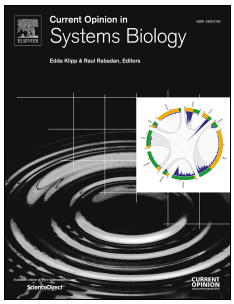
Metabolomics in systems medicine: an overview of methods and applications

Accepted Manuscript



Metabolomics in Systems Medicine: an overview of methods and applications

Effrosyni Karakitsou, Carles Foguet, Pedro de Atauri, Kim Kultima, Payam Emami Khoonsari, Vitor A.P. Martins dos Santos, Edoardo Saccenti, Antonio Rosato, Marta Cascante

PII: S2452-3100(18)30122-7

DOI: https://doi.org/10.1016/j.coisb.2019.03.009

Reference: COISB 239

To appear in: *Current Opinion in Systems Biology*

Received Date: 11 January 2019

Revised Date: 19 March 2019

Accepted Date: 20 March 2019

Please cite this article as: Karakitsou E, Foguet C, de Atauri P, Kultima K, Khoonsari PE, Martins dos Santos VAP, Saccenti E, Rosato A, Cascante M, Metabolomics in Systems Medicine: an overview of methods and applications, *Current Opinion in Systems Biology*, https://doi.org/10.1016/j.coisb.2019.03.009.

Metabolomics in Systems Medicine: an overview of methods and applications

Effrosyni Karakitsou[1,2†], Carles Foguet[1,3†], Pedro de Atauri[1,3], Kim Kultima[4], Payam Emami Khoonsari[4], Vitor A.P. Martins dos Santos[5,6], Edoardo Saccenti[5*], Antonio Rosato[7*], Marta Cascante[1,3*]

1. Department of Biochemistry and Molecular Biomedicine and Institute of Biomedicine (IBUB), Faculty of Biology, Universitat de Barcelona (UB), Barcelona, Spain
2. Institute of Cancer and Genomic Sciences and Centre for Computational Biology, University of Birmingham, B15 2TT, Birmingham, UK
3. Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas (CIBEREHD) and metabolomics node at INB-Bioinformatics Platform, Instituto de Salud Carlos III (ISCIII), Madrid, Spain
4. Department of Medical Sciences, Clinical Chemistry, Uppsala University, Uppsala, Sweden
5. Laboratory of Systems and Synthetic Biology, Wageningen University & Research, Stippeneng 4, 6708WE, Wageningen, the Netherlands
6. LifeGlimmer GmbH, Markelstraße 38, 12163 Berlin, Germany
7. Department of Chemistry and Magnetic Resonance Center, University of Florence, Italy

†These authors contributed equally to this work.

*Co-corresponding authors: Edoardo Saccenti, Antonio Rosato, Marta Cascante

E-mail: edoardo.saccenti@wur.nl, rosato@cerm.unifi.it, martacascante@ub.edu

Abstract

Patient-derived metabolomics offer valuable insights into the metabolic phenotype underlying diseases with a strong metabolic component. Thus, these datasets will be pivotal to the implementation of personalized medicine strategies in health and disease. However, to take full advantage of such data sets, they must be integrated with other omics within a coherent pathophysiological framework to enable improved diagnostics, to identify therapeutic interventions and to accurately stratify patients. Herein, we provide an overview of the state-of-the-art for different data analysis and modeling approaches applicable to metabolomics data and of their potential for systems medicine.

Highlights

- Metabolomics integrated with multi-omics datasets can strongly foster systems medicine
- Modeling and analysis of metabolic fluxes permit to decipher complex biological relationships
- Metabolomics-based systems medicine can improve diagnostics, therapeutics, and patient stratification

The brave new world of Systems Medicine

Systems biology treats biological systems as ensembles of networks at multiple levels, starting from the molecular level and from there gradually addressing more complex systems such as cells, tissues, organs, whole organisms and finally analyzing population-dynamics. Systems biology aims to describe and predict the behavior of groups of interacting components. To do so, it uses mathematical and computational tools to analyze measurements collected by systematic high-throughput technologies such as (post-)genomics, metabolomics or proteomics among others. The goal of systems approaches is to boost our understanding of biology by overcoming the limitations of reductive science, which addresses individual genes, proteins, metabolites, pathways or cells and thus does not account for the properties emerging from their interactions [1, 2].

Current medical science is mostly conducted using the reductionist approach [3, 4]. This limits our ability to grasp how multiple variables interact with one another to create emergent effects [5] and hampers our understanding of diseases, as well as our capability of delivering better treatments. Systems medicine can be regarded as the application of systems biology to human physiology in a clinical context [6, 7]. It addresses the above issues by applying iterative and reciprocal feedback between clinical research and practice through computational, statistical and mathematical multiscale analysis. This includes modeling of disease progression and remission, treatment responses and adverse events both at the epidemiological and patient level. This new paradigm of systems science and medicine strongly complements the traditional reductionist approach (Figure 1).

Figure 1: Overview of the core differences between reductionism and systems science, when analyzing the properties of a system; figure initially published in Tillmann et al. 2015 [8] under the terms of Creative Commons Attribution 2.0 license.

The functioning of the human body is regulated by the interaction and interdependencies of biological molecules at multiple levels (protein-protein, protein-RNA, protein-DNA networks, and metabolic networks) [9]. Therefore, it can only be efficiently analyzed by examining various omics concurrently. Systems medicine provides the appropriate framework to achieve this goal. The complementary perspectives offered by different datasets allow the genotype of an individual to be linked to its observed phenotype as a function of lifestyle and environmental conditions. Eventually, this could lead to defining how any healthy state can transition into a pathological one and vice versa and pave the way for personalized medicine.

Multi-omics data integration

The integration of multiple omics data (sometimes also called trans-omics) will further enhance the contribution of omics science to our understanding of biomedicine [10]. The example in Figure 2 shows the connections among genomics, transcriptomics, proteomics, and

metabolomics, thus providing an overview of the system from its potential (encoded in DNA) to the actual outcome (monitored by metabolomics).

Figure 2: Multi-Omics integration across different omics layers. Red arrows highlight the top-down flow of interactions across layers: genes are transcribed, transcripts determine enzyme concentrations, and finally enzymes act on metabolites. Purple arrows highlight bottom-up interactions, whereby metabolite levels modulate enzyme activities, the DNA/RNA binding affinities of regulators or DNA methylation. Note that metabolites can also interact directly with transcripts. Black arrows are intra-omic networks, which can be derived based on individual omic datasets (for a review of methods in metabolomics see Rosato et al. 2018 [11]). Intra-omic networks may describe direct physical interactions (e.g., protein-protein interactions) and correlations between their abundances (e.g., transcript levels or metabolite concentrations). Environmental stimuli (blue arrows) can affect all omic layers. For example, they can trigger DNA mutations, transcriptional events and modify protein activity. Additionally, the environment is also a source of metabolites and xenobiotic molecules. Overall, the different omics levels, which are a function of the environment and the omics interactions, determine the phenotype.

