Biotechnology and Bioprocess Engineering 19: 568-585 (2014)
DOI 10.1007/s12257-014-0172-8

REVIEW PAPER

Database and Tools for Metabolic Network Analysis

Lu Shi Jing, Farah Fathiah Muzaffar Shah, Mohd Saberi Mohamad, Nur Laily Hamran, Abdul Hakim Mohamed Salleh, Safaai Deris, and Hany Alashwal

Received: 18 March 2014 / Revised: 15 May 2014 / Accepted: 23 May 2014

Abstract Metabolic network analysis has attracted much attention in the area of systems biology. It has a profound role in understanding the key features of organism metabolic networks and has been successfully applied in several fields of systems biology, including *in silico* gene knockouts, production yield improvement using engineered microbial strains, drug target identification, and phenotype prediction. A variety of metabolic network databases and tools have been developed in order to assist research in these fields. Databases that comprise biochemical data are normally integrated with the use of metabolic network analysis tools in order to give a more comprehensive result. This paper reviews and compares eight databases as well as twenty one recent tools. The aim of this review is to study the different types of tools in terms of the features and usability, as well as the databases in terms of the scope and data provided. These tools can be categorised into three main types: standalone tools; toolbox-based tools; and web-based tools. Furthermore, comparisons of the databases as well as the tools are also provided to help software developers and users gain a clearer insight and a better understanding of metabolic network analysis. Additionally, this review also helps to provide useful information that can be used as guidance in choosing tools and databases for a particular research interest.

Lu Shi Jing, Farah Fathiah Muzaffar Shah, Mohd Saberi Mohamad*, Nur Laily Hamran, Abdul Hakim Mohamed Salleh, Safaai Deris
Artificial Intelligence and Bioinformatics Research Group, Faculty of Computing, Universiti Teknologi Malaysia, Skudai 81310, Malaysia
Tel: +60-7-553-3153; Fax: +60-7-556-5044
E-mail: saberi@utm.my

Hany Alashwal
College of Pharmacy, The University of Rhode Island, 7 Greenhouse Rd South Kingstown, RI 0288, USA

Keywords: metabolic network analysis tools, metabolic network reconstruction, Kyoto Encyclopedia of Genes and Genomes (KEGG), flux balance analysis (FBA), OptFlux, MetaFluxNet

1. Introduction

Metabolic network analysis has been growing remarkably ever since the field of systems biology gained prominence in the mid-1990s. The in-depth analysis, successful identification of complete sets of chemical reactions, and increasingly available complete genome sequences have driven the development of the metabolic network. Generally, the metabolic network can be defined as a large system of chemical reactions with two major roles [1]. First, it is responsible for converting energy sources in the environment into new energy forms useful to an organism; second, it is crucial in synthesis of amino acids that are essential for the growth of cells from nutrients or other chemical sources [1].

Analysing the metabolic network provides a better understanding of the systems biology of cellular metabolism within an organism. Cellular metabolism, by definition, is chemical and physical processes that are necessary for the maintenance of life [1]. Metabolic network analysis can be used to promote the processes of fermentation engineering, drug target identification, and microbial engineering [2]. In addition, metabolic network analysis has also been proven to be effective in many applications, for example, metabolic engineering [3-7], drug target identification [8-10], gene deletion predictions [11], and cellular regulatory network elucidation [12]. Recently, there has been a shift from a reductionist to a holistic approach for revealing the complexity of living systems and characterising the functional behaviour of a biological system. This is guided by automated

sequencing and high-throughput experimental techniques that focus on the systemic properties of metabolic networks.

Metabolic network reconstructions have become popular since early 90's. Many research on metabolic network reconstruction and modelling has been carried out on the available genome information. This reconstruction is usually based on the existing genome sequence information and molecular physiology (biochemical information). A list of genome scale models that have been reconstructed and experimentally validated is available at http://gcrg.ucsd.edu/InSilicoOrganisms/OtherOrganisms. In nature, the understanding of interactions between genome information and molecular physiology of an organism is a challenging task. The large amount of high-throughput data is hard to interpret and analyse. With the advancement of bioinformatics tools, many biological metabolic networks of different organisms can be reconstructed and analysed. Moreover, relevant information can be obtained by searching the databases available on the internet. Therefore, useful and intuitive tools will likely be developed to accomplish the task.

Several steps are involved in performing a metabolic network reconstruction [9]. Collections of resources and tools are needed to assist the metabolic network reconstructions: (1) drafting a model; (2) reconstruction of a detailed model; (3) mathematical representation of the model; (4) filling of gaps; and (5) simulation and visualisation. The first and second steps represent processes of identifying functional annotation of genes that can be accomplished by extracting and matching the biochemical reaction data from an existing database such as the Kyoto Encyclopedia of Genes and Genomes (KEGG) [12]. The functional annotations are then translated into appropriate biochemical reactions in order to build a metabolic model. Meanwhile, steps 3 ~ 5 can be achieved by using metabolic network simulation tools.

In this paper, we review the databases available for extracting biochemical reaction data and tools that can be used to aid metabolic network analysis. Comparisons among databases and software tools are also included to give a better insight and understanding on different approaches and data resources available for metabolic network analysis. In addition, a few examples of research conducted using the tools and databases discussed here are provided to show their capability in various aspects of metabolic network analysis. The reviewed databases are KEGG, BioCyc, MetaCyc, ENZYME, BRaunschweigENzymeDAtabase (BRENDA), Biochemical Genetic and Genomic (BiGG), Reactome, and KaPPA-View4. Besides being among the most commonly used databases, these eight were selected for the fact that they are great resources of biological pathway information with cross-references to plentiful external databases, and we only focused on these databases to give a more detailed comparison. Each of these databases has their own particular aim, scope, and range of data available that differ from each other. The comparisons are made to help users to choose the most suitable databases for the research being conducted and the types of data required.

2. Databases

The study of metabolic networks has been increasing day by day in the field of systems biology and has become a dynamic complex interactive non-linear system [13]. Thus, it is not surprising that numerous metabolic databases are available. Important information obtained from these databases can help in metabolic network reconstruction of the organism. Now, semi-automatic assembly of metabolic network reconstructions is crucial due to concerns of time and effort. For instance, resources such as PathoLogic or ERGO (tools for integrating data), in combination with MetaCyc, can be utilised to produce initial fast reconstructions automatically and the reconstructions can then be manually updated by utilising resources such as PathwayTools (tools for integrating data). These databases are important for further analysis and they can be an easy-to-use research platform. These existing tools are publicly available to be used by researchers, especially in the field of metabolic network reconstruction and analysis and other related fields.

2.1. KEGG

KEGG was founded by Ogata *et al*. [14] in 1999. It is one of the first pathway databases that were initiated to move from the existing gene catalogues to pathways catalogues. It is a bioinformatics database that contains information such as proteins, genes, pathways, and reactions. It is also a bioinformatics resource for linking genomes to life and the environment. In the KEGG organism area, it has been divided into two parts, eukaryotes and prokaryotes, which comprise much data such as gene and DNA information that can be simply searched by typing in the enzyme of choice.

KEGG is commonly applied in the area of *in silico* modelling of metabolic networks. In this study, the KEGG converter was applied as a web-based application, as suggested by Moutselos *et al*. [15]. It was utilised as a source of KEGG Markup Language (KGML) files for the purposes of constructing integrated pathway System Biology Markup Language (SBML) models that were fully functional for the simulation process. In the study by Stobbe *et al.* [16], KEGG was used to extract information of the human metabolic network. A number of groups have constructed high-quality human (metabolic) pathway

databases in almost the past 15 years that can be used in this effort.

Liu *et al.* [17] applied KEGG in the development of a pathway comparison tool in the analysis of bacteria genomes. The tool was based on the integration of the KEGG Orthology (KO), KEGG pathway database, and the Molecular Interaction (MINT) database. Using the proposed method, we developed a comparison tool for a pathway that is able to identify dissimilarities in term of proteins that are present or absent in a pathway [17]. This process can be used to identify various organisms by calculating the match similarity index, based on shared gene position, KO id, and protein-protein interactions. As an example of use of this tool, it was applied to compare a group of bacteria including *Bacillus cereus.*

KEGG is also applied in a web-based platform like MicrobesFlux for generating and reconstructing metabolic models for annotated microorganisms. MicrobesFlux is able to automatically download the metabolic network (including enzymatic reactions and metabolites) of ~1,200 species from the KEGG database and then convert it to a metabolic model draft. The platform also provides diverse customised tools such as gene knockouts and introduces heterogonous pathways for users to reconstruct the model network [6].

2.2. BioCyc

BioCyc was developed by Karp *et al*. [18]. It is a group of 161 Pathway or Genome DataBases (PGDBs) that presents information on genomes and cellular networks in an ordered manner, as well as allows powerful computational analysis and exploitation of the database. EcoCyc and MetaCyc are the databases that provide highly curated Tier 1 PGDBs at the core of BioCyc. They consist of many experimentally elucidated metabolic pathways originating from *Escherichia coli* and other kinds of organisms. BioCyc can be edited and viewed by using Pathway Tools. It is a kind of environment software built to display, query, and edit the information regarding each pathway. Furthermore, its component compounds, reactions, protein complexes, enzymes, operons, genes, and regulation are at the transcriptional and substrate levels.

Besides, BioCyc was applied in the area of querying and computing in the study by Krummenacker *et al*. [19]. In their study, they mentioned many ways to access and query the difficult and integrated cellular data in the family of the BioCyc database, such as accessing through several file formats, accessing through the Application Program Interfaces (APIs) for Locator or Identifier Separation Protocol (LISP), Java, and Perl, and also Structured Query Language (SQL), which is accessed through the relational database BioWarehouse.

BioCyc was also applied in a study by Choi *et al.* [20], which combined a data warehouse concept with web services for the establishment of the *Pseudomonas* systems biology database, which was SYSTOMONAS. Systems biology requires the incorporation of data from diverse sources and their united analysis by utilising dissimilar bioinformatics tools. The incorporation of dissimilar biological databases is often a problem because of their semantic and structural diversity. Furthermore, for an integration process, continuous updates of both the structure and content of a database present further challenges. Thus, in this field, a SYSTOMONAS database was established for SYSTems biology of pseudOMONAS by incorporating heterogeneous data from different external resources including BioCyc.

