BRIEFINGS IN FUNCTIONAL GENOMICS. VOL II. NO 6. 420–433 doi:10.1093/bfgp/els030

Structural and dynamical analysis of biological networks

Cecilia Klein, Andrea Marino, Marie-France Sagot, Paulo Vieira Milreu and Matteo Brilli

Advance Access publication date 20 August 2012

Abstract

Biological networks are currently being studied with approaches derived from the mathematical and physical sciences. Their structural analysis enables to highlight nodes with special properties that have sometimes been correlated with the biological importance of a gene or a protein. However, biological networks are dynamic both on the evolutionary time-scale, and on the much shorter time-scale of physiological processes. There is therefore no unique network for a given cellular process, but potentially many realizations, each with different properties as a consequence of regulatory mechanisms. Such realizations provide snapshots of a same network in different conditions, enabling the study of condition-dependent structural properties. True dynamical analysis can be obtained through detailed mathematical modeling techniques that are not easily scalable to full network models.

Keywords: *networks; structural analysis; centrality; mathematical modeling; flux balance analysis*

INTRODUCTION

High-throughput technologies have recently led to a new perspective in biology, where the cell is interpreted as a large and complex system composed of highly integrated subsystems. Interpretation of these systems as networks of interactions has spurred the application of analytical tools developed since long by mathematicians and physicists to analyze biological networks.

Different biological networks can be defined; detailed descriptions in addition to the approaches to their reconstruction are treated exhaustively in several publications (Supplementary Material File 1). In this review, we focus on gene regulatory, metabolic and protein–protein interaction networks (PPINs), which are at the basis of all cellular processes, sparsely citing other kinds of networks when interesting for the discussion. A few technical definitions are provided in the Supplementary Material File 2 for the terms underlined in the text.

A PPIN (Figure 1A) has nodes corresponding to proteins and edges indicating their physical interaction. When a protein has more than one partner, the network is not able to tell if the different interactions take place together (as in a protein complex), or if they correspond to interactions taking place at different times.

An MN may be interpreted and built in various ways (Figure 1B): nodes can be metabolites or reactions (respectively giving rise to the compound and

Corresponding author. Matteo Brilli, INRIA, Université de Lyon, F-69000, Lyon; Université Lyon 1; CNRS, UMR5558, Laboratoire de Biométrie et Biologie Evolutive, F-69622, Villeurbanne cedex, France. E-mail: matteo.brilli@univ-lyon1.fr; matteo.brilli.bip@gmail.com

Cecilia Klein is a PhD student at the University of Lyon 1 and at the INRIA, co-advised by Ana Tereza Vasconcelos and Marie-France Sagot. She is interested in computational biology and is currently studying the evolution of metabolic dialogs among symbiotic partners.

Andrea Marino is a PhD student at the University of Florence (Italy), advised by Pierluigi Crescenzi; he is visiting the INRIA Bamboo Team (Lyon, France). He is interested in algorithms and complexity, bioinformatics and complex networks analysis in general.

Marie-France Sagot is Director of Research at the French National Institute of Research in Computer Science and Control (INRIA) and head of the INRIA-CNRS-UCBL BAOBAB-BAMBOO teams.

Paulo Vieira Milreu is developing his PhD work at the UCBL (INRIA Bamboo Team), under the co-supervision of Vincent Lacroix, Christian Gautier and Marie-France Sagot. His interests are currently focused on understanding metabolic mechanisms through mathematical modeling strategies and algorithm design.

Matteo Brilli (post-doc INRIA, equipe Bamboo) is a biologist converted to bioinformatics: he is interested in comparative genomics, mainly to study the evolution of metabolism and gene regulation, mathematical modeling of integrated systems and biological network analysis.

the reaction graphs), and arcs (i.e. directed edges) can be reactions or shared metabolites. In both cases, the reconstruction may lead to a loss of fundamental information (Figure 1B). These limitations ask for a full treatment of complex reactions in an MN (discussed in detail e.g. in [1,2]): bipartite graphs and hypergraphs help to overcome these problems at the price of a higher algorithmic complexity. Hypergraphs are indeed generalizations of graphs and thus problems may become harder to solve (see [3] for some examples of hypergraphs applied to biological questions and the associated computational problems).

In a gene regulatory network (GRN; Figure 1C), nodes representing transcriptional regulators are connected to the nodes corresponding to their targets by signed arcs. The sign or weight of such arcs indicates the effect of the control. Because of combinatorial regulation whose output depends on the architecture of promoters which is not encoded in a basic GRN, an hypergraph representation could also represent a better choice for these networks [4–6].

With a biological network in hand, we can inspect many properties of the nodes or the edges/arcs searching for interesting features. Network metrics were mainly developed for nonbiological purposes, but in some cases they provided meaningful biological information (see sections below and Supplementary Material File 1). A more thorough description of the use of network metrics in biology is given in the following sections. Different measures focus on distinct properties of nodes or edges/arcs; hence, the choice of a meaningful metric depends on the type of network and on the question(s) asked. This task requires some knowledge on the biological processes modeled by the network because they strongly affect the interpretation or even the usefulness of a measure.

MNs can also be studied using quantitative constraint-based models that are able to identify the optimal distribution of fluxes in the network in a defined growth condition, at the expense of neglecting the dynamics to reach steady state [7]. The accessible structure of the network can therefore be proficiently used to obtain quantitative and testable information on the physiological state of a bacterium.

Although informative, the analysis of a static structure has its drawbacks. The first one is that we completely neglect any additional property the nodes (genes and proteins) may have, asking for an integration of those features into meaningful network metrics inspired by biology. The second drawback concerns the highly dynamic nature of biological networks: regulatory mechanisms active in different physiological states change the connectivity of the network, so that structural properties may be condition dependent. Another problem arises because a structural analysis is not always able to take into account regulatory mechanisms: the activity of enzymes is often regulated by one or more effector metabolites but since the latter are not consumed, the MN neglects such regulations (Figure 1B). This can have profound consequences because these regulations have important roles in stabilizing the metabolic states and in generating complex and biologically important dynamic behaviors [8–10].

These effectors are moreover able to cross the boundaries between different biological levels, such as metabolism and gene regulation. Building integrated models taking these cross-talks into account therefore represents a major challenge in systems biology. Previous modeling efforts have demonstrated that none of the different biological layers is truly isolated [11–13] and that enzymes also have regulatory functions, exerted through their control over the concentration of particular metabolites.

These considerations lead to a view of the cell as a network of networks, whose understanding requires considering regulatory interactions not only within, but also between biological networks.

STRUCTURAL ANALYSIS

In this section, we explore some topological metrics often used to analyze biological networks. In particular, we focus on centrality measures to predict essential genes, average distance (AD) and diameter to inspect the compactness of the network, assortativity and dyadicity to study the modularity of a network and any correlations between the properties of the nodes.

Before discussing these measures, let us stress that biases in the network reconstructions or manipulation can strongly affect the results of the analysis, confounding (if any exist) the observed correlations of biological and topological properties [14]. Consequently, we need to carefully interpret the topological measures obtained given that we only have a partial reconstruction in hand, and that

