Genome scale metabolic network modelling for metabolic profile predictions

Juliette Cooke, Maxime Delmas, Cecilia Wieder, Pablo Rodríguez Mier, Clément Frainay, Florence Vinson, Timothy Ebbels, Nathalie Poupin, Fabien Jourdan

PLOS COMPUTATIONAL BIOLOGY

RESEARCH ARTICLE

Genome scale metabolic network modelling for metabolic profile predictions

Juliette Cooke[1]*, Maxime Delmas[1,2¤a], Cecilia Wieder[3], Pablo Rodríguez Mier[1,4¤b], Clément Frainay[1], Florence Vinson[1,5], Timothy Ebbels[3], Nathalie Poupin[1], Fabien Jourdan[1,5]

1 Toxalim (Research Centre in Food Toxicology), Université de Toulouse, INRAE, ENVT, INP-Purpan, UPS, Toulouse, France, **2** Idiap Research Institute, Martigny, Switzerland, **3** Section of Bioinformatics, Division of Systems Medicine, Department of Metabolism, Digestion, and Reproduction, Faculty of Medicine, Imperial College London, London, United Kingdom, **4** Heidelberg University, Faculty of Medicine, Heidelberg University Hospital, Institute for Computational Biomedicine, Bioquant, Heidelberg, Germany, **5** MetaToul-MetaboHUB, National Infrastructure of Metabolomics and Fluxomics, Toulouse, France

¤a Current address: Idiap Research Institute, Martigny, Switzerland
¤b Current address: Heidelberg University, Faculty of Medicine, Heidelberg University Hospital, Institute for Computational Biomedicine, Bioquant, Heidelberg, Germany
* juliette.cooke@inrae.fr

Citation: Cooke J, Delmas M, Wieder C, Rodríguez Mier P, Frainay C, Vinson F, et al. (2024) Genome scale metabolic network modelling for metabolic profile predictions. PLoS Comput Biol 20(2): e1011381. https://doi.org/10.1371/journal.pcbi.1011381

Editor: Anders Wallqvist, US Army Medical Research and Materiel Command: US Army Medical Research and Development Command, UNITED STATES

Received: July 25, 2023

Accepted: January 25, 2024

Published: February 22, 2024

Data Availability Statement: The mGWAS data can be found at https://static-content.springer.com/esm/art%3A10.1038%2Fnature10354/MediaObjects/41586_2011_BFnature10354_MOESM279_ESM.pdf The IEM data is from: https://www.rsc.org/suppdata/mb/c2/c2mb25075f/c2mb25075f.xlsx The metabolic networks used in this paper are available at the following links: HumanGEM: https://github.com/SysBioChalmers/Human-GEM/releases/tag/v1.14.0 Recon2: https://

Abstract

Metabolic profiling (metabolomics) aims at measuring small molecules (metabolites) in complex samples like blood or urine for human health studies. While biomarker-based assessment often relies on a single molecule, metabolic profiling combines several metabolites to create a more complex and more specific fingerprint of the disease. However, in contrast to genomics, there is no unique metabolomics setup able to measure the entire metabolome. This challenge leads to tedious and resource consuming preliminary studies to be able to design the right metabolomics experiment. In that context, computer assisted metabolic profiling can be of strong added value to design metabolomics studies more quickly and efficiently. We propose a constraint-based modelling approach which predicts *in silico* profiles of metabolites that are more likely to be differentially abundant under a given metabolic perturbation (e.g. due to a genetic disease), using flux simulation. In genome-scale metabolic networks, the fluxes of exchange reactions, also known as the flow of metabolites through their external transport reactions, can be simulated and compared between control and disease conditions in order to calculate changes in metabolite import and export. These import/export flux differences would be expected to induce changes in circulating biofluid levels of those metabolites, which can then be interpreted as potential biomarkers or metabolites of interest. In this study, we present SAMBA (SAMpling Biomarker Analysis), an approach which simulates fluxes in exchange reactions following a metabolic perturbation using random sampling, compares the simulated flux distributions between the baseline and modulated conditions, and ranks predicted differentially exchanged metabolites as potential biomarkers for the perturbation. We show that there is a good fit between simulated metabolic exchange profiles and experimental differential metabolites detected in plasma, such as patient data from the disease database OMIM, and metabolic trait-SNP associations found in mGWAS studies. These biomarker recommendations can provide insight into the

github.com/opencobra/COBRA.papers/tree/master/2013_Recon2 The data produced using SAMBA and used to make Figures 4, 5 and 6 are at the following address: https://doi.org/10.5281/zenodo.8414404. The code developed for this study is available at https://forgemia.inra.fr/metexplore/cbm/samba-project/samba and at https://doi.org/10.5281/zenodo.8369624.

Funding: JC is supported by a state-funded PhD contract (Ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation, Grant 2020-2), and by MetaboHUB (Grant 11-INBS-0010). FJ is supported by the Agence Nationale de la Recherche (ANR) through MetaboHUB (Grant ANR-INBS-0010), as well as the MetClassNet project (ANR-19-CE45-0021 and Deutsche Forschungsgemeinschaft DFG: 431572533). 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.

underlying mechanism or metabolic pathway perturbation lying behind observed metabolite differential abundances, and suggest new metabolites as potential avenues for further experimental analyses.

Author summary

Associating diseases and other metabolic disruptions with physiological markers is key for diagnostic and personalised medicine. These markers can be metabolites—small molecules involved in every living being's metabolism, and can be measured in biofluids such as blood or urine using metabolic profiling (metabolomics). Nevertheless, this experimental metabolomics design needs to be tailor made for each disease to ensure that most relevant metabolites will be detected. The selection of metabolites to analyse for future experiments can be time-consuming and expensive. In this paper, we build upon an existing computational method for simulating metabolite changes in a human model. This provides a prediction of the change in biofluid abundance of every known metabolite involved in human metabolism in a potentially large number of metabolic situations. The newly introduced method produces a change score and a rank for each metabolite in each condition. We show the strong potential of the approach by comparing predictions with experimental results.

Introduction

Molecular biomarkers are measurable molecules directly or indirectly related to alterations of a certain physiological state. They can be used as diagnostic indicators, and can be measured in order to detect the presence or severity of many different diseases [1]. Among them, metabolic biomarkers are small molecules (metabolites) whose concentrations differ from a healthy state due to changes in the organism's or tissue's metabolism. In clinical settings, these biomarkers are traditionally detected using targeted bioassays which result in measurements for a small number of well characterised diagnostic metabolites (Fig 1A), such as glucose for diabetic disorders.

A more holistic approach to measuring metabolites experimentally is via metabolic profiling, also called metabolomics [2]. In this paradigm, the detection of metabolic modulations relies on a broader panel of molecules (the metabolic profile) than targeted biochemical assays, hence resulting in an increased predictive and explanatory potential (Fig 1A). The most common analytical techniques used to measure metabolites are Nuclear Magnetic Resonance (NMR) and Mass Spectrometry (MS) coupled with separation techniques such as Liquid or Gas Chromatography (LC or GC).

Both NMR and MS are unable to completely cover the entire metabolome (see [3] for metabolic coverage assessment of MS data). In fact, both analytical techniques have the ability to detect a portion of the metabolome depending on the physico-chemical properties of molecules present in the sample (e.g. polarity) [4]. Hence, it is important to select the right analytical method (e.g. separation, ionisation technique), or combination of methods [5], to ensure that measured metabolites will be relevant for a specific disease. This experimental design step can be time and resource consuming for experts.

