

RESEARCH ARTICLE

Δ FBA—Predicting metabolic flux alterations using genome-scale metabolic models and differential transcriptomic data

Sudharshan Ravi ^{1,2}, Rudiyanto Gunawan ^{1*}**1** Department of Chemical and Biological Engineering, University at Buffalo-SUNY, Buffalo, New York, United States of America, **2** Institute for Chemical and Bioengineering, ETH Zurich, Zurich, Switzerland* rgunawan@buffalo.edu**OPEN ACCESS**

Citation: Ravi S, Gunawan R (2021) Δ FBA—Predicting metabolic flux alterations using genome-scale metabolic models and differential transcriptomic data. *PLoS Comput Biol* 17(11): e1009589. <https://doi.org/10.1371/journal.pcbi.1009589>

Editor: Joerg Stelling, ETH Zurich: Eidgenossische Technische Hochschule Zurich, SWITZERLAND

Received: February 5, 2021

Accepted: October 25, 2021

Published: November 10, 2021

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pcbi.1009589>

Copyright: © 2021 Ravi, Gunawan. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data and codes are available at <https://github.com/CABSEL/DeltaFBA>.

Abstract

Genome-scale metabolic models (GEMs) provide a powerful framework for simulating the entire set of biochemical reactions in a cell using a constraint-based modeling strategy called flux balance analysis (FBA). FBA relies on an assumed metabolic objective for generating metabolic fluxes using GEMs. But, the most appropriate metabolic objective is not always obvious for a given condition and is likely context-specific, which often complicates the estimation of metabolic flux alterations between conditions. Here, we propose a new method, called Δ FBA (deltaFBA), that integrates differential gene expression data to evaluate directly metabolic flux differences between two conditions. Notably, Δ FBA does not require specifying the cellular objective. Rather, Δ FBA seeks to maximize the consistency and minimize inconsistency between the predicted flux differences and differential gene expression. We showcased the performance of Δ FBA through several case studies involving the prediction of metabolic alterations caused by genetic and environmental perturbations in *Escherichia coli* and caused by Type-2 diabetes in human muscle. Importantly, in comparison to existing methods, Δ FBA gives a more accurate prediction of flux differences.

Author summary

Metabolic alterations are often used as hallmarks of observable phenotypes. In this regard, reconstructed genome-scale metabolic models (GEMs) provide a rich and computable representation of the entire set of biochemical reactions in a cell. However, the performance of analytical tools for predicting metabolic reaction rates or fluxes using GEMs is sensitive to the assumed metabolic objective that is often unknown and likely context-specific. Here, we propose a novel method called Δ FBA that combines differential gene expression data and GEMs to evaluate differences in the metabolic fluxes between two conditions (perturbation vs. control) without the need for specifying a metabolic objective. In our demonstration, Δ FBA outperformed other existing methods in predicting metabolic flux alterations.

Funding: SR and RG received funding from Swiss National Science Foundation (grant number 163390 and 176279; www.snsf.ch). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

This is a *PLOS Computational Biology* Methods paper.

Introduction

In the post-genomic era, there has been intense efforts directed toward the reconstruction of genome-scale models of cellular networks. An important portion of these efforts focuses on metabolic networks due to the significance of cellular metabolism for understanding diseases such as cancer [1–4] as well as for metabolic engineering applications in biomanufacturing [5]. Recent advances in high-throughput sequencing technologies, gene functional annotation, and metabolic pathway databases, and developments of algorithms for mapping gene-protein-reaction (GPR) associations and identifying missing metabolic reactions systematically (gap-filling), have enabled the reconstruction of thousands of genome-scale metabolic models (GEMs), from single cell organisms to human [6,7]. A GEM provides GPR associations that encompass the set of metabolites and metabolic reactions in an organism as prescribed by its genome. Concurrent with these developments is the creation of efficacious algorithms that use GEMs to predict intracellular metabolic fluxes—the rates of metabolic reactions—and how these fluxes vary under different environmental, genetic, and disease conditions [8–10].

A prominent class of algorithms based on a constrained-based modeling technique called flux balance analysis (FBA) have flourished due to its ease of formulation and flexibility. FBA uses the stoichiometric coefficients of the metabolic reactions in a GEM, an assumed cellular objective such as maximization of biomass production, and experimental data on metabolic capabilities and constraints of the cells, to predict metabolic fluxes [11]. Although FBA is effective in handling large networks and predicting cell behavior in many metabolic engineering studies [12–15], considerable uncertainty still remains about the appropriate choice of cellular objective for different conditions and cell types, a choice that typically requires expert knowledge of the cells and their phenotype in a given condition. Such an issue is particularly prominent for complex organisms such as human. Moreover, multiple equivalent flux solutions exist that give the same cellular objective value [16]. Not to mention, the standard FBA often produces biologically unrealistic flux solutions [17,18].

Driven by the increasing ease and availability of whole-genome omics profiling data, a multitude of FBA-based algorithms have been developed to incorporate omics datasets to create context-specific metabolic networks and to improve flux prediction accuracy [19–26]. Several of these methods, such as GIMME (Gene Inactivity Moderated by Metabolism and Expression) [20], iMAT (integrative Metabolic Analysis Tools) [21], and MADE (Metabolic Adjustment by Differential Expression) [22], are based on maximizing the consistency between the predicted flux distribution and the mRNA transcript abundance of metabolic genes, where the higher the transcript level of an enzyme, the larger should the flux of the corresponding reactions. Recent methods use data of mRNA transcript abundance for setting the bounds on reaction fluxes, e.g. E-Flux [23], or in the biological objective function, e.g. Lee et al. [24] and RELATCH (RELATive Change) [25]. Meanwhile, others like GX-FBA [26] determine fluxes in a perturbed state using differential gene expression and FBA flux prediction for the control (reference) state. Interestingly, a systematic evaluation of different FBA methods that incorporate gene expression data revealed a surprisingly poorer performance of these methods when compared to FBA with growth maximization and parsimony criteria, referred to parsimonious FBA (pFBA) [27]. More recently, ME-model [28] and GECKO [29] combine FBA with an explicit modeling of enzyme/protein expression and thus, are able to directly account for protein abundance. Thermodynamics constraints have also been integrated with the FBA to eliminate thermodynamically infeasible fluxes, and at the same time enable the integration of

metabolite concentration data, as done in recent methods such as ETFL [30]. All of the aforementioned methods, however, revolve around using omics data to predict metabolic fluxes for a given condition. But, many a times we are interested in the metabolic alterations caused by a perturbation or a change in intra/extracellular conditions.

Thus far, only a handful of methods focus on using differential expression data between two conditions (e.g., perturbation vs. control) to predict metabolic alterations directly, which is a particular focus of our study. The method Relative Expression and Metabolic Integration (REMI) [31] used differential expression of transcriptome and metabolome to estimate metabolic flux profiles in *Escherichia coli* under varying dilution and genetic perturbations. The method relies on maximizing the agreement between the fold-changes of metabolic fluxes and the fold-changes of enzyme expressions between two conditions. The metabolome data, if available, are used to determine the flux directionality using reaction thermodynamics. Among the alternative flux solutions, the L1-norm minimal solution is adopted to give a representative flux distribution. Another method by Zhu *et al.* [32] employed a softer definition when assessing consistency between the metabolic fluxes and enzyme differential expressions, where only the sign of the differences needs to agree. The method provides a qualitative determination of metabolic flux changes by determining the maximum and minimum flux through each reaction in the GEM. Both of the above methods generate metabolic flux predictions for each of the conditions in comparison. Also, like the standard FBA, both methods require an assumption on the cell's metabolic objective. Generally, model prediction inaccuracy is amplified when evaluating the differences between two model predictions. Another related method MOOMIN [33] uses a Bayesian approach to integrate differential gene expression profiles with GEMs to predict the qualitative change in the metabolic fluxes—increased, decreased or no change.

In this work, we developed Δ FBA (deltaFBA) for predicting the metabolic flux difference given a GEM and differential transcriptomic data between two conditions. Δ FBA relies on a constrained-based model that governs the balance of flux difference in the GEM, while maximizing the consistency and minimizing inconsistency between the flux alterations and the gene expression changes. Δ FBA is developed as a MATLAB package that works seamlessly with the COntstraint-Based Reconstruction and Analysis (COBRA) toolbox [34]. We applied the Δ FBA to analyze the metabolic changes of *Escherichia coli* in response to environmental and genetic perturbations using data from the studies of Ishii *et al.* [35] and Gerosa *et al.* [36]. We compared the performance of Δ FBA in evaluating flux differences between conditions to that of REMI and eight FBA methods, including parsimonious FBA (pFBA) [19], GIMME [20], iMAT [21], MADE [22], E-Flux [23], Lee *et al.* [24], RELATCH [25], and GX-FBA [26]. We also demonstrated the application of Δ FBA to a human GEM, specifically evaluating the metabolic alterations associated with type-2-diabetes in skeletal muscle using myocyte-specific GEM [37].