It is commonly accepted that the relationships between genes, gene products, and metabolites participate in complex, interconnected networks (Figure 2). Various biological molecules can be represented as nodes in a network and the interactions connecting them as edges. For example, in metabolomics, metabolites would be nodes and the edges would represent the enzymatic reactions interconnecting them. Graph theory can be applied to analyze the complexity of the interactions within a biological network and link *a priori* knowledge from the literature and databases [12]. The application of network analysis allows the identification of nodes with a high degree of connectivity (“hubs”) and groups of highly interconnected nodes (“modules”) identifying molecules functionally related to a disease state [13-15].

It is possible to outline a general strategy to integrate various omic datasets based on network representations. First, the network scaffold is defined by defining how the individual components are interconnected. The structure of the network can be identified based on the data or prior knowledge (i.e., database information). Subsequently, the network itself can be separated into modules. Finally, all the information can be combined with computational models of the whole system to simulate and predict how the network determines the observed phenotype. In practice, if two omics elements share a common driver, or if one perturbs the other, they will exhibit correlation or association. Various specialized statistical approaches can be applied to measure these correlations. For example, a linear model taking into account age, gender, body mass index, and white blood cell count was used to find correlations between DNA methylation and metabolite concentrations in human blood serum[16]. An even broader study analyzed the genome, transcriptome, proteome, metabolome, and metabolic fluxes in *Escherichia coli* to understand how its metabolic state reacted to perturbations [17]. More recently, weighted gene correlation network analysis was used to identify connectivity-based gene modules highly correlated to pathways identified by metabolomics [18].

Connecting the metabolome layer and other omic layers

Metabolomics measure the metabolites present within a cell, tissue or organism. It is a core experimental omic within systems biology, as it delivers an integrated view of biochemistry[19, 20]. Current experimental approaches in metabolomics are mostly based on Nuclear Magnetic Resonance (NMR) and Mass Spectroscopy (MS) [21, 22]. Metabolomic studies can be divided into two major groups: targeted and untargeted.

Targeted metabolomics quantitatively measure the abundance of a pre-defined group of known, well-characterized metabolites in a sample. Usually, the aim is to identify novel associations between metabolites in the context of specific physiological states [23, 24]. On the other hand, untargeted metabolomics typically focus on capturing all the chemical compounds present in a sample, including metabolites of unknown chemical structure, thus generating notably large datasets. By comparing the metabolome of control and test groups and focusing on the differences between their metabolic profiles, the number of significant detected signals becomes more manageable. Finally, the compounds or metabolites identified are annotated using *in silico* libraries when possible, or by applying analytical chemistry methods to explore the newly observed structure [25].

One of the technical challenges in connecting the metabolome with other omic layers is matching the identities of the same objects in different layers (ID conversion). Various databases support this task: the Kyoto Encyclopedia of Genes and Genomes (KEGG) integrates one computationally generated and fifteen manually curated databases, allowing the users to link metabolites to reactions, enzymes, pathways, and genes [26]; BRENDA provides information on enzymes, such as kinetic parameters for enzymatic reactions, allosteric effectors and association to diseases [27]; Reactome is a database that organizes metabolites into biological pathways and processes, using reactions to define relationships **[28]**; MetaCyc is a database of metabolic pathways and enzymes whereas BioCyc (BioCyc.org) collects organism-specific genomes and computationally predicted metabolic networks [29]. For example, in a multi-omic study on the flow of insulin signal based on time-course data from the metabolome, phosphoproteome, and transcriptome, a global metabolism map was generated by mapping quantitatively changed metabolites and their corresponding metabolic enzymes to the KEGG database [30].

Finally, it is worth mentioning the Investigation Study Assay (ISA-Tab) format, which is a convenient standard to store the metadata and the results of experiments across the various omics and is already implemented in metabolomic platforms like MetaboLights or PhenoMeNal [31-33].

Metabolic models

The metabolic phenotype is defined by two complementary omics, the metabolome, and the fluxome. The first offers a static view of metabolism (snapshot-like), whereas the latter represents the rate at which metabolites are interconverted through metabolic pathways and

therefore provides a dynamic view of the metabolic phenotype [34]. The fluxome emerges from complex interactions among metabolites, enzymes and transmembrane carriers. Thus, the fluxome cannot be directly measured, and instead needs to be inferred through the analysis of other omic measurements. One of the most informative techniques to determine the fluxome are Stable Isotope-Resolved Metabolomics (SIRM). In SIRM, a biological system is incubated with a substrate labeled with a stable heavy isotope (e.g., ^{13}C) that propagates to metabolites in the network generating characteristic label patterns which are indicative of the underlying flux distribution [35].

Metabolic models, mathematical representations of metabolism, are the tools used by systems biology and systems medicine to integrate multiple layers of data and predict metabolic fluxes. Nowadays, the vast availability of genomic data and the functional annotations allows the reconstruction of Genome-Scale Metabolic Models (GSMMs). GSMMs are built starting from genome annotations, which are used to identify enzyme-coding genes. These can then be mapped to reactions using biochemical databases, such as KEGG, BRENDA or MetaCyc. The resulting network is then curated to account for misannotations and missing reactions. Finally, the built reconstruction is validated by simulating the known metabolic functions of the target organism [36]. In 2007, the first human GSMMs were reconstructed [37, 38]. They formed the basis for much more in-depth human genome-scale reconstruction models including Human Metabolic Reaction, Recon 2 and Recon3D [39-41].

Metabolic simulations based on these genome-scale networks, or a subset of them, are usually performed with either kinetic or constraint-based modeling (CBM) techniques. Kinetic models integrate kinetic properties of enzymes (e.g., their affinity for substrates, the number of catalytic cycles that they can undergo per unit of time and their regulation by activators or inhibitors) and allow to simulate the dynamic behavior of fluxes and metabolites. However, they are limited by the complexity to build and parameterize kinetics models for large networks. In contrast, CBM uses network stoichiometry and the assumption of metabolic pseudo-steady state (i.e., intracellular metabolite concentrations are constant in time) to simulate steady-state flux distributions. While CBM is easily applied to large networks like GSMMs, it has a more limited capacity when it comes to studying the dynamic behavior of metabolic networks than kinetic models.

Building large scale kinetic models

Kinetic models are systems of Ordinary Differential Equations (ODEs) where metabolic fluxes are computed as a function of metabolite concentrations through a set of defined kinetic equations. Each metabolite has an ODE equation representing its variation in time, and each reaction has a kinetic equation describing the dependency of reaction fluxes to metabolite and enzyme concentrations. Metabolomic data, taken at multiple time points, is the primary input to validate kinetic models and iteratively fit unknown parameters of the kinetic equations (Figure 3) [42].

