

RESEARCH ARTICLE

# p<sup>13</sup>CMFA: Parsimonious <sup>13</sup>C metabolic flux analysis

Carles Foguet<sup>1,2</sup>, Anusha Jayaraman<sup>1‡</sup>, Silvia Marin<sup>1,2‡</sup>, Vitaly A. Selivanov<sup>1,2</sup>, Pablo Moreno<sup>3</sup>, Ramon Messeguer<sup>4</sup>, Pedro de Atauri<sup>1,2\*</sup>, Marta Cascante<sup>1,2\*</sup>

**1** Department of Biochemistry and Molecular Biomedicine & Institute of Biomedicine of University of Barcelona, Faculty of Biology, Universitat de Barcelona, Barcelona, Spain, **2** Centro de Investigación Biomédica en Red de Enfermedades Hepáticas y Digestivas (CIBEREHD) and Metabolomics node at Spanish National Bioinformatics Institute (INB-ISCI-ES-ELIXIR), Instituto de Salud Carlos III (ISCI), Madrid, Spain, **3** European Molecular Biology Laboratory, European Bioinformatics Institute (EMBL-EBI), Cambridge, United Kingdom, **4** LEITAT Technological Center, Health & Biomedicine Unit, Barcelona, Spain

‡ These authors contributed equally to this work.  
\* [pde\\_atauri@ub.edu](mailto:pde_atauri@ub.edu) (PdA); [martacascante@ub.edu](mailto:martacascante@ub.edu) (MC)



**OPEN ACCESS**

**Citation:** Foguet C, Jayaraman A, Marin S, Selivanov VA, Moreno P, Messeguer R, et al. (2019) p<sup>13</sup>CMFA: Parsimonious <sup>13</sup>C metabolic flux analysis. *PLoS Comput Biol* 15(9): e1007310. <https://doi.org/10.1371/journal.pcbi.1007310>

**Editor:** Vassily Hatzimanikatis, Ecole Polytechnique Fédérale de Lausanne, SWITZERLAND

**Received:** October 23, 2018

**Accepted:** August 6, 2019

**Published:** September 6, 2019

**Copyright:** © 2019 Foguet et al. 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 are within the manuscript and its Supporting Information files.

**Funding:** This work was supported by European Commission (EC-654241, FP7-PEOPLE-264780), MINECO-European Commission FEDER funds – “Una manera de hacer Europa” (SAF2017-89673-R; SAF2015-70270-REDT), Agència de Gestió d’Ajuts Universitaris i de Recerca (AGAUR) – Generalitat de Catalunya (2017SGR1033) and Instituto de Salud Carlos III (CIBEREHD, CB17/04/00023). CF acknowledges the support received

## Abstract

Deciphering the mechanisms of regulation of metabolic networks subjected to perturbations, including disease states and drug-induced stress, relies on tracing metabolic fluxes. One of the most informative data to predict metabolic fluxes are <sup>13</sup>C based metabolomics, which provide information about how carbons are redistributed along central carbon metabolism. Such data can be integrated using <sup>13</sup>C Metabolic Flux Analysis (<sup>13</sup>C MFA) to provide quantitative metabolic maps of flux distributions. However, <sup>13</sup>C MFA might be unable to reduce the solution space towards a unique solution either in large metabolic networks or when small sets of measurements are integrated. Here we present parsimonious <sup>13</sup>C MFA (p<sup>13</sup>CMFA), an approach that runs a secondary optimization in the <sup>13</sup>C MFA solution space to identify the solution that minimizes the total reaction flux. Furthermore, flux minimization can be weighted by gene expression measurements allowing seamless integration of gene expression data with <sup>13</sup>C data. As proof of concept, we demonstrate how p<sup>13</sup>CMFA can be used to estimate intracellular flux distributions from <sup>13</sup>C measurements and transcriptomics data. We have implemented p<sup>13</sup>CMFA in Iso2Flux, our in-house developed isotopic steady-state <sup>13</sup>C MFA software. The source code is freely available on GitHub (<https://github.com/cfoguet/iso2flux/releases/tag/0.7.2>).

## Author summary

<sup>13</sup>C Metabolic Flux Analysis (<sup>13</sup>C MFA) is a well-established technique that has proven to be a valuable tool in quantifying the metabolic flux profile of central carbon metabolism. When a biological system is incubated with a <sup>13</sup>C-labeled substrate, <sup>13</sup>C propagates to metabolites throughout the metabolic network in a flux and pathway-dependent manner. <sup>13</sup>C MFA integrates measurements of <sup>13</sup>C enrichment in metabolites to identify the flux distributions consistent with the measured <sup>13</sup>C propagation. However, there is often a range of flux values that can lead to the observed <sup>13</sup>C distribution. Indeed, either when the metabolic network is large or a small set of measurements are integrated, the range of

through “Becas de la Caixa para estudios de doctorado en universidades españolas” funded by the “La Caixa” foundation. CF also acknowledges the support from the Spanish National Bioinformatics Institute (INB-ISCI-ES-ELIXIR). MC acknowledges the support received through the prize “ICREA Academia” for excellence in research, funded by ICREA foundation – Generalitat de Catalunya. 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.

valid solutions can be too wide to accurately estimate part of the underlying flux distribution. Here we propose to use flux minimization to select the best flux solution in the <sup>13</sup>C MFA solution space. Furthermore, this approach can integrate gene expression data to give greater weight to the minimization of fluxes through enzymes with low gene expression evidence in order to ensure that the selected solution is biologically relevant. The concept of using flux minimization to select the best solution is widely used in flux balance analysis, but it had never been applied in the framework of <sup>13</sup>C MFA. We have termed this new approach parsimonious <sup>13</sup>C MFA (p<sup>13</sup>CMFA).

## Introduction

Fluxomics is the omics field that analyses metabolic fluxes (i.e., reaction and transport rates in living cells) which are a close reflection of the metabolic phenotype. As such, quantitative tracking of metabolic fluxes is vital for deciphering the regulation mechanisms of metabolic networks subjected to perturbations, including disease states and drug-induced stress. However, unlike other omics data that can be quantified directly, the fluxome can only be estimated through an indirect interpretation of experimental data[1–3].

There are two main model-based approaches to quantifying metabolic fluxes, Flux Balance Analysis (FBA) and <sup>13</sup>C Metabolic Flux Analysis (<sup>13</sup>C MFA). Both methods use stoichiometric, thermodynamic and experimental constraints to find the range of feasible fluxes across a metabolic network and then find the flux distributions within that space that optimize a given objective function. However, both techniques differ in the type of objective function optimized.

In FBA, the objective function is a set of fluxes to be minimized or maximized. These fluxes must represent a biological objective deemed desirable in the conditions of study (e.g., synthesis of biomass components for proliferating systems)[4]. A significant limitation of FBA is that the choice of objective(s) can significantly influence the predicted flux distributions.

In <sup>13</sup>C MFA, the objective function is to minimize the difference between simulated and measured <sup>13</sup>C enrichment in metabolites [5,6]. <sup>13</sup>C enrichment is quantified in metabolic products and intermediates after incubating samples with metabolic substrates labeled with <sup>13</sup>C (tracers) and provides information about how carbons are redistributed along metabolic pathways[7]. Compared to FBA, <sup>13</sup>C MFA has a greater capacity to elucidate the fluxes of central carbon metabolism. However, <sup>13</sup>C MFA is more complex to solve than FBA due to the non-linear nature of the <sup>13</sup>C MFA objective.

A significant limitation of FBA is that there is generally a wide range of optimal flux distributions[8]. This is not usually the case with <sup>13</sup>C MFA which can generally determine flux distributions with a high degree of accuracy. <sup>13</sup>C MFA achieves this by integrating large sets of measured isotopologue fractions from parallel experiments with tracers optimized for different parts of the network[9–16]. However, when <sup>13</sup>C MFA is used in large metabolic networks and with a limited set of measurements, it can also suffer from the same limitation as FBA and result on a wide interval of flux values for part of the metabolic network[5,17–19].

On FBA, an approach to reduce the range of optimal solutions consists in running a second optimization step on the optimal solution range. One of such methods is parsimonious FBA (pFBA)[20]. This approach, which follows the principle of parsimony or simplicity, consists on finding the optimal value of the primary objective function through FBA and then running a second optimization step where the sum of reaction fluxes is minimized while maintaining the optimal primary objective. The GIMME (and its derivative GIM<sup>3</sup>E) algorithms[21,22] are

based on a similar principle as pFBA. Unlike standard pFBA, where all reactions fluxes are minimized with equal weight, GIMME integrates gene expression data to give greater weight to the minimization of fluxes through reactions catalyzed by lowly expressed enzymes. Different to FBA, for <sup>13</sup>C MFA, there is currently no approach that relies on a second optimization to reduce the solution space when experimental data is insufficient to constrain the system towards a unique solution.

