A Comprehensive View on Metabolic Pathway Analysis Methodologies

Article *in* Current Bioinformatics · May 2014

DOI: 10.2174/1574893609666140516005147

2 authors:

Dr. Namrata Tomar
Medical College of Wisconsin
54 PUBLICATIONS **404** CITATIONS
SEE PROFILE

Rajat Kumar De
Indian Statistical Institute
158 PUBLICATIONS **1,514** CITATIONS
SEE PROFILE

Current Bioinformatics, 2014, *9*, 295-305

A Comprehensive View on Metabolic Pathway Analysis Methodologies

Namrata Tomar* and Rajat K. De

Machine Intelligence Unit, Indian Statistical Institute, 203 B.T. Road, Kolkata 700108, India

Abstract: Advances in 'omics' high-throughput technologies have led to a vast amount and quality of available biological data. It has fostered the development of bioinformatics methods to interpret these data. In this regard, characterization of cellular metabolism is a useful task to understand the phenotypic capabilities of an organism. Several *in silico* approaches have emerged for analysis of metabolic pathways, including structural and stoichiometric analysis, metabolic flux analysis, metabolic control analysis, and several kinetic modeling based analysis. The present article provides the comprehensive survey on existing metabolic pathway analysis methodologies.

Keywords: Elementary flux mode analysis, extreme pathway analysis, flux balance analysis, metabolic pathway modeling.

1. INTRODUCTION

Biochemical pathways such as metabolic, regulatory or signal transduction pathways can be viewed as interconnected processes. These processes build an intricate network of functional and physical interactions between molecules in the cell. Understanding of the interactions among the cellular components plays an important role to interpret a living system. It leads to evolving of a branch known as Systems Biology. The objective of Systems Biology is to study the mechanisms underlying complex biological processes as integrated systems of many interacting components. There are three major kinds of biochemical pathways, signal transduction pathways (STPs), gene regulatory networks (GRNs) and metabolic pathways. STPs are the pathways of molecular interactions that provide communication between the cell membrane and intracellular end-points, leading to some changes in the cell. Signals are transduced by modification of activity of one protein by another one. A GRN provides how genes interact with one another indirectly through their RNA and protein expression products. In this way, it governs the rates at which genes in the network are transcribed into mRNA. GRNs determine whether or not a particular gene is expressed at a particular time. Metabolic pathways are the sequences of biochemical reactions, mostly being catalyzed by enzymes, where certain product molecules are formed from other substrates. Metabolites are usually small molecules while enzymes are proteins.

The first issue that needs to be addressed in order to be able to perform systematic analyses of metabolic pathway is the choice of an appropriate model to represent them. For this, metabolic pathway first maps onto a graph and the choice for mapping may partly depend on the purpose of the analysis and the algorithms used. The metabolic pathway analysis involves three key problems, *viz*., 1) Modeling and simulation of metabolic processes; 2) Reconstruction of metabolic pathways; and 3) Comparison of metabolism revealing the functional relationship between a set of metabolic pathways. To develop and investigate detailed mathematical models of metabolic processes, the main challenge is that, the modeled system is supposed to reproduce the given data. Current applications of metabolic pathways analysis include the tasks, like, finding pathways of maximum yield, such as for amino acid and antibiotic synthesis in the area of biotechnology; finding non-redundant pathways that are important in drug design; testing whether a set of enzymes can produce a desired product [1, 2].

In the present article, we have surveyed some of the approaches for metabolic pathway analysis based upon the kind of task, they perform, like, determining optimal pathways from a substrate to target metabolites through which the amount of targets being maximum [3]; deciding over the reactions to be active in a pathway [4]; linking cellular phenotype to corresponding genotype [5]; identifying essential genes and dominant metabolic processes [6,7]; analyzing redundancy [8,9]; gene knock out studies [10-12]; finding out alternative optimal solutions from which the same maximal objective can be fulfilled [13]; and rational strain design [14]. Several techniques are available for the modeling, simulation and analyses of metabolic pathways and networks, which we describe in detail. Models may vary in the level of details, which influence the accuracy of simulation results and the nature of insights obtained from them. The majority of kinetic parameters are not available; therefore, large genome scale metabolic modeling (GSMM) has been benefitted by developing new computational methods that are functional even without experimental parameters [15]. Identification and mathematical computation of biological constraints are found to be more feasible. While describing the basic details behind various algorithms/tools, we have also mentioned their corresponding applications.

There exist survey articles on *in silico* approaches for metabolic pathway modeling with different views. They surveyed only some of the existing metabolic pathway analysis approaches and thus contain very specific information. They give emphasis to a particular approach. For

*Address correspondence to this author at the Machine Intelligence Unit, Indian Statistical Institute, 203 Barrackpore Trunk Road, Kolkata 700108, India; Tel: +91-33-25753112; Fax: +91-33-25783357;
E-mail: namrata_t@isical.ac.in

2212-393X/14$58.00+.00

example, in [16], the emphasis is on isotope labeling experiments, while only elementary modes and extreme paths are compared and assessed in [17]. Price *et al.* [9] have evaluated constraints for genome scale metabolic models (GSMMs), graph based approaches for network analysis are discussed in [18], data models types with emphasis on structural aspects of biochemical networks are provided in [19], computational tools are for metabolic engineering are highlighted in [20], whereas constraint-based reconstruction and analysis (COBRA) methods are evaluated in [21]. In the present article, we have categorized the existing approaches for metabolic pathway analysis and discussed almost all, under these categories (as given in Table **1**). We have addressed challenges in terms of extraction/retrieval of omics data and automation tools for modeling approaches for experimental biologists in Section 3 (Discussion). Below is a categorization of the existing metabolic pathway analysis methodologies.

2. CATEGORIZATION OF THE EXISTING METABOLIC PATHWAY ANALYSIS METHODOLOGIES

The approaches for analysis of metabolic systems are mainly based on structural/stoichiometric modeling and kinetic data based modeling. Moreover, there exists a few hybrid modeling techniques. The categorization is given in Table **1**. Metabolic pathway information related to structure and topology can be translated into mathematical terms through stoichiometric matrix. The other kind of approach is kinetic modeling which describes dynamic biochemical reactions through nonlinear ordinary or partial differentiation equations.

2.1. Structural and Stoichiometric Modeling

The stoichiometric information of a metabolic network can be represented by a matrix, where underlying biological knowledge can be translated into mathematical terms. It assumes the pseudosteady state for the internal metabolites, and thereby it excludes dynamic intracellular behavior. One of the reasons behind the success of structural stoichiometric modeling is that one can avoid the problems in developing the kinetic models that involve a lot of intracellular experimental measurements [22]. Analysis of structural invariants of a biochemical network has been first applied to a system of chemical reactions [23]. Further, it has been generalized for investigating steady states and stabilities of dynamical systems with reaction stoichiometry [24,25]. Stoichiometric network analysis (SNA) is applied to study qualitatively, nonlinear dynamics of chemical reaction mechanisms including the others with stoichiometry [26]. SNA, based on convex analysis, was the pioneering work which identifies unique pathways. Stoichiometric models constitute the basic framework for fluxome quantification in the realm of metabolic engineering. Under stoichiometric modeling, there are various approaches, *viz.*, network/topology based and constraint-based approaches. There are a few methods/tools under each of these approaches.

2.1.1. Network/Topology Based Approaches

Systemic functions arise from the interactions of various metabolites and enzymes, and cannot be described by an individual components' characterization. Network/topology based approaches have evolved from traditional metabolic pathway definitions as they mathematically describe systemic functions through the existing knowledge of the cellular components and their connectivities. There are three basic methods under this approach- Bipartite graph and hypergraphs based methods, elementary flux mode analysis (EMA) and extreme pathway analysis (EPA).

Bipartite Graphs and Hypergraphs (BGH): Graphs are models for representing relationships between certain

Table 1. Categorization of the existing metabolic pathway analysis approaches.

