Constraining the metabolic genotype–phenotype relationship using a phylogeny of *in silico* methods

Nathan E. Lewis[1], Harish Nagarajan[2] and Bernhard O. Palsson[1]

Abstract | Reconstructed microbial metabolic networks facilitate a mechanistic description of the genotype–phenotype relationship through the deployment of constraint-based reconstruction and analysis (COBRA) methods. As reconstructed networks leverage genomic data for insight and phenotype prediction, the development of COBRA methods has accelerated following the advent of whole-genome sequencing. Here, we describe a phylogeny of COBRA methods that has rapidly evolved from the few early methods, such as flux balance analysis and elementary flux mode analysis, into a repertoire of more than 100 methods. These methods have enabled genome-scale analysis of microbial metabolism for numerous basic and applied uses, including antibiotic discovery, metabolic engineering and modelling of microbial community behaviour.

The genotype–phenotype relationship is fundamental to biology. For decades this relationship has been subjected to mostly argument, speculation and qualitative analysis. However, our ability to fundamentally understand the genotype–phenotype relationship began to change in the mid 1990s, on completion of the first bacterial genome-sequencing projects. Full genome sequences provide comprehensive, albeit not yet complete, information about the genetic elements that create an organism. A detailed understanding of some cellular processes, such as metabolism, has resulted in structured knowledge bases that can be mathematically represented[1–3]. This mathematical representation enables the computation of phenotypic states[4–7] based on genetic and environmental parameters. Remarkably, this provides a mechanistic representation of the microbial metabolic genotype–phenotype relationship.

Constraint-based models of genome-scale metabolic networks capture the genotype–phenotype relationship by simultaneously accounting for constraints that are imposed on phenotype by physicochemical laws and genetics. The realization that these quantitative genotype–phenotype relationships could be constructed from a genome has driven the emergence of this area of research, and the flood of increasingly rich high-throughput data has accelerated the evolution of constraint-based reconstruction and analysis (COBRA) methods from a set of basic tools for metabolic network analysis into a powerful analytical framework that is increasingly used. Here, we describe basic features of the COBRA framework, the 'phylogeny' of evolving COBRA methods and the COBRA 'ecology' (that is, how COBRA methods complement each other in answering larger questions in biology).

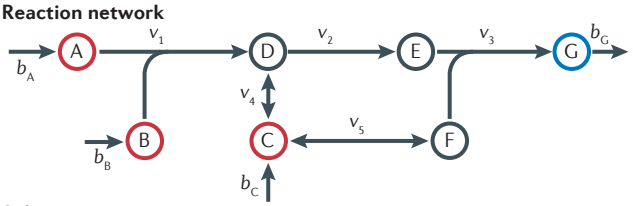
Constraint-based modelling defined

The COBRA approach is based on a few fundamental concepts. These concepts include the imposition of physicochemical constraints that limit computable phenotypes (FIG. 1a–d), the identification and mathematical description of evolutionary selective pressures (FIG. 1e), and a genome-scale perspective of cell metabolism that accounts for all metabolic gene products in a cell (FIG. 1d,f). These fundamental concepts are briefly described below.

Constraints on reaction networks. Metabolism is a complex network of biochemical reactions. The reaction occurrence is limited by three primary constraints: substrate and enzyme availability, mass and charge conservation, and thermodynamics. For metabolic reactions, substrates must be present in the microenvironment of the cells or produced from other reactions, and enzymes must be available. Mass conservation further limits the possible reaction products and their stoichiometry, and thermodynamics constrains reaction directionality. For a given organism, this information can be obtained

[1]*Department of Bioengineering, University of California.*
[2]*Bioinformatics and Systems Biology Graduate Program, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093–0412, USA.*
Correspondence to B.O.P.
e-mail: palsson@ucsd.edu
doi:10.1038/nrmicro2737
Published online
27 February 2012

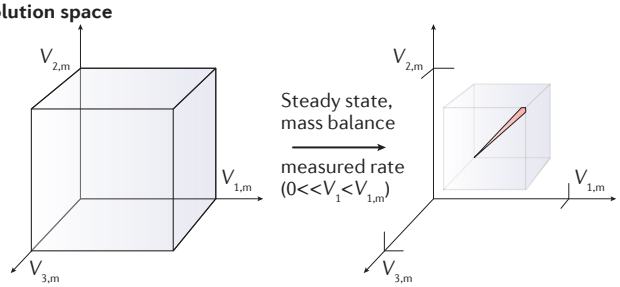
a Reaction network



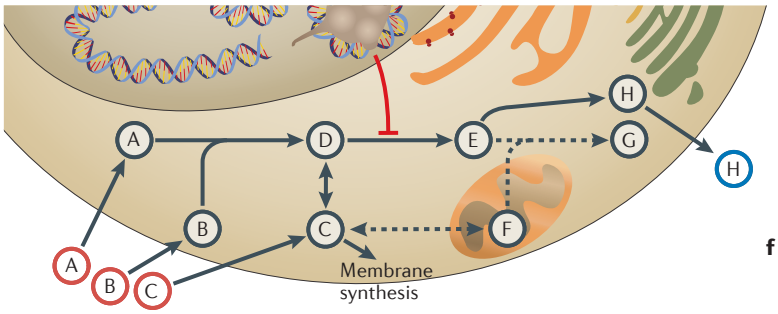
b Stoichiometric matrix

	v_1	v_2	v_3	v_4	v_5	b_A	b_B	b_C	b_G
A	-1	0	0	0	0	1	0	0	0
B	-1	0	0	0	0	0	1	0	0
C	0	0	0	-1	-1	0	0	1	0
D	1	-1	0	1	0	0	0	0	0
E	0	1	-1	0	0	0	0	0	0
F	0	0	-1	0	1	0	0	0	0
G	0	0	1	0	0	0	0	0	-1
Lower	0	0	0	$V_{m,r}$	$V_{m,r}$	0	0	0	0
Upper	V_m	V_m	V_m	V_m	V_m	∞	∞	∞	∞

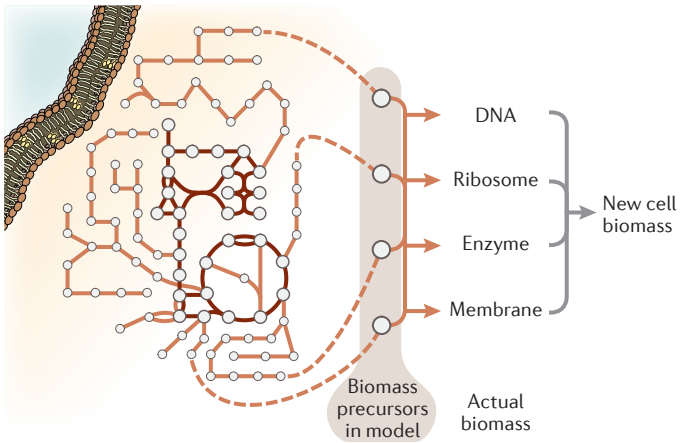
c Solution space



d Additional complexity



e Biomass objective function



f Gene–protein–reaction associations for data integration

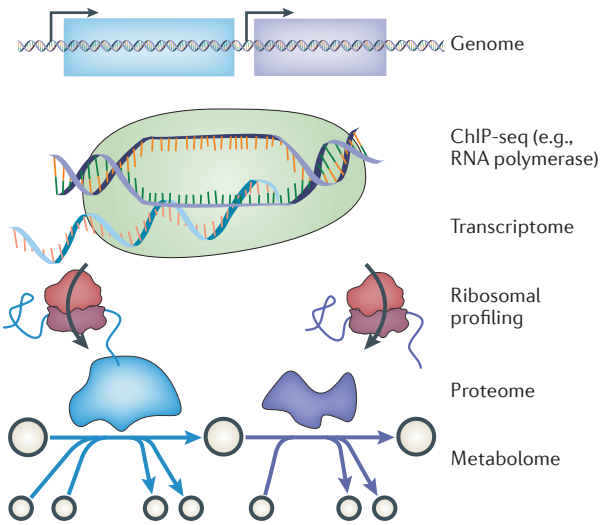


Figure 1 | **Fundamentals of the genome-scale metabolic genotype–phenotype relationship.** The constraint-based reconstruction and analysis (COBRA) approach is based on three primary fundamental concepts: network constraints (parts **a**–**d**), objective functions (part **e**) and the association of reactions with the genome (part **f**). **a** | A complex mixture of molecules (red) can react to yield end products (blue). In the model, metabolites enter the system through boundary pseudoreactions (*b*) and are metabolized in internal reactions (*v*) with a flux of *V*. **b** | The stoichiometry of this reaction network is described mathematically in a stoichiometric matrix, with each column representing the stoichiometry of a reaction. Negative and positive values represent reactants and products, respectively. Reaction flux is limited by thermodynamics and catalytic capacities (V_m and $V_{m,r}$ represent V_{max}, the velocity of the forward and reverse enzyme-catalysed reactions, respectively, at an infinite concentration of substrate), described by upper and lower bounds on flux for each reaction. **c** | Reaction constraints result in a 'solution space' that contains all feasible flux distributions. Additional constraints (for example, mass balance, the steady-state assumption and measured metabolite consumption rates) reduce the space of feasible flux distributions, as shown by the pink line. **d** | *In vivo* biochemical networks involve additional complexity. Gene regulation can change the abundance of catalysts (for example, the transformation of D to E). Often, components are also localized in different organelles (for example, E and F), thereby blocking reactions. **e** | The biomass objective function describes an evolutionary pressure for microbial growth, and describes the metabolic demands to make the basic metabolite building blocks for all cellular components (for example, membranes, macromolecules, ATP, and so on). **f** | The association of metabolism with the genome is carried out by mathematically linking the genome to transcripts, proteins and chemical reactions. The gene–protein–reaction schema is used to describe gene association in the models and provide an interface for the integration of high-throughput data. ChIP–seq, chromatin immunoprecpitation followed by sequencing.

from careful biochemical and genetic studies or inferred from related organisms, and then catalogued in metabolic reconstruction knowledge bases[1,2].

In the COBRA framework, a metabolic reconstruction is converted into an *in silico* model by mathematically describing the reactions and adding network inputs and outputs (for example, uptake and secretion of products). Similarly to a cell having one genome and many transcriptional states, an organism has one metabolic reconstruction from which context-specific models can be derived, each representing cellular functions under different conditions.

Physicochemical constraints on the metabolic network are mathematically described by a matrix representing the stoichiometric coefficients of each reaction[8] (FIG. 1a,b). Known upper and lower bounds on each reaction flux are imposed as additional constraints. Mathematically, these constraints define a multidimensional 'solution space' of allowable reaction flux distributions, and the actual expressed flux state resides in this solution space. Additional constraints can further shrink the solution space to focus in on the actual flux state of the network (FIG. 1c). These additional constraints may include enzyme capacity, spatial localization, metabolite sequestration, and multiple levels of gene, transcript and protein regulation (FIG. 1d).

A mathematical statement of cell objectives: a reflection of evolution. In non-biological chemical networks, the material flow through pathways can be predicted in a cause and effect manner, using mathematical models that describe the associated physical laws. This description is time invariant, as reproducing the physical conditions will always drive flux through the same pathways. By contrast, causation in biology is time variant. A plethora of chemical reactions may occur inside a cell, and many pathways can link a starting molecule to a given product. However, regulatory mechanisms have evolved to select when and where pathways will be used in an organism under a given set of conditions. Thus, if the cellular objectives that drive evolution are understood or can be inferred, optimal flux states of biochemical reaction networks can be predicted. In the COBRA framework these cellular objectives are described mathematically and used to compute phenotypic states.

Many cellular objectives can be defined in the context of metabolism. For example, as a proxy for growth, a biomass function[9] can be defined, containing all the necessary precursors for synthesizing the cell components for growth (for example, amino acids for proteins and nucleic acids for RNA) (FIG. 1e). The biomass function and other objective functions can be used with optimization algorithms such as linear programing[10,11] to predict metabolic pathway usage and cellular phenotypes[11]. As these objective functions mathematically state cellular aims and can predict phenotypes, they capture the pressures guiding evolution and therefore represent agents of causation in biology. The objective function is thus an important part of the COBRA framework. It is not directly based on fundamental physical principles, but is based on biological functions that are selected for over many generations.