In addition to model-based approaches (e.g., FBA or <sup>13</sup>C MFA), metabolic fluxes can also be analyzed through the direct or semidirect interpretation of <sup>13</sup>C data. This approach primarily consists of predicting the contribution of a labeled substrate to the synthesis of a given metabolite (nutrient contribution) and predicting the relative activity of pathways (pathway activity analysis). Pathway activity analysis assumes that the isotopologue fractions used as a surrogate for the pathways of interest are primarily generated through them. This assumption is generally based on the assertion that the pathways of interest are the most direct way to generate such fractions from the labeled substrate used in the experiment [2,7,23–25]. Unlike <sup>13</sup>C MFA, direct interpretation of <sup>13</sup>C data is generally not able to quantify network-wide flux maps. Instead, it provides a series of qualitative or semiquantitative flux predictions around each analyzed metabolite. Strategies that couple direct interpretation of <sup>13</sup>C data to regression and correlation analyses are widely applied to unveil the effect of an external perturbation, such as a therapeutic intervention, on central carbon metabolism [26–30].

Here we present parsimonious <sup>13</sup>C MFA (p<sup>13</sup>CMFA), a new model-based approach to flux estimation. p<sup>13</sup>CMFA first minimizes the difference between experimental and simulated <sup>13</sup>C enrichment in metabolites (<sup>13</sup>C MFA) and then applies the flux minimization principle to select the best solution among the solutions that fit experimental <sup>13</sup>C data. Hence, p<sup>13</sup>CMFA can be used to select the best flux map in instances where experimental <sup>13</sup>C measurements are not enough to fully constrain the <sup>13</sup>C MFA solution space. Furthermore, the minimization can be weighted by gene expression allowing seamless integration of <sup>13</sup>C with gene expression data (Fig 1).

We have implemented p<sup>13</sup>CMFA in Iso2Flux, our in-house developed isotopic steady-state <sup>13</sup>C MFA software (<https://github.com/cfoguuet/iso2flux/releases/tag/0.7.2>). As a proof of concept, we have applied it to the analysis of the metabolic flux distribution in HUVECs (Human umbilical vein endothelial cells) through the integration of a small set of <sup>13</sup>C enrichment measurements and transcriptomics data. Furthermore, we validated the predictive capacity of p<sup>13</sup>CMFA using data from a published study of HTC116 cells where fluxes had been estimated with a high degree of confidence [14]. Using only a small subset of the measurements from such study, p<sup>13</sup>CMFA was able to achieve significantly better flux predictions than both <sup>13</sup>C MFA and GIMME.

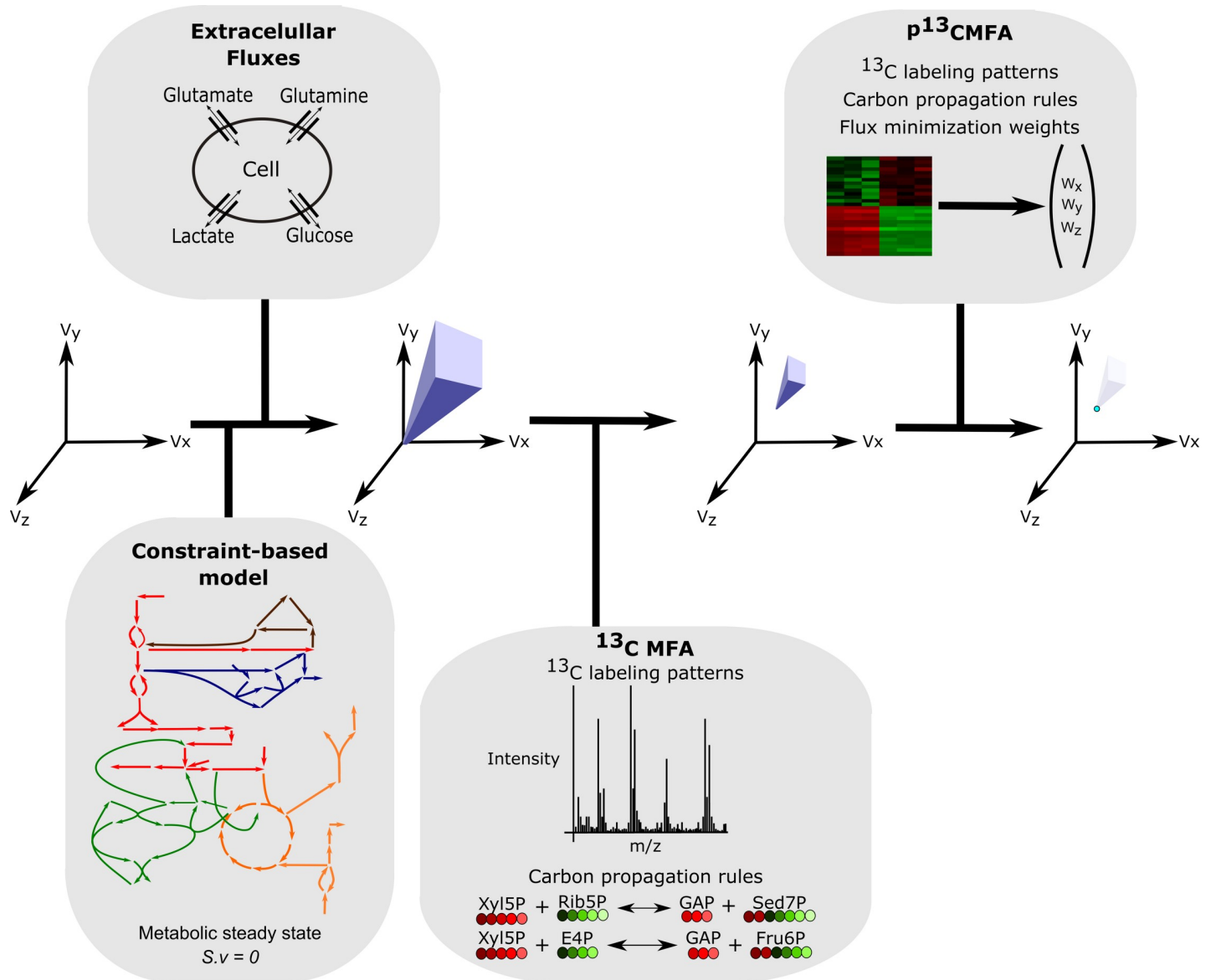
## Results

### Description of the p<sup>13</sup>CMFA approach

p<sup>13</sup>CMFA consists of two consecutive optimizations: first, the optimal solution to the <sup>13</sup>C MFA problem is identified (Eq 1); secondly, the weighted sum of reaction fluxes is minimized within the optimal solution space of <sup>13</sup>C MFA (Eq 2).

The <sup>13</sup>C MFA optimization (Eq 1) identifies the flux distribution that minimizes the difference between measured and simulated isotopologue fractions [5,7]:

$$X_{opt} = \min \sum_j \left( \frac{E_j - Y_j(v)}{\sigma_j} \right)^2 \quad (\text{Eq1})$$



**Fig 1. The conceptual basis of p<sup>13</sup>CMFA.** From an infinite space of flux ( $v$ ) solutions, a space of feasible solutions is obtained through the integration of stoichiometric and thermodynamic constraints (in the form of a constraint-based model) and the measured extracellular fluxes. Applying <sup>13</sup>C MFA to integrate experimental <sup>13</sup>C data can further reduce the solution space to those flux distributions that are consistent with such data. Through flux minimization, p<sup>13</sup>CMFA can identify the optimal flux distribution that lies on the edge of the <sup>13</sup>C MFA solution space. Such minimization can be weighted according to the gene expression evidence for each enzyme.

<https://doi.org/10.1371/journal.pcbi.1007310.g001>

$$\text{Subject to } S.v = 0, \quad lb \leq v \leq ub$$

where,

$v$  is a vector of flux values describing a valid steady-state flux distribution;

$X_{opt}$  is the optimal value of the <sup>13</sup>C MFA objective;

$E_j$  is the experimentally quantified fraction for isotopologue  $j$ ;

$Y_j(v)$  is the simulated isotopologue fraction for isotopologue  $j$  with flux distribution  $v$ . Such simulation is performed by solving a complex non-linear system of equations built around isotopologues balances [1].

$\sigma_j$  is the experimental standard deviation of the measurements of isotopologue  $j$ ;

$S$  is the stoichiometric matrix;

$lb$  and  $ub$  are vectors defining the upper and lower bounds for flux values. Flux bounds can be used to integrate experimental flux measurements;

Either in large metabolic networks or when small sets of <sup>13</sup>C measurements are integrated, the <sup>13</sup>C MFA problem can be undetermined and there can be a wide range of possible solutions. Such indetermination emerges from cycles and alternative pathways in the metabolic network, which lead to many possible flux combinations that can result in the measured <sup>13</sup>C label patterns. Furthermore, many of the <sup>13</sup>C MFA solutions can involve large fluxes through futile cycles, which are usually artifacts of the optimization process as *in vivo* enzyme activities cannot support such large flux values. Therefore, to select the best solution among the many solutions that fit experimental <sup>13</sup>C data, p<sup>13</sup>CMFA runs a second optimization where the weighted sum of fluxes is minimized (Eq 2):

$$\min \sum_i |v_i| \cdot w_i \quad (\text{Eq2})$$

$$\text{subject to } S \cdot v = 0, \quad lb \leq v \leq ub, \quad \sum_j \left( \frac{E_j - Y_j(v)}{\sigma_j} \right)^2 \leq X_{opt} + T$$

where:

$w_i$  is the weight given to the minimization of flux through reaction  $i$ ;

$T$  is the maximum value that the <sup>13</sup>C MFA objective can deviate from its optimal value (primary objective tolerance) when fluxes are minimized;

The difference between the optimal <sup>13</sup>C MFA objective function value and the objective function value when the total reaction flux is minimized can be assumed to follow a Chi<sup>2</sup>-distribution with one degree of freedom. Therefore, setting  $T$  to 3.84 gives a p<sup>13</sup>CMFA solution within the 95% confidence intervals of <sup>13</sup>C MFA[5].