Paradigm	Approach	Method
Structural and Stoichiometric modeling	Network/Topology based	Bipartite graph and Hypergraph EMA
	Constraints based	FBA EPA ROOM MOMA CoPE-FBA R-DFBA M-DFBA FVA FCA Hybrid MFA
Kinetic modeling	Discrete	PN
	Continuous	ODE based
Hybrid modeling	Structural kinetic modeling (SKM) Hybrid modeling Hybrid Functional PNs Isotope labeling Integration of high-through put data and metabolic models	

objects. Bipartite graphs and hypergraphs offer an unambiguous representation of the reactions and substrates in biochemical networks. The coverage is limited as possible controls of reactions, like, catalysis, inhibition cannot be explicitly represented [19]. These are appropriate when the analysis is limited to reactions and compounds. However, many applications such as topological properties, path finding, synthesis and prediction can be covered by these approaches. In a bipartite graph, there are two classes of nodes; no edges can relate nodes from the same set. In the context of biochemical networks, there are substrate and reaction nodes. Here, an edge relates a substrate node and a reaction node and can be undirected or directed. 31 metabolic networks of 43 organisms are modeled as bipartite graphs, permitting a systematic comparative analysis showing that these metabolic pathways have the same topological scaling properties [27]. Bipartite graphs have also been used for the analysis of global structural properties of metabolic networks stored in KEGG [28], where authors proposed different graph analysis operations, such as path finding, using classical graph algorithms.

A hypergraph is a generalization of a compound graph model, defined as G = (V, E), where V is a set of vertices or nodes, and E is the set of hyperedges. Each hyperedge can connect more that two nodes or we can say that it allows an arbitrary number of nodes to be connected. In biology, the nodes describe proteins, metabolites, genes, or other biological molecules, whereas the edges represent functional relationships among these nodes. In other words, hypergraph has nodes as a set of substrates and hyperedges connect a set of substrates to a set of products. One can get detailed information in [29] on hypergraph as applied to biochemical pathways. Graph representation of a biological pathway has been used for studying interesting properties like robustness, modularity, connectivity and finding motifs/clusters. Moreover, graph-theoretic concepts can be used to predict the structural and dynamical properties of an underlying network [18,30]. The objectives include analysis of network structure, shortest path finding, path enumeration, and reaction clustering. Classical path finding algorithms need some adaptation in order to yield meaningful results for metabolic networks, especially, the reversibility of a biochemical reaction and the necessity for being selective while navigating through intermediate compounds [28]. However, path finding methods do not consider reaction stoichiometry, and the path found might not be operated in sustained steady-state. For this reason, Pey *et al.* [31] have incorporated reaction stoichiometry into path-finding approaches, known as 'flux path', *via* mixed-integer linear programming, and shown improved prediction of topological and functional properties of metabolic networks. Path structure of a metabolic network is found to be changed after incorporating the stoichiometry.

2.1.2. Constraints Based Approaches

The constraint-based modeling describes a metabolic pathway system by a set of constraints which characterize its possible behaviors, when the biological information is typically incomplete. The classical starting point of constraint-based modeling is flux balance analysis (FBA) of a metabolic network at steady state. The theory applied to analyze metabolic networks is developed based on the first principle of mass conservation of internal metabolites within a system [32,25]. The unique set of systemically independent biochemical pathways/extreme pathways is identified using an algorithm, based on system stoichiometry and limited thermodynamics [33]. Here we describe various tools, along with FBA, under constraint based approaches.

Flux Balance Analysis (FBA): Flux balance analysis (FBA) is a mathematical method for analyzing the metabolic capacity of a cell. It does not require knowledge of metabolite concentration or details of the enzyme kinetics of the system. Thus, it greatly reduces the computer time required for simulations. The assumption is that the system being studied is in homeostasis and follows the law of mass conservation. This approach aims at finding out a target metabolite, and the set of metabolic fluxes that maximizes the growth rate of a target metabolite, given some known available nutrients. Describing some of the earliest work in FBA includes that of Papoutsakis [34], which has demonstrated the way to construct flux balance equations using a metabolic map. Watson [35,36], however, first introduced the idea of using linear programming (LP) and an objective function to solve for an optimal path. The key parts for model preparation for a metabolic pathway analysis under FBA are to create a metabolic network, add constraints and to decide over an objective function. In FBA, there are a large number of mathematically acceptable solutions to the steady-state problem*, i.e.*, $\mathbf{S.v} = 0$, but the ones that are biologically interesting are those that produce the desired metabolites in the correct proportion. In equation, $\mathbf{S.v} = 0$, the representation of a biochemical pathway can be in the form of a stoichiometric matrix, **S**, and is composed of stoichiometric coefficients that capture the underlying reactions of the biochemical pathway. The concentration of molecules (*e.g.*, transcription factors/enzymes) catalyzing various reactions in the network may not be expressed at the desired level. Thus, a modified formulation for the steady state given by $\mathbf{S(C.v)} = 0$ [37,38]. Here, **C** is an $n \times n$ diagonal matrix whose diagonal elements represent relative concentrations of enzymes catalyzing the reactions. A notable example of the success of FBA is the ability to accurately predict the growth rate of the prokaryote *E. coli* when cultured in different growth media [39], and to define precise minimal media for the culture of *Salmonella typhimurium* [40]. Thus the suitable organisms can be cultivated in media with defined concentrations of nutrients and their growth rates measured. Hence, the predictions of FBA can be compared with the experimental results. In contrast to elementary mode analysis and extreme pathway analysis, only a single solution results in the end through applying FBA. It should be mentioned here that FBA identifies only one optimal solution while alternative optimal solutions or suboptimal solutions may exist.

The scope of FBA has been widened as it computes metabolic fluxes in large metabolic networks and is important in network curation and structural analysis. Thus, FASIMU [41] (http://www.bioinformatics.org/fasimu), a command line oriented software has been developed, which is suitable enough to realize a wide variety of FBA algorithms and able to handle batch series of flux-balance optimizations. It computes flux distributions using a variety of FBA algorithms, including the first available implementation of weighted flux minimization, fitness

maximization for partially inhibited enzymes, and using the concentration-based thermodynamic feasibility constraint. It allows batch computation with different objectives and constraints, which are helpful in network pruning, leak analysis, Flux Variable Analysis, and systematic probing of metabolic objectives for network curation. FBA was used to analyze a cancer network model to validate which predicted growth-supporting genes. The knockdown of these genes would significantly reduce cellular proliferation rate [42]. FBA gives a huge number of fluxes of the same optimal performance, and thus the complete description of the resulting optimal solution spaces is difficult to be solved computationally. For this reason, Comprehensive Polyhedra Enumeration Flux Balance Analysis (CoPE-FBA) [43] has been developed, which indicates that the thousands to millions of optimal flux patterns result from a combinatorial explosion of flux patterns in just a few metabolic sub-networks. The entire optimal solution space can now be compactly characterized in terms of the topology of these sub-networks.

Extreme Pathway Analysis (EPA): It is an alternative method [44], which can be considered as a hybridization of stoichiometric network analysis and elementary mode analysis. Extreme pathways account for all the reaction steps a network must use to complete the synthesis process. In calculating extreme pathways, the analysis splits only the internal reversible reactions into two irreversible reactions while not decomposing reversible exchange reactions. Bell *et al.* [45] has computed extreme pathways in a program, called ExPa (available at http://systemsbiology.ucsd.edu) [44]. Extreme pathways can be characterized by their length and reaction participation. The length of an extreme pathway is the number of reactions that it comprises. Reaction participation is the percentage of extreme pathways that utilize a given reaction. These properties have been applied for the computation for the production of individual amino acids and protein production in *H. pylori*, and individual amino acid production in *H. influenzae* [8]. The distribution of lengths of extreme pathways in *H. pylori* and *H. influenzae* differs, which indicates a systemic difference between the two organisms, despite their overall similar metabolic networks. Reactions that participate in a large number of extreme pathways may represent potential targets for regulation. Thus, one can use extreme pathway lengths and reaction participation numbers characterizing the large-scale properties of metabolic networks. Another application of extreme pathway analysis has been done on human red blood cell (RBC) metabolic network, and interpreted RBC metabolism and its metabolic physiology [46]. Wiback *et al.* [47] has defined the α-spectrum, which is the allowable range of extreme pathway contributions to a given flux distribution. This α-spectrum was computed for two cases, *viz*., for optimal states of a skeleton representation of core metabolism including transcriptional regulation and for human red blood cell metabolism under various physiological and non-optimal conditions.

