation.<sup>(76)</sup> Probing network effects experimentally (e.g. by genetic<sup>(77)</sup> or physical<sup>(51)</sup> manipulation) in mutation accumulation lines may therefore prove illuminating. For adaptive evolution, there is much more experimental evidence, not least from Lenski's paradigmatic long-term (now over 20 years) experimental evolution of *E. coli*.<sup>(78)</sup> This has demonstrated that the networks (or network 'modules') where adaptive mutations lie can be very distinct from the selected phenotype. Thus, while selection principally concerns metabolic traits (growth on glucose as a sole carbon source), adaptive mutations are seen involving DNA superhelicity and the stringent response.<sup>(78)</sup> The reverse situation is also seen in shorter-term evolution in another bacterium, with selection for a non-metabolic trait being effected by mutations in a gene principally controlling metabolism.<sup>(72)</sup> In a social experimental evolution system,<sup>(79)</sup> the wide diversity of genes where mutations can result in social cheating is of itself predicted to provide a route to evolving network complexity via 'conflict-generated churning'.<sup>(80)</sup> This raises important issues, beyond our scope here, around the role and meaning of pleiotropy and

modularity in network evolution.<sup>(7,9,81)</sup> In practice, unpredictability of the loci involved in adaptive evolution, even of relatively well understood networks, makes it difficult to use the 'candidate gene' approaches that have been successful in *evo-devo* for identifying the genetic basis of evolution (e.g.<sup>(82)</sup>). However, falling costs of complete genome sequencing are removing the need for a candidate gene when identifying small numbers of changes in experimentally evolved lines. This makes experimental adaptive evolution an increasingly promising area of research.

## Genetics, eQTLs and correlation networks

Beyond experimental evolution, the genetic basis of network evolution is harder to unpick. *Evo-devo* has a focus on transcriptional GRNs and within that, *cis*-acting regulatory changes.<sup>(83)</sup> This focus has undoubtedly enabled insight into interspecific evolution of transcription-factor networks, particularly of key developmental systems.<sup>(83)</sup> However, it is by no means clear that these are uniquely relevant genetic changes, even within developmental networks.<sup>(84)</sup> Beyond development, transcription factors are in fact rather poorly represented among the genetic loci responsible for the evolution of yeast transcription.<sup>(85)</sup>

Whatever the scope of the evolutionary role of *cis*-acting regulatory changes, *evo-devo* expresses an important objective: to identify patterns among mutations with particular roles in evolution. At a broad level, some relationships have been identified, such as weak negative correlations between genes' evolutionary rate and the connectivity of their proteins in PPI networks.<sup>(86)</sup> However, a much more interesting challenge will be to understand patterns of individual evolutionary steps at the molecular level. To obtain such an understanding, researchers will need to develop both global and probabilistic views of individual mutations in the context of biological networks. One approach to gaining such an understanding for expression is given by eQTLs.<sup>(87)</sup>

All quantitative trait locus (QTL) analyses link genotype and phenotype in evolution by assessing statistical associations between genetic loci that have evolved differences (markers) and some quantitative phenotype(s) of interest, typically as both segregate in a cross between evolutionarily diverged individuals. For eQTLs, the phenotypes of interest are transcription levels of genes. Since both genotype and phenotype in an eQTL association are defined by identifiable genetic loci, it is possible to construct a network where the nodes are genes and the directed and signed edges indicate that evolved changes at one locus are associated with transcription levels at the other locus<sup>(88,89)</sup> (Fig. 1F), *i.e.* a network exclusively of evolutionary effects. Thus, while eQTL networks share with GRNs the superficial similarity of showing loci affecting one another's transcription, the

approach is almost diametrically opposite – GRNs typically focus on evolutionary conservation (for instance the concept of evolutionarily conserved GRN kernels advanced by Hinman and Davidson<sup>(90)</sup>), whereas eQTL networks show only loci that change or whose expression is changed in evolution, typically over intraspecific evolutionary distances. Different approaches may be more appropriate for studying processes of network evolution at different evolutionary scales. The challenge will be to define how these views relate to one another and hence to obtain a global evolutionary view of relationships among molecules of the cell. Pursuing this end will complement and may provide insight into non-network studies that demonstrate the non-equivalence of evolutionary conservation and functional importance, with many genomic regions showing experimentally demonstrable functional roles without evolutionary constraint.<sup>(91)</sup>

