

## Review Article

# Human Systems Biology and Metabolic Modelling: A Review—From Disease Metabolism to Precision Medicine

Claudio Angione <sup>1,2</sup>

<sup>1</sup>Department of Computer Science and Information Systems, Teesside University, UK

<sup>2</sup>Healthcare Innovation Centre, Teesside University, UK

Correspondence should be addressed to Claudio Angione; [c.angione@tees.ac.uk](mailto:c.angione@tees.ac.uk)

Received 1 November 2018; Revised 7 February 2019; Accepted 20 May 2019; Published 9 June 2019

Academic Editor: Stanley Brul

Copyright © 2019 Claudio Angione. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In cell and molecular biology, metabolism is the only system that can be fully simulated at genome scale. Metabolic systems biology offers powerful abstraction tools to simulate all known metabolic reactions in a cell, therefore providing a snapshot that is close to its observable phenotype. In this review, we cover the 15 years of human metabolic modelling. We show that, although the past five years have not experienced large improvements in the size of the gene and metabolite sets in human metabolic models, their accuracy is rapidly increasing. We also describe how condition-, tissue-, and patient-specific metabolic models shed light on cell-specific changes occurring in the metabolic network, therefore predicting biomarkers of disease metabolism. We finally discuss current challenges and future promising directions for this research field, including machine/deep learning and precision medicine. In the omics era, profiling patients and biological processes from a multiomic point of view is becoming more common and less expensive. Starting from multiomic data collected from patients and N-of-1 trials where individual patients constitute different case studies, methods for model-building and data integration are being used to generate patient-specific models. Coupled with state-of-the-art machine learning methods, this will allow characterizing each patient's disease phenotype and delivering precision medicine solutions, therefore leading to preventative medicine, reduced treatment, and *in silico* clinical trials.

## 1. Introduction

With the advent of bioinformatics and computational biology, computational and mathematical techniques can provide accurate simulation of biological processes. The most widely used approaches to analyze omics data mainly focus on genomics, transcriptomics, and proteomics, through differential expression or network-based coexpression analysis. However, genes and their expression alone do not always constitute a reliable indicator of cellular phenotype. When characterizing a phenotypic outcome, relying solely on gene or protein expression profiles will miss the highly nonlinear interaction between these biological layers. Such approaches often overlook the metabolic level, the dense network of biochemical reactions occurring in a cell with the aim of converting nutrients into energy and cellular building blocks.

Being the best-characterized network in biological systems and also the closest to the phenotype, metabolism is arguably the best indicator for the cell physiological state [1].

Once considered only a passive result of the state of a cell, it is now widely recognized as a main contributor to cellular behavior. More specifically, it is a key player in a number of diseases, including diabetes, neurodegenerative diseases, and cancer, where altered metabolism is now accepted as a hallmark [2].

The recent availability of high-throughput data regarding multiple layers of biological organization (omics) allows mapping cellular processes at the levels of genes, mRNA, proteins, and metabolites (Figure 1). In a single experiment, these measurements are often at both the “genotype” level (i.e., referring to the genetic elements on a genome) and at the “phenotype” level (the form and function of the cell). A fundamental question in systems biology is the definition and understanding of the genotype-phenotype relationship [3]. A mechanistic link between genotype and phenotype is offered by genome-scale metabolic models, which contain all known biochemical reactions occurring in a cell. Such models have been generated taking into account decades of studies in

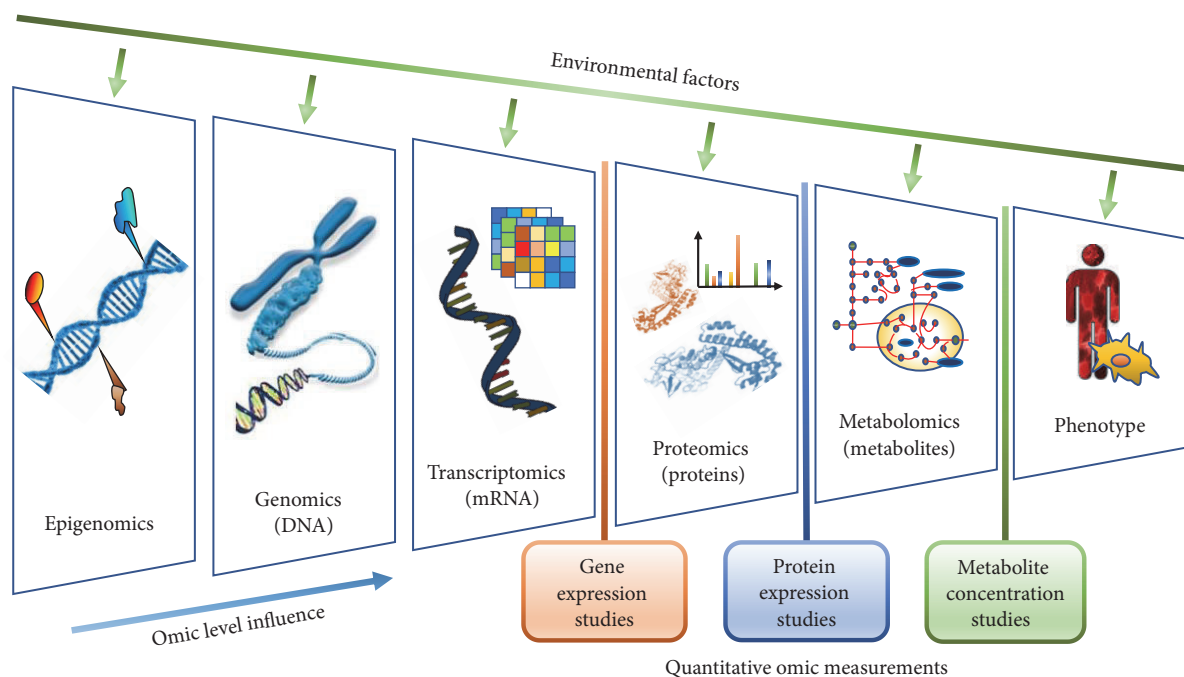


FIGURE 1: The observable phenotype of a cell is a result of complex interactions and feedback loops among several omic layers, each influenced by environmental perturbations. In order of “distance” from the cell phenotype, these are epigenomics (epigenetic markers that affect gene activity and expression), genomics (DNA containing the genetic code of the cell), transcriptomics (the RNA encoded by the genome), proteomics (the set of proteins produced as a result of gene expression and subsequent posttranslational modifications), and metabolomics (the set of metabolites and metabolic reactions taking place in the cell). Although each omic layer can be studied alone, no single-omic layer has achieved a satisfactory correlation with the phenotypic observables. As a result, in recent years, a multiomic approach has been adopted where all layers are considered together, and the effect of interactions and feedback is taken into consideration.

biochemistry and in most cases are able to predict the cellular phenotype with high accuracy.

Constraint-based modelling is the most widely used approach to model the behavior of metabolism, often assuming that cells have to fulfill a given task (e.g., ATP production, growth, or proliferation) or to optimize the production of a given compound. Such models have two main advantages: first, they do not need dynamic or kinetic data as they are based on mass balance across the metabolic network; second, they are suitable for integration of different omic layers at genome scale to improve their predictive performance. In particular, multiomic “vertical” integration methods have been proposed to include omic layers (mainly transcriptomics and proteomics) [4–6]. Conversely, “horizontal” integration methods have focused on modelling different environments, cancers or growth conditions starting from the same model [7–11]. Such multiomic integration in genome-scale models has provided a mechanistic link between the genotype and their phenotypic observables [12–15]. This is a key added feature that such models possess if compared with genome-wide association studies (GWAS), which are able to associate gene variations to phenotypic traits, but not to provide a mechanistic explanation of the associations observed [16].

In this review, we focus on metabolic systems biology, and in particular on human metabolic modelling. Originally tested and validated with microorganisms, metabolic systems

biology is now becoming widespread for human tissues and biomedical applications [17]. We provide a comprehensive review of the metabolic models developed so far for human cells, discussing their strengths and limitations. We show that the main focus of the metabolic modelling community has shifted from increasing the size of the models—a “size-focused modelathon approach”, often at the expense of reliability of predictions—to team-based curation approaches where gap filling algorithms and extensive manual curation improve the reliability of the predictions without necessarily increasing (and in fact often decreasing) the number of components in the model. We also outline key success stories, current challenges, and future directions for the human metabolic modelling field, including the integration with machine learning methods and precision medicine.

## 2. Metabolic Systems Biology

Understanding the role of individual components in a biological system is an important step to elucidate or predict its behavior. However, a major theme in systems biology is investigating these systems with an integrated approach. In fact, studying the interactions between different components also gives insights into the functioning and behavior of the components taken independently [18].

Metabolism is the biggest biological network fully described. It is a surrogate of the phenotype, being able to

elucidate phenotypic changes caused by single or combined omics or physiological factors. It also has the advantage of being the only system that can be entirely modelled at genome-scale. As a result, it plays a central role in elucidating the genotype-phenotype relation. Systems-based approaches have been successfully applied over the last decade to investigate metabolic networks, composed of a set of chemical reactions and a pool of metabolites. The first examples of metabolic systems biology appeared in 1999 and were focused on modelling, connecting, and simulating several cellular processes [19]. Whole-cell modelling, named the “grand challenge of the 21st century”, is still an active area of research [20–24]. In metabolic systems biology, the value of modelling cell metabolism is not merely explanatory of a biological process, but also predictive. A model can be used to suggest hypotheses that can be tested or to pinpoint unexpected behaviors that can be further investigated *in vitro*.

There are different methods to model a metabolic system: steady-state analysis (e.g., FBA) involves a set of linear equations, while kinetic simulations involve ordinary differential equations (ODEs). Each variable represents the variation of a metabolite concentration, in a dynamic or steady state, where the concentration depends on the rates of the reactions that produce and consume that metabolite [25]. Kinetic models do not assume steady state and therefore are able to model highly dynamic mechanisms, including allosteric and posttranslational regulation, metabolite concentrations, and thermodynamics. Such ODE-based systems contain a large number of equations (differential or algebraic) and require unique kinetic parameter values. They are highly effective at predicting the behavior of small systems where sufficient experimental data can be collected for model calibration and parameter estimation [26]. Furthermore, unlike standard FBA-based methods, reaction kinetics can be accounted for, and metabolite concentrations can be modelled explicitly, and therefore intracellular metabolomics data can be integrated directly [27].

For large systems, however, the use of kinetic modelling remains challenging. The increasing demand for systems-level genome-scale analyses has recently led to the widespread use of constraint-based steady-state models and their unsteady-state extensions. This has also been facilitated by the increasing availability of multiomics data and phenotypic information, used to constrain the model dynamically and therefore often compensating for the lack of regulatory and kinetic modelling.

Following a number of successful attempts at building multiscale kinetic and constraint-based models [28, 29], achieving the right combination (and therefore trade-off) of kinetic modelling and steady-state assumptions is likely the next step for the metabolic modelling community. Techniques like unsteady-state FBA (uFBA) [4] have been recently developed to relax the steady-state assumption in genome-scale models, with the goal of modelling dynamic cellular states derived from changes in the concentration of internal metabolites. An approach to integrate reaction kinetics with steady-state metabolic networks has also been proposed, where the genome-scale information of shared metabolites among different reactions is used to inform the interactions

between the reactions and predict metabolite concentrations in a network kinetics approach [30]. More generally, due to the larger range of predictions that kinetic models can perform compared to steady-state approaches, expanding them towards the genome scale or using information derived from genome-scale model simulations is a promising direction that will increase their spatial and temporal resolution.

**2.1. Genome-Scale Metabolic Models.** Genome-scale metabolic models contain all known metabolic reactions in an organism and can therefore serve as functional databases of cell-specific metabolism. A genome-scale model is built using the following process. First, a draft reconstruction is generated starting from the genome and including all the genome-encoded metabolic reactions. The draft reconstruction also includes annotated enzyme, reaction, and pathway data from databases like KEGG [31], BioCyc [32], and BRENDA [33]. Details on which genes control each reaction are also included. Then, a sequence of manual curation steps improves the draft reconstruction, by gathering evidence to prove or disprove the presence of a reaction in the network of the organism.