2.3. MetaCyc

MetaCyc is a well-known metabolic pathway database that contains information on almost 158 organism enzymes and pathways, developed by Karp *et al.* [21]. It includes highly curated databases such as those for *E. coli* and *Musmusculus* [6]. Besides, the MetaCyc database is a reference source that can be accessed online for metabolic data. This database covers the metabolic pathways, enzymes, substrates, reactions, and products of compounds. As a review-level database, each entry in MetaCyc is also combined with the knowledge from various literature sources, enabling the integration of data.

As a reference source in metabolic pathways, MetaCyc gives a number of goals in metabolic engineering. It provides an encyclopaedic listing of enzymes and their properties in order to allow a metabolic engineer to see which enzyme characteristics can contribute to solving an engineering problem within seconds [7]. Furthermore, it also acts as a reference for the database pathway to predict the metabolic pathway match of an organism from its genome. The MetaCyc database itself can be browsed through http://MetaCyc.org/. It is also accessible through a Sun workstation and local setting for a PC.

The viewpoint of MetaCyc is that the pathways reported from experimental literature are coded and labelled to their corresponding organisms which they are found to occured in. Thus, the method in which each pathway is recognised and tagged with the organism is actually based on experiments reported in the literature and provide cross organism references. For instance, MetaCyc was applied as multi-organism database of metabolic pathways and enzymes in a study conducted by Krieger *et al.* [22]. In this study, MetaCyc was used to compile a representative sample for each of the experimentally elucidated pathways. The universe of metabolism was used to produce a catalogue. Also, the MetaCyc database contains information about the

reactions, chemical compounds, and genes, as well as much information on the pathways and enzymes in MetaCyc, such as comments and literature citations.

Furthermore, MetaCyc was applied in the Metabolic Pathway Databases for Plant Research in research conducted by Zhang *et al*. [23]. The content of MetaCyc on biochemical pathways was used as a source for metabolic enzymes and pathways properties. Regarding the software Pathway Tools, the metabolic pathways in the MetaCyc database can be used to predict the complement of an annotated genome computationally. Thus, there are 60 plant-specific pathways, in order to enlarge the size of pathways and enzymes, and these have been added or reorganised in MetaCyc. Besides, AraCyc is a kind of specific database that contains enzymes and pathways and originates from the plant model of *Arabidopsis* (*Arabidopsisthaliana*). AraCyc was the first predicted plant that was used computationally in a metabolism database, and can be accessed at http://arabidopsis.org/tools/aracyc/ resulting from MetaCyc.

Also, MetaCyc was applied to enzymes and metabolic pathways from all domains of life in a study by Caspi *et al.* [24]. The primary scientific literature was curated from the pathways in MetaCyc which are experimentally built by metabolic pathways of small molecules. The MetaCyc pathway was annotated for one or more well-characterised enzymes in each reaction and it contains only uniquely high-quality resource for metabolic pathways and enzymes.

2.4. ENZYME

ENZYME is a nomenclature database developed by Bairoch [25]. It is a repository of information related to the nomenclature of enzymes. It has become an indispensable resource for the development of metabolic databases. It has also been integrated into the ExPASY proteomics server of the Swiss Institute of Bioinformatics. After an enzyme of interest in the database is searched, the associated reactions from a given sources will be displayed. This database has a direct link to other genome, enzyme, and literature databases such as KEGG, PubMed, and BRENDA.

The ENZYME database was applied in the prediction of interaction sites in enzymes in a study by Sheu *et al.* [26]. One of the most crucial functions of proteins is to serve as enzymes that catalyse biochemical reactions. The nomenclature of enzymes can be found in the List of Enzyme and the ENZYME database from the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (NC-IUBMB). This database provides information on the functions of enzymes. It has a wide collection of facts related to enzymes such as the specificity of reactions, functional parameters, substrates, products, and inhibitors. Furthermore, IntEnz is a relational database incorporating enzyme data from all three of these sources.

The ENZYME database was applied in a combination data warehouse concept, together with web services for the *Pseudomonas* systems biology database establishment of the SYSTOMONAS in a research by Krummenacker *et al.* [19]. It was done by incorporating various data from highly dissimilar external resources, including BioCyc, BRENDA, PRODORIC Pseudomonas Genome Database v2, KEGG, and ENZYME. Thus, together with the purpose and benefits given by the web services, the data warehouse concept was integrated. Because of the fast performance, this hybrid approach has the benefit of data consistency provided by the data warehouse system.

The ENZYME database was also implemented in the study by Andreini *et al.* [27], in which it was utilised in metal ions in biological catalysis. It is well known that some metal ions are important to life. In living systems, a major determinant of functional relevance is that a substantial fraction of enzymes require metals for their catalytic activity. Large and diverse metal-dependent enzymes that catalyse fundamental biological processes such as respiration, photosynthesis, and nitrogen fixation can be found in nature.

The ENZYME database was also used by Sharma *et al*. [28] to discover commercially useful enzymes in metagenomic datasets. For this analysis, the Enzyme Commission was entirely used as a reference for enzymes and function definition. The complete set of information on 4877 enzymes annotated with Enzyme Commission numbers was obtained from the nomenclature of the ENZYME database, which can be accessed at ExPASy. The Swiss-Prot sequences were obtained from the Swiss-Prot database.

2.5. BRENDA

The development of BRENDA began in 1987 at the German National Research Center for Biotechnology (GBF) and was continued at the Cologne University Bioinformatics Centre. BRENDA was initially published as books by Schomburg and Schomburg [29]. It is a protein function database, as stated by Schomburg *et al.* [30], which contains a huge amount of enzymatic and metabolic data and is updated and evaluated by extracting information from primary literature. It keeps information on functional and molecular properties of enzymes, including data on occurrence, kinetics, catalysed reaction, products and substrates, cofactors, inhibitors, activators, stability, and structure.

BRENDA was further updated by Schomburg *et al.* [31], including major new developments. It includes biochemical and molecular information on reaction and specificity, classification and nomenclature, occurrence, functional parameters, enzyme structure, engineering, application, disease, stability, preparation and isolation, links, and literature references. It provides various search modes for

Table 1. Data and Information Fields of BRENDA [29]

Information field	Total entries	Information field	Total entries
Enzyme nomenclature		**Functional parameters**	
EC number	3869	K_m value	28 134
Recommended name	3509	Turnover number	3986
Systematic name	3182	Specific activity	11 787
Synonyms	17 707	pH optimum	14 037
CAS registry number	3552	pH range	3929
Reaction	3518	Temperature optimum	6147
Reaction type	4123	Temperature range	908
Enzyme structure		**Molecular properties**	
Molecular weight	12 329	pH stability	2931
Subunits	7416	Temperature stability	6825
Sequence links	33 099	General stability	5398
Posttranslational modification	1112	Organic-solvent stability	311
Crystallization	1003	Oxidation stability	349
3D-structure, specific PDB links	6142	Storage stability	6505
		Purification	11 176
Enzyme-ligand interactions		Cloned	2015
Substrates-products	47 630	Engineering	797
Natural substrate	7668	Renatured	199
Cofactor	6217	Application	338
Activating compound	6217		
Metals-ions	13 173	**Organism-related information**	
Inhibitors	56 336	Organism	40 027
		Source tissue, organ	19347
Bibliographical data		Localization	7935
References	46 305		

overall or organism-specific queries and now includes a tool for substructure searches of ligands. There are additional new features that can be searched within TaxTree.

The BRENDA enzyme information system is a manually annotated repository for enzyme data. The BRENDA enzyme information system that can be browsed at http://www.brenda.unikoeln.de is the largest publicly available enzyme information system worldwide. The major part of its content was manually extracted from primary literature. It is not restricted to specific groups of enzymes, and includes information on all identified enzymes irrespective of the source of the enzyme [32]. Table 1 shows the data and information fields of BRENDA [30].

2.6. BiGG

BiGG is a knowledge base of biochemically, genomically, and genetically structured genome-scale and metabolic network reconstructions developed by Schellenberger *et al.* [33]. It is used for large-scale metabolic reconstructions. BiGG combines some of the published genome-scale metabolic networks, all into one resource that allows components to be compared across various organisms by using standard nomenclature. It can also be used to access the model content, visualise the metabolic pathway maps, and also export SBML files through external software packages of the models for advanced investigation. Users can go from BiGG to various external databases to obtain additional knowledge regarding genes, proteins, reactions, metabolites, and citations of interest. This database meets the need in the community of systems biology to access high-quality metabolic models and reconstructions. It can be accessed freely for academic purposes at http://BiGG.ucsd.edu.

BiGG was applied in the achievement of an infrastructure called Recognition Infrastructure (ReIn) [34]. This infrastructure can integrate a huge amount of two-dimensional (2D) and even three-dimensional (3D) object recognition to evaluate pose techniques in parallel as loadable plugins that are dynamic. It also provides an efficient design of data passing and offers the possibility to modify the parameters and initial settings of the techniques during the implementation. There are two new classifiers highlighted for robots need which is Binarized Gradient Grid Pyramids (BiGGPy) that is crucial for 2D classification and Viewpoint Feature Histograms (VFH) that are vital for 3D classification and pose. These two classifiers can be easily combined

using ReIn to solve object recognition and identification of pose problems.

2.7. Reactome

Reactome (http://www.reactome.org) is an open access, open source, and manually curated database of human reactions and pathways developed by collaborated groups between New York University School of Medicine, Ontario Institute for Cancer Research, The European Bioinformatics Institute, and Cold Spring Harbor Laboratory. The ultimate goal of the Reactome project is to systematically associate molecular and cellular functions of human proteins to develop human biological processes, pathways, and reactions so that they can be used as a systems biology platform and as an online encyclopaedia for data analysis and mining. As of June 2011, this database is supported by 8492 publications and contains 6248 human proteins organised into 1153 pathways and 4354 reactions [35].

The datasets of Reactome are highly reliable for the analysis of pathway databases but is limited by a small coverage of proteins. Thus, molecular interaction data and network information are integrated into Reactome pathway diagrams, causing the proteins to be displayed as interacting and overlaid with manually annotated protein components. Furthermore, the feature of 'Analyze, Annotate and Upload' on the Pathway Browser also helps to overlay all the pathway proteins' interactors [36].