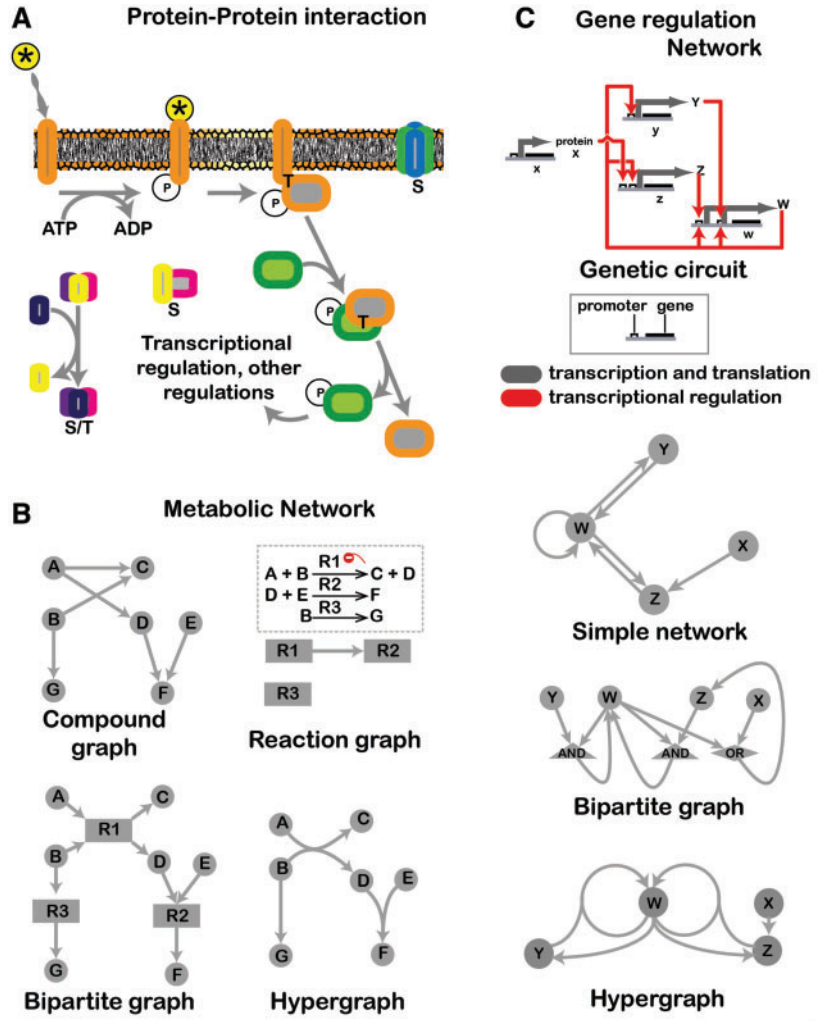

Figure 1: (**A**) An example of different kinds of interactions that build up a PPIN. A signal (asterisk) activates a receptor, which auto-phosphorylates and then passes the phosphate group to another protein (in Bacteria usually a Response Regulator), which is then able to regulate the activity of other proteins, or activate and repress gene expression. Interactions during this process are transient (T), therefore they are more difficult to detect using high-throughput technologies. Consequently, the PPIN is enriched in stable (S) interactions. (**B**) Graph models to represent an MN. Given three biochemical reactions (RI, R2, R3), metabolic graphs are built with metabolites as round nodes and reactions as square nodes. The enzyme catalyzing reaction RI has a metabolic regulatory feedback from compound C. The same system can be represented using different kinds of networks. Compound graph, where nodes are metabolites and there is an arc between a substrate and a product of a reaction; reaction graph, where nodes correspond to reactions and are connected when a product of one reaction is a substrate of the next one; bipartite graph: nodes are either compounds or reactions in which there is an arc between the substrate/reaction and reaction/product; hypergraph: nodes are compounds and a hyperarc links the substrate(s) to the product(s) of a reaction. The feedback from C to the enzyme catalyzing reaction RI is lost in all of these representations. Also, the compound and reaction graphs account for loss of information, e.g. reaction RI has two substrates (A and B) and two products (C and D), however, by looking at the corresponding compound graph one could imagine that the production of C only requires A, and by looking at the corresponding reaction graph we notice that the arc between RI and R2 exists only because of the compound D regardless of the presence of E. (**C**) A genetic circuit is a visual representation of a biological system and we provide three of its possible mathematical translations. The bipartite graph has nodes for proteins (circles) and different logical gates for combinatorial regulation: AND (triangle) requires the presence of both regulators to have transcription, while OR (diamond) can be activated by one of the regulators alone. The information on the promoter logics is lost in the Simple representation, while it is encoded in the hypergraph. The difference between these representations is evident if we suppose to remove regulator Z. By analyzing the Simple network, one may infer that the autoregulation of W continues to take place, which is not true, as correctly predicted by the bipartite graph and the hypergraph.

some of the measures described below are strongly affected by the sampling [15,16].

Centrality analysis

Given a network, it is natural to wonder how important each node is to its functionality. A number of graph measures have been developed for evaluating node centrality [17–21] and several tools allow to compute diverse network metrics, like CentiBiN [17], VisANT [22], Visone [23], Pajek [23], CentiScaPe [21] and CentiLib [24].

Centrality measures can be local (or neighborhood based) or global (distance or feedback based).

Local measures

With neighborhood-based measures, such as degree, the importance of the nodes is inferred from their local connectivity: the more connections a node has, the more central it is. Highly connected nodes (hubs) were found to possess special properties in the yeast PPIN: they are more often essential than non-hub proteins [25,26]), they tend to play a central role in the modular organization of a PPIN [27,28] and they seem to be evolutionarily more conserved [29]. Nevertheless, since then, several works have raised doubts on some of these associations [30,31].

There is no consensus in the literature on how to define a hub, and different criteria have been used: a given fraction of the highest degree nodes [32]; nodes with a given fraction of the total connectivity [33]; and a degree greater than an arbitrary threshold [28,34,35]. Recently, Vallabhajosyula *et al*. [36] proposed three objective functions allowing to define hubs in a PPIN in a rigorous way; unfortunately these are based on previous results on the properties of hubs in PPINs, limiting their applicability to other types of networks.

In order to have an indication about the homogeneity of the nodes of a network, it could be interesting to study the degree distribution that for most biological networks is well fitted by a power-law ($P(k) \sim k^{-\gamma}$) with $\gamma \sim 2$, where k is the degree. In these networks, a few hubs play a fundamental role for the integrity and navigability of the network [27], whereas a vast majority of the nodes has only a few connections. This degree distribution has been associated with robustness against random node removal. Robustness to the loss of a node in an MN indicates the presence of alternative pathways bypassing the missing reaction; in GRNs it may correspond to the presence of alternative ways of transmitting and controlling information. On the contrary, these networks are highly sensitive to attacks directed on hubs, because their removal deeply affects network functionality [37]. Even though much research has been done on the power-law distribution and its universality in biological networks, criticisms have been raised [38]. Power-law degree distributions indeed can be obtained through random sampling of networks with different topologies, indicating that it might not be possible to infer the true degree distribution from biological networks, for which complete reconstructions are usually not available [39].

The local connectivity of nodes can be studied in further detail by using either assortativity or dyadicity. The first measure represents the correlation between the degree of adjacent nodes [40]. Maslov and Sneppen [41] found that hubs in the yeast PPIN are mostly connected to non-hubs, and are therefore well separated from each other. Dyadicity [42] measures the degree to which the nodes of a network are connected to other nodes that share some characteristic (functional classification, essentiality, involvement in a disease and so on) and is therefore able to characterize the modular structure of a network by considering the distribution of the functions over the nodes and their connectivity [43]. A network is called heterophilic (heterophobic) when different categories are connected more (less) often than expected under a random model. It was recently used to study the potential coupling between structure and functionality in transcriptional and noncoding (nc) RNA–protein interaction networks [44]. The results showed that most transcriptional regulators and ncRNAs tend to connect to genes/proteins of other functional classes, suggesting that regulators do not really belong to a functional class but tend instead to coordinate several of them [44]. On the converse, in PPINs and MNs, the connections more often involve proteins of a same functional category.

Global measures

Closeness [45] and shortest path-based betweenness [46] reflect global properties of a network and use a distance measure between nodes, often the shortest path. The closeness of a node depends on its AD from the others and is of particular interest for information networks (such as signaling network and GRNs) as it measures how fast information flows from a node of interest to all the reachable nodes

[47]. It has been recently integrated with biological information in a parameter-free gene prioritization approach that computes the interconnectedness (ICN) between genes in a network [48]. ICN measures closeness of each candidate gene to genes possessing an interesting property by considering alternative paths in addition to the direct link and shortest one.

Shortest path-based betweenness depends on the number of shortest paths crossing a node. In PPINs, betweenness can be interpreted as the relevance of a protein to be intermediary in the interaction between other proteins, assuming that this interaction passes through shortest paths [21]. Bottlenecks are nodes with high betweenness centrality and were found to be key connectors with surprising functional and dynamical properties, often essential [49]. Bottleneck and hub genes were identified in coexpression networks inferred from experimental data, and found to be often essential for virulence in *Salmonella typhimurium* with the role of mediators of transitions between different cellular states or of sentinels that reflect the dynamics of these transitions [50]. Cell cycle checkpoints were found to be bottlenecks in a gene coexpression network of cell cycle regulated genes in the fission yeast [51].

Network metrics in general [52–54], and betweenness centrality in particular are also used for the rational prediction of drug targets [55]. Essential genes are preferred targets for drug design and central genes are more likely to be essential. Another constraint was imposed in this particular case: the gene must be essential for the pathogen but not for the host to reduce any side effects of the drug.