Figure 3: Kinetic model of upper glycolysis (Puigjaner et al., 1997 [43]). The network has four metabolites (Glc: Glucose, G6P: Glucose 6-phosphate, F6P: Fructose 6-phosphate) connected by three reactions (HK: Hexokinase, GPI: Glucose 6-phosphate isomerase, PFK: phosphofructokinase). HK has a Michaelis-Menten kinetic law with an uncompetitive inhibition by G6P, GPI a reversible Michaelis-Menten kinetic law and PFK a Hill cooperative kinetic law. Each kinetic law is parametrized from measurements of mice muscle extracts (V_{max}: maximal reaction rate, Km/Kms/Kmp/S_{halve}: concentration at which half of the V_{max} is achieved, Ki: Concentration at which half of the inhibition is achieved, h: hill cooperativity coefficient). From network stoichiometry, the parametrized kinetic laws are combined to build a system of ODEs, with each equation describing the dependent dynamic of a metabolite concentration. Starting with initial metabolomic values, solving the system of ODEs simulates time courses for metabolite concentrations and reaction fluxes, which can be compared with additional metabolomics data for validation.

There are two approaches to building large-scale kinetic models: the bottom-up or forward reconstruction and the top-down or inverse reconstruction. In the former method, the various subparts of the model are built individually and then put together to form the final model, whereas in the latter the entire model is reconstructed, and all the parameters are fitted at the same time. The major issues in large-scale kinetic model reconstruction are the many unknown parameters in the model and the lack of knowledge of regulatory information. Indeed, the greatest challenge to build large kinetic models is the parameter inference or fitting step. Over the last years, approaches like structural Kinetic Modelling (SKM) and Mass Action Stoichiometric Simulation (MASS) modeling have been developed to tackle this step.

SKM aims to quantitatively describe the dynamic performance of a system, rather than specifically define kinetic parameters, and constructs local linear approximations for each parameter according to experimental data and feasible biochemical states. Then, the reconstructed local linear models are used for the interrogation of a solution parameter space [44, 45]. On the other hand, MASS models try to combine constraint-based stochiometric reconstructions with matrix based kinetic modeling. More specifically, MASS uses large scale stochiometric network reconstructions as scaffolds, onto which fluxomic and metabolomic data measured *in vivo* are integrated and then kinetic parameters, explicit for the modeled steady-state of the system, are estimated. If simulations of growth conditions are performed, kinetic constants for the evolution of the system can be calculated, thus describing its dynamic behavior [46].

Constraint-Based modeling

CBM assumes a metabolic pseudo-steady state to build mass-balance constraints around metabolites and identify valid steady-state flux distributions. In this manner, the stoichiometry of the network can be represented as a system of linear equations and steady-state flux distributions can be simulated without the need of defining the kinetic equations for

each enzyme[37, 38, 47]. As the resulting system is usually underdetermined, additional constraints and optimizations need to be applied to reduce the solution space towards a unique solution (Figure 4) [39, 40].

For instance, GSMMs generally need to be constrained by integrating transcriptomics or proteomics data. This need arises because GSMMs define the entire metabolic potential for a given organism, whereas at any given cell and time point only a subset of enzymes are expressed and only a subset of reactions will be active. There are several approaches to integrate such data, but they are generally based on maximizing the consistency between the transcript and protein abundances of enzymes and the flux through reactions catalyzed by them. Integrating transcriptomics and proteomics allows to obtain maps of active/inactive reactions, as well as to characterize the changes in flux distributions between two or more different conditions or time points [48-53].

The range of feasible flux values can be further constrained by metabolomics data. Metabolomics from the extracellular media can be used to constrain extracellular fluxes (i.e., rates of uptake of secretion for extracellular metabolites). Concerning intracellular metabolomics, if a metabolite is detected, the model can be constrained to have at least one reaction active, where this metabolite is produced [51]. Furthermore, quantitative metabolomics of intracellular metabolites allows setting the rate at which intracellular metabolites must be synthesized to maintain a steady state in proliferating or growing systems [54]. Finally, SIRM-based metabolic flux analysis (MFA) can be applied to identify the range of flux values underlying a given set of SIRM measurements. The resulting flux ranges can be added to GSMM as flux bounds[35].

Even after integrating transcriptomics or proteomics and metabolomics, GSMMs are generally still undetermined. Flux Balance Analysis (FBA) aims to identify a unique optimal solution by maximizing or minimizing one or more fluxes in the metabolic network [55]. The choice of objective depends on the system under study, for instance, to study rapidly proliferating systems, like cancer cells, the synthesis of biomass is used as the objective, but other objectives can be set depending on the system of study [55-60].

Figure 4: Constraint-based modeling. First, the stoichiometry of the metabolic network is written as a stoichiometric matrix (*s*), where the $s_{m,r}$ element of the matrix is the stoichiometric coefficient of metabolite *m* in reaction *r*. From an infinite space of possible flux (*v*) solutions, a feasible solution space which contains possible steady-state solutions is obtained by applying the steady-state constraint ($s.v=0$) and defining the directionality of reactions. A condition-specific solution space can be obtained by integrating condition-specific omics like transcriptomics, proteomics or metabolomics. Finally, an optimization can be performed in the solution space to select the best solution(*s*). For instance, biomass production can be maximized so that the solution(s) that optimize growth efficiency can be selected.

Applications in systems medicine

The integration of multiple omics data in a systems medicine manner is an emerging field. Nevertheless, it has already provided new insights into the interplay among different regulatory layers.

For example, by studying the associations between SNPs with metabolomics measurements, it has been demonstrated that the variability of metabolite concentrations in the blood between individuals, is explained to a large extent by common genetic variants [61]. In another study, associations using Epigenome-Wide ASsociation data (EWAS) in combination with cytosine-guanine dinucleotide (CpG) methylation data and other multi-omics data suggested a causal effect of metabolite levels on methylation of obesity-associated CpG sites [62].

Furthermore, even if the reconstruction of large-scale kinetic models still poses a big challenge, several examples of kinetic models in systems medicine demonstrate their great potential. For instance, a kinetic model of human erythrocytes was used to identify metabolic targets that would selectively kill parasite *Trypanosoma brucei* with minimal collateral damage to human cells [63]. Berndt et al. reconstructed a kinetic model of the liver, and they used it to characterize the metabolic phenotype of hepatocytes and the metabolic reprogramming that they undergone during carcinogenesis [64]. Bordbar et al. (2015) have simulated individual responses to drug exposure including side effect incidence and demonstrated that enzyme activities and cellular dynamics, rather than metabolomics, are the most accurate representation of the genotype [65].

CBM has also been widely used in systems medicine to perform multi-omics data integration in the framework of GSMMs. For example, Mardinoglu et al. integrated proteomics and transcriptomics to build an adipocyte-specific GSMM and identified several putative therapeutic against obesity [66]. GSMMs have also been widely applied to identify genes or sets of genes that are essential for a disease related process [60, 67-69]. For instance, Folger et al. created a GSMM of cancer metabolism that predicted 52 cytostatic drug targets, 40% of which were targeted by known anticancer drugs [70]. Similarly, Agren et al. built 27 patient-specific GSMM of hepatocellular carcinoma and identified 101 potential drug targets, many of which had a strong correlation with disease progression [71]. GSMMs have also shown great potential in biomarker discovery for example in liver diseases and type two diabetes [72, 73].