A genome-wide basis for modelling metabolism. Constraint-based modelling has rapidly developed since the advent of whole-genome sequencing[12,13]. A genome provides the genetic basis for the metabolic network of an organism, and genome annotation defines the relationships between genes, enzymes and the reactions that they catalyse[14] (FIG. 1f). Annotated genomes and their associated biochemical and genetic data have facilitated the development of carefully curated metabolic network reconstructions containing thousands of reactions. When a reconstruction knowledge base for an organism is converted into a genome-scale model (GEM), the mathematical representation provides constraints, and the objective function can be used to represent the optimal biological functions that the organism strives to achieve. Thus, simulation of an organism's phenotypes can be performed using its GEM.

The genome-scale view of metabolism in these models has two primary implications. First, in principle, the models account for all known metabolic genes in a cell and their functions. Thus, when used in the analysis of genome-scale data sets (for example, proteomics, metabolomics, and so on)[15], they provide novel insight because they account for real chemical connections between components (FIG. 1f). Second, as metabolic genes are associated with the biochemical functions of their gene products, simulations of metabolite flow through the network can provide mechanistic predictions of how each gene product affects the metabolic network function. Thus, cell phenotypes can be computed and data can be interpreted with GEMs, providing mechanistic insight into how the cell genotype may contribute to the cell phenotype.

A phylogeny of constraint-based methods

COBRA methods have 'evolved' and 'diversified' over the past decade, leading to more than 100 different methods (see Supplementary information S1 (table) and http://sourceforge.net/apps/mediawiki/opencobra/), many of which have been implemented in available software packages (see Supplementary information S2 (table)). These developments may be likened to an evolutionary process in which specific scientific questions have selected for algorithmic innovations, yielding a phylogenetic tree of COBRA methods (FIG. 2). Below, we classify these methods into major groups and describe examples that address the broader scientific questions.

Global characterization of solution spaces. Metabolic pathways are conceptual abstractions that group reactions. However, sometimes these pathways fail to reflect actual metabolic network usage[16], as textbook pathways often reflect the order of enzyme discovery or pathway usage in one model organism. Fortunately, through computational analysis of metabolic networks, the required pathways for specific metabolic functions can be identified without biases from traditional pathway concepts. In constraint-based modelling, this is approached through unbiased and biased methods, represented by the two primary branches of the phylogenetic tree (FIG. 2).

Metabolic reconstruction
A carefully curated and biochemically validated knowledge base in which all known chemical reactions for an organism are detailed and catalogued.

Solution space
The feasible region satisfying a set of constraints. In constraint-based reconstruction and analysis (COBRA) models, this represents the feasible flux values for all of the reactions in the model.

Flux distributions
A set of steady-state fluxes for all of the reactions in a metabolic network.

Biomass function
A pseudo-reaction that is formed to aid in predicting the growth of a cell in constraint-based reconstruction and analysis (COBRA) models. It describes the rate at which and accurate proportions in which all of the biomass precursors are made.

Linear programing
A mathematical optimization technique that determines a way to maximize a particular linear objective under a given set of conditions (that is, when subjected to linear equalities and inequalities as constraints). Typically used in flux balance analysis, in which the objective is often the biomass function (growth) and the constraints represent the growth conditions.

Genome-scale model
(GEM). A condition-specific, mathematically described, computable derivative of a metabolic reconstruction, containing comprehensive knowledge of metabolism.

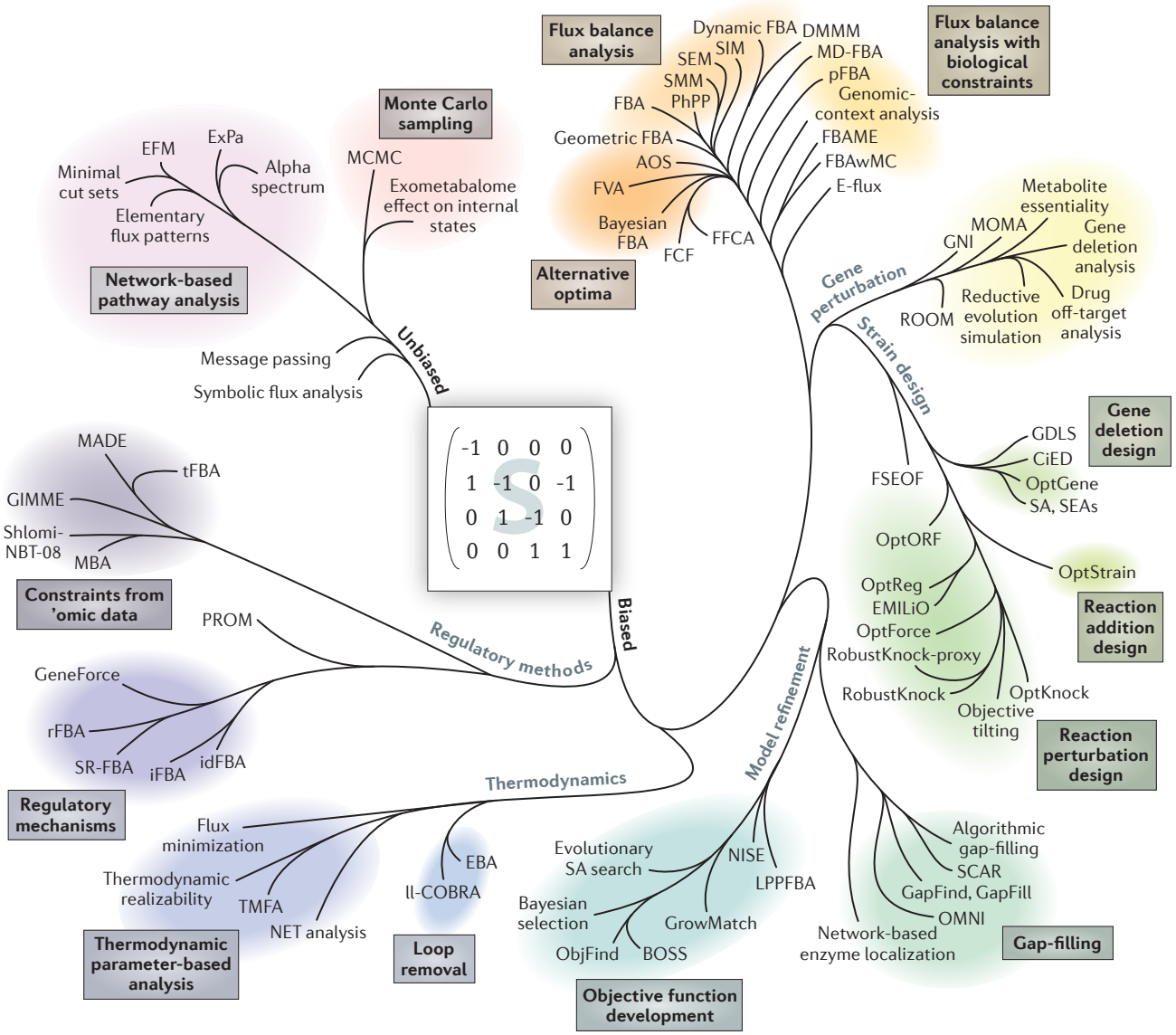


Figure 2 | **The 'phylogeny' of constraint-based modelling methods.** Over the past few years, the repertoire of tools for constraint-based reconstruction and analysis (COBRA) modelling has rapidly expanded. Because of the versatility and scalability of COBRA models, more than 100 methods have been developed for their prediction and analysis, all based on analysis of the underlying metabolic network structure (that is, the stoichiometric matrix). A phylogenetic tree is used to depict the similarities between the applications and uses of these methods, and the underlying algorithms for many of the methods. See Supplementary information S1 (table) for a more complete list of methods and descriptions, and for definitions of abbreviations.

Unbiased methods describe all steady-state flux distributions, including reaction sets that function together without belonging to the same traditional pathway concepts.

Two such unbiased approaches, elementary flux mode (EFM) analysis and extreme pathway (ExPa) analysis, globally characterize allowable phenotypes and have been reviewed and compared previously[17–19]. These methods identify reaction sets (that is, pathways) that achieve specific metabolic functions, and combinations of these reaction sets describe the entire solution space (that is, all steady-state phenotypes). These two unbiased methods have been used for many different ends. For example, in studies of *Escherichia coli* metabolism they have helped assess global pathway regulation[20], facilitated the design of an ethanol-secreting strain[21], identified synthetic-lethal gene interactions[22] and demonstrated the trade-off between reducing translation costs and rapidly responding to environmental changes[23]. These methods are usually applied to small models or portions of GEMs[24], because their computational complexity scales exponentially[25,26]. However, their use on larger models is becoming possible through simplifications that, for example, calculate a subset of potential pathways or find minimal pathways that accomplish a biological function[27–30].

Alternative approaches can also describe the entire solution space in an unbiased manner[31,32]. For example, Markov chain Monte Carlo (MCMC) sampling methods[32] characterize all feasible steady-state reaction fluxes. This provides a probability distribution of feasible fluxes for each reaction under the user-provided growth conditions. These methods have provided insight into several biological properties, such as the high-flux backbone of central metabolism in *E. coli*[33], condition-specific regulation of metabolism in yeast[34,35] and *E. coli*[36], and disease states in cardiac myocytes[37], erythrocytes[38] and the human brain[39].

Finding the 'optimal' metabolic state with flux balance analysis methods. EFM, ExPa and MCMC methods characterize all the flux states that a metabolic network can deploy. However, a cell does not use most of the possible flux states. Thus, biased COBRA methods include the optimization of an objective function to identify physiologically relevant flux distributions. Flux balance analysis (FBA) is the most basic and commonly used biased method for simulating genome-scale metabolism. In FBA, the cellular objective is defined, and metabolites in the media are supplied to the metabolic network. Using linear programing, an objective function (for example, the biomass objective function) is optimized subject to the constraints imposed by the metabolic network and metabolite uptake rates[10,11,40]. This calculation finds one solution in the solution space that is believed to best represent the true cellular phenotype. The solution includes a prediction of the optimal objective magnitude (for example, biomass yield or growth rate) and potential flux values for each reaction (FIG. 3a).

FBA successfully makes quantitative predictions using a few governing constraints on the model. For example, a pre-genome era application of FBA recapitulated the acetate overflow phenotype of *E. coli*[41], in which acetate is excreted at high growth rates. Using GEMs, FBA has since predicted growth rates[42], pathway usage[43,44] and the effect of gene expression noise on fitness[45]. It has allowed the analysis of complex phenotypes, such as metabolism in non-growing cells[46], and numerous variations on FBA have been developed to assess alternative optimal solutions or to account for additional constraints on metabolic flux in cells (FIG. 2).

Predicted flux values from FBA can vary owing to there being alternative optimal solutions (that is, the same objective value using different reactions) (FIG. 3b). Alternative optimal solutions are enumerated using mixed-integer linear programing (MILP)[47], and the ranges spanned by alternative optima are found for each reaction using flux variability analysis[48,49] (FIG. 3b). The consideration of all alternative optima is crucial when interpreting an FBA solution, as the flux through a single reaction can vary considerably depending on which solution is found. For example, the COBRA method minimization of metabolic adjustment (MOMA)[50] predicts a new flux vector and objective value after a perturbation (for example, a gene deletion). To do this, MOMA computes one 'wild-type' FBA solution and finds the nearest solution after perturbing the network (that is, the minimum change to the reaction fluxes from the FBA solution). As the new predicted flux vector and growth rate can differ considerably depending on which alternative optimal solution is used (FIG. 3b), all possible results from alternative optima must be assessed.

To identify realistic microbial phenotypes in FBA predictions, additional biologically relevant constraints have been proposed. These include constraints imposed by metabolite dilution[51], changes in transcript level[52,53] and economy in enzyme use[43,54–56]. These FBA refinements further decrease the range of feasible reaction fluxes to obtain solutions closely resembling cellular physiology under certain growth conditions. For example, constraints from enzyme crowding have been applied to FBA solutions (FBAwMC)[57,58]. In FBAwMC, reaction flux is constrained to reflect internal spatial limitations on enzyme abundance in the crowded cytoplasm. This method predicted that molecular crowding contributes to substrate preferences in *E. coli*[57]. In a medium with multiple carbon substrates, FBAwMC accurately predicted that glucose would be preferentially consumed, followed by mixed-substrate consumption and a late utilization of glycerol and the excreted acetate (FIG. 3c), suggesting that molecular crowding contributes to substrate preference. A similar variation on FBA, FBAME, accounts for cytoplasmic membrane crowding by limiting the flux through the glucose transporter and the three cytochromes in *E. coli*[59]. This constraint recapitulated the simultaneous use of respiratory and fermentative pathways and predicted the effect of glucose and oxygen availability on cytochrome oxidase expression. Thus, the imposition of crowding constraints on metabolic flux has provided additional insights into cell physiology[57–59].