QTL-based networks not only focus on evolution, they aim at a global unbiased view of relationships across a genome. Thus, they go some way towards identifying what biological networks are, as defined by evolution rather than biologists. This is an important issue in the light of the cases mentioned above, where the mutations underlying adaptive evolution occur outside the pre-defined network of the phenotype studied.<sup>(40,78)</sup> In eQTL networks, the network *is* the evolution, limiting evolutionary comparisons among networks. An alternative approach without this issue, but similarly aiming at a global unbiased view, is correlation network analysis. Assessing the correlation of a wide class of cellular components, most commonly metabolites, across a series of perturbations, commonly the minor environmental changes seen across biological replicates, identifies some components as strongly correlated and others not. These correlations may be interpreted as a fully connected, weighted graph, typically represented as a simpler discrete network by a process of edge removal based on the strength or significance of the correlations<sup>(92)</sup> (e.g. Fig. 1C).

Correlation networks represent co-regulation, rather than proximity in a biochemical network. For instance, metabolites adjacent in a metabolic network may not be the most closely correlated in a correlation network, an observation attested to experimentally<sup>(93)</sup> as well as theoretically.<sup>(94)</sup> Use of the term 'network' here is potentially problematic, in that its meaning is far looser than is usual for biological networks. While network graphs are generated to which graph theoretical approaches can be, and have been, applied, there is no direct physical interpretation of an edge in the graph in terms of a molecular interaction. However, correlation networks are very good for capturing and making sense of control relationships, potentially transient or otherwise elusive relationships, and how they change. For instance they have been successfully applied to understanding changes in currency markets.<sup>(95)</sup> They have also been used to look at genetic changes<sup>(96)</sup> including those changes responsible for adaptive evolu-

tion.<sup>(72)</sup> However, studying the evolution of such graphs is tricky – it depends on how the fully connected, weighted graph is simplified, which may involve arbitrary thresholds, or approaches such as minimal spanning trees (MSTs, Fig. 1C). MSTs may be used as a clustering tool, equivalent to single linkage clustering.<sup>(95)</sup> Thus the analysis of correlation networks merges into other, non-network analyses of correlation matrices. It may be that other such methods (e.g. approaches based on eigenvalues of correlation matrices<sup>(97)</sup>), while less visually striking and lacking the network buzzword, may be equally or more appropriate tools with which to tackle network evolution in terms of correlations among biological molecules.

Whether or not correlation networks are usefully classified as networks, they highlight an important point: in the evolution of biological networks, there are very general issues of approach that remain open. As we have seen, much current work stems from the popular concept of the cellular wiring diagram, and hence by implication the role of evolution is equivalent to that of re-wiring a radio.<sup>(33)</sup> Many of the analytical techniques discussed above are firmly rooted in this metaphor, which may be applied to network evolution in more or less biologically reasonable ways (see Fig. 2). The wiring diagram as a biological analogy clearly has value, primarily in the clarity it brings, enabling the application of engineering approaches, both analytical and synthetic. However, its adequacy as a biological analogy is particularly questionable when it comes to evolution. Actual wiring diagrams are descriptions of how individual, static elements designed to perform specific functions – transistors, resistors, *etc.* – are arranged in a design to perform specific roles in a higher level function, e.g. transduction of radio waves into sound. It is a very static view with design and purpose inherent in the image. It may be possible to navigate such issues, but they are particularly problematic when it comes to understanding network evolution.<sup>(31)</sup>