The construction of genome-scale model represents the starting point for flux balance analysis (FBA, see following subsections). Finally, the model is run and validated by comparing its predictions with existing experimental results, and new *in silico* experiments are performed to further improve and validate the model. For more details on how to build a metabolic model from the DNA sequence of an organism, the reader is referred to full protocols [34–36].

Most genome-scale models are annotated with curated gene-protein-reaction associations (GPR rules), linking genes with enzymes (Figure 2). Such annotations pave the way for overlaying multiomic data on the models, using them as omic-scaffolds (see Section 2.3). Since omic data can be quantified numerically and in a condition-, tissue-, and patient-specific way (e.g., transcriptomic profiles, protein levels, and metabolite concentrations), models with GPR rules can serve as a baseline for generating personalized metabolic models. For instance, personalized predictions using such models can lead to precise phenotypic characterization of patients (see Section 4).

**2.2. Constraint-Based Modelling and Flux Balance Analysis.** Flux balance analysis (FBA) is the most widely used constraint-based technique to predict flux distributions and network capabilities in genome-scale models [37]. FBA has proved useful thanks to its ability to handle large networks and predict genome-scale flux distributions. It requires information about biochemical reactions and stoichiometric coefficients but does not involve kinetic parameters. This makes it well suited to metabolic engineering studies that identify and characterize optimal perturbations such as different substrates or genetic interventions (e.g., knockouts) leading to obligatory coupling between the growth rate and the overproduction of the desired metabolite [38–42]. In general, FBA is a powerful tool for predictions of cell behavior under different metabolic conditions.

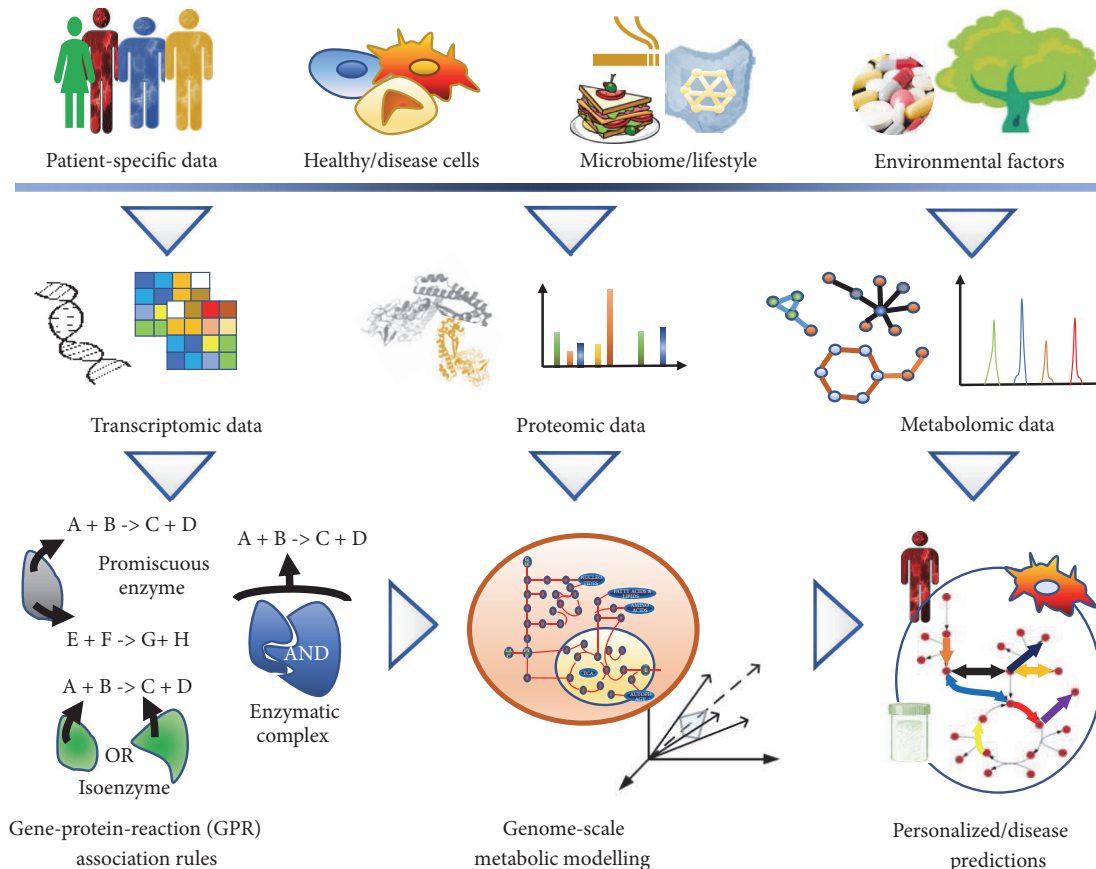


FIGURE 2: The integration of different types of omics data can be used to infer tissue- and condition-specific intracellular metabolic flux distributions. Intracellular metabolic reactions provide the cell with basic biochemical building blocks, as well as energy and a thermodynamically favorable environment to sustain its life. Patient-specific data, molecular information, lifestyle, and environmental factors affect different omic levels. As a consequence, transcriptomic, proteomic, and metabolomic data need to be integrated to determine gene-protein association rules and to build genome-scale models used for personalized predictions. Given the large effect of environmental factors on omics level, determination of system-level changes in intracellular metabolic fluxes is important for understanding the fundamental mechanisms of metabolic responses to perturbations. Indeed, environmental factors affect omics data on different levels, from epigenomics to the cell phenotype. Omic-augmented genome-scale metabolic reconstructions have proved successful due to the ability to integrate omic measurements at genome scale and to give mechanistic insights into the genotype-phenotype relationship.

FBA is a linear programming technique that models the steady-state condition in a chemical reaction network [43]. The FBA representation of a genome-scale model is built based on a stoichiometric matrix  $S$ , containing the stoichiometric coefficients of each metabolite (on the rows) in each reaction of the network (on the columns). A stoichiometric coefficient of a metabolite in a given reaction is positive if the metabolite is produced by the reaction and negative if consumed. An underdetermined linear system of equations  $Sv = 0$  is defined from the stoichiometric matrix, where the unknowns are represented by the vector  $v$  of reaction flux rates. Additional constraints are included as lower and upper bounds of the fluxes in  $v$  ( $v^{\min} \leq v \leq v^{\max}$ ). These account for the growth or physiological condition, and can be used to incorporate omics data. One or more cellular objectives (usually growth- or energy-related, e.g., biomass or ATP), or a linear combination thereof, are finally selected to be maximized or minimized under the above-mentioned constraints, therefore solving a linear program.

FBA is therefore based on two main assumptions:

- (i) Homeostatic assumption: the organism has reached a steady state where the metabolite concentrations are constant and a set of nutrients are being constantly converted to generate biomass.
- (ii) (Multilevel) optimality: in each state, the organism tends to maximize one or multiple objectives, usually related to growth, biotechnologically relevant compound production (e.g., acetate exchange) and important energy-carrying molecules (e.g., ATP).

FBA is widely used in systems biology to quantify the entire metabolic steady state of a cell and calculate its flux distribution. All known metabolic reactions in a given cell are considered, and they are mathematically described in a way that allows simulation of various states and configurations of the chemical reaction network. Intuitively, the steady-state constraints used in FBA can be thought of as Kirchoff's

laws applied to any node representing a metabolite in the network: the flux through each metabolite in the network must be constant, namely the input flux must equal the output flux. The combination of steady-state constraints and capacity constraints on reaction fluxes is a system of linear homogeneous equations and inequalities; thus, its solution space is a convex polyhedral cone representing the feasible flux distributions.

The assumptions of FBA and the reduction of the problem to a linear program can cause some limitations. First, incorporating or predicting metabolite concentrations is challenging, and requires a dynamic FBA [44], relaxing the steady state [4], flux-sum methods [45], or thermodynamic approaches [46]. Second, the reliability of the flux distributions is highly dependent on the objective function chosen (see Section 2.4), on the quality of the reconstruction and on the method used to obtain the solution. Regularized FBA methods help alleviate this issue, as discussed later in this section and in Section 5. Finally, FBA lacks the ability to directly model regulatory effects or posttranscriptional regulation of expression levels. Likewise, changes occurring over quick transients (e.g., perturbations to the cell microenvironment) cannot be modelled dynamically, but can be approximated through step-wise before/after simulations [47]. Although the steady-state rule has been challenged and probabilistic approaches have been proposed to relax the steady-state equality [48], this assumption enables the use of linear systems and linear programming, lowering the computational requirements and enabling fast simulation of genome-scale models in a variety of growth or physiological conditions.

When solving linear programs for FBA, different solvers can give different solutions due to numerical implementation differences and to the existence of multiple alternate optimal solutions [49–51]. While the value of the objective function is the same in all the optimal solutions, the other flux rates can vary. Having a unique solution is therefore important when the full flux distribution is used for further analysis or as a feature of predictive algorithms. To avoid the degenerate solutions provided by standard FBA approaches, parsimonious FBA (pFBA) has been proposed with the aim of minimizing the overall flux carried by the metabolic network after maximization of the main objective [52]. However, in some cases, pFBA has produced less plausible results in central carbon metabolism and in the glycolytic pathway compared to standard FBA methods [53]. Furthermore, pFBA makes use of the L1-norm (minimization of the sum of absolute flux values) which does not guarantee a unique solution as it is not strictly convex [54]. For solving such minimization problems with a guarantee of a unique solution, the L2-norm should be used instead (minimization of the sum of squared flux values). Additional approaches to alleviate the problem of alternate optimal solutions include geometric FBA [55] and sampling [56].

Several tools are available for metabolic model reconstruction, constraint-based modelling, FBA and related analyses. These include COBRA [57], its Python version COBRAPy [58], RAVEN [59], PathwayTools [60], and FAME [61]. For a complete list, the reader is referred to the related reviews and comparison papers [62, 63].

**2.3. Multiomic Flux Balance Analysis.** Although some biological outcomes can be elucidated through single-omic analysis (e.g., protein-protein interaction or gene networks), the vast majority of phenotypic observables are a result of more comprehensive interconnections among multiple omics [64]. Regulatory mechanisms also take place at different omic levels (including transcription, translation and metabolic reactions). The complex interplay between these levels is responsible for cell behavior. Multiomic analyses have been proposed to integrate single-omic networks, based on coexpression or interaction networks [65]. However, such methods often omit mechanistic links between omic layers, derived from previous molecular knowledge.

The idea of multiomic FBA is that by mapping omic data onto *in silico* models of metabolism, it is possible to obtain a metabolism-enriched view of any given omic profile. In this context, as well as permitting the simulation of large, usually genome-scale, systems in a few seconds of CPU time, FBA also has the advantage of facilitating the introduction of additional omic layers of experimental data that can be overlaid onto the model, using GPR rules or metabolite annotations to place constraints on the flux bounds. Omic-informed metabolic models consider the mechanistic relation between omics, therefore correctly representing the prior information available on biochemical networks [66]. Multiomic FBA allows assigning the importance of a gene or enzyme through its predicted function, therefore avoiding approximations or statistical methods merely based on the value of its expression level.

Following this approach, FBA and its multiomic modifications have been used to predict the metabolic response to a given condition in light of the multiple cellular objectives that a cell is required to meet [67, 68]. Recent studies have been also aimed at improving the predicting capability of a metabolic model and elucidating the genotype-phenotype relationship through computational analyses across multiple omic levels. Predictions of flux distributions after multiomic data integration, combined with experimental methods, have been successfully used to formulate novel biological hypotheses. These techniques reduce the problem of determining the flux distribution through all reactions in the system, under a given growth or physiological condition, to a tractable linear program, under the assumptions of steady state and optimality. Due to their scalability and precision, these methods have been used widely, e.g., to predict bacterial growth phenotypes in specific environmental conditions [69, 70], to identify novel therapeutic targets against infections [71, 72], to characterize cancer metabolism of different cell lines [73–76], and to generate cancer-vs-normal tissue-specific models for 17 tissues [77]. Further examples of how this approach has been used to characterize disease metabolism are provided in Section 4. For a comprehensive review on multiomic integration techniques in FBA models, the reader is referred to the recent reviews on omic-informed metabolic modelling [78–81].