One problem with shortest path-based measures is that communication between biological entities is assumed to pass along those paths, which is often not plausible: from the point of view of MNs, the shortest path might be defined on the basis of the energy/cofactor requirements instead of the number of steps, whereas in GRNs and PPINs all active connections will take place, not only the shortest ones. In the case of GRNs, the targets with different shortest paths to a common regulator may exhibit hierarchical gene expression patterns as is the case for flagellar genes [56].

To overcome the limitation of shortest paths, a node can be considered central when it is crossed by many random walks: this is the case of the random walk-based betweenness centrality [57]. Some feedback-based measures implicitly rely on random walks, like eigenvector [58] and spectral centrality [59]. Eigenvector centrality has been applied to several MNs [60] and was shown to outperform other metrics for the identification of essential proteins in the PPIN of yeast [61], together with subgraph centrality [62].

Distance analysis

The diameter of a network is an overall indication of its compactness. Despite the fact that real networks sometimes exhibit the small-world property and that shorter diameters may be beneficial to some networks (e.g. for rapid information flow), it was shown that several biological networks have larger diameters than their randomizations. One possible reason for this is their modular nature [63] leading to the suggestion that modularity may be a universal characteristic of real networks, due to the advantages it brings to multi-functionality, robustness and evolvability. On one hand, high modularity reduces pleiotropic effects improving the evolvability of the system. On the other, numerical experiments also demonstrated that modularization provides robustness against random perturbations in network structure, i.e. evolutionary change [64].

The distribution of distances and the AD may be more informative than the diameter about the global properties of a network [63]. The small AD commonly observed in biological networks pertains to the so-called small-world effect [65]. The AD ranged between 3 and 5 in 43 MNs of 200–800 nodes [66], showing that all nodes are quite close to each other. Although several groups confirmed the small-world property of the MN of different organisms [67–71], Arita [72] heavily criticized the way the pathways are computed in those works since they do not conserve their structural moieties. When this problem is accounted for correctly, the analysis revealed that the average path length of the *Escherichia coli* metabolism is much longer than previously thought [72,73].

Quantitative structural analysis

Flux Balance Analysis (FBA; Figure 2) is a quantitative modeling technique that relies on a validated reconstruction of an MN, the steady-state assumption and additional constraints [74–76].

The target of the method is obtaining the flux distribution within the MN under specified growth conditions (Figure 2).

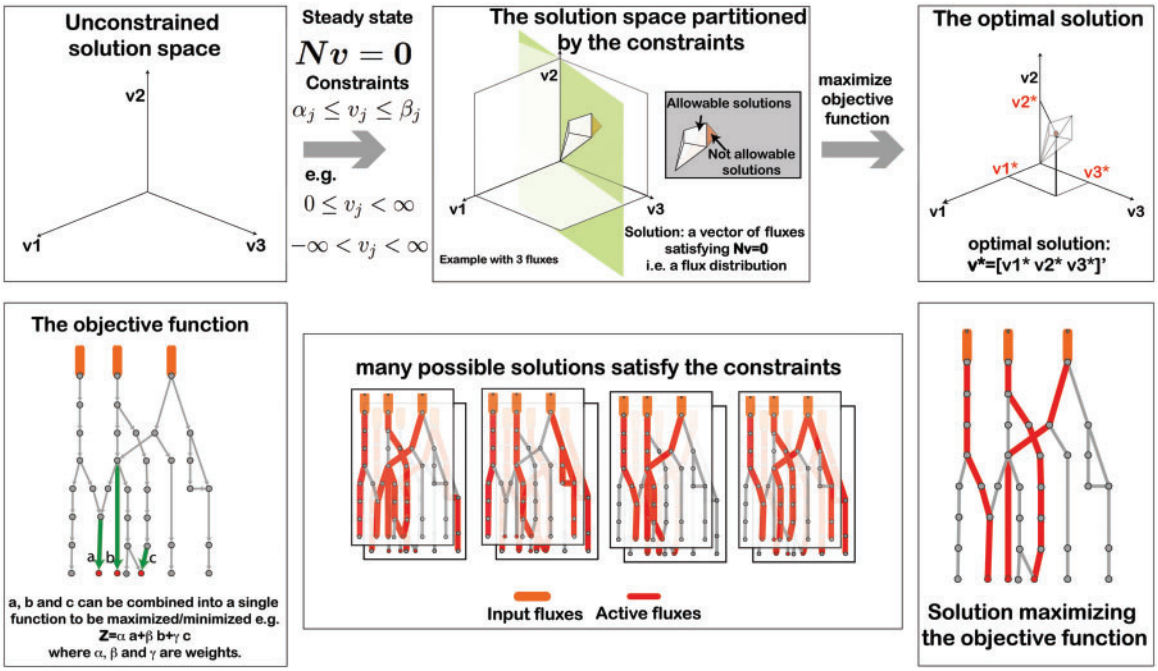

Figure 2: FBA is a constraint-based model based on the stoichiometric modeling of an MN, a (quasi) steady-state condition and an objective function. The constraints are the reaction set of the network encoded in the stoichiometric matrix N and additional thermodynamic and environmental constraints. The steady-state condition for MNs corresponds to a regime where the intracellular fluxes and metabolite concentrations are constant in time ($Nv = 0$), where v is a vector representing a flux distribution for the reactions. There are many flux distributions satisfying the steady-state condition and the other constraints. In FBA experiments, the interest is the identification of the flux distribution that maximizes/minimizes a given objective function.

The stoichiometry of the reactions encode the mass conservation rules, and a modeling of the environment through transport reactions impose constraints on the possible flux distributions satisfying the steady-state condition; additional constraints may also be added such as reaction reversibility and maximum velocity of enzymes. Since the solution space for such models is very large even under the constraints used, FBA seeks an optimal flux distribution with respect to a carefully chosen objective function using optimization techniques. The assumption behind FBA is that metabolism maximizes some objective, but there may exist many suboptimal flux distributions that help the organism during adaptation to specific environmental conditions. This led to elementary mode analysis [77], which seeks for the solutions satisfying the above constraints regardless of the objective function. Elementary modes can be loosely defined as the smallest subnetworks allowing an MN to function in steady state [78,79]. According to Stelling *et al.* [79], they can be used to understand cellular objectives for an overall MN.

The objective function plays a fundamental role in FBA as it provides a way to choose one optimal solution: assuming that the objective of *E. coli* in rich medium is to grow at maximum speed, we may formulate an objective function that combines fluxes exiting the MN to produce biomass. Optimization through integer linear programming [7,80] then allows to identify one optimal solution which is a physiological steady state of the MN of an organism in that condition. When the target is maximization of the production of some compound, the compound is usually included in the objective function to enforce solutions where its production is active. Other formulations for the objective function may be designed to mimic disparate growth conditions, not necessarily focusing on fast growth [81–91].

Biologically speaking, solutions obtained through FBA describe a partition of the input fluxes into the different branches of the network to produce the compounds required by growth (through the objective function).

One of the most appealing properties of constraint-based models is that they provide a way to explore the consequences of genetic manipulations on the whole MN: one or more reactions can be eliminated (simulating knock-out mutants) [92–95]

or otherwise manipulated, and simulations can be run to see if and how the objective function can be improved with respect to the wild-type model [96]. By coupling two levels of optimization, it is possible to predict the best engineering strategy to have mutants that maximize some by-product of interest, such as ethanol [96] or lactate [97], while growing. A recent survey on FBA and its applications can be found in [98].

Dynamic analysis

Dynamic analysis of structural properties

In general, we look at biological networks as static entities, but it should be stressed that they are instead very dynamic at widely different time-scales. They are dynamic in evolutionary time like any other biological structure, and even more on short time-scales, since regulatory connections and feedbacks change the connectivity of the network depending on the physiological state (Figure 3). Consequently, we should interpret most of the currently available biological network reconstructions as potential networks, where all the possible connections are indicated. By the term potential, we highlight the fact that edges/arcs and nodes in this network will be hardly present all together *in vivo*. If we consider for instance a PPIN, not all interaction partners of a protein will be expressed in a given condition, reducing the number of actual partners. Conversely, we may speak of network realizations when focusing on the active subgraph of a potential network, defined on the basis of experimental data [28,99–101]. The dynamic nature of biological networks is also at the basis of differential network analysis [102], which aims at capturing the subgraphs specific of a given network realization.