Materials and methods

Method formulation

Δ FBA generates a prediction for metabolic flux differences between a pair of conditions, for example, treated vs. untreated or mutant vs. wild-type strains. In the following, we use the superscript C to denote the control (reference) condition and P to denote the perturbed condition. In the standard FBA, writing mass balance around every metabolite and applying the steady state assumption give a linear equation $\mathbf{S}\mathbf{v} = \mathbf{0}$, where $\mathbf{S} \in \mathbb{R}^{m \times n}$ denotes the stoichiometric matrix for m metabolites that are involved in n metabolic reactions and transports in the GEM and $\mathbf{v} \in \mathbb{R}^n$ denotes the vector of n fluxes (rates). In Δ FBA, the steady state flux balance

is assumed for each condition, and consequently, the flux difference $\Delta \mathbf{v} = (\mathbf{v}^P - \mathbf{v}^C)$ satisfies the following balance equation:

$$\mathbf{S}\Delta \mathbf{v} = \mathbf{S}(\mathbf{v}^P - \mathbf{v}^C) = \mathbf{S}\mathbf{v}^P - \mathbf{S}\mathbf{v}^C = \mathbf{0}$$

where $\mathbf{v}^C \in \mathbb{R}^n$ and $\mathbf{v}^P \in \mathbb{R}^n$ denote the vectors of metabolic fluxes in C and P, respectively, and $\Delta \mathbf{v} \in \mathbb{R}^n$ denotes the vector of metabolic flux differences. The prediction of $\Delta \mathbf{v}$ is based on maximizing the consistency while minimizing the inconsistency between the flux changes $\Delta \mathbf{v}$ and the differential reaction expressions, constrained by among other things, the flux balance equation above. The following constrained mixed integer linear programming (MILP) gives the main formulation for ΔFBA:

$$\max_{z^U, z^D} \Phi = \max_{z^U, z^D} \sum_{i \in R^U} w_i^U (z_i^U - z_i^D) + \sum_{j \in R^D} w_j^D (z_j^D - z_j^U) \tag{1}$$

subject to:

$$\mathbf{S}\Delta \mathbf{v} = \mathbf{0} \tag{2}$$

$$\Delta \mathbf{v}_{min} \leq \Delta \mathbf{v} \leq \Delta \mathbf{v}_{max} \tag{3}$$

$$\Delta \mathbf{v} - M\mathbf{z}^U \geq \boldsymbol{\mu} - M\mathbf{1} \tag{4}$$

$$\Delta \mathbf{v} - M\mathbf{z}^U \leq \boldsymbol{\mu} \tag{5}$$

$$\Delta \mathbf{v} + M\mathbf{z}^D \leq -\boldsymbol{\eta} + M\mathbf{1} \tag{6}$$

$$\Delta \mathbf{v} + M\mathbf{z}^D \geq -\boldsymbol{\eta} \tag{7}$$

$$\Delta \mathbf{v} + M\mathbf{z}^0 \leq M\mathbf{1} \tag{8}$$

$$\Delta \mathbf{v} - M\mathbf{z}^0 \geq -M\mathbf{1} \tag{9}$$

$$z_k^0 + z_k^U \geq +z_{k'}^D \tag{10}$$

$$z_k^0 + z_k^D \geq +z_{k'}^U \tag{11}$$

Eqs (2) and (3) ensure that the flux difference $\Delta \mathbf{v}$ satisfy the flux balance equation while staying within acceptable lower and upper bounds. The constrained MILP produces the optimal binary vectors $\mathbf{z}^U \in \mathbb{Z}^n$ and $\mathbf{z}^D \in \mathbb{Z}^n$ that maximize the consistency and minimize the inconsistency between the flux differences and the differential gene expressions. When $z_i^U = 1$, Δv_i takes a positive value beyond the threshold μ_i , as specified by the constraints in Eqs (4) and (5). When $z_i^D = 1$, Δv_i takes a negative value beyond a threshold η_i , as specified by Eqs (6) and (7). Clearly, z_i^U and z_i^D cannot simultaneously be equal to 1. Meanwhile, the binary variable $\mathbf{z}^0 \in \mathbb{Z}^n$ is used to force certain user-selected reactions, if any, to have zero flux change value, as specified by Eqs (8) and (9). Note that all reversible reactions in the GEM are written as two separate irreversible reactions, whose indices are denoted by k and k' , the former for the forward and the latter for the backward direction. For all half pairs of reversible reactions, Eqs (10) and (11) ensure that the forward and reverse reactions are prevented to simultaneously have non-zero values, which is done to reduce degeneracy of the flux change solution $\Delta \mathbf{v}$.

Finally, the constant M in Eqs (4)–(9) should be set to a large value (default = 10^5), following the Big M method in linear programming [38].

The set of upregulated reactions \mathbf{R}^U and downregulated reactions \mathbf{R}^D are user-defined inputs. More specifically, the sets \mathbf{R}^U and \mathbf{R}^D include indices of reactions with significant increase and decrease in gene expression between the perturbed condition and the control, respectively. The non-negative weighting coefficients $w_i^U \in \mathbb{R}$ and $w_j^D \in \mathbb{R}$ (default value = 1) in the objective function allow users to prioritize certain reactions for consistency among those in the sets \mathbf{R}^U and \mathbf{R}^D , respectively. For example, the reaction corresponding to a gene deletion should be assigned a high w_j^D to force the corresponding flux change to be negative. The upper and lower bounds for the flux differences in Eq (3) are also user-defined parameters that can be set based on experimental data (e.g., the difference of experimentally determined biomass production or growth rates) or based on the flux bounds from each condition. For the latter, given the lower and upper flux bounds for the i -th flux in the perturbed ($v_{min,i}^p \in \mathbb{R}$ and $v_{max,i}^p \in \mathbb{R}$, respectively) and the control condition ($v_{min,i}^c \in \mathbb{R}$ and $v_{max,i}^c \in \mathbb{R}$, respectively), the bounds for the flux difference can be set as follows:

$$\Delta v_{min,i} = (v_{min,i}^p - v_{max,i}^c) \tag{12}$$

$$\Delta v_{max,i} = (v_{max,i}^p - v_{min,i}^c) \tag{13}$$

Finally, the thresholds μ_i and η_i for the positive and negative flux differences, respectively, are user-defined parameters. In the case studies, we used the same constant threshold value ϵ (default = 0.1% of the largest flux bound magnitude in the two conditions). These thresholds serve as a lower (upper) bound for which a positive (negative) flux difference is deemed to be upregulated (downregulated).

Given the degrees of freedom in GEMs for Δv , many equivalent optimal solutions often exist that give the same objective function value Φ^* as specified in Eq (1). By assuming parsimony for Δv —that is, Δv is minimal between the perturbed and control condition—a two-step optimization procedure is implemented in ΔFBA. The first step is to maximize consistency with gene expression changes as prescribed in Eqs (1)–(11) to determine the maximum objective function value, denoted by Φ^* . The second step is to produce an L2 norm minimal solution for Δv , as follows:

$$\min_{\Delta v, z^U, z^D} \|\Delta v(z^U, z^D)\|_2^2 \tag{14}$$

subject to the same constraints in Eqs (2)–(11) while achieving the same level of consistency Φ^* , implemented by the following additional constraint:

$$\sum_{i \in \mathbf{R}^U} w_i^U (z_i^U - z_i^D) + \sum_{j \in \mathbf{R}^D} w_j^D (z_j^D - z_j^U) = \Phi^* \tag{15}$$

The L2 minimization is based on the premise that the flux differences should be small between the conditions, which is similar to the method called Minimization of Metabolic Adjustment (MOMA) [39]. An alternative to L2-norm minimization is L1-norm minimization, which is analogous to maximizing sparsity of Δv . The L1-norm minimization was previously used in parsimonious FBA (pFBA) method [19], but such an approach often still leads to multiple degenerate solutions. On the other hand, the L2-norm minimization will produce a unique solution. However, the mixed integer quadratic optimization that is required to find the minimum L2-norm solution may have high computational requirement.

Δ FBA is available as MATLAB scripts and are compatible with the COBRA toolbox [34]. Δ FBA requires Gurobi optimizer (<http://www.gurobi.com>) as a pre-requisite. Δ FBA has been tested on a Windows PC using a 6-core Intel Xeon (2146G) Processor with 16 GB RAM.

Gene-protein-reaction mapping

For mapping fold-change gene expression data to fold-change reaction expression, we utilized the gene-protein-reaction (GPR) associations that are built into the GEM. These associations do not follow a one-to-one relationship since metabolic enzymes include isozymes (multiple enzymes mapping to the same reaction), promiscuous enzymes (a single enzyme participating in multiple reactions), and enzyme complexes (multiple genes required for an enzyme). Here, we used the Min/Max GPR rule [21,40,41]. The fold-change gene expression is first mapped to fold-change protein/enzyme expression. When multiple genes are required to form an enzyme complex, the fold-change enzyme expression is set to the minimum fold-change expression of the participating genes. Otherwise, the fold-change enzyme expression is equal to the fold-change gene expression. Then, the fold-change enzyme expression is mapped to the fold-change reaction expression. Here, when isozymes are involved in a reaction, the fold-change reaction expression is set to the maximum fold-change expression of the isoenzymes. Otherwise, the fold-change reaction expression is equal to the fold-change enzyme expression. Finally, based on the reaction expression, we prescribed the set of upregulated reactions \mathbf{R}^U and the set of downregulated reactions \mathbf{R}^D based on the fold-change reaction expression.