Elementary Mode Analysis (EMA): Elementary Mode Analysis, introduced by Schuster *et al.* [5], is a useful metabolic pathway analysis tool to identify the structure of a metabolic network. The analysis decomposes an intricate metabolic network comprising highly interconnected reactions, which is a minimal set of enzymes into uniquely organized pathways. All metabolic capabilities in steady states are composed of elementary flux modes that can individually generate valid steady states [48,49]. In this way, it links cellular phenotype to the corresponding genotype. Elementary Mode Analysis calculates all the solutions in the admissible flux space by solving equation $\mathbf{S.v} = \mathbf{0}$ in conjunction with the thermodynamic constraint and an additional non-decomposability constraint. Here $\mathbf{S}$ is the stoichiometry matrix of dimension $m \times n$, where m is the number of metabolites involved in a metabolic network comprising n reactions. An element of $\mathbf{S}$, $s$ij, represents the stoichiometry coefficient of ith metabolite involved in jth reaction, and v is the flux vector of dimension n. Each solution represents an elementary flux mode. The advantage of using the non-decomposability constraint is that each elementary mode is unique up to a positive scalar factor because deletion of any reaction in an elementary mode will disrupt the whole pathway. Valid states of the network are linear superpositions of the elementary modes forming a polyhedral cone/flux cone. Thus it is clear that computing elementary flux modes is equivalent to extreme ray enumeration of polyhedral cones.

A new recursive enumeration strategy with bit pattern trees for adjacent rays is the ancestors of extreme rays, has been formulated by Terzer *et al.* [50,51]. Distinguishing extreme rays from normal (composite) vectors is a aspect of the algorithm, and it is roughly one order of magnitude faster than previous methods. This method has been applied to a central metabolism of *Escherichia coli*, and resulted in 26 million elementary flux modes and also analyzed the simultaneous production of non-essential amino acids in a genome scale metabolic network of *Helicobacter pylori*. It is found that all steady states may not be interpreted as real biological systems. Thus another approach has been introduced, based on quadratic programming (QP), where the decomposition is performed in a way so that the resulting modes are closest to the actual state of the system. The method has been applied to a yeast glycolysis model [52].

Elementary flux mode computation in large-scale metabolic networks is a challenging issue due to its combinatorial complexity. For this reason, a novel procedure based on linear integer programing was presented to determine the K-shortest EFMs in large scale metabolic networks [53]. K-shortest EFMs can be applied for several biological applications as shorter pathways are better targets for genetic manipulation and can carry higher fluxes [54, 55]. A novel optimization-based method, i.e., a convex basis was presented for determining a minimal generating set of elementary flux modes, and was shown that a subset of elements of this convex basis could be effectively computed even in large metabolic networks [56]. A convex basis is typically referred to as a minimal generating set and the elements of this basis as generating flux modes, GFMs [57]. A method has been developed for the decomposition of metabolic flux distributions into elementary flux modes that easily operate on large genome-scale metabolic networks [58]. It is useful to large networks as it does not require all relevant modes, of the metabolic network, to be generated. Its application is shown for metabolic engineering of *E. coli* and for understanding the survival of *M. tuberculosis* during infection. It has also been said that this method can equally be applied to flux distributions obtained by FBA, fluxes

derived from experiments, MOMA [10] or ROOM [11]. Some of the publicly available softwares for calculating elementary modes are METATOOL [59,60], GEPASI [61], COPASI [62] and FluxAnalyzer [17].

Regulatory On/Off Minimization (ROOM): Predicting the metabolic state of an organism after a gene knockout is a challenging task, as the regulatory system governs a series of transient metabolic changes. Thus, a constraint-based algorithm, called Regulatory on-off minimization (ROOM), has been introduced for predicting the metabolic steady state after gene knockouts [11]. It aims at minimizing the number of significant flux changes (hence on-off) with respect to the wild type. It has been shown to accurately predict steady-state metabolic fluxes, having similarity with experimental flux measurements, and to correctly identify short alternative pathways used for rerouting metabolic flux in response to gene knockouts. ROOM-based DFBA approach (R-DFBA) is developed by combining the principle of ROOM with DFBA [63]. R-DFBA extends the ROOM approach by considering the minimization of the total number of significant changes of metabolite concentrations. It allows a large concentration change at a few time points (orthogonal roots) and a relatively constant concentration of the metabolites.

Minimization of Metabolic Adjustment (MOMA): The assumption of optimality for a wild-type bacterium is justifiable, but it cannot be the same for mutants/knockouts. Knockout metabolic fluxes undergo a minimal redistribution with respect to the flux configuration of the wild type. In order to address this issue, minimization of metabolic adjustment (MOMA) [10], a quadratic programming (QP) based method has been introduced to identify a point in flux space, which is closest to a wild-type point, compatible with the gene deletion constraint. Comparing MOMA and FBA predictions to experimental flux data for *E. coli* pyruvate kinase mutant PB25, it has been found that MOMA displays a significantly higher correlation than FBA. MOMA is further supported by experimental data for *E. coli* knockout growth rates. Thus it can be used for networks, whose growth performance is in general suboptimal. MOMA does not assume optimality of growth or of any other metabolic function, in contrast to FBA. Instead, MOMA approximates metabolic phenotype by performing distance minimization in flux space for perturbations. In case of lethality of some *E. coli* gene deletions, MOMA, unlike FBA, has been able to predict correctly.

In MOMA-based DFBA (M-DFBA) approach, DFBA has been combined with MOMA and shown to minimize the fluctuations in metabolic concentrations over time [64]. It has predicted the dynamics of photosynthetic metabolism and enabled to make hypotheses about its robustness under different CO2 and water conditions. In contrast to R-DFBA, which prevents a large number of small (significant) changes over time, M-DFBA captures smooth changes over time using Euclidean distance.

Flux Variability Analysis (FVA): It has already been pointed out that FBA identifies only one optimal solution while alternative optimal solutions or suboptimal solutions may exist. In general, FBA calculates metabolic flux vectors based on limited experimental data and requires objective functions. If one can measure more fluxes, it will lead to the more accurate computational determination of the flux vectors [49]. A key issue of constraint based models is the existence of alternate optimal solutions for the same maximal objective, which can be achieved through different flux distributions. Thus Flux Variability Analysis (FVA), an LP-based method has been described to calculate the range of flux variability that achieves optimal as well as suboptimal objective states [13]. It is used to find the minimum and maximum fluxes for reactions in a network, maintaining some states of the network. Flux variability analysis (FVA) is often used to determine the robustness of metabolic pathways under various simulation conditions, but it is limited by large computation time. Therefore, an open source implementation of FVA, called fastFVA, has been developed [65] for the exploration of alternative optima of the equation $\mathbf{S.v} = 0$, studying flux distributions under suboptimal growth, investigating network redundancy, optimal strain design and optimization of process feed formulation for antibiotic production. The performance of fastFVA has been evaluated on six biochemical network systems comprising 650-13,700 reactions.

