

A metabolic network approach for the identification and prioritization of antimicrobial drug targets

Arvind K. Chavali^{1*}, Kevin M. D'Auria^{1*}, Erik L. Hewlett², Richard D. Pearson^{2,3} and Jason A. Papin¹

¹ Department of Biomedical Engineering, University of Virginia, Charlottesville, VA 22908, USA

² Department of Medicine, Division of Infectious Diseases and International Health, University of Virginia, Charlottesville, VA 22908, USA

³ Department of Pathology, University of Virginia, Charlottesville, VA 22908, USA

For many infectious diseases, novel treatment options are needed in order to address problems with cost, toxicity and resistance to current drugs. Systems biology tools can be used to gain valuable insight into pathogenic processes and aid in expediting drug discovery. In the past decade, constraint-based modeling of genome-scale metabolic networks has become widely used. Focusing on pathogen metabolic networks, we review *in silico* strategies used to identify effective drug targets and highlight recent successes as well as limitations associated with such computational analyses. We further discuss how accounting for the host environment and even targeting the host may offer new therapeutic options. These systems-level approaches are beginning to provide novel avenues for drug targeting against infectious agents.

Systems biology and pathogen metabolism

Systems biology methods have been applied extensively to the study of infectious diseases across multiple scales of biological organization to generate predictions ranging from pathogen gene lethality in particular microenvironments [1,2] to dynamics involved in the host immune response to infection [3]. Utilizing the predictive power of computational modeling and systems analysis, wide-ranging questions related to pathogen virulence, disease progression and host response can be explored to generate hypotheses for more thorough experimental investigation.

The increasing availability of high-quality genome-scale metabolic reconstructions [4] presents an opportunity for the rational and systematic identification of metabolic drug targets in a pathogen of interest. Built bottom-up from functional genome annotations (and a variety of other data sources) and analyzed with computational methods such as flux balance analysis (FBA; see Glossary), these biochemical networks can account for hundreds to thousands of metabolites participating in enzymatic reactions across a range of metabolic subsystems (e.g. carbohydrate, amino acid, lipid, nucleotide and energy metabolism) and cellular localizations (e.g. extracellular space, cytosol and

Glossary

Biomass: an objective reaction consisting of important metabolites that the cell needs in order to grow (e.g. amino acids, lipids and carbohydrates).

Correlated reaction sets: groups of reactions whose fluxes always change in relation to one another.

Exchange reaction: a reaction that transports metabolites into or out of the system.

Feasible flux space: the possible combination of allowable fluxes given a defined set of constraints which bound reaction fluxes, for example reversible and irreversible reactions.

Flux balance analysis (FBA): a linear programming problem formulated to maximize the flux through an objective reaction, which the analyst defines, under steady-state flux constraints that are derived from the stoichiometric matrix of the metabolic network.

Flux distribution: the set of all reaction fluxes within a metabolic network.

Flux variability analysis (FVA): a linear programming-based method which determines the minimum and maximum reaction fluxes that allow for optimal or near-optimal flux through the objective reaction.

Gene-protein-reaction (GPR) relationship: the combination of proteins or protein components sufficient to carry out an enzymatic reaction and the combination of genes sufficient to express each of the protein components.

In silico: in contrast to *in vivo* or *in vitro*, the term indicates a computational process in a simulated environment.

Linear programming: also termed linear optimization, an area of mathematics developed for maximizing a linear combination of variables (e.g. $A \cdot v_1 + B \cdot v_2 + C \cdot v_3$), such that the variables are constrained by many linear equalities and inequalities (e.g. the constraint $v_1 - 2 \cdot v_2 = 0$ implies that flux v_1 is constrained to be twice flux v_2).

Mass balance: a requirement that the mass entering the system or any pathway within the system equals the mass exiting the system or pathway; a crucial characteristic of metabolic reconstructions from which functional mathematical models can be derived.

Metabolic network reconstruction: a manually-curated computational network of the metabolism of an organism with all the GPRs assembled from a functionally annotated genome, biochemical data, and literature, that are compiled into a stoichiometric matrix, which serves as the framework for further computational analysis.

Objective flux: the flux through the objective reaction

Objective reaction: a reaction which sets a demand for particular metabolites in the network. The typical goal in formulating this reaction is to simulate a biological objective, whether it be growth, energy, virulence, or a combination of other factors.

Reaction flux: moles consumed in a reaction per unit time.

Steady state: with respect to flux, a key assumption in FBA that the reaction fluxes within the system, and therefore the amounts of each metabolite, do not change over time, an assumption often justified by the very short time-scale of metabolic reactions compared to the time necessary for changes in cell phenotype.

Stoichiometric matrix (S matrix): a mathematical formalization of a metabolic reconstruction; a matrix in which each element contains the stoichiometric coefficient for a metabolite (row) participating in the corresponding reaction (column).

Corresponding author: Papin, J.A. (papin@virginia.edu)

* These authors contributed equally.

compartments specific to particular organisms) (Figure 1a) [5]. Since the initial genome-scale reconstructions of *Escherichia coli* [6,7] and *Haemophilus influenzae* [8,9], the metabolic networks of over 50 organisms (bacteria, archaea and eukaryotes) have been reconstructed (reviewed in [4]). Elements of this network reconstruction process have been automated, allowing the preliminary analysis of hundreds of draft network reconstructions [10]. Among these, metabolic networks have been reconstructed for several pathogenic organisms (Table 1). Indeed, the study of pathogen metabolism – for the elucidation of high-priority drug targets and metabolic factors contributing to pathogenicity – is an exciting application for metabolic network modeling and systems biology.

In this review, we explore several techniques and approaches used to predict antimicrobial drug targets from metabolic network modeling using FBA. Where possible we present examples that have led to novel data, drug targets, or drugs. Metabolic network modeling is still in its infancy, but has allowed for predictions that align with previous data and has provided many hypotheses that

continue to be developed. We first discuss the fundamental aspects of network analysis and FBA in particular. Subsequently, we delve into how computational metabolic reconstructions can be used to prioritize drug target predictions. Furthermore, we review recent developments on model-guided pipelines for drug target discovery against pathogens. Finally, we extend the discussion to include host cell metabolism and propose directions for future modeling efforts in infectious disease.

Reconstructing the metabolic network and defining an objective

A metabolic network reconstruction is assembled piece-by-piece by compiling data on known enzymes, genes encoding these enzymes, and the stoichiometry of the reactions catalyzed by these enzymes ([11] for a list of databases containing such data). Gene–protein–reaction (GPR) relationships, in the form of Boolean logic statements, define which genes are necessary for each enzyme and which enzymes are necessary for each reaction [5] (Figure 1b). The information for all the reactions in a network reconstruction with