With p<sup>13</sup>CMFA, the activity through cycles is minimized to the minimum amount needed to account for experimental measurements. Furthermore, gene expression measurements can be integrated to give greater weight to the minimization of fluxes through reactions catalyzed by lowly expressed enzymes. Then, in instances where multiple pathways can result in similar label patterns, those pathways with stronger gene expression evidence are selected. Hence, p<sup>13</sup>CMFA reduces the solution space towards a unique solution without requiring a simplification of the metabolic network or additional <sup>13</sup>C measurements (Fig 1).

### Example of p<sup>13</sup>CMFA usage

As an example of a potential application of p<sup>13</sup>CMFA, we applied it to analyze the metabolic flux distribution in HUVECs using a publicly available dataset not large enough to make meaningful flux predictions with conventional <sup>13</sup>C MFA.

In this study, available in the MetaboLights repository[31] (accession number MTBLS412), HUVECs were incubated in the presence of the tracer [1,2-<sup>13</sup>C<sub>2</sub>]-glucose, and the relative abundance of <sup>13</sup>C isotopologues was quantified in glycogen, ribose, lactate, and glutamate. The rates of production/consumption of glucose, glycogen, lactate, glutamate, and glutamine were also quantified. The data were integrated into a stoichiometric model of central metabolism which includes glycolysis, glycogen metabolism, pentose phosphate pathway (PPP), tricarboxylic acid (TCA) cycle, fatty acid synthesis, and energy and redox metabolism (S1 ZIP).

To predict the flux distribution using conventional <sup>13</sup>C MFA, 95% confidence intervals were computed for each predicted flux value. From this analysis, the space of flux solutions

consistent with the measured <sup>13</sup>C enrichment was estimated. The resulting space of solution was still mostly undetermined and, in general, <sup>13</sup>C MFA was unable to significantly constraint the flux ranges emerging from the stoichiometric and thermodynamic constraints and the measured extracellular fluxes (Fig 2, S1 Table). For instance, despite integrating measurements of <sup>13</sup>C enrichment in ribose, it was not possible to conclude whether the oxidative branch of the pentose phosphate pathway contributed more to *de novo* ribose synthesis than the non-oxidative branch or vice versa.

Nevertheless, p<sup>13</sup>CMFA can be applied to select the best solution in the <sup>13</sup>C MFA solution space. With this aim, transcriptomic data taken from the literature[32] were used to add additional penalties to the flux through lowly expressed enzymes. Indeed, by applying p<sup>13</sup>CMFA, we can now conclude that, under the condition of the study, glucose is mostly directed towards lactate production except for a small part going to the TCA cycle through pyruvate dehydrogenase (Fig 2, Fig 3). Glutamine is mainly metabolized to glutamate or directed to glycolysis through the TCA cycle and phosphoenolpyruvate carboxykinase. In the PPP, the non-oxidative branch contributes to roughly 60% of the net ribose synthesis. Only the glycogen phosphorylase/glycogen synthase futile cycle is predicted to be active, while the remaining futile cycles (i.e., the hexokinase/glucose 6-phosphatase, phosphofructokinase/fructose bis-phosphatase, pyruvate carboxylase/phosphoenolpyruvate carboxykinase, and glutaminase/glutamine synthase cycles) are predicted to be inactive. Concerning redox metabolism, most of the reduced NAD<sup>+</sup> (NADH) produced in the mitochondria is exported to the cytosol through the malate-aspartate shuttle, where it is used to reduce pyruvate to lactate.

To evaluate the contribution of <sup>13</sup>C MFA to the p<sup>13</sup>CMFA solution, GIMME (i.e., flux minimization weighted by gene expression without integrating <sup>13</sup>C data) was also performed (Fig 2, S1 Table). Lacking <sup>13</sup>C data, GIMME does not predict any activity in the oxidative branch of the pentose phosphate pathway, nor on the glycogen phosphorylase/glycogen synthase futile cycle. Furthermore, GIMME predicts a significantly larger flux through pyruvate dehydrogenase than p<sup>13</sup>CMFA. Interestingly, p<sup>13</sup>CMFA predicts an increased activity of the TCA cycle compared to the GIMME solution. This increased activity is fueled by alternative sources of acetyl-CoA such as fatty acid oxidation or catabolism of ketogenic amino acids. Hence, p<sup>13</sup>CMFA is able to take advantage of measured <sup>13</sup>C enrichments and predict significantly different flux maps than those derived from flux minimization alone.

## Validation of the p<sup>13</sup>CMFA approach

To validate the p<sup>13</sup>CMFA method, we used data from a metabolic characterization of the colon cancer cell line HCT 116 published by Tarrado-Castellarnau *et al.* [14]. In this study, 25 direct flux measurements and 24 sets of isotopologue fractions, measured after incubation with either [1,2-<sup>13</sup>C<sub>2</sub>]-glucose or [U-<sup>13</sup>C<sub>5</sub>]-glutamine, had been integrated in the framework of <sup>13</sup>C MFA. With such a large set of experimental measurements, <sup>13</sup>C MFA had been able to estimate the flux through 62 reactions with a high degree of accuracy. In the same study, transcriptomics data were also collected.

From this large data set, we selected a partial data set consisting of 7 experimental flux measurements (the rates of uptake/secretion of glucose, lactate, glutamine, glutamate and, oxygen and the rate of protein and glycogen synthesis) and 4 sets of isotopologue fractions (isotopologue fractions in ribose, lactate, glutamate and glycogen measured after incubation with 1,2-<sup>13</sup>C<sub>2</sub>]-glucose). Those are the sets of isotopologues and fluxes that were analyzed in the HUVECs case study with the addition of the rate of protein synthesis and oxygen consumption which Tarrado-Castellarnau *et al.* described as key determinants of the metabolic phenotype of HCT 116 cells. The partial data set was used to apply pFBA, GIMME, <sup>13</sup>C MFA and