Case studies: Data and implementation

The first case study involved the response of *E. coli*'s metabolism to genetic (single-gene deletions) and environmental perturbations (dilution rates) performed by Ishii *et al.* [35]. The study provided ^{13}C -based flux data and RT-PCR mRNA abundances for the central carbon metabolism, pentose phosphate pathway (PPP), and the tricarboxylic acid (TCA) cycle for wild-type K12 *E. coli* culture in chemostat under different dilution rates (0.1, 0.2, 0.4, 0.5, and 0.7 hours $^{-1}$) and for 24 single-gene perturbations along the glycolysis and PPP [35]. The global transcriptional response was only captured for 5 of the 24 single-gene deletions (*pgm*, *pgi*, *gapC*, *zwf* and *rpe*) and two of the 4 dilution conditions (0.5 and 0.7 hours $^{-1}$). The differential (fold-change) gene expression levels were computed with respect to the control condition that was set to be wild-type K12 *E. coli* cultured at a dilution rate of 0.2 h $^{-1}$. The differential (fold-change) reaction expressions were subsequently evaluated based on the fold-change gene expression using the GPR Max/Min rule in the COBRA toolbox (MATLAB) [40]. For samples with only RT-PCR mRNA abundance data, the set of up- and downregulated reactions included all reactions with fold-change reaction expressions higher than 1 and those with fold-change lower than 1, respectively. In the additional analyses for samples with whole-genome transcriptome data, the set of up- and downregulated reactions were taken from the top and bottom 5th percentile of the differential reaction expressions. The differences of the measured cell specific glucose uptake rates between perturbed and control experiments were used as constraints. Δ FBA was applied using the two-step optimization with the L2 norm minimization, as described above.

The second case study came from a study of *E. coli* growth on 8 different carbon sources performed by Gerosa *et al.* [36]. Unprocessed global transcriptomic data were obtained from ArrayExpress (E-MTAB-3392), and differential expression analyses between every pair of carbon sources were evaluated using the *Limma* package in R [42]. We only included the set of genes with significance fold-change expression at FDR < 0.05. As before, the fold-change reaction expressions were computed based on fold-change in the global gene expression using the

Max-Min GPR rule using COBRA toolbox [40]. The up- and downregulated set of reactions were taken from the top and bottom 5th percentile of the differential reaction expressions. In addition, cell culture data on specific growth rates were used to compute the bounds for flux difference for biomass production rate. The uptake rates of the carbon source changes were also incorporated as constraints. We implemented the two-step optimization of Δ FBA using L2 norm minimization.

The third case study came from two studies of skeletal muscle tissue metabolism in type-2 diabetes (T2D) patients by van Tienen *et al.* [43] and Jin *et al.* [44]. The microarray gene expression datasets were obtained from GEO (GSE19420 [43] and GSE25462 [44,45], respectively) and the differential (fold-change) expression of genes for each dataset were computed using the *Limma* package in R [42]. We only included the set of genes with significance fold-change expression at FDR < 0.05. The fold-change reaction expressions were computed based on the differential gene expression using the Max/Min GPR rule [40]. In the absence of additional constraints in the form of exchange fluxes or growth characteristics, we set the up- and downregulated reactions from the top and bottom 25th percentile in differential reaction expressions, rather than the 5th percentile threshold used in *E. coli* case studies above, so as to incorporate more differentially expressed transcripts. We implemented an L1-norm minimization in the second step of Δ FBA to reduce computational complexity (time) due to the large number of constraints associated with the differential reaction expressions.

Implementation of comparative methods

Among the comparative methods in this work, the method Relative Expression and Metabolic Integrations (REMI) was specifically developed for predicting individual flux distributions of a pair of conditions (\mathbf{v}^P and \mathbf{v}^C) using multi-omics dataset, and thus more comparable to Δ FBA. The toolbox was downloaded from <https://github.com/EP-LCSB/remi>. The differential gene expressions in each case study were obtained as described above. The mapping from differential gene expression to the corresponding reaction expressions were done using the procedure detailed in REMI [31]. Briefly, the authors followed the implementation of Fang *et al.* [34] to translate gene expression ratios to obtain reaction expression ratios. When several enzyme subunits are required for a reaction, a geometric mean of expression ratios is chosen to represent the reaction ratio. In the case where multiple isozymes catalyze a reaction, the arithmetic mean of the individual expression ratios of the isozymes is used for the reaction ratio. The set of up- and down-regulated reactions \mathbf{R}^U and \mathbf{R}^D were taken from the computed differential reaction expressions as in Δ FBA implementation. Unlike Δ FBA, REMI produces solutions for the metabolic fluxes of perturbed \mathbf{v}^P and control condition \mathbf{v}^C . For comparison, we evaluated the flux change predicted by REMI by taking the difference: $\Delta\mathbf{v} = \mathbf{v}^P - \mathbf{v}^C$.

We also considered 8 additional FBA methods with transcriptome data integration, including parsimonious FBA (pFBA) [19], GIMME [20], iMAT [21], MADE [22], E-Flux [23], Lee *et al.* [24], RELATCH [25], and GX-FBA [26]. The implementation of each of these 8 methods was described in a previous systematic comparison [27]. For performance evaluation, we again evaluated the differences of flux predictions: $\Delta\mathbf{v} = \mathbf{v}^P - \mathbf{v}^C$.

Performance evaluation

The agreement between the predicted flux changes $\Delta\mathbf{v}^*$ and the ground truth ¹³C-based flux difference $\Delta\mathbf{v}^M$ was assessed by using two accuracy metrics: uncentered Pearson correlation coefficient and normalized root mean square error (NRMSE). The uncentered Pearson

correlation coefficient ρ was computed as follows

$$\rho = \frac{\Delta \mathbf{v}^M \cdot \Delta \mathbf{v}^*}{\|\Delta \mathbf{v}^M\|_2 \|\Delta \mathbf{v}^*\|_2} \quad (16)$$

Meanwhile, the NRMSE was according to the following equation—using *tdStats* package in R:

$$NRMSE = \frac{1}{\Delta v_{max}^M - \Delta v_{min}^M} \sqrt{\frac{\|\Delta \mathbf{v}^M - \Delta \mathbf{v}^*\|_2^2}{n_M}} \quad (17)$$

where n_M is the number of measured fluxes. Besides the quantitative agreement in flux changes, we also evaluated the qualitative agreement by comparing the signs of the flux changes between experimental measurements and predictions. To this end, we discretized the measured and predicted flux changes into +1, 0, and -1, to describe upregulated, no change, and downregulated reactions, respectively. The agreement in the direction of the flux changes was evaluated as the number of correct sign predictions divided by the total number of fluxes.

Metabolic subsystem enrichment analysis

The flux differences obtained from applying Δ FBA were first filtered according to the directionality of their change. The significantly altered fluxes ($|\Delta v_i| > \epsilon$) were grouped based on the subsystem to which the fluxes belong. A Fisher exact test (*fisher.test* function in the R-package) was used in determining over-represented subsystems in upregulated (positive change) and downregulated (negative change) fluxes. The statistical significance p -values were corrected for multiple hypothesis testing using the *p.adjust* function in R.

Results

Escherichia coli response to genetic and environmental variations

Ishii *et al.* [35] studied the robustness of *E. coli* K12 metabolism in chemostat in response to changes in dilution rates and to gene deletions. The study generated multi-omics data, including transcriptomic, proteomic, metabolomic, and ^{13}C metabolic fluxes, and demonstrated the remarkable ability of *E. coli* to reroute its metabolic fluxes to maintain metabolic homeostasis in response to environmental and genetic perturbations. But, only a small fraction of variation in the measured flux ratios can be explained by the fold-change in reaction expressions, as indicated by the low coefficient of determinations R^2 ($R^2 = 0.088 \pm 0.059$). The low agreement between reaction expressions and metabolic fluxes suggests that metabolic fluxes are only weakly controlled by the gene expression. The formulation of Δ FBA is driven by two main assumptions: (1) first and foremost, that metabolic flux differences are balanced—an assumption that follows directly from steady-state flux balances in the control and perturbed conditions, and in addition (2) that the flux differences should be maximally consistent with the gene expression changes. Note that Δ FBA allows for inconsistency between differential gene expression and flux difference—for example, the gene expression is downregulated, but the flux difference is positive—but such inconsistency is kept low through a constrained MILP optimization.