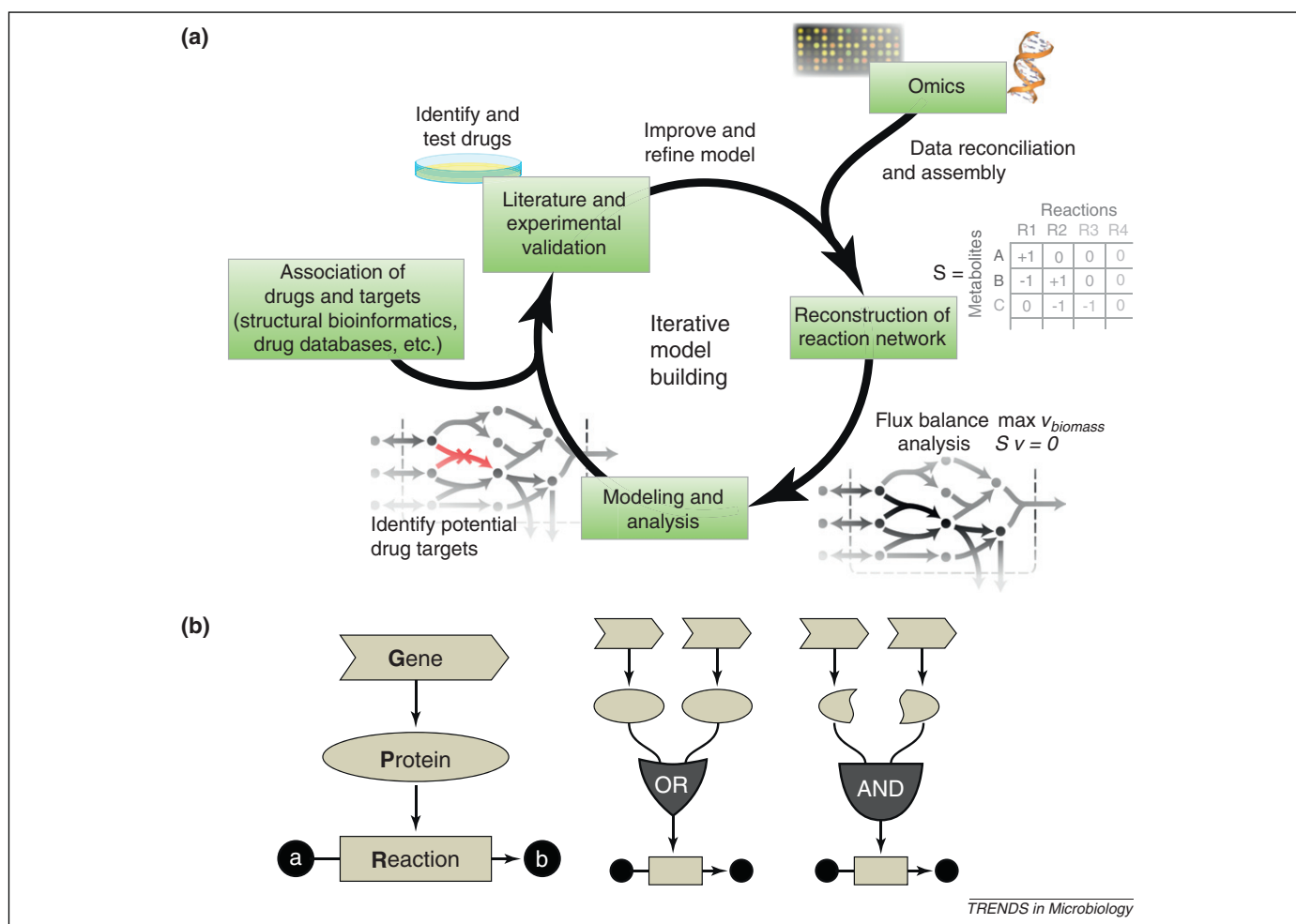


Figure 1. The iterative process of model building and refinement. (a) A functionally annotated genome together with data from the biochemical literature are used to assemble a network reconstruction. Flux balance analysis (FBA) allows for modeling and simulation of the reconstructed network. Advanced network analyses (such as gene essentiality or flux variability) allow for identifying potential anti-microbial drug targets. These targets can then be associated with drugs using bioinformatics approaches and obtaining target drug information from a variety of publicly available databases (e.g. STITCH or DrugBank). Predictions involving targets and drugs can be experimentally validated. Any discrepancies between computational predictions and experimental validation can be informative to improving upon and refining the original reconstruction and modeling platform. (b) Gene–protein–reaction (GPR) relationships, central to the assembly of a metabolic reconstruction, define the genes and gene products needed for each enzymatic reaction. Isozymes can be represented with ‘OR’ statements, whereas enzyme subunits required to function together to catalyze a particular reaction can be represented with ‘AND’ statements.

Table 1. Drug targeting-related analysis of pathogen metabolic networks

Pathogen	Disease	Drug targeting-related <i>in silico</i> analysis								Validation		Novel compounds identified	Refs
		Gene/reaction essentiality	Minimal media prediction	Conditional essentiality	Synthetic lethality	Flux variability analysis	Enzyme robustness	Metabolite essentiality	Correlated reaction sets	Literature-derived	Novel experimental validation		
<i>Acinetobacter baumannii</i>	Opportunistic; nosocomial infection	X						X		X			[24]
<i>Burkholderia cenocepacia</i>	Opportunistic; cepacia syndrome	X		X						X			[55]
<i>Francisella tularensis</i>	Tularemia	X		X		X	X			X	X		[21]
<i>Haemophilus influenzae</i>	Otitis media and respiratory infections	X	X		X					X			[8,9]
<i>Helicobacter pylori</i>	Gastritis; peptic ulceration; gastric cancer	X	X	X	X					X			[56,57]
<i>Klebsiella pneumoniae</i>	<i>Klebsiella</i> pneumonia; urinary tract infection										X		[58]
<i>Mycobacterium tuberculosis</i>	Tuberculosis	X		X	X	X	X		X	X	X		[1,2,19]
<i>Neisseria meningitidis</i>	Meningitis; meningococcal septicemia		X								X		[59]
<i>Porphyromonas gingivalis</i>	Periodontal disease	X	X							X			[17]
<i>Pseudomonas aeruginosa</i>	Opportunistic; nosocomial infection	X								X			[47]
<i>Salmonella Typhimurium</i>	Gastroenteritis; diarrhea	X			X	X				X	X		[60–62]
<i>Staphylococcus aureus</i>	Opportunistic; nosocomial infection	X	X		X					X	X		[63–65]
<i>Vibrio vulnificus</i>	Cellulitis; septicemia							X		X	X	X	[25]
<i>Yersinia pestis</i>	Bubonic, pneumonic, and septicemic plague	X	X		X					X			[66]
<i>Cryptosporidium hominis</i>	Cryptosporidiosis	X								X			[67]
<i>Leishmania major</i>	Leishmaniasis	X	X		X		X			X			[18]
<i>Plasmodium falciparum</i>	Malaria	X			X					X	X	X	[37,68]
<i>Trypanosoma cruzi</i>	Chagas disease	X				X				X			[69]

m metabolites and n reactions can be stored in an m by n table or matrix, the stoichiometric or S matrix. Each element or cell in this matrix corresponds to the stoichiometric coefficient of one particular metabolite in one particular reaction [5,12]. The S matrix enables strict accounting for the underlying biochemistry and allows a quantitative description of complex interactions between metabolites that are responsible for driving a cellular phenotype. This matrix

formalism facilitates interrogation of the structural and functional properties of the network.

The application of FBA to a network reconstruction results in the identification of combinations of reaction fluxes that correspond to a maximum flux through a targeted reaction (an objective) while requiring that constraints are satisfied, for example that the mass entering the network is equal to the mass exiting the

Box 1. A brief primer on FBA

Nutrient availability, restrictions on surrounding environmental pH, and temperature are examples of basic constraints that, when imposed upon cells, can affect the resulting phenotypes [12]. Such constraints can be mathematically described, and they serve to narrow the operating range of the cell and yield a set of feasible reaction fluxes for a metabolic network. In other words, the constraints (which can be physicochemical, topological, environmental and regulatory) restrict the number of possible phenotypes, which then allows the function of the biochemical network to be characterized [12,14].

Flux balance analysis (FBA) is one constraint-based method that has been extensively applied in the study of prokaryotic and eukaryotic metabolic networks [47,70]. First, reactions of a metabolic network are assembled into a stoichiometric or S matrix, whose elements correspond to the stoichiometric coefficients describing the conversions from reactants to products (Figure 1) [5,12]. This simple matrix formalism permits quantitative description of the complex interactions between metabolites. Following network reconstruction, the concentrations of metabolites and fluxes through reactions can be represented as follows (Equation I):

$$dC/dt = Sv \quad \text{[I]}$$

Here, C is a vector of concentrations of metabolites, t is time, S is the stoichiometric matrix consisting of m rows of metabolites and n columns of reactions, and v is a vector of fluxes through the corresponding reactions. Invoking the steady-state assumption so

that the rate of production of every metabolite equals its rate of consumption yields the following (Equation II):

$$Sv = 0 \quad \text{[II]}$$

Limits can be applied to individual fluxes as follows:

$$v_{\min} \leq v \leq v_{\max}$$

Particular reactions can have set upper limits (v_{\max}) that may align closely with experimental enzyme capacity measurements, whereas other irreversible reactions will have v_{\min} set to 0.