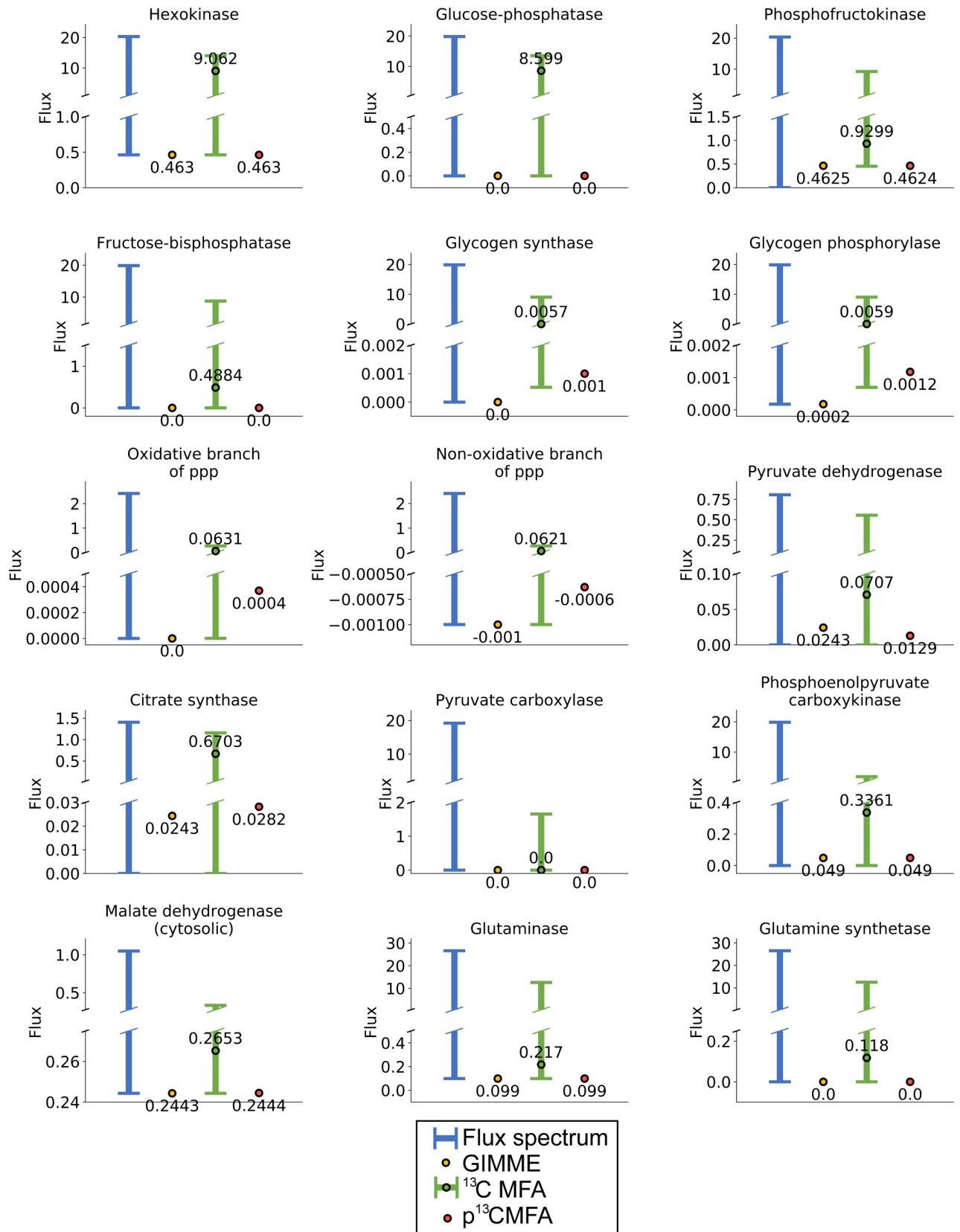


Fig 2. Flux spectrum, GIMME solutions, <sup>13</sup>C MFA flux ranges, and p<sup>13</sup>CMFA solutions for some key reaction fluxes in the HUVECs case study. Flux spectrum represents the feasible flux ranges considering only the stoichiometric and thermodynamic constraints and the measured extracellular fluxes. GIMME flux values are obtained when total reaction flux is minimized weighted by gene expression without integrating <sup>13</sup>C data. For <sup>13</sup>C MFA,

the flux values obtained after the <sup>13</sup>C MFA optimization and the range of the 95% confidence intervals for such values are shown. The p<sup>13</sup>CMFA flux values are obtained when total reaction flux is minimized within the <sup>13</sup>C MFA solution space. Fluxes are expressed in  $\mu\text{mol}\cdot\text{h}^{-1}\cdot\text{million-cells}^{-1}$ .

<https://doi.org/10.1371/journal.pcbi.1007310.g002>

p<sup>13</sup>CMFA in the framework of the metabolic network defined by Tarrado-Castellarnau *et al.* [14] (S2 Zip). p<sup>13</sup>CMFA was applied both with and without integrating gene expression data (p<sup>13</sup>CMFA+ge and p<sup>13</sup>CMFA-ge, respectively). Two complementary metrics, Pearson's correlation and Euclidian distance, were used to evaluate the similarity between the predicted flux distributions and the flux maps estimated by Tarrado-Castellarnau *et al.* using the full dataset [14] (Fig 4, S2 Table). The results show that p<sup>13</sup>CMFA-ge yields a significantly more accurate flux prediction than both pFBA (i.e., flux minimization without integrating <sup>13</sup>C data), and <sup>13</sup>C MFA. Interestingly, while integrating gene expression significantly enhances the accuracy of p<sup>13</sup>CMFA (p<sup>13</sup>CMFA+ge compared to p<sup>13</sup>CMFA-ge), such effect is less marked than the effect of adding gene expression to standard flux minimization (GIMME compared to pFBA). This is due to the fact that p<sup>13</sup>CMFA-ge flux predictions have already a remarkable level of accuracy; hence, less information can be gained by adding transcriptomics data. Nevertheless, even if GIMME achieves flux predictions of similar accuracy to p<sup>13</sup>CMFA-ge, p<sup>13</sup>CMFA+ge results on flux predictions that are significantly more accurate than those obtained with GIMME. Hence, in instances where only a limited number of <sup>13</sup>C measurements are available, p<sup>13</sup>CMFA is a valid method for obtaining accurate flux estimations, regardless of the availability of gene expression data.

## Discussion

<sup>13</sup>C MFA is a well-established technique and has proven to be an extremely valuable tool in quantifying metabolic fluxes [9–18]. However, to fully determine fluxes through a large metabolic network, parallel labeling experiments must be performed and <sup>13</sup>C propagation must be quantified in many metabolites in the network [19]. Indeed, when applying <sup>13</sup>C MFA either with a small set of experimental data or with a large metabolic network, part of the <sup>13</sup>C MFA solution space can be too wide to draw meaningful conclusions about the underlying flux distribution. This solution space can be reduced by removing degrees of freedom from the system, for instance, by removing reactions from the network or making reactions irreversible. However, this can introduce an arbitrary bias in the resulting flux distribution.

Here we describe p<sup>13</sup>CMFA, a new approach for <sup>13</sup>C data integration which can overcome these limitations of <sup>13</sup>C MFA and estimate a realistic solution within an undetermined <sup>13</sup>C MFA solution space. This solution will be the flux distribution within the <sup>13</sup>C MFA solution space that minimizes the weighted sum of reaction fluxes. Thus, it will be the most enzymatically efficient solution. In that regard, p<sup>13</sup>CMFA is partially based on a similar principle as pathway activity analysis (i.e., the assumption that specific fractions of isotopologues are primarily generated through the simplest combinations of pathways). However, unlike pathway activity analysis, p<sup>13</sup>CMFA is able to integrate all quantified isotopologue fractions and flux measurements (e.g. rates of metabolite uptake and secretion) to generate network-wide flux maps consistent with such data. Furthermore, p<sup>13</sup>CMFA is highly flexible; for instance, here we show that it can be used to seamlessly integrate gene expression data by giving higher weight to the minimization of the fluxes through lowly expressed enzymes.

As a proof of concept, we exemplified how p<sup>13</sup>CMFA can be used to estimate flux distributions integrating only limited sets of <sup>13</sup>C measurements in a test case where traditional <sup>13</sup>C MFA was unable to provide a narrow solution space. Furthermore, we demonstrated that, when a limited set of measurements are integrated, p<sup>13</sup>CMFA can yield more accurate flux predictions than both <sup>13</sup>C MFA and GIMME.



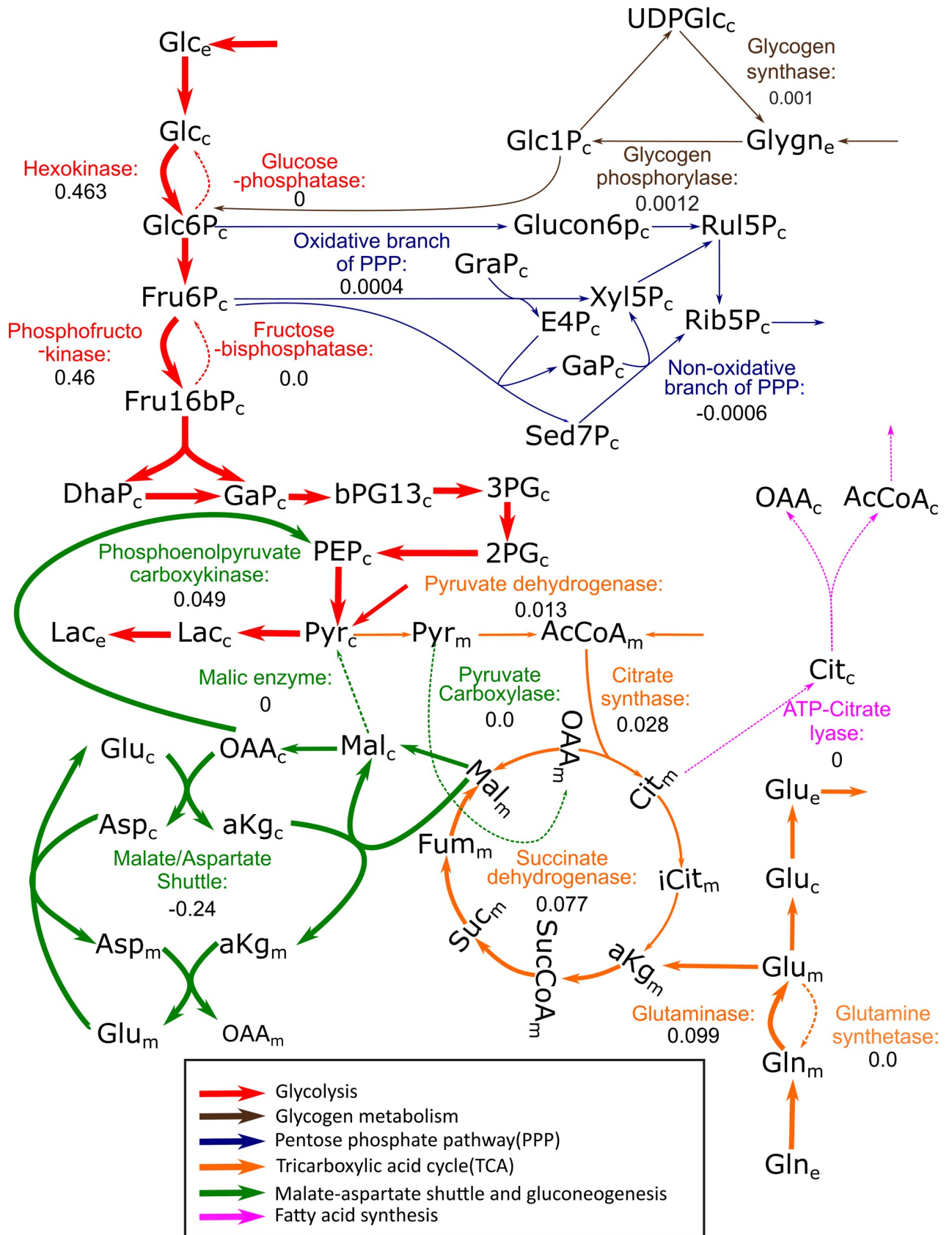


Fig 3. Schematic representation of the central carbon metabolism flux map obtained with p<sup>13</sup>CMFA in the HUVECs case study. Reaction fluxes are indicated for some key reactions in μmol·h<sup>-1</sup>·million-cells<sup>-1</sup>. Arrows indicate net flux direction, and line width is representative of flux magnitude.