Modelling genetic perturbations. As genome-scale metabolic networks capture the activities of hundreds of enzymes, mutant phenotypes can be assayed through *in silico* gene perturbation and simulation. For the first GEMs[12,13], such approaches demonstrated the predictive power of COBRA methods when metabolic genes were 'knocked out' in the model by restricting flux through their associated reactions. When the growth of *E. coli* mutants was simulated with FBA, 86% of the mutant phenotypes (that is, growth or no growth) were accurately predicted[13]. This success rate exceeded any other phenotype-predicting algorithm available at the time. Subsequent studies have identified growth conditions[60] and genetic backgrounds[61] for which genes in *Saccharomyces cerevisiae* are conditionally essential. For example, combinations of gene knockouts were simulated and tested for essentiality. This demonstrated that 74% of yeast metabolic genes contribute to essential metabolic processes, and most of these are masked by isozymes and alternative pathways[61]. To address additional questions concerning gene deletion, new methods have been introduced such as MOMA[50], regulatory on–off minimization (ROOM)[62] and metabolite essentiality analysis (MEA)[63] (FIG. 2).

Gene and reaction perturbation studies have aided health-related applications such as prediction of the metabolic side effects of off-target protein–drug

Mixed-integer linear programing
(MILP). Similar to linear programing, but some of the constraints are integer values. Used for applications such as enumerating alternative optimal solutions, strain design, eliminating loops, and so on.

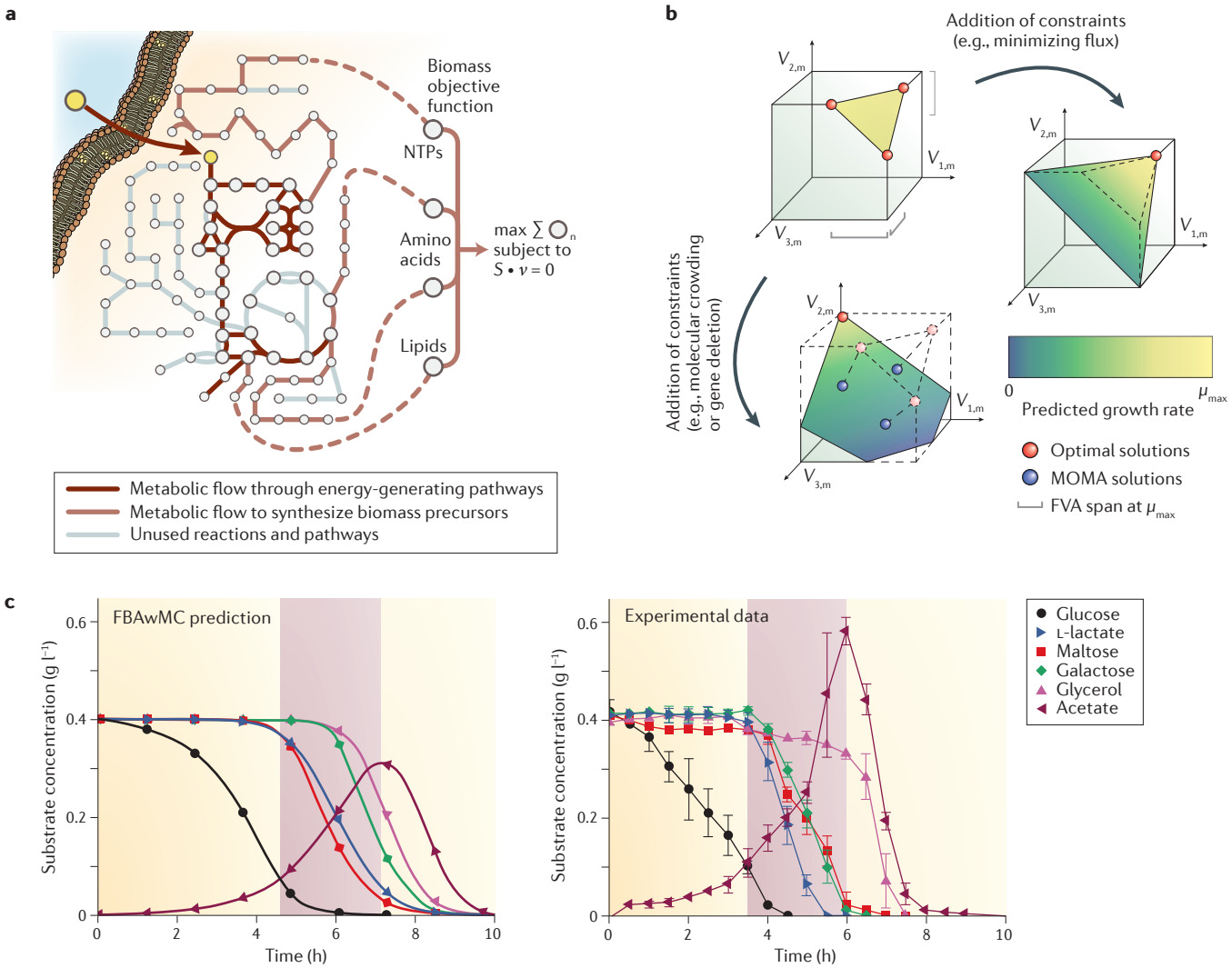


Figure 3 | **Flux balance analysis.** **a** | In flux balance analysis (FBA), a cellular objective (for example, growth, as represented by the sum of all biomass precursors in the biomass function) is optimized. This provides the predicted flux for each reaction in the network. *S* is the stoichiometric matrix, *v* is a vector of all reaction fluxes, and $S \cdot v = 0$ is the steady state assumption showing that all reaction fluxes are balanced. **b** | FBA solutions are typically not unique — that is, there are alternative optimal solutions that use different pathways to achieve the same objective value (for example, growth rate). Additional constraints can be applied to reduce the solution space size, and these may remove competing optimal solutions (right) or change the optimal solution (bottom). If the optimal solution is moved, then the choice of the new optimal solution may depend on the solver and/or algorithm. For example, FBS will find the new optimal objective value (red) following a change in constraints. However, the minimization of metabolic adjustment (MOMA)[50] method may predict various objective values (blue) depending on which alternative optimal value (pink) is initially used. **c** | The addition of constraints can enhance predictions. For example, when constraints on molecular crowding are added to FBA (FBAwMC), the model-predicted order of substrate metabolism is consistent with experimental observation. NTPs, nucleotide triphosphates; FVA, flux variability analysis; V_m, the velocity of the enzyme-catalysed reaction at an infinite concentration of substrate; μ_{max}, predicted maximum growth rate. Panel **c** image is reproduced, with permission, from REF. 57 © (2007) US National Academy of Sciences.

interactions[64] and prediction of novel antimicrobial targets[65]. For example, MEA[63] was applied to the *Vibrio vulnificus* GEM[66] to identify potential antibiotic targets for this pathogenic relative of *Vibrio cholerae*. MEA was used because it identifies metabolites that, if removed, inhibit biomass production. These metabolites could possibly be blocked *in vivo* with analogues that bind or modify active sites on their associated enzymes. This analysis identified five metabolites as potential antibiotic targets, and only 352 analogues had to be tested for antimicrobial properties, allowing for a smaller screen than would commonly be required for drug discovery. One of the screened molecules with antimicrobial properties was subjected to additional analysis, and this candidate molecule considerably out-performed sulphamethoxazole, an existing therapeutic for *V. vulnificus* infection. Although additional drug safety assessment and optimization is required for this candidate drug, this study

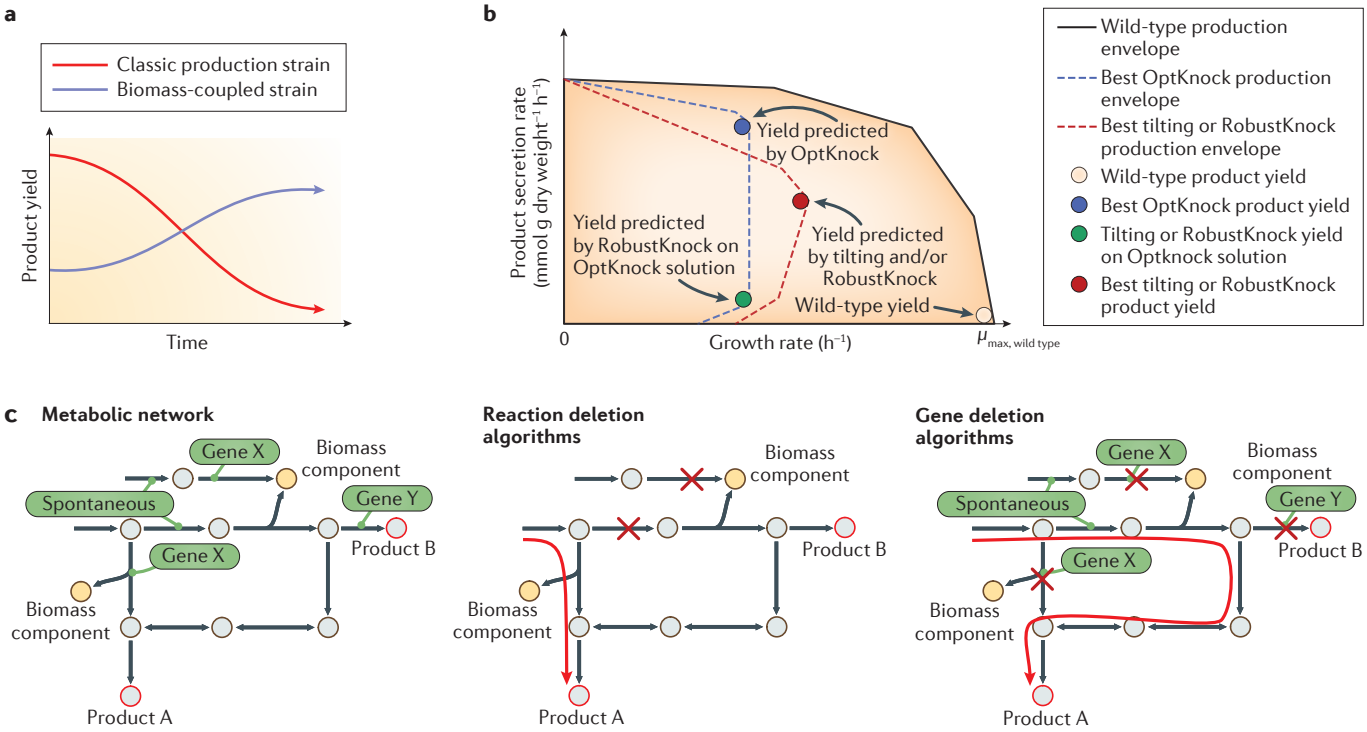


Figure 4 | **Principles of model-guided strain design.** **a** | Non-growth-coupled (classic) production strains exhibit a decrease in product yield over time, whereas growth-coupled (biomass-coupled) strains can enhance product yield. **b** | Designs for growth-coupled strains are predicted to force product secretion while allowing optimal growth of the organism. Several methods have been developed to predict strains that undergo growth-coupled production, and these methods model reaction deletion, gene deletion or reaction addition. Different reaction deletion algorithms, such as OptKnock[68], objective tilting[69] and RobustKnock[70], can provide different optimal growth-coupled-strain designs owing to algorithmic differences. **c** | Many algorithms predict the set of reactions that must be blocked (or deleted) to obtain a desired product. However, methods such as OptGene[71] and genetic design through local search (GDLS)[72] provide a more realistic view by modelling genetic modifications, as some genes catalyse multiple reactions and some reactions are spontaneous. μ_{max}, predicted maximum growth rate.

demonstrates how COBRA methods can guide antibiotic screens and provide immediate insight into the mode of action of a drug.