Even if a technique is in principle able to measure a relevant metabolite, it still requires an identification step to ensure the measured feature is a metabolite of interest. In MS for

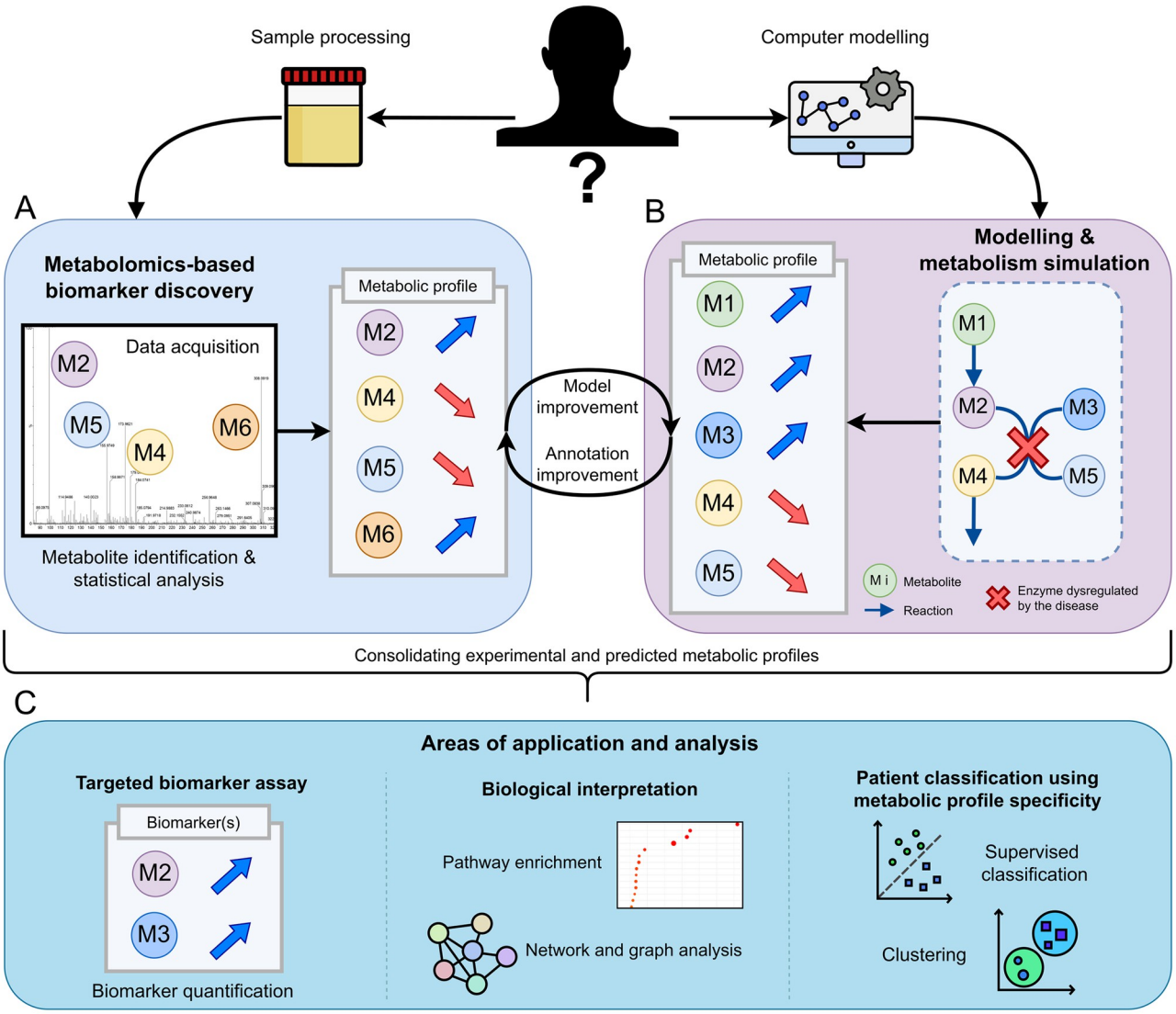

Fig 1. Combining metabolomics profiling with simulations of metabolism. A: Experimental-based biomarker discovery produces metabolic profiles containing detected and annotated metabolites along with a concentration or fold-change value. B: Metabolic disruptions can be modelled to simulate metabolic profiles similar to those generated using metabolomics. C: By combining information from both types of metabolic profiles and improving both experimental annotation and *in silico* models, various approaches can be used to improve our knowledge of given biomarker sets, affected pathways, and patient disease classes.

https://doi.org/10.1371/journal.pcbi.1011381.g001

instance, the exact mass is not sufficient to confidently identify a metabolite. It requires confronting measurements with at least an orthogonal approach and confirming the identification by comparing the observed spectra with the one of the pure molecule (reference spectra of a standard, level 1 identification [6]). However, the reference spectra of the metabolite might not be available, which requires the laboratory to buy and inject new standards into mass spectrometers.

Computational solutions can be used to fill the gaps in experimental observations by providing a recommendation list of metabolites which are expected to be altered in the studied condition. This list can be used at different steps of the metabolomics process. Firstly, it can be used upstream of the experiment to select "the most suitable" analytical platform and set-up

(e.g. if mostly lipids are expected to be affected, a lipidomics setup will be favoured). The benefit of predicting profiles is also downstream of the analysis for annotation purposes. Raw data can be mined to look directly at the predicted metabolites which will accelerate the process of identification. This prediction can be used to select the right set of standards to be analysed to reach level 1 annotation. Finally, SAMBA can be of added value to fill gaps in biochemical interpretation by suggesting metabolites (and related pathways) which could be of interest for the biological comprehension of a disease.

A well suited class of methods for global metabolic modelling is Constraint-Based Modelling (CBM). CBM is a modelling framework which uses genome-scale metabolic networks, under the formalism of a stoichiometric matrix, to compute steady-state metabolic fluxes (the flow of metabolites) through biochemical reactions [7]. These networks aim to encompass all known metabolic genes, reactions and metabolites as well as the interactions between them for a given organism [8–10]. They can also be built to model a specific tissue or cell type [11], or even multiple tissues linked together [12].

CBM can be used to predict fluxes at steady-state under various conditions. This is achieved by defining metabolism as a system of linear mass balance equations, composed of the reaction flux vectors for each metabolite. These fluxes exist under defined flux constraints (setting upper and lower bounds) to model different metabolic states and reaction directionality. Controlling these bounds can for instance be used to simulate the complete knock-out (KO) (Fig 1B) of one or multiple gene(s) or reaction(s) (like for genetic diseases), or the reduction of the flux through a reaction or multiple reactions (knock-down), representing reduced enzyme activity due to some effect of treatment, regulation, or exposure to xenobiotics.

By using an organism-specific metabolic network in conjunction with a metabolic disruption, this CBM methodology can be used to predict which metabolites will be more or less released in biofluids. Indeed, in metabolic networks, some metabolites can be transported in and out from the internal compartment (cell or tissue) to the external compartment (e.g. biofluid or cell culture medium) usually using a single specific exchange reaction. For the *in silico* prediction of biomarkers these exchange reactions can be used to model the in/out flux of metabolites between tissues and circulating biofluids like blood or urine. This is why, in the context of metabolic profile prediction, the focus must be on these specific exchange reactions from the metabolic network in order to predict the equivalent of "biofluid metabolite level changes" using CBM. A break-down of this methodology is shown in Fig 2, using a simple metabolic network to compare flux simulations in healthy and disease conditions, and resulting in a ranked list of metabolites which change the most between the two conditions.

The overall methodology introduced in Shlomi *et al.* [13] for biomarker prediction consists in comparing the flux of exchange reactions between two conditions: one corresponding to a standard metabolic state (wild type) and one to a disease state. The wild type (WT) model is obtained by running the flux simulation on the chosen genome-scale metabolic network with default constraints, forcing the reaction of interest to have a non-zero flux (see Methods for more details). Then, the disease or mutant state model (MUT) is obtained by simulating a knock-out or knock down state using a list of genes or reactions to constrain (e.g. imposing null flux for knocked out reactions). Flux simulation is then also run on this MUT model (Fig 2B). In both states, fluxes are assessed simultaneously for all of the reactions in the network, however in the case of metabolic profile prediction, we only consider the exchange reaction fluxes. In constraint-based metabolic models a positive exchange reaction flux value represents an export of the corresponding metabolite, and a negative flux value means that the metabolite is imported. These flux values are represented in Fig 2B as the thickness of the arrows, and in Fig 2C as single flux values for each metabolite's exchange reaction. Then, the flux values for every exchange reaction in the network can be compared between the WT and MUT