Flux Coupling Analysis (FCA): Flux coupling analysis (FCA) finds dependencies among fluxes of a metabolic network at steady-state. FCA classifies reactions into subsets of coupled reaction sets in which activity of one reaction implies activity of another reaction. If a non-zero flux through a reaction in steady-state implies a non-zero flux through another reaction then the first reaction is said to be coupled with the second one. Flux Coupling Finder (FCF) method is based on solving LPs to find out coupled reactions (directionally, partially or fully coupled) and blocked reactions in genome-scale metabolic systems. It helps in elucidating the topological and flux connectivity features of genome-scale metabolic networks [66]. If the maximum possible value of a particular flux is zero, then the reaction is said to be blocked. Linear fractional programming is employed to identify the maximum and minimum flux ratios for every pair of fluxes. It has been demonstrated on genome-scale metabolic reconstructions of *H. pylori, E. coli* and *S. cerevisiae*. It also identifies the essential core of reactions whose activity is required for cellular growth for a given condition. Thus the method allows evaluating the lethality of any knockout by examining the essential core of metabolic reactions.

Another method of FCA is based on Minimal Metabolic Behaviors, MMB-FCA [67], which computes a minimal set of generating vectors of the flux cone. Then, the coupling relation for any pair of reactions is inferred based on the co-appearance of nonzero fluxes in the generating vectors. FCA was applied to compare biological subsystems and genome scale metabolic networks [68]. When applied to real world metabolic system models, as observed in the case of plastid subsystem, it has been found that a non-negligible number of reaction pairs in plastids may have altered flux coupling relations when the plastids are studied in isolation. Flux coupling was also used to explain metabolic gene coregulation that was earlier ignored by graph-theoretical techniques [69]. FCA was also applied to photosynthetic bacterium Synechocystis sp. PCC6803 under four different growth conditions, *viz*., autotrophy, mixotrophy, heterotrophy, and light-activated heterotrophy (LH) [70]. It was integrated with transcriptomic data to identify reporter

flux coupling pairs and reporter flux coupling groups, *i.e.*, regulatory hot spots during metabolic shifts triggered by the availability of light.

Feasibility-Based Flux Coupling Analysis (FFCA, http://www.bioinformatics.org/ffca/) [71] is based on checking the feasibility of a system under the constraint of linear inequalities. The theoretical basis of FFCA is finding a feasible solution, which is faster than computing an optimal solution. This method does not distinguish between partial and full coupling, since flux patterns only contain the information about the activity or inactivity of the fluxes, but not regarding the flux values. Both FCF and FFCA are based on FCA and solve linear programs (LPs) for flux coupling analysis, but FFCA is faster than FCF. In FFCA, finding the first feasible solution is sufficient, while the LPs should be solved for optimality in case of the FCF algorithm. In FCF, every reversible reaction, in contrast to FFCA, is divided into two (forward and backward) irreversible reactions. This step slows down the procedure and increases the size of the LPs to be solved.

Hybrid Metabolic Flux Analysis (Hybrid MFA): Although, stoichiometric models constitute the basic framework for fluxome quantification, a hybrid method has been developed combining classical metabolic flux analysis (MFA) [72] and projection to latent structures (PLS) [73]. Hybrid metabolic flux analysis (Hybrid MFA) [74] extends the study of Bernal *et al.* [75] in which they have developed rational strategies for baculovirus production optimization in insect cell cultures, based on classical Metabolic Flux Analysis (MFA). These results are combined with the ability of hybrid metabolic flux analysis (hybrid MFA) to assign individual fluxes/pathways of central metabolism to cell specific functions that cannot be completely defined in a stoichiometric network analysis. Hybrid MFA is a suitable tool for metabolic identification and quantification in incomplete metabolic networks.

As mentioned earlier, computing elementary modes in larger metabolic networks is expensive as the number of elementary modes grows exponentially with network size. Thus the concept of elementary flux patterns considers the possible fluxes through the entire network, while analyzing the steady-state fluxes through a subnetwork. Elementary flux patterns can be defined as a set of reactions within a subsystem of a larger network, which represents the basic routes of each steady-state flux of a large network through a subnetwork. Kaleta *et al.* [76] have suggested that elementary flux patterns can be used for characterizing flux coupling relations. They have analyzed a system of the tricarboxylic acid cycle (TCA) and adjacent reactions in *E. coli*, and detected several pathways that can be used as alternative routes to some central metabolic pathways. This can be further applied to the computation of minimal media, the development of knockout strategies, and the analysis of combined genome-scale networks. In this way, one can say that elementary flux pattern is an application derived from the concept of elementary mode analysis to genome-scale metabolic networks without the drawbacks that arise due to exchange fluxes.

2.2. Kinetic Modeling

Biological processes are time dependent, which identify the flow of mass between cellular components. Quantitative description of the concentration changes due to active pathways in a protein interaction network is a challenging task. Computational models help quantify the reaction dynamics and lead to study the dynamics of metabolic behavior over time. A kinetic reaction network consists of reactions present in a biological pathway and conventionally, the set of reactions is described by nonlinear ordinary or partial differential equations (ODEs or PDEs). Kinetic models are important for modeling behaviors such as oscillations and bi-stability.

2.2.1. Discrete Approach

Biochemical reactions take place in continuous time process, whereas experimental data are measured by sampling the continuous biochemical reaction output at discrete time points. Thus we can say that the discrete time models can act as an interface between kinetic reactions, experiments and computer based simulation. Discrete approaches consist of time-consuming processes, and therefore, the logical activation functions for updating the network must be explicit. A discrete-time model not only helps in better understanding of pathway reaction dynamics and reproduces the experimental data, but also generates predictions which rules out the expensive and time consuming lab experiments.

Petri Net (PN) Modeling: A Petri Net (PN) is a formal description for modeling concurrent systems developed by Carl Adam in 1960 [77]. A PN is defined as a 5-tuple, (P, T, F, W, M_0), where $P = \{p_1, p_2,..., p_m\}$ is the set of *m* places, $T = \{t_1, t_2,..., t_n\}$ is the set of n transitions and $F \subseteq (P \times T) S (T \times P)$ is the set of arcs. W: $F \rightarrow N$, N being the set of natural numbers, is called the weighting function, and M_0: $P \rightarrow \{0, 1,..., m\}$ stands for the number of tokens in a place. A PN can also be represented as an incidence matrix **C**, equivalent to the stoichiometric matrix of a metabolic network. All matrix entries are either 0, 1, or -1; these quantities specify the absence or presence of a connecting edge between two places, as well as its direction.

Concentrations of molecules at an instant of time are expressed as the discrete number of tokens, assigned to a place. A transition is said to fire when its input places have the minimum number of tokens specified in the corresponding arc weights (enabled transition). One can model and represent complex metabolic network through PNs in a way which enable one to manipulate and analyze the functionality of a cell [78,79]. Barjis and Gehlot [80] have modeled and simulated the process of carbohydrate metabolism in general and the pentose phosphate pathway in particular. A computational model has been developed for citric acid cycle (Krebs cycle), which accurately decomposes a complex network into a complete family of principal sub circuits to study the minimal circuits *via* PNs [81]. Some combinations of these principal sub circuits can exhibit every operational behavior of the network under consideration. Basic PNs have also been used for modeling [82] *Trypanosoma brucei* glycolysis in which they have illustrated the biological meaning of PN properties, like siphons, traps, deadlocks and liveness. PN theory can be used for the structural validation of metabolic networks through qualitative analysis [83], where sucrose breakdown metabolism in a potato tuber has been studied. It has been

shown that T-invariants have related to combinations of subpathways in a given network. Colored PNs (CPNs) for simulating enzymatic reaction chains have also been used in which the associated color is a pair encompassing the name and the concentration of the related substrate [84]. A systematic approach for modeling regulated metabolic pathways with PNs has been developed to study the biosynthesis of tryptophan in *E. coli* [85]. Here, two separate PN models of the metabolic and regulatory pathways have been built, and then they have been integrated into a unique structure.