We applied Δ FBA using *E. coli*'s *iJO1366* GEM to predict the metabolic flux shifts from the control condition (wild-type K12 at 0.2 hour $^{-1}$ dilution rate), caused by alterations in dilution rates (0.1, 0.4, 0.5, and 0.7 hours $^{-1}$) and by 24 single-gene deletions (*galM*, *glk*, *pgm*, *pgi*, *pfkA*, *pfkB*, *fbp*, *fbaB*, *gapC*, *gpmA*, *gpmB*, *pykA*, *pykF*, *ppsA*, *zwf*, *pgl*, *gnd*, *rpe*, *rpiA*, *rpiB*, *tktA*, *tktB*, *talA*, and *talB*), one condition at a time. For each single-gene deletion experiment, the knocked-out reaction was included in the set \mathbf{R}^D and was assigned weighting $w_j^D = 10$, keeping

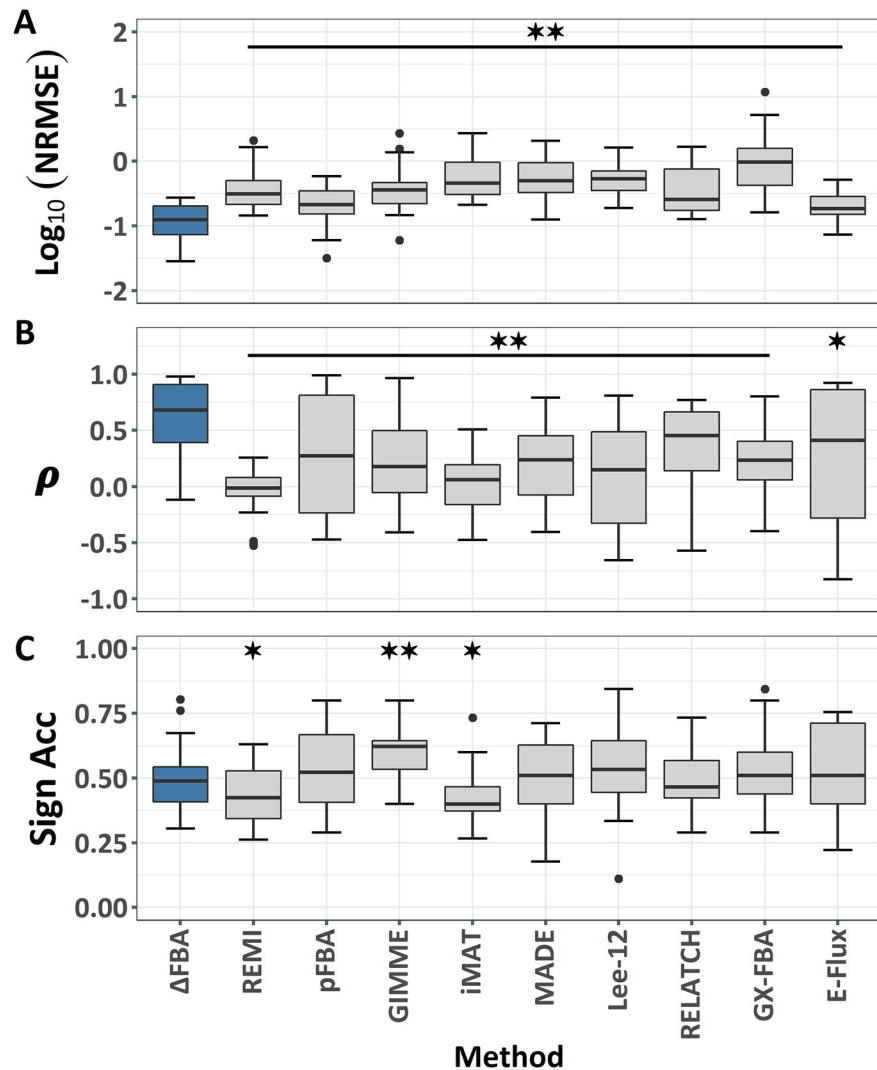


Fig 1. Comparison of the performance of Δ FBA and 9 comparative FBA methods, including REMI [31], parsimonious FBA (pFBA) [19], GIMME [20], iMAT [21], MADE [22], E-Flux [23], Lee *et al.* [24], RELATCH [25], and GX-FBA [26], in predicting *E. coli* metabolic response to environmental (dilution rates) and genetic (single gene deletions) perturbations in the Ishii *et al.* [35] study. (A) Normalized Root Mean Square Error (NRMSE) of the predicted flux differences; (B) Uncentered Pearson's Correlation Coefficient (ρ); and (C) Sign Accuracy (Sign Acc) between the predicted and measured flux differences. Statistical significance was done using two-sided paired *t*-test. * indicates *p*-value < 0.05 and ** indicates *p*-value < 0.01.

<https://doi.org/10.1371/journal.pcbi.1009589.g001>

the other weight coefficients to their default values of 1. However, we noted that using the default $w_j^D = 1$ for the knocked-out reaction produced the same outcome as using $w_j^D = 10$ in this case study. We compared the predicted flux differences using Δ FBA with the measured differences of 46 metabolic fluxes along the central carbon metabolism by incorporating the enzyme expression obtained from RT-PCR. Fig 1 depicts NRMSE, uncentered Pearson correlations, and sign accuracy of the flux differences from Δ FBA, indicating a good agreement between the prediction and the ground truth. The performance of Δ FBA is robust with respect to the thresholds used in Eqs (4)–(9) (see S1 Text and S1 and S2 Figs) and to the cut-off for differential gene expression in specifying the sets \mathbf{R}^U and \mathbf{R}^D (see S3 Fig). The results of Δ FBA

using the whole-genome gene expression profiles for a subset of perturbation experiments are comparable with those using RT-PCR data (see [S4 Fig](#)). The prediction accuracy for individual reactions is given in [S5 Fig](#), demonstrating that metabolic reactions that form futile cycles, including reactions along the glycolysis and citric acid cycle, were associated with higher prediction errors. The difficulty in predicting metabolic fluxes in futile cycles is not surprising since such cycles generate degeneracy in FBA [\[27\]](#).

We compared the prediction accuracy of flux differences by Δ FBA with REMI [\[31\]](#) and eight other FBA methods: parsimonious FBA (pFBA) [\[19\]](#), GIMME [\[20\]](#), iMAT [\[21\]](#), MADE [\[22\]](#), E-Flux [\[23\]](#), Lee et al. [\[24\]](#), RELATCH [\[25\]](#), and GX-FBA [\[26\]](#). Except for pFBA, all of these methods integrated gene expression data for the flux predictions. As illustrated in [Fig 1](#), Δ FBA outperforms the other methods in predicting the flux differences by having statistically significantly lower NRMSE and higher Pearson correlations. Meanwhile, the sign accuracies for all methods are comparable with each other. We noted that roughly 18% of the measured flux differences are exactly 0, while methods generally do not produce any zero flux differences. Here, GIMME performed better than Δ FBA (and other methods) in sign accuracy while having worse NRMSE and Pearson correlation, since the method is more readily able to produce zero flux differences than Δ FBA (e.g., when a reaction is removed from the metabolic network model [\[20\]](#)).

Another study, carried out by Gerosa *et al.* [\[36\]](#), looked at how *E. coli*'s central carbon metabolism adapts to 8 different carbon sources: acetate, fructose, galactose, glucose, glycerol, gluconate, pyruvate and succinate. The study generated ^{13}C metabolic flux, metabolite concentration and microarray gene expression data from exponentially growing *E. coli* under each carbon source. The study found that only a small subset of the numerous transcriptome changes translates to notable shifts in the corresponding metabolic fluxes, indicating non-trivial relationships between transcriptional regulations and metabolic fluxes. We applied Δ FBA to predict flux changes between every pair of the carbon sources, treating one as the perturbation and another as the control condition. [Fig 2](#) describes the good agreement between the flux difference predictions by Δ FBA with the measured differences of 34 metabolic fluxes between any two carbon sources, specifically in terms of NRMSE (mean: 0.15), uncentered Pearson correlation (mean: 0.61), and sign accuracy (mean: 0.66). The findings from Ishii *et al.* [\[35\]](#) and Gerosa *et al.* [\[36\]](#) highlight the ability of Δ FBA in accurately predicting metabolic flux alterations using transcriptomic data for both environmental (e.g., dilution rates, carbon sources) and genetic perturbations.

Dysregulation of skeletal muscle metabolism in type-2 diabetes

In this case study, we looked at metabolic alterations of human muscle using the myocyte GEM *iMyocyte2419* [\[37\]](#) and gene expression datasets from two type-2 diabetes (T2D) studies, one by van Tienen *et al.* [\[43\]](#) and another by Jin *et al.* [\[44\]](#). The study by van Tienen *et al.* [\[43\]](#) compared long term T2D patients with age-matched cohort, and reported the downregulation of gene expression related to substrate transport into mitochondria, conversion of pyruvate into acetyl-CoA, aspartate-malate shuttle in mitochondria, glycolysis, TCA cycle, and electron transport chain. Similarly, Jin *et al.* [\[44\]](#) reported a significant enrichment of pathways involved the oxidative phosphorylation among the downregulated genes in their T2D cohort when compared to control. Jin *et al.* [\[44\]](#) further identified the transcription factor SRF and its cofactor MKL1 among the top-ranking enriched gene sets with increased expression. But, the correlation between the differential gene expressions in the two studies is only modest. [\[37\]](#)

We applied Δ FBA to predict the flux changes based on the differential gene expressions in each of the two studies above (see [Materials and methods](#)). We grouped the reactions based on