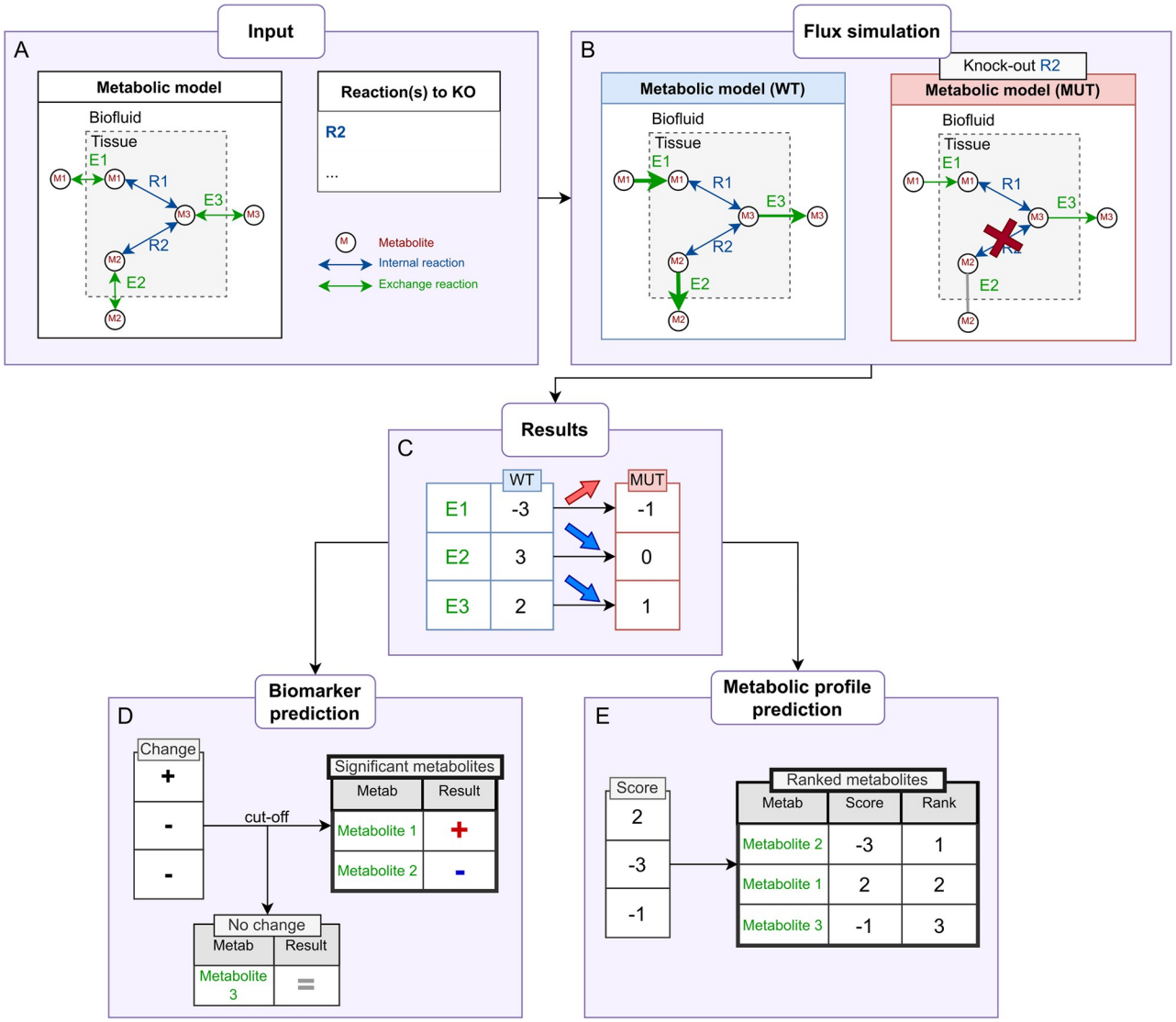

Fig 2. Methodology for the comparison of flux values and prediction of metabolite ranks. Using a simple network (A) in two conditions (B), with single flux values (C). The methodology from Shlomi *et al.* [13] is shown in (D): each exported metabolite will have an associated change score for a given pair of conditions. (E) shows our methodology of scoring and ranking by absolute value among all of the metabolites in the network.

https://doi.org/10.1371/journal.pcbi.1011381.g002

conditions (Fig 2C) to determine a change (Fig 2D). The change of a given metabolite's exchange reaction flux provides information on the corresponding metabolite's production/consumption in both conditions, which leads over time to an increase, decrease or unchanged concentration in the biofluids. This is based on the assumption that the metabolites are not being depleted or produced through alternate external metabolic processes, or, at the very least, the rate of depletion/production is negligible enough to have any noticeable effect on the extracellular metabolic concentrations.

In Shlomi *et al.* and Thiele *et al.* [13, 14], the authors show that CBM modelling can be used to predict biomarkers for diseases associated with gene deletions, validating these predictions with data from targeted clinical assays (focusing on specific classes of compounds like amino acids). The resulting predictions are boolean values associated with metabolites (i.e., the

metabolite is a biomarker or is not a biomarker (Fig 2D)), which does not allow the assessment of expected levels of changes in terms of concentrations.

In this article, we propose to go beyond previous work by predicting metabolic changes for 1497 metabolites from the Human1 metabolic network. The method can also be used to pinpoint metabolites not detected in the selected analytical setup, therefore guiding metabolomics researchers in the selection and design of assays. To do so, we implement a measure which ranks metabolites based on their likelihood to be modulated under a given genetic, environmental, or otherwise metabolic stress (Fig 2E). To highlight the strong potential of our approach, we demonstrate how predicted metabolite rankings fit data from human metabolic profiling studies (both targeted and untargeted), hence paving the way for combining both in silico and experimental metabolic profiling.

This recommendation system is implemented in a freely available computational workflow at the following addresses: https://forgemia.inra.fr/metexplore/cbm/samba-project/samba or https://doi.org/10.5281/zenodo.8369624.

Results

From metabolic network flux sampling to improved metabolite change predictions

Sampling based flux estimation. Metabolic network models are usually undetermined, i.e., there are not enough constraints in the model to determine a unique solution for the mass balance system of linear equations [7]. For example, in the WT (Fig 2B), any other value for E1 would still ensure the steady state, as long as the other reactions' flux values are changed accordingly to compensate. This is why it is difficult to define any one exact value for a reaction, and it is generally necessary to evaluate the range of possible flux values for all reactions through various CBM methods.

Fig 3 highlights two CBM methods for assessing the variability of flux values, again in two different conditions (WT and MUT). Flux Variability Analysis (FVA) [15] predicts minimum and maximum flux bounds for each reaction in the network (Fig 3B) under the predefined constraints. This interval represents the range of possible values that the fluxes of a given reaction can take. In an extreme case, like a complete KO or a single flux value, both the upper and lower bounds of the involved reaction will have null or close to zero values, like R2 in the MUT state in Fig 3B.

In Shlomi *et al.* and Thiele *et al.* [13, 14], the authors used FVA to predict biomarkers by comparing two conditions: WT, and inborn errors of metabolism as gene knock-outs. If the FVA intervals of an exchange reaction shift between the two conditions, then the corresponding metabolite is considered to be a potential biomarker for the disease state (like for R2 in Fig 3B).

A second method for determining the most frequent flux values for a given reaction is by assessing the possible flux configurations (distributions) which fit the constraints defined in the model. These constraints on reaction fluxes (minimum and maximum possible values) define boundaries in the space of all possible flux values: the solution space. The solution space contains an infinite number of feasible flux values, hence it is impossible to fully explore the entire solution space without an infinite number of samples. Markov Chain Monte Carlo methods can be utilised for uniformly sampling the constrained space of feasible solutions, thereby offering a more comprehensive characterization of the solution space [16–18].