Drawbacks of Kinetic Modeling Based Approaches: The dynamic models encompass a number of profound difficulties. Kinetic models rely on the enzymatic rate equations and their associated parameter values. Some of the parameter values can be available from the literature, but some of them may also depend on many other factors, like, tissue type or experimental conditions. It has also been found that the most enzyme-kinetic rate laws have been determined *in vitro*, but it cannot be described well if a particular rate law is appropriate *in vivo*.

2.2.2. Continuous Approach

Continuous modeling approach simulates a system pathway quantitatively using time. It uses differential equations specifying how the concentrations of various molecular species evolve over time. Complex numerical computation is one of its drawbacks.

Ordinary Differential Equations (ODEs) Based Analysis: To understand the functionality of complex biological pathways, it is necessary to know its evolution and modeling. It is commonly done through ordinary differential equations (ODEs). The mathematical rate equation that describes the flux v must be chosen carefully so that it describes the mechanistic aspects of the system. The nonlinearity in an enzyme is usually catalyzed metabolic reaction have been commonly captured by mass action or Michaelis-Menten like kinetic models. A model has been developed to study the diauxic growth of *E. coli* in glucose and lactose, using a detailed model of the lac operon regulation (with 13 state variables) which includes catabolite repression or inducer exclusion [86]. A Cellerator language extension, kMech has been introduced that describes a suite of enzyme mechanisms [87]. Each enzyme mechanism is parsed by kMech into a set of fundamental association-dissociation reactions that are translated by Cellerator into ODEs that are numerically solved by MathematicaTM. It requires kinetic measurements to estimate rate constants required to solve these ODEs. The values of the rate constants of all the metabolic reactions and the initial concentration levels of the metabolites are needed to simulate the model. But only a few parameters are known experimentally. The remaining parameters need to be estimated by fitting the model to experimental data. Such data is limited and noisy, and also estimating parameters of continuous ODEs is a computationally costly.

2.3. Hybrid Modeling Approaches

Certain biochemical processes deal with information in different time-scales. Therefore, we can say that they have a clear translation into a hybrid system framework. In other words, certain complex biological systems show continuous and discrete dynamic behaviors simultaneously. Thus hybrid system based modeling approaches are commonly used to model dynamic systems showing both of these behaviors. It deals with different time-scales to account for dynamic modes [88].

2.3.1. Structural Kinetic Modeling (SKM)

While stoichiometric network analysis (SNA) has been proved to be immensely effective to address the functional capabilities of large metabolic networks, it fails to incorporate dynamic aspects into the description of a system. Moreover, the steady state equation allows no conclusions about the stability or possible instability of a metabolic state. Structural kinetic modeling (SKM) has the advantages of stoichiometric analysis, as well as it incorporates the dynamic aspects of the system [89,90]. SKM involves a parametric representation of a metabolic system, in terms of a Jacobian matrix, at each possible point in parameter space. One needs to know the eigenvalues of the associated Jacobian matrix. Here, each element is accessible even without explicit knowledge of the functional form of the rate equations. This parametric representation informs about the quantitative dynamical capabilities of the metabolic system at a given state. The objective of this approach is the statistical evaluation of the Jacobian matrix, with each element of the Jacobian constraint representing the available experimental information. Instead of evaluating a single model, it evaluates an ensemble of possible models with each instance reflecting the available experiments.

2.3.2. Hybrid Functional Petri Nets (HFPNs)

The idea behind Hybrid Functional Petri Nets (HFPNs) [91] includes the definition of functions on discrete and continuous arcs and transitions. The notions of functions on arcs and transitions seem very useful to model dynamic biological systems, since they enable establishment of links between places which are not directly related in the model. Discrete transitions can be set to fire with an associated time-delay. Thus, this idea is well suited for modeling cellular processes having different time-scales and describes the systems dynamics. Chen *et al.* [92] has used it to study the regulation of urea cycle in liver. The results show that the defects of the enzymes in the urea cycle can be treated by limiting the input of ammonia or by replacing the missing intermediates from the cycle, by supplementing with arginine or citrulline.

2.3.3. Isotope Labeling

The use of stable isotopes is found to be a powerful method for flux determination in complex biological systems [93]. Metabolic conversion of isotopically labeled substrates generates molecules with distinct labeling patterns, called isotope isomers (isotopomers). They can be detected by mass spectrometry and nuclear magnetic resonance spectroscopy [94, 95]. Later, an improved version, 13CFLUX2 (www.13cflux.net), with faster optimization algorithms process spectra of MS, MS/MS, 1H-NMR and 13CNMR. There also exists OpenFlux platform (http://openflux.sourceforge.net/) under Matlab [96] that has many 13C flux analysis tools. The

isotopomer composition of metabolic intermediates is fully and uniquely determined by the cell's flux state and the administered isotopic label at metabolic and isotopic steady state. Mathematical models are required for quantitative interpretation of isotopomer data to describe the relationship between metabolic fluxes and the observed isotopomer abundances [97]. Moreover, there also exists 13C metabolic flux analysis (MFA) that aims to compute *in vivo* metabolic fluxes in terms of metabolite balancing extended by carbon isotopomer balances, quantitatively tracks metabolic pathway activity and determines overall enzymatic function in cells [98]. It utilizes 13C labeling patterns of metabolic products and provides detailed information about intracellular pathway fluxes *in vivo* [99]. 13C MFA has certain advantages over stoichiometric MFA. For example, it does not require the assumption of an optimization objective. It uses isotopomer data, and therefore, usually provides more accurate flux estimations. However, 13C MFA is expensive due to the high price of labeled feed isotopes [98, 100].

2.3.4. Integration of High-Throughput Biological Data into Computational Models

Sequencing techniques lead to the rapid increase in the amount of reconstructed metabolic networks. The integration of high-throughput biological data into computational models can be used to generate hypotheses, find inconsistencies, give way to new biological discovery [101], and to lead to effective identification of new components and interactions in biological networks. GSMMs use sequence data to produce detailed and quantitative predictions of organism behavior. Based on the available literature and databases, the first integrated genome-scale computational model of a transcriptional regulatory and metabolic network contains 1, 010 genes in *Escherichia coli*, including 104 regulatory genes [102]. This model was able to predict the outcomes of high-throughput growth phenotyping and gene expression experiments, knowledge gaps, and identify previously unknown components and interactions in the regulatory and metabolic networks.

The model SEED (http://www.theseed.org/models/) has been developed for high-throughput data generation, optimization and analysis of GSMMs [103]. Authors have generated 130 genome-scale metabolic models that represent a taxonomically diverse set of bacteria through applying model resource SEED. Jerby *et al.* [104] have put light on the use of genome scale metabolic method to model human metabolism under healthy and diseased conditions. They have described the use of these models to elucidate the metabolic alterations that accompany cancer progression, and aid the experimental efforts in the identification of drug targets and metabolic biomarkers. MetabolIc Reconstruction *via* functionAl GEnomics (MIRAGE), based on gap filling approach, identifies missing network reactions by integrating metabolic flux analysis and functional genomics data [105]. There is a quadratic programming (QP) based method, *viz*., integrative omics metabolic analysis (IOMA) that quantitatively integrates proteomic and metabolomic data with GSMMs, to more accurately predict metabolic flux distributions [106]. Authors have successfully predicted the metabolic state of human erythrocytes through IOMA, as compared to kinetic model simulations and have shown a significant advantage over FBA and MOMA. Therefore, we can say that GSMMs would be used as a platform for the analysis of high throughput transcriptomic, proteomic and metabolomic data to provide insight into metabolic activity of an organism.