The wiring-diagram analogy for biological systems was borrowed from one of many forms of network analysed in the physical sciences. A surprising subject for network analysis to emerge from the physical sciences more recently is the Eurovision song contest.<sup>(98)</sup> We suggest that this may be a more appropriate analogy to borrow when it comes to the evolution of biological networks. The nature of the analogy is sketched in Table 1. Briefly, like an organism, the Eurovision song contest is a system that 'works', as evidenced by its survival. There is a role for the effectiveness of the individual network elements (countries, *cf.* biological molecules) at doing their apparent jobs in the system (producing music, *cf.* biological functions). But while necessary, these functions do not determine how the system as a whole functions or evolves over time. System function and evolution is underlain by a variety of relationships among the elements, of varying strengths and degrees of permanence. This implies that



**Table 1.** Comparison of analogies for subcellular biological networks

	<b>Biological network</b>	<b>Wiring diagram</b>	<b>Eurovision song contest</b>
<b>System</b>	Living organism	Electronic device (e.g. radio)	Music competition
<b>Elements</b>	Biological molecules	Electronic components	Countries' representatives
<b>Nature of elements</b>	Complex chemical entities with evolutionary histories	Minimal elements designed to perform simple tasks	Complex decision-making units with historical continuity
<b>Clearly defined element functions</b>	Performing biological functions (e.g. as described in GO terms)	Component specifications	Performing songs
<b>Other element features involved in network function</b>	Genetic location	None	Geographical proximity
	Transcriptional, translational, post-translational and degradation control		Politics
	More or less specific PPIs		Cultural history
	Other known and unknown relationships		Other unknown relationships
<b>Performance measure</b>	Inclusive fitness of an individual within a population	Performance of one or more pre-defined functions	International TV audience figures
<b>Evolutionary step</b>	DNA mutation	Addition or removal of connection or component	Change in individual voting behaviour

individual network changes (particular genetic changes in a cell, *cf.* switches among cliques of countries voting for each other in the Eurovision analogy) may be very difficult to attribute to specific causes and of limited significance in themselves. Nonetheless, analysis of network evolution remains a route to understanding system function (e.g. analysis of changes in Eurovision voting patterns successfully identified unofficial cliques of countries<sup>(98)</sup>).

The analogy used for biological systems affects the sorts of questions, experiments and analyses expected to be most fruitful. Thus, the Eurovision analogy (Table 1), while less appropriate to engineering than a wiring diagram, presupposes a more dynamic and complex system. As with the Eurovision analysis itself,<sup>(98)</sup> this suggests eschewing static approaches, such as degree distribution analysis, in favour of network analyses designed to deal with dynamically changing network topology. Thus, the Eurovision analogy fits better than the traditional one with dynamic evolutionary network phenomena discussed above such as developmental system drift<sup>(46)</sup> and conflict-generated churning.<sup>(80)</sup>

## Conclusions

The application of network approaches to biological systems has gone through something of a hype cycle,<sup>(99)</sup> where the 'peak of inflated expectations' corresponded to the widespread excitement, now passed, about scale-free networks and purely topology-based network approaches more generally, but a 'plateau of productivity' has not yet been reached. Part of the

excess hype comprised the over-zealous application of ideas across very different fields, when the hypothesis (power-law degree distributions) was both weak and in many cases incorrect.<sup>(16)</sup> However, that does not preclude useful cross-fertilisation of network ideas and techniques from different branches of biology. Research on the evolution of biological networks seems ideally placed to help steer a course between graph theory devoid of biological realism and experimental 'ridicologram' networks devoid of theory.