In silico *design of production strains.* Metabolic engineering approaches often perturb and screen cells for desired phenotypes. However, engineered strains can decrease product yield over time, as products drain cellular resources. Thus, several COBRA methods aim to predict the perturbations (for example, gene deletions or additions) that would force the strain to couple product yield to a cellular objective. For example, the secretion of a product can be coupled to growth if the product precursor provides an essential biomass component and if pathways that would metabolize the desired product are removed. Thus, as cells grow exponentially, they can actually increase productivity[67] (FIG. 4a).

Most COBRA methods of strain design systematically identify reactions that, when perturbed, may couple a product to a selective pressure (FIG. 2). The production capabilities of a strain can be shown with a production envelope, which is a plot of the bounds of all possible product secretion rates and their associated feasible growth rates (FIG. 4b). As there is selection for fast growing cells that minimize product secretion (FIG. 4b, white circle), OptKnock[68] employs MILP on a wild-type model to find reaction deletions that force product secretion under the maximum growth rate for the mutant, indicated by a non-zero production rate in the production envelope when growth rate is maximized (FIG. 4b, blue circle). However, as OptKnock optimizes both the biomass objective function and product yield, strain designs occasionally have alternative optima with other secretion products, leading to a potential low product secretion rate (FIG.4b, green circle). To avoid this, the product can be added to the biomass function (called objective tilting[69]) or MILP can be applied (using RobustKnock[70]) to find designs that provide the maximum lower bound on product yield while maximizing growth (FIG.4b, red circle).

For algorithmic simplicity, most strain design methods perturb reactions. However, strain designs that are based on reactions can require additional gene deletions to remove relevant isozymes. Moreover, predictions are occasionally not feasible when they require the removal of one reaction that is catalysed by a multispecific enzyme (FIG. 4c). To avoid such predictions, heuristic approaches such as OptGene[71] and genetic design through local

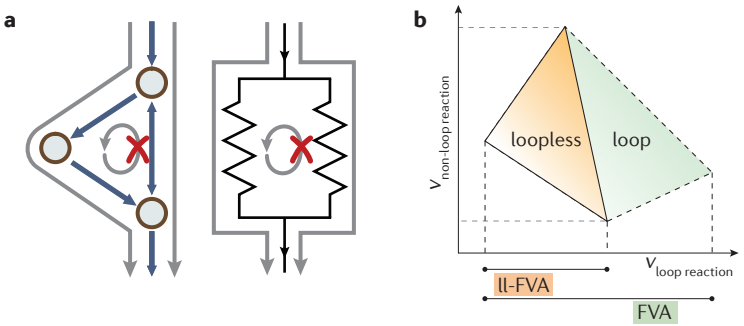


Figure 5 | **Refining thermodynamic constraints.** Thermodynamic constraints in constraint-based reconstruction and analysis (COBRA) models can be refined. **a** | For example, when a metabolic network is not adequately constrained, metabolites can cycle infinitely in loops. Akin to Kirchhoff's loop law for electrical circuits, this property is thermodynamically infeasible (crosses). **b** | Thus, methods such as loopless flux variability analysis (ll-FVA), which uses the loopless COBRA[96] constraints on FVA, are able to systematically remove these loops by adding a constraint that limits flux to the solution space regions that are not involved in these loops. *V*, reaction flux.

search (GDLS)[72] identify strain designs that directly involve gene deletions to give growth-coupled production of a desired product. Thus, these strain designs are more realistic and easier to test *in vivo*.

Strain design predictions are not limited to manipulations of the metabolic pathways encoded by the host cell. The repertoire of products may be expanded *in silico* by adding genes from other organisms to confer novel metabolic functions. *In silico* methods have used graph theoretical approaches[73,74] or kinetic parameters[75] to build novel biosynthetic pathways, which were subsequently tested or ranked using COBRA methods. Unfortunately, without accounting for the metabolic network of the host, these approaches cannot guarantee growth-coupled-strain designs. Thus, without further engineering (for example, using scaffolds that physically couple enzymes[76]), predicted biosynthetic pathways may not yield a product *in vivo*. However, this concern has been addressed through a few approaches, such as by manually removing genes to couple the new pathways to growth[75], or systematically following pathway prediction with OptKnock[77]. OptStrain goes further by conducting a novel pathway search within the metabolic network of the host to optimize the balance between reaction addition and deletion[78]. Thus, COBRA approaches allow the coupling of non-native product synthesis to a cellular objective.

The concept of designing strains that couple a product to a defined selective pressure is intriguing, and a few COBRA-based *in silico* predictions have in fact been implemented *in vivo*[67,77]. It is anticipated that these tools will continue to aid metabolic engineering projects.

Refining representations of biological causation. Simulating cell phenotypes requires accurate representations of metabolic network stoichiometry and objective functions. Although metabolic reconstructions are usually carefully built and rigorously tested, they are often incomplete and may contain a few errors in stoichiometry, thermodynamics, gene associations or biomass composition as a result of ambiguities in the associated biochemical studies[79] or genome annotation[80]. Moreover, biomass composition and cellular objectives can vary between environments[81,82], especially under nutrient limitation or stress and during stationary phase[46,83]. To address these concerns, phenotypic screens have been analysed with gap-filling COBRA methods (FIG. 2) to predict missing pathways[84,85], to identify incorrect reaction directionality or inclusion[79,86–88] and to suggest subcellular reaction localizations in microorganisms with multiple organelles[89]. Complementary COBRA methods also improve the definition of cellular objectives by integrating data to systematically assess[90–92], predict[93] or modify objective functions[79,81,87].

Recently, high-throughput genetic-interaction screens have helped refine metabolic networks and the biomass objective function of yeast[79,94]. For example, model-predicted epistasis in *S. cerevisiae* was compared with 176,821 experimentally measured genetic interaction pairs. Although the COBRA model predictions were enriched for high-confidence measured genetic interactions, this method did not predict many epistatic interactions. An algorithm was developed to reconcile discrepancies between model-predicted and experimentally measured interactions, and several of the resulting predicted model improvements were experimentally validated. For example, it was found that the formation of quinolinate from aspartate had been mistakenly included in the yeast reconstruction. In addition, the algorithm predicted that glycogen should be removed as an essential component in the biomass objective function, as it is not essential for growth. Thus, this study demonstrates that COBRA methods could be deployed to improve the model of the yeast metabolic network so that it better predicts the *in vivo* phenotype and contains condition-specific updates to the biomass objective function.

Thermodynamics. COBRA methods provide quantitative predictions without detailed parameterization of each reaction, beyond declaring the direction of a reaction to reflect reaction thermodynamics. Reaction directionality is often determined from biochemical assays, but such assays may not recapitulate the conditions and metabolite concentrations inside the cell. Therefore, the reaction directionality *in vitro* may be inconsistent with *in vivo* flux. In addition, unrealistic fluxes can be predicted *in silico* if a reaction is reversible in a model but irreversible *in vivo*. Thus, methods are now applying more rigorous thermodynamic constraints (FIG. 2) by removing thermodynamically infeasible pathway uses[95–97] or constraining flux according to Gibbs free energy calculations[54,98,99]. Methods are also being used to infer thermodynamic parameters[100].

Most COBRA models contain sets of reactions that can cycle metabolites among themselves (FIG. 5a). In these cases, FBA cannot predict flux values for the reactions, because their metabolites are cycled infinitely. Such 'loops' are biologically unrealistic, as no net thermodynamic driving force exists, akin to Kirchhoff's second law for electrical circuits. Thus, the net flux around

Epistasis
The interaction between two genes such that the phenotypic effect of one gene is masked by that of the other. Usually identified by the phenotype of the double mutant relative to the phenotypes of the two single mutants.

these loops should be zero[95]. Although these loops do not often affect the non-loop reaction flux, their existence can upset some model predictions. Approaches to systematically remove loops have been proposed[95–97]. For example, loopless COBRA[96] improves FBA solutions by employing MILP to cancel out loop flux (FIG. 5b).

Although loop removal methods can be easily deployed without extra parameterization, detailed thermodynamic approaches may provide more biologically meaningful reaction flux predictions. Thermodynamic parameters for many metabolites are not known. Fortunately, recent advances in group contribution theory provide estimates of the Gibbs free energy of formation for metabolites in COBRA models[101]. With these predicted values, the standard Gibbs free energy change can be predicted for each reaction. These values can help determine the direction of reactions[54,102], predict reasonable concentration levels[98] and allow the use of metabolite concentrations[103] and of ranges for kinetic parameters[99] as constraints. A recent study[104] used estimated metabolite free energy values with experimentally measured equilibrium constants to quantitatively assign directions to reactions. This approach also incorporated *in vivo* pH, temperature and ionic strength to quantitatively assign reaction directionality to the *E. coli* metabolic network. When the model-predicted and experimentally measured growth rates were compared, the quantitative assignment of directionality was found to result in model predictions that matched experimental data, and qualitative directionality assignment was required only for certain reactions (for example, ABC and proton-coupled transporters). As thermodynamics represents one primary model constraint that is necessary for accurate COBRA predictions, it is expected that further developments in this area will be of great importance to the field.

Incorporating regulatory constraints and signalling. Transcriptional regulation and signalling networks interface extensively with metabolism to produce cellular phenotypes. By incorporating regulatory and signalling constraints into metabolic network models, interactions between the systems can be captured to enhance COBRA predictions. There are two primary paradigms for how regulatory constraints are implemented in constraint-based models (FIG. 2). Either experimental data are used to constrain flux through specific reactions[52,53,105–108] (FIG. 6a), or a mathematical representation of transcription regulation[109,110] or signalling[111,112] is interfaced directly with the metabolic network to aid in modelling (FIG. 6b).

Not all pathways are active under all growth conditions, and 'omic data can be used to constrain models accordingly[52,53,105–108] (FIG. 6a). Methods such as GIMME[105], Shlomi-NBT-08 (REF. 106) and model-building algorithm (MBA)[107] each remove pathways that lack expression in 'omic data to obtain functional models that are consistent with cellular gene or protein expression. These approaches have provided novel insights into and discoveries about tissue-specific human metabolism[39,64,113,114]. However, they were also recently used to model metabolic interactions between *Mycobacterium tuberculosis* and a macrophage[81].

To expand model predictions beyond metabolism, regulatory mechanisms are being integrated with metabolic models (FIG. 6b). Such integrated metabolic and regulatory models can improve phenotype predictions and even suggest novel regulatory interactions. This was carried out for the nutrient-controlled transcription-regulatory network of *S. cerevisiae*[115], which includes Boolean regulatory interactions between 55 transcription factors and 750 metabolic genes. This integrated regulatory–metabolic network could simulate growth under different environmental and genetic perturbations using regulatory FBA (rFBA). The model predicted new transcription-regulatory interactions, and elucidated regulatory cascades using chromatin immunoprecipitation data and the binding motifs of transcription factors. Although integrated models of metabolism and transcription regulation provide improved phenotype predictions, this study shows they can also expand our knowledge of regulatory interactions. It is anticipated that such models will further aid in simulating metabolism under conditions for which 'omic data are not available.

Variations on rFBA have been suggested[110,111]. Despite their success, rFBA and related methods have two main weaknesses. First, they assume binary responses for all transcription-regulatory interactions, when real biological systems exhibit a range of behaviour in transcription regulation, from binary to continuous. Second, few organisms have been studied enough to provide adequate regulatory information for rFBA. However, a method called probabilistic regulation of metabolism (PROM) addresses these concerns[116]. When ample transcriptomic data and candidate regulatory interactions (for example, from chromatin immunoprecipitation followed by sequencing (ChIP–seq)) are available, PROM can build a probabilistic model of the transcription-regulatory network of an organism and integrate it with the metabolic network, yielding an improved regulatory–metabolic network model. Moreover, PROM can apply intermediate responses (as opposed to binary ones), as it uses conditional probabilities for modelling transcription regulation instead of hard Boolean rules (FIG. 6c,d).

PROM was deployed to build the first integrated regulatory–metabolic network of *M. tuberculosis*[116]. Each transcription factor that modulates expression of metabolic genes was systematically deleted from the model, and *in silico* growth phenotypes were compared with experimentally measured phenotypes. PROM correctly predicted 96% of the transcription factor-knockout phenotypes, including those for five of the six transcription factors that were essential for optimal growth. This indicates that this method may help predict antibiotic targets from both regulatory and metabolic genes. Furthermore, the connections between the regulatory network and metabolism may represent novel regulatory targets for uncharacterized transcription factors.