Ongoing modelling efforts are aimed at incorporating further biological knowledge in the models. To account for enzyme promiscuity when overexpressing or underexpressing a gene, GPR rules have been proposed as additional

rows of the stoichiometric matrix, as an enzyme-by-reactions submatrix [53]. For instance, if multiple reactions are catalyzed by the same promiscuous enzyme, the change in flux as a result of underexpression/overexpression of the enzyme is distributed among the reactions (also considering the other enzymes participating in the reactions), rather than triggering *a priori* an equivalent flux increase/decrease in all the reactions. Codon usage can also be integrated through the GPR rules [71]. Splice isoform annotations, initially lost and then simply ignored in human reconstructions, have also received recent attention [82, 83].

Soft constraints can be used to integrate omics data in a reaction-specific fashion, rather than with a single mathematical rule modifying the constraints for all the reactions [84]. Further constraints have been also added to integrate transcriptional regulatory models and metabolic models. The idea is that a constraint models the correlation between the expression level of a target gene and that of its regulating genes [6]. To further analyze the reaction flux distributions predicted by multiomic FBA, sensitivity analysis can be carried out on genes, reactions or pathways. Definitions of sensitivity scores have been based on effects of gene knockouts [85], real-valued gene perturbations [69], or the role of each metabolite in reducing thermodynamic uncertainty in the model [86].

**2.4. Choosing an Objective Function and Accounting for Multiple Cellular Goals.** The selection of an appropriate objective function is still a challenge in metabolic modelling. A common assumption in systems biology is that cells tend to optimize their metabolic network in order to maximize the growth rate (biomass). This is however still a matter of debate [87], both in terms of its composition [88] and because in many cases the best objective for a cell is not growth-related [89]. While models of microorganisms can assume that the cell aims at maximizing growth rate, human cells might not necessarily aim for maximum biomass (although this is widely accepted as an objective for cancer cells). Furthermore, it is now evident that the cellular goal can change between different cells in a tissue, between tissues, and also over time for the same cell. For tissue-specific studies, starting from a generic human metabolic reconstruction (e.g. Recon) and using the general-purpose biomass composition of generic human models can prove unreliable. Recent studies are therefore moving towards a cell-specific estimation of the biomass compositions [13, 90]. Algorithms for generating and standardizing the biomass reaction [91, 92], and for automatically generating an objective function have also been proposed [93].

Especially in the context of human metabolism, the question ‘*what does a particular cell do?*’ has often more than one correct answer. There is increasing evidence that cells have to cope with multiple, usually competing, objectives to optimize simultaneously [94]. A single FBA objective function is not able to capture all of them. It is also likely that metabolism is not fully optimized for any particular objective [95], and evolution has shaped cells in order to reach an optimal trade-off between all objectives [96].

Optimization processes have therefore been proposed that take into account multiple objectives (e.g., protein or energy production, detoxification, proliferation) with the advantage of ensuring metabolic flexibility for possible reorganizations performed during adaptations to changes in the environmental conditions [97]. These methods include a probabilistic approach [98], lexicographic ordering [99], and the reduction to a single objective through the definition of weights for the objectives [100, 101]. Approaches based on these ideas are fast and can generate a Pareto front; however, they miss suboptimal solutions, solutions in nonconvex regions [102, 103], or they give preference to one objective [104], therefore addressing a multilevel problem rather than a strictly multiobjective problem. Although slower, methods based on evolutionary algorithms are able to explore nonconvex trade-offs without requiring the combination of the objectives into a single objective function [105–107]. Exploring a set of trade-off solutions between competing objectives, rather than a single biomass-maximizing solution, also accounts for suboptimal solutions [108].

### 3. 15 Years of Human Metabolic Modelling

The study of human metabolism is becoming increasingly important for biomedical applications as an approach for understanding many diseases and aspects of health [109, 110]. A systems-level understanding of metabolic behavior is enabled by the availability of high-quality genome-scale reconstructions integrating extensive metabolic information from various resources. The first genome-scale metabolic reconstruction efforts focused on bacteria (*Haemophilus influenzae* [111], followed by *Escherichia coli* [112]), due to their simplicity, the available genome sequences, and the well-known mechanisms of substrate utilization. Reconstructions of human metabolism, which required a much larger number of pathways and a larger pool of essential nutrients, were attempted at a later stage, after the publication of the draft and complete human genome sequences in 2001 and 2004 [113, 114].

HumanCyc [115] (in 2004) and Reactome knowledgebase [116, 117] (in 2005) were the first successful attempts to build a curated collection of biochemical reactions in human cells. The first models of human metabolism published in 2007—Recon 1 [118] and EHMN (Edinburgh human metabolic network reconstruction, and its subsequent compartmentalized version in 2010) [119, 120]—achieved a better coverage of the human metabolic network, especially in compartments other than the cytosol. Recon 1 contains approximately 1.4 times more unique metabolites and 2.2 times more reactions than HumanCyc.

In 2013, a large improvement of the number of metabolic processes covered in the reconstruction was achieved with Recon 2 [121]. This is a consensus model that includes all reactions from EHMN, Recon 1, HepatoNet1 [122], and a module for acylcarnitine and fatty-acid oxidation [123]. Compared to Recon 1, Recon 2 represents a major improvement with approximately 1.7 times more unique metabolites and 2 times more reactions. Recon 2 was then refined in 2015 with

updated gene-reaction associations. Simultaneously, a larger reconstruction named HMR (Human Metabolic Reaction database) [73] was built independently, based on HepatoNet1, Recon1, EHMN, Reactome, HumanCyc, KEGG [31], and the Human Metabolic Atlas [124]. HMR was then extended in 2014 to include lipid metabolism, therefore generating HMR2 [125].

Although containing a large set of reactions and being a consensus model, the predictions from Recon 2 were incorrect in some cases, mainly due to a number of incorrectly balanced reactions. At this point, the main focus of the community shifted to curation of existing models. Between 2014 and 2015, a series of minor updates and corrections were published for Recon 2 [126–129]. In 2016, Swainston et al. performed an extensive curation effort on Recon 2 and its updates, resulting in Recon 2.2 [130]. For the first time, a significant improvement of an existing model contained fewer genes than the original. In fact, due to the removal of duplicates and pseudogenes, and to the new unique HGNC ID convention, Recon 2.2 contains 112 enzyme-encoding genes less than Recon 2, and for the first time with unified identifiers (grouping splice isoforms into a single gene). Recon 2.2 is a major improvement for the energy-related metabolism, with a new compartment simulating the mitochondrial intramembrane space. As a result, the ATP and biomass flux predictions were greatly improved, achieving for the first time accurate predictions of maximum ATP yield under 14 carbon sources, both in aerobic and anaerobic conditions, and under 20 additional amino acid carbon sources in aerobic conditions [130].

In 2017, a new human metabolic model named iHsa [131] was obtained as an expansion and manual curation of HMR2. iHsa is a result of a reconciliation with the rat metabolic model iRno, also presented in the same paper. Both models were generated in parallel; rat-specific reactions, thermodynamically infeasible reaction loops, and other incorrect reactions present in HMR2 and Recon 2 were removed from iHsa, while new reactions were added from KEGG and MetaCyc/BioCyc [132]. For the first time, rat-specific reactions were removed from a human metabolic reconstruction. In fact, both models were reconciled and manually curated focusing on species-specific metabolic differences. The clear definition of differences between the two metabolic networks, coupled with the presence of unique reactions in both models, opens opportunities for human/rat cross-species comparison.

In 2018, Recon 3D [133] was developed from Recon 2 by incorporating HMR2 and a number of additional reaction sets, including reactions modelling host-microbe interaction, reactions for simulating drug effects on human metabolism, reactions for absorption of dietary compounds, reactions of lipid metabolism and reactions from metabolomics datasets. 3D protein structures, pharmacogenomics data and atom-atom mappings were also included in the model. After extensive manual curation steps and refinement of GPR rules, the model was tested for consistency when replicating 431 essential functions of the human body. To date, Recon 3D is the largest metabolic reconstruction available for human metabolism, containing 1.1 times more unique metabolites and 1.4 times more reactions than Recon 2.

Figure 3 summarizes the progress made by the metabolic modelling community over the last 15 years. The number of genes, metabolites and reactions is reported in the figure. To better quantify the connectedness of a model, we computed the number of genes whose knockout yields measurable effects in the model as measured by the predicted biomass flux. To account for the tolerance of the linear solver, we defined a gene *essential* if its knockout produces biomass  $< 10^{-10} \text{ h}^{-1}$ ; we defined a gene *nonnegligible* if its knockout causes a biomass variation  $> 10^{-10} \text{ h}^{-1}$ . Compared to Recon 2, the increase in the number of genes whose variation has a nonnegligible effect on the growth rate shows that efforts have been successfully made towards curation, gap filling, and consistency checks, as detailed in Section 3.1.

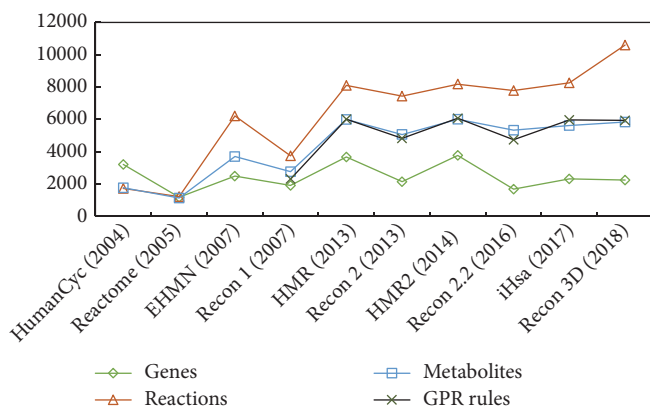
### 3.1. From Size-Focused “Modelathons” to Manual Curation.

When the first Recon 1 [118] and EHMN [119] human metabolic reconstructions were published in 2007, Recon 1 was preferred in many studies because it contained more reactions. In fact, this was only due to the compartmentalization of Recon 1, with metabolites repeated in different compartments and transport reactions between compartments. The number of unique reactions was indeed greater in EHMN (1028 more reactions and 1202 more metabolites), and a subsequent compartmentalized version of EHMN was generated [120].

These frequent comparisons among different models indicate that the main focus was initially to include as many reactions as possible—a size-focused “modelathon” approach—often at the expenses of curation and accuracy. Conversely, the approach of including all the available biological components in a model is now always coupled with extensive team-based curation, consistency checks, uniformity of annotations and gap filling. It is not uncommon to achieve a better model by removing reactions from an existing one. This encouraged the community to focus on the accuracy of the models and solvers used, rather than only on the model size.

Manual curation is therefore considered one of the main steps in generating a metabolic model. Tools like MetaNetX [134] simplify curation efforts and suggest identifier mappings to unify reactions and metabolites identifiers. Such curation efforts are likely to become even more important in the near future given the recent advances of automated reconstruction tools, now able to generate a full working draft of a model, e.g., MicrobesFlux [135], Pathway Tools [136], PathwayBooster [137], CoReCo [138], and Merlin [139]. Given the importance of obtaining reconciled and high-quality models, MetaNetX also provides a database of models generated after reconciliation of metabolites and biochemical reactions [134]. More recently, CarveMe [140] and RAVEN 2.0 [59] have shown high potential in automating manual steps, achieving for the first time an accuracy directly comparable with manually curated models when predicting experimental phenotypes. Tools for checking inconsistencies and for visually inspecting the model are also available [141].