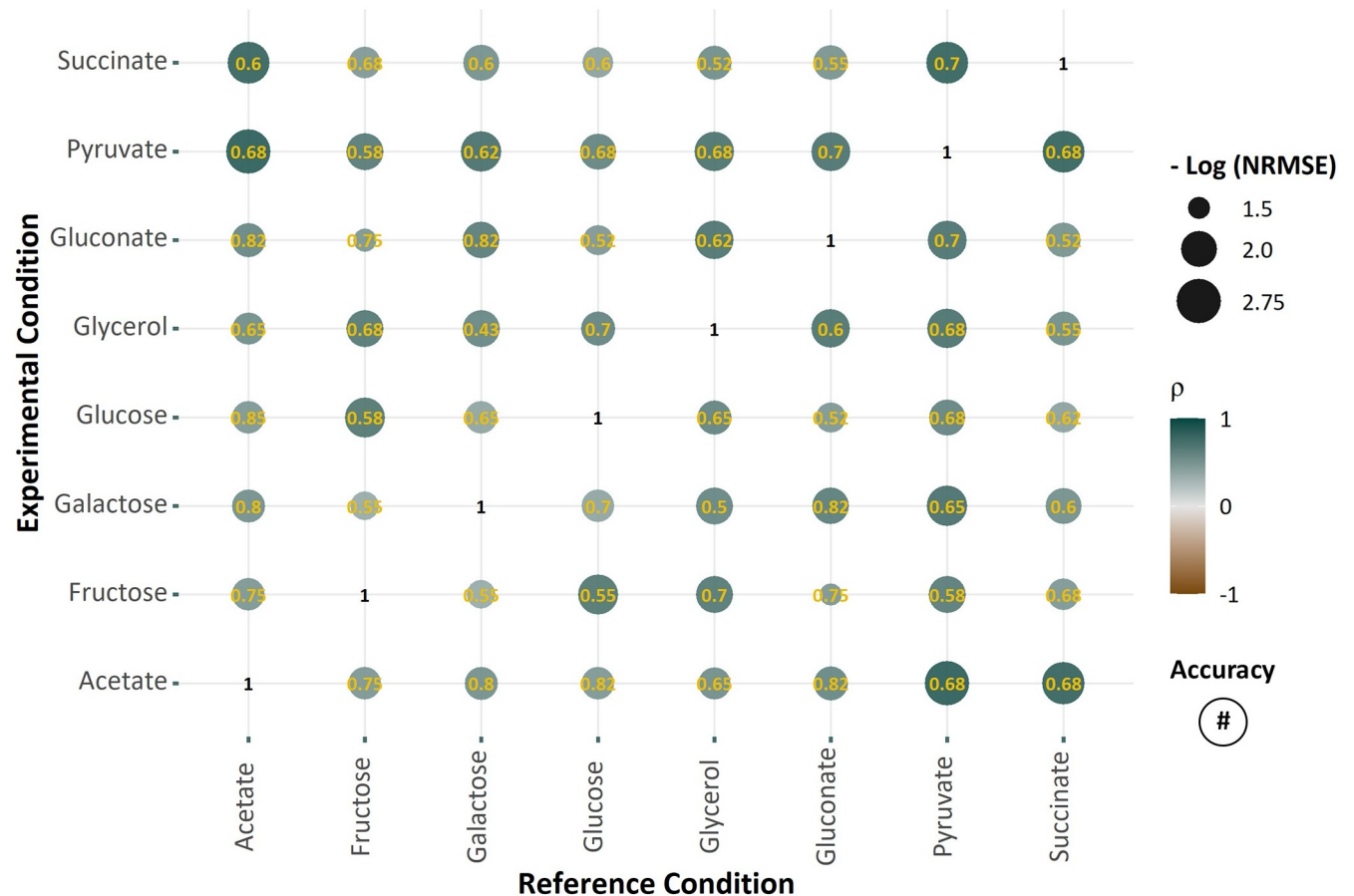


Fig 2. Prediction of metabolic flux changes in *E. coli* caused by changes in the carbon source using Δ FBA. The horizontal axis reports the reference carbon source (control) and the vertical axis shows the altered (perturbed) carbon source. Uncentered Pearson's Correlation Coefficient (ρ) is shown by the color of the markers. NRMSE is represented by the size of the markers—the larger the markers, the smaller is the NRMSE. Finally, the directional (sign) accuracy of the flux perturbation predictions is shown by the numbers inside the markers.

<https://doi.org/10.1371/journal.pcbi.1009589.g002>

whether the predicted flux differences are positive or negative, denoted by up- and down-reactions, respectively. We performed metabolic subsystem enrichment analysis using the subsystems defined in myocyte specific GEM *iMyocyte2419* [37] to identify over-represented metabolic subsystems among the up- and down-reactions (see [Materials and methods](#)). As summarized in [Fig 3](#), the enrichment analysis of metabolic changes in the van Tienen *et al.* study shows a significant over-representation of β -oxidation and BCAA (branched-chain amino acids) metabolism among the down-reactions, and of extracellular transport and lipid metabolism among the up-reactions. The enrichment analysis of flux differences in the Jin *et al.* study also indicates an over-representation of lipid metabolism among the up-reactions in T2D patients, as well as an over-representation of β -oxidation pathway among the down-reactions (see [Fig 3](#)).

Furthermore, we evaluated the difference in the flux throughput for every metabolite irrespective of its compartmental location by computing the difference in the total production flux of each metabolite. Metabolites with a large difference in the flux throughput are of particular interest for disease biomarkers. In the following, we focused on metabolites that have a flux throughput change above a threshold ($|\Delta v_i| > 1\%$ of the largest flux bounds) and excluded

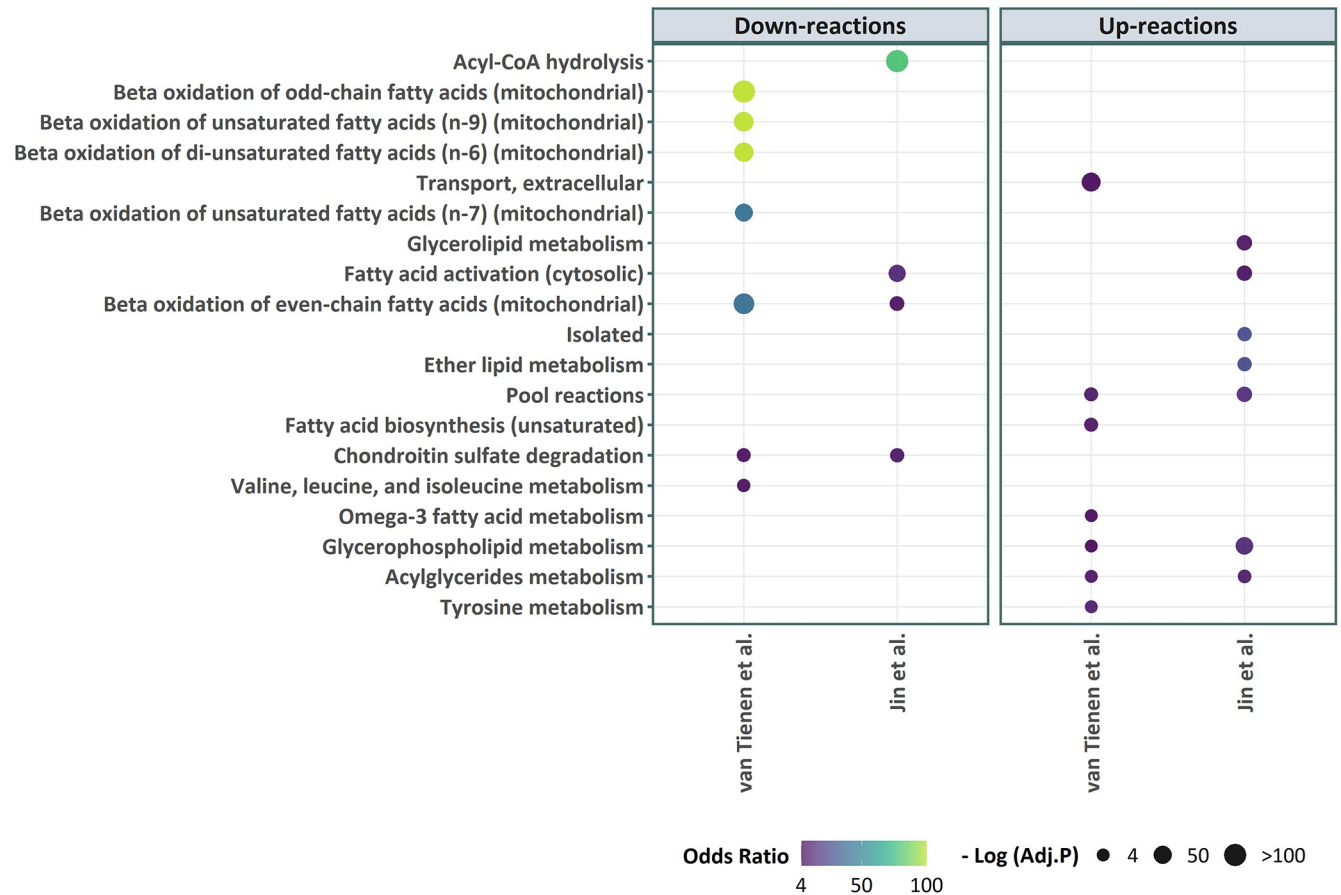


Fig 3. Enriched metabolic subsystems (FDR<0.05) among the in T2D patients based on flux changes predicted using Δ FBA. The flux changes were computed based on the transcriptome datasets from two T2D studies: van Tienen *et al.* [43] (GSE19420) and Jin *et al.* [44] (GSE25462). The statistical significance of the over-representation is shown by the size of the markers—larger markers have smaller adjusted p -values—while the odds ratio is shown by the color of the markers.