Reactome can be used to annotate anti-cancer therapeutics and cancer variants. In a study conducted by Milacic *et al.* [37], new data classes were added to the Reactome database, allowing the capture of pathways that promote cell growth and cellular differentiation, as well as pathways that inhibit cell division. This study has successfully annotated the recombinant antibody cetuximab and small tyrosine kinase inhibitors that help in the inhibition of Epidermal Growth Factor Receptor (EGFR) kinase activity in cancer [38]. Small-molecule recombinant therapeutics and antibodies are believed to have potential for cancer treatments impelled by increased activity of Fibroblast Growth Factor Receptor (FGFR), EGFR, or AKT.

Aside from cancer, the Reactome database is also applicable to various disease pathways. "Influenza Infection", "HIV Infection", "Latent Infection with Mycobacterium Tuberculosis", and "Botulinum Neurotoxicity" are a few examples of diseases currently featured by Reactome. Metabolic genetic diseases like "Mucopolysaccharidoses" and "Abnormal Metabolism in Phenylketonuria" are also published with the help of Reactome [37].

2.8. KaPPA-View4

KaPPA-View4 is one of the most recent metabolic pathway databases that is believed to be the only database that can overview all correlations on pathway maps, and was developed by Sakurai *et al.* [39]. It is a unique and a very good database system as it enables the characterisation of all correlations on metabolic pathway maps. Its ability to overlay metabolite-to-metabolite and gene-to-gene correlations as curves on either a single metabolic pathway map or on a combined map of a maximum of four has contributed to the discovery of genes with unknown function and regulatory mechanisms of metabolic pathways, as well as functional diversities of family members of a gene.

Another significant improvement made to this database compared with its previous version is that it can instantly display results. The viewer functions of this 'omics' database were extended to receive data uploaded from external systems. Two versions of the system are provided, KaPPA-View4 Classic and KaPPA-View4 KEGG. KaPPA-View4 Classic (http://kpv.kazusa.or.jp/kpv4/) is highly recommended for plant scientists as its pathway maps are primarily based on *Arabidopsis*. The initial version of this database was based on approximately 150 leaves of *Arabidopsis* pathway maps [40]. Information for genome-sequenced plant species and for plant species with available AffymetrixGeneChips and correlated data are implemented, including *Arabidopsis*, *Lotus japonicus*, and rice, as well as wheat, tomato, barley, maize, soybean, and wine grape. With AffymetrixGeneChips probes, the best hits for each target sequence were determined as those with the minimum e-values among candidates, with the threshold of 1×10^{-30}.

Meanwhile, KaPPA-View4 KEGG (http://kpv.kazusa.or.jp/kpv4-kegg/) is primarily for general users and also plant scientists who require information on microorganisms, animals, and plants. For the time being, information on only 15 species is provided, namely rat, mouse, *Drosophila melanogaster*, human, castor bean, maize, sorghum, rice, wine grape, *Arabidopsis*, poplar, budding yeast, *E. coli*, *Physcomitrella patens* subsp. *patens*, and *Caenorhab ditiselegans*. One of the significant advantages and improvements made on this version is the ability to share the latest results of gene curation of KEGG in terms of their categorisations, assignments, and descriptions on maps [39].

KaPPA-View4 is implemented in the omics studies of *Jatropha curcas*, a type of shrub that has significant roles in animal nutrition [41], biofuels [42], and medical applications [43]. This database helps to visualise the co-expression data between pathway maps, allowing the discovery of key genes that control the production of novel end products in *Jatropha* [44]. This research is believed to be the first to deliver a pathway analysis environment for *Jatropha* omics data, and no Jatropha gene expression data have been deposited in Gene Expression Omnibus (GEO) applications [45].

Table 2. Comparison of the scope of each database

Database	Scope				
	Enzymes	Genes	Reactions	Pathways	Biochemistry
KEGG	√	√	√	√	
BioCyc				√	
MetaCyc	√			√	
ENZYME	√		√		
BRENDA	√				
BiGG		√		√	√
Reactome	√	√		√	√
KaPPA-View4		√		√	√

Table 3. Comparison of metabolic network databases

Database	Website	Advantage	Disadvantage
KEGG	www.genome.ad.jp/kegg	• Well-known database • Organised in a hierarchy • Has a list of references • Publicly available	• No summaries • Data download is complicated
BioCyc	biocyc.org/	• Has visualisation services • Can run as a desktop application	• Not readily accessible from the SYSTOMONAS database
MetaCyc	www.metacyc.org/	• Available free of charge • Free of charge to non-profit organisations • Non-redundant reference database	• MetaCyc pathways are typically smaller
ENZYME	enzyme.expasy.org/	• Can help in the development of computer programs • Useful to those who work with enzymes	• Not frequently updated
BRENDA	www.brenda-enzymes.info/	• Not limited to a specific aspect of the enzyme • Covers organism-specific information on functional and molecular properties • Data are continuously updated	• Not frequently updated
BiGG	BiGG.ucsd.edu/	• Capable of browsing and exporting whole reconstructions	• Needs maintenance and support
Reactome	www.reactome.org	• Provide detailed description of components • Heavily curated with links to more specific information from other sites	• Does not give a complete graphical view of the pathway • Not easy to export data
KaPPA-View4	kpv.kazusa.or.jp/kpv4/	• Able to share gene descriptions, categorisations, and assignments with KEGG • Able to view all the correlations on pathway maps	• Not ready for integrative omics

2.9. Comparison of databases

Eight different databases were applied in this study. Table 2 shows the comparison of the scope of each database. Table 3 illustrates the advantages and disadvantages of all eight databases and Table 4 shows a brief description on the usage of each database by different scientists.

3. Tools

The tools used to perform metabolic network analysis had been grouped into three classes, which were standalone tools, toolbox-based tools, and web-based tools. From the analysis, it can be seen that MetaFluxNet, BioSPICE (Biological Simulation Program for Intra-and-Inter-Cellular Evaluation), FluxExplorer, Systems Biology Research Tool (SBRT), Pathway Analyser, OptFlux, BioOpt, SurreyFBA, and Flux-balance Analysis based SIMUlations (FASIMU) were categorised into the standalone category. Toolbox-based tools included FluxAnalyzer/CellNetAnalyzer, SNA, OpenFlux, FBA SimVis, and COstraint-based Reconstruction and Analysis Toolbox (COBRA), while web-based tools included GSMN-TB, CycSim, WEbcoli, Model SEED, Acorn, Flux Analysis and Modeling Environment (FAME), and MicrobesFlux.

Compared with the review done by Lakshamanan *et al.*

Table 4. Examples of research that incorporates the discussed databases

Database	Usage/ application	Reference
KEGG	To create new genome scale model for *Clostridiumacetobutylicum* ATCC 824, named iCAC794 model, to allow *in silico* metabolic engineering of clostridial metabolism	[46]
	To assist in the curation process to verify directionality and stoichiometry of all reactions in *Arabidopsis thaliana* for improvement of terpenoid production.	[47]
BioCyc	Biochemical reactions for reconstruction and analysis of *Klebsiella oxytoca* to produce 2,3-butanediol.	[48]
MetaCyc	To remodel iBB814 to simulate *Scheffersomycesstipitis* metabolism for optimisation of ethanol production from xylose or glucose mixtures.	[49]
	As a reference database for PathoLogic component of Pathway Tools software.	[50]
	To predict organism metabolic network.	[51]
ENZYME	Information on enzyme function as it has a wide collection of facts related to enzymes including specificity of reaction, functional parameters, substrates, products, and inhibitors.	[52]
	Implemented in the study of effects of metal ions on enzyme catalysis.	[53]
	To discover commercially useful enzymes in metagenomic datasets.	[54]
BRENDA	To identify non-native production routes with the help of a graph-based algorithm for the production of microbial 1-butanol.	[55]
BiGG	Reconstruction of the human metabolic network at the genome scale, *Homo sapiens Recon 1*	[56]
	To fill in knowledge gaps in orphan reactions by listing each reaction and exhibiting other reconstructions that use the same reactions.	[57]
Reactome	To annotate anti-cancer therapeutics and cancer variants.	[37]
KaPPA-View4	To visualise the co-expression data between pathway maps.	[41]

Table 5. Reference for reliability and significant of the studied tools

Name	Link	Reference
MetaFluxNet	http://www.biomedcentral.com/1752-0509/2/55	[83]
BioSPICE	http://bvsalud.org/portal/resource/en/mdl-14683613	[60]
FluxEXPLORER	http://www.fluxexplorer.com/about.html	-
Pathway Analyser	http://bib.oxfordjournals.org/content/10/4/435.full	[84]
OptFlux	http://www.biomedcentral.com/1752-0509/4/45	[64]
BioOpt	http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3361690/	[85]
SurreyFBA	http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3534579/	[86]
FASIMU	http://mitpress.mit.edu/sites/default/files/titles/content/ecal13/978-0-262-31709-2-ch175.pdf	[87]
FluxAnalyzer/ CellNetAnalyzer	http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0071909	[88]
SNA	http://udini.proquest.com/view/unraveling-the-regulation-of-mint-goid:304441015/	[89]
OpenFlux	http://www.ncbi.nlm.nih.gov/pubmed/19409084	[72]
FBA-SimVis	http://bioinformatics.oxfordjournals.org/content/25/20/2755.full	[73]
COBRA	http://nar.oxfordjournals.org/content/25/12/2532.full	[90]
GSMN-TB	http://genomebiology.com/2007/8/5/r89	[9]
CycSim	http://bioinformatics.oxfordjournals.org/content/25/15/1987.full	[76]
WEbcoli	http://bioinformatics.oxfordjournals.org/content/25/21/2850.full	[77]
Model SEED	http://www.theseed.org/wiki/Home_of_the_SEED	-
Acorn	http://www.biomedcentral.com/1471-2105/12/196	[79]
FAME	http://www.biomedcentral.com/1752-0509/6/8	[80]
MicrobesFlux	http://www.biomedcentral.com/1752-0509/6/94	[91]

[58], we provide a more detailed discussion of each tool by giving a brief description of each and every one of them. This is to provide a little background on the tools and some other related information. In addition, we also introduce and discuss four new tools, namely BioSPICE, Pathway Analyzer, FluxExplorer, and Openflux, a more recent version of Metaflux, which is not discussed in this review. The authors have been discuss the reliability and the

significant of the mentioned tools in their papers, and are listed in Table 5.