Over the last few years, the curation efforts have yielded a reduction in the number of unbalanced reactions, as well as in



Model	Genes	Metabolites	Reactions	GPR rules
HumanCyc (2004)	3209	1761	1716	-
Reactome (2005)	1180	1131	1216	-
EHMN (2007)	2492	3695	6216	-
Recon 1 (2007)	1905	2766	3744	2307
HMR (2013)	3668	6000	8100	6015
Recon 2 (2013)	2140	5063	7440	4821
HMR2 (2014)	3765	6006	8181	6071
Recon 2.2 (2016)	1675	5324	7785	4742
iHsa (2017)	2315	5620	8264	5961
Recon 3D (2018)	2248	5835	10600	5938

Model	Essential	Non-negligible
Recon 2 (2013)	0.09%	0.14%
HMR2 (2014)	0.90%	1.99%
Recon 2.2 (2016)	0.90%	3.28%
iHsa (2017)	3.33%	3.33%
Recon 3D (2018)	0.04%	1.78%

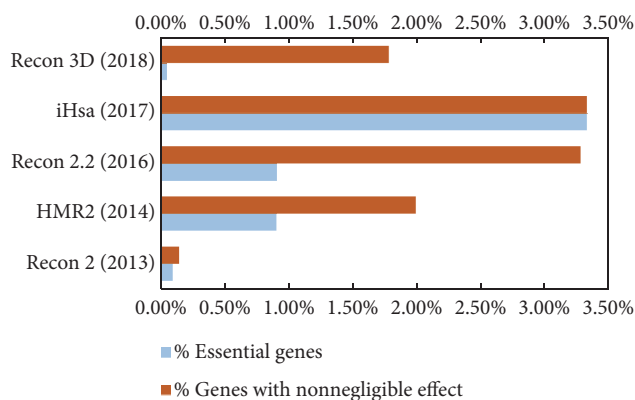


FIGURE 3: (Top, left) Number of genes, metabolites, reactions, and GPR rules in human metabolic models. The number of genes in the table refers to the number of transcripts in the model. (Number of enzyme-encoding genes: Recon 1: 1490, Recon 2: 1789, and Recon 2.2: 1675.) Recon 2 was considered in its latest Recon 2.04 iteration. The GPR rules in HMR have been extracted as all the nonempty “listOfModifiers” fields via a Matlab custom script. The boundary metabolites were excluded from the comparison, therefore metabolite counts represent the size of the stoichiometric matrix on which the model is built. (Bottom, right) Gene essentiality in the five most recent metabolic models. A gene is deemed essential if, when knocked out, it induces a biomass value lower than  $10^{-10} \text{ h}^{-1}$ . A gene is nonnegligible if, when knocked out, it affects the biomass by more than  $10^{-10} \text{ h}^{-1}$ . Although the improvement in the number of genes and metabolites has been limited in the past five years, the curation and gap filling efforts have produced more reliable models. As a result, compared to Recon 2, there has been a sharp increase in the percentage of genes whose knockout yields measurable effects on the predicted biomass.

the number of blocked reactions and dead-end metabolites. For instance, Recon 2.2 has no unbalanced reactions apart from the biomass objective functions, and Recon 3 has now less than 15 blocked reactions and dead-ends. Furthermore, as generating context-specific models using omics relies on GPR rules, the accuracy of these models strictly depends on the accuracy of the GPR in the baseline model. As shown in Figure 3, GPR coverage is increasing. Finally, the agreement between theoretical and model-predicted ATP yields has been improved in the most recent models. We calculated the root mean squared error (RMSE) between the theoretical and model-predicted maximum ATP yields per unit of carbon source, which has been reduced from  $\infty$  (Recon 2 and earlier models) to 5.17 (in Recon 3) and 1.67 (in Recon 2.2). The table of ATP yields and the RMSE calculations are provided as Supplementary Material.

#### 4. Tissue- and Patient-Specific Insights into Human Disease Metabolism

Starting from human generic metabolic reconstructions, a number of successful efforts have been published on reconstructing tissue-specific models. These include brain [142],

adipocytes [143], breast cancer [144], heart [145], kidney [146], myocytes [147], and hepatocytes [125, 148]. A further set of 32 tissue- and organ-specific models has been generated by mapping protein expression onto the generic human model HMR2 [125].

Tissue-specific models are usually built as a reduction of a generic human model. The reactions that are removed in the construction of the tissue-specific model are found to be not active in that tissue. Reactions are removed according to transcriptomic or proteomic data collected in that tissue, and therefore the tissue-specific model contains fewer reactions than the generic one. Unsurprisingly, in most cases they outperform the generic counterpart where they stemmed from. At this point, a new tissue-specific biomass equation needs to be established, or the model needs to be set so that the cell can find alternative pathways to sustain its life. Although many methods for generating tissue-specific models set the goal of creating the minimal metabolic model that satisfies viability or a set of metabolic tasks, the use of nonminimalistic methods for generating tissue-specific models should be preferred. For instance, considering essentiality before removing reactions from a generic model yields a nonminimal tissue-specific model, but improves model functionality and agreement with experimental data [149].



Since the generic models are a superset of the tissue-specific models, predictions of generic models may still be correct in tissue-specific problems, but less accurate.

Context-specific human metabolic models have shed light on disease onset and progression in a range of recent case studies. In this context, the renewed interest in disease metabolism is due to the fact that a holistic genome-scale view, rather than a single-gene approach, is necessary to fully characterize most diseases. For instance, in cancer, the differences in the metabolic pathways between cancer cells and their parent tissue have been characterized using omics data and a human metabolic model [75, 150, 151]. As a result, tissue- and cell-specific metabolic models have been successfully used to identify—and successively validate—specific drug targets that inhibit cancer proliferation but do not affect normal cell proliferation [74, 152]. With a similar approach, submodels built from human genome-scale models have been used to generate several testable hypotheses, e.g., to compare wild-type and Fh1-deficient kidney mouse cells, and to predict further gene knockouts that affect growth in the Fh1-deficient cells but do not affect the wild-type cells, therefore suggesting targets for treating hereditary leiomyomatosis and renal-cell cancer [153].

For specific diseases, tissue-specific models have been used successfully to identify biomarkers and therapeutic targets [125]. Using RNA-Seq expression levels in combination with genome-scale models has enabled the reconstruction of cancer cell line-specific models, enabling the discovery of metabolites supporting proliferation, and antimetabolites leading to cell death [154]. (An antimetabolite is a compound that simultaneously inhibits those enzymes involved in metabolizing the associated endogenous metabolite. Antimetabolites can affect multiple enzymes at the same time and can reduce proliferation, and are therefore used as anticancer drugs.) After integration with extracellular metabolomic data and transcriptomic profiles, metabolic modelling has reliably characterized intracellular metabolism of lymphoblastic leukemia cell lines [155]. A combination of different omics data and flux splits have been used to generate cancer-specific models of the NCI60 panel, achieving correlation close to 1 in predicting the remaining flux rates [13].

Following the same direction, personalized models (e.g., using patients' omic data to constrain the generic reconstruction) hold promise to become key for precision medicine (Figure 2). The largest study to date with patient-specific models is the Human Pathology Atlas [77]. The authors built personalized genome-scale models of cancer in each patient across 17 tissues. This allowed investigating the metabolic differences between different cancers, as well as patient-specific biological functions. In another recent study [156], 86 patients with nonalcoholic fatty liver disease were recruited, and their personalized hepatocyte genome-scale metabolic models were built using patient-specific experimental data on lipoprotein fluxes. A new molecular mechanism of the disease was elucidated as a result of personalized metabolic modelling.

Genotyping coupled with patient-specific metabolic modelling offers a new opportunity for personalized medicine. In fact, new biomarkers can be predicted in a

patient-specific fashion, and personalized therapies can therefore be designed and subsequently assessed in terms of their metabolic mechanistic effects [157]. As shown in a recent case of arginase deficiency (a urea cycle disorder), different individuals can respond differently to the same disease and treatment, and this can often be flagged observing their individual metabolic response [158].

Due to the importance of the gut microbiome composition in human health, a set of modelling approaches has been proposed recently [159]. Using host-microbiome modelling, e.g., combining metabolic community modelling with human metabolic models, the interplay among the gut bacteria and their interaction with the surrounding human cells can be investigated [160, 161]. Personalized models of gut microbiome have been built from patient-specific metagenomic data in order to elucidate individual-specific bile acid production in microbiomes of healthy individuals and patients with inflammatory bowel disease [162]. Although several modelling challenges remain to be solved [163], an approach based on metabolic modelling is likely to shed light on the role of human gut microbial communities in human health, therefore suggesting potential dietary changes or personalized intervention on gut composition [164, 165].

## 5. Discussion and Perspective

Research in computational biology has led to detailed models for a better understanding of the individual biological components, but arguably to a less clear picture of the interactions among the components that result in a given phenotype [166]. Genome-scale systems biology studies can effectively address this issue. For instance, in biomedical applications, this holistic view is necessary to characterize a patient's disease phenotype and deliver precision medicine [17].

Since metabolic homeostasis and observable phenotype are strictly linked, metabolism is nowadays considered diagnostic of the phenotype, and therefore arguably the best indicator of the functional state of a cell. Metabolism can also be used to prioritize genes and assess their function and the role of gene perturbations (including knockouts). Without such integrated analysis, a gene may e.g., incorrectly be regarded as important only due to its highly variable expression value.

Metabolic models are increasingly being used to construct multiscale, multicellular or multitissue models. In 2012, a research effort by Karr et al. [21] provided the first whole-cell computational model of the life cycle of a small pathogenic bacterium, *Mycoplasma genitalium*. The model includes metabolism, replication of the genome, and cell division. Several metabolic models can be combined in frameworks to investigate the metabolic exchanges between individual cells and the emerging community behavior [167], with applications ranging from microbial communities to host-pathogen interactions and cancer proliferation [168–170].

The availability of tissue-specific models has also led to the generation of multitissue metabolic models [171].

Further steps in this direction will soon allow studying whole-body metabolism models. The first study towards this goal used dynamic parsimonious FBA to combine steady-state FBA with differential equations for the concentration of metabolites in each organ [172]. As a result, a whole-body and inherently multiscale model was generated with 14 organs plus the human serum. More recently, the first whole-body metabolic models were generated named Harvey and Harvetta, where the metabolic network for every organ/tissue was built simultaneously from Recon3D [173]. Manufacturing organ-on-a-chip and ultimately body-on-a-chip devices with metabolic models (see e.g., [174]) seems the next step in this direction.

Given the recently renewed interest of the scientific community in understanding disease metabolism, it is highly likely that cancer metabolism (and metabolism in other diseases) will become the main research topic in drug development. This will complement rather than replace standard transcriptomic-only studies, and will provide a proxy for the phenotypic and observable outcome. The idea is that to improve phenotypic predictions, the signatures need to be taken from sources of data that are closer to the phenotype. The integration of omics data into metabolic models has enabled the prediction and successively the validation of biomarkers and therapeutic targets. In drug design, although genomics, transcriptomics and proteomics are often deemed sufficient, metabolomics can elucidate mechanisms that are not visible from genes and protein activity. More importantly, it can address cases where the genes responsible for a disease are well known but not druggable, but their corresponding downstream reactions are [175].

Genomics analysis is usually able to identify known diseases through gene mutations. However, it may not be able to flag variants of known diseases, or identify novel diseases. This can happen, for instance, if (i) a gene mutation is not flagged as important; (ii) the gene is not screened at all; (iii) the disease is not a direct effect of a gene perturbation; (iv) individuals respond differently to the same mutation [158]. Genome-scale models can elucidate mechanistic modes of drug action, side effects (both off-target drug binding and downstream transcriptional effects), and potential toxicity of drugs by linking omics data to the phenotype through a condition-specific model [8, 176]. In a patient-specific framework, such biological data and markers can be predicted in a personalized fashion, paving the way for *in silico* clinical trials [177].

The efforts of the metabolic modelling community are far from being complete. In this regard, we envisage a joint effort from metabolic modelling and genome-wide association studies (GWAS) communities to identify gene-metabolite interactions that are currently not included in metabolic models using state-of-the-art association mappings [178]. Computational tools combining both approaches would go towards genome-scale detection of errors and missing enzymatic reactions, remarkably improving the predictive ability of metabolic models. For instance, this could be achieved by minimizing the error between predictions of flux coupling and experimental coexpression data integrated with GWAS [179, 180].