In this essay we have highlighted some ways in which the evolution of biological networks is currently being studied, drawing on examples from diverse fields. It is clear that various issues of network evolution are common across these research areas, including challenges such as the need to move from discrete wiring diagram networks to quantitative weighted graphs, a change that may be assisted by moving away from the wiring diagram metaphor or analogy itself (Table 1). A major focus has to be on the genetic basis of network evolution and how to take as broad and unbiased a view as possible of the networks in question. The approaches highlighted show how an explicitly evolutionary view of biological networks helps us to draw links between fields and to frame wider questions about the meaning of biological networks. The paradigm of understanding biological networks caricatured in Fig. 3A, with evolution as a downstream appendix, is clearly inadequate. Alternatives are being developed. However, there is much work to be done before we reach the more tightly integrated paradigm envisioned in Fig. 3B, where we can make sense of biological networks in the light of their evolution.

**Acknowledgments:** Thanks to Michael Stumpf for a critical reading of this manuscript. J. W. P. is supported by a University Research Fellowship from the Royal Society and C. G. K. is supported by a Research Career Development Fellowship from the Wellcome Trust.

## References

1. Kitano, H., Systems biology: a brief overview. *Science* 2002. **295**: 1662–1664.
2. Dobzhansky, T., Biology, molecular and organismic. *Am Zool* 1964. **4**: 443–452.
3. Dunne, J. A., Williams, R. J., Martinez, N. D., Wood, R. A. and Erwin, D. H., Compilation and network analyses of cambrian food webs. *PLoS Biol* 2008. **6**: e102.
4. Schneeberger, A., Mercer, C. H., Gregson, S. A., Ferguson, N. M., Nyamukapa, C. A., *et al.* Scale-free networks and sexually transmitted diseases: a description of observed patterns of sexual contacts in Britain and Zimbabwe. *Sex Transm Dis* 2004. **31**: 380–387.
5. Proulx, S. R., Promislow, D. E. L. and Phillips, P. C., Network thinking in ecology and evolution. *Trends Ecol Evol* 2005. **20**: 345–353.
6. Wilkins, A. S., Between “design” and “bricolage”: genetic networks, levels of selection, and adaptive evolution. *Proc Natl Acad Sci U S A* 2007. **104**: 8590–8596.
7. Wagner, G. P., Pavlicev, M. and Cheverud, J. M., The road to modularity. *Nat Rev Genet* 2007. **8**: 921–931.
8. Kitano, H., Biological robustness. *Nat Rev Genet* 2004. **5**: 826–837.
9. Tyler, A. L., Asselbergs, F. W., Williams, S. M. and Moore, J. H., Shadows of complexity: what biological networks reveal about epistasis and pleiotropy. *BioEssays* 2009. **31**: 220–227.
10. Newman, M., The structure and function of complex networks. *SIAM Rev* 2003. **45**: 167.
11. Mason, O. and Verwoerd, M., Graph theory and networks in Biology. *IET Syst Biol* 2007. **1**: 89–119.