FVA may conceal the fact that the values present in most of the solution combinations (e.g. the most frequent flux values) are, for instance, close to one of the bounds as opposed to at the halfway point. For example, by comparing Fig 3B and 3C, the most frequent flux values for

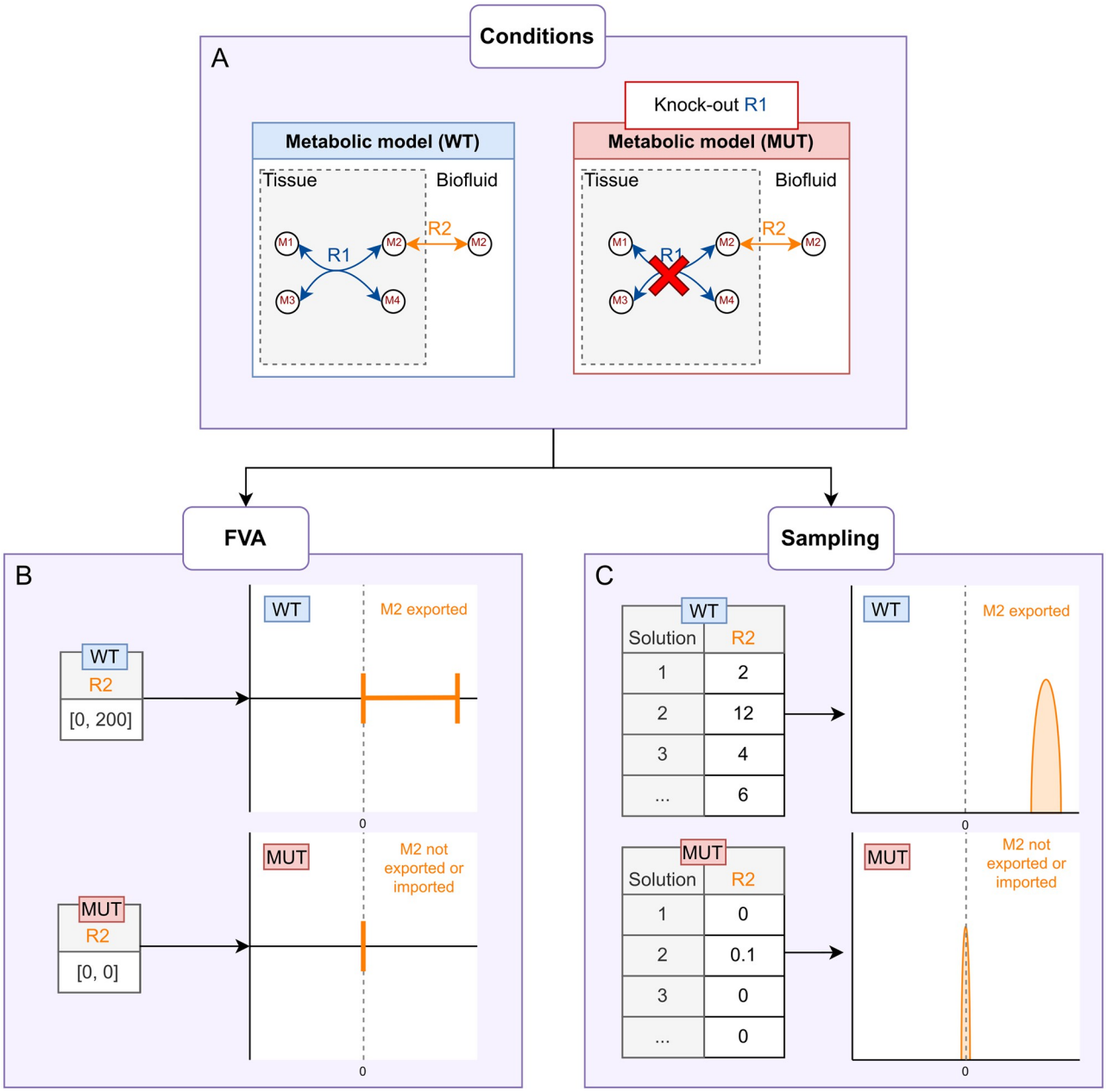

Fig 3. Flux Variability Analysis (FVA) and sampling. For simulating fluxes in different conditions (A). The resulting flux values to be compared differ depending on the method used. FVA generates minimum and maximum possible flux values, shown as intervals (B), whereas sampling generates many values within those bounds, shown as distributions (C).

https://doi.org/10.1371/journal.pcbi.1011381.g003

WT are much closer to the upper FVA bound than the centre of the bounds. It is also possible for two reactions to have the same extreme bounds but with flux values which are spread out differently in between the bounds. In other words, FVA is not sufficiently precise to assess the most frequent flux values. FVA does not provide insight into the distribution of flux values between those bounds since the most frequent flux value for a reaction could be located anywhere in between the two bounds.

By exploring the many combinations of flux values that satisfy the model constraints, sampling provides an overview of the most frequently valid flux values for every reaction in the network. Indeed, any one possible solution is a specific combination of flux values for each reaction, and some flux values are more frequent than others, meaning that they appear more often in different valid solutions. These many solutions can therefore outline each reaction's flux distribution and thus reveal the most frequent fluxes along with the variability of the values, as shown in Fig 3C.

In this study, we propose an application we developed specifically for predicting metabolic profiles using random sampling, called SAMBA (SAMpling Biomarker Analysis) (using the optGPSampler algorithm).

Scoring and ranking metabolites according to their predicted out/influx changes. While previous modelling methods had been used to predict biomarkers, the goal here is to predict entire metabolic profiles by capturing, for each metabolite, its amplitude of variation between the control and the condition under study. This variability can be evaluated, and metabolites can be scored and ranked in order to prioritise the ones to be measured and annotated during metabolic profiling.

Using FVA to predict biomarkers reveals another underlying limit regarding the comparison of bounds between conditions: there are multiple ways of defining the comparison between two intervals. This can become especially difficult when the intervals are both positive and negative (reversible reactions): for example, if a reaction has bounds of [-100, 100] in one condition and [-50, 300] in another, can this be considered an increase or a decrease of flux? The total flux range of this interval has increased (from 200 to 350) but the flux in one direction has decreased (from -100 to -50). This is one of many cases where flux intervals can be ambiguous in determining an increase or decrease.

Comparing sampling distributions directly can help overcome this difficulty in comparing intervals. SAMBA is used to sample flux distributions for every exchange reaction in the network, in two network states (WT and MUT). A score is calculated for each exchange reaction in the model by comparing the samples between both states. We propose the use of a z-score to evaluate the shift in distributions weighted by their variance, based on the z-score used in Mo *et al.* [19]. Z-scores are calculated for each metabolite's exchange reaction ex_i between the WT and MUT. First, we sample a number (by default the total number of samples) of random pairs of values from the WT and MUT distributions. The collections of all MUT samples and WT samples for metabolite *i*'s exchange reaction are MUT_i and WT_i respectively. A "difference distribution" dd_i is calculated by subtracting random pairs of values from both MUT_i and WT_i (Eq 1). These random samples are not matched: the two WT and MUT values are not necessarily from the same sample step. The final z-score z_i is calculated by dividing the mean μ_{dd_i} by the standard deviation σ_{dd_i} of dd_i.

For an exchange reaction ex_i:

$$z_i = \frac{\mu(MUT_i - WT_i)}{\sigma(MUT_i - WT_i)} = \frac{\mu_{dd_i}}{\sigma_{dd_i}} \quad (1)$$

The z-score z_i is directional: a negative z-score indicates a decreased shift in the flux distributions from WT to MUT, and a positive z-score indicates an increased shift. A z-score close to 0 means that there is little difference between the distributions in the WT and MUT conditions. A z-score therefore represents the intensity and direction of one metabolite's shift in a specific condition.

While a z-score was chosen for this metabolic profile prediction approach, other scoring methods were analysed and compared to make sure it was the best suited method, such as the