The principle physicochemical constraint in Equation II represents a set of linear equations. The number of equations is usually less than the number of unknown fluxes [12], and a linear programming optimization can be used to generate a single flux distribution for the entire network by choosing an appropriate objective (or goal): $\text{Max } v_{\text{biomass}}$.

Traditionally, maximization of biomass has been chosen as the objective of choice in FBA [12]. A set of metabolites (e.g. amino acids, lipids, nucleotides and carbohydrates) that are necessary for the cell or organism to grow are typically included in the biomass reaction.

Therefore, the optimization problem can be summarized as simply:

$$\begin{array}{ll} \text{Max } v_{\text{biomass}} \\ \text{Subject to} & Sv = 0 \\ & v_{\min} \leq v_i \leq v_{\max} \quad \text{for all } i \end{array}$$

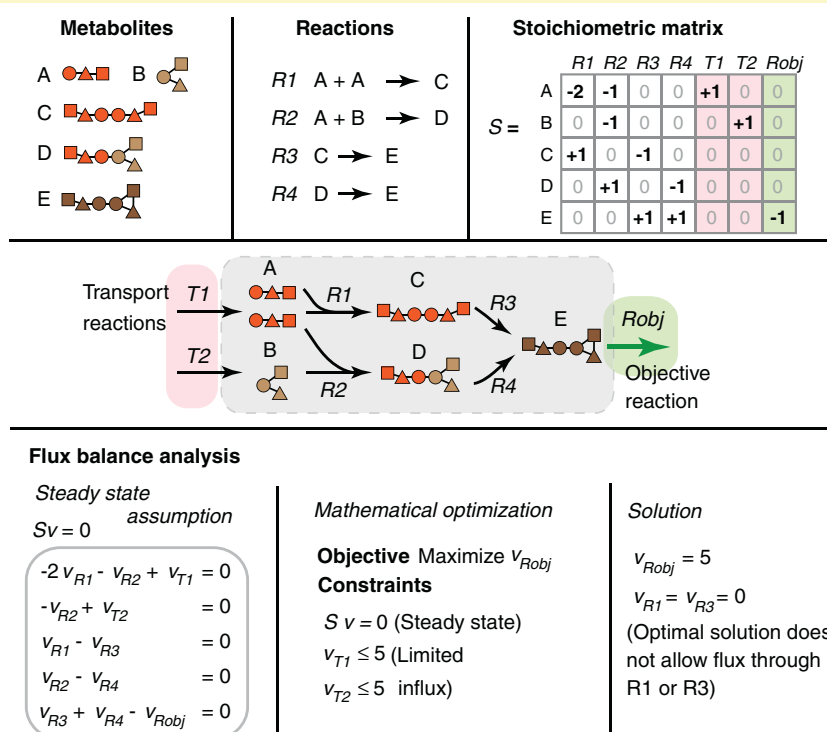


Figure 1. A generic metabolic model.

network. In more mathematical terms, FBA involves the use of a linear programming formulation wherein an objective is optimized subject to a set of governing constraints (Box 1) [4,12,13]. In addition to requiring mass balance for every reaction, thermodynamic, topological, environmental, and regulatory data may provide additional constraints that dictate the feasible flux space [14]. An objective often used with FBA is biomass production, which is represented *in silico* as a drain capturing crucial metabolites necessary for growth of the organism [15,16]. With the ultimate goal of identifying antimicrobial drug targets (and associated drugs) to slow or stop the growth of a pathogen, a biomass objective is often very appropriate for computational modeling efforts. Network reconstructions of most pathogenic organisms have incorporated biomass reactions as their objectives (Table 1).

The inability of the metabolic network model to synthesize even one metabolite of the biomass reaction will result in a predicted value of zero for the objective (biomass) flux, and analogously no growth. Therefore, growth predictions are sensitive to metabolites that are placed in the biomass reaction. FBA can be used to investigate the ability of the model to produce each metabolite within biomass. In the *Porphyromonas gingivalis* metabolic network, the ability of the model to produce each of the 52 metabolites within the biomass objective was evaluated after systematic reaction deletions [17]. Crucial groups of reactions were identified that were responsible for lipopolysaccharide (LPS) production, coenzyme A production, glycolysis, or purine and pyrimidine biosynthesis [17]. Corresponding enzymes that are essential for growth, as well as for the production of important bacterial components such as LPS, could serve as potential drug targets. In addition, in a study of *Leishmania major* metabolism, the contribution of minimal media components to the synthesis of individual biomass constituents was analyzed [18]. The study found that the absence of cysteine and oxygen in the minimal media had a drastic impact on the overall metabolic network, limiting the generation of 30 of the 40 biomass constituents [18]. Network analyses as delineated in these examples of

P. gingivalis and *L. major* may permit the formulation of a hypothesis for the role of specific metabolites and their influence on growth. Therefore, defining an appropriate biomass reaction (Boxes 2 and 3) is crucial for useful predictions and identification of vulnerable parts of the metabolism of a pathogen.

The quest for drug targets in metabolic networks of pathogens

Gene essentiality analysis

The most common method to identify potential drug targets has been through the prediction of essential genes (Table 1). The enzymes encoded by essential genes are typically hypothesized as drug targets. Gene knockouts might lead to a redistribution of flux through the network if the perturbed gene or gene product affects the removal of a particular flux-carrying reaction [4]. A GPR aids in mapping the effects of a genetic (or pharmacological) perturbation on the associated reactions, and thus the network. Gene-level perturbations that result in reduced or zero flux through a biomass reaction correspond to growth-reducing or lethal gene knockouts, respectively (Figure 2). For example, in a reconstruction of *Mycobacterium tuberculosis*, five previously known drug targets were encoded by genes predicted to be essential from computational analysis [19]. For metabolic reconstructions of pathogens, enzymes that are predicted to be essential will offer new experimental hypotheses and avenues for drug discovery. In the following subsections we discuss several other approaches for drug targeting using metabolic network analysis. These other approaches may provide a separate list of potential targets; however, they can also be used in tandem with gene essentiality analysis for step-by-step prioritization of drug targets.

Enzyme robustness and flux variability

In addition to an analysis of gene essentiality, assessing enzyme robustness could identify vulnerable or sensitive portions of a metabolic network suitable for drug targeting. To determine the robustness of a metabolic network to the

Box 2. Cellular objectives of pathogenic organisms

A biomass objective reaction is not applicable under all conditions, and very often growth alone may not be a realistic objective. Other objectives such as maximizing or minimizing ATP or maximizing the production of particular cellular by-products (e.g. lactate or pyruvate) can also be used.

The metabolism of an organism may be adapted for increased virulence or pathogenicity. Pathogens and host cells may also temporarily opt for alternative objectives while under selective pressures (e.g. changes in nutrients or environmental influences of secreted toxins). In addition, different morphological stages of a pathogen (e.g. the sporozoite stage of *Plasmodium falciparum* in mosquitoes vs the merozoite stage in humans) may be characterized by varying metabolic requirements. Consequently, the objective function must be appropriately defined to find relevant enzyme targets crucial in particular stages of infection and/or environmental conditions.

To explore the effects of targeted perturbations (pharmacological or genetic), different objectives can be explored. It may be that the evolutionary pressures that dictate wild-type cells are different for knockout mutants. Additionally, mutants may not have the ability for immediate regulation of fluxes that allow for optimal growth. Based

on these ideas, an approach termed minimization of metabolic adjustment (MOMA) was developed. The requirement for optimal growth is relaxed for gene deletions in MOMA. Instead, MOMA assumes that the overall flux distribution of a gene-knockout mutant will probably not change significantly from that of the corresponding wild type [71]. In terms of flux values, the gene-knockout mutant will remain as close as possible (in Euclidean distance) to the wild-type optimal flux state. MOMA aided in correcting gene essentiality predictions associated with knockouts of fructose-1,6-bisphosphate aldolase, triosephosphate isomerase and phosphofructokinase in the *E. coli* central metabolic model [71]. These genes were predicted to be nonessential when biomass was used as the objective for *E. coli* growth on glucose, which was inconsistent with supporting literature evidence. MOMA yielded a suboptimal flux distribution for a knockout mutant that would not necessarily equal the optimum as dictated by traditional FBA [71].