<https://doi.org/10.1371/journal.pcbi.1009589.g003>

intermediary metabolites that participate in linear reaction sequences. Fig 4 shows the flux throughput differences predicted by Δ FBA for various metabolites. Among the metabolites with a large drop in the flux throughput in both studies are Coenzyme A (CoA), Acetyl-CoA and AMP (Adenosine monophosphate), all of which have been previously identified as metabolite reporters of diabetes [37,46]. Other metabolic biomarkers that have been previously proposed for T2D, such as repression of FAD (Flavin adenine dinucleotide), FADH₂ and NADH by van Tienen *et al.* study [43] and increased glycerol by Jin *et al.* study [44], are confirmed by Δ FBA (see Fig 4). Väre *et al.* [37] had identified these markers of T2D using gene-reaction associations and consensus gene-set analysis in the GEM, *iMyocyte2419*. Besides the above confirmatory observations, Δ FBA results of the two studies further suggest that arachidonate and palmitate are candidate metabolic biomarkers for T2D, both of which have a large positive flux throughput change in the two T2D studies. These metabolites are undetected by simple gene-set analysis using GPR associations in the GEM, but have important roles in the progression and cause of T2D [47–49]. The results above showcase the ability of Δ FBA in elucidating metabolic flux alterations in a complex human GEM and identifying key metabolites of interest in human diseases.

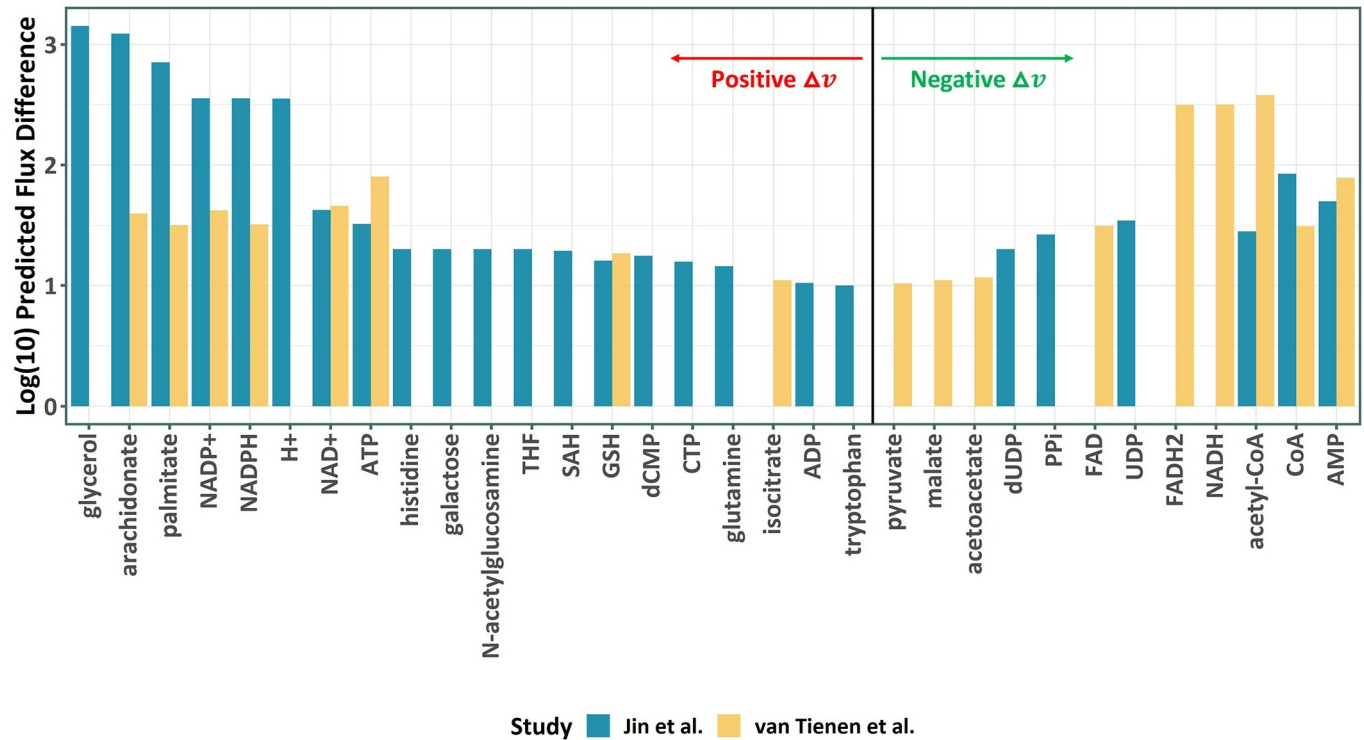


Fig 4. Alterations in metabolite flux throughput in T2D patients as predicted by Δ FBA.

<https://doi.org/10.1371/journal.pcbi.1009589.g004>

Discussions

GEMs and constraint-based modeling using FBA and the myriad FBA variants have proven to be important enabling tools for establishing genotype-phenotype relationship [10,50,51]. The increasing availability of omics data have driving the development of FBA-based strategies that are able to use such data to improve the accuracy of predictions of intracellular metabolic fluxes. In this work, we present a new FBA-based method, called Δ FBA, built for the purpose of analyzing the metabolic alterations between two conditions given data on differential gene expression. Δ FBA does not require the specification of the metabolic objective, and thus, eliminates any potential pitfalls that are associated with an incorrect selection of this objective. Note that Δ FBA does not generate the flux prediction for a given condition; rather, the method produces differences of metabolic fluxes between two conditions. Differential flux predictions are indispensable in formulating hypothesis and in understanding the physiological response of cells to changes in the environment. Δ FBA can be easily integrated and have been tested to work with the widely popular COBRA toolbox [34].

We showed the applicability and performance of Δ FBA for predicting metabolic flux changes in an array of experimental perturbations and in both simple prokaryotic *E. coli* and complex multicellular human muscle cells. In comparison to other relevant FBA methods, Δ FBA show a markedly better accuracy in prediction metabolic flux changes in *E. coli*. Further, the application of Δ FBA to two T2D studies shed light on the rewiring of muscle metabolism associated with type-2 diabetes that leads to the repression of β -oxidation and activation of glycerolphospholipids, pointing to increased lipid metabolism in the T2D patients. Interestingly, serum metabolic profiling of T2D patients showed increased glycerolphospholipids when compared to healthy controls [52]. Besides, clinical and experimental studies have demonstrated

the association between phospholipids and insulin resistance [53]. Furthermore, by looking at the changes in the flux throughput of metabolites, the results of Δ FBA suggest two fatty acids, arachidonate and palmitate, for candidate biomarkers of T2D.

There are several limitations of Δ FBA, the most obvious of which is that the method does not produce flux predictions for individual conditions under comparison. If separate flux predictions for control and perturbed conditions are desired, Δ FBA can be applied synergistically with another FBA method that is capable of predicting single-condition metabolic fluxes. Many of such methods, such as GIMME [20] and iMAT [21], transform gene expression data to a binary state (active/inactive, high/low) and produce metabolic flux prediction for a single condition. But, as shown in Fig 1, pFBA often works as well, if not better, without using gene expression data. When deciding the reference (control) condition, the more well-characterized metabolic state (e.g., more experimental data, more obvious metabolic objective) should be used to generate the reference flux distribution. Δ FBA flux differences can be combined with the reference flux values by simple algebra to evaluate metabolic fluxes of the other (perturbed) condition. Such a strategy may be advantageous since once the metabolic flux distribution for the baseline condition is accurately determined (and ideally experimentally validated), one can use Δ FBA and differential gene expression datasets for various perturbation experiments to generate accurate prediction for metabolic fluxes of the perturbed conditions. Note that for many gene expression profiling technologies the relative (differential) expressions are often more reliable and informative of the underlying cellular alterations than the absolute expression because of technical and biological considerations.

Finally, while in the formulation and the application of Δ FBA we considered only differential gene expression data, the method can also accommodate other omics dataset, such as proteomics, by appropriate mapping of the data to changes in reaction expressions. Metabolomics data can also be accommodated in Δ FBA via thermodynamics constraints, as done in REMI [31], in which certain reactions can only proceed in one direction.

Supporting information

S1 Text. Threshold criteria for minimum flux change magnitudes.

(PDF)

S1 Fig. Comparison of Δ FBA predictions of *E. coli* metabolic response in Ishii *et al.* study [35] using original (relaxed) thresholding in Equation (S1) and stringent thresholding using fold-change reaction expression in Equations (S2)-(S3) (see S1 Text). (Left) Normalized root mean square error (NRMSE), (Middle) uncentered Pearson's Correlation Coefficient (ρ), (Right) Sign accuracy (Sign Acc) between the predicted flux difference and the measured flux change. The error bars show standard deviation across for 4 dilution rates (0.1, 0.4, 0.5, and 0.7 hours⁻¹) and 24 single-gene deletions (*galM*, *glk*, *pgm*, *pgi*, *pfkA*, *pfkB*, *fbp*, *fbaB*, *gapC*, *gpmA*, *gpmB*, *pykA*, *pykF*, *ppsA*, *zwf*, *pgl*, *gnd*, *rpe*, *rpiA*, *rpiB*, *tktA*, *tktB*, *talA*, and *talB*). The difference in performance is not statistically significant.

(TIF)

S2 Fig. Comparison of FBA performance for different ϵ . Accuracy of Δ FBA predictions of *E. coli* metabolic shifts in response to environmental and genetic perturbations in the Ishii *et al.* study [35]. The default ϵ is 0.1% of the largest flux in the metabolic model under growth maximization and parsimony criteria. The error bars show standard deviation across flux difference predictions for 4 dilution rates and 24 single-gene deletions. The result indicates that the performance of Δ FBA is relatively insensitive to ϵ between 0.01% and 1%.