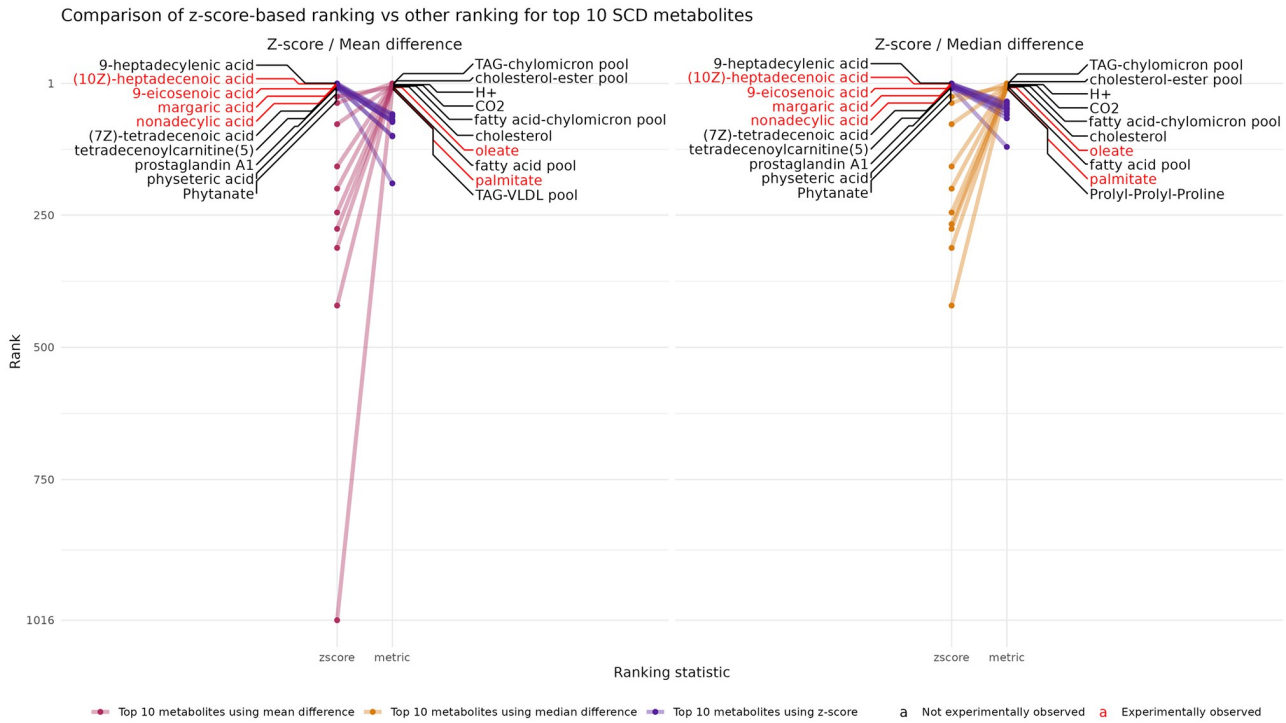

Fig 4. Z-score based ranks (left side of each plot) vs using the difference between means (right side of left plot) or medians (right side of right plot) to rank metabolites. The metabolite labels are all ranked in the top 10 by each method. The metabolites highlighted in red are differentially abundant in the example dataset, whereas those in black are the rest of the top 10 metabolites which are not experimentally associated with the SNP.

https://doi.org/10.1371/journal.pcbi.1011381.g004

subtraction of distribution means. In the following figures, the SCD example from the following sections was used for illustrative purposes. Briefly, the example is from a metabolic Genome Wide Association Study (mGWAS) cohort which analyses Single Nucleotide Polymorphisms (SNPs) significantly associated with various experimentally measured metabolites. When modelling one example SNP, the simulation predicts the 1497 exchanged metabolites from the network and compares them to the list of 20 experimentally significant metabolites (shown in red in the following figures). The list of 1497 metabolites is ranked based on the z-score as well as other methods, and then compared between each scoring method.

Fig 4 shows the comparison of using a z-score to compare distributions, versus subtracting the means of the distributions (left panel) or medians of the distributions (right panel). For each pair of metrics, the top ten metabolites for each given metric are displayed with their labels (red being those in the experimental signature) and with dots located on the y-axis. The y-axis corresponds to the position in the ranking of these metabolites among the 1497 metabolites using the corresponding metric (the top being the first metabolite while the bottom is the last metabolite for the given ranking). For example, both TAG-chylomicron pool and oleate are ranked in the top 10 when using the mean difference as a ranking metric instead of the z-score (right side of the left plot), and TAG-chylomicron pool is in black because it is not observed to be experimentally abundant, whereas oleate (in red) is significantly abundant in the example experimental data.

Using the z-score as a ranking metric predicts more experimentally observed metabolites in the top 10 than when using the difference of the means or medians to rank metabolites. Two experimentally observed metabolites, oleate and palmitate, that were not in the z-score top 10

are highly ranked when using the mean- or median-based rank. However, most of the other top 10 ranked metabolites using the mean differences appear to be unrelated and non-specific (H+, CO2) or vague (metabolite pools) in modelling terms. A similar plot using fold changes of means and other metrics is shown in Fig E in S1 Text, and leads to the conclusion that using ratios of means is less adapted than both z-score and mean difference based ranking metrics. For the rest of this study, we used the z-score to rank metabolite changes as it provided the best results when looking at the top of the ranked list. Indeed, using the bottom half of the list can be complex due to the similarity in z-scores for the low-ranking metabolites (see Fig G in S1 Text).

Therefore, we propose the use of z-scores of all the exchange reactions in the network as a basis to rank all exchanged metabolites based on the intensity of the changes. Z-scores can also be used as-is or used with a threshold (see Discussion section). Since both increased and decreased metabolites are of potential interest, this ranking is based on the absolute values of the z-scores (Fig 2D). This reveals the metabolites whose import/export behaviour changes the most between the WT and MUT, relative to every other exchange metabolite in the network.

Furthermore, ranking the z-scores by absolute value provides insight via the comparison of the list of the top ranked metabolites between different scenarios, as ranks are relative to the entire list of exchange metabolites and not quantitative. In this paper, the rank calculated by SAMBA for each metabolite is referred to as SAMBARank.

In order to show the benefits of this methodology, we selected two complementary applications in human health. The goal was twofold: first to show the ability of SAMBA to retrieve known biomarkers, and then to show in an untargeted study that the newly introduced SAMBA ranking system is coherent with metabolomics data and indicate other potential metabolites of interest chemically related to the observed metabolites. First, the benefit of sampling for metabolic profile prediction is illustrated on a specific Inborn Error of Metabolism (IEM) and associated biomarkers. Then, we use a mGWAS example to show how ranking these metabolite predictions can align with untargeted studies like metabolomics, and to recommend new potential metabolites of interest.

Illustrating the benefits of sampling through the prediction of Inborn Errors of Metabolism biomarkers

Predicting Xanthinuria type I biomarkers. Xanthinuria type I is a rare genetic disease caused by a mutation in the XDH gene [20], and is characterised by kidney stones (urolithiasis), urinary tract infections, and rarely kidney failure [21]. In patients with this disease, a decrease in urate and an increase in hypoxanthine has been observed (from OMIM [22]).

Here, we applied both FVA and sampling in order to compare the information which can be drawn from both techniques. Both the FVA bounds and the sampling distributions are displayed on the same plot for both of the expected biomarker metabolites (Fig 5). Expected biomarkers are defined as metabolites with observed significant changes in patients with the disease according to the original dataset. The flux simulations were run using Recon 2 [14], a human genome-scale metabolic network, by knocking out the XDH gene, which knocks out 7 reactions (see Table 1 in Methods). The flux values are reported on a log scale for clarity.

For urate, the FVA bounds were the same in both conditions (Fig 5A), which is interpreted as a metabolite not considered as a biomarker in Shlomi *et al.* [13], and thus the FVA prediction does not agree with the observed decrease. On the other hand, the sampling distributions correctly show a decreasing shift from WT to MUT. For hypoxanthine, both methods are able to predict the expected increase in metabolite export, as shown in Fig 5B, via the shift in