Approaches such as MOMA that consider alternate hypotheses for the objective of metabolic networks provide a basis for understanding a potential biological goal for pathogenic organisms, especially considering the complexity of the environment surrounding the pathogen of interest.

Box 3. Knowledge gaps and caveats to metabolic network analysis

Multiple steps in the model-building process and subsequent analyses are prone to errors that may greatly affect flux and growth predictions and, consequently, predicted drug targets. FBA provides a quality-assurance check that ensures mass balance. Growth rates and gene essentiality predictions are validated against experimental data to ensure the model truly reflects biological processes. Gene essentiality predictions (commonly between 55% and 90%) provide confidence that downstream analyses are based on a high-quality model [1,72]. To ensure the usefulness of any computational pipeline, drug target predictions should be compared to known targets from the literature. Below we discuss more of the particular difficulties and limitations encountered when using metabolic models to identify drug targets.

Genome annotation

In any metabolic reconstruction, there may be hundreds of putative metabolic enzymes with no experimentally identified function. These reactions and associated GPR relationships may be assembled strictly based on existing functional annotations of the genome or based on evidence from related organisms. Even some very well characterized enzymes may have other unexpected activities. In such cases, misannotated enzymes may yield incorrect model predictions leading to errors in drug targeting. For instance, an enzyme may be incorrectly predicted to be essential if the activity of another enzyme, which is not included in the network, can account for the same function. Therefore, this represents one important knowledge gap in the assembly of metabolic networks, and the inclusion of more refined enzyme annotations will directly improve drug target predictions [73,74].

Nutrient availability

An important challenge in reconstructing and modeling metabolic networks is determining the composition of *in silico* media. Nutrients available in host environments are poorly characterized. In addition, for any nutrients that are identified, quantitative data on uptake rates are unavailable. Knowledge of the transporters of an organism elucidates which metabolites are transported into the cell. However, information on transporters is particularly limited and, in general, transport reactions lack any experimental evidence or gene associations

supporting their presence. Instead, transport reactions are added for proper functioning of the computational model. Because model predictions are dependent on the media environment, nutrients and transporters must be carefully defined.

Objective function

As stated previously, in FBA-based modeling of microbial organisms, a biomass objective reaction has often been used. The purpose of the reaction is to ensure a drain of metabolites that are deemed essential to support the growth of the organism. Starting with the estimated weight fraction of important macromolecular components of the cell (e.g. protein, lipid, RNA, DNA and carbohydrate), the relative abundance of metabolites comprised in each group (e.g. amino acids, phospholipids, nucleotides) can be computed [15]. Among several available, a few experimental methods to measure biomass components include chloroform–methanol extraction (lipids), colorimetric protein assays, and gas chromatography–mass spectrometry (protein content) [75]. Fluxes measured directly by metabolic flux analysis, which tracks the movement of ^{13}C from an initial ^{13}C -labeled substrate, may also help to define the objective function [76,77].

An objective function is most likely to cause errors if metabolites are entirely missing or incorrectly included; the relative amounts of each metabolite in the objective reaction (i.e. the stoichiometric coefficients) do not greatly affect FBA results [16,78]. Therefore, logically deducing a biomass objective is often adequate to estimate the growth of an organism. However, additional experimental data can reveal interesting peculiarities that may be used to design a specific biomass objective for a particular organism. Data on growth- and non-growth-associated ATP maintenance can be included in the biomass reaction [79]. Under *in vivo* conditions when a pathogen is interacting with its host, an aspect that is often unclear is which biomass component(s) to include in the reaction. The composition of the biomass reaction is likely to vary under different physiological conditions, and the choice of metabolites can directly influence model predictions regarding drug targets. For example, failure to include a particular cell-wall component will not necessarily direct flux through reactions that may be crucial *in vivo*, and therefore the associated enzymes will not be targeted.

inhibition of an enzyme-catalyzed reaction, the flux of a reaction may be constrained to a fraction of its wild-type flux (simulating partial to complete inhibition), and the effects on an objective flux (e.g. biomass production) can be evaluated [20] (Figure 2). Such an approach was used in analyzing the network reconstruction of *Francisella tularensis*, which revealed that the growth rate was sensitive to changes in H^+ and NH_4^+ flux in a simulated *in vitro* medium but not in a simulated *in vivo* medium that mimicked the environment during infection [21]. Analyzing enzyme robustness with different constraints provides a detailed view of possibly non-lethal reactions whose change in flux has a strong effect on the objective of a network under different environmental conditions, therefore suggesting important reactions during an infection or perturbations for drug targeting.

Alternatively, the objective function could be constrained to a fixed percentage of its wild-type flux and the allowable range of flux for each reaction can be determined using an approach termed flux variability analysis (FVA; Box 4) [22]. A recognized shortcoming of FBA is that the typical implementation calculates only one of many possible solutions that optimize flux through the objective reaction. Consequently, there may be many possible routes through the network that achieve the same optimal flux for a given objective [14,22]. In an effort to circumvent this shortcoming, FVA was developed to determine the range of

fluxes over which a particular reaction operates, while still allowing for optimal, or near-optimal, objective flux [22]. FVA also identifies blocked reactions (reactions incapable of carrying any flux in a given model under specified constraints) or reactions with different flux ranges in various media. Enzymes that catalyze reactions with little

Box 4. A brief primer on FVA

After the application of various constraints on the biochemical network, the number of allowable network states (or possible phenotypes) is typically large. Depending on the size and inter-connectedness of the cellular network, there may be several alternative optimal phenotypes [14]. FBA calculates one of many feasible solutions that result in the same optimal value of the cellular objective.

Flux variability analysis (FVA) calculates the range of flux in each reaction that allows for the same optimal flux through the objective reaction. The objective flux of the reaction is specified as an additional constraint and multiple optimizations are performed to compute the maximum and minimum flux for every reaction in the network:

$$\begin{array}{ll} \text{Max/Min} & v_i \\ \text{Subject to} & Sv = 0 \\ & v_{\text{biomass}} = Z_{\text{obj}} \\ & v_{\text{min}} \leq v_i \leq v_{\text{max}} \quad \text{for } i = 1 \dots n \end{array}$$

Here, n refers to the number of reactions in the biochemical network, and Z_{obj} is the optimal value of the cellular objective (v_{biomass}) as obtained by FBA [22].