(TIF)

S3 Fig. Comparison of Δ FBA performance for different fold-change (FC) expression cut-off for assigning up- and down-regulated reactions (the sets R^U and R^D). The default FC cut-off is 1. The error bars show standard deviation across for 4 dilution rates and 24 single-gene deletions in the Ishii et al. study [35]. The difference in performance is not statistically significant (Mean NRMSE—FC cutoff of 1 = 0.14, FC cutoff of 2 = 0.13; Mean ρ —FC cutoff of 1 = 0.61, FC cutoff of 2 = 0.63; Mean sign accuracy—FC cutoff of 1 = 0.49, FC cutoff of 2 = 0.48).

(TIF)

S4 Fig. Comparison of Δ FBA performance in predicting *E. coli* metabolic shifts using whole-genome transcriptome data versus using RT-PCR mRNA data. Directional (Sign Accuracy) agreement and uncentered Pearson's Correlation Coefficient (ρ) between the predicted and measured flux differences have little difference between the incorporation of the two transcriptomic sources using Δ FBA (Mean NRMSE: whole-genome = 0.15, RT-PCR = 0.16; Mean ρ : whole-genome = 0.57; RT-PCR = 0.54; Mean sign accuracy: whole-genome = 0.53, RT-PCR = 0.53).

(TIF)

S5 Fig. Normalized prediction errors of flux differences by Δ FBA across 46 individual reactions in *E. coli* central carbon metabolism in Ishii et al. study [35]. The NRMSE for the full flux differences is shown in blue (leftmost box plot). The remaining box plots in red show the distribution of the normalized error (NE) for each flux i : $NE_i = \frac{\sqrt{(\Delta v_i^M - \Delta v_i^*)^2}}{\Delta v_{max}^M - \Delta v_{min}^M}$, across 28 conditions (4 dilution rates and 24 single-gene deletions).

(TIF)

Acknowledgments

The authors would like to acknowledge the University at Buffalo's Center for Computational Research for computational support.

Author Contributions

Conceptualization: Sudharshan Ravi, Rudiyanto Gunawan.

Data curation: Sudharshan Ravi.

Formal analysis: Sudharshan Ravi, Rudiyanto Gunawan.

Funding acquisition: Rudiyanto Gunawan.

Investigation: Sudharshan Ravi, Rudiyanto Gunawan.

Methodology: Sudharshan Ravi, Rudiyanto Gunawan.

Project administration: Rudiyanto Gunawan.

Software: Sudharshan Ravi.

Supervision: Rudiyanto Gunawan.

Validation: Sudharshan Ravi.

Visualization: Sudharshan Ravi, Rudiyanto Gunawan.

Writing – original draft: Sudharshan Ravi, Rudiyanto Gunawan.

Writing – review & editing: Sudharshan Ravi, Rudiyanto Gunawan.

References

1. Granata I, Troiano E, Sangiovanni M, Guarracino MR. Integration of transcriptomic data in a genome-scale metabolic model to investigate the link between obesity and breast cancer. *BMC Bioinformatics*. 2019; 20: 162. <https://doi.org/10.1186/s12859-019-2685-9> PMID: 30999849
2. Agren R, Mardinoglu A, Asplund A, Kampf C, Uhlen M, Nielsen J. Identification of anticancer drugs for hepatocellular carcinoma through personalized genome-scale metabolic modeling. *Mol Syst Biol*. 2014; 10: 721. <https://doi.org/10.1002/msb.145122> PMID: 24646661
3. Nilsson A, Nielsen J. Genome scale metabolic modeling of cancer. *Metabolic Engineering*. Academic Press Inc.; 2017. pp. 103–112. <https://doi.org/10.1016/j.ymben.2016.10.022> PMID: 27825806
4. Lewis NE, Abdel-Haleem AM. The evolution of genome-scale models of cancer metabolism. *Front Physiol*. 2013; 4: 237. <https://doi.org/10.3389/fphys.2013.00237> PMID: 24027532
5. Zhang C, Hua Q. Applications of genome-scale metabolic models in biotechnology and systems medicine. *Front Physiol*. 2016; 6: 413. <https://doi.org/10.3389/fphys.2015.00413> PMID: 26779040
6. Gu C, Kim GB, Kim WJ, Kim HU, Lee SY. Current status and applications of genome-scale metabolic models. *Genome Biol*. 2019; 20: 121. <https://doi.org/10.1186/s13059-019-1730-3> PMID: 31196170
7. Oberhardt MA, Palsson B, Papin JA. Applications of genome-scale metabolic reconstructions. *Molecular Systems Biology*. 2009. <https://doi.org/10.1038/msb.2009.77> PMID: 19888215
8. Santos F, Boele J, Teusink B. A practical guide to genome-scale metabolic models and their analysis. *Methods in Enzymology*. Academic Press Inc.; 2011. pp. 509–532. <https://doi.org/10.1016/B978-0-12-385118-5.00024-4> PMID: 21943912
9. Bordbar A, Monk JM, King ZA, Palsson BO. Constraint-based models predict metabolic and associated cellular functions. *Nat Rev Genet*. 2014; 15: 107–120. <https://doi.org/10.1038/nrg3643> PMID: 24430943
10. Lewis NE, Nagarajan H, Palsson BO. Constraining the metabolic genotype-phenotype relationship using a phylogeny of in silico methods. *Nature Reviews Microbiology*. Nature Publishing Group; 2012. pp. 291–305. <https://doi.org/10.1038/nrmicro2737> PMID: 22367118
11. Orth JD, Thiele I, Palsson BO. What is flux balance analysis? *Nat Biotechnol*. 2010; 28: 245–248. <https://doi.org/10.1038/nbt.1614> PMID: 20212490
12. Park JH, Lee SY. Towards systems metabolic engineering of microorganisms for amino acid production. *Current Opinion in Biotechnology*. 2008. pp. 454–460. <https://doi.org/10.1016/j.copbio.2008.08.007> PMID: 18760356
13. Kim TY, Sohn SB, Kim HU, Lee SY. Strategies for systems-level metabolic engineering. *Biotechnology Journal*. 2008. pp. 612–623. <https://doi.org/10.1002/biot.200700240> PMID: 18246579
14. Nevoigt E. Progress in Metabolic Engineering of *Saccharomyces cerevisiae*. *Microbiol Mol Biol Rev*. 2008; 72: 379–412. <https://doi.org/10.1128/MMBR.00025-07> PMID: 18772282
15. Kim HU, Kim TY, Lee SY. Metabolic flux analysis and metabolic engineering of microorganisms. *Mol Biosyst*. 2008; 4: 113–120. <https://doi.org/10.1039/b712395g> PMID: 18213404
16. Ebrahim A, Almaas E, Bauer E, Bordbar A, Burgard AP, Chang RL, et al. Do genome-scale models need exact solvers or clearer standards? *Mol Syst Biol*. 2015; 11: 831. <https://doi.org/10.15252/msb.20156157> PMID: 26467284
17. Richelle A, Chiang AWT, Kuo C-C, Lewis NE. Increasing consensus of context-specific metabolic models by integrating data-inferred cell functions. Ouzounis CA, editor. *PLOS Comput Biol*. 2019; 15: e1006867. <https://doi.org/10.1371/journal.pcbi.1006867> PMID: 30986217
18. Hyde DR, Lewis NE, Palsson BO. Analysis of omics data with genome-scale models of metabolism. *Molecular BioSystems*. NIH Public Access; 2013. pp. 167–174. <https://doi.org/10.1039/c2mb25453k> PMID: 23247105
19. Lewis NE, Hixson KK, Conrad TM, Lerman JA, Charusanti P, Polpitiya AD, et al. Omic data from evolved *E. coli* are consistent with computed optimal growth from genome-scale models. *Mol Syst Biol*. 2010; 6: 390. <https://doi.org/10.1038/msb.2010.47> PMID: 20664636
20. Becker SA, Palsson BO. Context-specific metabolic networks are consistent with experiments. Sauro HM, editor. *PLoS Comput Biol*. 2008; 4: e1000082. <https://doi.org/10.1371/journal.pcbi.1000082> PMID: 18483554
21. Zur H, Rupp E, Shlomi T. iMAT: an integrative metabolic analysis tool. *Bioinformatics*. 2010; 26: 3140–2. <https://doi.org/10.1093/bioinformatics/btq602> PMID: 21081510
22. Jensen PA, Papin JA. Functional integration of a metabolic network model and expression data without arbitrary thresholding. *Bioinformatics*. 2011; 27: 541–547. <https://doi.org/10.1093/bioinformatics/btq702> PMID: 21172910