Reactions and metabolites of redox and energy metabolism have been omitted from this figure for clarity. 2PG: 2-Phosphoglycerate. 3PG: 3-Phosphoglycerate. AcCoA: Acetyl-CoA. aKG:  $\alpha$ -Ketoglutarate. Asp: Aspartate. bPG13: 1,3-Bisphosphoglycerate. Cit: Citrate. DhaP: Dihydroxyacetone phosphate. Fru16bP: Fructose 1,6-bisphosphate. Fru6P: Fructose 6-phosphate. Fum: Fumarate. GaP: Glyceraldehyde-3-Phosphate. Glc: Glucose. Glc1P: Glucose 1-phosphate. Glc6P: Glucose 6-phosphate. Gln: Glutamine. Glu: Glutamate. Glucon6P: Gluconate 6-phosphate. Glygn: Glycogen. iCit: Isocitrate. Lac: Lactate. Mal: Malate. OAA: Oxaloacetate. PEP: Phosphoenolpyruvate. Pyr: Pyruvate. Rib5P: Ribose 5-phosphate. Rul5P: Ribulose 5-phosphate. Sed7P: Sedoheptulose 7-phosphate. Suc: Succinate. SucCoa: Succinyl-CoA. UDPGlc: Uridine diphosphate glucose. The subscripts *e*, *c*, and *m* denote the extracellular, cytosolic and mitochondrial compartments, respectively.

<https://doi.org/10.1371/journal.pcbi.1007310.g003>

p<sup>13</sup>C MFA does not aim to be a replacement of <sup>13</sup>C MFA; instead, it seeks to supplement it by identifying the more straightforward solution in parts of the network that cannot be uniquely determined. In that regard, it can be used to quantitatively study flux distributions in instances where not enough information can be obtained with conventional <sup>13</sup>C MFA. Nor does it aim to replace the direct interpretation of <sup>13</sup>C data. The latter is still a suitable technique when the goal of the analysis is to compare the relative activity of well-established pathways across conditions or quantify substrate contributions rather than to generate complete flux maps.

<sup>13</sup>C data has been widely used to assist in drug discovery. In this regard, tracer analysis coupled with regression and correlation analyses is frequently used to characterize drug response [26–29]. Such approach uses regression and correlation statistics with binary, numeric and visual analysis to integrate drug dosage, time points, as well as all necessary biological variables in order to diagnose disturbed stable isotope labeled matrices[29]. p<sup>13</sup>CMFA could further expand the role of <sup>13</sup>C in drug discovery by allowing the integration of <sup>13</sup>C and transcriptomic data in the framework of genome-scale metabolic models. In the framework of such models, drug targets are identified by systematically simulating the effect of reactions or genes knock out to cell function[34]. This is usually attained by applying the ROOM[35] or MOMA[36] algorithms, which take a unique flux solution as input (wild-type flux distribution) to predict the most likely effect of a gene KO. Hence, p<sup>13</sup>CMFA results could be potentially used as ROOM/MOMA inputs allowing to take full advantage of the flux information derived from both <sup>13</sup>C and transcriptomics data to predict new drug targets. With atom mappings now available on a genome-scale[37], the main obstacle to applying p<sup>13</sup>CMFA at a genome-scale is the high computational complexity of solving the resulting non-linear problem which increases with the size of the network. Hence, the next challenge for p<sup>13</sup>CMFA will be optimizing its implementation for genome-scale networks.

## Methods

### Flux spectrum

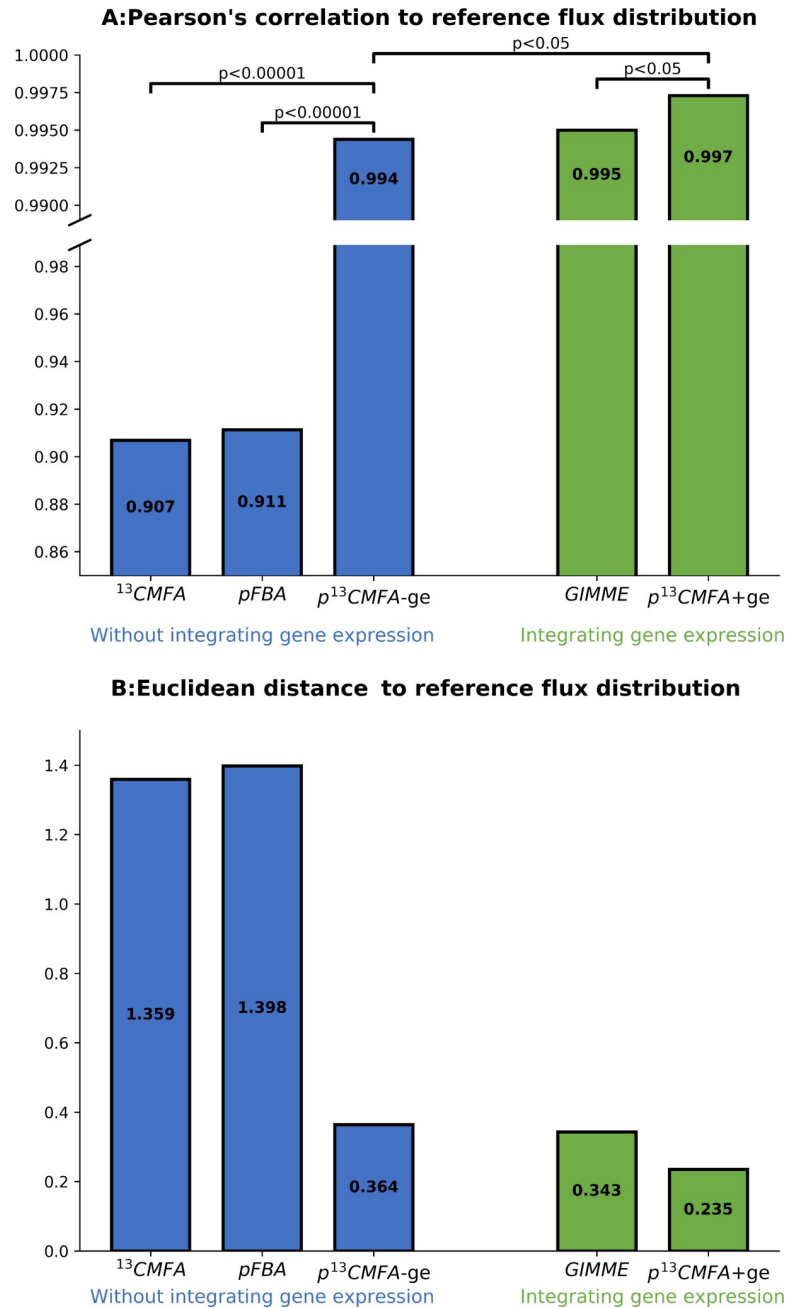
The flux spectrum[38] (i.e., the feasible range of fluxes for a given set of stoichiometric, thermodynamic and flux boundary constraints) was determined using flux variability analysis [8]. Under this approach, each flux is minimized (Eq 3) and maximized (Eq 4) subject to constraints to find the minimum ( $v_{\min_i}^{FS}$ ) and maximum ( $v_{\max_i}^{FS}$ ) feasible values for each flux:

$$v_{\min_i}^{FS} = \min v_i \tag{Eq3}$$

$$\text{subject to } S.v = 0, lb \leq v \leq ub$$

$$v_{\max_i}^{FS} = \max v_i \tag{Eq4}$$

$$\text{subject to } S.v = 0, lb \leq v \leq ub$$



**Fig 4. Comparison of the predictive power of <sup>13</sup>C MFA, pFBA, GIMME, and p<sup>13</sup>CMFA.** A: Pearson's correlation coefficients between the reference flux distribution and the flux maps obtained from <sup>13</sup>C MFA (optimal solution), pFBA, GIMME, and p<sup>13</sup>CMFA. p<sup>13</sup>CMFA was applied both with and without integrating gene expression data (p<sup>13</sup>CMFA+ge and p<sup>13</sup>CMFA-ge, respectively). The statistical significance of the difference between correlation coefficients was evaluated using the Fisher r-to-z transformation[33]. B: Euclidean distances between the reference flux distribution and the flux maps obtained from <sup>13</sup>C MFA (optimal solution), pFBA, GIMME, and p<sup>13</sup>CMFA.