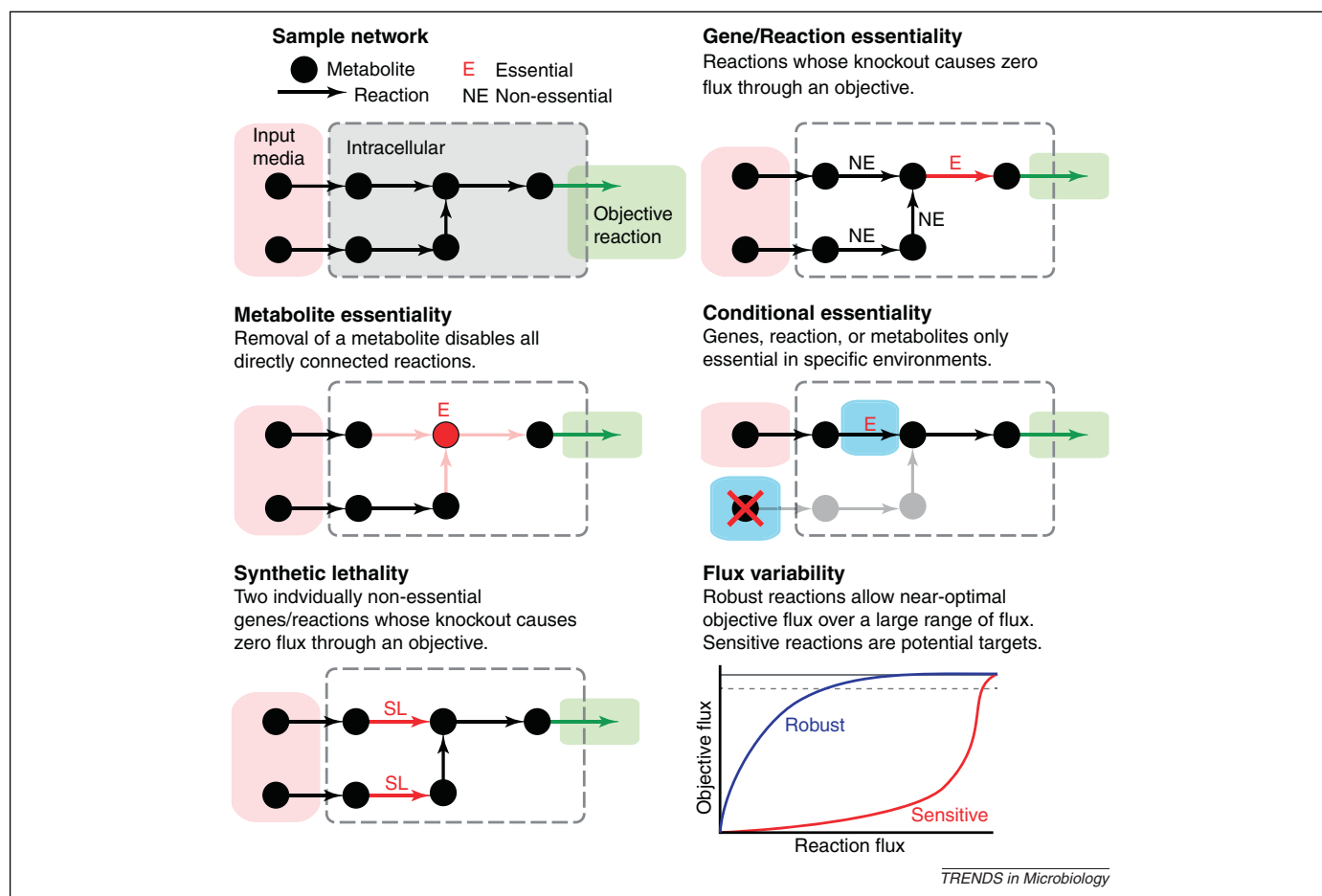


Figure 2. Drug targeting in metabolic networks. Various strategies are illustrated for identifying drug targets by performing FBA on metabolic reconstructions. The sample network shows an input media that represents the environment and exchange reactions, intracellular reactions, and an objective reaction that drains metabolites out of the system. An essential reaction and metabolite that, when removed, block any flux through the objective are highlighted in red. In the conditionally essential panel, the absence of the metabolite highlighted in blue causes the highlighted reaction to become essential in the selected media. One of the synthetic lethal pairs of the network is denoted by 'SL'. The dashed line in the flux variability illustration may represent 'near-optimal' objective flux. A robust reaction maintains near-optimal objective flux over a larger range of reaction fluxes.

to no variability in flux for a given objective could be selected as potential drug targets given that the network may be sensitive to even modest inhibition of its activity. In analyzing the reconstruction of *M. tuberculosis*, the average of the highest and lowest allowable flux for each reaction was used as a point of comparison for the model under constraints for two growth rates [19]. In the study, the reaction catalyzed by isocitrate lyase – an enzyme important for persistence in a host – was predicted to have increased flux during slow growth, and the activity of the enzyme was experimentally confirmed to be greater in slow-growing cells of the closely related *Mycobacterium bovis* [19]. Hence, flux variability and enzyme robustness aid in further prioritizing drug targets identified from other methods such as gene essentiality.

Metabolite essentiality

Many current drugs have high similarity to natural metabolites and compete for and/or inhibit normal enzymatic activity [23]. Therefore, another very interesting avenue towards identifying drug targets in metabolic networks is via the prediction of essential metabolites. In contrast to the more traditional individual gene knockouts mentioned previously, metabolites in the S matrix can also be systematically removed. Consequently, all reactions in which

a given metabolite participates are removed, and the resultant effects on the objective are assessed (Figure 2). A total of 211 essential metabolites in the *Acinetobacter baumannii* reconstruction were narrowed to 9 following removal of (i) currency metabolites (ATP, NADH, H₂O, etc.); (ii) metabolites present in the human metabolic network; and (iii) metabolites participating in reactions catalyzed by enzymes with human homologs [24]. Enzymes that catalyze reactions involved in the consumption or production of these essential metabolites may be considered as drug targets. Moreover, structural analogs of essential metabolites may be considered as test compounds for experimental evaluation, thus sidestepping extensive screening or computational predictions [25].

Combination gene and reaction perturbations

Many anti-infectives on the market act on multiple targets. This multiplicity of targets was exemplified in a network analysis of 890 FDA-approved drugs targeting 394 human proteins (derived from the DrugBank database), where approximately 38% of the drugs were associated with more than one protein target, and a few drugs were associated with as many as 14 targets [26]. In that regard, a drug-discovery strategy incorporating compounds known to act simultaneously on multiple targets can be adopted for

microbial pathogens. FBA allows predictions involving the perturbation of multiple genes or reactions in a rapid time-frame (minutes or hours) – a single *in silico* combination taking only a fraction of a second [27]. In a reconstruction of *M. tuberculosis* that accounted for features specific to an *in vivo* environment, all non-trivial double-deletion mutants (synthetic lethal pairs) were tested for *in silico* growth using FBA [2]. Of two experimentally-characterized double gene deletions, the model accurately predicted reduced *in silico* growth in both *in vitro* and *in vivo* environmental conditions [2]. A combination of drugs that affect synthetic lethal targets may act synergistically to inhibit growth of a pathogen, thereby paving the way for model-guided predictions of drug synergy. Experimentally screening for all possible drug combinations against a particular pathogen is costly and is often not feasible. Therefore, predicted combinations of drugs associated with synthetic lethal targets can direct more specific experiments and perhaps reveal entirely novel treatment strategies.

Groups of targets and network topology

Other approaches characterizing the structure of a genome-scale network reconstruction identify sets of reactions that act together and therefore may be targeted as an entire pathway. Sets of correlated reactions (or Co-sets) consist of groups of reactions whose fluxes are linked, and which represent functional modules within a biochemical network [28]. Co-sets, which can aid in suggesting alternative drug targets by identifying reactions that are functionally related to each other, can be divided into several categories. A perfect Co-set consists of a group of reactions such that, for any given pair in the group, a non-zero flux in one reaction implies a non-zero flux in the other, with a fixed ratio [29]. Other categories include partial Co-sets (pairs of linked reactions, but with a variable flux ratio) and directional Co-sets (a non-zero flux in one reaction implies a non-zero flux in the other, but the converse is not true) [29]. By calculating hard-coupled reaction (HCR) sets – a subgroup of perfect Co-sets where sets of reactions are defined by participating metabolites sharing a consumption to production connectivity of 1:1 (i.e. two reactions are linked by a metabolite that is connected to no other reaction) – one study found 25 of 147 HCR sets contained previously identified drug targets in *M. tuberculosis* [1]. Because an altered flux in one reaction results in an altered flux of all reactions within a HCR set, only one enzyme needs to be targeted. This approach aids in prioritizing the list of potential drug targets by identifying linked enzymatic reactions. Hence, analyses of the topology of a metabolic network can reveal key local and structural features that may be important drug targets when data related to environment fluxes or appropriate objectives are not necessarily available.

Environment and conditional essentiality

Finally, the metabolic phenotype is dependent on the media environment and the exchange of metabolites into and out of the system (see Box 3 for discussion of knowledge gaps associated with nutrient availability). By constraining uptake or secretion fluxes, a minimal set of metabolites that allows flux through the objective can be computed [18]. Moreover, enzymes or metabolites that are necessary for

growth in various environments (e.g. minimal media, defined media, or rich media with an abundance of nutrients and carbon sources) can be predicted. In the reconstruction of *F. tularensis*, genes that were essential in a simulated macrophage environment and five other environmental conditions were considered to be unconditionally essential genes (in other words, they represented a core set of genes that were essential regardless of the medium). Of the 17 virulence factors cataloged from previous literature, eight were unconditionally essential genes [21]. By contrast, enzymes that may be necessary for growth under one condition, but not another, are conditionally essential (Figure 2). With a careful consideration of an objective and appropriate nutrient uptake, gene essentiality analysis may reveal new drug targets specific to particular growth conditions and environments in which a pathogen must survive. Such an analysis can also inform strategies for manipulating the environment of the pathogen that could be effective as a treatment option.