Likewise, the methods for integration of omics data in genome-scale models still show room for improvement. For instance, in yeast and *Escherichia coli*, single-omic integration in FBA has been reported to give similar accuracy to pFBA without omic integration [78]. Reassuringly, reaction-specific rules to constrain flux rates can halve the normalized error of pFBA [84]. In multicellular organisms, a better correlation between gene expression and metabolic fluxes can be expected [78]. In mammalian cells, the main contributors to the overall protein expression level are mRNA levels [181, 182], while in most normal and cancer cell lines, mRNA and protein levels were found to correlate positively [183, 184]. While different methods (and their parameters) can yield different context-specific models, the accuracy in predicting essential genes is almost always higher in context-specific models than in the generic human models [81]. Therefore, metabolic models can be regarded as useful tools to mechanistically link transcriptomic data with flux rates. Using multiple omics data and thermodynamic constraints simultaneously [185, 186], or a combination of regularized FBA methods and omics data, can improve the reliability of the predictions [187].

Compared to the well-characterized microbial metabolic reconstructions, gaps still present in our knowledge of human metabolism—including characterization of enzymes, the definition of cell-specific metabolic functions and tissue-specific growth mechanisms—make it more difficult to test (and compare) metabolic models and omic integration methods. Human metabolic models and methods have been tested and cross-compared for gene essentiality through CRISPR-Cas9-mediated loss-of-function screens, and for their ability to predict growth rate and recapitulate known metabolic functions [81]. A dataset including both omics data and measured metabolic information is the NCI-60 panel, with metabolite uptake and secretion rates published in 2012 [188]. While *E. coli* and yeast fluxomic data is publicly available from several experiments [189–191], to the best of our knowledge NCI-60 is the only publicly available human-cell dataset containing expression levels, metabolic flux rates and proliferation rates, and therefore suitable for validating FBA methods. Due to the lack of fluxomic datasets, the difficulty in choosing a reliable objective function, and the larger size compared to bacterial models, further experimental validation is needed for human metabolic models, especially when using the full flux distribution to inform decisions or subsequent algorithmic steps.

Finally, despite many recent advances, gathering insights from the data generated through omic-informed models remains a bottleneck in systems biology. We therefore envisage that genome-scale metabolic models will be increasingly investigated with machine/deep learning algorithms in a patient-specific fashion [13, 192]. In disease modelling, the usefulness of general biomarker discovery is debatable when information on patients is not taken into account, because most biological components would be flagged as perturbed in a general “disease versus normal” analysis. Conversely, if the appropriate person-specific data is integrated with a model, biomarkers can become the central part of precision medicine, where data acquired on patients drive predictions, analysis and therapeutics [17]. In a modelling

analogy, machine learning algorithms alone cannot provide mechanistic information in the biological processes they simulate or mimic. However, if machine learning is coupled with multiomic genome-scale modelling, the combination of experimentally and model-generated omic data can predict—and explain mechanistically—personalized therapy predictions by including key biological information in the learning process.

## Conflicts of Interest

The author declares that they have no conflicts of interest.

## Supplementary Materials

Agreement between theoretical and model-predicted ATP yields for the most recent human metabolic models. We calculate the root mean squared error (RMSE) between the theoretical and model-predicted maximum ATP yields (per unit of carbon source) for the most recent human metabolic models. The RMSE has improved from  $\infty$  (Recon 2 and earlier models) to 5.17 (in Recon 3) and 1.67 (in Recon 2.2). (*Supplementary Materials*)

## References

- [1] G. Piedrafito, M. Keller, and M. Ralser, “The impact of non-enzymatic reactions and enzyme promiscuity on cellular metabolism during (oxidative) stress conditions,” *Biomolecules*, vol. 5, no. 3, pp. 2101–2122, 2015.
- [2] N. N. Pavlova and C. B. Thompson, “The emerging hallmarks of cancer metabolism,” *Cell Metabolism*, vol. 23, no. 1, pp. 27–47, 2016.
- [3] J. T. Yurkovich and B. O. Palsson, “Solving puzzles with missing pieces: The power of systems biology [Point of View],” *Proceedings of the IEEE*, vol. 104, no. 1, pp. 2–7, 2016.
- [4] A. Bordbar, J. T. Yurkovich, G. Paglia, O. Rolfsson, Ó. E. Sigurjónsson, and B. O. Palsson, “Elucidating dynamic metabolic physiology through network integration of quantitative time-course metabolomics,” *Scientific Reports*, vol. 7, p. 46249, 2017.
- [5] C. Angione, M. Conway, and P. Lió, “Multiplex methods provide effective integration of multi-omic data in genome-scale models,” *BMC Bioinformatics*, vol. 17, p. 257, 2016.
- [6] E. Motamedian, M. Mohammadi, S. A. Shojaosadati, and M. Heydari, “TRFBA: An algorithm to integrate genome-scale metabolic and transcriptional regulatory networks with incorporation of expression data,” *Bioinformatics*, vol. 33, no. 7, pp. 1057–1063, 2017.
- [7] F. Eyassu and C. Angione, “Modelling pyruvate dehydrogenase under hypoxia and its role in cancer metabolism,” *Royal Society Open Science*, vol. 4, no. 10, Article ID 170360, 2017.
- [8] V. Dougherty, T. J. Moutinho Jr, and J. Papin, “Accelerating the drug development pipeline with genome-scale metabolic network reconstructions,” *Systems Biology*, vol. 6, 2017.
- [9] E. Yaneske and C. Angione, “The poly-omics of ageing through individual-based metabolic modelling,” *BMC Bioinformatics*, vol. 19, p. 415, 2018.
- [10] L. Tobalina, J. Pey, A. Rezola, and F. J. Planes, “Assessment of FBA based gene essentiality analysis in cancer with a fast context-specific network reconstruction method,” *PLoS ONE*, vol. 11, no. 5, p. e0154583, 2016.
- [11] A. Occhipinti, F. Eyassu, T. J. Rahman, P. K. Rahman, and C. Angione, “In silico engineering of *Pseudomonas* metabolism reveals new biomarkers for increased biosurfactant production,” *PeerJ*, vol. 6, p. e6046, 2018.
- [12] Z. Qi and E. O. Voit, “Inference of cancer mechanisms through computational systems analysis,” *Molecular BioSystems*, vol. 13, no. 3, pp. 489–497, 2017.
- [13] D. C. Zielinski, N. Jamshidi, A. J. Corbett, A. Bordbar, A. Thomas, and B. O. Palsson, “Systems biology analysis of drivers underlying hallmarks of cancer cell metabolism,” *Scientific Reports*, vol. 7, 2017.
- [14] Z. Dai and J. W. Locasale, “Understanding metabolism with flux analysis: From theory to application,” *Metabolic Engineering*, vol. 43, pp. 94–102, 2017.
- [15] M. A. Keibler, T. M. Wasylenko, J. K. Kelleher, O. Iliopoulos, M. G. Vander Heiden, and G. Stephanopoulos, “Metabolic requirements for cancer cell proliferation,” *Cancer & Metabolism*, vol. 4, article no. 16, 2016.
- [16] D. Welter, J. MacArthur, J. Morales et al., “The NHGRI GWAS Catalog, a curated resource of SNP-trait associations,” *Nucleic Acids Research*, vol. 42, no. 1, pp. D1001–D1006, 2014.
- [17] J. Nielsen, “Systems biology of metabolism: a driver for developing personalized and precision medicine,” *Cell Metabolism*, vol. 25, no. 3, pp. 572–579, 2017.
- [18] H. Kitano, “Systems biology: a brief overview,” *Science*, vol. 295, pp. 1662–1664, 2002.
- [19] M. Tomita, K. Hashimoto, K. Takahashi et al., “E-CELL: Software environment for whole-cell simulation,” *Bioinformatics*, vol. 15, no. 1, pp. 72–84, 1999.
- [20] M. Tomita, “Whole-cell simulation: A grand challenge of the 21st century,” *Trends in Biotechnology*, vol. 19, no. 6, pp. 205–210, 2001.
- [21] J. R. Karr, J. C. Sanghvi, D. N. MacKlin et al., “A whole-cell computational model predicts phenotype from genotype,” *Cell*, vol. 150, no. 2, pp. 389–401, 2012.
- [22] A. P. Goldberg, B. Szigeti, Y. H. Chew, J. A. Sekar, Y. D. Roth, and J. R. Karr, “Emerging whole-cell modeling principles and methods,” *Current Opinion in Biotechnology*, vol. 51, pp. 97–102, 2018.
- [23] A. C. Babbie and M. P. H. Stumpf, “How to deal with parameters for whole-cell modelling,” *Journal of the Royal Society Interface*, vol. 14, Article ID 20170237, 2017.
- [24] P. Palumbo, M. Vanoni, F. Papa, S. Busti, and L. Alberghina, *Whole yeast model: what and why*, 2018.
- [25] A. Khodayari, A. R. Zomorodi, J. C. Liao, and C. D. Maranas, “A kinetic model of *Escherichia coli* core metabolism satisfying multiple sets of mutant flux data,” *Metabolic Engineering*, vol. 25, pp. 50–62, 2014.
- [26] K. Tummaller and E. Klipp, “The discrepancy between data for and expectations on metabolic models: How to match experiments and computational efforts to arrive at quantitative predictions?” *Current Opinion in Systems Biology*, vol. 8, pp. 1–6, 2018.
- [27] C. Ramon, M. Gollub, and J. Stelling, “Integrating -omics data into genome-scale metabolic network models: principles and challenges,” *Essays in Biochemistry*, vol. 62, no. 4, pp. 563–574, 2018.
- [28] A. Khodayari and C. D. Maranas, “A genome-scale *Escherichia coli* kinetic metabolic model k-ecoli457 satisfying flux data