Conclusions and future perspective

The primary goal of systems medicine is to explain, predict and prevent the progression of disease based on clinical, environmental and multi-omics data. Given the inherent network structure of metabolic processes, network modeling and the analysis of multi-omics data provide powerful and flexible inference tools to decipher the complex interactions in biological systems. However, consensus models built from samples from many individuals, albeit informative, might fail to capture the heterogeneity that is present in a population [74]. This

limits the elucidation of the molecular drivers for an individual-specific phenotype (either healthy or pathological), which result from the differential regulation or dysfunction of individual-specific networks.

Towards that end, methods are being proposed to build patient-specific networks that capture the subject's specificity of clinical manifestation with the goal of understanding diseases at the individual level and providing targeted and personalized treatments [75-78]. In principle, a personalized database could be generated for each individual, containing his/hers omics information (e.g., genomics, urine and blood metabolomics, gut microbiome), together with lifestyle data across time. This information, if properly analyzed, can provide the means to build patient-specific networks in order to identify the best diagnostic, therapeutic and prevention strategies for each individual and enable P4 (Predictive, Preventive, Personalized and Participatory) medicine [79, 80].

Acknowledgments

AR acknowledges financial support from C.I.R.M.M.P.

VAPMDS and ES acknowledge financial support from EU FP7 funded project INFECT (contract no. 305340; www.fp7infect.eu/)

MC acknowledges support from MINECO-European Commission FEDER funds – "Una manera de hacer Europa" (SAF2017-89673-R and SAF2015-70270-REDT), Instituto de Salud Carlos III and Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas (CIBEREHD CB17/04/00023), AGAUR – Generalitat de Catalunya (2017SGR-1033) and through the prize "ICREA Academia" for excellence in research, funded by ICREA foundation – Generalitat de Catalunya.

References

1. Ideker, T., T. Galitski, and L. Hood, *A NEW APPROACH TO DECODING LIFE: Systems Biology.* Annual Review of Genomics and Human Genetics, 2001. **2**(1): p. 343-372.
2. Kirschner, M.W., *The Meaning of Systems Biology.* Cell, 2005. **121**(4): p. 503-504.
3. Ahn, A.C., et al., *The limits of reductionism in medicine: could systems biology offer an alternative?* PLoS medicine, 2006. **3**(6): p. e208-e208.
4. Miles, A., *On a Medicine of the Whole Person: away from scientistic reductionism and towards the embrace of the complex in clinical practice*.* Journal of Evaluation in Clinical Practice, 2009. **15**(6): p. 941-949.
5. Ahn, A.C., et al., *The Limits of Reductionism in Medicine: Could Systems Biology Offer an Alternative?* PLOS Medicine, 2006. **3**(6): p. e208.
6. Auffray, C., Z. Chen, and L. Hood, *Systems medicine: the future of medical genomics and healthcare.* Genome Med, 2009. **1**(1): p. 2.
7. Hood, L., R. Balling, and C. Auffray, *Revolutionizing medicine in the 21st century through systems approaches.* Biotechnology Journal, 2012. **7**(8): p. 992-1001.
8. Tillmann, T., et al., *Systems Medicine 2.0: potential benefits of combining electronic health care records with systems science models.* Journal of medical Internet research, 2015. **17**(3): p. e64-e64.
9. Barabási, A.-L., N. Gulbahce, and J. Loscalzo, *Network medicine: a network-based approach to human disease.* Nature reviews. Genetics, 2011. **12**(1): p. 56-68.
10. Joyce, A.R. and B.Ø. Palsson, *The model organism as a system: Integrating 'omics' data sets.* Nature Reviews Molecular Cell Biology, 2006. **7**: p. 198.
11. Rosato, A., et al., *From correlation to causation: analysis of metabolomics data using systems biology approaches.* Metabolomics : Official journal of the Metabolomic Society, 2018. **14**(4): p. 37-37.
12. Ma'ayan, A., *Introduction to Network Analysis in Systems Biology.* Science signaling, 2011. **4**(190): p. tr5-tr5.
13. Barabási, A.-L., *Scale-Free Networks: A Decade and Beyond.* Science, 2009. **325**(5939): p. 412.
14. Vidal, M., M.E. Cusick, and A.-L. Barabási, *Interactome networks and human disease.* Cell, 2011. **144**(6): p. 986-998.
15. Emilsson, V., et al., *Genetics of gene expression and its effect on disease.* Nature, 2008. **452**: p. 423.
16. Petersen, A.-K., et al., *Epigenetics meets metabolomics: an epigenome-wide association study with blood serum metabolic traits.* Human molecular genetics, 2014. **23**(2): p. 534-545.
17. Ishii, N., et al., *Multiple High-Throughput Analyses Monitor the Response of 'E. coli' to Perturbations.* Science, 2007. **316**(5824): p. 593-597.
18. Liu, C., et al., *Arachidonic Acid Metabolism Pathway Is Not Only Dominant in Metabolic Modulation but Associated With Phenotypic Variation After Acute Hypoxia Exposure.* Frontiers in Physiology, 2018. **9**(236).
19. Clish, C.B., *Metabolomics: an emerging but powerful tool for precision medicine.* Cold Spring Harbor molecular case studies, 2015. **1**(1): p. a000588-a000588.