From target to drug and the development of model-guided pipelines for drug discovery

Computational analysis of metabolic processes in pathogens can yield a ranking of predicted drug targets. Bioinformatics and network analyses were performed to yield a high-confidence list of targets against *M. tuberculosis* [30]. By implementing a multilayered approach, targets that did not pass sequential cut-off values were removed (e.g. elimination of enzymes with human homologs or targets with no computationally predicted binding pocket) [30]. In another proof-of-concept study it was noted that essential type II fatty acid biosynthesis (FAS II) reactions in the *E. coli* MG1655 metabolic network were also essential in several *Staphylococcus aureus* strains [31]. Following network analysis, a virtual screening strategy was employed whereby small molecules from a library of approximately 10^6 compounds were docked to enzymes catalyzing essential reactions, and 41 inhibitors of FAS II enzymes were selected for experimental validation [31]. In cell viability assays, six of the inhibitors had growth-retarding effects against *E. coli* and *S. aureus* strains in standard LB agar plates [31]. Finally, following the identification of 163 essential metabolites, a third study used a layered approach to prioritize five essential metabolites in the metabolic network of the opportunistic pathogen *Vibrio vulnificus* [25]. Currency metabolites, metabolites consumed by a single reaction, metabolites present in the human metabolic network, and metabolites associated with enzymes with human homologs were removed. The study screened 352 compounds found to be structurally similar to one of the five essential metabolites and identified one compound that most potently inhibited growth, more so than a currently used drug [25]. These studies provide various examples of model-guided pipelines to drug discovery by primarily using network analyses to identify and prioritize drug targets. Additional constraints such as enzyme druggability and elevated gene expression can also be used to prioritize drug targets, which can then guide the screening and selection of compounds.

A common approach in proposing drug targets using metabolic networks has been to rule out targets that

overlap with host cell metabolism – the idea being that off-targets can be minimized and drug interference plus subsequent complications with the host can be avoided. However, there are arguments to be made in favor of retaining targets that overlap with human metabolism. First, accounting for the drug selectivity between host and pathogen targets at the respective binding sites may preclude off-target influences [32]. Second, if the goal is to discover drugs against infectious diseases quickly, then the best option may be to focus on finding new clinical indications for existing FDA-approved drugs (i.e. pursuing drug repurposing strategies) instead of developing new investigational compounds that are subject to regulatory hurdles [33]. Also, the majority of FDA-approved drugs target human proteins. Hence, eliminating pathogen targets that overlap with human proteins reduces the number of potential drugs that could be evaluated experimentally.

A host cell perspective

Interaction with a host cell is often crucial to the metabolism and survival of a pathogen. For instance, the kinetoplastid parasite *L. major* is unable to synthesize several essential amino acids and therefore obtains them from the host macrophage [34]. As another example, *Legionella pneumophila*, the bacterium responsible for Legionnaire's disease, ceases to replicate inside a host macrophage when it cannot access or process threonine [35]. Consequently, identifying the particular niche of nutrients and resources in the host cell required by a pathogen is vital to discovering treatment options that specifically target host–pathogen interactions [36].

A systems-level analysis of pathogen metabolism interfaced with cell type-specific host metabolic networks can also be conducted. Because *P. falciparum* invades mature erythrocytes to establish infection in its human host, a metabolic model of the human erythrocyte was built in conjunction with the *P. falciparum* reconstruction to make predictions which aligned closer to known conditions in the infected erythrocyte [37]. This model, modifying an existing approach from Shlomi *et al.* [38], integrated previous gene expression data where several enzymes were constrained to be 'on' and 'off' during specific life-cycle stages. The combined erythrocyte-blood stage *P. falciparum* network correctly predicted metabolite exchanges between the microbe and host [37]. A recently active area of research has been the development of algorithms to create cell type-specific metabolic networks by integrating gene and protein expression data with existing human metabolic reconstructions [38–40]. Inclusion of host-specific factors into pathogen metabolic-network reconstructions or developing systems-level models of host and pathogen networks will continue to enable investigations into the complexities of the host–pathogen interplay.

Finally, inhibiting host pathways and perturbing the flux of metabolites in the host cell may alter the ambient environment and require pathogens to adapt their metabolic needs. Therefore, targeting the machinery of a host cell at the host–pathogen interface can provide new therapeutic approaches [41]. For example, host proteins hijacked for viral replication are potentially important drug targets. Recently, a high-throughput screening assay

identified a lipophilic compound – NA255 – that inhibits the host serine palmitoyltransferase, an enzyme needed for association of hepatitis C virus (HCV) with host lipid rafts [42]. Moreover, to characterize transformations in host functions, data specific to the host cell pre- and post-infection must be obtained. The analysis of transcription profiles is one approach that has been successfully implemented to identify genes in the host cell that are differentially regulated due to pathogenic infection [43,44]. Similarly, profiling the proteome and lipidome of a hepatocyte over the time-course of infection and integrating these data with protein–protein interaction networks revealed multiple lipids and enzymes differentially regulated in HCV-infected cells [45]. A third approach for identifying factors in the host necessary for establishing infection involves the use of genetic screens in which large-scale insertional mutagenesis is performed to develop null mutants in a human cell line [46]. Ultimately, the use of new experimental technologies along with metabolic modeling will be vital to discovering host components crucial to the survival of a pathogen.

Next steps

Advanced meta-network analyses, such as comparative modeling of metabolic reconstructions across multiple strains and species or community-based modeling of metabolism across differing pathogenic organisms, have broad implications for understanding and investigating infectious diseases. Below, we highlight several future directions in this realm and provide a few examples of efforts already underway.

The recent completion of metabolic reconstructions of the pathogen *Pseudomonas aeruginosa* [47] and the related non-pathogen *Pseudomonas putida* [48] creates new opportunities for investigating species-specific differences in metabolism and the metabolic basis for virulence of *P. aeruginosa*. Towards that end, a reconciliation of the two reconstructions was completed such that any differences in the metabolic networks of *P. aeruginosa* and *P. putida* would be indicative of true biological variations as opposed to artifacts of the reconstruction and modeling process [49]. In the reconciliation study, the model for each organism was analyzed to characterize the tradeoffs of producing biomass versus the production of individual metabolites. Compared to *P. putida*, *P. aeruginosa* was able to produce a small proportion of the shared virulence factor precursors with only a slight decrease in biomass production. In general, the metabolic flexibility analysis suggested that the virulence of *P. aeruginosa* is complex and highly multifactorial, and has more flexibility than *P. putida* in many metabolic pathways. This computational analysis paves the way for future modeling efforts of other infectious disease-causing agents and their basis for establishing virulence.

As another example, syntrophic mutualism between a sulfate-reducing bacterium, *Desulfovibrio vulgaris*, and a methanogen, *Methanococcus maripaludis*, was investigated by performing FBA on a compartment-based model involving the metabolic reconstructions of both organisms and a culture medium [50]. In another study, gene expression data were integrated with the genome-scale metabolic reconstruction of *P. aeruginosa* in the context of a chronic

cystic fibrosis lung infection over a 44-month time-course [51]. This analysis provided a systems-level view of bacterial adaptations in a cystic fibrotic lung environment over time. Subsequent studies can shed light on the interactions between multiple organisms over the time course of an infection.

Finally, automated reconstruction platforms such as ModelSEED permit the rapid reconstruction of hundreds of draft bacterial metabolic networks [52,53]. Integration of many such reconstructed networks may help elucidate interactions within the host microbiome and partially explain the development of opportunistic infections that occur primarily because of an altered bacterial flora and environment. For example, an integrative metabolic analysis of organisms in the human gut microbiome could aid understanding of the intricate balance between non-pathogenic and potentially pathogenic organisms during healthy and infectious states in the gastrointestinal tract. An expected outcome of such analyses could result in the selection of drugs or drug cocktails that specifically target pathogens without eliminating non-pathogenic members.