3.1. Standalone tools

Standalone tools are open source software that are independently installed on computers and executed locally. These locally executed tools enable users to perform their task independently, at any time and any place. A total of nine standalone tools are discussed in this paper, which are MetaFluxNet [59], BioSPICE [60], FluxExplorer [61], SBRT [62], Pathway Analyzer [63], OptFlux [64], BioOpt [65], SurreyFBA [66], and FASIMU [67]. A brief description of each toolis provided accordingly.

3.1.1. MetaFluxNet

MetaFluxNet (http://mbel.kaist.ac.kr/) is a metabolic flux analysis (MFA) tool developed by Lee *et al.* [59] that is used to manage metabolic information and perform metabolic fluxes analysing using customised approaches. It allows users to interpret and examine cellular metabolic networks in response to gene modification and different environments. Moreover, users can gain a better understanding of metabolic status and metabolic engineering strategy design by using quantified *in silico* simulation of metabolic pathways. The tool provides a well-developed model construction environment, a user friendly interface for metabolic flux analysis, and an automated pathway layout. The main metabolic flux analysis approach applied in this tool is MFA. This comparative MFA allows the generation of automated pathway layouts. In this study, the usefulness of the tool was demonstrated by applying it to *E. coli* metabolic pathways.

3.1.2. BioSPICE

BioSPICE (http://biospice.sourceforge.net/) is an open source tool designed to help in the study of systems biology, intended to assist researchers in the modelling and simulation of spatiotemporal processes in living cells, and was developed by Price *et al*. [13] and Segre *et al*. [68]. Garvey *et al*. [60] claimed that different investigators could provide models, data sources, simulation engines, and output in the current dashboard. Besides, BioSPICE also gives a parallel working plane that allows it to be run on different machines across a heterogeneous and distributed network. The software uses a combination of Open Agent Architecture (OAA) and Netbeans to build a graphical environment. The OAA server allows servers that are registered with the same agents to communicate using different computing machines. In addition, the graphical user interface was developed using Java programming language, which allows users to select different simulation parameters. Perl programming is used to read the output. Besides, BioSPICE provides a graphical user interface in which a tabular format is implemented for data visualisation.

3.1.3. FluxExplorer

FluxExplorer (http://www.scbit.org/eng/mmp.php) was developed by Luo *et al*. [61] of Shanghai Centre for Bioinformation Technology to perform modelling of metabolic networks. It was developed as a publicly available tool for systems biology searches, and is convenient to use due to its completely graphical operation, such as using circles to represent metabolites, short vertical lines to represent reactions, and arrows to represent the direction of the reaction. Users can manipulate the metabolic network by moving the nodes in the graphs and set the parameters using the equation editor in the software. This platform consists of various analytic approaches, and has been applied to mammalian mitochondria metabolic network reconstruction, producing useful results. In general, it is a powerful and a very convenient tool for metabolic network modelling and analysis.

3.1.4. SBRT

SBRTis (http://www.ieu.uzh.ch/wagner/software/SBRT/) an open source tool that enables storage, retrieval, analysis, and collecting of data from high-throughput experiments, and was developed by Wright and Wagner [62]. It is a free and user friendly tool that facilitates the modelling of metabolic network analysis with integration of geometry data, graph theory, and combinatorics, as well as algebra information. Besides analysis of the data, it also provides management of the data, for example, translating the data file into different formats to be used by researchers. In addition, it is a highly versatile tool that allows developers to edit or add new functionality to the tool if necessary. The SBRT can currently perform 35 methods for analysing stoichiometric networks and 16 methods from other fields such as graph theory, geometry, algebra, and combinatorics. New computational techniques can be added to the SBRT *via* process plug-ins, providing a high degree of evolvability and a unified framework for tool development in systems biology.

3.1.5. Pathway Analyser

Pathway Analyser (http://sourceforge.net/projects/pathwayanalyser/) is a tool designed for biologists to analyse metabolic pathways, especially for FBA and simulations of SBML models, and was developed by Raman and Chandra [63]. It is a statistical package for the analysis of whole-genome expression data. Pathway Analyser uses GNU Linear Programming Kit (GLPK) for linear programming,

as well as Taylor expansion for high-precision simulation and Object-Oriented Quadratic Programming (OOQP) for quadratic programming in MOMA. It can provide a comprehensive report on gene deletions from the SBML model input. Currently, the tool is just a command-line toolkit and is only compatible with the Linux operating system. However, a graphical user interface suitable for all platforms will be promoted in the future. The Pathway Analyser only allows data input in the SBML format.

3.1.6. OptFlux

OptFlux (http://www.optflux.org/) is an open source and modular tool aimed to provide user friendly computational tools for metabolic engineering applications, and was developed by Rocha *et al*. [64]. It aims to provide good usability, with a high value placed on the simplicity and intuitiveness of the tool. It is the first tool that provides a user-friendly interface that allows strain optimisation to be performed. There are four major functional modules in this tool, which are model handling, simulation, optimisation, and pathway analysis. Besides, it allows the graphical visualisation of pathways *via* Bio Visualizer, a visual plugin. OptFlux has the ability to associate numerical values to the different types of nodes and edges. Moreover, the flux values can be exported to Cell Designer. The main metabolic network analysis approaches employed in this tool are FBA, MOMA, regulatory on/off minimisation (ROOM), and MFA.

3.1.7. BioOpt

BioOpt (http://129.16.106.142/tools.php?c=bioopt) is one of the tools included in the BioMet Toolbox, and was developed by Cvijovic *et al*. [65]. The BioOpt tool is applied for *in silico* metabolic network prediction of target models. It focuses on FBA to represent the mathematical stoichiometric model and maximise the objective function. In general, this tool calculates the internal flux distributions by applying constraints and the objective function defined by the user. Users can also adjust the parameters in this tool in order to obtain useful output for system model analysis. The tool is also applied in the study of *in silico* screening and to perform fully standardised simulations. The models included in this tool are fungi and bacteria. The input file has to be in SBML format, BioOpt format, or an Excel file. Furthermore, a converter is provided for users to convert the custom model in SBML format into BioOpt format (SBML2BioOpt).

3.1.8. SurreyFBA

SurreyFBA (http://sysbio3.fhms.surrey.ac.uk/) is an integration of command line and graphical user interface tool for constraint-based modelling and network map visualisation. It is a free and standalone tool developed by Gevorgyan *et al*. [66]. It aims to perform basic simulation operations such as analysis of substrate production rate and the prediction of nutrient requirements. Besides, it is useful in the application of metabolic engineering and gene essentiality identification. SurreyFBA includes JyMet, which is a desktop graphical user interface (GUI) that presents model structure in a spreadsheet and provides a menu-based interface to simulation methods. It is a command line interface program that integrates GLPK solver to be distributed for three major types of operating systems. It was developed in C++ language, which allows easy scripting. It also provides efficient iterative task implementation for the linear programming solver.

3.1.9. FASIMU

FASIMU (http://www.bioinformatics.org/fasimu/downloads/) is a command line tool that allows metabolic flux prediction in large metabolic networks. It also plays an important role in model structural analysis and network curation. This tool was developed by Hoppe *et al*. [67] and has been integrated with a new plugin for BiNa, which provides a clearer visualisation for the metabolic flux distributions. It is a free open source application, is available online, and can work with either commercial or free solvers. It is available for the implementation of: (i) the concentration-based thermodynamic feasibility constraint; (ii) weighted flux minimisation; and (iii) fitness maximisation for partially inhibited enzymes. Besides, it allows batch computation with various objectives and constraints suited to network pruning, leak analysis, flux-variability analysis, and systematic probing of metabolic objectives for network curation.

3.2. Toolbox-based tools

Toolbox-based applications refer to the add-on libraries installed in general purpose computation such as MATLAB, Visualization and Analysis of Networks, containing Experimental Data(VANTED), and Mathematica. A toolbox in computer science can be defined as a set of precompiled routines for use in writing new programs. In other words, the term toolbox is used in computing to represent a set of subroutines (or functions) and global variables. Typically, these implement missing functionality using the capabilities available in the core tool [58]. In this paper, five toolbox-based tools for metabolic network analysis are discussed briefly, which are FluxAnalyzer/CellNetAnalyzer [69,70], SNA [71], OpenFlux [72], FBA SimVis [73], and COBRA [74].

3.2.1. FluxAnalyzer/CellNetAnalyzer

FluxAnalyzer or CellNetAnalyzer (http://www.mpi-magdeburg.mpg.de/projects/cna/cna.html) was developed by Klamt *et*

al. [69,70] at the Max Planck Institute for Dynamics of Complex Technical Systems in Magdeburg. It is a package for MATLAB that is used for reconstruction and functional analyses of metabolic networks using a comprehensive user-friendly environment. In order to visualise the interaction of cellular networks, the abstract network is linked to visual networks. Users can key in the input data and display the analysis algorithms by utilising a collection of tools on the graphical user interface, including metabolic flux analysis, flux optimisation, detection of topological features, and pathway analysis by elementary flux modes or extreme pathway. CellNetAnalyzer is the successor of FluxAnalyzer. CellNetAnalyzer was developed to extend the function of FluxAnalyzer so that it can be used for structural analysis of mass-flow (metabolic) networks. This tool has been developed and promoted for many years and is widely used in metabolic network analysis.

3.2.2. SNA Toolbox

The SNA Toolbox (http://bioinformatics.org/project/?group_id?546) was developed in 2006 by Urbanczik [71]. It is an interactive toolbox that is used to analyse metabolic networks at steady state. This can be done by calculating the elementary vectors of flux distribution and conversion cones. In particular, SNA is used to construct the abstract data type of a metabolic net based on a list of reactions. This toolbox provides a large set of functions, such as joining metabolic nets, extracting sub networks, and changing the role of metabolites in a network. The user interface of this toolbox is provided by SNAsym, a Mathematica package that gives a flexible environment for metabolic network analysis and interpretation. The Mathematica package is called SNAsym and the matrix level computational geometry routines is executed by SNAmat, which is a lower package of SNAsym. It is used to analyse large metabolic networks, and this toolbox runs in a PC-Linux or MATLAB environment.