These considerations are important since they affect the analysis of biological networks. As there are many condition-specific realizations of a biological network, they plausibly have different structural properties. It was indeed shown that random subgraphs of a network do not necessarily maintain the same-degree distribution as the entire network [103], suggesting that other structural properties may also change (Figure 4).

Therefore, it is not clear if we can look for 'universal' properties of biological networks by analyzing potential networks, or whether we should instead define as 'universal' those properties that characterize most realizations.

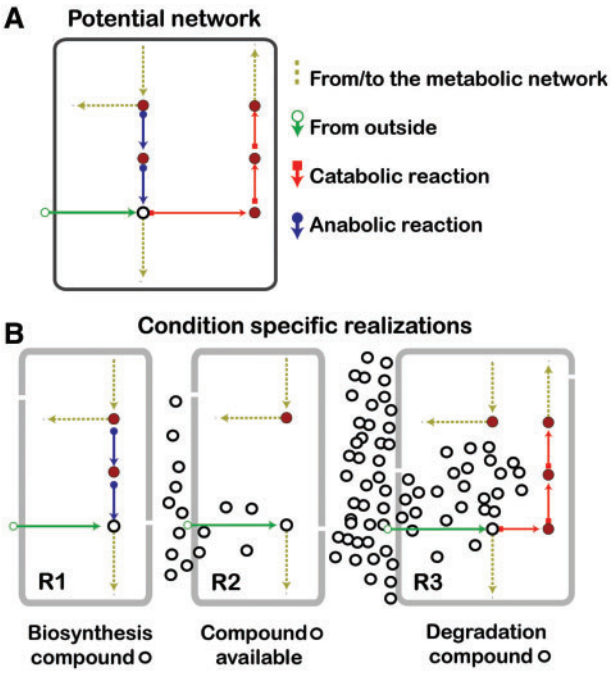

Figure 3: Illustrative example on the potential and realization concept concerning the anabolic and catabolic pathways of a same compound (4). (**A**) The potential network. (**B**) The realizations are shown for different physiological states: RI, biosynthetic state for compound o. R2 compound o is available and its biosynthetic route is off. R3 catabolic state: a degradation pathway is activated to reduce the intracellular concentration of the compound.

Han *et al.* [28] estimated the temporal connectivity of hubs in the yeast PPIN by using gene expression data: the correlation in gene expression between two connected nodes in the potential network allowed to define two types of hubs: party hubs, interacting with their partners simultaneously; and date hubs, which bind their different partners at different times or locations. It is then plausible to do the same for other measures: genes may be central in the potential network and frequently or not in the realizations (party and date centers); party and date bottlenecks may be defined in the same way, and so on. This additional level of complexity may allow a deeper understanding of how physiological transitions are driven by topological changes.

Gene expression was integrated in a centrality measure called Pec [104], which was used to identify essential genes in yeast. This measure exploits the strength of the connectivity between two adjacent nodes based on an Edge Clustering Coefficient [105], weighted by the co-expression between genes in experimental data.

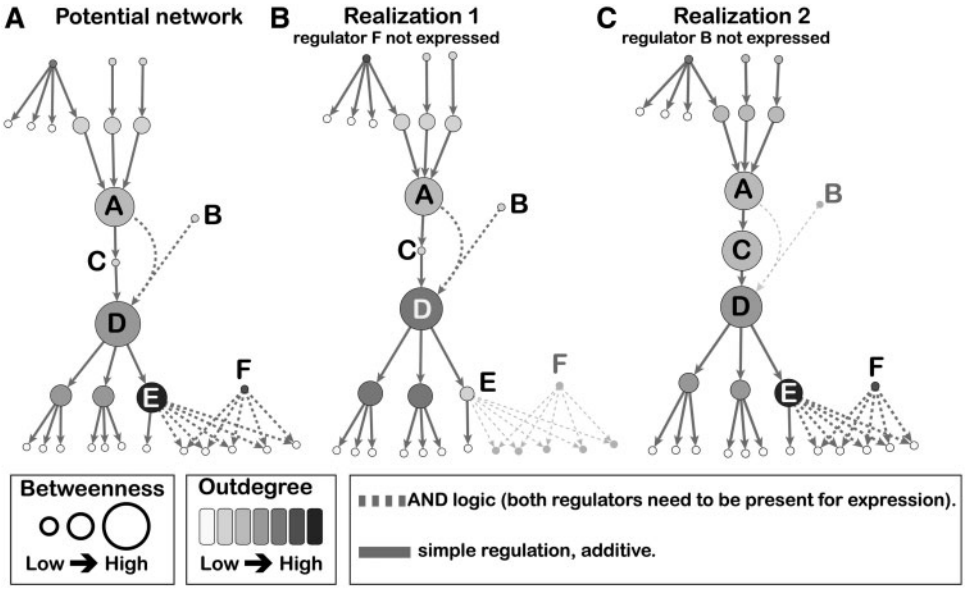

Figure 4: Centrality measures change in GRN realizations. Nodes have a size proportional to the betweenness centrality measure and the color of a node changes according to the outdegree. The pairs of regulators A and B as well as E and F are both required for the activation of the target gene(s). (**A**) The potential network, where regulators A and D are central following betweenness centrality, and E with respect to outdegree centrality. Now let us suppose to use experimental data to obtain two realizations of this potential network. In (**B**) regulator F is not expressed, and regulator E has consequently a low outdegree. In (**C**) regulator B is inactive, imposing a remarkable change in the betweenness centrality value of regulator C.

This reasoning also affects the evolutionary interpretation of network properties, for instance when concluding that evolution promoted the fixation of a given structural feature of the potential network. Luscombe *et al.* [99] analyzed the structural properties of the yeast GRN in different conditions. Starting from a validated GRN, they used gene expression data to extract the subnetworks supposed to be active during environmental stress or the cell cycle, highlighting important differences: the cell cycle subnetwork has long shortest paths and combinatorial regulation is common, whereas short paths and mainly single-input regulations characterize the stress condition. The length of a path may be relevant in the context of a GRN because it can be interpreted as a measure of the delay to have a response once the top regulator is activated (Figure 1B). The short paths for the stress conditions suggest evolution of a fast response to stressors, whereas cell cycle evolved under the necessity for fine regulations giving the correct temporal ordering of events, which explains the combinatorial regulation (information integration) and the longer paths (check points). Performing the analysis on the potential network, these differences would not have been noticed. The previous work has however been heavily criticized [99], but both studies conclude that realization networks can be largely different in their structural properties (see also [28,101]).

The use of realization networks is currently limited by the need for high-quality and high-throughput experimental data, today available only for a few organisms. Nevertheless, large-scale experimental data will be more easily obtained in the future, giving the occasion to develop the algorithms required for a similar approach.

Kinetic modeling of full-scale networks

In the previous section, we discussed how to explore the structural properties of a biological network using experimental data to define the active subgraphs in a potential network. However, the analysis is not really dynamic, but gives instead only a snapshot of the steady states of a network in different conditions. To move forward with the dynamic analysis of networks, we discuss the mathematical modeling of biochemical reaction networks from the perspective of building large, network-scale models able to predict the dynamics between different states. Many different modeling strategies were devised and