<https://doi.org/10.1371/journal.pcbi.1007310.g004>

### <sup>13</sup>C MFA confidence intervals

The <sup>13</sup>C MFA solution space is estimated by computing the confidence intervals for each flux. Such intervals are obtained by minimizing (Eq 5) and maximizing (Eq 6) each flux subject to

constraints[5].

$$vmin_i = \min v_i \tag{Eq5}$$

$$subject\ to\ S.v = 0, lb \leq v \leq ub, \sum_j \left( \frac{E_j - Y_j(v)}{\sigma_j} \right)^2 \leq X_{opt} + T$$

$$vmax_i = \max v_i \tag{Eq6}$$

$$subject\ to\ S.v = 0, lb \leq v \leq ub, \sum_j \left( \frac{E_j - Y_j(v)}{\sigma_j} \right)^2 \leq X_{opt} + T$$

where,

*vmin<sub>i</sub>*: is the lower bound of the confidence interval for flux *i* with tolerance T;

*vmax<sub>i</sub>*: is the upper bound of the confidence interval for flux *i* with tolerance T;

Provided that the same primary objective tolerance (T) is used in computing both the p<sup>13</sup>CMFA solution and the <sup>13</sup>C MFA confidence intervals, the p<sup>13</sup>CMFA solution will always fall within the boundaries of <sup>13</sup>C MFA confidence intervals (*vmin<sub>i</sub>* ≤ *v<sub>i</sub>* ≤ *vmax<sub>i</sub>*).

### GIMME and pFBA

To apply GIMME and pFBA, the sum of fluxes is minimized subject only to network stoichiometry and flux boundaries (Eq 7).

$$\min \sum_i |v_i| \cdot w_i \tag{Eq7}$$

$$subject\ to\ S.v = 0, lb \leq v \leq ub$$

In GIMME, flux minimization weights are derived from gene expression measurements, whereas in pFBA all reactions are given the same minimization weight[20,22].

### Transcriptomic analysis

Transcriptomic data of HUVECs and HCT 116 cells published by Weigand *et al.*[32] and Tarado-Castellarnau[14] *et al.*, respectively, were obtained from the Gene Expression Omnibus repository[39]. A Robust Multichip Analysis gene-level normalization[40] was performed with the Oligo package for R[41].

Using gene-protein-reaction rules, normalized transcript intensities were mapped to each enzyme-catalyzed reaction or protein-facilitated transport process. The weight given to the minimization of fluxes was assigned according to the following equation:

$$w_i = 1 + \max(Th - ge_i, 0) \tag{Eq8}$$

where,

*ge<sub>i</sub>* is the gene expression value assigned to reaction *i*;

*Th* is the gene expression threshold. Fluxes through reactions with gene expression levels below this threshold are given additional minimization weight;

Using the same criteria as GIM<sup>3</sup>E[22], Th was set at the maximum gene expression value found in the set of genes mapped to the metabolic network (Eq 9):

$$Th = \max(ge) \quad (\text{Eq9})$$

Using this threshold, the information gained from integrating available gene expression measurements is maximized. Other Th values were tested in the validation case study[14] and using the maximum gene expression as the threshold was found to yield the most accurate flux predictions (S3 Table).

### p<sup>13</sup>CMFA implementation

p<sup>13</sup>CMFA was implemented in Iso2Flux, our in-house developed <sup>13</sup>C MFA software (<https://github.com/cfoguete/iso2flux/releases/tag/0.7.2>).

Iso2Flux computes steady-state flux distributions as the product of the null space of the stoichiometric matrix and the vector of free fluxes. Reversible reactions are split into forward and reverse reactions. For each reversible reaction, a turnover variable ( $t_i$ ) is introduced defining the flux that is common to the forward ( $v_i^f$ ) and reverse ( $v_i^r$ ) reactions. These variables are used to assign values to the fluxes of the forward and reverse reactions as a function of the steady-state net flux ( $v_i$ ).

$$v_i^f = t_i + \max(v_i, 0) \quad (\text{Eq10})$$

$$v_i^r = t_i - \min(v_i, 0) \quad (\text{Eq11})$$

Iso2flux uses the Elementary Metabolite Unit (EMU) framework[1] to build the <sup>13</sup>C propagation model. This framework is based on a highly efficient decomposition method that identifies the minimum amount of isotopologue transitions required to simulate the experimentally quantified isotopologues according to the defined carbon propagation rules. The isotopologue transitions are grouped into decoupled systems based on isotopologue size. Balance equations are built around each isotopologue fraction under the assumption of isotopic steady state (S1 Fig). Using the steady-state flux distribution as an input, systems of equations around isotopologues balances are solved sequentially starting with the smallest isotopologue size [1] using the fsolve function of the SciPy library (<https://scipy.org/scipylib/index.html>). Solving such system predicts the isotopologue distribution associated with a given steady-state flux distribution ( $Y_j(v)$ ).

The self-adaptive differential evolution (SADE) algorithm from PyGMO (Python Parallel Global Multiobjective Optimizer, <https://github.com/esa/pagmo2>) was used to find the optimal solution of the <sup>13</sup>C MFA (Eq 1) and p<sup>13</sup>CMFA (Eq 2) problems. SADE was parallelized using the generalized island-model paradigm. Under such implementation, SADE is run in parallel in different CPU processes (islands). After a given number of SADE iterations (generations), the best solutions (individuals) in each SADE process (island) are shared to parallel SADE processes (migrate to adjacent islands). To prevent bias from the starting solutions (starting populations), the islands are seeded through random sampling of all variables. Free fluxes variables are sampled using the optGpSampler implemented into COBRAPy[42,43]. Turnover variables are sampled using the random.uniform function built into python. The algorithm was run with 7 islands, each with a population of 60, and with migrations between islands set to occur every 400 generations. For the analyzed <sup>13</sup>C MFA and p<sup>13</sup>CMFA problems, repeated iterations of the algorithm were shown to reliably converge towards the same minimal objective function value.

## Accommodating large metabolite pools

At the beginning of a <sup>13</sup>C experiment, all internal metabolites are unlabeled (m<sub>0</sub>). Progressively, these products are enriched in <sup>13</sup>C, with the subsequent decrease in m<sub>0</sub>. Isotopic steady state is quickly reached for small pools of metabolites but not necessarily for larger pools such as those of fatty acids, glycogen or metabolites present in large concentrations in the external medium[44]. For these larger pools, unlabeled isotopologues m<sub>0</sub> are oversized and might not quickly decrease to the theoretical value that should be reached at steady-state.

However, it is possible to represent the effect of large pools in the framework of steady-state <sup>13</sup>C MFA through the addition of a virtual reaction. This reaction replaces labeled isotopologues by unlabeled isotopologues in metabolites with large pools. With p<sup>13</sup>CMFA, the flux through this virtual reaction can be minimized. Effectively, this allows correcting steady-state <sup>13</sup>C simulations for large pools while identifying the solutions that require the minimum amount of correction.

## Evaluating the significance of the difference between correlation coefficients

The statistical significance of the difference between correlation coefficients was evaluated using the Fisher r-to-z transformation[33]. Following this approach, Pearson's correlation coefficients (r) can be converted to a z-score (r')

$$r' = \frac{1}{2} \cdot \ln\left(\frac{1+r}{1-r}\right) \quad (\text{Eq12})$$

The variance of z (S<sub>z</sub>) will depend only on the sample size (n):

$$S_z = \sqrt{\frac{1}{n-3}} \quad (\text{Eq13})$$

From Eq 12 and Eq 13, the significance of the difference between two correlation coefficients (r<sub>1</sub> and r<sub>2</sub>) can be evaluated by computing the z score corresponding to such difference (Eq 14) and its associated p-value.

$$Z = \frac{r'_1 - r'_2}{\sqrt{\frac{1}{n_1-3} + \frac{1}{n_2-3}}} \quad (\text{Eq14})$$

## Experimental methods

Human Umbilical Vein Endothelial Cells (HUVECs-pooled, Lonza) were maintained on 1% gelatin-coated flasks at 37°C in a humidified atmosphere of 5% CO<sub>2</sub> and 95% air in MCDB131 (Gibco) medium, supplemented with the recommended quantity of endothelial growth medium (EGM) SingleQuots (Lonza), 10% fetal bovine serum (FBS) (Gibco), 2 mM glutamine (Gibco) and 0.1% Streptomycin (100 µg/mL)/Penicillin (100 units/mL) (S/P) (Gibco). 1 × 10<sup>6</sup> HUVECs were seeded in 1% gelatin-coated cell culture plates for 6h, and then the maintenance medium was replaced with the MCDB131 basal medium, supplemented with 2% FBS, 2 mM glutamine and 0.1% S/P and cells were incubated overnight for nutrient deprivation. After nutrient deprivation, the medium was replaced with a restricted medium containing MCDB131 medium supplemented with 2% FBS, 2 mM glutamine and 0.1% S/P with 10 mM of 50% [1,2-<sup>13</sup>C<sub>2</sub>]-glucose (Sigma-Aldrich) and cells were incubated for 40h in a humidified atmosphere with 5% CO<sub>2</sub> and 1% O<sub>2</sub> at 37°C. Both at the beginning (t = 0h) and the end (t = 40h) of incubation, media and pellets were collected. On the one hand, media and cell

pellets were used for analyzing isotopologue abundances for glucose, lactate, glutamate, RNA ribose and glycogen. Raw data are publicly available in the MetaboLights repository at <http://www.ebi.ac.uk/metabolights> [31], with accession number MTBLS412. Isolation, derivatization and analysis details are described in MetaboLights. Glucose, lactate, glutamate, and glutamine concentrations were determined in media samples for estimation of secretion or uptake rates of these metabolites using spectrophotometric methods[45]. Also, the net rate of glycogen re-utilization into glucose was estimated by quantifying glycogen content at initial and final time points using [U-<sup>13</sup>C-D<sub>7</sub>]-glucose as recovery standard[46]. All biochemical data were normalized by cell number, and by incubation time (h). The resulting rates—expressed in micromoles of metabolite consumed/produced/transformed per hour per million cells ( $\mu\text{mol}\cdot\text{h}^{-1}\cdot\text{million-cells}^{-1}$ )—were 0.463, 0.099, 0.050 and 1.169 for glucose uptake, glutamine uptake, glutamate secretion, and lactate secretion, respectively, and a net transformation of glycogen of 0.000175.