3.2.3. OpenFlux

OpenFlux (http://openflux.sourceforge.net/) is a MATLAB-based modelling tool for small- and large-scale ^{13}C meta-flux analysis using mass isotopomer distribution data. This application which is developed by Quek *et al*. [72], is based on the new Elementary Metabolite Unit (EMU) framework that has been proven able to enhance the computation speed for flux calculation. The user interface is based on command line. It allows the user to utilise a versatile spreadsheet-based interface to control the underlying metabolite and isotopomer balance models used for flux analysis. Besides, it provides the implementation of large-scale metabolic networks in which users can choose either by accompanying algorithm package for flux estimation and sensitivity analysis or applying alternative numerical approaches for flux analysis. OpenFlux is implemented using Java programming language.

3.2.4. FBA SimVis

FBASimVis (http://fbasimvis.ipk-gatersleben.de/) was developed by Grafahrend-Belau *et al.* [73]. This is a VANTED plugin for metabolic network analysis using constraint-based modelling. It focuses on the visualisation of metabolic flux distribution output. It provides a user friendly interface for metabolic network modelling, metabolic network reconstructions, and an environment for interactive visualisation of simulation results. A drag-and-drop mechanism is provided for network constructions and the management of data is provided in the textfield. Various metabolic network analysis approaches have been integrated into FBASimVis, such as FBA, knockout analysis, robustness analysis, and flux variability analysis. Besides, various graph theoretical analysis techniques such as cycle detection, connectivity analysis, and shortest path length, as well as statistical analysis techniques such as correlation analysis, *t*-test, and cluster analysis are employed in this tool.

3.2.5. COBRA Toolbox

COBRA Toolbox (http://opencobra.sourceforge.net/openCOBRA/Welcome.html) was developed by Becker *et al*. [74]. It is run in a MATLAB environment and the metabolic network model information is entered in SBML format. The COBRA approach focuses on simulating, analysing, and predicting a variety of metabolic phenotypes by employing physicochemical constraints to a biological network in a given condition. These constraints include mass conservation, thermodynamic directionality, compartmentalisation, and molecular crowding. This powerful tool is embedded with several analytical functions such as gene knockout analysis, dynamic optimisation of production in a steady-state system, network modular analysis, and system robustness. Recently, there is an upgraded version of COBRA Toolbox that expands the *in silico* analysis methods, including network gap filling, metabolic engineering, visualisation, omics-guided analysis, and ^{13}C analysis. In addition, the usability of documentation has been improved in this newer version of the COBRA Toolbox, as updated by Schellenberger *et al.* [75].

3.3. Web-based tools

Web-based tools are accessible by users over a network such as the internet or an intranet, which require web browsers to render the application executable. In general, web-based tools are more popular due to the ubiquity of

web browsers and the convenience of using a browser as a client. In addition, it is flexible as users can access files and work on these tools from any computer that has access to the internet. Another important reason is because of the ability to update and maintain web-based tools without distributing and installing tools on potentially thousands of client computers. In this paper, a total of seven web-based tools are reviewed. Brief descriptions of each tool are provided, including GSMN-TB [9], CycSim [76], WEbcoli [77], Model SEED [78], Acorn [79], FAME [80], and MicrobesFlux [6].

3.3.1. GSMN-TB

GSMN-TB (http://sysbio3.fhms.surrey.ac.uk/cgi-bin/fba/fbapy) refers to a metabolic scale metabolic network modelling server that contains the total number of reactions, metabolites, and SMBL files of four models, *Streptomyces coelicolor* in Borodina *et al*. [81], *Streptomyces lividans* and *Mycobacterium tuberculosis* in Beste *et al*. [9], and *Neisseria meningitidis*. It is useful to investigate the cellular physiology of these models. The GSMN server provides metabolic network simulation based on FBA of these models. Flux variability analysis and reaction essentiality scan, as well as gene essentiality prediction are also provided. In conclusion, there are four analysis protocols available for the simulation of metabolic networks. The results are formatted as HyperTextMarkup Language (HTML) and sent to the user's browser. This tool provides four main functional features, including prediction of maximum growth rate, flux variability analysis, reaction essentiality scan, and gene essentiality prediction.

3.3.2. CycSim

CycSim (http://www.genoscope.cns.fr/cycsim/org.nemostudio.web.gwt.App/App.html) is a web application developed by Fevre *et al.* [76] to simulate constraint-based models of metabolism, coupled with the information from metabolic network databases KEGG and BioCyc. Specifically, this web application provides an intractable environment to perform gene knockout experiments, prediction of growth phenotype under single or multiple gene deletions under specified environment conditions, comparison of predicted results, and visualisation of both predictions and experimental results on metabolic maps. This pathway genome simulator also acts as an online repository of genome-scale metabolic models. It was developed using an AndroMDA framework with a MySQL backend deployed on a Java application server. Each analysis setting is saved in the server in order to ensure the availability of sufficient computational resources for users. The users can retrieve the results of analysis through a unique identifier.

3.3.3. WEbcoli

WEbcoli (http://webcoli.org/main.jsp) is another web application for *in silico* design, analysis, and engineering of *E. coli* metabolism, and was developed by Jung *et al*. [77]. This application implements Ajax, Java Web Start, and Java Servlet, which are advanced web technologies to enhance usability and accessibility of the application. It provides a user friendly interface that allows the user to engineer the *E. coli* strains and structure their pathways by using a graph editor. Besides, the simulation of metabolic flux distribution can be conducted through constraint-based analysis. There are three main functional features included in this tool, which are *in silico* analysis of *E. coli* metabolic pathways, virtual design and modelling, and comprehensive representation of metabolic networks. In addition, pathways can be exported in SBML format in order to facilitate communication with other systems biology tools.

3.3.4. Model SEED

Model SEED (http://seed-viewer.theseed.org/seedviewer.cgi?page%C2%BCModelView) is an online resource for reconstruction, annotation, comparison, analysis, and curation of genome-scale metabolic models developed by Henry *et al*. [78]. This application supplies the genome annotation into Model SEED by using Rapid Annotation Using Subsystems Technology (RAST) to facilitate the drafting of metabolic models. Model SEED simulates the model metabolism by using flux balance analysis. It then automatically constructs a metabolic map that contains metabolic reactions, gene-protein reactions, and biomass composition of each genome. It is mostly applied to speed up the creation of new metabolic modes, especially of non-compartmentalised species such as bacteria. There are three main types of metabolic models available, which are pathogens, mammalian cells, and industrially relevant organisms. Overall, 130 metabolic models have been published by the SEED project.

3.3.5. Acorn

Acorn (http://sysbio3.fhms.surrey.ac.uk:8080/acorn/homepage.jsf) refers to an open source grid computing system for constraint-based modelling and visualisation of genome-scale metabolic reaction networks using an interactive web interface. It was developed by Sroka *et al.* [79]. This application uses efficient web technologies of Ajax to improve the model browsing and search functions. It employs flux variability analysis, which allows the execution of computational intensive metabolic network analysis. The advantage of this tool is that it can share the selected models and simulation results between different users and make them publicly available. Besides, the constructed

pathway map that is used to visualise numerical results can be imported into the server using the integrated editor for desktop. The main application can be installed on any platform running Java VM, the Glassfish application server, and relational database management system.

3.3.6. FAME

FAME (http://f-a-m-e.org/ajax/page1.php), developed by Boele *et al.* [80], is a web-based application that aims to perform metabolic network modelling, for example, creating, running, analysing, and visualising models into a single program. In general, it is a one stop shop stoichiometric modelling web-based program that provides a user-friendly interface. The goal of this tool is to perform model reconstruction, result generation, and result interpretation, including sensitivity analysis, pathway visualisation, and metabolite connectivity. This application implements PHP script version 5 (PHP5) and JavaScript to enable faster browsing. In addition, this web-based application can run or edit the existing model, build a new model, or merge models. The input model format for this tool is SBML. Furthermore, FAME integrates the Python-based PySCeS-CBM as a linear solver. This toolkit is able to handle mathematical operations and returns the results to FAME.

3.3.7. MicrobesFlux

MicrobesFlux (http://tanglab.engineering.wustl.edu/static/MicrobesFlux.html) is a web-based platform developed by Feng *et al.* [6]. It provides high-throughput metabolic drafting and constraint-based modelling for both steady and non-steady states. It is semi-automated for constructing metabolic network models. It provides gene knockout protocols and heterologous pathways for metabolic model reconstructions. Then, the constructed model can be formulated into a stoichiometric model to predict and perform metabolic model network simulation by using flux balance analysis or dynamic flux balance analysis. In summary, the functional features of the tools include metabolic model reconstruction, customised metabolic model drafting, and flux distribution analysis. Besides, it provides integration into KEGG in which the metabolic networks of more than 1200 species in KEGG can be downloaded automatically. The results of the analysis can be exported in SBML format and the metabolic networks can also be viewed in Scalable Vector Graphics (SVG) format in a web browser.

3.4. Comparison of tools

Comparison of the three classes of tools are discussed in this section. Table 6 shows the comparison of their dependency and usability, whereas Table 7 shows the comparison in terms of functionality support. In Table 6, two strain design algorithm namely Optknock [3] and Optgene [82] are also taken account as some of the tools provide the functionality. OptKnock is a bilevel programming framework to identify best set of genes to be knockout in order to increase the production of metabolites whereas OptGene is the extension of OptKnock which utilize genetic algorithm to increase the prediction capability. These two tables are modified from Tables 2 and 3 in Lakshmanan *et al*. [58]. Table 8 showed the features supported by the mentioned tools.

4. Discussion and Conclusion

With the rapid development of genomics and various successful genome projects, biologists have deciphered the genome sequence and metabolic networks of many organisms, such organisms can be faithfully reconstructed from the available genome information. Thus, the analysis of metabolic networks has become essential in further studies, in which such analysis could help us to obtain a better understanding of the topology and biological functions for different organisms, enabling us to utilise cellular metabolic processes to assist the development of fermentation technology, medicines, and agriculture. Successful development in these areas would not only benefit economics, but also allow us to understand the biological evolution.

This review discusses a number of tools and databases that can be useful in metabolic network analysis. The brief description of features and capabilities of each of the tools and databases is given in order to give a clear view of their scope and help researchers to choose the ones that are suitable for a particular research project and to differentiate the goal for each of the tools and databases. For the tools, there is a choice of a wide range of software prior to conducting the experiment. Here, we have grouped the different tools that are commonly used to perform metabolic network analysis into three classes, which are standalone application, toolbox-based libraries, and web-based applications. Each tool has its own advantages as well as disadvantages that can become a limitation. Therefore, choosing the right tool will determine the efficiency of the research and maintain optimal resource allocations. Integration of tools is sometimes required when it comes to a problem that requires a wide range of capabilities, for example, in interdisciplinary research.