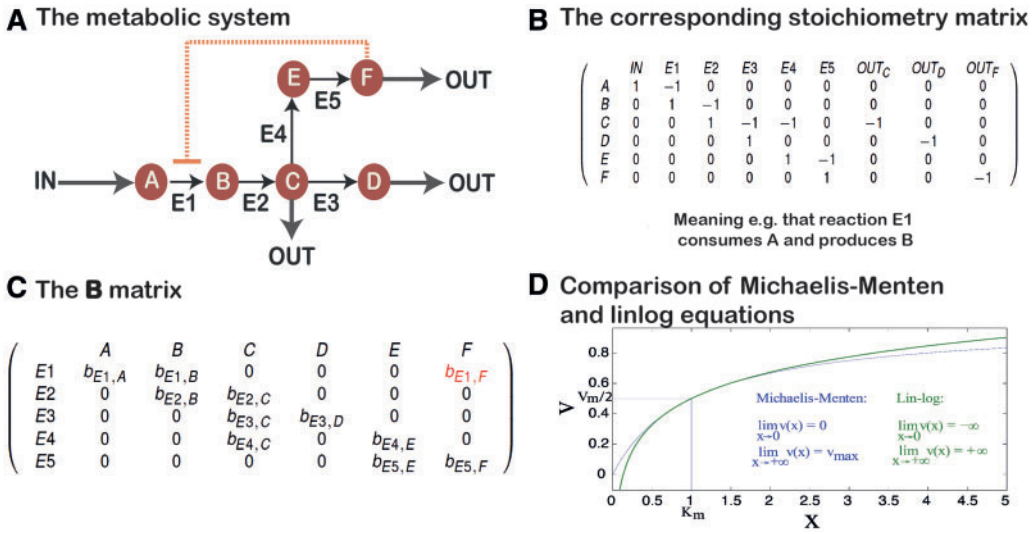

Figure 5: (**A**) A metabolic system. (**B**). The corresponding stoichiometry matrix N. The evolution in time of the six metabolite concentrations is given by: $dx/dt = Nv(e,f(x,p))$, where x is the vector of metabolite concentrations and $v(e,f(x,p))$ is a vector of rates, functions depending on enzyme levels e and on metabolites in x, including the effectors. The latter dependencies are not encoded in the stoichiometry matrix. $f(x,p)$ can take many different forms, e.g. mass action, Michaelis–Menten or linlog. (**C**) The parameter matrix of the linlog approximation of the entire system; all the rate functions have the same standard format, a linear combination of logarithmic metabolite concentrations i.e. $v = \text{diag}(e)\ (A + B \log X)$, with A and B a vector and a matrix of parameters, respectively. (**D**) Comparison of the irreversible Michaelis–Menten (V_{max} $[S]/(K_m+[S])$) and corresponding linlog: linlog is not saturable for large substrate concentrations, and gives minus infinite fluxes when one of the metabolites in a given reaction goes to zero.

described elsewhere [4,8,106–119]; here we briefly discuss the modeling of biochemical networks (MN and GRN) and its application to cellular scale systems. Some of the discussions also apply to signaling systems, which combine different types of regulation (protein–protein interaction, phosphorylation and transcriptional regulation).

Kinetic metabolic models are traditionally based on systems of ordinary differential equations where the rates modeling the activity of an enzyme are mechanistic, nonlinear and more or less precisely describe the catalytic mechanism of an enzyme. The activity of promoters in gene regulation is usually modeled using sigmoid functions as suggested by experimental data [120,121], and combination thereof in the presence of combinatorial regulation [4]. The parameters of these models are usually derived from *in vitro* (rarely *in vivo*) experiments but the large differences between *in vivo* and *in vitro* conditions have called into question this approach [122–125], and *in vivo* experiments should be preferred [126]. The main drawback of building such detailed models is therefore that it is very time-consuming for the amount of good quality and informative experimental data required to perform parameter identification. Mechanistic models have been consequently applied mainly to well-studied systems, and only recently models for less studied ones have started being implemented [127–131].

All these limitations make it impossible at the moment to build mechanistic models at a full network scale. The only exception for MNs is a work by Jamshidi and Palsson [132], who use mass action kinetics to build a model of the MN of red blood cells with 100 chemical reactions (catalytic or regulatory), and 95 variables. To overcome the limits imposed by mechanistic models, approximative nonmechanistic rate equations have been developed for both metabolic (e.g. [113–115]) and gene regulation systems [4]. The main advantage of approximated formalisms is that they require less parameters, reducing as well the experimental effort for parameter identification. One of these approximations is called linlog, and was recently used to model a network-scale MN of yeast [133]. The parameters were obtained from a model repository (see Figure 5 for more details on this

approximation). The resulting model contains 956 metabolic reactions and 820 metabolites; the key steps were identified using metabolic control analysis. This modeling framework may be considered a stepping-stone towards the long-term goal of a fully parameterized model of genome-scale metabolism even if its performance needs to be improved.

GRNs also cannot be modeled at a full scale, since much of the information required is not available, and approximated formalisms were proposed [4]. We stress that obtaining a GRN is much more difficult than obtaining an MN; the methods give moreover very partial reconstructions that strongly affect the structural analysis [16].

Modeling network scale integrated systems

An important and ambitious challenge in systems biology is building integrated models where the interactions between different biological layers are explicitly taken into account. We here consider the case of integrated models where metabolism is modeled together with the gene regulation system, but it should be noticed that increasing experimental evidence suggests further integration of signaling pathways and GRNs with regulation mediated by ncRNAs [134–138]. On one hand, integration of metabolism and gene regulation might allow to study a much wider range of situations using a same model, and on the other, it allows to study more in detail the importance of the cross-talk between the two systems. A first effort to measure the effect of regulation in FBA predictions through the addition of Boolean logic time-dependent constraints modeling transcriptional regulatory events is regulatory FBA (rFBA; [139]). rFBA changes the shape of the solution space considerably with respect to FBA, finding physiologically relevant solutions [139]. These initial methods were improved by several recent works such as steady-state regulatory FBA (SR-FBA), which is an integrated regulatory-metabolic model for predicting gene expression and metabolic fluxes [140], integrated FBA (iFBA) that combines rFBA and inferred ordinary differential equations [141], OptFlux which is a software for strain prediction through metabolic/regulatory integrated data [142], and hybrid modeling [143]. For a more detailed review on different coupled regulatory/metabolic models, we refer to [144].

CONCLUSIONS

Structural analysis allows the identification of important nodes within a network and for this reason, has become very popular in many disciplines. However, in the biological domain, the importance of a node can be defined in many different ways so that identifying the most appropriate network measures is an important preliminary step that can radically change the output of an analysis. It is then essential to understand the meaning of a given measure with respect to the specific network at hand.

Besides discussing some of the most informative metrics for biological networks analysis, we stress the importance of a biologically meaningful interpretation of any measure, which is not always intuitive and can change for different networks.

The dynamical nature of biological networks indicates that it may be better to perform structural analysis on what we have defined as the realizations of a network. The risk when studying a potential network is confounding the signals encoded in the network by putting everything together. Are we sure that a metabolic hub is a hub in every realization of the network? What if it is lowly connected with different nodes in every realization? This approach is today limited by the availability of experimental data, but databases are growing fast and a similar analysis would be feasible for several prokaryotes, as well as for a few eukaryotes.

Concerning the more biologically oriented interpretation of the metrics, it requires to move the collaboration between computational and experimental biologists to a higher level. It would also contribute to the integration of biological information in network analysis, which is a topical challenge in the field. Let us take the example of hubs in a GRN. From the biological point of view, it is clearly different if the hub controls a single cellular function or affects widely different processes. Since a GRN transmits information, a similar approach would require being able to define the scope of a regulator by also taking into account indirect targets (similarly to [6]). This example illustrates the need for biologically oriented network metrics that are able to take into account the heterogeneous information associated with biological entities. As pointed out by Keller [145], Watts and Strogatz (65) have proficiently used simple mathematical models to study social networks, but some of their most interesting results emerged only after they took into account the

property that sociologists consider as fundamental to social dynamics: social identity. The challenge is to do the same with biological networks, which requires an effort to develop meaningful metrics able to account for and integrate biological properties.

SUPPLEMENTARY DATA

Supplementary data are available online at http://bfg.oxfordjournals.org/.

Key Points

- Structural analysis of biological networks allows to identify genes and proteins playing important roles in cellular physiology.
- Biological networks are dynamic; the structural properties of genes and proteins are consequently also dynamic, i.e. the importance of a protein might change depending on the growth condition.
- The dynamics of biological systems can be studied using detailed mathematical modeling, but they are not easily scalable at the network level and approximations have been provided that might simplify the task.

FUNDING

This work was funded by the French project ANR MIRI BLAN08-1335497, and the ERC Advanced Grant Sisyphe held by M.-F.S. It was supported by the INRIA International Partnership AMICI.