## Supporting information

**S1 Fig. Example of isotopologue balance equations in a toy metabolic network.** In this toy metabolic network, two mono-carbon metabolites ( $C_a$  and  $C_b$ ) are condensed into a bi-carbon metabolite ( $C_a-C_b$ ) through a reaction with a flux  $v_1$ . Metabolite  $C_a-C_b$  is removed from the system at a rate of  $v_2$ . For each metabolite, isotopologue fractions ( $M_x$ ) are defined as the relative abundance of the metabolite with  $x$  number of <sup>13</sup>C substitutions. Isotopologue balances for metabolite  $C_a-C_b$  are indicated. Under the assumption of isotopic steady state (i.e., isotopologue fractions are constant in time) and given  $v_1$  and  $v_2$ , and a set of isotopologue fractions for  $C_a$  and  $C_b$  (assumed a constant input), the system can be solved to identify the steady-state isotopologue fractions for metabolite  $C_a-C_b$ .  
(TIF)

**S1 Table. Flux spectrum, GIMME, <sup>13</sup>C MFA and p<sup>13</sup>CMFA flux solutions for all net reaction fluxes in the HUVECs case study.** Fluxes are expressed in  $\mu\text{mol}\cdot\text{h}^{-1}\cdot\text{million-cells}^{-1}$ .  
(XLSX)

**S2 Table. Comparison between the reference flux map in HCT 116 cells and the flux maps computed from the partial data set using <sup>13</sup>C MFA, pFBA, GIMME, and p<sup>13</sup>CMFA.** Fluxes are indicated in  $\mu\text{mol}\cdot\text{h}^{-1}\cdot\text{million-cells}^{-1}$ .  
(XLSX)

**S3 Table. Comparison between the reference flux map in HCT 116 cells and the flux maps computed from the partial data set with p<sup>13</sup>CMFA using different gene expression percentiles as thresholds for adding additional weight to flux minimization.** Fluxes are indicated in  $\mu\text{mol}\cdot\text{h}^{-1}\cdot\text{million-cells}^{-1}$ .  
(XLSX)

**S1 ZIP. Files describing the metabolic network, carbon propagation rules, and experimental data used for the HUVECs case study.** The files are inputs for running p<sup>13</sup>CMFA on Iso2Flux.  
(ZIP)

**S2 ZIP. Files describing the metabolic network, carbon propagation rules, and experimental data used for the HCT 116 cells case study.** The files are inputs for running p<sup>13</sup>CMFA on Iso2Flux.  
(ZIP)

## Author Contributions

**Conceptualization:** Carles Foguet, Vitaly A. Selivanov, Pedro de Atauri, Marta Cascante.

**Data curation:** Anusha Jayaraman, Silvia Marin, Vitaly A. Selivanov, Pablo Moreno, Ramon Messeguer.

**Formal analysis:** Carles Foguet.

**Funding acquisition:** Marta Cascante.

**Investigation:** Anusha Jayaraman, Silvia Marin, Ramon Messeguer.

**Methodology:** Carles Foguet.

**Project administration:** Marta Cascante.

**Software:** Carles Foguet, Vitaly A. Selivanov, Pablo Moreno, Pedro de Atauri.

**Supervision:** Pedro de Atauri, Marta Cascante.

**Writing – original draft:** Carles Foguet, Silvia Marin, Pedro de Atauri.

**Writing – review & editing:** Anusha Jayaraman, Silvia Marin, Vitaly A. Selivanov, Pablo Moreno, Ramon Messeguer, Pedro de Atauri, Marta Cascante.

## References

1. Antoniewicz MR, Kelleher JK, Stephanopoulos G. Elementary metabolite units (EMU): a novel framework for modeling isotopic distributions. *Metab Eng.* 2007; 9: 68–86. <https://doi.org/10.1016/j.ymben.2006.09.001> PMID: 17088092
2. Buescher JM, Antoniewicz MR, Boros LG, Burgess SC, Brunengraber H, Clish CB, et al. A roadmap for interpreting (13)C metabolite labeling patterns from cells. *Curr Opin Biotechnol.* 2015; 34: 189–201. <https://doi.org/10.1016/j.copbio.2015.02.003> PMID: 25731751
3. Zamboni N, Saghatelian A, Patti GJ. Defining the metabolome: size, flux, and regulation. *Mol Cell.* 2015; 58: 699–706. <https://doi.org/10.1016/j.molcel.2015.04.021> PMID: 26000853
4. Orth JD, Thiele I, Palsson BØ. What is flux balance analysis? *Nat Biotechnol.* 2010; 28: 245–8. <https://doi.org/10.1038/nbt.1614> PMID: 20212490
5. Antoniewicz MR, Kelleher JK, Stephanopoulos G. Determination of confidence intervals of metabolic fluxes estimated from stable isotope measurements. *Metab Eng.* 2006; 8: 324–337. <https://doi.org/10.1016/j.ymben.2006.01.004> PMID: 16631402
6. Niedenführ S, Wiechert W, Nöh K. How to measure metabolic fluxes: A taxonomic guide for 13C fluxomics. *Curr Opin Biotechnol.* 2015; 34: 82–90. <https://doi.org/10.1016/j.copbio.2014.12.003> PMID: 25531408
7. Balcells C, Foguet C, Tarragó-Celada J, de Atauri P, Marin S, Cascante M. Tracing metabolic fluxes using mass spectrometry: Stable isotope-resolved metabolomics in health and disease. *TrAC Trends Anal Chem.* 2019; <https://doi.org/10.1016/j.trac.2018.12.025>
8. Gudmundsson S, Thiele I. Computationally efficient flux variability analysis. *BMC Bioinformatics.* 2010; 11: 489. <https://doi.org/10.1186/1471-2105-11-489> PMID: 20920235
9. Niklas J, Sandig V, Heinzle E. Metabolite channeling and compartmentation in the human cell line AGE1.HN determined by <sup>13</sup>C labeling experiments and 13C metabolic flux analysis. *J Biosci Bioeng.* 2011; 112: 616–623. <https://doi.org/10.1016/j.jbiosc.2011.07.021> PMID: 21865082
10. Walther JL, Metallo CM, Zhang J, Stephanopoulos G. Optimization of 13C isotopic tracers for metabolic flux analysis in mammalian cells. *Metab Eng.* 2012; 14: 162–171. <https://doi.org/10.1016/j.ymben.2011.12.004> PMID: 22198197
11. Metallo CM, Gameiro PA, Bell EL, Mattaini KR, Yang J, Hiller K, et al. Reductive glutamine metabolism by IDH1 mediates lipogenesis under hypoxia. *Nature.* 2012; 481: 380–384. <https://doi.org/10.1038/nature10602> PMID: 22101433
12. Grassian AR, Parker SJ, Davidson SM, Divakaruni AS, Green CR, Zhang X, et al. IDH1 mutations alter citric acid cycle metabolism and increase dependence on oxidative mitochondrial metabolism. *Cancer Res.* 2014; 74: 3317–3331. <https://doi.org/10.1158/0008-5472.CAN-14-0772-T> PMID: 24755473