20. Trivedi, D.K., K.A. Hollywood, and R. Goodacre, *Metabolomics for the masses: The future of metabolomics in a personalized world.* New horizons in translational medicine, 2017. **3**(6): p. 294-305.
21. Dettmer, K., P.A. Aronov, and B.D. Hammock, *Mass spectrometry-based metabolomics.* Mass spectrometry reviews, 2007. **26**(1): p. 51-78.
22. Wishart, D.S., *Quantitative metabolomics using NMR.* TrAC Trends in Analytical Chemistry, 2008. **27**(3): p. 228-237.
23. Bingol, K., *Recent Advances in Targeted and Untargeted Metabolomics by NMR and MS/NMR Methods.* High-throughput, 2018. **7**(2): p. 9.
24. Roberts, L.D., et al., *Targeted Metabolomics.* Current Protocols in Molecular Biology, 2012. **98**(1): p. 30.2.1-30.2.24.
25. Alonso, A., S. Marsal, and A. Julià, *Analytical Methods in Untargeted Metabolomics: State of the Art in 2015.* Frontiers in Bioengineering and Biotechnology, 2015. **3**(23).
26. Kanehisa, M., et al., *KEGG: new perspectives on genomes, pathways, diseases and drugs.* Nucleic Acids Research, 2017. **45**(D1): p. D353-D361.
27. Placzek, S., et al., *BRENDA in 2017: new perspectives and new tools in BRENDA.* Nucleic Acids Research, 2017. **45**(D1): p. D380-D388.
28. Fabregat, A., et al., *The Reactome Pathway Knowledgebase.* Nucleic acids research, 2018. **46**(D1): p. D649-D655.
29. Caspi, R., et al., *The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases.* Nucleic acids research, 2016. **44**(D1): p. D471-D480.
30. Yugi, K., et al., *Reconstruction of Insulin Signal Flow from Phosphoproteome and Metabolome Data.* Cell Reports, 2014. **8**(4): p. 1171-1183.
31. Sansone, S.-A., et al., *The First RSBI (ISA-TAB) Workshop: "Can a Simple Format Work for Complex Studies?".* OMICS: A Journal of Integrative Biology, 2008. **12**(2): p. 143-149.
32. Kale, N.S., et al., *MetaboLights: An Open-Access Database Repository for Metabolomics Data.* Current Protocols in Bioinformatics, 2016. **53**(1): p. 14.13.1-14.13.18.
33. Peters, K., et al., *PhenoMeNal: Processing and analysis of Metabolomics data in the Cloud.* GigaScience, 2018: p. giy149-giy149.
34. Niedenführ, S., W. Wiechert, and K. Nöh, *How to measure metabolic fluxes: a taxonomic guide for 13C fluxomics.* Current Opinion in Biotechnology, 2015. **34**: p. 82-90.
35. Balcells, C., et al., *Tracing metabolic fluxes using mass spectrometry: stable isotope-resolved metabolomics in health and disease.* TrAC Trends in Analytical Chemistry, 2019.
36. Faria, J.P., et al., *Methods for automated genome-scale metabolic model reconstruction.* Biochemical Society Transactions, 2018: p. BST20170246.
37. Duarte, N.C., et al., *Global reconstruction of the human metabolic network based on genomic and bibliomic data.* Proceedings of the National Academy of Sciences, 2007. **104**(6): p. 1777-1782.
38. Ma, H., et al., *The Edinburgh human metabolic network reconstruction and its functional analysis.* Molecular Systems Biology, 2007. **3**(1): p. 135.
39. Thiele, I., et al., *A community-driven global reconstruction of human metabolism.* Nature biotechnology, 2013. **31**(5): p. 419-425.
40. Pornputtapong, N., I. Nookaew, and J. Nielsen, *Human metabolic atlas: an online resource for human metabolism.* Database : the journal of biological databases and curation, 2015. **2015**: p. bav068-bav068.
41. Brunk, E., et al., *Recon3D enables a three-dimensional view of gene variation in human metabolism.* Nature Biotechnology, 2018. **36**: p. 272.

42. Foguet, C., et al., *HepatoDyn: A Dynamic Model of Hepatocyte Metabolism That Integrates 13C Isotopomer Data.* PLOS Computational Biology, 2016. **12**(4): p. e1004899.
43. Puigjaner, J., et al., *Comparison of control analysis data using different approaches: modelling and experiments with muscle extract.* FEBS Letters, 1997. **418**(1): p. 47-52.
44. Steuer, R., et al., *Structural kinetic modeling of metabolic networks.* Proceedings of the National Academy of Sciences, 2006. **103**(32): p. 11868.
45. Grimbs, S., et al., *The stability and robustness of metabolic states: identifying stabilizing sites in metabolic networks.* Molecular Systems Biology, 2007. **3**(1).
46. Jamshidi, N. and B.Ø. Palsson, *Mass action stoichiometric simulation models: incorporating kinetics and regulation into stoichiometric models.* Biophysical journal, 2010. **98**(2): p. 175-185.
47. Nilsson, A. and J. Nielsen, *Genome scale metabolic modeling of cancer.* Metabolic Engineering, 2017. **43**: p. 103-112.
48. Zur, H., E. Ruppin, and T. Shlomi, *iMAT: an integrative metabolic analysis tool.* Bioinformatics, 2010. **26**(24): p. 3140-3142.
49. Jensen, P.A. and J.A. Papin, *Functional integration of a metabolic network model and expression data without arbitrary thresholding.* Bioinformatics (Oxford, England), 2011. **27**(4): p. 541-547.
50. Agren, R., et al., *Reconstruction of Genome-Scale Active Metabolic Networks for 69 Human Cell Types and 16 Cancer Types Using INIT.* PLOS Computational Biology, 2012. **8**(5): p. e1002518.
51. Schmidt, B.J., et al., *GIM3E: condition-specific models of cellular metabolism developed from metabolomics and expression data.* Bioinformatics (Oxford, England), 2013. **29**(22): p. 2900-2908.
52. Yizhak, K., et al., *Model-based identification of drug targets that revert disrupted metabolism and its application to ageing.* Nature Communications, 2013. **4**: p. 2632.
53. Galhardo, M., et al., *Integrated analysis of transcript-level regulation of metabolism reveals disease-relevant nodes of the human metabolic network.* Nucleic acids research, 2014. **42**(3): p. 1474-1496.
54. Reimers, A.-M. and A.C. Reimers, *The steady-state assumption in oscillating and growing systems.* Journal of Theoretical Biology, 2016. **406**: p. 176-186.
55. Orth, J.D., I. Thiele, and B.Ø. Palsson, *What is flux balance analysis?* Nature biotechnology, 2010. **28**(3): p. 245-248.
56. Ramakrishna, R., et al., *Flux-balance analysis of mitochondrial energy metabolism: consequences of systemic stoichiometric constraints.* American Journal of Physiology-Regulatory, Integrative and Comparative Physiology, 2001. **280**(3): p. R695-R704.
57. Gille, C., et al., *HepatoNet1: a comprehensive metabolic reconstruction of the human hepatocyte for the analysis of liver physiology.* Molecular systems biology, 2010. **6**: p. 411-411.
58. Bordbar, A., et al., *Constraint-based models predict metabolic and associated cellular functions.* Nature Reviews Genetics, 2014. **15**: p. 107.
59. Rienksma, R.A., et al., *Modeling the Metabolic State of Mycobacterium tuberculosis Upon Infection.* Frontiers in Cellular and Infection Microbiology, 2018. **8**(264).
60. Wang, T., et al., *Gene Essentiality Profiling Reveals Gene Networks and Synthetic Lethal Interactions with Oncogenic Ras.* Cell, 2017. **168**(5): p. 890-903.e15.
61. Suhre, K., et al., *Human metabolic individuality in biomedical and pharmaceutical research.* Nature, 2011. **477**(7362): p. 54-60.