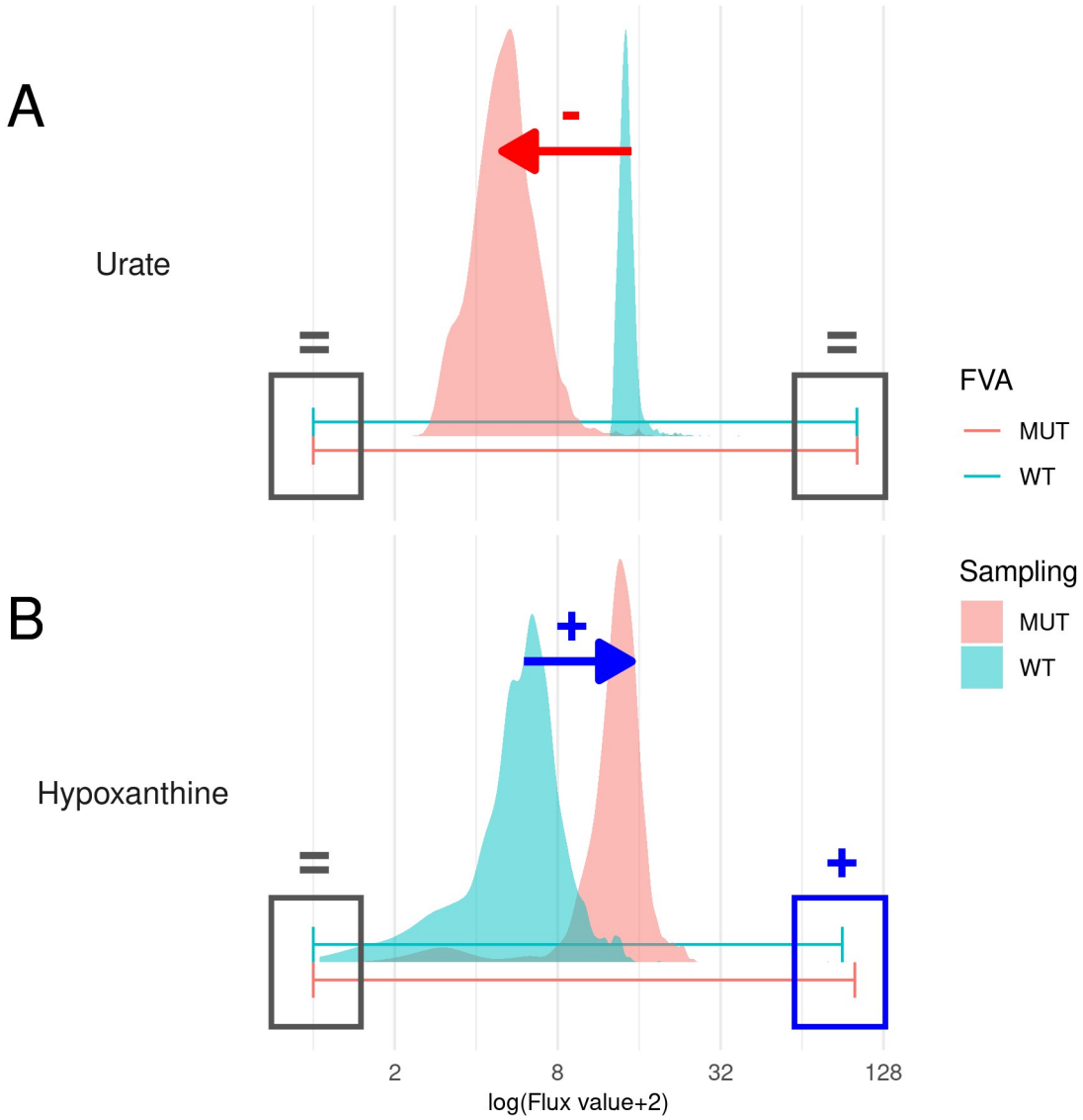

Fig 5. Flux bounds (FVA) and distributions (sampling) for urate and hypoxanthine. The WT state is shown in light blue, and the MUT state in red. MUT here corresponds to the knock-out of the XDH gene. Highlighted in grey, red, and dark blue are the absences of shifts (=), decreases (-), and increases (+) respectively between WT and MUT.

https://doi.org/10.1371/journal.pcbi.1011381.g005

distributions for sampling and the change in upper bounds for FVA although this change is very small (10.3%) compared to the total feasible range.

Comparison with FVA-based methods on 49 IEMs. Xanthinuria type I is one of many IEM from an entire IEM—biomarker dataset which was curated in Sahoo *et al.* [23]. We used

Table 1. Genes and reactions knocked-out to simulate Xanthinuria Type I in Recon2.

Condition	Model	Gene KO	Reaction KO
Xanthinuria type I	Recon2	XDH	XANDp XAO2x XAOx r0424 r0425 r0546 r0547
SCD	Human1	SCD, SCD5, FADS6	MAR02281 MAR02282 MAR02284 MAR02286 MAR02287 MAR02292 MAR02293 MAR02294 MAR02295 MAR02296 MAR02288 MAR02289 MAR00144 MAR00146 MAR00147 MAR00148 MAR02126 MAR02128

https://doi.org/10.1371/journal.pcbi.1011381.t001

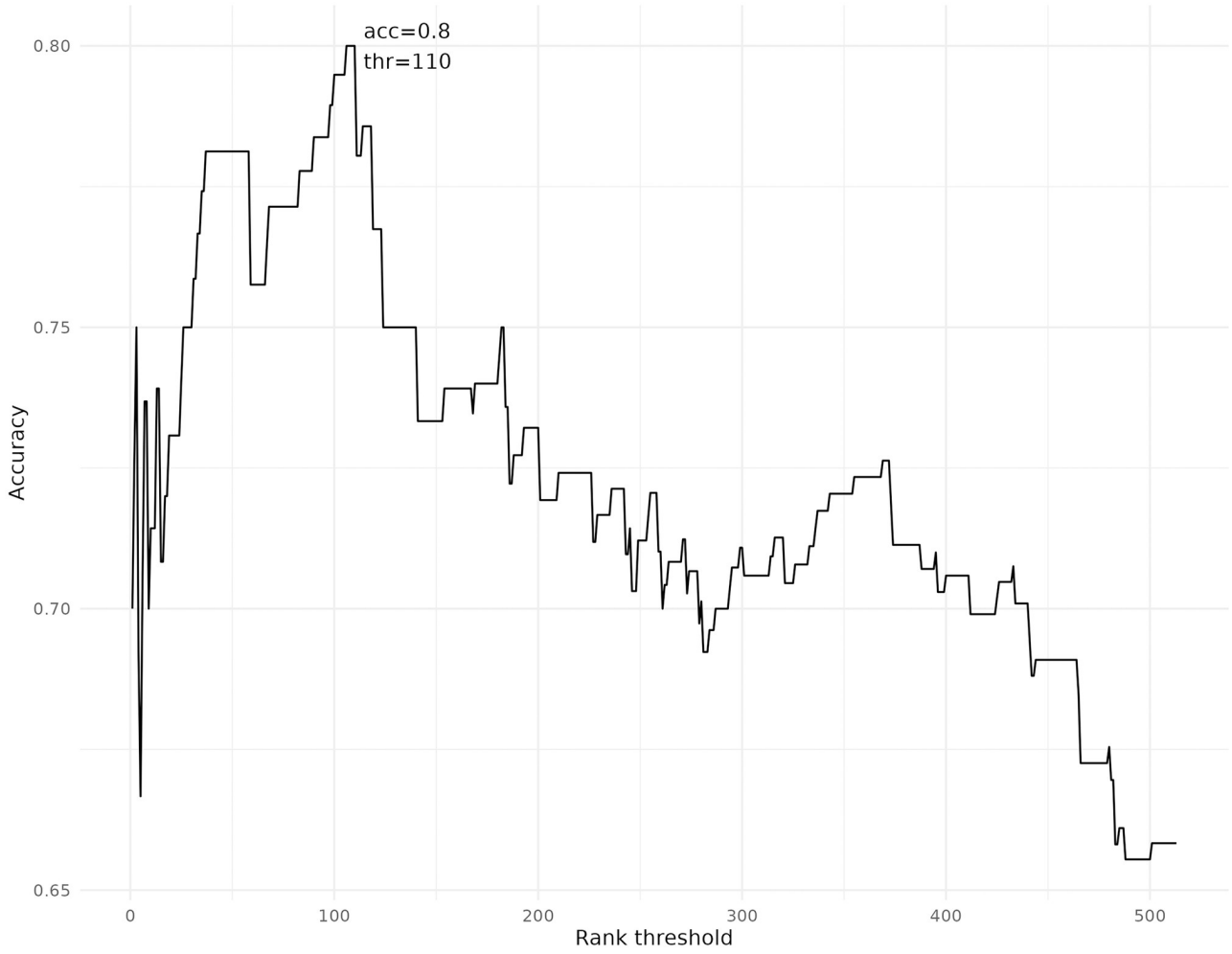

Fig 6. Accuracy plot for every rank cut-off threshold for the IEM dataset.

https://doi.org/10.1371/journal.pcbi.1011381.g006

a subset of this dataset, the same used in Thiele *et al.* [14], to run our analyses by knocking out each gene responsible for each disease: we ran both FVA and sampling on the 49 IEM for 54 metabolites on Recon 2. Heatmap figures containing the entire set of predictions for both FVA and sampling are included in supplementary data Fig A in S1 Text and Fig F in S1 Text, and overlaps are shown in Fig B in S1 Text.