12. Price, D. J., Networks of scientific papers. *Science* 1965. **149**: 510–515.
13. Barabasi, A. L. and Albert, R., Emergence of scaling in random networks. *Science* 1999. **286**: 509–512.
14. Jeong, H., Tombor, B., Albert, R., Oltvai, Z. N. and Barabási, A. L., The large-scale organization of metabolic networks. *Nature* 2000. **407**: 651–654.
15. Jeong, H., Mason, S. P., Barabási, A. L. and Oltvai, Z. N., Lethality and centrality in protein networks. *Nature* 2001. **411**: 41–42.
16. Keller, E. F., Revisiting “scale-free” networks. *BioEssays* 2005. **27**: 1060–1068.
17. Stumpf, M., Ingram, P., Nouvel, I. and Wiuf, C., Statistical model selection methods applied to biological networks. *Trans Comput Syst Biol* 2005. **3**: 65–77.
18. Siegal, M. L., Promislow, D. E. and Bergman, A., Functional and evolutionary inference in gene networks: does topology matter? *Genetica* 2007. **129**: 83–103.
19. Lin, N. and Zhao, H., Are scale-free networks robust to measurement errors? *BMC Bioinf* 2005. **6**: 119.
20. Han, J. D., Dupuy, D., Bertin, N., Cusick, M. and Vidal, M., Effect of sampling on topology predictions of protein–protein interaction networks. *Nat Biotechnol* 2005. **23**: 839–844.
21. Hakes, L., Pinney, J. W., Robertson, D. L. and Lovell, S. C., Protein–protein interaction networks and biology—what’s the connection? *Nat Biotechnol* 2008. **26**: 69–72.
22. Dorogovtsev, S. and Mendes, J., Evolution of networks. *Adv Phys* 2002. **51**: 1079–1187.
23. Lesne, A., Complex Networks: from Graph Theory to Biology. *Lett Math Phys* 2006. **78**: 235–262.
24. Vazquez, A., Flammini, A., Maritan, A. and Vespignani, A., Modeling of protein interaction networks. *ComPlexUs* 2003. **1**: 38–44.
25. Chung, F., Lu, L., Dewey, T. G. and Galas, D. J., Duplication models for biological networks. *J Comput Biol* 2003. **10**: 677–687.
26. Farid, N. and Christensen, K., Evolving networks through deletion and duplication. *New J Phys* 2006. **8**: 212.
27. Chung, W. Y., Albert, R., Albert, I., Nekrutenko, A. and Makova, K. D., Rapid and asymmetric divergence of duplicate genes in the human gene coexpression network. *BMC Bioinf* 2006. **7**: 46.
28. Evlampiev, K. and Isambert, H., Evolution of protein interaction networks by whole genome duplication and domain shuffling. *BMC Syst Biol* 2007. **1**: 49.
29. Wiuf, C., Brameier, M., Hagberg, O. and Stumpf, M. P., A likelihood approach to analysis of network data. *Proc Natl Acad Sci U S A* 2006. **103**: 7566–7570.
30. Levasseur, A., Pontarotti, P., Poch, O. and Thompson, J. D., Strategies for reliable exploitation of evolutionary concepts in high throughput biology. *Evol Bioinf* 2008.
31. Lynch, M., The evolution of genetic networks by non-adaptive processes. *Nat Rev Genet* 2007. **8**: 803–813.
32. Kanehisa, M., Araki, M., Goto, S., Hattori, M., Hirakawa, M., *et al.* KEGG for linking genomes to life and the environment. *Nucleic Acids Res* 2008. **36**: D480–D484.