62. Zaghlool, S.B., et al., *Deep molecular phenotypes link complex disorders and physiological insult to CpG methylation.* Human molecular genetics, 2018. **27**(6): p. 1106-1121.
63. Haanstra, J.R., et al., *Targeting pathogen metabolism without collateral damage to the host.* Scientific Reports, 2017. **7**: p. 40406.
64. Berndt, N., et al., *HEPATOKIN1 is a biochemistry-based model of liver metabolism for applications in medicine and pharmacology.* Nature Communications, 2018. **9**(1): p. 2386.
65. Bordbar, A., et al., *Personalized Whole-Cell Kinetic Models of Metabolism for Discovery in Genomics and Pharmacodynamics.* Cell Systems, 2015. **1**(4): p. 283-292.
66. Mardinoglu, A., et al., *Integration of clinical data with a genome-scale metabolic model of the human adipocyte.* Molecular systems biology, 2013. **9**: p. 649-649.
67. von Kamp, A. and S. Klamt, *Enumeration of Smallest Intervention Strategies in Genome-Scale Metabolic Networks.* PLOS Computational Biology, 2014. **10**(1): p. e1003378.
68. Pratapa, A., S. Balachandran, and K. Raman, *Fast-SL: an efficient algorithm to identify synthetic lethal sets in metabolic networks.* Bioinformatics, 2015. **31**(20): p. 3299-3305.
69. Zhan, T. and M. Boutros, *Towards a compendium of essential genes – From model organisms to synthetic lethality in cancer cells.* Critical Reviews in Biochemistry and Molecular Biology, 2016. **51**(2): p. 74-85.
70. Folger, O., et al., *Predicting selective drug targets in cancer through metabolic networks.* Molecular systems biology, 2011. **7**: p. 501-501.
71. Agren, R., et al., *Identification of anticancer drugs for hepatocellular carcinoma through personalized genome-scale metabolic modeling.* Molecular systems biology, 2014. **10**(3): p. 721-721.
72. Mardinoglu, A., et al., *Genome-scale metabolic modelling of hepatocytes reveals serine deficiency in patients with non-alcoholic fatty liver disease.* Nature Communications, 2014. **5**: p. 3083.
73. Väremo, L., et al., *Proteome- and Transcriptome-Driven Reconstruction of the Human Myocyte Metabolic Network and Its Use for Identification of Markers for Diabetes.* Cell Reports, 2015. **11**(6): p. 921-933.
74. Kuiper, J.S., et al., *Social relationships and risk of dementia: A systematic review and meta-analysis of longitudinal cohort studies.* Ageing Research Reviews, 2015. **22**: p. 39-57.
75. Hamburg, M.A. and F.S. Collins, *The Path to Personalized Medicine.* New England Journal of Medicine, 2010. **363**(4): p. 301-304.
76. Liu, X., et al., *Identifying disease genes and module biomarkers by differential interactions.* Journal of the American Medical Informatics Association : JAMIA, 2012. **19**(2): p. 241-248.
77. Liu, R., et al., *Early Diagnosis of Complex Diseases by Molecular Biomarkers, Network Biomarkers, and Dynamical Network Biomarkers.* Medicinal Research Reviews, 2014. **34**(3): p. 455-478.
78. Zhang, W., et al., *Diagnosing phenotypes of single-sample individuals by edge biomarkers.* Journal of Molecular Cell Biology, 2015. **7**(3): p. 231-241.
79. Chen, R. and M. Snyder, *Systems biology: personalized medicine for the future?* Current Opinion in Pharmacology, 2012. **12**(5): p. 623-628.
80. Flores, M., et al., *P4 medicine: how systems medicine will transform the healthcare sector and society.* Personalized Medicine, 2013. **10**(6): p. 565-576.

Annotations

(*) Zaghlool, S. B., Mook-Kanamori, D. O., Kader, S., Stephan, N., Halama, A., Engelke, R., Sarwath, H., Al-Dous, E. K., Mohamoud, Y. A., Roemisch-Margl, W., Adamski, J., Kastenmüller, G., Friedrich, N., Visconti, A., Tsai, P. C., Spector, T., Bell, J. T., Falchi, M., Wahl, A., Waldenberger, M., Peters, A., Gieger, C., Pezer, M., Lauc, G., Graumann, J., Malek, J. A., … Suhre, K. (2018). Deep molecular phenotypes link complex disorders and physiological insult to CpG methylation. Human molecular genetics, 27(6), 1106-1121.

Zaghool et al., combined Epigenome-Wide ASsociation data (EWAS) with cytosine-guanine dinucleotide (CpG) methylation and other multi-omics datasets and revealed a causal effect of metabolite levels on methylation of obesity-associated CpG sites.

(*)Haanstra, J.R., et al., *Targeting pathogen metabolism without collateral damage to the host.* Scientific Reports, 2017. **7**: p. 40406.

This study highlights how kinetic modelling can assist drug design by identifying targets against pathogen metabolism with have minimal side effects on the host.

(**) Bordbar, Aarash & Taylor, Douglas & Zielinski, Daniel & Sonnenschein, Nikolaus & Jamshidi, Neema & Palsson, Bernhard O.. (2015). Personalized Whole-Cell Kinetic Models of Metabolism for Discovery in Genomics and Pharmacodynamics. Cell Systems. 1. 283-292. 10.1016/j.cels.2015.10.003.

Bordbar et al., constructed personalized whole cell kinetic models of erythrocyte and showed that personalized kinetic rate constants are the best representation of the genotype. They were also able to identify individuals at risk for a drug side effect.

(*) Berndt, N., Bulik, S., Wallach, I., Wünsch, T., König, M., Stockmann, M., Meierhofer, D., … Holzhütter, H. G. (2018). HEPATOKIN1 is a biochemistry-based model of liver metabolism for applications in medicine and pharmacology. *Nature communications*, *9*(1), 2386. doi:10.1038/s41467-018-04720-9

In this study, the authors build a large-scale kinetic model of hepatocyte metabolism and integrated proteomics to highlight the metabolic differences in the metabolic phenotype hepatocytes compared to adenoma and hepatocellular carcinoma cells.

(**) Nilsson, Avlant & Nielsen, Jens. (2016). Genome Scale Metabolic Modeling of Cancer. Metabolic Engineering. 43. 10.1016/j.ymben.2016.10.022.

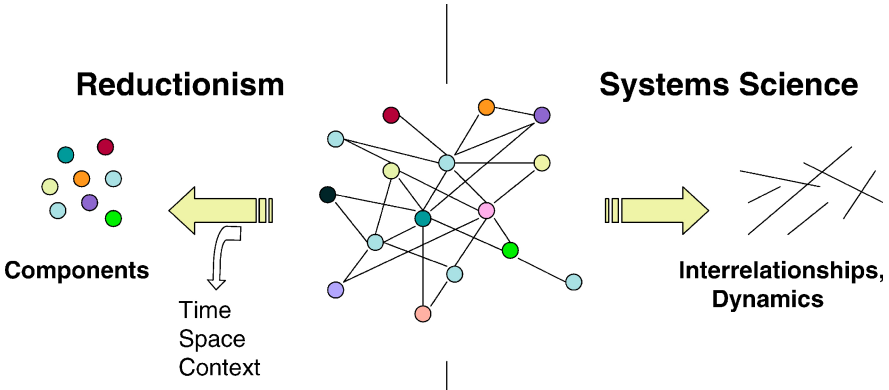
A comprehensive review of GSMMs focused on their application to the study of Cancer metabolism.