References

1. Lacroix V, Cottret L, Thébault P, *et al*. An introduction to metabolic networks and their structural analysis. *IEEE/ACM Transact Comput Biol Bioinform* 2008;**5**:594–617.
2. Cottret L, Jourdan F. Graph methods for the investigation of metabolic networks in parasitology. *Parasitology* 2010;**137**:1393–1407.
3. Klamt S, Haus U-U, Theis F. Hypergraphs and cellular networks. *PLoS Comput Biol* 2009;**5**:e1000385.
4. de Jong H. Modeling and simulation of genetic regulatory systems: a literature review. *J Comput Biol* 2002;**9**:67–103.
5. Klamt S, Saez-Rodriguez J, Lindquist JA, *et al*. A methodology for the structural and functional analysis of signaling and regulatory networks. *BMC Bioinformatics* 2006;**7**:56.
6. Wang R, Albert R. Elementary signaling modes predict the essentiality of signal transduction network components. *BMC Syst Biol* 2011;**5**:44.
7. Edwards JS, Ibarra RU, Palsson BØ. In silico predictions of Escherichia coli metabolic capabilities are consistent with experimental data. *Nat Biotechnol* 2001;**19**:125–30.
8. Steuer R, Gross T, Selbig J, *et al*. Structural kinetic modeling of metabolic networks. *Proc Natl Acad Sci USA* 2006;**103**:11868–73.
9. Grimbs S, Selbig J, Bulik S, *et al*. The stability and robustness of metabolic states: identifying stabilizing sites in metabolic networks. *Mol Syst Biol* 2007;**3**:146.
10. Steuer R. Computational approaches to the topology, stability and dynamics of metabolic networks. *Phytochemistry* 2007;**68**:2139–51.
11. Baldazzi V, Ropers D, Markowicz Y, *et al*. The carbon assimilation network in Escherichia coli is densely connected and largely sign-determined by directions of metabolic fluxes. *PLoS Comput Biol* 2010;**6**:e1000812.
12. Baldazzi V, Ropers D, Geiselmann J, *et al*. Importance of metabolic coupling for the dynamics of gene expression following a diauxic shift in Escherichia coli. *J Theor Biol* 2011;**295**:100–15.
13. Kotte O, Zaugg JB, Heinemann M. Bacterial adaptation through distributed sensing of metabolic fluxes. *Mol Syst Biol* 2010;**6**:355.
14. Coulomb S, Bauer M, Bernard D, *et al*. Gene essentiality and the topology of protein interaction networks. *Proc Biol Sci Roy Soc* 2005;**272**:1721–5.
15. Costenbader E, Valente TW. The stability of centrality measures when networks are sampled. *Soc Network* 2003;**25**:283–307.
16. de Silva E, Thorne T, Ingram P, *et al*. The effects of incomplete protein interaction data on structural and evolutionary inferences. *BMC Biol* 2006;**4**:39.
17. Junker BH, Koschützki D, Schreiber F. Exploration of biological network centralities with CentiBiN. *BMC Bioinformatics* 2006;**7**:219.
18. Koschützki D, Schreiber F. Centrality analysis methods for biological networks and their application to gene regulatory networks. *Gene Regulat Syst Biol* 2008;**2**:193–201.
19. Pavlopoulos GA, Secrier M, Moschopoulos CN, *et al*. Using graph theory to analyze biological networks. *BioData Mining* 2011;**4**:10.
20. Mason O, Verwoerd M. Graph theory and networks in Biology. *IET Syst Biol* 2007;**1**:89.
21. Scardoni G, Laudanna C. Centralities based analysis of complex networks. In: Zhang Y (ed). *New Frontiers in Graph Theory*. InTech Open, 2012.
22. Hu Z, Hung J-H, Wang Y, *et al*. VisANT 3.5: multi-scale network visualization, analysis and inference based on the gene ontology. *Nucl Acids Res* 2009;**37**:W115–21.
23. Baur M, Benkert M, Brandes U, *et al*. Visone - software for visual social network analysis. *Proceeings of the 9th Internatinal Symposium on Graph Drawing (GD '01)*, 2001.
24. Grassler J, Koschutzki D, Schreiber F. CentiLib: comprehensive analysis and exploration of network centralities. *Bioinformatics* 2012;**28**:1178–9.
25. Jeong H, Mason SP, Barabási A-L, *et al*. Lethality and centrality in protein networks. *Nature* 2001;**411**:41–2.
26. He X, Zhang J. Why do hubs tend to be essential in protein networks? *PLoS Genet* 2006;**2**:e88.
27. Albert R, Jeong H, Barabasi A. Error and attack tolerance of complex networks. *Nature* 2000;**406**:378–82.
28. Han JDJ, Bertin N, Hao T, *et al*. Evidence for dynamically organized modularity in the yeast protein–protein interaction network. *Nature* 2004;**430**:88–93.
29. Wuchty S, Almaas E. Peeling the yeast protein network. *Proteomics* 2005;**5**:444–9.

30. Wuchty S. Interaction and domain networks of yeast. *Proteomics* 2002;**2**:1715–23.
31. Zotenko E, Mestre J, O'Leary DP, *et al*. 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.
32. Batada NN, Reguly T, Breitkreutz A, *et al*. Stratus not altocumulus: a new view of the yeast protein interaction network. *PLoS Biol* 2006;**4**:e317.
33. Reguly T, Breitkreutz A, Boucher L, *et al*. Comprehensive curation and analysis of global interaction networks in Saccharomyces cerevisiae. *J Biol* 2006;**5.**
34. Ekman D, Light S, Björklund AK, *et al*. What properties characterize the hub proteins of the protein-protein interaction network of Saccharomyces cerevisiae? *Genome Biol* 2006;**7**:R45.
35. Aragues R, Sali A, Bonet J, *et al*. Characterization of protein hubs by inferring interacting motifs from protein interactions. *PLoS Comput Biol* 2007;**3**:1761–71.
36. Vallabhajosyula RR, Chakravarti D, Lutfeali S, *et al*. Identifying hubs in protein interaction networks. *PloS One* 2009;**4**:e5344.
37. Barabási A-lászló, Oltvai ZN. Network biology: understanding the cell's functional organization. *Nat Rev Genet* 2004;**5**:101–13.
38. Lima-Mendez G, van Helden J. The powerful law of the power law and other myths in network biology. *Mol Bio Sys* 2009;**5**:1482–93.
39. Han J-DJ, Dupuy D, Bertin N, *et al*. Effect of sampling on topology predictions of protein-protein interaction networks. *Nat Biotechnol* 2005;**23**:839–44.
40. Newman MJ. Assortative mixing in networks. *Phys Rev Lett* 2002;**89**:1–4.
41. Maslov S, Sneppen K. Specificity and stability in topology of protein networks. *Science* 2002;**296**:910–3.
42. Park J, Barabási A-L. Distribution of node characteristics in complex networks. *Proc Natl Acad Sci USA* 2007;**104**: 17916–20.
43. Jiang X, Liu B, Jiang J, *et al*. Modularity in the genetic disease-phenotype network. *FEBS Lett* 2008;**582**: 2549–54.
44. Nacher J, Araki N. On the relation between structure and biological function in transcriptional networks and ncRNA-mediated interactions. *Intl Conf Biosci Biochem Bioinform* 2011;**5**:348–352.
45. Latora V, Marchiori M. A measure of centrality based on network efficiency. *New J Phys* 2007;**9**:188.
46. Freeman L. A set of measures of centrality based on betweenness. *Sociometry* 1977;**40**:35–40.
47. Li X-li, Ng S-K. Biological data mining in protein interaction networks. Hershey: IGI Global, 2009.
48. Hsu C-L, Huang Y-H, Hsu C-T, *et al*. Prioritizing disease candidate genes by a gene interconnectedness-based approach. *BMC Genomics* 2011;**12**(Suppl. 3):S25.
49. Yu H, Kim PM, Sprecher E, *et al*. The importance of bottlenecks in protein networks: correlation with gene essentiality and expression dynamics. *PLoS Comput Biol* 2007;**3**:e59.
50. McDermott JE, Taylor RC, Yoon H, *et al*. Bottlenecks and hubs in inferred networks are important for virulence in Salmonella typhimurium. *J Comput Biol* 2009;**16**:169–80.
51. Caretta-Cartozo C, De Los Rios P, Piazza F, *et al*. Bottleneck genes and community structure in the cell cycle network of S. pombe. *PLoS Comput Biol* 2007;**3**:e103.
52. Vallabhajosyula RR, Raval A. Computational modeling in systems biology. *Syst Biol Drug Discov Dev* 2010;**662.**
53. Chavali AK, Blazier AS, Tlaxca JL, *et al*. Metabolic network analysis predicts efficacy of FDA-approved drugs targeting the causative agent of a neglected tropical disease. *BMC Syst Biol* 2012;**6**:27.
54. Chen L-C, Yeh H-Y, Yeh C-Y, *et al*. Identifying co-targets to fight drug resistance based on a random walk model. *BMC Syst Biol* 2012;**6**:5.
55. Rahman SA, Schomburg D. Observing local and global properties of metabolic pathways: "load points" and "choke points" in the metabolic networks. *Bioinformatics* 2006;**22**:1767–74.
56. Smith TG, Hoover TR. Deciphering bacterial flagellar gene regulatory networks in the genomic era. *Adv Appl Microbiol* 2009;**67**:257–95.
57. Newman MJ. A measure of betweenness centrality based on random walks. *Soc Network* 2005;**27**:39–54.
58. Bonacich P. Some unique properties of eigenvector centrality. *Soc Net* 2007;**29**:555–64.
59. Perra N, Fortunato S. Spectral centrality measures in complex networks. *Phys Rev E* 2008;**78**:1–10.
60. Ding D-wu, He X-qing, Science C. Application of eigenvector centrality in metabolic networks. In *2010 2nd International Conference on Computer Engineering and Technology*, 2010. 89–91.
61. Estrada E. Virtual identification of essential proteins within the protein interaction network of yeast. *Proteomics* 2006;**6**: 35–40.
62. Estrada E, Rodríguez-Velázquez JA. Subgraph centrality in complex networks. *Phys Rev E* 2005;**71**:056103.
63. Zhang Z, Zhang J. A big world inside small-world networks. *PloS One* 2009;**4**:e5686.
64. Kashtan N, Alon U. Spontaneous evolution of modularity and network motifs. *Proc Natl Acad Sci USA* 2005;**102**: 13773–8.
65. Watts DJ, Strogatz SH. Collective dynamics of "small-world" networks. *Nature* 1998;**393**:440–2.
66. Jeong H, Tombor B, Albert R, *et al*. The large-scale organization of metabolic networks. *Nature* 2000;**407**:651–4.
67. Fell DA, Wagner A. The small world of metabolism. *Nat Biotechnol* 2000;**18**:1121–2.
68. Wagner A, Fell DA. The small world inside large metabolic networks. *Proc Biol Sci Roy Soc* 2001;**268**:1803–10.
69. Ravasz E, Somera AL, Mongru DA, *et al*. Hierarchical organization of modularity in metabolic networks. *Science* 2002;**297**:1551–5.
70. Ma H, Zeng AP. Reconstruction of metabolic networks from genome data and analysis of their global structure for various organisms. *Bioinformatics* 2003;**19**:270.
71. Yook S-H, Oltvai ZN, Barabási A-L. Functional and topological characterization of protein interaction networks. *Proteomics* 2004;**4**:928–42.
72. Arita M. The metabolic world of Escherichia coli is not small. *Proc Natl Acad Sci USA* 2004;**101**:1543–7.
73. Pitakänen E, Rantanen A, Rousu J, *et al*. Finding feasible pathways in metabolic networks. *Proceedings of the 10th*