3. DISCUSSION

Mathematical models of cellular metabolism serve as a basis to investigate questions of major biotechnological importance, such as the effects of directed modifications of enzymatic activities to improve a desired property of the system. All *in silico* simulations of metabolic pathways involve a certain level of mathematical abstraction that should be guided by biochemical knowledge and experimental accessibility of system parameters. An important aspect of pathway analysis and reconstruction is the availability of the correct data. Data incompleteness and its quality can have impact on the methodology used. Missing data information usually leads to a large parameter solution space with many possible sets of parameter values, whereas noise and poor data give error.

Constraint based analysis methods, *e.g.*, rFBA, involve an optimization procedure within a polyhedral cone, and define the metabolic and regulatory constraints. In this way, it cuts down wrong information and gives only required flux distribution solution [107]. Moreover, GSMMs have an ability to build novel hypotheses that can lead to generate new experiments. Although, modern biology experimental techniques and computational approaches establish the large experimental genome-scale omics databases, the major issue is with the extraction of all the relevant information from these databases and employs it for designing efficient industrial processes. Hence, the major challenge of metabolic pathway analysis and engineering is to broaden its design methodologies through incorporating different kinds of omics data and approaches. For example, Medina *et al.* [108] have shown stoichiometrically the role of glycerol as a redox sink for anaerobic growth of *S. cerevisiae*, which can be fully replaced by a linear pathway for NADH-dependent reduction of acetate to ethanol. Here, they have the potential of metabolic engineering in the form of introducing an enzyme from *E. coli* into *S. cerevisiae*. Authors have expressed an NAD^+ dependent, acetylating acetaldehyde dehydrogenase (EC: 1.2.1.10) from *E. coli* into a gpd1 Δ gpd2 Δ strains of *S. cerevisiae*, and completed the linear pathway for acetic acid reduction. A detailed description on comparison of available metabolic pathway analysis methodologies can be found in [109], along with the emphasis on metabolic engineering.

The limitation of accessing kinetic information makes the use of the mathematical approach inconvenient in describing the process. ODEs are commonly used with known reaction rates, however, they deal with large computational costs and lacks robustness required to handle partial knowledge. On the other hand, discrete-event approaches for modeling and simulation can use all the available data, and can include analytic methods, if required. Although there is a good amount of progress in structural/topological analysis of metabolic pathway systems, dynamic cellular metabolic models still lack the realistic complexity. We can suggest hybrid approaches, like, coupling of FBA-derived and kinetic models approaches, flux distributions derived

through stoichiometric computations can simultaneously be used with available experimental data. Given the current state of art, we can say that hybrid modeling approaches should be improved to meet the need of integration of different kinds of data that experimental System Biology offers. Moreover, in our view, there is a specific requirement to fully automate the predictive network models, which can consider global changes in network structures and demonstrate the potential effects of system perturbations on the quality of model. This kind of automation will tremendously help experimental biologists to build and test novel hypotheses through lab experiments.

CONCLUSION

Analysis of biochemical pathways is one of the key topics in the post-genomic era that has a great potential for biotechnology and metabolic engineering. This leads to a better understanding of the cellular metabolism and to find possible targets for manipulation. In order to understand the cellular mechanisms from the predicted metabolic pathways, one has to develop and implement analysis methodologies, algorithms and tools. Well described mathematical pathway models can be served as a virtual laboratory which signify the system and give insights into basic principles of cellular functions. Advances in network topology theory and visualization tools might enable biologists to assemble data into network models that better present the kinetics of molecular interactions. Simulations based on these models should enhance and refine the experimental designs, thereby speeding up application of the Systems Biology cycle. In other words, a model should be able to raise new questions to give directions for next experimental work. Moreover, the integration of cellular metabolism and regulation often gives a complex picture. The availability of complete genome sequences for several microorganisms has provided an opportunity to develop genomic scale metabolic models. Since, it has been found that cellular metabolism is often altered in disease; therefore, metabolic analysis can be used for drug discovery. Metabolic engineering and related biotechnological applications benefit immensely from a systems view of the metabolism.

CONFLICT OF INTEREST

Authors have declared no conflict of interest.

ACKNOWLEDGEMENTS

Ms. Namrata Tomar, one of the authors, gratefully acknowledges CSIR, India for providing her a Senior Research Fellowship (9/93 (0145)/12, EMR-I).

REFERENCES

[1] Bower JM, Bolouri H. Computational modeling of genetic and biochemical networks. Massachusetts Inst Technol 2001.

[2] 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: 326-332.

[3] Varma A. Palsson B. Metabolic flux balancing: basic concepts, scientific and practical use. Biotechnology 1994; 12: 994-998.

[4] Beasley JE, Planes FJ. Recovering metabolic pathways *via* optimization. Bioinformatics 2007; 23: 92-98.

[5] Schuster S, Hilgetag S. On elementary flux modes in biochemical reaction systems at steady state. J Biol Syst 1994; 2: 165-182.

[6] Schwartz J, Kanehisa M. Quantitative elementary mode analysis of metabolic pathways: the example of yeast glycolysis. BMC Bioinformatics 2006; 7.

[7] Edwards J, Palsson B. Metabolic flux balance analysis and the *in silico* analysis of *Escherichia coli* K-12 gene deletions. BMC Bioinformatics 2000; 1.

[8] Papin J, Price N, Edwards J, Palsson B. The genome-scale metabolic extreme pathway structure in *Haemophilus influenzae* shows significant network redundancy. J Theor Biol 2002; 215: 67-82.

[9] Price N, Papin J, Palsson B. Determination of redundancy and systems properties of the metabolic network of *Helicobacter pylori* using genome-scale extreme pathway analysis. Genome Res 2002; 12: 760-769.

[10] Segre D, Vitkup D, Church G. Analysis of optimality in natural and perturbed metabolic networks. Proc Natl Acad Sci 2002; 99: 15112-15117.

[11] Shlomi T, Berkman O, Ruppin E. Regulatory on/off minimization of metabolic flux changes after genetic perturbations. Proc Natl Acad Sci 2005; 102: 7695-7700.

[12] Burgard A, Pharkya P, Maranas C. Optknock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. Biotechnol Bioengin 2003; 84: 647-657.

[13] Mahadevan R, Schilling C. The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. Metabolic Engin 2003; 5: 264-276.

[14] Carlson R, Fell D, Srienc F. Metabolic pathway analysis of a recombinant yeast for rational strain development. Biotechnol Bioengin 2002; 79: 121-134.

[15] Terzer M, Maynard ND, Covert MW, Stelling J. Genome-scale metabolic networks. WIREs Syst Biol Med 2009; 1: 285-297.

[16] Shimizu A.A review on metabolic pathway analysis with emphasis on isotope labeling approach. Biotechnol Bioprocess Engin2002; 7: 237-251.

[17] Klamt S, Stelling J, Ginkel M, Gilles ED. FluxAnalyzer: exploring structure, pathways, flux distributions in metabolic networks on interactive flux maps. Bioinformatics 2003; 22: 26-269.

[18] Aittokallio T, Schwikowski B. Graph-based methods for analysing networks in cell biology. Brief Bioinform 2006; 7: 243-255.

[19] Deville Y, Gilbert D, van Helden J, Wodak SJ. An overview of data models for the analysis of biochemical pathways. Brief Bioinform 2003; 4(3): 246-59.

[20] Copeland WB, Bartley BA, Chandran D, *et al.* Computational tools for metabolic engineering. Metabolic Engineering 2012; 14: 270-280.

[21] Lewis N, Nagarajan H, Palsson B. Constraining the metabolic genotype-phenotype relationship using a phylogeny of *in silico* methods. Nat Rev Microbiol 2012; 10: 291-305.

[22] Llaner F, Pico J. Stoichiometric modelling of cell metabolism. J Biosci Bioengin 2008; 105: 1-11.

[23] Milner P. The possible mechanisms of complex reactions involving consecutive steps. J Electrochem Soc 1964; 111: 228-232.