33. Lazebnik, Y., Can a biologist fix a radio?—Or, what I learned while studying apoptosis. *Cancer Cell* 2002. **2**: 179–182.
34. Andrianantoandro, E., Basu, S., Karig, D. K. and Weiss, R., Synthetic biology: new engineering rules for an emerging discipline. *Mol Syst Biol* 2008. **2**: 0028.
35. Feist, A. M. and Palsson, B. O., The growing scope of applications of genome-scale metabolic reconstructions using *Escherichia coli*. *Nat Biotechnol* 2008. **26**: 659–6567.
36. Pal, C., Papp, B. and Lercher, M. J., Adaptive evolution of bacterial metabolic networks by horizontal gene transfer. *Nat Genet* 2005. **37**: 1372–1375.
37. Harrison, R., Papp, B., Pal, C., Oliver, S. G. and Delneri, D., Plasticity of genetic interactions in metabolic networks of yeast. *Proc Natl Acad Sci U S A* 2007. **104**: 2307–2312.
38. Pal, C., Papp, B., Lercher, M. J., Csermely, P., Oliver, S. G., *et al.* Chance and necessity in the evolution of minimal metabolic networks. *Nature* 2006. **440**: 667–670.
39. Ibarra, R. U., Edwards, J. S. and Palsson, B. O., *Escherichia coli* K-12 undergoes adaptive evolution to achieve in silico predicted optimal growth. *Nature* 2002. **420**: 186–189.
40. Herring, C. D., Raghunathan, A., Honisch, C., Patel, T., Applebee, M. K., *et al.* Comparative genome sequencing of *Escherichia coli* allows observation of bacterial evolution on a laboratory timescale. *Nat Genet* 2006. **38**: 1406–1412.
41. Isalan, M., Lemerle, C., Michalodimitrakis, K., Horn, C., Beltrao, P., *et al.* Evolvability and hierarchy in rewired bacterial gene networks. *Nature* 2008. **452**: 840–845.
42. Bateson, W., Facts limiting the theory of heredity. *Science* 1907. **26**: 649–660.
43. Muller, G. B., Evo-devo: extending the evolutionary synthesis. *Nat Rev Genet* 2007. **8**: 943–949.
44. Salazar-Ciudad, I. and Jernvall, J., A gene network model accounting for development and evolution of mammalian teeth. *Proc Natl Acad Sci U S A* 2002. **99**: 8116–8120.
45. Wagner, G. P., The developmental genetics of homology. *Nat Rev Genet* 2007. **8**: 473–479.
46. True, J. R. and Haag, E. S., Developmental system drift and flexibility in evolutionary trajectories. *Evol Dev* 2001. **3**: 109–119.
47. Johnson, N. A. and Porter, A. H., Evolution of branched regulatory genetic pathways: directional selection on pleiotropic loci accelerates developmental system drift. *Genetica* 2007. **129**: 57–70.
48. Munteanu, A. and Sole, R. V., Neutrality and robustness in evo-devo: emergence of lateral inhibition. *PLoS Comput Biol* 2008. **4**: e1000226.
49. Cowperthwaite, M. C. and Meyers, L. A., How mutational networks shape evolution: lessons from RNA models. *Annu Rev Ecol Evol Syst* 2007. **38**: 203–230.
50. Johnson, N. A., The Micro-evolution of development. *Genetica* 2007. **129**: 1–5.
51. Felix, M. A., Cryptic quantitative evolution of the vulva intercellular signaling network in *Caenorhabditis*. *Curr Biol* 2007. **17**: 103–114.