Panhellenic Conference on Informatics (PCI'2005), Lecture Notes in Computer Science, 2005.

74. Price N. Genome-scale microbial in silico models: the constraints-based approach. *Trends Biotechnol* 2003;**21**: 162–9.
75. Varma A, Palsson BO. Metabolic flux balancing: basic concepts, scientific and practical use. *Nat Biotechnol* 1994;**12**: 994–8.
76. Kauffman KJ, Prakash P, Edwards JS. Advances in flux balance analysis. *Curr Opin Biotechnol* 2003;**14**:491–6.
77. Schuster S, Dandekar T, Fell DA. Detection of elementary flux modes in biochemical networks: a promising tool for pathway analysis and metabolic engineering. *Trends Biotechnol* 1999;**17**:53–60.
78. Schuster S, Fell DA, Dandekar T. A general definition of metabolic pathways useful for systematic organization and analysis of complex metabolic networks. *Nat Biotechnol* 2000;**18**:326–32.
79. Stelling J, Klamt S, Bettenbrock K, *et al*. Metabolic network structure determines key aspects of functionality and regulation. *Nature* 2002;**420**:190–3.
80. Chvatal V. Linear Programming. New York: W. H. Freeman, 1983.
81. Perumal D, Samal A, Sakharkar KR, *et al*. Targeting multiple targets in Pseudomonas aeruginosa PAO1 using flux balance analysis of a reconstructed genome-scale metabolic network. *J Drug Target* 2011;**19**:1–13.
82. Boyle NR, Morgan JA. Flux balance analysis of primary metabolism in Chlamydomonas reinhardtii. *BMC Syst Biol* 2009;**3**:4.
83. Puchalka J, Oberhardt MA, Godinho M, *et al*. Genome-scale reconstruction and analysis of the Pseudomonas putida KT2440 metabolic network facilitates applications in biotechnology. *PLoS Comput Biol* 2008;**4**:e1000210.
84. Resendis-Antonio O, Reed JL, Encarnación S, *et al*. Metabolic reconstruction and modeling of nitrogen fixation in Rhizobium etli. *PLoS Comput Biol* 2007;**3**:1887–95.
85. Radhakrishnan D, Rajvanshi M, Venkatesh KV. Phenotypic characterization of Corynebacterium glutamicum using elementary modes towards synthesis of amino acids. *Syst Synth Biol* 2010;**4**:281–91.
86. Oh Y-K, Palsson BØ, Park SM, *et al*. Genome-scale reconstruction of metabolic network in Bacillus subtilis based on high-throughput phenotyping and gene essentiality data. *J Biol Chem* 2007;**282**:28791–9.
87. Mahadevan R, Palsson BØ, Lovley DR. In situ to in silico and back: elucidating the physiology and ecology of Geobacter spp. using genome-scale modelling. *Nat Rev Microbiol* 2010;**9.**
88. Wittmann C, Heinzle E. Modeling and experimental design for metabolic flux analysis of lysine-producing Corynebacteria by mass spectrometry. *Metabolic Eng* 2001; **3**:173–91.
89. Poolman MG, Venkatesh KV, Pidcock MK, *et al*. A method for the determination of flux in elementary modes, and its application to Lactobacillus rhamnosus. *Biotechnol Bioeng* 2004;**88**:601–12.
90. Risso C, Sun J, Zhuang K, *et al*. Genome-scale comparison and constraint-based metabolic reconstruction of the facultative anaerobic Fe(III)-reducer Rhodoferax ferrireducens. *BMC Genomics* 2009;**10**:447.
91. Sun J, Sayyar B, Butler JE, *et al*. Genome-scale constraint-based modeling of Geobacter metallireducens. *BMC Syst Biol* 2009;**3**:15.
92. Suthers PF, Zomorrodi A, Maranas CD. Genome-scale gene/reaction essentiality and synthetic lethality analysis. *Mol Syst Biol* 2009;**5**:301.
93. Wunderlich Z, Mirny L. Using topology of the metabolic network to predict viability of mutant strains. *Genome Biol* 2005;**6**:P15.
94. Pharkya P, Burgard AP, Maranas CD. OptStrain: a computational framework for redesign of microbial production systems. *Genome Res* 2004;**14**:2367–76.
95. Pharkya P, Burgard AP, Maranas CD. Exploring the overproduction of amino acids using the bilevel optimization framework OptKnock. *Biotechnol Bioeng* 2003;**84**:887–99.
96. Trinh CT, Unrean P, Srienc F. Minimal Escherichia coli cell for the most efficient production of ethanol from hexoses and pentoses. *Appl Env Microbiol* 2008;**74**:3634–43.
97. Fong SS, Burgard AP, Herring CD, *et al*. In silico design and adaptive evolution of Escherichia coli for production of lactic acid. *Biotechnol Bioeng* 2005;**91**:643–8.
98. Raman K, Chandra N. Flux balance analysis of biological systems: applications and challenges. *Brief Bioinform* 2009;**10**: 435–49.
99. Luscombe NM, Babu MM, Yu H, *et al*. Genomic analysis of regulatory network dynamics reveals large topological changes. *Nature* 2004;**431**:308–12.
100. Konagurthu AS, Lesk AM. Single and multiple input modules in regulatory networks. *Proteins* 2008;**73**:320–4.
101. Gopalacharyulu PV, Velagapudi V, Lindfors E. Dynamic network topology changes in functional modules predict responses to oxidative stress in yeast. *Mol Bio Syst* 2009.
102. Ideker T, Krogan NJ. Differential network biology. *Mol Syst Biol* 2012;**8**:1–9.
103. Stumpf MPH, Wiuf C, May RM. Subnets of scale-free networks are not scale-free: sampling properties of networks. *Proc Natl Acad Sci USA* 2005;**102**:4221–4.
104. Li M, Zhang H, Wang J, *et al*. A new essential protein discovery method based on the integration of protein-protein interaction and gene expression data. *BMC Syst Biol* 2012;**6**:15.
105. Wang J, Li M, Wang H, *et al*. Identification of essential proteins based on edge clustering coefficient. *IEEE/ACM Transact Comput Biol Bioinform* 2011;**6.**
106. Zheng Y, Sriram G. Mathematical modeling: bridging the gap between concept and realization in synthetic biology. *J Biomed Biotechnol* 2010;**2010**:541609.
107. Bellouquid A, Delitala M. Mathematical modeling of complex biological systems - a kinetik theory approach 2006;**188.**
108. Allman ES, Rhodes JA. Mathematical models in biology, an introduction. Cambridge (UK): Cambridge University press, 2004.
109. Goutsias J, Kim S. A nonlinear discrete dynamical model for transcriptional regulation: construction and properties. *Biophys J* 2004;**86**:1922–45.
110. Gao J, Li L, Wu X, *et al*. BioNetSim: a Petri net-based modeling tool for simulations of biochemical processes. *Protein Cell* 2012;**3**:225–9.
111. Pozo C, Marin-Sanguino A, Alves R, *et al*. Steady-state global optimization of metabolic non-linear dynamic