An ecosystem of constraint-based methods

Individual COBRA methods can answer numerous scientific questions. However, multiple methods can be deployed in parallel to obtain additional insights into a question of interest. Moreover, different models can be

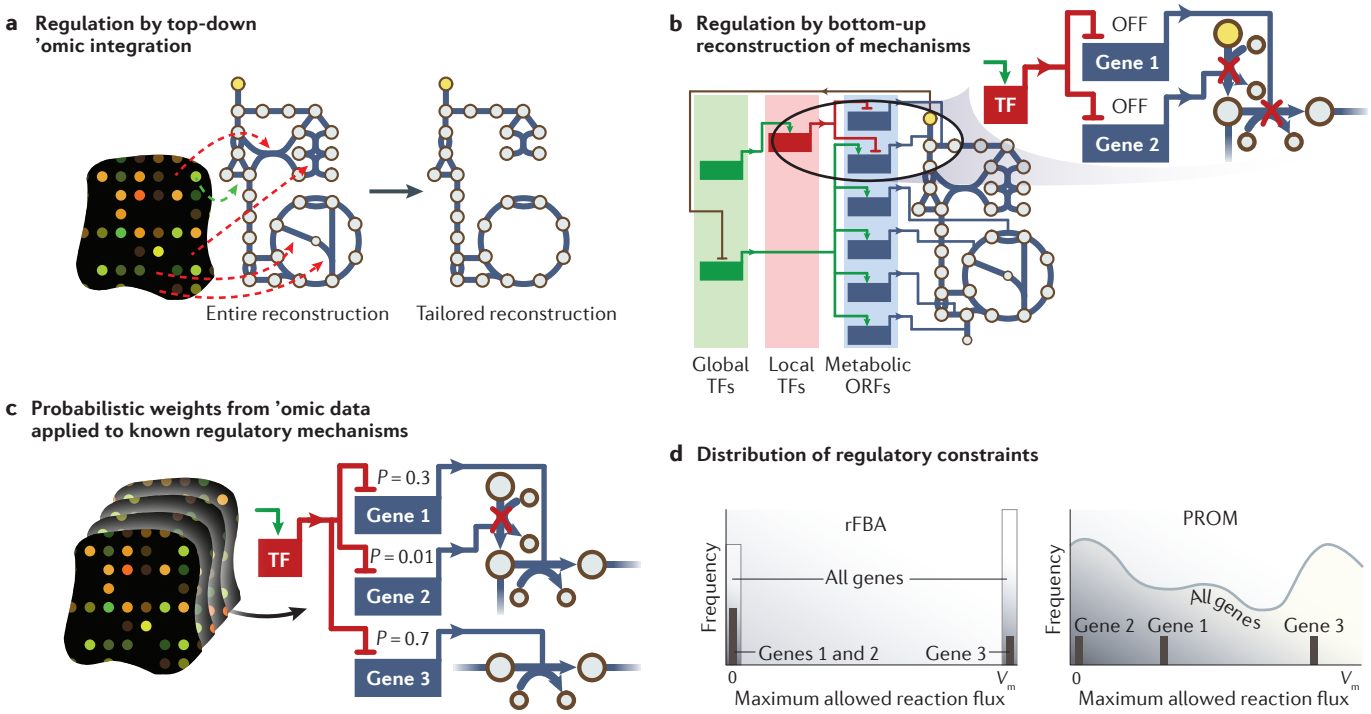


Figure 6 | **Incorporating regulation.** Two primary paradigms exist in constraint-based reconstruction and analysis (COBRA) modelling for integrating transcription regulation and metabolism. **a** | Algorithms such as GIMME[105] and model-building algorithm (MBA)[107] use high-throughput data and model simulations to identify which pathways were likely to be expressed and active in the cells when they were sampled. This results in a tailored, context-specific representation of the metabolic network. **b** | Algorithms such as regulatory flux balance analysis (rFBA)[109], integrated FBA (iFBA)[111] and steady-state regulatory FBA (SR-FBA)[110] incorporate detailed mathematical representations of the known molecular mechanisms of transcription regulation. These approaches contain binary regulatory logic that dictates, under a specific signal, which metabolic pathways are suppressed and cannot carry flux. **c** | Hybrid methods such as probabilistic regulation of metabolism (PROM)[116] are arising, in which transcriptomic and transcription factor (TF)-binding data are used to build a probabilistic regulatory–metabolic network. This allows for the elucidation of novel regulatory interactions and their immediate incorporation into model simulations. PROM also uses probabilistic measures to allow for a more continuous regulation of reaction flux. For example, gene 2 in the model shown is tightly regulated by a TF. Thus, when the TF is activated by a signal (green arrow), the reaction flux mediated by gene 2 is more tightly constrained than that mediated by gene 1, which is only loosely regulated. **d** | Whereas rFBA provides only binary, on–off constraints on reaction flux for these genes, the probabilistic measures in PROM allow for a more continuous regulation, as shown for genes 1 and 3. V_m, the velocity of the enzyme-catalysed reaction at an infinite concentration of substrate.

easily swapped or combined to test hypotheses relevant to different species. Thus, by using a community of methods and several data types, deeper insights into more profound questions may be attained. For example, COBRA methods have complemented each other and provided insight into the interactions in microbial communities.

The community structure in the microenvironment of an organism can shape the use of metabolic pathways. Organisms compete for scarce resources or depend on the metabolic capabilities of their cohabitants, and evolution often selects for cells that exploit this community structure[117]. COBRA methods are now characterizing the role for metabolism in microbial community structure[118–120], and these studies are providing insight into mutualism[121], competition[122], parasitism[81,123] and community evolution[117,124].

Mutualism. Synthetic mutualism between auxotrophic *E. coli* mutants was recently studied using COBRA methods[121]. The researchers grew pairs of auxotrophic mutants and then modelled their coupled metabolism using MOMA to identify mutant pairs that exchange essential metabolites to improve growth (FIG. 7a). FBA shadow prices demonstrated the balance between the cost (from metabolite loss) and the benefit (from receiving missing essential metabolites) to each rescued auxotroph. The cooperative efficiency (that is, the ratio of uptake benefit to production cost) recapitulated the observed growth of the co-cultures. Substantial increases in growth (FIG. 7b) were witnessed in co-cultures that exchanged beneficial but less costly metabolites (that is, in co-cultures that had higher cooperative efficiency). Although it is difficult to directly measure metabolite exchange between the auxotrophs, the computed cooperation efficiency provides an indirect quantitative assessment of the metabolite cross-feeding in this mutualistic system.

Competition. Metabolic competition for scarce nutrients has also been assessed with COBRA methods. Dynamic multispecies metabolic modelling (DMMM)

Shadow prices
A mathematical term that refers to the dual aspects of the linear programing problem. It represents the rate at which the objective value (for example, growth rate) changes as the supply of a particular resource (for example, a metabolite) increases.

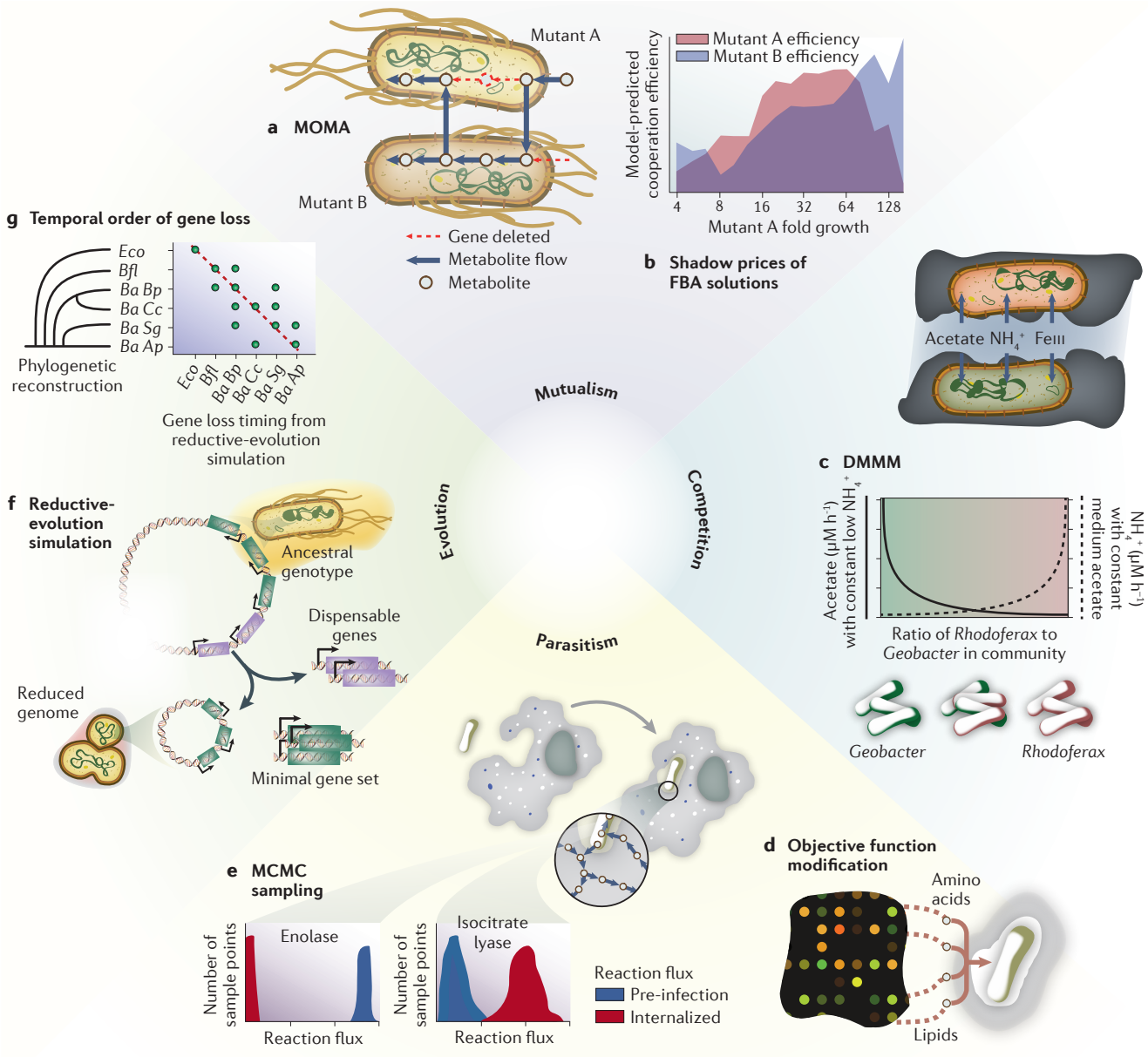


Figure 7 | **Integrating constraint-based reconstruction and analysis methods to study community interactions.** Constraint-based reconstruction and analysis (COBRA) methods are providing insight into the metabolic interactions in various types of microbial communities. **a** | To study the mutualistic behaviour of co-dependent mutant *Escherichia coli*, researchers used minimization of metabolic adjustment (MOMA)[50] to simulate synergistic growth of pairs of auxotrophic *E. coli*. **b** | Shadow prices from flux balance analysis (FBA) simulations of these pairs were used to compute cooperation efficiencies between strains, which were subsequently compared with measured fitness improvements. **c** | Competition in communities was modelled using dynamic multispecies metabolic modelling (DMMM)[122] to understand how communities of *Geobacter sulphurreducens* and *Rhodoferax ferrireducens* compete for resources, and how the demographics vary under different nutrient ratios, thereby affecting the efficiency of bioremediation efforts. **d** | Host–pathogen interactions between *Mycobacterium tuberculosis* and a human macrophage were studied using COBRA. Although transcriptomic data were employed to build host–pathogen models at different stages of infection, the cellular objective of internalized *M. tuberculosis* is not known, so refinements to the objective function were predicted from transcriptomic data to account for changes in required amounts of compounds such as lipids and amino acids. **e** | This information was used to compute flux states of internalized *M. tuberculosis* with Markov chain Monte Carlo (MCMC) sampling[32]. This demonstrated a suppression of central metabolism and activation of the glyoxylate shunt, represented here by enolase and isocitrate lyase, respectively. **f**,**g** | The role of communities in evolution has been studied using reductive-evolution simulations[117]. In particular, this method predicted the minimal set of genes needed for *Buchnera aphidicola* (*Ba*) to grow in the rich innards of the aphid. The predicted minimal gene sets (part **f**) and temporal order of gene loss (part **g**) were consistent with the gene content and phylogenetic structure of several *B. aphidicola* variants. *Ap*, *Acyrthosiphon pisum*; *Bfl*, '*Candidatus* Blochmannia floridanus'; *Bp*, *Baizongia pistaciae*; *Cc*, *Cinara cupressi*; *Eco*, *Escherichia coli*; *Sg*, *Schizaphis graminum*.

was used to characterize the competition for acetate, FeIII and ammonia between *Geobacter sulphurreducens* and *Rhodoferax ferrireducens* in subsurface anoxic environments[122] (FIG. 7c). DMMM simulated the growth rate of both organisms and the rates of change of external metabolites to dynamically predict population changes in the community. The community composition was predicted under geochemically distinct conditions of low, medium and high acetate flux. DMMM predicts that *R. ferrireducens* dominates the community when sufficient ammonia is available under low acetate flux, whereas *G. sulphurreducens* dominates under low ammonia concentrations and high acetate flux. This difference was attributed to the nitrogen fixation abilities of *G. sulphurreducens*, as well as to its acetate uptake rate being higher than that of *R. ferrireducens*. Moreover, it was also predicted that, under nitrogen-fixing conditions, *G. sulphurreducens* increases its respiration at the expense of biomass production, thus showing how balancing community structure can affect the efficacy of uranium bioremediation in low-ammonium zones.