- for multiple mutant strains,” *Nature Communications*, vol. 7, p. 13806, 2016.
- [29] A. A. Mannan, Y. Toya, K. Shimizu, J. McFadden, A. M. Kierzek, and A. Rocco, “Integrating kinetic model of *E. coli* with genome scale metabolic fluxes overcomes its open system problem and reveals bistability in central metabolism,” *PLoS ONE*, vol. 10, no. 10, Article ID e0139507, 2015.
- [30] A. Zelezniak, S. Sheridan, and K. R. Patil, “Contribution of network connectivity in determining the relationship between gene expression and metabolite concentration changes,” *PLoS Computational Biology*, vol. 10, Article ID e1003572, 2014.
- [31] M. Kanehisa, M. Furumichi, M. Tanabe, Y. Sato, and K. Morishima, “KEGG: new perspectives on genomes, pathways, diseases and drugs,” *Nucleic Acids Research*, vol. 45, no. 1, pp. D353–D361, 2017.
- [32] P. D. Karp, R. Billington, R. Caspi et al., “The BioCyc collection of microbial genomes and metabolic pathways,” *Briefings in Bioinformatics*, vol. 1, p. 9, 2017.
- [33] S. Placzek, I. Schomburg, A. Chang et al., “BRENDA in 2017: New perspectives and new tools in BRENDA,” *Nucleic Acids Research*, vol. 45, no. 1, pp. D380–D388, 2017.
- [34] D. A. Cuevas, J. Edirisinghe, C. S. Henry, R. Overbeek, T. G. O’Connell, and R. A. Edwards, “From DNA to FBA: How to build your own genome-scale metabolic model,” *Frontiers in Microbiology*, vol. 7, 2016.
- [35] M. Fondi and P. Liò, “Genome-scale metabolic network reconstruction,” *Bacterial Pangenomics: Methods and Protocols*, pp. 233–256, 2015.
- [36] I. Thiele and B. Ø. Palsson, “A protocol for generating a high-quality genome-scale metabolic reconstruction,” *Nature Protocols*, vol. 5, no. 1, pp. 93–121, 2010.
- [37] Ø. Palsson, *Systems Biology: Constraint-Based Reconstruction and Analysis*, Cambridge University Press, 2015.
- [38] M. Conway, C. Angione, and P. Liò, “Iterative multi level calibration of metabolic networks,” *Current Bioinformatics*, vol. 11, no. 1, pp. 93–105, 2016.
- [39] F. Shen, R. Sun, J. Yao et al., “OptRAM: In-silico strain design via integrative regulatory-metabolic network modeling,” *PLoS Computational Biology*, vol. 15, Article ID e1006835, 2019.
- [40] A. von Kamp and S. Klamt, “Enumeration of smallest intervention strategies in genome-scale metabolic networks,” *PLoS Computational Biology*, vol. 10, 2014.
- [41] C. Angione, J. Costanza, G. Carapezza, P. Lió, and G. Nicosia, “Analysis and design of molecular machines,” *Theoretical Computer Science*, vol. 599, pp. 102–117, 2015.
- [42] P. Schneider and S. Klamt, “Characterizing and ranking computed metabolic engineering strategies,” *Bioinformatics*, 2019.
- [43] J. D. Orth, I. Thiele, and B. O. Palsson, “What is flux balance analysis?” *Nature Biotechnology*, vol. 28, no. 3, pp. 245–248, 2010.
- [44] R. Mahadevan, J. S. Edwards, and F. J. Doyle III, “Dynamic flux balance analysis of diauxic growth in *Escherichia coli*,” *Biophysical Journal*, vol. 83, no. 3, pp. 1331–1340, 2002.
- [45] N. Töpfer, S. Kleessen, and Z. Nikoloski, “Integration of metabolomics data into metabolic networks,” *Frontiers in Plant Science*, vol. 6, p. 49, 2015.
- [46] P. Salvy, G. Fengos, M. Ataman et al., “pyTFA and matTFA: a python package and a matlab toolbox for thermodynamics-based flux analysis,” *Bioinformatics*, pp. 167–169, 2018.
- [47] E. Brunk, R. L. Chang, J. Xia et al., “Systemic post-translational control of bacterial metabolism regulates adaptation in dynamic environments,” 2018.
- [48] M. MacGillivray, A. Ko, E. Gruber, M. Sawyer, E. Almaas, and A. Holder, “Robust analysis of fluxes in genome-scale metabolic pathways,” *Scientific Reports*, vol. 7, p. 268, 2017.
- [49] E. Ebrahim, E. Almaas, A. Bauer et al. et al., “Do genome-scale models need exact solvers or clearer standards?” *Molecular systems biology*, vol. 11, article no. 831, 2015.
- [50] L. Chindelevitch, J. Trigg, A. Regev, and B. Berger, “Reply to “do genome-scale models need exact solvers or clearer standards?”,” *Molecular Systems Biology*, vol. 11, no. 10, 2015.
- [51] P. D. Karp, D. Weaver, and M. Latendresse, “How accurate is automated gap filling of metabolic models?” *BMC Systems Biology*, vol. 12, p. 73, 2018.
- [52] N. E. Lewis, K. K. Hixson, T. M. Conrad et al., “Omic data from evolved *E. coli* are consistent with computed optimal growth from genome-scale models,” *Molecular Systems Biology*, vol. 6, p. 390, 2010.
- [53] D. Machado, M. J. Herrgård, I. Rocha, and K. R. Patil, “Stoichiometric representation of gene–protein–reaction associations leverages constraint-based analysis from reaction to gene-level phenotype prediction,” *PLoS Computational Biology*, vol. 12, no. 10, p. e1005140, 2016.
- [54] M. K. Kim, A. Lane, J. J. Kelley, and D. S. Lun, “E-Flux2 and sPOT: Validated methods for inferring intracellular metabolic flux distributions from transcriptomic data,” *PLoS ONE*, vol. 11, Article ID e0157101, 2016.
- [55] K. Smallbone and E. Simeonidis, “Flux balance analysis: a geometric perspective,” *Journal of Theoretical Biology*, vol. 258, no. 2, pp. 311–315, 2009.
- [56] N. D. Price, J. Schellenberger, and B. O. Palsson, “Uniform sampling of steady-state flux spaces: Means to design experiments and to interpret enzymopathies,” *Biophysical Journal*, vol. 87, no. 4, pp. 2172–2186, 2004.
- [57] J. Schellenberger, R. Que, R. M. T. Fleming et al., “Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0,” *Nature Protocols*, vol. 6, no. 9, pp. 1290–1307, 2011.
- [58] A. Ebrahim, J. A. Lerman, B. Ø. Palsson, and D. R. Hyduke, “COBRAPy: COntstraints-based reconstruction and analysis for python,” *BMC Systems Biology*, vol. 7, no. 1, article 74, 2013.
- [59] H. Wang, S. Marčišauskas, B. J. Sánchez et al., “RAVEN 2.0: A versatile toolbox for metabolic network reconstruction and a case study on *Streptomyces coelicolor*,” *PLoS Computational Biology*, vol. 14, no. 10, p. e1006541, 2018.
- [60] M. Latendresse, M. Kruppenacker, M. Trupp, and P. D. Karp, “Construction and completion of flux balance models from pathway databases,” *Bioinformatics*, vol. 28, no. 3, pp. 388–396, 2012.
- [61] J. Boele, B. G. Olivier, and B. Teusink, “FAME, the flux analysis and modeling environment,” *BMC Systems Biology*, vol. 6, article no. 8, 2012.
- [62] J. P. Faria, M. Rocha, I. Rocha, and C. S. Henry, “Methods for automated genome-scale metabolic model reconstruction,” *Biochemical Society Transactions*, vol. 46, no. 4, pp. 931–936, 2018.
- [63] M. Lakshmanan, G. Koh, B. K. S. Chung, and D.-Y. Lee, “Software applications for flux balance analysis,” *Briefings in Bioinformatics*, vol. 15, no. 1, pp. 108–122, 2014.
- [64] L. K. Reed, C. F. Baer, and A. S. Edison, “Considerations when choosing a genetic model organism for metabolomics studies,” *Current Opinion in Chemical Biology*, vol. 36, pp. 7–14, 2017.

- [65] Y. Hasin, M. Seldin, and A. Lusic, "Multi-omics approaches to disease," *Genome Biology*, vol. 18, p. 83, 2017.
- [66] L. Heirendt, S. Arreckx, T. Pfau et al., "Creation and analysis of biochemical constraint-based models: the cobra toolbox v3. 0," <https://arxiv.org/abs/1710.04038>, 2017.
- [67] C. F. Arias, P. Catalán, S. Manrubia, and J. A. Cuesta, "ToyLIFE: A computational framework to study the multi-level organisation of the genotype-phenotype map," *Scientific Reports*, vol. 4, article no. 7549, 2014.
- [68] S. S. Fong, A. R. Joyce, and B. Ø. Palsson, "Parallel adaptive evolution cultures of *Escherichia coli* lead to convergent growth phenotypes with different gene expression states," *Genome Research*, vol. 15, no. 10, pp. 1365–1372, 2005.
- [69] C. Angione and P. Lió, "Predictive analytics of environmental adaptability in multi-omic network models," *Scientific Reports*, vol. 5, p. 15147, 2015.
- [70] J. Y. Ryu, H. U. Kim, and S. Y. Lee, "Reconstruction of genome-scale human metabolic models using omics data," *Integrative Biology*, vol. 7, no. 8, pp. 859–868, 2015.
- [71] S. S. Kashaf, C. Angione, and P. Lió, "Making life difficult for *Clostridium difficile*: Augmenting the pathogen's metabolic model with transcriptomic and codon usage data for better therapeutic target characterization," *BMC Systems Biology*, vol. 11, p. 25, 2017.
- [72] M. Larocque, T. Chénard, and R. Najmanovich, "A curated *C. difficile* strain 630 metabolic network: prediction of essential targets and inhibitors," *BMC Systems Biology*, vol. 8, p. 117, 2014.
- [73] A. Mardinoglu, R. Agren, C. Kampf et al., "Integration of clinical data with a genome-scale metabolic model of the human Adipocyte," *Molecular Systems Biology*, vol. 9, p. 649, 2013.
- [74] O. Folger, L. Jerby, C. Frezza, E. Gottlieb, E. Ruppín, and T. Shlomi, "Predicting selective drug targets in cancer through metabolic networks," *Molecular Systems Biology*, vol. 7, p. 501, 2011.
- [75] K. Yizhak, B. Chaneton, E. Gottlieb, and E. Ruppín, "Modeling cancer metabolism on a genome scale," *Molecular Systems Biology*, vol. 11, p. 817, 2015.
- [76] N. E. Lewis and A. M. Abdel-Haleem, "The evolution of genome-scale models of cancer metabolism," *Frontiers in Physiology*, vol. 4, 2013.
- [77] M. Uhlen, C. Zhang, S. Lee et al., "A pathology atlas of the human cancer transcriptome," *Science*, vol. 357, no. 6352, Article ID eaan2507, 2017.
- [78] D. Machado and M. Herrgård, "Systematic evaluation of methods for integration of transcriptomic data into constraint-based models of metabolism," *PLoS Computational Biology*, vol. 10, no. 4, Article ID e1003580, 2014.
- [79] J. S. Cho, C. Gu, T. H. Han, J. Y. Ryu, and S. Y. Lee, "Reconstruction of context-specific genome-scale metabolic models using multiomics data to study metabolic rewiring," *Current Opinion in Systems Biology*, vol. 15, pp. 1–11, 2019.
- [80] S. Vijayakumar, M. Conway, P. Lió, and C. Angione, "Seeing the wood for the trees: a forest of methods for optimization and omic-network integration in metabolic modelling," *Briefings in Bioinformatics*, 2017.
- [81] S. Opdam, A. Richelle, B. Kellman, S. Li, D. C. Zielinski, and N. E. Lewis, "A systematic evaluation of methods for tailoring genome-scale metabolic models," *Cell Systems*, vol. 4, no. 3, pp. 318–329, 2017.
- [82] C. Angione, "Integrating splice-isoform expression into genome-scale models characterizes breast cancer metabolism," *Bioinformatics*, vol. 34, no. 3, pp. 494–501, 2018.
- [83] J. Y. Ryu, H. U. Kim, and S. Y. Lee, "Framework and resource for more than 11,000 gene-transcript-protein-reaction associations in human metabolism," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 114, no. 45, pp. E9740–E9749, 2017.
- [84] M. Tian and J. L. Reed, "Integrating Proteomic or Transcriptomic Data into Metabolic Models Using Linear Bound Flux Balance Analysis," *Bioinformatics*, vol. 34, p. 3882, 2018.
- [85] J. Costanza, G. Carapezza, C. Angione, P. Lió, and G. Nicosia, "Robust design of microbial strains," *Bioinformatics*, vol. 28, no. 23, pp. 3097–3104, 2012.
- [86] A. Kiparissides and V. Hatzimanikatis, "Thermodynamics-based Metabolite Sensitivity Analysis in metabolic networks," *Metabolic Engineering*, vol. 39, pp. 117–127, 2017.
- [87] E. W. Birch, M. Udell, and M. W. Covert, "Incorporation of flexible objectives and time-linked simulation with flux balance analysis," *Journal of Theoretical Biology*, vol. 345, pp. 12–21, 2014.
- [88] J. C. Xavier, K. R. Patil, and I. Rocha, "Integration of biomass formulations of genome-scale metabolic models with experimental data reveals universally essential cofactors in prokaryotes," *Metabolic Engineering*, vol. 39, pp. 200–208, 2017.
- [89] A. M. Feist and B. O. Palsson, "What do cells actually want?" *Genome Biology*, vol. 17, no. 1, 2016.
- [90] B. Dikicioglu, S. G. Krdar, and Oliver., "Biomass composition: the elephant in the room of metabolic modelling," *Metabolomics*, vol. 11, pp. 1690–1701, 2015.
- [91] S. H. J. Chan, J. Cai, L. Wang, M. N. Simons-Senftle, and C. D. Maranas, "Standardizing biomass reactions and ensuring complete mass balance in genome-scale metabolic models," *Bioinformatics*, vol. 33, no. 22, pp. 3603–3609, 2017.
- [92] J.-C. Lachance, J. M. Monk, C. J. Lloyd et al., "Bofdat: generating biomass objective function stoichiometric coefficients from experimental data," *BioRxiv*, Article ID 243881, 2018.
- [93] L. Yang, J. Bento, J.-C. Lachance, and B. O. Palsson, "Estimating Cellular Goals from High-Dimensional Biological Data," <https://arxiv.org/abs/1807.04245>, 2018.
- [94] R. Schuetz, N. Zamboni, M. Zampieri, M. Heinemann, and U. Sauer, "Multidimensional optimality of microbial metabolism," *Science*, vol. 336, no. 6081, pp. 601–604, 2012.
- [95] G. A. Parker and J. M. Smith, "Optimality theory in evolutionary biology," *Nature*, vol. 348, no. 6296, pp. 27–33, 1990.
- [96] P. Zakrzewski, M. H. Medema, A. Gevorgyan, A. M. Kierzek, R. Breitling, and E. Takano, "MultiMetEval: comparative and multi-objective analysis of genome-scale metabolic models," *PLoS ONE*, vol. 7, no. 12, Article ID e51511, 2012.
- [97] S. Vijayakumar, M. Conway, P. Lió, and C. Angione, "Optimization of multi-omic genome-scale models: Methodologies, hands-on tutorial, and perspectives," in *Metabolic Network Reconstruction and Modeling*, pp. 389–408, Springer, 2018.
- [98] A. L. Knorr, R. Jain, and R. Srivastava, "Bayesian-based selection of metabolic objective functions," *Bioinformatics*, vol. 23, no. 3, pp. 351–357, 2007.
- [99] J. O. H. Sendín, O. Exler, and J. R. Banga, "Multi-objective mixed integer strategy for the optimisation of biological networks," *IET Systems Biology*, vol. 4, no. 3, pp. 236–248, 2010.
- [100] Y.-G. Oh, D.-Y. Lee, S. Y. Lee, and S. Park, "Multiobjective flux balancing using the NISE method for metabolic network analysis," *Biotechnology Progress*, vol. 25, no. 4, pp. 999–1008, 2009.
- [101] G. Xu, "An iterative strategy for bi-objective optimization of metabolic pathways," in *Proceedings of the 2011 Fourth*