The accuracy score calculated in [14] to evaluate directional biomarker predictions for various IEM can also be applied to sampling predictions for the same set of metabolites and diseases. It must be noted that the accuracy score evaluates the capacity of the method to predict the correct change direction among all observations, and not the capacity to predict correctly among all possible predictions. This means that false positives and false negatives are disregarded when using this score, and the focus is on evaluating how well the method can predict the metabolite's direction of dysregulation when a prediction matches an observation. Fig 6 shows these accuracy scores for when using SAMBA predictions with different rank cutoffs (only the top N metabolites out of all SAMBA predictions for each IEM are considered).

The best rank cut-off for this dataset is 110, with an accuracy of 80%. In this case, SAMBA's accuracy score is slightly higher than the accuracy with the FVA-based method, which was

76.8%. Despite predicting less total metabolites when using a rank cut-off, SAMBA can provide a similar level of accuracy in terms of metabolite change direction prediction to the FVA-based method used in previous studies. SAMBA's main advantage lies in extending metabolic profiles to provide an improved understanding of the perturbation, as shown in the following sections.

Overall, sampling not only complements FVA by providing new correct predictions, but also attributes more meaning to the scores of the predictions for each metabolite (see Fig A in S1 Text). Sampling z-scores, as opposed to the binary increase/decrease indicators of FVA, can be used to rank, filter and gain insight on the intensity of changes. Indeed, once generated, distributions can be used in multiple ways, one of which is calculating z-scores and ranks as shown in this paper; in contrast FVA interval boundaries represent only the flux extremes instead of the general flux behaviour. This is illustrated by comparing the sampling z-scores in Fig G in S1 Text with the FVA boundary differences in Fig H in S1 Text. Indeed, this comparison demonstrates that a ranking metric or threshold would not be feasible on FVA intervals due to the almost binary nature of the results: either at or close to zero difference between conditions, or an extreme shift in bounds.

Generating coherent recommendations and identifying novel metabolites of interest associated with SNPs using SAMBA

Simulating the metabolic profiles associated with two SNPs, SCD and ACADS. For this study, we chose a case example for the comparison of mGWAS data and SAMBA ranking system. In general, Genome Wide Association Study (GWAS) datasets are composed of traits associated with SNPs, which are germline genetic substitutions of one nucleotide, present at a specific DNA position in at least 1% of the population. Specifically, mGWAS data consists of SNP-to-metabolic trait associations. One type of metabolic trait consists of single metabolite fold changes between non-SNP and SNP individuals (e.g. the fold change of margarate). The second type associated with SNPs in the study is ratios of two different metabolite levels, again compared between non-SNP and SNP individuals (e.g. the ratio of margarate / palmitoleate). Other examples of these types of data are shown in Table 2 in Methods. We used data from Suhre *et al.* [24], extracted SNPs associated with significant metabolites, and mapped them onto the Human 1 metabolic network.

Among the 37 SNPs present in Supplementary Table 3 of Suhre *et al.* [24], 17 were SNPs of 17 metabolic genes (one SNP per gene) present in the metabolic model Human 1 (version 1.10). The 20 other SNPs were impossible to simulate since they do not correspond to metabolic genes in Human 1. Human 1 was used to run sampling on 2 of these 17 SNPs: SNPs

Table 2. Examples of mGWAS data.

Ratio	beta' meta	P meta	p-gain meta
myristate (14:0) / myristoleate (14:1n5)	0.124	$2.9*10^{-57}$	$1.2*10^{48}$
myristate (14:0) / palmitoleate (16:1n7)	0.131	$1.4*10^{-48}$	$1.0*10^{39}$
margarate (17:0) / palmitoleate (16:1n7)	0.157	$2.1*10^{-42}$	$6.6*10^{32}$
margarate (17:0)	0.06	$4.9*10^{-08}$	1
myristoleate (14:1n5)	−0.075	$3.3*10^{-09}$	1

3 significant metabolite ratios and 2 significant single metabolites. Beta' represents the relative difference per copy of the minor allele (SNP) for the metabolic trait compared to the estimated mean of the non SNP population. The p-gain statistic quantifies the decrease in P value for the association with the ratio compared to the P values of the two separate corresponding metabolite concentrations.

https://doi.org/10.1371/journal.pcbi.1011381.t002

affecting the -CoA 9-desaturase (SCD) gene and the ACADS (Acyl-CoA Dehydrogenase Short chain) gene. Human1 is one of the most recent and largest reconstructions of the human metabolic network, also showing that the method can scale to this larger network (13 024 reactions, 8 363 metabolites). The 15 remaining SNPs with corresponding genes in Human 1 were not analysed due to the manual curation needed to confirm genetic, enzymatic and metabolic matches.

We chose to focus on the SCD SNP specifically because i) the gene and reactions are present in the network, and ii) there are many measured metabolites present in the network, which is not the case for all of the SNPs, as some SNPs only have one or two significantly associated metabolites, or the associated metabolites do not exist in the network. It therefore serves as a good proof of concept application for the methodology. Furthermore, the selection of the correct genes to KO in the model for each SNP requires manual curation to make sure the Gene Protein Reaction (GPR) relationships correctly represent the enzyme and corresponding gene. Additionally, mapping the metabolite names from the study to model metabolites is a time consuming manual step. Results for SCD and ACADS (another SNP) are shown in the main text.

In contrast with IEM data, where mutations always result in an enzyme defect, an SNP might reduce enzyme activity (knock-down), enhance enzyme activity, or have an effect on a different gene. Some of the SNPs from the Suhre *et al.* [24] study are well known to be associated with loss-of-function phenotypes such as enzyme deficiencies (e.g. the ACADS gene in ACADS-deficiency), and others have not been studied enough to confirm the effect of the SNP on gene function. As one example of an understudied SNP phenotype, the SCD gene (SNP rs603424 [25]) codes for the enzyme Stearoyl-CoA 9-desaturase, involved in fatty acid metabolism. The hypothesis is that the SNP mutation in the gene affects the corresponding enzyme negatively, which leads to no SCD enzyme activity, represented in the network by knocking-out the SCD gene and therefore blocking the corresponding reactions. This is suggested in Illig *et al.* [25] by drawing a parallel between known loss of function SNPs leading to severe disorders, and newly identified SNPs. Additionally, the SNP mutation in the SCD enzyme-coding gene is predicted to be in an intronic (i.e. non-coding) region, using ensembl's VEP (Variant Effect Predictor) [26]. When simulating a scenario, the effect of the gene mutation should always be checked in order to generate the most accurate metabolic condition possible.

In Human 1, there are 19 reactions linked to the SCD gene, most of which involve the desaturation of stearoyl-CoA, palmitoyl-CoA and myristoyl-CoA into corresponding mono-unsaturated fatty acids. Following the GPR relationships in the model, knocking out SCD only affects 4 reactions (due to the fact that SCD can be compensated by another gene). However, SCD also shares 14 GPRs with two other genes: SCD5 and FADS6, whose functions are not well described. We decided to knock out these extra 14 reactions in order to block the enzymatic function related to SCD completely.

The SCD SNP has two types of significantly associated metabolic traits: single metabolite changes, and ratios of two different metabolite concentrations. The single significant metabolites measured for the mGWAS study for this SNP are margarate, palmitoleate, myristoleate, stearate and 1-palmitoleoylglycerophosphocholine. These are the main "expected" metabolites, which will be compared with the SAMBA recommended metabolites.