[24] Feinberg M, Horn FJM.Dynamics of open chemical systems and the algebraic structure of the underlying reaction network. Chem Engin Sci1974; 29: 775-787.

[25] Clarke BL. Complete set of steady states for the general stoichiometric dynamical system. J Chem Phys 1981; 75: 4970-4979.

[26] Clarke BL. Stoichiometric network analysis. Cell Biophys 1988; 237: 237-253.

[27] Jeong H, Tombor B, Albert R, Oltvai Z, Barabasi A. The large-scale organization of metabolic networks. Nature 2000; 406: 651-654.

[28] Helden JV, Wernisch L, Gilbert D, Wodal S.Graph-based analysis of metabolic networks. Ernst Schering Research Foundation Workshop, Springer Verlag, 2002;38: 245-274.

[29] Klamt S, Haus U.-U, Theis F. Hypergraphs and cellular networks. PloS Comput Biol 2009; 5: 1-6.

[30] Albert R. Network inference, analysis, modeling in systems biology. Plant Cell 2007; 19: 3327-3338.

[31] Pey J, Prada J, Beasley JE, Planes FJ.Path finding methods accounting for stoichiometry in metabolic networks. Genome Biol 2011; 12: R49.

[32] Reder C.Metabolic control theory: a structural approach. J Theor Biol 1988; 135: 175-201.
[33] Roels J.Energetics and Kinetics in Biotechnology. Elsevier Biomedical Press; Amsterdam; New York 1983.
[34] Papoutsakis E.Equations and calculations for fermentations of butyric acid bacteria. Biotechnol Bioengin 1984; 26: 174-187.
[35] Watson MR. Metabolic maps for the Apple II. Biochem Soc Trans 1984; 12: 1093-1094.
[36] Watson MR. A discrete model of bacterial metabolism. Comput Appl Biosci 1986; 2: 23-27.
[37] De RK, Das M, Mukhopadhyay S.Incorporation of enzyme concentrations into FBA and identification of optimal metabolic pathways. BMC Syst Biol 2008; 2.
[38] Das M, Mukhopadhyay S, De RK. Gradient descent optimization in gene regulatory pathways. PLoS ONE 2010; 5: e12475.
[39] Edwards J, Ibarra Palsson B. *In silico* predictions of *Escherichia coli* metabolic capabilities are consistent with experimental data. Nat Biotechnol 2001; 19: 125-130.
[40] Raghunathan A, Reed J, Shin S, Palsson B, Daefler S. Constraint-based analysis of metabolic capacity of *Salmonella typhimurium* during host-pathogen interaction. BMC Syst Biol 2009; 3.
[41] Hoppe A, Hoffmann S, Gerasch A, Gille C, Holzhtter H.-G. FASIMU: flexible software for flux-balance computation series in large metabolic networks. BMC Bioinformatics 2011; 12.
[42] Folger O, Jerby L, Frezza C, Gottlieb E, Ruppin E, Shlomi T.Predicting selective drug targets in cancer through metabolic networks. Mol Syst Biol 2011; 7: 1-10.
[43] Kelk SM, Olivier BG, Stougie L, Bruggeman FJ. Optimal flux spaces of genome-scale stoichiometric models are determined by a few subnetworks. Nature 2012; 2: 1-7.
[44] Schilling C, Letscher D, Palsson B. Theory for the systemic definition of metabolic pathways and their use in interpreting metabolic function from a pathway-oriented perspective. J Theor Biol 2000; 203: 229-248.
[45] Bell SL, Palsson BO.expa: a program for calculating extreme pathways in biochemical reaction networks. Bioinformatics 2004; 21: 1739-1740.
[46] Wiback SJ, Palsson BO. Extreme pathway analysis of human red blood cell metabolism. Biophys J 2002;83: 808-818.
[47] Wiback SJ, Mahadevana R, Palsson BO. Reconstructing metabolic flux vectors from extreme pathways: defining the α -spectrum. J Theor Biol 2003; 224: 313-324.
[48] 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.
[49] Trinh CT, Wlaschin A, Srienc F. Elementary mode analysis: a useful metabolic pathway analysis tool for characterizing cellular metabolism. Appl Microbiol Biotechnol 2009; 81: 813-826.
[50] Terzer M, Stelling J. Elementary flux modes- state-of-the-art implementation and scope of application. BMC Syst Biol 2007;1.
[51] Terzer M, Stelling J. Large-scale computation of elementary flux modes with bit pattern trees. Bioinformatics 2008; 24: 2229-2235.
[52] Schwartz J.-M, Kanehisa M. A quadratic programming approach for decomposing steady-state metabolic flux distributions onto elementary modes. Bioinformatics 2005; 21(Suppl. 2): ii204-ii205.
[53] de Figueiredo LF, Podhorski A, Rubio A, *et al.* Computing the shortest elementary flux modes in genome-scale metabolic networks. Bioinformatics 2009; 25: 3158-3165.
[54] Melendez-Hevia E, Wadell TG, Montero F.Optimization of metabolism: the evolution of metabolic pathways toward simplicity through the game of the pentose phosphate cycle. J Theor Biol 1994; 166: 201-220.
[55] Pfeiffer T, Bonhoeffer S. Evolution of cross-feeding in microbial populations. Am Nat 2004; 163, E126-E135.
[56] Rezola A, de Figueiredo L, Brock M, *et al.* Exploring metabolic pathways in genome-scale networks *via* generating flux modes. Bioinformatics 2011; 7: 534-540.
[57] Wagner C, Urbanczik R. The geometry of the flux cone of a metabolic network. Biophys J 2005; 89: 3837-3845.
[58] Ip K, Colijn C, Lun DS. Analysis of complex metabolic behavior through pathway decomposition. BMC Syst Biol 2011;5.
[59] Pfeiffer T, Sanchez-Valdenebro I, Nuno J, Montero F, Schuster S. METATOOL: for studying metabolic networks. Bioinformatics 1999; 15: 251-257.
[60] von Kamp A, Schuster S. Metatool 5.0: fast and flexible elementary modes analysis. Bioinformatics 2006; 22: 1930-1931.
[61] Mendes P. GEPASI: a software package for modelling the dynamics, steady states and control of biochemical and other systems. Comput Appl Biosci 1993; 9: 563-571.
[62] Hoops S, Sahle S, Gauges R, *et al.* COPASI-a COmplex PAthway SImulator. Bioinformatics 2006; 22: 3067-3074.
[63] Kleessen S, Nikoloski Z. Dynamic regulatory on/off minimization for biological systems under internal temporal perturbations. BMC Syst Biol 2012; 6.
[64] Luo RY, Liao S, Tao GY, *et al.* Dynamic analysis of optimality in myocardial energy metabolism under normal and ischemic conditions. Mol Syst Biol 2006; 2: 0031.
[65] Gudmundsson S, Thiele I. Computationally efficient flux variability analysis. BMC Bioinformatics 2010; 11.
[66] Burgard AP, Nikolaev EV, Schilling CH, Maranas CD. Flux coupling analysis of genome-scale metabolic network reconstructions. Genome Res 2004;14: 301-312.
[67] Larhlimi A, Bockmayr A. A new constraint-based description of the steady-state flux cone of metabolic networks. Discrete Appl Math 2009; 157: 2257-2266.
[68] Marashi S.-A, LaszloDavid Bockmayr A. On flux coupling analysis of metabolic subsystems. J Theor Biol 2012; 302: 62-69.
[69] Notebaart RA, Teusink B, Siezen RJ, Papp B. Co-regulation of metabolic genes is better explained by flux coupling than by network distance. Plos Comput Biol 2008; 4: 1-7.
[70] Montagud A, Zelezniak A, Navarro E, de Cordoba PF, Urchueguia JF, Patil KR. Flux coupling and transcriptional regulation within the metabolic network of the photosynthetic bacterium Synechocystis sp. pcc6803. Biotechnology J 2011; 6: 1-13.
[71] David L, Marashi S.-A, Larhlimi A, Mieth B, Bockmayr A. FFCA: a feasibility-based method for flux coupling analysis of metabolic networks. BMC Bioinformatics 2011;12.
[72] Provost A, Bastin G. Metabolic Flux Analysis: an approach for solving non-stationary undetermined systems. Proceedings 5th MATHMOD, Vienna, February, 2006; pp. 48-55.
[73] Wold S, Sjostroma M, Eriksson, L. PLS-regression: a basic tool of chemometrics. Chemometrics Intel Lab Syst 2001; 58: 109-130.
[74] Carinhas N, Bernal V, Teixeira AP, Carrondo MJ, Alves PM, Oliveira R. Hybrid metabolic flux analysis: combining stoichiometric and statistical constraints to model the formation of complex recombinant products. BMC Syst Biol 2011; 5.
[75] Bernal V, Carinhas N, Yokomizo A, Carrondo M, Alves P. Cell density effect in the baculovirus-insect cells system: a quantitative analysis of energetic metabolism. Biotechnol Bioengin 2009; 104: 162-180.
[76] Kaleta C, de Figueiredo LF, Schuster S. Can the whole be less than the sum of its parts? pathway analysis in genome-scale metabolic networks using elementary flux patterns. Genome Res 2009; 19: 1872-1883.
[77] Petri CA. Kommunikation mit Automaten. Ph.D. thesis, Institut fur Instrumentelle Mathematik, Bonn 1962.
[78] Chaouiya C. Petri net modelling of biological networks. Brief Bioinform 2007; 8: 210-219.
[79] Baldan P, Cocco N, Marin A, Simeoni M. Petri nets for modelling metabolic pathways: a survey. Nat Comput 2010; 9: 955-989.
[80] Barjis I, Gehlot V. Petri net based description and modeling of metabolic pathway. Proceeding SCSC Proceedings of the 2007 summer computer simulation conference Society for Computer Simulation International, 2007; San Diego, CA, USA.
[81] Oliveira JS, Bailey CG, Jones-Oliveira JB, Dixon DA, Gull DW, Chandler ML. A computational model for the identification of biochemical pathways in the krebs cycle. J Comput Biol 2003; 10: 57-82.
[82] Zevedei-Oancea I, Schuster S. Topological analysis of metabolic networks based on Petri net theory. *In Silico* Biol 2003; 3: 323-345.
[83] Koch I, Junker B, Heiner M. Application of Petri net theory for modelling and validation of the sucrose breakdown pathway in the potato tuber. Bioinformatics 2005; 21: 1219-1226.
[84] Hartmann Genrich RK, Voss K. Executable Petri net models for the analysis of metabolic pathways. Intl J Software Tools Technol Transfer 2001; 3: 394-404.
[85] Simao E, Remy E, Thieffry D, Chaouiya C. Qualitative modelling of regulated metabolic pathways: application to the tryptophan biosynthesis in *E. coli*. Bioinformatics 2005; 21: ii190-ii196.