- International Joint Conference on Computational Sciences and Optimization*, pp. 587-588, IEEE Computer Society, 2011.
- [102] G. M. de Hijas-Liste, E. Klipp, E. Balsa-Canto, and J. R. Banga, "Global dynamic optimization approach to predict activation in metabolic pathways," *BMC Systems Biology*, vol. 8, no. 1, 2014.
- [103] D. Nagrath, M. Avila-Elchiver, F. Berthiaume, A. W. Tilles, A. Messac, and M. L. Yarmush Martin L., "Soft constraints-based multiobjective framework for flux balance analysis," *Metabolic Engineering*, vol. 12, no. 5, pp. 429-445, 2010.
- [104] A. R. Zomorodi and C. D. Maranas, "OptCom: a multi-level optimization framework for the metabolic modeling and analysis of microbial communities," *PLoS Computational Biology*, vol. 8, no. 2, Article ID e1002363, 2012.
- [105] B. A. Boghigian, H. Shi, K. Lee, and B. A. Pfeifer, "Utilizing elementary mode analysis, pathway thermodynamics, and a genetic algorithm for metabolic flux determination and optimal metabolic network design," *BMC Systems Biology*, vol. 4, article no. 49, 2010.
- [106] C. Angione, J. Costanza, G. Carapezza, P. Lió, and G. Nicosia, "Multi-target analysis and design of mitochondrial metabolism," *PLoS ONE*, vol. 10, no. 9, Article ID e0133825, 2015.
- [107] E. Balsa-Canto, D. Henriques, A. Gábor, and J. R. Banga, "AMIGO2, a toolbox for dynamic modeling, optimization and control in systems biology," *Bioinformatics*, vol. 32, no. 21, pp. 3357-3359, 2016.
- [108] E. H. Wintermute, T. D. Lieberman, and P. A. Silver, "An objective function exploiting suboptimal solutions in metabolic networks," *BMC Systems Biology*, vol. 7, p. 98, 2013.
- [109] R. J. Deberardinis and C. B. Thompson, "Cellular metabolism and disease: What do metabolic outliers teach us?" *Cell*, vol. 148, no. 6, pp. 1132-1144, 2012.
- [110] B. H. Goodpaster and L. M. Sparks, "Metabolic flexibility in health and disease," *Cell Metabolism*, vol. 25, no. 5, pp. 1027-1036, 2017.
- [111] C. H. Schilling and B. Ø. Palsson, "Assessment of the metabolic capabilities of *Haemophilus influenzae* Rd through a genome-scale pathway analysis," *Journal of Theoretical Biology*, vol. 203, no. 3, pp. 249-283, 2000.
- [112] B. P. Edwards, "The *Escherichia coli* mg1655 in silico metabolic genotype: its definition, characteristics, and capabilities," in *Proceedings of the National Academy of Sciences* 97, pp. 5528-5533, 2000.
- [113] C. Venter, M. D. Adams, E. W. Myers et al. et al., "The sequence of the human genome," *Science*, vol. 291, pp. 1304-1351, 2001.
- [114] H. G. S. Consortium et al., "Finishing the euchromatic sequence of the human genome," *Nature*, vol. 431, p. 931, 2004.
- [115] P. Romero, J. Wagg, M. L. Green, D. Kaiser, M. Krummenacker, and P. D. Karp, "Computational prediction of human metabolic pathways from the complete human genome," *Genome Biology*, vol. 6, no. 1, p. R2, 2005.
- [116] G. Joshi-Tope, M. Gillespie, I. Vastrik et al., "Reactome: a knowledgebase of biological pathways," *Nucleic Acids Research*, vol. 33, pp. D428-D432, 2005.
- [117] A. Fabregat, K. Sidiropoulos, P. Garapati et al., "The reactome pathway knowledgebase," *Nucleic Acids Research*, vol. 44, no. 1, pp. D481-D487, 2016.
- [118] N. C. Duarte, S. A. Becker, N. Jamshidi et al., "Global reconstruction of the human metabolic network based on genomic and bibliomic data," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 6, pp. 1777-1782, 2007.
- [119] H. Ma, A. Sorokin, A. Mazein et al., "The Edinburgh human metabolic network reconstruction and its functional analysis," *Molecular Systems Biology*, vol. 3, article 135, 2007.
- [120] T. Hao, H.-W. Ma, X.-M. Zhao, and I. Goryanin, "Compartmentalization of the Edinburgh human metabolic network," *BMC Bioinformatics*, vol. 11, article no. 393, 2010.
- [121] I. Thiele, N. Swainston, R. M. T. Fleming et al., "A community-driven global reconstruction of human metabolism," *Nature Biotechnology*, vol. 31, no. 5, pp. 419-425, 2013.
- [122] C. Gille, C. Bölling, A. Hoppe et al., "HepatoNet1: A comprehensive metabolic reconstruction of the human hepatocyte for the analysis of liver physiology," *Molecular Systems Biology*, vol. 6, p. 411, 2010.
- [123] S. Sahoo, L. Franzson, J. J. Jonsson, and I. Thiele, "A compendium of inborn errors of metabolism mapped onto the human metabolic network," *Molecular BioSystems*, vol. 8, no. 10, pp. 2545-2558, 2012.
- [124] N. Pornputtpong, I. Nookaew, and J. Nielsen, "Human metabolic atlas: An online resource for human metabolism," *Database*, vol. 2015, Article ID bav068, 2015.
- [125] A. Mardinoglu, R. Agren, C. Kampf, A. Asplund, M. Uhlen, and J. Nielsen, "Genome-scale metabolic modelling of hepatocytes reveals serine deficiency in patients with non-alcoholic fatty liver disease," *Nature Communications*, vol. 5, p. 3083, 2014.
- [126] Smallbone, "Striking a balance with recon 2.1," <https://arxiv.org/abs/1311.5696>, 2013.
- [127] L.-E. Quek, S. Dietmair, M. Hanscho, V. S. Martínez, N. Borth, and L. K. Nielsen, "Reducing Recon 2 for steady-state flux analysis of HEK cell culture," *Journal of Biotechnology*, vol. 184, pp. 172-178, 2014.
- [128] S. Sahoo, M. K. Aurich, J. J. Jonsson, and I. Thiele, "Membrane transporters in a human genome-scale metabolic knowledgebase and their implications for disease," *Frontiers in Physiology*, vol. 5, p. 91, 2014.
- [129] S. Sahoo, H. S. Haraldsdóttir, R. M. T. Fleming, and I. Thiele, "Modeling the effects of commonly used drugs on human metabolism," *FEBS Journal*, vol. 282, no. 2, pp. 297-317, 2015.
- [130] N. Swainston, K. Smallbone, H. Hefzi et al., "Recon 2.2: from reconstruction to model of human metabolism," *Metabolomics*, vol. 12, pp. 1-7, 2016.
- [131] E. M. Blais, K. D. Rawls, B. V. Dougherty et al., "Reconciled rat and human metabolic networks for comparative toxicogenomics and biomarker predictions," *Nature Communications*, vol. 8, p. 14250, 2017.
- [132] R. Caspi, R. Billington, L. Ferrer et al., "The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases," *Nucleic Acids Research*, vol. 44, no. 1, pp. D471-D480, 2016.
- [133] E. Brunk, S. Sahoo, D. C. Zielinski et al., "Recon3D enables a three-dimensional view of gene variation in human metabolism," *Nature Biotechnology*, vol. 36, no. 3, pp. 272-281, 2018.
- [134] S. Moretti, O. Martin, T. Van Du Tran, A. Bridge, A. Morgat, and M. Pagni, "MetaNetX/MNXref - Reconciliation of metabolites and biochemical reactions to bring together genome-scale metabolic networks," *Nucleic Acids Research*, vol. 44, no. 1, pp. D523-D526, 2016.
- [135] X. Feng, Y. Xu, Y. Chen, and Y. J. Tang, "MicrobesFlux: a web platform for drafting metabolic models from the KEGG database," *BMC Systems Biology*, vol. 6, article no. 94, 2012.