models through recasting into power-law canonical models. *BMC Syst Biol* 2011;**5**:137.

112. Maus C, Rybacki S, Uhrmacher AM. Rule-based multi-level modeling of cell biological systems. *BMC Syst Biol* 2011;**5**:166.
113. Visser D, Schmid JW, Mauch K, *et al*. Optimal re-design of primary metabolism in Escherichia coli using linlog kinetics. *Metabolic Eng* 2004;**6**:378–90.
114. Alves R, Sorribas A. In silico pathway reconstruction: Iron-sulfur cluster biogenesis in Saccharomyces cerevisiae. *BMC Syst Biol* 2007;**1**:10.
115. Liebermeister W, Uhlendorf J, Klipp E. Modular rate laws for enzymatic reactions: thermodynamics, elasticities and implementation. *Bioinformatics* 2010;**26**:1528–34.
116. Ay A, Arnosti DN. Mathematical modeling of gene expression: a guide for the perplexed biologist. *Crit Rev Biochem Mol Biol* 2011;**46**:137–51.
117. Savageau M. Biochemical systems analysis. 3. Dynamic solutions using a power-law approximation. *J Theor Biol* 1970;**26**:215–26.
118. Savageau M. Biochemical systems analysis. I. Some mathematical properties of the rate law for the component enzymatic reactions. *J Theor Biol* 1969;**25**:365–9.
119. Savageau M. Biochemical systems analysis. II. The steady-state solutions for an n-pool system using a power-law approximation. *J Theor Biol* 1969;**25**:370–9.
120. Yagil G, Yagil E. On the relation between effector concentration and the rate of induced enzyme synthesis. *Biophys J* 1971;**11**:11–27.
121. Yagil G. Quantitative aspects of protein induction. *Curr Topic Cell Reg* 1975;**9**:183–236.
122. Teusink B, Passarge J, Reijenga CA, *et al*. Can yeast glycolysis be understood in terms of in vitro kinetics of the constituent enzymes? Testing biochemistry. *Eur J Biochem/FEBS* 2000;**267**:5313–29.
123. Wright B, Kelly P. Kinetic models of metabolism in intact cells, tissues, and organisms. *Curr Topic Cell Reg* 1981;**19**:103–58.
124. Rizzi M, Baltes M, Theobald U, *et al*. In vivo analysis of metabolic dynamics in Saccharomyces cerevisiae: II. Mathematical model. *Biotechnol Bioeng* 1997;**55**:592–608.
125. Vaseghi S, Baumeister A, Rizzi M, *et al*. In vivo dynamics of the pentose phosphate pathway. *Metabol Eng* 1999;**140.**
126. Visser D, Heijnen JJ. Dynamic simulation and metabolic re-design of a branched pathway using linlog kinetics. *Metabol Eng* 2003;**5**:164–76.
127. Anbumathi P, Bhartiya S, Venkatesh KV. Mathematical modeling of fission yeast Schizosaccharomyces pombe cell cycle: exploring the role of multiple phosphatases. *Syst Synth Biol* 2011;**5**:115–29.
128. Caldara M, Dupont G, Leroy F, *et al*. Arginine biosynthesis in Escherichia coli: experimental perturbation and mathematical modeling. *J Biol Chem* 2008;**283**:6347–58.
129. Rabouille S, Staal M, Stal LJ, *et al*. Modeling the dynamic regulation of nitrogen fixation in the cyanobacterium Trichodesmium sp. *Appl Environ Microbiol* 2006;**72**:3217–27.
130. Dräger A, Kronfeld M, Ziller MJ, *et al*. Modeling metabolic networks in C. glutamicum: a comparison of rate laws in combination with various parameter optimization strategies. *BMC Syst Biol* 2009;**3**:5.
131. Singh VK, Ghosh I. Kinetic modeling of tricarboxylic acid cycle and glyoxylate bypass in Mycobacterium tuberculosis, and its application to assessment of drug targets. *Theor Biol Med Model* 2006;**3**:27.
132. Jamshidi N, Palsson BØ. Mass action stoichiometric simulation models: incorporating kinetics and regulation into stoichiometric models. *Biophys J* 2010;**98**:175–85.
133. Smallbone K, Simeonidis E, Swainston N, *et al*. Towards a genome-scale kinetic model of cellular metabolism. *Syst Biol* 2010;**4.**
134. Beisel CL, Storz G. Base pairing small RNAs and their roles in global regulatory networks. *FEMS Microbiol Rev* 2010;**34**:866–82.
135. DiChiara JM, Contreras-Martinez LM, Livny J, *et al*. Multiple small RNAs identified in Mycobacterium bovis BCG are also expressed in Mycobacterium tuberculosis and Mycobacterium smegmatis. *Nucleic Acids Res* 2010;**38**:4067–78.
136. Mraheil MA, Billion A, Kuenne C, *et al*. Comparative genome-wide analysis of small RNAs of major Gram-positive pathogens: from identification to application. *Microbial Biotechnol* 2010;**3**:658–76.
137. Storz G, Vogel J, Wassarman KM. Regulation by small RNAs in bacteria: expanding frontiers. *Mol Cell* 2011;**43**:880–91.
138. Bradley ES, Bodi K, Ismail AM, *et al*. A genome-wide approach to discovery of small RNAs involved in regulation of virulence in Vibrio cholerae. *PLoS Pathogen* 2011;**7**:e1002126.
139. Covert MW, Palsson BØ. Transcriptional regulation in constraints-based metabolic models of Escherichia coli. *J Biol Chem* 2002;**277**:28058–64.
140. Shlomi T, Eisenberg Y, Sharan R, *et al*. A genome-scale computational study of the interplay between transcriptional regulation and metabolism. *Mol Syst Biol* 2007;**3**:101.
141. Covert MW, Xiao N, Chen TJ, *et al*. Integrating metabolic, transcriptional regulatory and signal transduction models in Escherichia coli. *Bioinformatics* 2008;**24**:2044–50.
142. Vilaça P, Rocha I, Rocha M. A computational tool for the simulation and optimization of microbial strains accounting integrated metabolic/regulatory information. *Bio Syst* 2011;**103**:435–41.
143. Kim J, Varner J, Ramkrishna D. A hybrid model of anaerobic E. coli GJT001: combination of elementary flux modes and cybernetic variables. *Biotechnol Prog* 2008;**24**:993–1006.
144. Tenazinha N, Vinga S. A survey on methods for modeling and analyzing integrated biological networks. *IEEE/ACM Transact Comput Biol Bioinform* 2011;**8**:943–58.
145. Keller EF. Revisiting "scale-free" networks. *BioEssays: news and reviews in molecular, cellular and developmental biology* 2005;**27**:1060–8.