In this review, eight commonly used metabolic pathway and enzyme databases are also discussed. It is crucial for researchers to obtain the data before using these tools. Online databases are easy to use, save time, and can be frequently accessed. The most important thing is that they are free to access. These databases cover a wide range of

Table 6. Dependency and usability of the tools [58]

Tool	Platform supported	Additional software requirements	Solver	Price	Input file format
MetaFluxNet	Windows	≥NET framework 1.1, ≥ Java JRE 1.4.2 NA	Qsopt LP	Free	FBA text File
BioSPICE	Windows, Linux and Mac OS	Open Agent Architecture, Java SDK 1.4.5	ESS	Free	System Biology Markup Language (SBML) file
FluxExplorer	Windows	NA	NA	Free	System Biology Markup Language (SBML) file
SBRT	Windows	≥ Java JRE 1.5.x	GLPK	Free	System Biology Markup Language (SBML) file
Pathway Analyser	Linux, UNIX	NA	GLPK, OOQP	Free	PathoLogic format file
OptFlux	Windows and Linux	Java JRE 1.6.x	GLPK	Free	System Biology Markup Language (SBML) file
BioOpt	Windows	NA	GLPK	Free	BioOpt text file
SurreyFBA	Windows, Linux and Mac OS X	Java JRE 1.6.x (4)	GLPK	Free	Microsoft access file
FASIMU	Linux and Mac OS X	Optimisation solver	GLPK, CPLEX, LINDO, LP	Free	Biolofical Network Analyzer text file
FluxAnalyzer/ CellNetAnalyzer	Windows, Linux and Mac OS X	≥MATLAB 7.1, optimisation solver, SBML Toolbox 3.0.0	GLPK, Optimisation Toolbox (MATLAB)	Commercial	System Biology Markup Language (SBML) file
SNA	Linux	≥Mathematica 5	NA	Free	
OpenFlux	Windows	Java^{TM} 5 or 6, ≥MATLAB 6.5, Optimization and Statistics Toolbox^{TM}	Roe-type Riemann	Free	Microsoft excel file
FBA SimVis	Windows	≥VANTED 1.8, Java JRE 1.6.x	GLPK	Free	Geography Markup Language (GML) file
COBRA	Windows and Mac OS X	≥MATLAB 6.5, optimisation solver, ≥libSBML 4.0, ≥SBML Toolbox 3.0.0	GLPK, TOMLAB.CPLEX, LINDO, Gurobi	Free	System Biology Markup Language (SBML) file
GSMN-TB	All Platforms	Web browser	GLPK	Free	Microsoft excel file
CycSim	All Platforms	Web browser ≥Firefox 2.0	-	Free	System Biology Markup Language (SBML) file
WEbcoli	All Platforms	Web browser ≥IE 6.0 or ≥Firefox 2.0	ILOG.CPLEX	Free	System Biology Markup Language (SBML) file
Model SEED	All Platforms	Web browser ≥Firefox 2.0	ILOG.CPLEX	Free	System Biology Markup Language (SBML) file
Acorn	All Platforms	Web browser, Java JRE 1.6.x	GLPK	Free	System Biology Markup Language (SBML) file
FAME	All Platforms	Web browser Google Chrome, Firefox	GLPK/ ILOG.CPLEX	Free	System Biology Markup Language (SBML) file
MicrobesFlux	All Platforms	Web browser	IPOPT	Free	(KEGG Markup Language) file

Qsopt LP, Qsopt Linear Programming; **ESS**, Exact Stochastic Simulator; **GLPK**, GNU Linear Programming Toolkit; **OOQP**, Object Oriented quadratic Programming; **LP,** Linear Programming; **IPOPT**, Interior Point OPTimizerm; **IE**, Internet Explorer; **Java JRE**, Java Runtime Environment.

Table 7. Functionality support by tools

Tool	Support for model reconstruction		Metabolic network modelling/analysis										Strain design algorithm	
	Pathway Import from DBs	Assistance in gap filling	FBA	MOMA	ROOM	MFA	DFBA	FCA	FVA	SVD	EMA	EPA	OptKnock [3]	OptGene [82]
MetaFluxNet		√	√			√			√		√			
BioSPICE				√										
FluxExplorer			√	√						√		√		
SBRT		√	√			√			√		√		√	√
Pathway Analyser			√	√										
OptFlux			√	√	√	√			√		√		√	√
BioOpt			√						√					
SurreyFBA			√						√		√			
FASIMU		√	√	√	√	√		√	√					
FluxAnalyzer/ CellNetAnalyzer			√			√			√		√			
SNA			√								√			
OpenFlux						√								
FBA SimVis	√		√						√					
COBRA		√	√	√			√		√				√	√
GSMN-TB			√						√					
CycSim			√								√			
WEbcoli			√											
Model SEED	√	√	√						√					
Acorn			√						√					
FAME	√		√						√					
MicrobesFlux	√		√											

FBA, flux balance analysis; **MOMA**, minimization of metabolic adjustment; **ROOM**, regulatory on/off minimisation; **MFA**, metabolic flux analysis; **DFBA**, dynamic flux balance analysis; **FCA**, flux coupling analysis; **FVA**, flux variability analysis; **SVD**, singular value decomposition; **EMA**, extreme mode analysis; **EPA**, extreme pathway analysis.

Table 8. Features support by tools

	Integrate omics (such as genomics) data with the metabolic model	Dynamic-flux balance analysis or integrate data	Combine different databases to fill the gaps in the model	Use the genome sequence to draft the metabolic model automatically
MetaFluxNet	√	√	√	√
BioSPICE	√	√	√	√
FluxExplorer	×	√	×	√
SBRT	√	√	√	×
Pathway Analyser	√	√	×	×
OptFlux	√	√	√	×
BioOpt	√	√	√	√
SurreyFBA	√	√	×	×
FASIMU	√	√	√	×
FluxAnalyzer/ CellNetAnalyzer	√	√	√	×
SNA	√	√	×	×
OpenFlux	√	√	√	√
FBA SimVis	√	√	√	×
COBRA	√	√	√	√
GSMN-TB	√	√	×	√
CycSim	√	√	×	√
WEbcoli	√	√	×	×
Model SEED	√	√	√	×
Acorn	√	√	×	√
FAME	√	√	×	×
MicrobesFlux	√	√	√	√

Symbol ' √' indicates the feature is consisted in the tool. Symbol 'x' indicates the feature is not consisted in the tool.

information and present the information differently. It is possible that there are inconsistencies in the data provided. Thus, in some cases cross-referencing between different databases is necessary. However, due to the various ways in which the information is presented, the data curated by different groups of people, and the format of the data supported, it is rather a difficult task to integrate and directly compare these databases. Here, a general comparison showing the advantages and disadvantages is presented to give the reader a general idea of the tools to help them to make an appropriate choice for their research.

Future development of these databases should incorporate the ability to integrate data from various sources with different formats so that the information obtained by a query will generate more detailed results. Furthermore, the databases should include a search engine for searching by level, such as pathway or genome, instead of the traditional gene, reaction, compound, and enzyme queries, and the query should produce fast results. This will give a bigger picture of the metabolic network that will allow more exploration for the users and reduce the time consumed. The range of organisms covered in the database should also be expanded and the annotation and naming of the components such as genes and pathways should be standardised for easier reference. Software tools that can retrieve information from the databases and allow more control for the user to view, explore, and modify the metabolic network will be essential in the next generation of metabolic network analysis.

Acknowledgements

We would like to thank the Malaysian Ministry of Science, Technology and Innovation for supporting this research by an e-science research grant (Grant number: 01-01-06-SF1234). This research is also funded by an Exploratory Research Grant Scheme (Grant number: R.J130000.7807.4L096) and a Fundamental Research Grant Scheme (Grant number: R.J130000.7807.4F190) from the Malaysian Ministry of Higher Education.