52. Patel, M. N., Knight, C. G., Karageorgi, C. and Leroi, A. M., Evolution of germ-line signals that regulate growth and aging in nematodes. *Proc Natl Acad Sci U S A* 2002. **99**: 769–774.
53. Navas, A., Cobas, G., Talavera, M., Ayala, J. A., Lopez, J. A., et al. Experimental validation of Haldane's hypothesis on the role of infection as an evolutionary force for Metazoans. *Proc Natl Acad Sci U S A* 2007. **104**: 13728–13731.
54. Koonin, E. V., Orthologs, paralogs, and evolutionary genomics. *Annu Rev Genet* 2005. **39**: 309–338.
55. Briand, F., Environmental control of food web structure. *Ecology* 1983. **64**: 253–263.
56. Madan Babu, M., Teichmann, S. A. and Aravind, L., Evolutionary dynamics of prokaryotic transcriptional regulatory networks. *J Mol Biol* 2006. **358**: 614–633.
57. Montoya, J. M., Pimm, S. L. and Sole, R. V., Ecological networks and their fragility. *Nature* 2006. **442**: 259–264.
58. Vazquez, D. P., Melian, C. J., Williams, N. M., Bluthgen, N., Krasnov, B. R., et al. Species abundance and asymmetric interaction strength in ecological networks. *Oikos* 2007. **116**: 1120–1127.
59. Bell, G., The evolution of trophic structure. *Heredity* 2007. **99**: 494–505.
60. Santamaria, L. and Rodriguez-Girones, M. A., Linkage rules for plant-pollinator networks: trait complementarity or exploitation barriers? *PLoS Biol* 2007. **5**: e31.
61. Rezende, E. L., Lavabre, J. E., Guimaraes, P. R., Jordano, P. and Bascompte, J., Non-random coextinctions in phylogenetically structured mutualistic networks. *Nature* 2007. **448**: 925–928.
62. Pinney, J. W., Amoutzias, G. D., Rattray, M. and Robertson, D. L., Reconstruction of ancestral protein interaction networks for the bZIP transcription factors. *Proc Natl Acad Sci U S A* 2007. **104**: 20449–20453.
63. Bluthgen, N., Menzel, F., Hovestadt, T. and Fiala, B., Specialization, constraints, and conflicting interests in mutualistic networks. *Curr Biol* 2007. **17**: 341–346.
64. Nielsen, A. and Bascompte, J., Ecological networks, nestedness and sampling effort. *J Ecol* 2007. **95**: 1134–1141.
65. Bluthgen, N. and Menzel, F., Measuring specialization in species interaction networks. *BMC Ecol* 2006. **6**: 9.
66. Felix, M. A. and Sternberg, P. W., Two nested gonadal inductions of the vulva in nematodes. *Development* 1997. **124**: 253–259.
67. Kauffman, K. J., Prakash, P. and Edwards, J. S., Advances in flux balance analysis. *Curr Opin Biotechnol* 2003. **14**: 491–496.
68. Gombert, A. K. and Nielsen, J., Mathematical modelling of metabolism. *Curr Opin Biotechnol* 2000. **11**: 180–186.
69. Rojas, I., Golebiewski, M., Kania, R., Krebs, O., Mir, S., et al. SABIO-RK: a database for biochemical reactions and their kinetics. *BMC Syst Biol* 2007. **1**: S1: S6.
70. Stumpf, M. P., Kelly, W. P., Thorne, T. and Wiuf, C., Evolution at the system level: the natural history of protein interaction networks. *Trends Ecol Evol* 2007. **22**: 366–373.
71. Elena, S. F. and Lenski, R. E., Evolution experiments with microorganisms: the dynamics and genetic bases of adaptation. *Nat Rev Genet* 2003. **4**: 457–469.
72. Knight, C. G., Zitzmann, N., Prabhakar, S., Antrobus, R., Dwek, R., et al. Unraveling adaptive evolution: how a single point mutation affects the protein coregulation network. *Nat Genet* 2006. **38**: 1015–1022.
73. Lynch, M., The frailty of adaptive hypotheses for the origins of organismal complexity. *Proc Natl Acad Sci U S A* 2007. **104**: 8597–8604.
74. Soyer, O. S. and Bonhoeffer, S., Evolution of complexity in signaling pathways. *Proc Natl Acad Sci U S A* 2006. **103**: 16337–16342.
75. Papp, B., Teusink, B. and Notebaart, R. A., A critical view of metabolic network adaptations. *HFSP J* 2009. **3**: 24–35.
76. Lynch, M., Sung, W., Morris, K., Coffey, N., Landry, C. R., et al. A genome-wide view of the spectrum of spontaneous mutations in yeast. *Proc Natl Acad Sci U S A* 2008. **105**: 9272–9277.
77. Cooper, T. F., Remold, S. K., Lenski, R. E. and Schneider, D., Expression profiles reveal parallel evolution of epistatic interactions involving the CRP regulon in *Escherichia coli*. *PLoS Genet* 2008. **4**: e35.