* Ref. 13, entitled "Epigenetics meets metabolomics: an epigenome-wide association study with blood serum metabolic traits."

This is the first epigenome-wide association study (EWAS) between DNA methylation and metabolic types in human blood, showing that that DNA methylation plays an important role in regulating human metabolism.

** Ref. 14, entitled "Multiple high-throughput analyses monitor the response of E. coli to perturbations."

This study integrated multiple omics measurements on Escherichia coli. It showed that in most cases the disruption of genes encoding metabolic enzymes led to surprisingly small changes in messenger RNA and proteins and did not sizably alter metabolite levels. This is due to the rerouting of metabolic fluxes. In contrast, E. coli actively controlled enzyme levels to maintain a stable metabolic state in response to changes in growth rate.

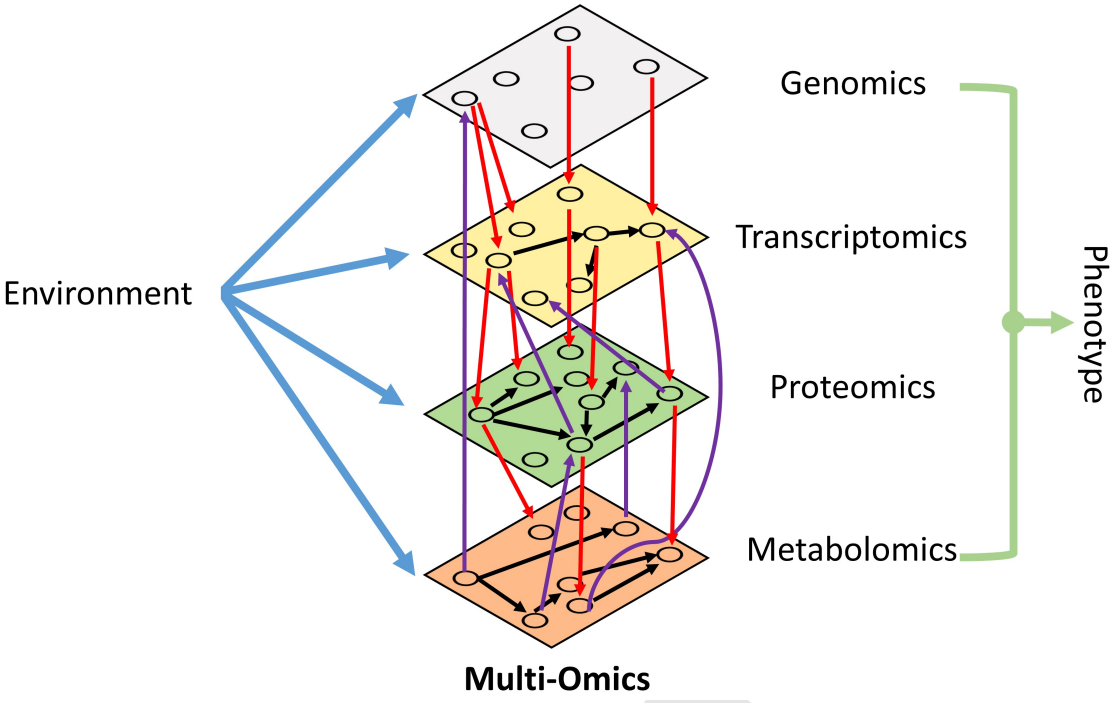


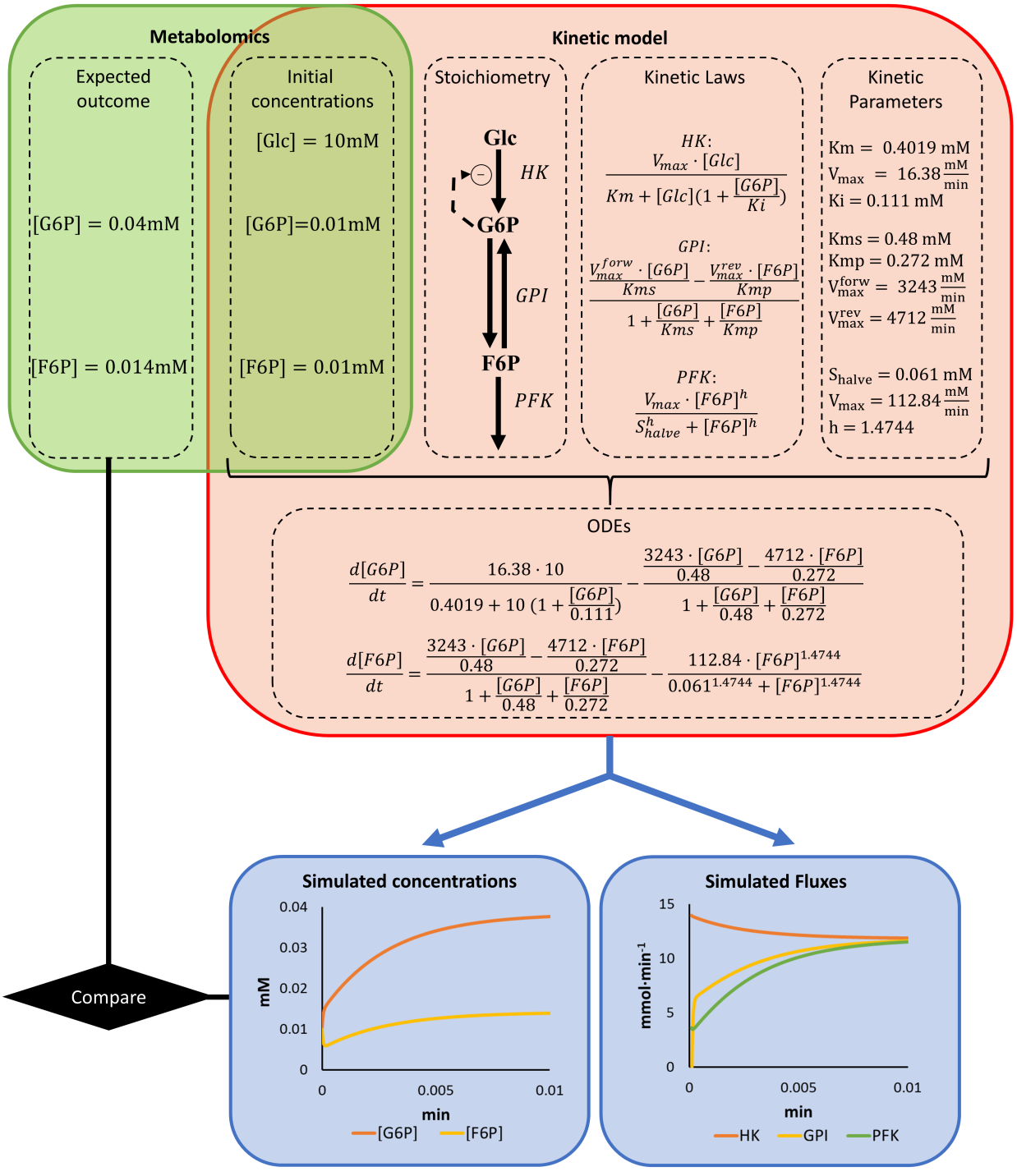
Medical Treatments

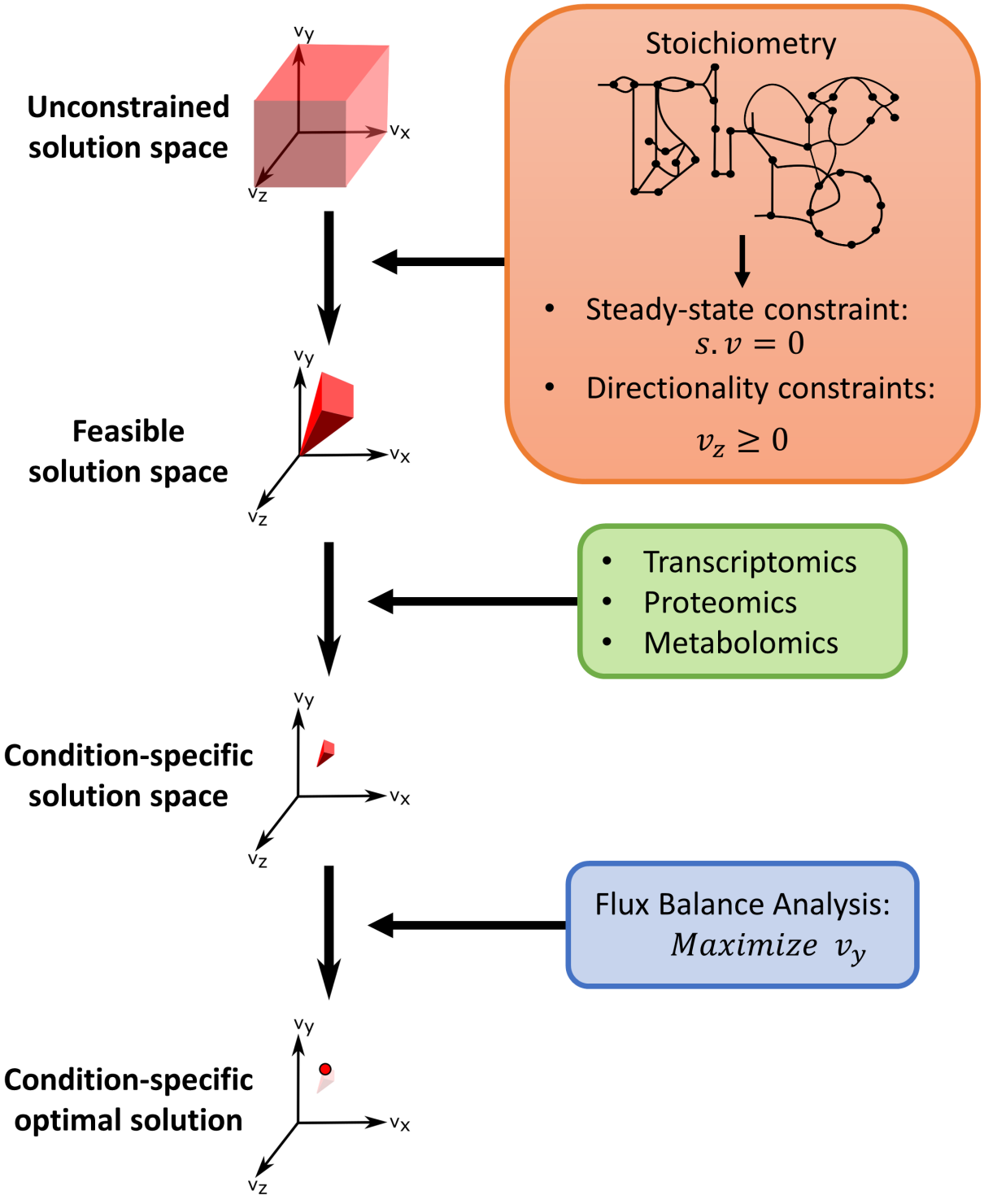
Disease-driven
Aimed for normalcy (normal range)
Additive

Medical Treatments

Individualized
Multidimensional use of drugs
Time-sensitive
Space-sensitive
Synergistic




Metabolomics
Expected outcome
[G6P] = 0.04mM
[F6P] = 0.014mM
Initial concentrations
[Glc] = 10mM
[G6P]=0.01mM
[F6P] = 0.01mM
Kinetic model
Stoichiometry
Glc
HK
G6P
GPI
F6P