23. Colijn C, Brandes A, Zucker J, Lun DS, Weiner B, Farhat MR, et al. Interpreting Expression Data with Metabolic Flux Models: Predicting Mycobacterium tuberculosis Mycolic Acid Production. *PLOS Comput Biol*. 2009; 5: e1000489. <https://doi.org/10.1371/journal.pcbi.1000489> PMID: 19714220
24. Lee D, Smallbone K, Dunn WB, Murabito E, Winder CL, Kell DB, et al. Improving metabolic flux predictions using absolute gene expression data. *BMC Syst Biol*. 2012; 6: 1–9.
25. Kim J, Reed JL. RELATCH: relative optimality in metabolic networks explains robust metabolic and regulatory responses to perturbations. *Genome Biol*. 2012; 13: 1–12. <https://doi.org/10.1186/gb-2012-13-9-r78> PMID: 23013597
26. Navid A, Almaas E. Genome-level transcription data of *Yersinia pestis* analyzed with a New metabolic constraint-based approach. *BMC Syst Biol*. 2012; 6: 1–18.
27. Machado D, Herrgård M. Systematic Evaluation of Methods for Integration of Transcriptomic Data into Constraint-Based Models of Metabolism. Maranas CD, editor. *PLoS Comput Biol*. 2014; 10: e1003580. <https://doi.org/10.1371/journal.pcbi.1003580> PMID: 24762745
28. O'Brien EJ, Lerman JA, Chang RL, Hyduke DR, Palsson B. Genome-scale models of metabolism and gene expression extend and refine growth phenotype prediction. *Mol Syst Biol*. 2013; 9: 693. <https://doi.org/10.1038/msb.2013.52> PMID: 24084808
29. Sánchez BJ, Zhang C, Nilsson A, Lahtvee P, Kerkhoven EJ, Nielsen J. Improving the phenotype predictions of a yeast genome-scale metabolic model by incorporating enzymatic constraints. *Mol Syst Biol*. 2017; 13: 935. <https://doi.org/10.15252/msb.20167411> PMID: 28779005
30. Salvy P, Hatzimanikatis V. The ETFL formulation allows multi-omics integration in thermodynamics-compliant metabolism and expression models. *Nat Commun*. 2020; 11: 30. <https://doi.org/10.1038/s41467-019-13818-7> PMID: 31937763
31. Pandey V, Hadadi N, Hatzimanikatis V. Enhanced flux prediction by integrating relative expression and relative metabolite abundance into thermodynamically consistent metabolic models. Patil KR, editor. *PLoS Comput Biol*. 2019; 15: e1007036. <https://doi.org/10.1371/journal.pcbi.1007036> PMID: 31083653
32. Zhu L, Zheng H, Hu X, Xu Y. A computational method using differential gene expression to predict altered metabolism of multicellular organisms. *Mol Biosyst*. 2017; 13: 2418–2427. <https://doi.org/10.1039/c7mb00462a> PMID: 28972214
33. Pusa T, Ferrarini MG, Andrade R, Mary A, Marchetti-Spaccamela A, Stougie L, et al. MOOMIN—Mathematical exploration of 'Omics data on a Metabolic Network. *Bioinformatics*. 2020; 36: 514–523. <https://doi.org/10.1093/bioinformatics/btz584> PMID: 31504164
34. Heirendt L, Arreckx S, Pfau T, Mendoza SN, Richelle A, Heinken A, et al. Creation and analysis of biochemical constraint-based models using the COBRA Toolbox v.3.0. *Nat Protoc*. 2019; 14: 639–702. <https://doi.org/10.1038/s41596-018-0098-2> PMID: 30787451
35. Ishii N, Nakahigashi K, Baba T, Robert M, Soga T, Kanai A, et al. Multiple high-throughput analyses monitor the response of *E. coli* to perturbations. 2007; 316. <https://doi.org/10.1126/science.1132067> PMID: 17379776
36. Gerosa L, Haverkorn Van Rijsewijk BRB, Christodoulou D, Kochanowski K, Schmidt TSB, Noor E, et al. Pseudo-transition Analysis Identifies the Key Regulators of Dynamic Metabolic Adaptations from Steady-State Data. *Cell Syst*. 2015; 1: 270–282. <https://doi.org/10.1016/j.cels.2015.09.008> PMID: 27136056
37. Våremo L, Scheele C, Broholm C, Mardinoglu A, Kampf C, Asplund A, et al. Proteome- and Transcriptome-Driven Reconstruction of the Human Myocyte Metabolic Network and Its Use for Identification of Markers for Diabetes. *Cell Rep*. 2015; 11: 921–933. <https://doi.org/10.1016/j.celrep.2015.04.010> PMID: 25937284
38. Griva I, Nash S (Stephen G., Sofer A. Linear and nonlinear optimization. 2009; 742.
39. Segrè D, Vitkup D, Church GM. Analysis of optimality in natural and perturbed metabolic networks. *Proc Natl Acad Sci U S A*. 2002; 99: 15112–15117. <https://doi.org/10.1073/pnas.232349399> PMID: 12415116
40. Opdam S, Richelle A, Kellman B, Li S, Zielinski DC, Lewis NE Correspondence, et al. A Systematic Evaluation of Methods for Tailoring Genome-Scale Metabolic Models Article A Systematic Evaluation of Methods for Tailoring Genome-Scale Metabolic Models. *Cell Syst*. 2017; 4: 318–329. PMID: 28215528
41. Richelle A, Joshi C, Lewis NE. Assessing key decisions for transcriptomic data integration in biochemical networks. *PLOS Comput Biol*. 2019; 15: e1007185. <https://doi.org/10.1371/journal.pcbi.1007185> PMID: 31323017
42. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015; 43: e47. <https://doi.org/10.1093/nar/gkv007> PMID: 25605792

43. Van Tienen FHJ, Praet SFE, De Feyter HM, Van Den Broek NM, Lindsey PJ, Schoonderwoerd KGC, et al. Physical activity is the key determinant of skeletal muscle mitochondrial function in type 2 diabetes. *J Clin Endocrinol Metab.* 2012; 97: 3261–3269. <https://doi.org/10.1210/jc.2011-3454> PMID: [22802091](https://pubmed.ncbi.nlm.nih.gov/22802091/)
44. Jin W, Goldfine AB, Boes T, Henry RR, Ciaraldi TP, Kim EY, et al. Increased SRF transcriptional activity in human and mouse skeletal muscle is a signature of insulin resistance. *J Clin Invest.* 2011; 121: 918–929. <https://doi.org/10.1172/JCI41940> PMID: [21393865](https://pubmed.ncbi.nlm.nih.gov/21393865/)
45. Lerin C, Goldfine AB, Boes T, Liu M, Kasif S, Dreyfuss JM, et al. Defects in muscle branched-chain amino acid oxidation contribute to impaired lipid metabolism. *Mol Metab.* 2016; 5: 926–936. <https://doi.org/10.1016/j.molmet.2016.08.001> PMID: [27689005](https://pubmed.ncbi.nlm.nih.gov/27689005/)
46. Misra P, Chakrabarti R. The role of AMP kinase in diabetes. *Indian Journal of Medical Research.* *Indian J Med Res;* 2007. pp. 389–398. Available: <https://pubmed.ncbi.nlm.nih.gov/17496363/> PMID: [17496363](https://pubmed.ncbi.nlm.nih.gov/17496363/)
47. Ly LD, Xu S, Choi S-K, Ha C-M, Thoudam T, Cha S-K, et al. Oxidative stress and calcium dysregulation by palmitate in type 2 diabetes. *Exp Mol Med.* 2017. <https://doi.org/10.1038/emmm.2016.157> PMID: [28154371](https://pubmed.ncbi.nlm.nih.gov/28154371/)
48. Igoillo-Esteve M, Marselli L, Cunha DA, Ladrière L, Ortis F, Grieco FA, et al. Palmitate induces a pro-inflammatory response in human pancreatic islets that mimics CCL2 expression by beta cells in type 2 diabetes. *Diabetologia.* 2010. <https://doi.org/10.1007/s00125-010-1707-y> PMID: [20369226](https://pubmed.ncbi.nlm.nih.gov/20369226/)
49. Das UN. Arachidonic acid in health and disease with focus on hypertension and diabetes mellitus: A review. *Journal of Advanced Research.* 2018. <https://doi.org/10.1016/j.jare.2018.01.002> PMID: [30034875](https://pubmed.ncbi.nlm.nih.gov/30034875/)
50. King ZA, Lloyd CJ, Feist AM, Palsson BO. Next-generation genome-scale models for metabolic engineering. *Current Opinion in Biotechnology.* Elsevier Ltd; 2015. pp. 23–29. <https://doi.org/10.1016/j.copbio.2014.12.016> PMID: [25575024](https://pubmed.ncbi.nlm.nih.gov/25575024/)
51. O'Brien EJ, Monk JM, Palsson BO. Using genome-scale models to predict biological capabilities. *Cell.* *Cell Press;* 2015. pp. 971–987. <https://doi.org/10.1016/j.cell.2015.05.019> PMID: [26000478](https://pubmed.ncbi.nlm.nih.gov/26000478/)
52. Xu F, Tavintharan S, Sum CF, Woon K, Lim SC, Ong CN. Metabolic signature shift in type 2 diabetes mellitus revealed by mass spectrometry-based metabolomics. *J Clin Endocrinol Metab.* 2013. <https://doi.org/10.1210/jc.2012-4132> PMID: [23633210](https://pubmed.ncbi.nlm.nih.gov/23633210/)
53. Chang W, Hatch GM, Wang Y, Yu F, Wang M. The relationship between phospholipids and insulin resistance: From clinical to experimental studies. *Journal of Cellular and Molecular Medicine.* 2019. <https://doi.org/10.1111/jcmm.13984> PMID: [30402908](https://pubmed.ncbi.nlm.nih.gov/30402908/)