Parasitism. Host–pathogen interactions have also been studied with COBRA methods[123]. A recent study modelled the metabolic interactions between a human alveolar macrophage and *M. tuberculosis*[81]. Context-specific models of infection were built with GIMME[105] and Shlomi-NBT-08 (REF. 106) using transcriptomic data from three types of *M. tuberculosis* infection. Next, the *M. tuberculosis* objective function was revised using infection-specific gene expression data to better represent the metabolic activity of the internalized pathogen (FIG. 7d). Gene deletion analysis was compared with *in vivo* gene essentiality data, and MCMC sampling was also used to demonstrate a substantial alteration in metabolic pathway use in *M. tuberculosis* during macrophage infection, including a suppression of glycolysis and an increased dependency on glyoxylate metabolism (FIG. 7e). This constraint of central metabolism during *M. tuberculosis* infection was also suggested by differential producibility analysis, a method that is related to FBA and identifies genes that affect the production of each metabolite in the metabolic network[125]. The suppression of certain metabolic pathways and the concurrent increased dependency on normally latent pathways may provide novel antibiotic targets.

Community evolution. In evolution, genetic drift and selective pressures cause organisms to optimize their cellular machinery for a particular niche[126]. This assumption that cellular optimization occurs has made COBRA methods useful tools for investigating hypotheses concerning organismal evolution, as reviewed recently[6]. In nature, the optimization of microbial metabolism is a multispecies affair, as demonstrated by the aphid endosymbiont *Buchnera aphidicola*. This member of the Enterobacteriaceae family has suffered drastic loss of genomic material as it evolved in the nutrient-rich innards of its host. *B. aphidicola* is related to *E. coli*, so a reductive-evolution simulation (a derivative of gene deletion analysis)[117] was carried out on the *E. coli* model to provide predictions of the minimal metabolic gene set. These predicted minimal sets are highly consistent with the metabolic gene content of *B. aphidicola* (FIG. 7f). In addition, the predicted temporal order of gene loss is consistent with the phylogenetically reconstructed gene loss timing among the genomes of five *B. aphidicola* strains[124] (FIG. 7g), suggesting that the bacterium optimized its pathway use for its new rich habitat. Interestingly, metabolic pathways that are retained in the computed minimal gene sets highlight the role of this bacterium in symbiotic evolution. These retained pathways contain reactions that are needed for producing riboflavin and essential amino acids that are lacking from the aphid diet, thereby highlighting their role in the symbiotic relationship[117]. Thus, COBRA methods are helping to describe how the community shapes the gene content in evolving symbiotic communities[6].

Future directions

Constraint-based modelling has rapidly evolved over the past two decades and now forms a foundation for achieving a genome-scale understanding of microbial metabolism. Before 2004, studies in this field focused on its conceptualization and algorithmic development. Thus, the methods developed were largely conceptual and used for studying the fundamental properties of metabolic networks, such as robustness, alternative optima and the functional consequences of metabolic network topology. After 2004, the field expanded to provide tools for addressing both basic and applied scientific questions focused on issues such as strain design, gap-filling[85] and evolution[6]. Despite the limitations in constraint-based modelling, its scope and uses are growing. GEMs and their corresponding analytical methods are expanding in scope beyond microbial metabolism, facilitating the analysis of 'omic data and directing scientific inquiry.

COBRA methods have gained rapid acceptance, because their focus on the governing constraints facilitates genome-scale analysis. However, the simplifying assumptions can also limit the scope of these tools. COBRA methods focus on steady-state flux, so the resultant models do not address metabolite concentrations, changes in biochemistry from pH and single-nucleotide polymorphisms, temporal metabolic changes or spatial constraints. Initial efforts are addressing some of the limitations and providing insight into these properties of metabolism[58,103,105,122,127], and additional efforts will further address these and other limitations.

Metabolism is involved in most cell processes and phenotypes. However, genome-scale models are extending beyond microbial metabolism to include transcription regulation[109,110,116], protein and transcript synthesis[128,129], signalling[112], plant and animal metabolism[39,58,64,113,114,130,131], and host–pathogen interactions[81,123,132]. The advances beyond microbial metabolism provide additional targets for drug discovery and metabolic engineering[133], and allow studies on medicine and crop engineering. This expansion of models and applications is requiring further evolution of COBRA methods, as well as theoretical breakthroughs to integrate non-stoichiometric networks

(for example, transcription regulation) with metabolism and to account for interactions with spatial constraints (for example, multicell metabolism[39,81,134]).

The past decade has witnessed a deluge of high-throughput data, including phenotypic screens, sequencing data, proteomics, metabolomics, and so forth. Recent studies have demonstrated that novel insights can be gained when these data are analysed in the context of GEMs[34,39,64,79,113,125,135]. As the models expand, they will increasingly aid in data interpretation, because they provide a structured context for high-throughput data analysis. Moreover, the biochemical mechanisms in these models will make use of 'omic analysis to inform experimental work.

Constraint-based modelling is already guiding discovery[85] by identifying missing metabolic and regulatory functions[84,86,94,115,116,136], predicting enzyme localization[89], suggesting novel drug targets[65,66,114], and aiding in strain design for chemical production[67,77,137–141] and biosensor development[142]. These studies are now increasingly directing experimental work. As models expand and are used to integrate 'omic data, COBRA methods will increasingly be deployed to guide scientific inquiry.