[86] Wong P, Gladney S, Keasling J. Mathematical Model of the lac Operon: Inducer Exclusion, Catabolite Repression, Diauxic Growth on Glucose and Lactose. Biotechnol Prog 1997; 13: 132-143.

[87] Yang C.-R, Shapiro BE, Mjolsness ED, Hatfield GW. An enzyme mechanism language for the mathematical modeling of metabolic pathways. Bioinformatics 2005; 21: 774-780.

[88] Lincoln P, Tiwari A. Symbolic systems biology: hybrid modeling and analysis of biological networks. Alur R, Pappas G. eds., Proceedings of Hybrid Systems: Computation and Control, LNCS, Springer, 2004; 2993: pp. 660-672.

[89] Steuer R, Gross T, Selbig J, Blasius B. Structural kinetic modeling of metabolic networks. Proc Natl Acad Sci 2006; 103: 11868-11873.

[90] Steuer R, Junker B. Computational Models of Metabolism: Stability and Regulation in Metabolic Networks, John Wiley and Sons 2009; 142.

[91] Matsuno H, Tanaka Y, Aoshima H, Doi A, Matsui M, Miyano S. Biopathways representation and simulation on hybrid functional petri Net. *In Silico* Biol 2003; 3: 389-404.

[92] Chen M, Hofestadt R. Quantitative petri net model of gene regulated metabolic networks in the Cell. *In Silico* Biol 2003; 3: 347-365.

[93] Wiechert W. 13c metabolic flux analysis. Metab Engin 2001; 3: 195-206.

[94] Szyperski T. Biosynthetically directed fractional 13C-labeling of proteinogenic amino acids. An efficient analytical tool to investigate intermediary metabolism. Eur J Biochem 1995; 232: 433-448.

[95] Szyperski T. 13C-NMR, MS and metabolic flux balancing in biotechnology research. Quaterly Rev Biophys 1998; 31: 41-106.

[96] Queck L, Wittmann C, Nielsen L, Kromer J. OpenFLUX: efficient modelling software for 13C-based metabolic flux analysis. Microbial Cell Factories 2009; 8.

[97] Antoniewicz M, Kelleher J, Stephanopoulos G. Determination of confidence intervals of metabolic fluxes estimated from stable isotope measurements. Metab Engin 2006; 8: 324-337.

[98] Tang Y, Martin H, Myers S, Rodriguez S, Baidoo E, Keasling J. Advances in analysis of microbial metabolic fluxes *via* 13C isotopic labeling. Mass Spectrometry Rev 2009; 362-375.

[99] Antoniewicz MR, Kelleher JK, Stephanopoulos G. Elementary metabolite units EMU;: a novel framework for modeling isotopic distributions. Metab Engin 2007; 9: 68-86.

[100] Wiechert W, de Graaf A. Bidirectional reaction steps in metabolic networks I. Modeling and simulation of carbon isotope labeling experiments. Biotechnol Bioeng 1997; 55: 101-117.

[101] Palsson B. The challenges of *in silico* biology. Nat Biotechnol 2000; 18: 1147-1150.

[102] Covert MW, Knight EM, Reed JL, Herrgard MJ, Palsson BO. Integrating high-throughput and computational data elucidates bacterial networks. Nature 2004; 429: 92-96.

[103] Henry CS, DeJongh M, Best AA, Frybarger PM, Linsay B, Stevens RL. High-throughput generation, optimization and analysis of genome-scale metabolic models. Nat Biotechnol 2010; 28: 977-982.

[104] Jerby L, Ruppin E. Predicting drug targets and biomarkers of cancer *via* genome-scale metabolic modeling. Clin Cancer Res 2012; 18: 5572-5584.

[105] Vitkin E, Shlomi T. MIRAGE: a functional genomics-based approach for metabolic network model reconstruction and its application to cyanobacteria networks. Genome Biol 2012; 13.

[106] Yizhak K, Benyamini T, Liebermeister W, Ruppin E, Shlomi T. Integrating quantitative proteomics and metabolomics with a genome-scale metabolic network model. Bioinformatics 2010; 26: i255-i260.

[107] Tenazinha N, Vinga S. A survey on methods for modeling and analyzing integrated biological networks. IEEE/ACM Trans Comput Biol Bioinform 2011; 8: 943-958.

[108] Medina VG, Almering MJH, van Maris AJA, Pronk JT. Elimination of glycerol production in anaerobic cultures of a *Saccharomyces cerevisiae* strain engineered to use acetic acid as an electron acceptor. Appl Environ Microbiol 2010; 76: 190-195.

[109] Tomar N, De RK, Comparing methods for metabolic network analysis and an application to metabolic engineering. Gene 2013; 521(1): 1-14. doi: 10.1016/j.gene.2013.03.017.

Received: April 24, 2013 Revised: June 4, 2013 Accepted: June 5, 2013