- [136] P. D. Karp, M. Latendresse, S. M. Paley et al., "Pathway Tools version 19.0 update: software for pathway/genome informatics and systems biology," *Briefings in Bioinformatics*, vol. 17, no. 5, pp. 877–890, 2016.
- [137] R. Liberal, B. K. Lisowska, D. J. Leak, and J. W. Pinney, "PathwayBooster: A tool to support the curation of metabolic pathways," *BMC Bioinformatics*, vol. 16, article no. 86, 2015.
- [138] E. Pitkänen, P. Jouhten, J. Hou et al., "Comparative genome-scale reconstruction of gapless metabolic networks for present and ancestral species," *PLoS Computational Biology*, vol. 10, no. 2, Article ID e1003465, 2014.
- [139] O. Dias, M. Rocha, E. C. Ferreira, and I. Rocha, "Reconstructing genome-scale metabolic models with merlin," *Nucleic Acids Research*, vol. 43, no. 8, pp. 3899–3910, 2015.
- [140] D. Machado, S. Andrejev, M. Tramontano, and K. R. Patil, "Fast automated reconstruction of genome-scale metabolic models for microbial species and communities," *Nucleic Acids Research*, vol. 46, no. 15, pp. 7542–7553, 2018.
- [141] N. Martyushenko and E. Almaas, "ModelExplorer - software for visual inspection and inconsistency correction of genome-scale metabolic reconstructions," *BMC Bioinformatics*, vol. 20, p. 56, 2019.
- [142] E. Özcan and T. Çakir, "Reconstructed metabolic network models predict flux-level metabolic reprogramming in glioblastoma," *Frontiers in Neuroscience*, vol. 10, 2016.
- [143] A. Mardinoglu, C. Kampf, A. Asplund et al., "Defining the human adipose tissue proteome to reveal metabolic alterations in obesity," *Journal of Proteome Research*, vol. 13, no. 11, pp. 5106–5119, 2014.
- [144] L. Jerby, L. Wolf, C. Denkert et al., "Metabolic associations of reduced proliferation and oxidative stress in advanced breast cancer," *Cancer Research*, vol. 72, no. 22, pp. 5712–5720, 2012.
- [145] A. Karlstädt, D. Fliegner, G. Kararigas, H. S. Ruderisch, V. Regitz-Zagrosek, and H.-G. Holzhütter, "CardioNet: A human metabolic network suited for the study of cardiomyocyte metabolism," *BMC Systems Biology*, vol. 6, p. 114, 2012.
- [146] S. Sohrabi-Jahromi, S.-A. Marashi, and S. Kalantari, "A kidney-specific genome-scale metabolic network model for analyzing focal segmental glomerulosclerosis," *Mammalian Genome*, vol. 27, no. 3-4, pp. 158–167, 2016.
- [147] L. Väre, C. Scheele, C. Broholm et al., "Proteome- and transcriptome-driven reconstruction of the human myocyte metabolic network and its use for identification of markers for diabetes," *Cell Reports*, vol. 11, pp. 921–933, 2015.
- [148] R. Pagliarini, R. Castello, F. Napolitano et al., "In silico modeling of liver metabolism in a human disease reveals a key enzyme for histidine and histamine homeostasis," *Cell Reports*, vol. 15, no. 10, pp. 2292–2300, 2016.
- [149] A. Schultz, A. A. Qutub, and C. D. Maranas, "Reconstruction of tissue-specific metabolic networks using CORDA," *PLoS Computational Biology*, vol. 12, Article ID e1004808, 2016.
- [150] R. Agren, S. Bordel, A. Mardinoglu, N. Pornputtpong, I. Nookaew, and J. Nielsen, "Reconstruction of genome-scale active metabolic networks for 69 human cell types and 16 cancer types using INIT," *PLoS Computational Biology*, vol. 8, no. 5, Article ID e1002518, 2012.
- [151] Y. Wang, J. A. Eddy, and N. D. Price, "Reconstruction of genome-scale metabolic models for 126 human tissues using mCADRE," *BMC Systems Biology*, vol. 6, p. 153, 2012.
- [152] K. Yizhak, E. Gaude, S. Le Dévédec et al., "Phenotype-based cell-specific metabolic modeling reveals metabolic liabilities of cancer," *eLife*, vol. 3, Article ID e03641, 2014.
- [153] C. Frezza, L. Zheng, O. Folger et al., "Haem oxygenase is synthetically lethal with the tumour suppressor fumarate hydratase," *Nature*, vol. 477, no. 7363, pp. 225–228, 2011.
- [154] P. Ghaffari, A. Mardinoglu, A. Asplund et al., "Identifying anti-growth factors for human cancer cell lines through genome-scale metabolic modeling," *Scientific Reports*, vol. 5, p. 8183, 2015.
- [155] M. K. Aurich, G. Paglia, Ó. Rolfsson et al., "Prediction of intracellular metabolic states from extracellular metabolomic data," *Metabolomics*, vol. 11, no. 3, pp. 603–619, 2015.
- [156] A. Mardinoglu, E. Bjornson, C. Zhang et al., "Personal model-assisted identification of NAD<sup>+</sup> and glutathione metabolism as intervention target in NAFLD," *Molecular Systems Biology*, vol. 13, no. 916, 2017.
- [157] W. Zhu, J. C. Gregory, E. Org et al., "Gut microbial metabolite TMAO enhances platelet hyperreactivity and thrombosis risk," *Cell*, vol. 165, no. 1, pp. 111–124, 2016.
- [158] M. J. Miller, A. D. Kennedy, A. D. Eckhart et al., "Untargeted metabolomic analysis for the clinical screening of inborn errors of metabolism," *Journal of Inherited Metabolic Disease*, vol. 38, no. 6, pp. 1029–1039, 2015.
- [159] A. Blanco-Míguez, F. Fdez-Riverola, B. Sánchez, and A. Lourenço, "Resources and tools for the high-throughput, multi-omic study of intestinal microbiota," *Briefings in Bioinformatics*, 2017.
- [160] A. Heinken and I. Thiele, "Systematic prediction of health-relevant human-microbial co-metabolism through a computational framework," *Gut Microbes*, vol. 6, no. 2, pp. 120–130, 2015.
- [161] S. Magnúsdóttir, A. Heinken, L. Kutt et al., "Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota," *Nature Biotechnology*, vol. 35, no. 1, pp. 81–89, 2017.
- [162] D. A. Heinken, F. Ravcheev, L. Baldini, R. M. Heirendt, and I. Fleming, "Thiele, Personalized modeling of the human gut microbiome reveals distinct bile acid deconjugation and biotransformation potential in healthy and ibd individuals," *bioRxiv*, Article ID 229138, 2017.
- [163] P. Babaei, S. Shoaie, B. Ji, and J. Nielsen, "Challenges in modeling the human gut microbiome," *Nature Biotechnology*, vol. 36, no. 8, pp. 682–686, 2018.
- [164] A. Di Stefano, M. Scatà, S. Vijayakumar et al., "Social dynamics modeling of chrono-nutrition," *PLoS Computational Biology*, 2019.
- [165] K. C. H. van der Ark, R. G. A. van Heck, V. A. P. Martins Dos Santos, C. Belzer, and W. M. de Vos, "More than just a gut feeling: constraint-based genome-scale metabolic models for predicting functions of human intestinal microbes," *Microbiome*, vol. 5, p. 78, 2017.
- [166] A. Bordbar, J. M. Monk, Z. A. King, and B. O. Palsson, "Constraint-based models predict metabolic and associated cellular functions," *Nature Reviews Genetics*, vol. 15, no. 2, pp. 107–120, 2014.
- [167] E. Bosi, G. Bacci, A. Mengoni, and M. Fondi, "Perspectives and challenges in microbial communities metabolic modeling," *Frontiers in Genetics*, vol. 8, p. 88, 2017.
- [168] C. S. Henry, H. C. Bernstein, P. Weisenhorn et al., "Microbial community metabolic modeling: a community data-driven network reconstruction," *Journal of Cellular Physiology*, vol. 231, no. 11, pp. 2339–2345, 2016.
- [169] P. P. Sadhukhan and A. Raghunathan, "Investigating host-pathogen behavior and their interaction using genome-scale

- metabolic network models,” *Immunoinformatics*, pp. 523–562, 2014.
- [170] C. Damiani, M. Di Filippo, D. Pescini, D. Maspero, R. Colombo, and G. Mauri, “PopFBA: Tackling intratumour heterogeneity with Flux Balance Analysis,” *Bioinformatics*, vol. 33, no. 14, pp. i311–i318, 2017.
- [171] A. Bordbar, A. M. Feist, R. Usaita-Black, J. Woodcock, B. O. Palsson, and I. Famili, “A multi-tissue type genome-scale metabolic network for analysis of whole-body systems physiology,” *BMC Systems Biology*, vol. 5, p. 180, 2011.
- [172] M. K. Toroghi, W. R. Cluett, and R. Mahadevan, “A multi-scale model of the whole human body based on dynamic parsimonious flux balance analysis,” *IFAC-PapersOnLine*, vol. 49, no. 7, pp. 937–942, 2016.
- [173] S. Thiele, A. Sahoo, L. Heinken et al., “When metabolism meets physiology: Harvey and harvetta,” *BioRxiv*, Article ID 255885, 2018.
- [174] T. Ono, T. Fujii, Y. Sakai, and H. Kimura, “Development of organs-on-a-chip with metabolism model, in: Micro-NanoMechatronics and Human Science (MHS),” in *Proceedings of the 2016 International Symposium on*, pp. 1–3, IEEE, 2016.
- [175] Y. Zhang, V. Berka, A. Song et al., “Elevated sphingosine-1-phosphate promotes sickling and sickle cell disease progression,” *The Journal of Clinical Investigation*, vol. 124, no. 6, pp. 2750–2761, 2014.
- [176] D. C. Zielinski, F. V. Filipp, A. Bordbar et al., “Pharmacogenomic and clinical data link non-pharmacokinetic metabolic dysregulation to drug side effect pathogenesis,” *Nature Communications*, vol. 6, 2015.
- [177] M. Viceconti, A. Henney, and E. Morley-Fletcher, “In silico clinical trials: how computer simulation will transform the biomedical industry,” *International Journal of Clinical Trials*, vol. 3, no. 2, pp. 37–46, 2016.
- [178] T. Fuhrer, M. Zampieri, D. C. Sévin, U. Sauer, and N. Zamboni, “Genomewide landscape of gene–metabolome associations in *Escherichia coli*,” *Molecular Systems Biology*, vol. 13, article no. 907, 2017.
- [179] Z. Hosseini and S.-A. Marashi, “Discovering missing reactions of metabolic networks by using gene co-expression data,” *Scientific Reports*, vol. 7, 2017.
- [180] G. M. Calabrese, L. D. Mesner, J. P. Stains et al., “Integrating GWAS and co-expression network data identifies bone mineral density genes *sptbn1* and *mark3* and an osteoblast functional module,” *Cell Systems*, vol. 4, no. 1, pp. 46–59, 2017.
- [181] J. J. Li, P. J. Bickel, and M. D. Biggin, “System wide analyses have underestimated protein abundances and the importance of transcription in mammals,” *PeerJ*, vol. 2, no. 1, p. e270, 2014.
- [182] M. Jovanovic, M. S. Rooney, P. Mertins et al. et al., “Dynamic profiling of the protein life cycle in response to pathogens,” *Science*, vol. 347, Article ID 1259038, 2015.
- [183] I. Kosti, N. Jain, D. Aran, A. J. Butte, and M. Sirota, “Cross-tissue analysis of gene and protein expression in normal and cancer tissues,” *Scientific Reports*, vol. 6, 2016.
- [184] F. Edfors, F. Danielsson, B. M. Hallström et al., “Gene-specific correlation of RNA and protein levels in human cells and tissues,” *Molecular Systems Biology*, vol. 12, no. 10, p. 883, 2016.
- [185] N. Pandey and V. Hadadi, “Hatzimanikatis, Enhanced flux prediction by integrating relative expression and relative metabolite abundance into thermodynamically consistent metabolic models,” *bioRxiv*, Article ID 481499, 2018.
- [186] D. H. Pandey, A. C. Gardiol, and V. Pepe, “Tex-fba: A constraint-based method for integrating gene expression, thermodynamics, and metabolomics data into genome-scale metabolic models,” *bioRxiv*, Article ID 536235, 2019.
- [187] K. Jensen, S. Gudmundsson, and M. J. Herrgård, “Enhancing Metabolic Models with Genome-Scale Experimental Data,” in *Systems Biology*, RNA Technologies, pp. 337–350, Springer, 2018.
- [188] M. Jain, R. Nilsson, S. Sharma et al., “Metabolite profiling identifies a key role for glycine in rapid cancer cell proliferation,” *Science*, vol. 336, no. 6084, pp. 1040–1044, 2012.
- [189] N. Ishii, K. Nakahigashi, T. Baba et al., “Multiple high-throughput analyses monitor the response of *E. coli* to perturbations,” *Science*, vol. 316, no. 5824, pp. 593–597, 2007.
- [190] A. K. Holm, L. M. Blank, M. Oldiges et al., “Metabolic and transcriptional response to cofactor perturbations in *Escherichia coli*,” *The Journal of Biological Chemistry*, vol. 285, no. 23, pp. 17498–17506, 2010.
- [191] E. Rintala, M. Toivari, J.-P. Pitkänen, M. G. Wiebe, L. Ruohonen, and M. Penttilä, “Low oxygen levels as a trigger for enhancement of respiratory metabolism in *Saccharomyces cerevisiae*,” *BMC Genomics*, vol. 10, p. 461, 2009.
- [192] S. G. Wu, Y. Wang, W. Jiang et al., “Rapid prediction of bacterial heterotrophic fluxomics using machine learning and constraint programming,” *PLoS Computational Biology*, vol. 12, Article ID e1004838, 2016.