References

1. Wagner, A. (2012) Metabolic networks and their evolution. *Adv. Exp. Med. Biol.* DOI 10.1007/978-1-4614-3567-92.
2. Chen, X. W., A. P. Alonso, D. K. Allen, J. L. Reed, and Y. Shachar-Hill (2010) Synergy between 13C-metabolic flux analysis and flux balance analysis for understanding metabolic adaption to anaerobiosis in *E. coli*. *Metab. Eng.* 13: 38-48.
3. Burgard, A. P., P. Pharkya, and C. D. Maranas (2003) OptKnock: A bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. *Biotechnol. Bioeng.* 84: 647-657.
4. Koffas, M. A., G. Y. Jung, and G. Stephanopoulos (2003) Engineering metabolism and product formation in corynebacterium glutamicum by coordinated gene overexpression. *Metab. Eng.* 5: 32-41.
5. Alpher, H., Y. S. Jin, J. F. Moxley, and G. Strephanopoulos (2005) Identifying gene target for the metabolic engineering of lycopene biosynthesis in *Escherichia coli*. *Metab. Eng.* 7: 155-164.
6. Feng, X. Y., Y. Xu, Y. X. Chen, and Y. J. J. Tang (2012) MicrobesFlux: A web platform for drafting metabolic models from the KEGG Database. *BMC Syst. Biol.* 6: 94.
7. Karp, P. D. and R. Caspi (2011) A survey of metabolic databases emphasizing the MetaCyc family. *Arch. Toxicol*. 85: 1015-1033.
8. Raman, K., P. Rajagopalan, and N. Chandra (2005) Flux balance analysis of mycolic acid pathway: Targets for anti-tubercular drugs. *PLoS. Comput. Biol.* 1: 349-358.
9. Beste, D. J. V., T. Hooper, G. Stewart, B. Bonde, C. Avignone-Rossa, M. E. Bushell, P. Wheeler, S. Klamt, A. M. Kierzek, and J. McFadden (2007) GSMN-TB: A web-based genome-scale network model of *Mycobacterium tuberculosis* metabolism. *Genome Biol.* 8: 89.
10. Jamshidi, N. and B. Palsson (2007) Investigating the metabolic capabilities of *Mycobacterium tuberculosis* H37Rv using the *in silico* strain iNJ661 and proposing alternative drug targets. *BMC Syst. Biol.* 1: 26.
11. Larhlimi, A. and A. Bockmayr, (2007) Constraint-based Analysis of Gene Deletion in A Metabolic Network. *Workshop on Constraint based Methods for Bioinformatics.* 48-55.
12. Covert, M. W., J. L. Reed, E. M. Knight, M. J. Herrgard, and B. O. Palsson (2004) Integrating high-throughput and computational data elucidates bacterial networks. *Nature* 429: 92-96.
13. Price, N. D., J. L. Reed, and B. Palson (2004) Genome-scale models of microbial cells: Evaluating the consequences of constraints. *Nat. Rev. Microbiol.* 2: 886-897.
14. Ogata, H., S. Goto, K. Sato, W. Fujibuchi, H. Bono, and M. Kanehisa (1999) KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 27: 29-34.
15. Moutselos, K., I. Kanaris, A. Chatziioannou, I. Maglogiannis, and F. N. Kolisis (2009) KEGG converter: A tool for the *In-silico* modelling of metabolic networks of the KEGG pathways database. *BMC Bioinformat.* 10: 324.
16. Stobbe, M. D., S. M. Houten, G. A. Janse, A. H. C. V. Kampen, and P. D. Moerland (2011) Critical assessment of human metabolic pathway databases: A stepping stone for future integration. *BMC Syst. Biol.*5: 165.
17. Liu, H., J. Wang, C. Zhuang, N. Han, B. Wei, and S. Rayner (2010) Development of a pathway comparison tool for analysis of bacteria genomes. *Bioinformatics and Biomedicine Workshops (BIBMW), 2010 IEEE International Conference*. December 18-18. Hongkong.
18. Karp, P. D., C. A. Ouzounis, C. Moore-Ko*chlacs, L. Gol*dovsky, P. Kaipa, D. Ahren, S. Tsoka, N. Darzentas, V. Kunin, and N. Lopez-Bigas (2005) Expansion of the BioCyc collection of pathway/genome databases to 160 genomes. *Nucleic Acids Res*. 33: 6083-6089.
19. Krummenacker, M., S. Paley, L. Mueller, T. Yan, and P. D. Karp (2005) Querying and computing with BioCyc databases. 21: 3454-3455.
20. Choi, C., R. Münch, B. Bunk, J. Barthelmes, C. Ebeling, D. Schomburg, M. Schobert, and D. Jahn (2007) Combination of a data warehouse concept with web services for the establishment of the *Pseudomonas* systems biology database SYSTOMONAS. *J. Integrative Bioinformat*.. 4: 48.
21. Karp, P. D., M. Riley, S. M. Paley, and A. Pellegrini-Toole (2002) The MetaCyc database. *Nucleic Acids Res.* 30: 59-61.
22. Krieger, C. J., P. Zhang, L. A. Mueller, A. Wang, S. Paley, M.

Arnaud, J. Pick, S. Y. Rhee, and P. D. Karp (2004) MetaCyc: A multiorganism database of metabolic pathways and enzymes. *Nucleic Acids Res*. 32: 439-442.
23. Zhang, P., H. Foerster, C. P. Tissier, L. Mueller, S. Paley, P. D. Karp, and S. Y. Rhee (2005) MetaCyc and AraCyc. Metabolic pathway databases for plant research. *Plant Physiol*.138: 27-37.
24. Caspi, R., H. Foerster, C. A. Fulcher, P. Kaipa, M. Krummenacker, M. Latendresse, S. Paley, S. Y. Rhee, A. G. Shearer, C. Tissier, T. C. Walk, P. Zhang, and P. D. Karp (2008) The metacyc database of metabolic pathways and enzymes and the biocyc collection of pathway or genome databases. *Nucleic Acids Res*.36: 623-631.
25. Bairoch, A. (2000) The ENZYME database in 2000. *Nucleic Acids Res*. 28: 304-305.
26. Sheu, S., D. R. L. Jr, K. H. Clodfelter, M. R. Landon, and S. Vajda (2005) PRECISE: A database of predicted and consensus interaction sites in enzymes. *Nucleic Acids Res*. 33: 206-211.
27. Andreini C., I. Bertini, G. Cavallaro, G. L. Holliday, and J. M. Thornton (2008) Metal ions in biological catalysis: From enzyme databases to general principles. *J. Biol. Inorg. Chem.* 13: 1205-1218.
28. Sharma, V. K., N. Kumar, T. Prakash, and T. D. Taylor (2010) MetaBioME: A database to explore commercially useful enzymes in metagenomic datasets. *Nucleic Acids Res*. 38: 468-472.
29. Schomburg, D. and I. Schomburg (2001) *Springer Handbook of Enzymes.* 2 nd ed. Springer, Heidelberg, Germany.
30. Schomburg, I., A. Chang, O. Hofmann, C. Ebeling, F. Ehrentreich, and D. Schomburg (2002) BRENDA: A resource for enzyme data and metabolic information. *Trends Biochem. Sci.* 27: 54-56.
31. Schomburg, I., A. Chang, C. Ebeling, M. Gremse, C. Heldt, G. Huhn, and D. Schomburg (2004) BRENDA: The Enzyme Database: Updates and major new developments. *Nucleic Acids Res.*. 32: 431-433.
32. Barthelmes, J., C. Ebeling, A. Chang, I. S. Chomburg, and D. Schomburg (2006) New developments at the brenda enzyme information system. 211-225.
33. Schellenberger, J., J. O. Park, T. M. Conrad, and B. Palsson (2010) BiGGe: A biochemical genetic and genomic knowledgebase of large scale metabolic reconstructions. *BMC Bioinformat.* 11: 213-222.
34. Muja, M., R. B. Rusuy, G. Bradskiy, and D. G. Lowe (2010) REIN - A Fast, Robust, Scalable REcognition Infrastructure. 1-8.
35. Haw, R. A., D. Croft, C. K. Yung, N. Ndegwa, P. D. Eustachio, H. Hermjakob, and L. D. Stein (2011) The Reactome BioMart. *Database: The J. Biol. Databases and Curat.* 2011: bar031.
36. Crof, D., G. O'Kelly, G. Wu, R. Haw, M. Gillespie, S. L. Matthew, M. Caudy, P. Garapati, G. Gopinath, B. Jassal, S. Jupe, I. Kalatskaya, S. Mahajan, B. May, N. Ndegwa, E. Schmidt, V. Shamovsky, C. Yung, E. Birney, H. Hermjakob, P. D'Eustachio, and L. Stein (2011) Reactome: A database of reactions, pathways and biological processes. *Nucleic Acids Res.* 39: 691-697.
37. Milacic, M., R. Haw, K. Rothfels, G. Wu, D. Croft, H. Hermjakob, P. D. Eustachio, and L. Stein (2012) Annotating cancer variants and anti-cancer therapeutics in reactome. *Cancers* 4: 1180-1211.
38. Pao, W. and J. Chmielecki (2010) Rational, biologically based treatment of EGFR-mutant non-small-cell lung cancer. *Nat. Rev. Cancer* 10: 760-774.
39. Sakurai, N., T. Ara, Y. Ogata, R. Sano, T. Ohno, K. Sugiyama, A. Hiruta, K. Yamazaki, K. Yano, K. Aoki, A. Aharoni, K. Hamada, K. Yokoyama, S. Kawamura, H. Otsuka, T. Tokimatsu, M. Kanehisa, H. Suzuki, K. Saito, and D. Shibata (2011) KaPPA-View4: A metabolic pathway database for representation and analysis of correlation networks of gene co-expression and metabolite co-accumulation and omics data. *Nucleic Acids Res.* 39: 677-684.
40. Tokimatsu, T., N. Sakurai, H. Suzuki, H. Ohta, K. Nishitani, T. Koyama, T. Umezawa, N. Misawa, K. Saito, and D. Shibata (2005) KaPPA-View: A web-based analysis tool for integration of transcript and metabolite data on plant metabolic pathway maps. *Plant Physiol.* 138: 1289-1300.
41. Devappa, R. K., H. P. Makkar , and K. Becker (2010) Nutritional, biochemical, and pharmaceutical potential of proteins and peptides from jatropha: Review. *J. Agric. Food Chem*. 58: 6543-6555.
42. Abdulla, R., E. S. Chan, and P. Ravindra (2011) Biodiesel production from Jatrophacurcas: A critical review. *Crit. Rev. Biotechnol.* 31: 53-64.
43. Thomas, R., N. K. Sah, and P. B. Sharma, (2008) Therapeutic biology of Jatrophacurcas: A mini review. *Curr. Pharm. Biotechnol.* 9: 315-324.
44. Sakurai, N. , Y. Ogata, T. Ara, R. Sano, N. Akimoto, A. Hiruta, H. Suzuki, M. Kajikawa, U. Widyastuti, S. Suharsono, A. Yokota, K. Akashi, J. Kikuchi, and D. Shibata (2012) Development of KaPPA-View4 for omics studies on Jatropha and a database system KaPPA-Loader for construction of local omics databases. *Plant Biotechnol*. 29: 131-135.
45. Barrett, T., D. B. Troup, S. E. Wilhite, P. Ledoux, C. Evangelista, I. F. Kim, M. Tomashevsky, K. A. Marshall, K. H. Phillippy, P. M. Sherman, R. N. Muertter, M. Holko, O. Ayanbule, A. Yefanov, and A. Soboleva (2011) NCBI GEO: Archive for functional genomics data sets- 10 years on. *Nucleic Acids Res*. 39: 1005-1010.
46. McAnulty, M. J., J. Y. Yen, B. G. Freedman, and R. S. Senger (2012) Genome-scale modelling using flux ratio constraints to enable metabolic engineering of clostridial metabolism *in silico*. *BMC Syst. Biol*. doi: 10.1186/1752-0509-6-42.
47. Chung, B. K., M. Lakshmanan, M. Klement, B. Mohanty, and D. Y. Lee (2013) Genome-scale *in silico* modeling and analysis for designing synthetic terpenoid-producing microbial cell factories. *Chem. Eng. Sci*. 103: 100-108.
48. Park, J. M., H. Song, H. J. Lee, and D. Seung (2013) Genome-scale reconstruction and *in silico* analysis of *Klebsiella oxytoca* for 2,3-butanediol production. *Microbial. Cell Fact.* 12: 1-11.
49. Hanly, T. J. and M. A. Henson (2013) Dynamic metabolic modeling of a microaerobic yeast co-culture: Predicting and optimizing ethanol production from glucose/xylose mixtures. *Bioetchnol. Biofuels* 6: 1-16.
50. Karp, P. D., S. M. Paley, M. Krummenacker, M. Latendresse, J. M. Dale, T. J. Lee, P. Kaipa, F. Gilham, A. Spaulding, and L. Popescu (2010) Pathway Tools version 13.0: Integrated software for pathway/genome informatics and systems biology. *Brief Bioinform.* 11: 40-79.
51. Dale, J. M., L. Popescu, and P. D. Karp (2010) Machine learning methods for metabolic pathway prediction. *BMC Bioinformat*. 11: 15.
52. Sheu, S., Jr, D. R. L., K. H., Clodfelter, M. R. Landon, and S. Vajda (2005) PRECISE: A database of predicted and consensus interaction sites in enzymes. *Nucleic Acids Res*. 33: 206-211.
53. Andreini, C., I. Bertini, G. Cavallaro, G. L. Holliday, and J. M. Thornton (2008) Metal ions in biological catalysis: From enzyme databases to general principles. *J. Biol. Inorg. Chem.* 13: 1205-1218.
54. Sharma, V. K., N. Kumar, T. Prakash, and T. D. Taylor (2010) MetaBioME: A database to explore commercially useful enzymes in metagenomic datasets. *Nucleic Acids Res*. 38: 468-472.
55. Ranganathan, S. and C. D. Maranas (2010) Microbial 1-butanol production: Identification of non-native production routes and *in silico* engineering interventions. *Biotechnol. J.* 5: 716-725.
56. Duarte, N. C., S. A. Becker, N. Jasmshidi, I. Thiele, M. L. Mo, T. D. Vo, R. Srivas, and B. Palsson (2006) Global reconstruction of