Concluding remarks

As reconstructions of metabolic networks become more standard and automated [5], the need for computational tools to characterize these networks becomes more apparent. In addition, the generation and management of large datasets pertaining to both host cell and pathogen intracellular processes of metabolism, signal transduction or regulation has necessitated a systems approach and, therefore, the computational methods used to analyze these data are becoming increasingly important. Experimental methods will continue to improve, thereby generating data that have so far been either impossible or prohibitively laborious to obtain, and which have constrained the value of some model predictions. For example, TraDIS (a new experimental method used to identify all essential genes simultaneously) directly measures gene essentiality that the model could only predict [54]. However, the iterative relationship between modeling and experiment will always permit the generation of novel hypotheses and the contextualization of large datasets, often in a quicker and more cost-efficient manner (e.g. rapid essentiality prediction of all double gene knockouts). Network-based approaches such as genome-scale metabolic reconstructions have been effective in drug target prediction and will continue to expand in scope and applicability. In addition, integration of networks and data into more standard pipelines that traverse the spectrum from computational prediction to experimental evaluation and back again will speed the process dramatically. By including many types of data sources that have yet to be coupled, entirely new classes of drug targets or treatment strategies may be found. The already enormous amount of data is only increasing, and the use of systems biology approaches will be vital to driving future research of drug and drug target discovery against infectious diseases.

Acknowledgments

The authors are grateful to Glynis Kolling and Erwin Gianchandani for reading the manuscript and providing insightful feedback. The authors

wish to acknowledge funding from the National Science Foundation CAREER program (grant number 0643548 to J.P.), National Institute of Allergy and Infectious Diseases (5U54 AI057168 to E.H. and J.P.), National Institutes of Health (R01 GM088244 to J.P.), and Cystic Fibrosis Research Foundation (grant 1060 to J.P.). K.M.D. is a trainee on the Infectious Diseases Training Grant to the University of Virginia Division of Infectious Diseases and International Health (5T32 AI0070465).

References

- Jamshidi, N. and Palsson, B.O. (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
- Fang, X. *et al.* (2010) Development and analysis of an *in vivo*-compatible metabolic network of *Mycobacterium tuberculosis*. *BMC Syst. Biol.* 4, 160
- Segovia-Juarez, J.L. *et al.* (2004) Identifying control mechanisms of granuloma formation during *M. tuberculosis* infection using an agent-based model. *J. Theor. Biol.* 231, 357–376
- Oberhardt, M.A. *et al.* (2009) Applications of genome-scale metabolic reconstructions. *Mol. Syst. Biol.* 5, 320
- Thiele, I. and Palsson, B.O. (2010) A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nat. Protoc.* 5, 93–121
- Edwards, J.S. and Palsson, B.O. (2000) The *Escherichia coli* MG1655 *in silico* metabolic genotype: its definition, characteristics, and capabilities. *Proc. Natl. Acad. Sci. U.S.A.* 97, 5528–5533
- Reed, J.L. *et al.* (2003) An expanded genome-scale model of *Escherichia coli* K-12 (iJR904 GSM/GPR). *Genome Biol.* 4, R54
- Edwards, J.S. and Palsson, B.O. (1999) Systems properties of the *Haemophilus influenzae* Rd metabolic genotype. *J. Biol. Chem.* 274, 17410–17416
- Schilling, C.H. and Palsson, B.O. (2000) Assessment of the metabolic capabilities of *Haemophilus influenzae* Rd through a genome-scale pathway analysis. *J. Theor. Biol.* 203, 249–283
- Henry, C.S. *et al.* (2010) High-throughput generation, optimization and analysis of genome-scale metabolic models. *Nat. Biotechnol.* 28, 977–982
- Durot, M. *et al.* (2009) Genome-scale models of bacterial metabolism: reconstruction and applications. *FEMS Microbiol. Rev.* 33, 164–190
- Lee, J.M. *et al.* (2006) Flux balance analysis in the era of metabolomics. *Brief Bioinform.* 7, 140–150
- Gianchandani, E.P. *et al.* (2010) The application of flux balance analysis in systems biology. *Wiley Interdiscip. Rev. Syst. Biol. Med.* 2, 372–382
- Price, N.D. *et al.* (2004) Genome-scale models of microbial cells: evaluating the consequences of constraints. *Nat. Rev. Microbiol.* 2, 886–897
- Feist, A.M. and Palsson, B.O. (2010) The biomass objective function. *Curr. Opin. Microbiol.* 13, 344–349
- Varma, A. and Palsson, B.O. (1993) Metabolic capabilities of *Escherichia coli*. 1. Synthesis of biosynthetic precursors and cofactors. *J. Theor. Biol.* 165, 477–502
- Mazumdar, V. *et al.* (2009) Metabolic network model of a human oral pathogen. *J. Bacteriol.* 191, 74–90
- Chavali, A.K. *et al.* (2008) Systems analysis of metabolism in the pathogenic trypanosomatid *Leishmania major*. *Mol. Syst. Biol.* 4, 177
- Beste, D.J. *et al.* (2007) GSMN-TB: a web-based genome-scale network model of *Mycobacterium tuberculosis* metabolism. *Genome Biol.* 8, R89
- Edwards, J.S. and Palsson, B.O. (2000) Robustness analysis of the *Escherichia coli* metabolic network. *Biotechnol. Prog.* 16, 927–939
- Raghunathan, A. *et al.* (2010) Systems approach to investigating host-pathogen interactions in infections with the biothreat agent *Francisella*. Constraints-based model of *Francisella tularensis*. *BMC Syst. Biol.* 4, 118
- Mahadevan, R. and Schilling, C.H. (2003) The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. *Metab. Eng.* 5, 264–276
- Dobson, P.D. *et al.* (2009) 'Metabolite-likeness' as a criterion in the design and selection of pharmaceutical drug libraries. *Drug Discov. Today* 14, 31–40
- Kim, H.U. *et al.* (2010) Genome-scale metabolic network analysis and drug targeting of multi-drug resistant pathogen *Acinetobacter baumannii* AYE. *Mol. Biosyst.* 6, 339–348