PFK
Kinetic Laws
HK:
$\frac{V_{max} \cdot [Glc]}{Km + [Glc](1 + \frac{[G6P]}{Ki})}$
GPI:
$\frac{\frac{V_{max}^{forw} \cdot [G6P]}{Kms} - \frac{V_{max}^{rev} \cdot [F6P]}{Kmp}}{1 + \frac{[G6P]}{Kms} + \frac{[F6P]}{Kmp}}$
PFK:
$\frac{V_{max} \cdot [F6P]^h}{S_{halve}^h + [F6P]^h}$
Kinetic Parameters
Km = 0.4019 mM
V_{max} = 16.38 mM/min
Ki = 0.111 mM
Kms = 0.48 mM
Kmp = 0.272 mM
V_{max}^{forw} = 3243 mM/min
V_{max}^{rev} = 4712 mM/min
S_{halve} = 0.061 mM
V_{max} = 112.84 mM/min
h = 1.4744
ODEs
$\frac{d[G6P]}{dt} = \frac{16.38 \cdot 10}{0.4019 + 10\,(1 + \frac{[G6P]}{0.111})} - \frac{\frac{3243 \cdot [G6P]}{0.48} - \frac{4712 \cdot [F6P]}{0.272}}{1 + \frac{[G6P]}{0.48} + \frac{[F6P]}{0.272}}$
$\frac{d[F6P]}{dt} = \frac{\frac{3243 \cdot [G6P]}{0.48} - \frac{4712 \cdot [F6P]}{0.272}}{1 + \frac{[G6P]}{0.48} + \frac{[F6P]}{0.272}} - \frac{112.84 \cdot [F6P]^{1.4744}}{0.061^{1.4744} + [F6P]^{1.4744}}$
Compare
Simulated concentrations
mM
0.04
0.03
0.02
0.01
0
0
0.005
0.01
min
[G6P]
[F6P]
Simulated Fluxes
mmol·min-1
15
10
5
0
0
0.005
0.01
min
HK
GPI
PFK


Unconstrained solution space
Stoichiometry
• Steady-state constraint: $s.v = 0$
• Directionality constraints: $v_z \geq 0$
Feasible solution space
• Transcriptomics
• Proteomics
• Metabolomics
Condition-specific solution space
Flux Balance Analysis: $Maximize\ v_y$
Condition-specific optimal solution
vy
vx
vz