the human metabolic network based on genomic and bibliomic data. *Proc. Nat. Acad. Sci*. 104: 1777-1782.
57. Schellenberger, J., J. P. Park, T. M. Conrad, and B. Palsson (2010) BiGG: A biochemical genetic and genomic knowledge base of large scale metabolic reconstructions. *BMC Bioinform*. 11: 213.
58. Lakshmanan, M., G. Koh, B. Chung, and D. Lee (2012) Software applications for flux balance analysis. *Brief. Bioinforma.* 15: 108-112.
59. Lee, S. Y., D. Y. Lee, S. H. Hong, T. Y. Kim, H. Yun, Y. G. Oh, and S. Park (2003) MetaFluxNet: A program package for metabolic pathway construction and analysis, and it used in large-scale metabolic flux analysis of *Escherichia Coli*. *Genome Inform*. 14: 23-33.
60. Garvey, T. D., P. Lincoln, C. J. Pederson, D. Martin, and M. Johnson (2003) BioSPICE: Access to the most current computational tools for biologists. OMICS: *A J. Integrat. Biol*. 7: 411-420.
61. Luo, R. Y., S. Liao, and S. Q. Zeng (2006) FluxExplorer: A general platform for modeling and analyses of metabolic networks based on stoichiometry. *Chin. Sci. Bull*. 51: 689-696.
62. Wright, J. and A. Wagner (2008) The systems biology research tool: Evolvable open-source software. *BMC Syst. Biol.* 2: 55.
63. Raman, K. and N. Chandran (2008) Pathway analyser: A systems biology tool for flux analysis of metabolic pathways. *Nat. Prec.* 2: 38.
64. Rocha, I., P. Maia, P. Evangelista, P. Vilaca, S. Soares, J. P. Pinto, J. Nielsen, K. R. Patil, E. C. Ferreira, and M. Rocha (2010) OptFlux: An open-source software platform for *in silico* metabolic engineering. *BMC Syst. Biol*. 4: 45.
65. Cvijovic, M., R. Olivares-Hernandez, R. Agren, N. Dahr, W. Vongsangnak, I. Nookaew, K. R. Patil, and J. Nielsen (2010) BioMet toolbox: Genome-wide analysis of metabolism. *Nucleic Acid Res.* 38: 144-149.
66. Gevorgyan, A., M. E. Bushell, C. Avignone-Rossa, and A. M. Kierzek (2011) SurreyFBA: A command line tool and graphical user interface for constraint-based modeling of genome-scale metabolic reaction network. *Bioinform.* 27: 433-434.
67. Hoppe, A., S. Hoffmann, A. Gerasch, C. Gille, and H. G. Holzhutter (2011) FASIMU- Flexible software for flux-balnace computation series in large metabolic networks. *BMC Bioinform.* 12: 28.
68. Segre, D., J. Zucker, J. Katz, X. X. Lin, P. D'haeseleer, W. P. Rindone, P. Kharchenko, D. H. Nguyen, M. A. Wright, and G. M. Church (2003) From annotated genomes to metabolic flux models and kinetics parameter fitting. *OMICS* 7: 301-316.
69. Klamt, S., J. Saez-Rodriguez, and E. D. Gilles (2007) Structural and functional analysis of cellular networks with CellNetAnalyzer. *BMC Syst. Biol.* 1: 2.
70. Klamt, S., J. Stelling, M. Ginkel, and E. D. Gilles (2003) FluxAnalyzer: Exploring structure, pathways and flux distributions in metabolic networks on interative flux maps. *Bioinformat.* 19: 261-269.
71. Urbanczik, R. (2006) SNA- A toolbox for the stoichiometric analysis of metabolic networks. *BMC Bioinformat.* 7: 129.
72. Quek, L. E., C. Wittmann, L. K. Nielsen, and J. O. Kromer (2009) OpenFLUX: Efficient modelling software for 13C-based metabolic flux analysis. *Microbial. Cell Fact.* 8: 25.
73. Grafahrend-Belau, E., C. Klukas, B. H. Junker, and F. Schreiber (2009) FBA-SimVis: Interactive visualization of constraint-based metabolic models. *Bioinformat. Syst. Biol.* 25: 2755-2757.
74. Becker, S. A., A. M. Feist, L. M. Monica, G. Hannum, B. Palsson, and M. J. Herrgard (2007) Quantitative prediction of cellular metabolism with constraint-based models: The COBRA toolbox. *Nat. Protocols* 2: 727-738.
75. Schellenberger, J., R. Que, R. M. T. Fleming, I. Thiele, J. D. Ortho, A. M. Feist, D. C. Zielinski, A. Bordbar, N. E. Lewis, S. Rahmanian, J. Kang, D. R. Hyduke, and B. Palsson (2011) Quantitative prediction of cellular metabolism with constraint-based models: The COBRA Toolbox v2.0. *Nat. Protocols* 2: 727-738.
76. Fevre, F. L., S. Smidtas, C. Combe, M. Durot, F. d'Alche-Buc, and V. Schachter (2009) CycSim- An online tool for exploring and experimenting with genome-scale metabolic models. *Bioinformat.* 25: 1987-1988.
77. Jung, T. S., H. C. Yeo, S. G. Reddy, W. S. Cho, and D. Y. Lee (2009) WEbcoli: An interactive and asynchronous web application for *in silico* design and analysis of genome-scale *E. coli* model. *Bioinformat.* 25: 2850-2852.
78. Henry, C. S., M. Dejongh, A. A. Best, P. M. Frybarger, B. Linsay, and R. L. Stevens (2010) High-throughput generation, optimization and analysis of genome-scale metabolic models. *Nat. Biotechnol*. 28: 977-982.
79. Sroka, J., L. Bieniasz-Krzywiec, S. Gwozdz, D. Leniowski, J. Lacki, M. Markowski, C. Avignone-Rossa, M. E. Bushell, J. McFadden, and A. M. Kierzek (2011) Acorn: A grid computing system for constraint based modeling and visualization of the genome scale metabolic reaction networks *via* a web interface. *BMC Bioinformat.* 12: 196.
80. Boele, J., B. G. Olivier, and B. Teusink (2012) FAME, the flux analysis and modeling environment. *BMC Syst. Biol.* 6: 8.
81. Borodina, I., P. Kraben, and J. Nielsen (2005) Genome-scale analysis of streptomyces coelicolorA3(2) metabolism. *Genome Res.* 15: 820-829.
82. Patil, K. R., I. Rocha, J. Forster, and J. Nielsen (2005) Evolutionary programming as a platform for *in silico* metabolic engineering. *BMC Bioinformat*. 6: 308.
83. Wright, J. and W. Andreas (2008) The systems biology research tool: Evolvable open-source software. *BMC Syst. Biol.* 2: 55.
84. Raman, K. and C. Nagasuma (2009) Flux balance analysis of biological systems: Applications and challenges. *Brief. Bioinformat.* 10: 435-449.
85. Copeland, W. B., B. A. Bartley, D. Chandran, M. Galdzicki, K. H. Kim, S. C. Sleighta, C. D. Maranasc, and H. M. Sauroa (2012) Computational tools for metabolic engineering. *Metabol. Eng.* 14: 270-280.
86. Abrusán, G. (2012) Somatic transposition in the brain has the potential to influence the biosynthesis of metabolites involved in Parkinson's disease and schizophrenia. *Biol. Direct.* 7: 41.
87. Flamm, C., C. Hemmingsen, and D. Merkle (2013) Barrier trees for metabolic adjustment landscapes. *Adv. Artificial Life, ECAL.* 12: 1-8.
88. Quirós, M., R. Martínez-Moreno, J. Albiol, P. Morales, F. Vázquez-Lima, A. Barreiro-Vázquez, P. Ferrer, and R. Gonzalez (2013) Metabolic flux analysis during the exponential growth phase of *Saccharomyces cerevisiae* in wine fermentations. *PloS one.* 8: e71909.
89. Rios-Estepa, R. (2008) *Unraveling the regulation of mint monoterpene biosynthesis: Development and experimental testing of kinetic mathematical models*. Dissertation. Washington State University, USA.
90. Xiong, Z. and W. L. Peter (1997) COBRA: A sensitive and quantitative DNA methylation assay. *Nucleic Acids Res.* 25: 2532-2534.
91. Feng, X., Y. Xu, Y. Chen, and Y. J. Tang (2012) MicrobesFlux: A web platform for drafting metabolic models from the KEGG database. *BMC Syst. Biol*. 6: 94.