- 25 Kim, H.U. *et al.* (2011) Integrative genome-scale metabolic analysis of *Vibrio vulnificus* for drug targeting and discovery. *Mol. Syst. Biol.* 7, 460
- 26 Yildirim, M.A. *et al.* (2007) Drug-target network. *Nat. Biotechnol.* 25, 1119–1126
- 27 Becker, S.A. *et al.* (2007) Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox. *Nat. Protoc.* 2, 727–738
- 28 Papin, J.A. *et al.* (2004) Hierarchical thinking in network biology: the unbiased modularization of biochemical networks. *Trends Biochem. Sci.* 29, 641–647
- 29 Xi, Y. *et al.* (2010) Comparative study of computational methods to detect the correlated reaction sets in biochemical networks. *Brief Bioinform.* 12, 132–150
- 30 Raman, K. *et al.* (2008) targetTB: a target identification pipeline for *Mycobacterium tuberculosis* through an interactome, reactome and genome-scale structural analysis. *BMC Syst. Biol.* 2, 109
- 31 Shen, Y. *et al.* (2010) Blueprint for antimicrobial hit discovery targeting metabolic networks. *Proc. Natl. Acad. Sci. U.S.A.* 107, 1082–1087
- 32 Hopkins, A.L. (2008) Network pharmacology: the next paradigm in drug discovery. *Nat. Chem. Biol.* 4, 682–690
- 33 Chong, C.R. and Sullivan, D.J., Jr (2007) New uses for old drugs. *Nature* 448, 645–646
- 34 McConville, M.J. *et al.* (2007) Living in a phagolysosome; metabolism of *Leishmania amastigotes*. *Trends Parasitol.* 23, 368–375
- 35 Sauer, J.D. *et al.* (2005) The phagosomal transporter A couples threonine acquisition to differentiation and replication of *Legionella pneumophila* in macrophages. *Proc. Natl. Acad. Sci. U.S.A.* 102, 9924–9929
- 36 Brown, S.A. *et al.* (2008) Revisiting the host as a growth medium. *Nat. Rev. Microbiol.* 6, 657–666
- 37 Huthmacher, C. *et al.* (2010) Antimalarial drug targets in *Plasmodium falciparum* predicted by stage-specific metabolic network analysis. *BMC Syst. Biol.* 4, 120
- 38 Shlomi, T. *et al.* (2008) Network-based prediction of human tissue-specific metabolism. *Nat. Biotechnol.* 26, 1003–1010
- 39 Becker, S.A. and Palsson, B.O. (2008) Context-specific metabolic networks are consistent with experiments. *PLoS Comput. Biol.* 4, e1000082
- 40 Jensen, P.A. and Papin, J.A. (2011) Functional integration of a metabolic network model and expression data without arbitrary thresholding. *Bioinformatics* 27, 541–547
- 41 Schwegmann, A. and Brombacher, F. (2008) Host-directed drug targeting of factors hijacked by pathogens. *Sci. Signal.* 1, re8
- 42 Sakamoto, H. *et al.* (2005) Host sphingolipid biosynthesis as a target for hepatitis C virus therapy. *Nat. Chem. Biol.* 1, 333–337
- 43 Handley, S.A. *et al.* (2006) Histamine signaling through the H(2) receptor in the Peyer's patch is important for controlling *Yersinia enterocolitica* infection. *Proc. Natl. Acad. Sci. U.S.A.* 103, 9268–9273
- 44 Hossain, H. *et al.* (2006) Host gene expression profiling in pathogen-host interactions. *Curr. Opin. Immunol.* 18, 422–429
- 45 Diamond, D.L. *et al.* (2010) Temporal proteome and lipidome profiles reveal hepatitis C virus-associated reprogramming of hepatocellular metabolism and bioenergetics. *PLoS Pathog.* 6, e1000719
- 46 Carette, J.E. *et al.* (2009) Haploid genetic screens in human cells identify host factors used by pathogens. *Science* 326, 1231–1235
- 47 Oberhardt, M.A. *et al.* (2008) Genome-scale metabolic network analysis of the opportunistic pathogen *Pseudomonas aeruginosa* PAO1. *J. Bacteriol.* 190, 2790–2803
- 48 Puchalka, J. *et al.* (2008) Genome-scale reconstruction and analysis of the *Pseudomonas putida* KT2440 metabolic network facilitates applications in biotechnology. *PLoS Comput. Biol.* 4, e1000210
- 49 Oberhardt, M.A. *et al.* (2011) Reconciliation of genome-scale metabolic reconstructions for comparative systems analysis. *PLoS Comput. Biol.* 7, e1001116
- 50 Stolyar, S. *et al.* (2007) Metabolic modeling of a mutualistic microbial community. *Mol. Syst. Biol.* 3, 92
- 51 Oberhardt, M.A. *et al.* (2010) Metabolic network analysis of *Pseudomonas aeruginosa* during chronic cystic fibrosis lung infection. *J. Bacteriol.* 192, 5534–5548
- 52 Overbeek, R. *et al.* (2005) The subsystems approach to genome annotation and its use in the project to annotate 1000 genomes. *Nucleic Acids Res.* 33, 5691–5702
- 53 Aziz, R.K. *et al.* (2008) The RAST Server: rapid annotations using subsystems technology. *BMC Genomics* 9, 75
- 54 Langridge, G.C. *et al.* (2009) Simultaneous assay of every *Salmonella* Typhi gene using one million transposon mutants. *Genome Res.* 19, 2308–2316
- 55 Fang, K. *et al.* (2011) Exploring the metabolic network of the epidemic pathogen *Burkholderia cenocepacia* J2315 via genome-scale reconstruction. *BMC Syst. Biol.* 5, 83
- 56 Thiele, I. *et al.* (2005) Expanded metabolic reconstruction of *Helicobacter pylori* (iT341 GSM/GPR): an *in silico* genome-scale characterization of single- and double-deletion mutants. *J. Bacteriol.* 187, 5818–5830
- 57 Schilling, C.H. *et al.* (2002) Genome-scale metabolic model of *Helicobacter pylori* 26695. *J. Bacteriol.* 184, 4582–4593
- 58 Liao, Y.C. *et al.* (2011) An experimentally validated genome-scale metabolic reconstruction of *Klebsiella pneumoniae* MGH 78578, iYL1228. *J. Bacteriol.* 193, 1710–1717
- 59 Baart, G.J. *et al.* (2007) Modeling *Neisseria meningitidis* metabolism: from genome to metabolic fluxes. *Genome Biol.* 8, R136
- 60 AbuOun, M. *et al.* (2009) Genome scale reconstruction of a *Salmonella* metabolic model: comparison of similarity and differences with a commensal *Escherichia coli* strain. *J. Biol. Chem.* 284, 29480–29488
- 61 Thiele, I. *et al.* (2011) A community effort towards a knowledge-base and mathematical model of the human pathogen *Salmonella* Typhimurium LT2. *BMC Syst. Biol.* 5, 8
- 62 Raghunathan, A. *et al.* (2009) Constraint-based analysis of metabolic capacity of *Salmonella* Typhimurium during host-pathogen interaction. *BMC Syst. Biol.* 3, 38
- 63 Lee, D.S. *et al.* (2009) Comparative genome-scale metabolic reconstruction and flux balance analysis of multiple *Staphylococcus aureus* genomes identify novel antimicrobial drug targets. *J. Bacteriol.* 191, 4015–4024
- 64 Heinemann, M. *et al.* (2005) *In silico* genome-scale reconstruction and validation of the *Staphylococcus aureus* metabolic network. *Biotechnol. Bioeng.* 92, 850–864
- 65 Becker, S.A. and Palsson, B.O. (2005) Genome-scale reconstruction of the metabolic network in *Staphylococcus aureus* N315: an initial draft to the two-dimensional annotation. *BMC Microbiol.* 5, 8
- 66 Navid, A. and Almaas, E. (2009) Genome-scale reconstruction of the metabolic network in *Yersinia pestis*, strain 91001. *Mol. Biosyst.* 5, 368–375
- 67 Vanee, N. *et al.* (2010) A genome-scale metabolic model of *Cryptosporidium hominis*. *Chem. Biodivers.* 7, 1026–1039
- 68 Plata, G. *et al.* (2010) Reconstruction and flux-balance analysis of the *Plasmodium falciparum* metabolic network. *Mol. Syst. Biol.* 6, 408
- 69 Roberts, S.B. *et al.* (2009) Proteomic and network analysis characterize stage-specific metabolism in *Trypanosoma cruzi*. *BMC Syst. Biol.* 3, 52
- 70 Duarte, N.C. *et al.* (2007) Global reconstruction of the human metabolic network based on genomic and bibliomic data. *Proc. Natl. Acad. Sci. U.S.A.* 104, 1777–1782
- 71 Segre, D. *et al.* (2002) Analysis of optimality in natural and perturbed metabolic networks. *Proc. Natl. Acad. Sci. U.S.A.* 99, 15112–15117
- 72 Feist, A.M. *et al.* (2007) A genome-scale metabolic reconstruction for *Escherichia coli* K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. *Mol. Syst. Biol.* 3, 121
- 73 Hsiao, T.L. *et al.* (2010) Automatic policing of biochemical annotations using genomic correlations. *Nat. Chem. Biol.* 6, 34–40
- 74 Szappanos, B. *et al.* (2011) An integrated approach to characterize genetic interaction networks in yeast metabolism. *Nat. Genet.* 43, 656–662
- 75 Boyle, N.R. and Morgan, J.A. (2009) Flux balance analysis of primary metabolism in *Chlamydomonas reinhardtii*. *BMC Syst. Biol.* 3, 4
- 76 Chen, X. *et al.* (2011) Synergy between ¹³C-metabolic flux analysis and flux balance analysis for understanding metabolic adaptation to anaerobiosis in *E. coli*. *Metab. Eng.* 13, 38–48
- 77 Blank, L.M. *et al.* (2005) Large-scale ¹³C-flux analysis reveals mechanistic principles of metabolic network robustness to null mutations in yeast. *Genome Biol.* 6, R49
- 78 Varma, A. and Palsson, B.O. (1994) Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type *Escherichia coli* W3110. *Appl. Environ. Microbiol.* 60, 3724–3731
- 79 Forster, J. *et al.* (2003) Genome-scale reconstruction of the *Saccharomyces cerevisiae* metabolic network. *Genome Res.* 13, 244–253