1. Feist, A. M., Herrgård, M. J., Thiele, I., Reed, J. L. & Palsson, B. Ø. Reconstruction of biochemical networks in microorganisms. *Nature Rev. Microbiol.* **7**, 129–143 (2009).
This review provides the detailed concepts of metabolic network reconstruction.
2. Thiele, I. & Palsson, B. Ø. A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nature Protoc.* **5**, 93–121 (2010).
3. Henry, C. S. *et al.* High-throughput generation, optimization and analysis of genome-scale metabolic models. *Nature Biotech.* **28**, 977–982 (2010).
4. Feist, A. M. & Palsson, B. Ø. The growing scope of applications of genome-scale metabolic reconstructions using *Escherichia coli*. *Nature Biotech.* **26**, 659–667 (2008).
5. Oberhardt, M. A., Palsson, B. Ø. & Papin, J. A. Applications of genome-scale metabolic reconstructions. *Mol. Syst. Biol.* **5**, 320 (2009).
6. Papp, B., Notebaart, R. A. & Pal, C. Systems-biology approaches for predicting genomic evolution. *Nature Rev. Genet.* **12**, 591–602 (2011).
A thorough review of how COBRA methods aid in the study of evolution.
7. Mahadevan, R., Palsson, B. Ø. & Lovley, D. R. *In situ* to *in silico* and back: elucidating the physiology and ecology of *Geobacter* spp. using genome-scale modelling. *Nature Rev. Microbiol.* **9**, 39–50 (2011).
8. Palsson, B. Ø. *Systems Biology: Properties of Reconstructed Networks* (Cambridge Univ. Press, 2006).
9. Feist, A. M. & Palsson, B. O. The biomass objective function. *Curr. Opin. Microbiol.* **13**, 344–349 (2010).
A description of how the biomass objective function is formulated.
10. Fell, D. A. & Small, J. R. Fat synthesis in adipose tissue. An examination of stoichiometric constraints. *Biochem. J.* **238**, 781–786 (1986).
11. Watson, M. R. Metabolic maps for the Apple II. *Biochem. Soc. Trans.* **12**, 1093–1094 (1984).
12. Edwards, J. S. & Palsson, B. O. Systems properties of the *Haemophilus influenzae* Rd metabolic genotype. *J. Biol. Chem.* **274**, 17410–17416 (1999).
13. Edwards, J. S. & Palsson, B. O. The *Escherichia coli* MG1655 *in silico* metabolic genotype: its definition, characteristics, and capabilities. *Proc. Natl Acad. Sci. USA* **97**, 5528–5533 (2000).
14. Reed, J. L., Famili, I., Thiele, I. & Palsson, B. O. Towards multidimensional genome annotation. *Nature Rev. Genet.* **7**, 130–141 (2006).
15. Kim, T. Y., Kim, H. U. & Lee, S. Y. Data integration and analysis of biological networks. *Curr. Opin. Biotechnol.* **21**, 78–84 (2010).
16. Sauer, U. Metabolic networks in motion: 13C-based flux analysis. *Mol. Syst. Biol.* **2**, 62 (2006).
17. Papin, J. A. *et al.* Comparison of network-based pathway analysis methods. *Trends Biotechnol.* **22**, 400–405 (2004).
An assessment of the differences between EFM and ExPa analysis.
18. Trinh, C. T., Wlaschin, A. & Srienc, F. Elementary mode analysis: a useful metabolic pathway analysis tool for characterizing cellular metabolism. *Appl. Microbiol. Biotechnol.* **81**, 813–826 (2009).
19. Llaneras, F. & Pico, J. Which metabolic pathways generate and characterize the flux space? A comparison among elementary modes, extreme pathways and minimal generators. *J. Biomed. Biotechnol.* **2010**, 753904 (2010).
20. Stelling, J., Klamt, S., Bettenbrock, K., Schuster, S. & Gilles, E. D. Metabolic network structure determines key aspects of functionality and regulation. *Nature* **420**, 190–193 (2002).
21. Trinh, C. T., Unrean, P. & Srienc, F. Minimal *Escherichia coli* cell for the most efficient production of ethanol from hexoses and pentoses. *Appl. Environ. Microbiol.* **74**, 3634–3643 (2008).
22. Imielinski, M. & Belta, C. Exploiting the pathway structure of metabolism to reveal high-order epistasis. *BMC Syst. Biol.* **2**, 40 (2008).
23. Wessely, F. *et al.* Optimal regulatory strategies for metabolic pathways in *Escherichia coli* depending on protein costs. *Mol. Syst. Biol.* **7**, 515 (2011).
24. Schilling, C. H. & Palsson, B. Ø. Assessment of the metabolic capabilities of *Haemophilus influenzae* Rd through a genome-scale pathway analysis. *J. Theor. Biol.* **203**, 249–283 (2000).
25. Yeung, M., Thiele, I. & Palsson, B. Ø. Estimation of the number of extreme pathways for metabolic networks. *BMC Bioinformatics* **8**, 363 (2007).
26. Klamt, S. & Stelling, J. Combinatorial complexity of pathway analysis in metabolic networks. *Mol. Biol. Rep.* **29**, 233–236 (2002).
27. Kaleta, C., de Figueiredo, L. F. & Schuster, S. Can the whole be less than the sum of its parts? Pathway analysis in genome-scale metabolic networks using elementary flux patterns. *Genome Res.* **19**, 1872–1883 (2009).
28. Rezola, A. *et al.* Exploring metabolic pathways in genome-scale networks via generating flux modes. *Bioinformatics* **27**, 534–540 (2011).
29. Ip, K., Colijn, C. & Lun, D. S. Analysis of complex metabolic behavior through pathway decomposition. *BMC Syst. Biol.* **5**, 91 (2011).
30. Chan, S. H. & Ji, P. Decomposing flux distributions into elementary flux modes in genome-scale metabolic networks. *Bioinformatics* **27**, 2256–2262 (2011).
31. Braunstein, A., Mulet, R. & Pagnani, A. Estimating the size of the solution space of metabolic networks. *BMC Bioinformatics* **9**, 240 (2008).
32. Schellenberger, J. & Palsson, B. Ø. Use of randomized sampling for analysis of metabolic networks. *J. Biol. Chem.* **284**, 5457–5461 (2009).
33. Almaas, E., Kovacs, B., Vicsek, T., OItvai, Z. N. & Barabasi, A. L. Global organization of metabolic fluxes in the bacterium *Escherichia coli*. *Nature* **427**, 839–843 (2004).
34. Bordel, S., Agren, R. & Nielsen, J. Sampling the solution space in genome-scale metabolic networks reveals transcriptional regulation in key enzymes. *PLoS Comput. Biol.* **6**, e1000859 (2010).
35. Mo, M. L., Palsson, B. Ø. & Herrgård, M. J. Connecting extracellular metabolomic measurements to intracellular flux states in yeast. *BMC Syst. Biol.* **3**, 37 (2009).
36. Barrett, C. L., Herrgard, M. J. & Palsson, B. Decomposing complex reaction networks using random sampling, principal component analysis and basis rotation. *BMC Syst. Biol.* **3**, 30 (2009).
37. Thiele, I., Price, N. D., Vo, T. D. & Palsson, B. Ø. Candidate metabolic network states in human mitochondria. Impact of diabetes, ischemia, and diet. *J. Biol. Chem.* **280**, 11683–11695 (2005).
38. Price, N. D., Schellenberger, J. & Palsson, B. O. Uniformsampling of steady-state flux spaces: means to design experiments and to interpret enzymopathies. *Biophys. J.* **87**, 2172–2186 (2004).
39. Lewis, N. E. *et al.* Large-scale *in silico* modeling of metabolic interactions between cell types in the human brain. *Nature Biotech.* **28**, 1279–1285 (2010).
40. Orth, J. D., Thiele, I. & Palsson, B. Ø. What is flux balance analysis? *Nature Biotech.* **28**, 245–248 (2010).
A primer to the FBA method.
41. Varma, A. & Palsson, B. O. Stoichiometric flux balance models quantitatively predict growth and metabolic by-product secretion in wild-type *Escherichia coli* W3110. *Appl. Environ. Microbiol.* **60**, 3724–3731 (1994).
42. Edwards, J. S., Ibarra, R. U. & Palsson, B. O. *In silico* predictions of *Escherichia coli* metabolic capabilities are consistent with experimental data. *Nature Biotech.* **19**, 125–130 (2001).
43. Lewis, N. E. *et al.* Omic data from evolved *E. coli* are consistent with computed optimal growth from genome-scale models. *Mol. Syst. Biol.* **6**, 390 (2010).
44. Teusink, B., Wiersma, A., Jacobs, L., Notebaart, R. A. & Smid, E. J. Understanding the adaptive growth strategy of *Lactobacillus plantarum* by *in silico* optimisation. *PLoS Comput. Biol.* **5**, e1000410 (2009).
45. Wang, Z. & Zhang, J. Impact of gene expression noise on organismal fitness and the efficacy of natural selection. *Proc. Natl Acad. Sci. USA* **108**, E67–E76 (2011).
46. Goffin, P. *et al.* Understanding the physiology of *Lactobacillus plantarum* at zero growth. *Mol. Syst. Biol.* **6**, 413 (2010).
47. Lee, S., Palakornkule, C., Domach, M. M. & Grossmann, I. E. Recursive MILP model for finding all the alternate optima in LP models for metabolic networks. *Comput. Chem. Eng.* **24**, 711–716 (2000).
48. Gudmundsson, S. & Thiele, I. Computationally efficient flux variability analysis. *BMC Bioinformatics* **11**, 489 (2010).
49. Burgard, A. P., Vaidyaraman, S. & Maranas, C. D. Minimal reaction sets for *Escherichia coli* metabolism under different growth requirements and uptake environments. *Biotechnol. Prog.* **17**, 791–797 (2001).
50. Segre, D., Vitkup, D. & Church, G. M. Analysis of optimality in natural and perturbed metabolic networks. *Proc. Natl Acad. Sci. USA* **99**, 15112–15117 (2002).
51. Benyamini, T., Folger, O., Ruppin, E. & Shlomi, T. Flux balance analysis accounting for metabolite dilution. *Genome Biol.* **11**, R43 (2010).
52. Colijn, C. *et al.* Interpreting expression data with metabolic flux models: predicting *Mycobacterium tuberculosis* mycolic acid production. *PLoS Comput. Biol.* **5**, e1000489 (2009).
53. van Berlo, R. J. P. *et al.* Predicting metabolic fluxes using gene expression differences as constraints. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **8**, 206–216 (2011).
54. Holzhutter, H. G. The principle of flux minimization and its application to estimate stationary fluxes in metabolic networks. *Eur. J. Biochem.* **271**, 2905–2922 (2004).
55. Ponce de Leon, M., Cancela, H. & Acerenza, L. A strategy to calculate the patterns of nutrient consumption by microorganisms applying a two-level optimisation principle to reconstructed metabolic networks. *J. Biol. Phys.* **34**, 73–90 (2008).
56. Murabito, E., Simeonidis, E., Smallbone, K. & Swinton, J. Capturing the essence of a metabolic network: a flux balance analysis approach. *J. Theor. Biol.* **260**, 445–452 (2009).