78. Philippe, N., Crozat, E., Lenski, R. E. and Schneider, D., Evolution of global regulatory networks during a long-term experiment with *Escherichia coli*. *BioEssays* 2007. **29**: 846–860.
79. Foster, K. R., Parkinson, K. and Thompson, C. R., What can microbial genetics teach sociobiology? *Trends Genet* 2007. **23**: 74–80.
80. Santorelli, L. A., Thompson, C. R., Villegas, E., Svetz, J., Dinh, C., et al. Facultative cheater mutants reveal the genetic complexity of cooperation in social amoebae. *Nature* 2008. **451**: 1107–1110.
81. Beaumont, H. J. E., Gehrig, S. M., Kassen, R., Knight, C. G., Malone, J., et al. The genetics of phenotypic innovation., In: Logan, N. A., Lappin-Scott, H. M. and Oyston, P. C. F. editors. *Prokaryotic Diversity: Mechanisms and Significance (Society for General Microbiology Symposium 66)*. Cambridge, Cambridge University Press, 2006. p. 91–104.
82. Sucena, E. and Stern, D. L., Divergence of larval morphology between *Drosophila sechellia* and its sibling species caused by cis-regulatory evolution of *ovo/shaven-baby*. *Proc Natl Acad Sci U S A* 2000. **97**: 4530–4534.
83. Erwin, D. H. and Davidson, E. H., The evolution of hierarchical gene regulatory networks. *Nat Rev Genet* 2009. **10**: 141–148.
84. Alonso, C. R. and Wilkins, A. S., The molecular elements that underlie developmental evolution. *Nat Rev Genet* 2005. **6**: 709–715.
85. Yvert, G., Brem, R. B., Whittle, J., Akey, J. M., Foss, E., et al. Trans-acting regulatory variation in *Saccharomyces cerevisiae* and the role of transcription factors. *Nat Genet* 2003. **35**: 57–64.
86. Koonin, E. V. and Wolf, Y. I., Evolutionary systems biology: links between gene evolution and function. *Curr Opin Biotechnol* 2006. **17**: 481–487.
87. Rockman, M. V. and Kruglyak, L., Genetics of global gene expression. *Nat Rev Genet* 2006. **7**: 862–872.
88. Bing, N. and Hoeschele, I., Genetical genomics analysis of a yeast segregant population for transcription network inference. *Genetics* 2005. **170**: 533–542.
89. Keurentjes, J. J., Fu, J., Terpstra, I. R., Garcia, J. M., van den Ackerveken, G., et al. Regulatory network construction in *Arabidopsis* by using genome-wide gene expression quantitative trait loci. *Proc Natl Acad Sci U S A* 2007. **104**: 1708–1713.
90. Hinman, V. F. and Davidson, E. H., Evolutionary plasticity of developmental gene regulatory network architecture. *Proc Natl Acad Sci U S A* 2007. **104**: 19404–19409.
91. Birney, E., Stamatoyannopoulos, J. A., Dutta, A., Guigo, R., Gingeras, T. R., et al. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* 2007. **447**: 799–816.
92. Steuer, R., On the analysis and interpretation of correlations in metabolomic data. *Brief Bioinf* 2006. **7**: 151–158.
93. Muller-Linow, M., Weckwerth, W. and Hutt, M. T., Consistency analysis of metabolic correlation networks. *BMC Syst Biol* 2007. **1**: 44.
94. Camacho, D., de la Fuente, A. and Mendes, P., The origin of correlations in metabolomics data. *Metabolomics* 2005. **1**: 53–63.
95. McDonald, M., Suleman, O., Williams, S., Howison, S. and Johnson, N. F., Detecting a currency's dominance or dependence using foreign exchange network trees. *Phys Rev E Stat Nonlin Soft Matter Phys* 2005. **72**: 046106.
96. Weckwerth, W., Loureiro, M. E., Wenzel, K. and Fiehn, O., Differential metabolic networks unravel the effects of silent plant phenotypes. *Proc Natl Acad Sci U S A* 2004. **101**: 7809–7814.
97. Baier, G., Muller, M., Stephani, U. and Muhle, H., Characterizing correlation changes of complex pattern transitions: the case of epileptic activity. *Phys Lett A* 2007. **363**: 290–296.
98. Fenn, D., Suleman, O., Efstathiou, J. and Johnson, N. F., How does Europe make its mind up? Connections, cliques, and compatibility between countries in the Eurovision Song Contest. *Phys. A* 2006. **360**: 576–598.
99. Geels, F. W., Pieters, T. and Snelders, S., Cultural enthusiasm, resistance and the societal embedding of new technologies: psychotropic drugs in the 20th century. *Technol Anal Strateg Manage* 2007. **19**: 145–165.
100. Schuetz, R., Kuepfer, L. and Sauer, U., Systematic evaluation of objective functions for predicting intracellular fluxes in *Escherichia coli*. *Mol Syst Biol* 2007. **3**: 119.