13. Crown SB, Kelleher JK, Rouf R, Muoio DM, Antoniewicz MR. Comprehensive metabolic modeling of multiple <sup>13</sup>C-isotopomer data sets to study metabolism in perfused working hearts. *Am J Physiol Heart Circ Physiol*. 2016; 311: H881–H891. <https://doi.org/10.1152/ajpheart.00428.2016> PMID: 27496880
14. Tarrado-Castellarnau M, de Atauri P, Tarragó-Celada J, Perarnau J, Yuneva M, Thomson TM, et al. De novo MYC addiction as an adaptive response of cancer cells to CDK4/6 inhibition. *Mol Syst Biol*. 2017; 13: 940. <https://doi.org/10.15252/msb.20167321> PMID: 28978620
15. Carinhas N, Koshkin A, Pais DAM, Alves PM, Teixeira AP. <sup>13</sup>C-metabolic flux analysis of human adenovirus infection: Implications for viral vector production. *Biotechnol Bioeng*. 2017; 114: 195–207. <https://doi.org/10.1002/bit.26063> PMID: 27477740
16. DeWaal D, Nogueira V, Terry AR, Patra KC, Jeon S-M, Guzman G, et al. Hexokinase-2 depletion inhibits glycolysis and induces oxidative phosphorylation in hepatocellular carcinoma and sensitizes to metformin. *Nat Commun*. 2018; 9: 446. <https://doi.org/10.1038/s41467-017-02733-4> PMID: 29386513
17. Gopalakrishnan S, Maranas CD. <sup>13</sup>C metabolic flux analysis at a genome-scale. *Metab Eng*. 2015; 32: 12–22. <https://doi.org/10.1016/j.ymben.2015.08.006> PMID: 26358840
18. Foguet C, Marin S, Selivanov VA, Fanchon E, Lee W-NP, Guinovart JJ, et al. HepatoDyn: A Dynamic Model of Hepatocyte Metabolism That Integrates <sup>13</sup>C Isotopomer Data. Lewis NE, editor. *PLoS Comput Biol*. 2016; 12: e1004899. <https://doi.org/10.1371/journal.pcbi.1004899> PMID: 27124774
19. Crown SB, Long CP, Antoniewicz MR. Optimal tracers for parallel labeling experiments and <sup>13</sup>C metabolic flux analysis: A new precision and synergy scoring system. *Metab Eng*. 2016; 38: 10–18. <https://doi.org/10.1016/j.ymben.2016.06.001> PMID: 27267409
20. 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
21. 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
22. Schmidt BJ, Ebrahim A, Metz TO, Adkins JN, Palsson BØ, Hyduke DR. GIM3E: condition-specific models of cellular metabolism developed from metabolomics and expression data. *Bioinformatics*. 2013; 29: 2900–8. <https://doi.org/10.1093/bioinformatics/btt493> PMID: 23975765
23. Antoniewicz MR. A guide to <sup>13</sup>C metabolic flux analysis for the cancer biologist. *Exp Mol Med*. 2018; 50: 19. <https://doi.org/10.1038/s12276-018-0060-y> PMID: 29657327
24. Dong W, Keibler MA, Stephanopoulos G. Review of metabolic pathways activated in cancer cells as determined through isotopic labeling and network analysis. *Metabolic Engineering*. 2017. pp. 113–124. <https://doi.org/10.1016/j.ymben.2017.02.002> PMID: 28192215
25. Bruntz RC, Lane AN, Higashi RM, Fan TWM. Exploring cancer metabolism using Stable isotope-resolved metabolomics (SIRM). *J Biol Chem*. 2017; 292: 11601–11609. <https://doi.org/10.1074/jbc.R117.776054> PMID: 28592486
26. Harrigan GG, Colca J, Szalma S, Boros LG. PNU-91325 increases fatty acid synthesis from glucose and mitochondrial long chain fatty acid degradation: A comparative tracer-based metabolomics study with rosiglitazone and pioglitazone in HepG2 cells. *Metabolomics*. 2006; 2: 21–29. <https://doi.org/10.1007/s11306-006-0015-5> PMID: 24489530
27. Beger RD, Hansen DK, Schnackenberg LK, Cross BM, Fatollahi JJ, Lagunero FT, et al. Single valproic acid treatment inhibits glycogen and RNA ribose turnover while disrupting glucose-derived cholesterol synthesis in liver as revealed by the [U-<sup>13</sup>C<sub>6</sub>]-d-glucose tracer in mice. *Metabolomics*. 2009; 5: 336–345. <https://doi.org/10.1007/s11306-009-0159-1> PMID: 19718458
28. Cantoria MJ, Boros LG, Meuillet EJ. Contextual inhibition of fatty acid synthesis by metformin involves glucose-derived acetyl-CoA and cholesterol in pancreatic tumor cells. *Metabolomics*. 2014; 10: 91–104. <https://doi.org/10.1007/s11306-013-0555-4> PMID: 24482631
29. Boros LG, Beger RD, Meuillet EJ, Colca JR, Szalma S, Thompson PA, et al. Targeted <sup>13</sup>C-Labeled Tracer Fate Associations for Drug Efficacy Testing in Cancer. *Tumor Cell Metabolism*. Vienna: Springer Vienna; 2015. pp. 349–372. [https://doi.org/10.1007/978-3-7091-1824-5\\_15](https://doi.org/10.1007/978-3-7091-1824-5_15)
30. Varma V, Boros LG, Nolen GT, Chang CW, Wabitsch M, Beger RD, et al. Metabolic fate of fructose in human adipocytes: a targeted <sup>13</sup>C tracer fate association study. *Metabolomics*. 2015; 11: 529–544. <https://doi.org/10.1007/s11306-014-0716-0> PMID: 25972768
31. Haug K, Salek RM, Conesa P, Hastings J, de Matos P, Rijnbeek M, et al. MetaboLights—an open-access general-purpose repository for metabolomics studies and associated meta-data. *Nucleic Acids Res*. 2013; 41: D781–6. <https://doi.org/10.1093/nar/gks1004> PMID: 23109552

32. Weigand JE, Boeckel J-N, Gellert P, Dimmeler S. Hypoxia-Induced Alternative Splicing in Endothelial Cells. Preiss T, editor. *PLoS One*. 2012; 7: e42697. <https://doi.org/10.1371/journal.pone.0042697> PMID: 22876330
33. Weaver B, Wuensch KL. SPSS and SAS programs for comparing Pearson correlations and OLS regression coefficients. *Behav Res Methods*. 2013; 45: 880–895. <https://doi.org/10.3758/s13428-012-0289-7> PMID: 23344734
34. de Mas IM, Aguilar E, Jayaraman A, Polat IH, Martín-Bernabé A, Bharat R, et al. Cancer cell metabolism as new targets for novel designed therapies. *Future Med Chem*. 2014; 6: 1791–1810. <https://doi.org/10.4155/fmc.14.119> PMID: 25574531
35. Shlomi T, Berkman O, Ruppin E. Regulatory on/off minimization of metabolic flux changes after genetic perturbations. *Proc Natl Acad Sci U S A*. 2005; 102: 7695–700. <https://doi.org/10.1073/pnas.0406346102> PMID: 15897462
36. Segre D, Vitkup D, Church GM. Analysis of optimality in natural and perturbed metabolic networks. *Proc Natl Acad Sci*. 2002; 99: 15112–15117. <https://doi.org/10.1073/pnas.232349399> PMID: 12415116
37. Brunk E, Sahoo S, Zielinski DC, Altunkaya A, Dräger A, Mih N, et al. Recon3D enables a three-dimensional view of gene variation in human metabolism. *Nat Biotechnol*. 2018; 36: 272–281. <https://doi.org/10.1038/nbt.4072> PMID: 29457794
38. Llaneras F, Picó J. An interval approach for dealing with flux distributions and elementary modes activity patterns. *J Theor Biol*. 2007; 246: 290–308. <https://doi.org/10.1016/j.jtbi.2006.12.029> PMID: 17292923
39. Edgar R. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res*. 2002; 30: 207–210. <https://doi.org/10.1093/nar/30.1.207> PMID: 11752295
40. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics*. 2003; 4: 249–64. <https://doi.org/10.1093/biostatistics/4.2.249> PMID: 12925520
41. Carvalho BS, Irizarry RA. A framework for oligonucleotide microarray preprocessing. *Bioinformatics*. 2010; 26: 2363–2367. <https://doi.org/10.1093/bioinformatics/btq431> PMID: 20688976
42. Ebrahim A, Lerman JA, Palsson BO, Hyduke DR. COBRApy: COstraints-Based Reconstruction and Analysis for Python. *BMC Syst Biol*. 2013; 7: 74. <https://doi.org/10.1186/1752-0509-7-74> PMID: 23927696
43. Megchelenbrink W, Huynen M, Marchiori E. optGpSampler: An Improved Tool for Uniformly Sampling the Solution-Space of Genome-Scale Metabolic Networks. Rogers S, editor. *PLoS One*. 2014; 9: e86587. <https://doi.org/10.1371/journal.pone.0086587> PMID: 24551039
44. Selivanov VA, Meshalkina LE, Solovjeva ON, Kuchel PW, Ramos-Montoya A, Kochetov GA, et al. Rapid simulation and analysis of isotopomer distributions using constraints based on enzyme mechanisms: an example from HT29 cancer cells. *Bioinformatics*. 2005; 21: 3558–64. <https://doi.org/10.1093/bioinformatics/bti573> PMID: 16002431
45. Benito A, Polat IH, Noé V, Ciudad CJ, Marin S, Cascante M. Glucose-6-phosphate dehydrogenase and transketolase modulate breast cancer cell metabolic reprogramming and correlate with poor patient outcome. *Oncotarget*. 2017; 8: 106693–106706. <https://doi.org/10.18632/oncotarget.21601> PMID: 29290982
46. Vizán P, Sánchez-Tena S, Alcarraz-Vizán G, Soler M, Messeguer R, Pujol MD, et al. Characterization of the metabolic changes underlying growth factor angiogenic activation: identification of new potential therapeutic targets. *Carcinogenesis*. 2009; 30: 946–52. <https://doi.org/10.1093/carcin/bgp083> PMID: 19369582