57. Beg, Q. K. *et al.* Intracellular crowding defines the mode and sequence of substrate uptake by *Escherichia coli* and constrains its metabolic activity. *Proc. Natl Acad. Sci. USA* **104**, 12663–12668 (2007).
58. Vazquez, A., Markert, E. K. & Oltvai, Z. N. Serine biosynthesis with one carbon catabolism and the glycine cleavage system represents a novel pathway for ATP generation. *PLoS ONE* **6**, e25881 (2011).
59. Zhuang, K., Vemuri, G. N. & Mahadevan, R. Economics of membrane occupancy and respiro-fermentation. *Mol. Syst. Biol.* **7**, 500 (2011).
60. Papp, B., Pal, C. & Hurst, L. D. Metabolic network analysis of the causes and evolution of enzyme dispensability in yeast. *Nature* **429**, 661–664 (2004).
61. Deutscher, D., Meilijson, I., Kupiec, M. & Ruppin, E. Multiple knockout analysis of genetic robustness in the yeast metabolic network. *Nature Genet.* **38**, 993–998 (2006).
62. Shlomi, T., Berkman, O. & Ruppin, E. Regulatory on/off minimization of metabolic flux changes after genetic perturbations. *Proc. Natl Acad. Sci. USA* **102**, 7695–7700 (2005).
63. Kim, P. J. *et al.* Metabolite essentiality elucidates robustness of *Escherichia coli* metabolism. *Proc. Natl Acad. Sci. USA* **104**, 13638–13642 (2007).
64. Chang, R. L., Xie, L., Xie, L., Bourne, P. E. & Palsson, B. Ø. Drug off-target effects predicted using structural analysis in the context of a metabolic network model. *PLoS Comput. Biol.* **6**, e1000938 (2010).
65. Shen, Y. *et al.* Blueprint for antimicrobial hit discovery targeting metabolic networks. *Proc. Natl Acad. Sci. USA* **107**, 1082–1087 (2010).
A study in which metabolic networks are used to search for antimicrobials.
66. Kim, H. U. *et al.* Integrative genome-scale metabolic analysis of *Vibrio vulnificus* for drug targeting and discovery. *Mol. Syst. Biol.* **7**, 460 (2011).
67. Fong, S. S. *et al. In silico* design and adaptive evolution of *Escherichia coli* for production of lactic acid. *Biotechnol. Bioeng.* **91**, 643–648 (2005).
68. Burgard, A. P., Pharkya, P. & Maranas, C. D. Optknock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. *Biotechnol. Bioeng.* **84**, 647–657 (2003).
69. Feist, A. M. *et al.* Model-driven evaluation of the production potential for growth-coupled products of Escherichia coli. *Metab. Eng.* **12**,173–186 (2010).
70. Tepper, N. & Shlomi, T. Predicting metabolic engineering knockout strategies for chemical production: accounting for competing pathways. *Bioinformatics* **26**, 536–543 (2010).
71. Patil, K. R., Rocha, I., Forster, J. & Nielsen, J. Evolutionary programming as a platform for *in silico* metabolic engineering. *BMC Bioinformatics* **6**, 308 (2005).
72. Lun, D. S. *et al.* Large-scale identification of genetic design strategies using local search. *Mol. Syst. Biol.* **5**, 296 (2009).
73. Yousofshahi, M., Lee, K. & Hassoun, S. Probabilistic pathway construction. *Metab. Eng.* **13**, 435–444 (2011).
74. Rodrigo, G., Carrera, J., Prather, K. J. & Jaramillo, A. DESHARKY: automatic design of metabolic pathways for optimal cell growth. *Bioinformatics* **24**, 2554–2556 (2008).
75. Bar-Even, A., Noor, E., Lewis, N. E. & Milo, R. Design and analysis of synthetic carbon fixation pathways. *Proc. Natl Acad. Sci. USA* **107**, 8889–8894 (2010).
76. Delebecque, C. J., Lindner, A. B., Silver, P. A. & Aldaye, F. A. Organization of intracellular reactions with rationally designed RNA assemblies. *Science* **333**, 470–474 (2011).
77. Yim, H. *et al.* Metabolic engineering of *Escherichia coli* for direct production of 1,4-butanediol. *Nature Chem. Biol.* **7**, 445–452 (2011).
A detailed study in which several computational and experimental technologies are used to engineer a microorganism to synthesize 1,4-butanediol using.
78. Pharkya, P., Burgard, A. P. & Maranas, C. D. OptStrain: a computational framework for redesign of microbial production systems. *Genome Res.* **14**, 2367–2376 (2004).
79. Szappanos, B. *et al.* An integrated approach to characterize genetic interaction networks in yeast metabolism. *Nature Genet.* **43**, 656–662 (2011).
80. Hsiao, T. L., Revelles, O., Chen, L., Sauer, U. & Vitkup, D. Automatic policing of biochemical annotations using genomic correlations. *Nature Chem. Biol.* **6**, 34–40 (2010).
81. Bordbar, A., Lewis, N. E., Schellenberger, J., Palsson, B. Ø. & Jamshidi, N. Insight into human alveolar macrophage and *M. tuberculosis* interactions via metabolic reconstructions. *Mol. Syst. Biol.* **6**, 422 (2010).
82. Schuster, S., Pfeiffer, T. & Fell, D. A. Is maximization of molar yield in metabolic networks favoured by evolution? *J. Theor. Biol.* **252**, 497–504 (2008).
A critical assessment of the assumptions in FBA.
83. Milne, C. B. *et al.* Metabolic network reconstruction and genome-scale model of butanol-producing strain *Clostridium beijerinckii* NCIMB 8052. *BMC Syst. Biol.* **5**, 130 (2011).
84. Reed, J. L. *et al.* Systems approach to refining genome annotation. *Proc. Natl Acad. Sci. USA* **103**, 17480–17484 (2006).
85. Orth, J. D. & Palsson, B. Ø. Systematizing the generation of missing metabolic knowledge. *Biotechnol. Bioeng.* **107**, 403–412 (2010).
86. Satish Kumar, V., Dasika, M. S. & Maranas, C. D. Optimization based automated curation of metabolic reconstructions. *BMC Bioinformatics* **8**, 212 (2007).
87. Kumar, V. S. & Maranas, C. D. GrowMatch: an automated method for reconciling *in silico/in vivo* growth predictions. *PLoS Comput. Biol.* **5**, e1000308 (2009).
88. Suthers, P. F., Zomorrodi, A. & Maranas, C. D. Genome-scale gene/reaction essentiality and synthetic lethality analysis. *Mol. Syst. Biol.* **5**, 301 (2009).
89. Mintz-Oron, S., Aharoni, A., Ruppin, E. & Shlomi, T. Network-based prediction of metabolic enzymes' subcellular localization. *Bioinformatics* **25**, i247–i252 (2009).
90. Burgard, A. P. & Maranas, C. D. Optimization-based framework for inferring and testing hypothesized metabolic objective functions. *Biotechnol. Bioeng.* **82**, 670–677 (2003).
91. Knorr, A. L., Jain, R. & Srivastava, R. Bayesian-based selection of metabolic objective functions. *Bioinformatics* **23**, 351–357 (2007).
92. Schuetz, R., Kuepfer, L. & Sauer, U. Systematic evaluation of objective functions for predicting intracellular fluxes in *Escherichia coli*. *Mol. Syst. Biol.* **3**, 119 (2007).
93. Gianchandani, E. P., Oberhardt, M. A., Burgard, A. P., Maranas, C. D. & Papin, J. A. Predicting biological system objectives *de novo* from internal state measurements. *BMC Bioinformatics* **9**, 43 (2008).
94. Zomorrodi, A. R. & Maranas, C. D. Improving the iMM904 *S. cerevisiae* metabolic model using essentiality and synthetic lethality data. *BMC Syst. Biol.* **4**, 178 (2010).
95. Beard, D. A., Liang, S. D. & Qian, H. Energy balance for analysis of complex metabolic networks. *Biophys. J.* **83**, 79–86 (2002).
96. Schellenberger, J., Lewis, N. E. & Palsson, B. Ø. Elimination of thermodynamically infeasible loops in steady-state metabolic models. *Biophys. J.* **100**, 544–553 (2011).
97. Price, N. D., Thiele, I. & Palsson, B. Ø. Candidate states of *Helicobacter pylori*'s genome-scale metabolic network upon application of "loop law" thermodynamic constraints. *Biophys. J.* **90**, 3919–3928 (2006).
98. Henry, C. S., Broadbelt, L. J. & Hatzimanikatis, V. Thermodynamics-based metabolic flux analysis. *Biophys. J.* **92**, 1792–1805 (2007).
99. Fleming, R. M., Thiele, I., Provan, G. & Nasheuer, H. P. Integrated stoichiometric, thermodynamic and kinetic modelling of steady state metabolism. *J. Theor. Biol.* **264**, 683–692 (2010).
100. Kummel, A., Panke, S. & Heinemann, M. Putative regulatory sites unraveled by network-embedded thermodynamic analysis of metabolome data. *Mol. Syst. Biol.* **2**, 2006.0034 (2006).
101. Jankowski, M. D., Henry, C. S., Broadbelt, L. J. & Hatzimanikatis, V. Group contribution method for thermodynamic analysis of complex metabolic networks. *Biophys. J.* **95**, 1487–1499 (2008).
102. Henry, C. S., Jankowski, M. D., Broadbelt, L. J. & Hatzimanikatis, V. Genome-scale thermodynamic analysis of *Escherichia coli* metabolism. *Biophys. J.* **90**, 1453–1461 (2006).
103. Hoppe, A., Hoffmann, S. & Holzhutter, H. G. Including metabolite concentrations into flux balance analysis: thermodynamic realizability as a constraint on flux distributions in metabolic networks. *BMC Syst. Biol.* **1**, 23 (2007).
104. Fleming, R. M., Thiele, I. & Nasheuer, H. P. Quantitative assignment of reaction directionality in constraint-based models of metabolism: application to *Escherichia coli*. *Biophys. Chem.* **145**, 47–56 (2009).
105. Becker, S. A. & Palsson, B. O. Context-specific metabolic networks are consistent with experiments. *PLoS Comput. Biol.* **4**, e1000082 (2008).
106. Shlomi, T., Cabili, M. N., Herrgård, M. J., Palsson, B. Ø. & Ruppin, E. Network-based prediction of human tissue-specific metabolism. *Nature Biotech.* **26**, 1003–1010 (2008).
107. Jerby, L., Shlomi, T. & Ruppin, E. Computational reconstruction of tissue-specific metabolic models: application to human liver metabolism. *Mol. Syst. Biol.* **6**, 401 (2010).
108. Jensen, P. A. & Papin, J. A. Functional integration of a metabolic network model and expression data without arbitrary thresholding. *Bioinformatics* **27**, 541–547 (2011).
109. Covert, M. W., Knight, E. M., Reed, J. L., Herrgard, M. J. & Palsson, B. O. Integrating high-throughput and computational data elucidates bacterial networks. *Nature* **429**, 92–96 (2004).
110. Shlomi, T., Eisenberg, Y., Sharan, R. & Ruppin, E. A genome-scale computational study of the interplay between transcriptional regulation and metabolism. *Mol. Syst. Biol.* **3**, 101 (2007).
111. Covert, M. W., Xiao, N., Chen, T. J. & Karr, J. R. Integrating metabolic, transcriptional regulatory and signal transduction models in *Escherichia coli*. *Bioinformatics* **24**, 2044–2050 (2008).
112. Lee, J. M., Gianchandani, E. P., Eddy, J. A. & Papin, J. A. Dynamic analysis of integrated signaling, metabolic, and regulatory networks. *PLoS Comput. Biol.* **4**, e1000086 (2008).
113. Folger, O. *et al.* Predicting selective drug targets in cancer through metabolic networks. *Mol. Syst. Biol.* **7**, 501 (2011).
114. Frezza, C. *et al.* Haem oxygenase is synthetically lethal with the tumour suppressor fumarate hydratase. *Nature* **477**, 225–228 (2011).
115. Herrgård, M. J., Lee, B. S., Portnoy, V. & Palsson, B. Ø. Integrated analysis of regulatory and metabolic networks reveals novel regulatory mechanisms in *Saccharomyces cerevisiae*. *Genome Res.* **16**, 627–635 (2006).
116. Chandrasekaran, S. & Price, N. D. Probabilistic integrative modeling of genome-scale metabolic and regulatory networks in *Escherichia coli* and *Mycobacterium tuberculosis*. *Proc. Natl Acad. Sci. USA* **107**, 17845–17850 (2010).
An approach for integrating regulatory networks with metabolic modelling.
117. Pal, C. *et al.* Chance and necessity in the evolution of minimal metabolic networks. *Nature* **440**, 667–670 (2006).
118. Stolyar, S. *et al.* Metabolic modeling of a mutualistic microbial community. *Mol. Syst. Biol.* **3**, 92 (2007).
119. Taffs, R. *et al. In silico* approaches to study mass and energy flows in microbial consortia: a syntrophic case study. *BMC Syst. Biol.* **3**, 114 (2009).
120. Klitgord, N. & Segre, D. Environments that induce synthetic microbial ecosystems. *PLoS Comput. Biol.* **6**, e1001002 (2010).
121. Wintermute, E. H. & Silver, P. A. Emergent cooperation in microbial metabolism. *Mol. Syst. Biol.* **6**, 407 (2010).
122. Zhuang, K. *et al.* Genome-scale dynamic modeling of the competition between *Rhodoferax* and *Geobacter* in anoxic subsurface environments. *ISME J.* **5**, 305–316 (2011).
123. Huthmacher, C., Hoppe, A., Bulik, S. & Holzhutter, H. G. Antimalarial drug targets in *Plasmodium falciparum* predicted by stage-specific metabolic network analysis. *BMC Syst. Biol.* **4**, 120 (2010).
124. Yizhak, K., Tuller, T., Papp, B. & Ruppin, E. Metabolic modeling of endosymbiont genome reduction on a temporal scale. *Mol. Syst. Biol.* **7**, 479 (2011).
125. Bonde, B. K., Beste, D. J., Laing, E., Kierzek, A. M. & McFadden, J. Differential producibility analysis (DPA) of transcriptomic data with metabolic networks: deconstructing the metabolic response of *M. tuberculosis*. *PLoS Comput. Biol.* **7**, e1002060 (2011).
126. Nam, H., Conrad, T. M. & Lewis, N. E. The role of cellular objectives and selective pressures in metabolic pathway evolution. *Curr. Opin. Biotechnol.* **22**, 595–600 (2011).
127. Srinivasan, K. & Mahadevan, R. Characterization of proton production and consumption associated with microbial metabolism. *BMC Biotechnol.* **10**, 2 (2010).

128. Thiele, I., Jamshidi, N., Fleming, R. M. & Palsson, B. Ø. Genome-scale reconstruction of *Escherichia coli*'s transcriptional and translational machinery: a knowledge base, its mathematical formulation, and its functional characterization. *PLoS Comput. Biol.* **5**, e1000312 (2009).
129. Thiele, I., Fleming, R. M., Bordbar, A., Schellenberger, J. & Palsson, B. Ø. Functional characterization of alternate optimal solutions of *Escherichia coli*'s transcriptional and translational machinery. *Biophys. J.* **98**, 2072–2081 (2010).
130. Sigurdsson, M. I., Jamshidi, N., Steingrimsson, E., Thiele, I. & Palsson, B. Ø. A detailed genome-wide reconstruction of mouse metabolism based on human Recon 1. *BMC Syst. Biol.* **4**, 140 (2010).
131. Sweetlove, L. J. & Ratcliffe, R. G. Flux-balance modelling of plant metabolism. *Frontiers Plant Sci.* **2**, 38 (2011).
132. Gille, C. *et al.* HepatoNet1: a comprehensive metabolic reconstruction of the human hepatocyte for the analysis of liver physiology. *Mol. Syst. Biol.* **6**, 411 (2010).
133. Kim, J. & Reed, J. L. OptORF: Optimal metabolic and regulatory perturbations for metabolic engineering of microbial strains. *BMC Syst. Biol.* **4**, 53 (2010).
134. Bordbar, A. *et al.* A multi-tissue type genome-scale metabolic network for analysis of whole-body systems physiology. *BMC Syst. Biol.* **5**, 180 (2011).
135. Zhang, Y. *et al.* Three-dimensional structural view of the central metabolic network of *Thermotoga maritima*. *Science* **325**, 1544–1549 (2009).
136. Barua, D., Kim, J. & Reed, J. L. An automated phenotype-driven approach (GeneForce) for refining metabolic and regulatory models. *PLoS Comput. Biol.* **6**, e1000970 (2010).
137. Lee, K. H., Park, J. H., Kim, T. Y., Kim, H. U. & Lee, S. Y. Systems metabolic engineering of *Escherichia coli* for L-threonine production. *Mol. Syst. Biol.* **3**, 149 (2007).
138. Alper, H., Jin, Y. S., Moxley, J. F. & Stephanopoulos, G. Identifying gene targets for the metabolic engineering of lycopene biosynthesis in *Escherichia coli*. *Metab. Eng.* **7**, 155–164 (2005).
139. Kennedy, C. J., Boyle, P. M., Waks, Z. & Silver, P. A. Systems-level engineering of nonfermentative metabolism in yeast. *Genetics* **183**, 385–397 (2009).
140. Xu, P., Ranganathan, S., Fowler, Z. L., Maranas, C. D. & Koffas, M. A. Genome-scale metabolic network modeling results in minimal interventions that cooperatively force carbon flux towards malonyl-CoA. *Metab. Eng.* **13**, 578–587 (2011).
141. Chemler, J. A., Fowler, Z. L., McHugh, K. P. & Koffas, M. A. Improving NADPH availability for natural product biosynthesis in *Escherichia coli* by metabolic engineering. *Metab. Eng.* **12**, 96–104 (2010).
142. Tepper, N. & Shlomi, T. Computational design of auxotrophy-dependent microbial biosensors for combinatorial metabolic engineering experiments. *PLoS ONE* **6**, e16274 (2011).

Acknowledgements
The authors thank M. Abrams, J. Monk, D. Taylor, E. O'Brien, A. Feist, J. Lerman, D. Zielinski, R. Chang, K. Zengler, A. Bordbar, R. Fleming, N. Price and E. Ruppin for discussions and comments on this Review, and numerous members of the COBRA community for help with compiling methods. This work was funded by the US National Institutes of Health (grant R01GM057089) and the Office of Science (Biological and Environmental Research) for the US Department of Energy (grant DE-SC0004485).

Competing interests statement
The authors declare no competing financial interests.

FURTHER INFORMATION

Bernhard O. Palsson's homepage: http://sbrg.ucsd.edu
COBRA software:
http://sourceforge.net/apps/mediawiki/opencobra/

SUPPLEMENTARY INFORMATION

See online article: S1 (table) | S2 (table)

ALL LINKS ARE ACTIVE IN THE ONLINE PDF