SAMBA returned z-scores for the 1497 unblocked metabolite exchange reactions in Human 1. The distribution of these z-scores is shown in Fig G in S1 Text, and highlights the difference between the extreme high-ranking metabolites and the low-ranking metabolites in the centre. A metabolic profile this large is difficult to compare with the data from the mGWAS study as no raw data was included in the original study: only the significantly associated metabolites were reported, as well as the total list of 295 measured metabolites (but not

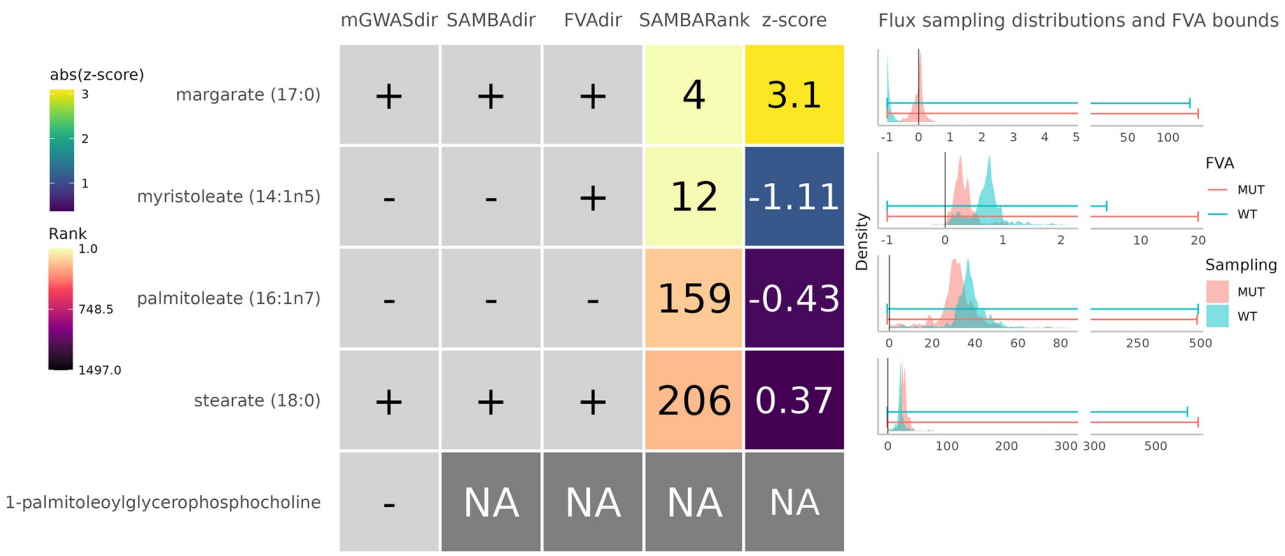

Fig 7. Observed and predicted changes for the five metabolites significantly associated with the rs603424 SNP. The first column shows the observed change directions from the mGWAS study. The second column shows the predicted change direction using SAMBA (SAMBAdir). The third column shows the predicted change direction using FVA (FVAdir). The fourth column shows the SAMBA predicted rank out of the 1497 metabolites in the network (SAMBARank). The fifth column shows the SAMBA predicted z-score, with the colour scale as the absolute value of the z-score. The NAs represent metabolites for which SAMBA was unable to predict fluxes for one of the following reasons: (i) the metabolite is not in the network, (ii) the metabolite is in the network but has no exchange reaction, or (iii) the metabolite's exchange reaction can carry no flux (=blocked). Sampling distributions and FVA predicted bounds for each metabolite's exchange reaction in WT and MUT are shown on the right.

https://doi.org/10.1371/journal.pcbi.1011381.g007

their fold changes for each SNP). We also calculated the FVA bounds for each metabolite for the same metabolic condition as the sampling. Here, we compared the 5 significant metabolites reported in the mGWAS study with their simulated SAMBA metabolite ranks and FVA bounds to see the biggest effect this KO has on metabolite exports and imports. No rank or z-score threshold was used for Figs 7 and 8 as the metabolites were selected based on their presence in the significant results of the Suhre *et al.* [24] dataset.

Fig 7 shows the five metabolites identified in the mGWAS study along with the corresponding SAMBA ranks and the FVA predictions. Both in Figs 7 and 8, the metabolite(s) marked with "NA" in the SAMBARank column have no flux values because either they aren't present as a metabolite in the network, don't have an exchange reaction in the network, or have a blocked exchange reaction, meaning no flux can be carried through it in the current metabolic state.

Four out of the five expected metabolites are present with an exchange reaction in Human 1, and the SAMBA predicted change directions match the expected mGWAS experimental changes. The directions of change predicted by FVA are correct except for myristoleate, which was predicted to be increased instead of decreased using the FVA bounds. Their ranks are shown in the column SAMBARank and these ranks are to be compared with the total number of exchange metabolites present in Human 1, i.e. 1497. These four metabolites are in the top 13%, two of which are in the top 1%.

The significant metabolite ratios linked to SCD include many different combinations of pairs of metabolites. The assumption here is that at least one of the two metabolites involved in each ratio must change for the ratio to be significantly changed. Fig 8 shows the metabolites present in at least one ratio significantly associated with SCD and their associated predicted SAMBARanks.

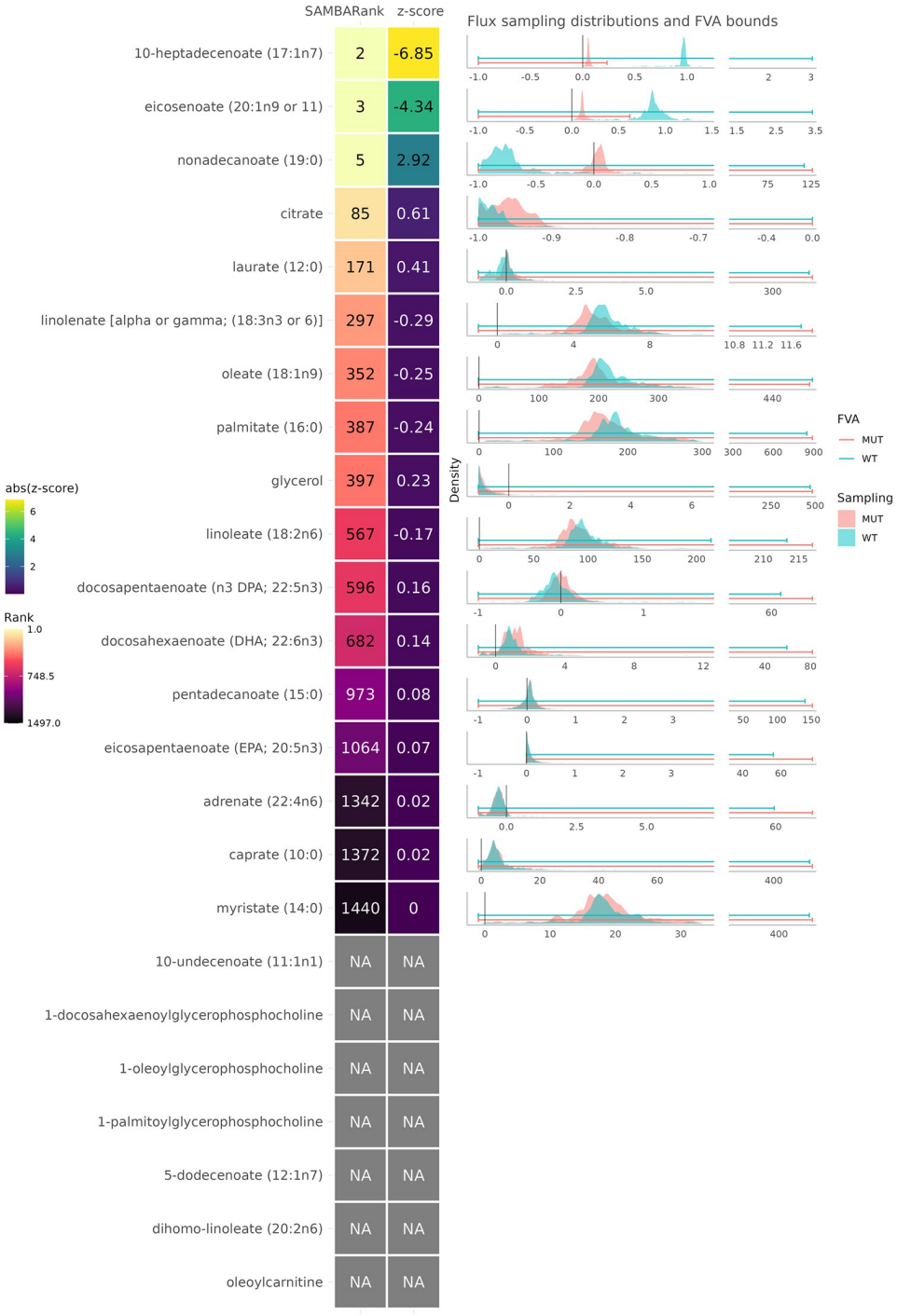

Fig 8. Predicted ranks for the metabolites present in a ratio significantly associated with the rs603424 SNP. The first column shows the predicted rank out of the 1497 metabolites in the network. The second column shows the SAMBA predicted z-score, with the colour scale as the absolute value of the z-score. The NAs represent metabolites for which SAMBA was unable to predict fluxes for one of the following reasons: (i) the metabolite is not in the network, (ii) the metabolite is in the network but has no exchange reaction, or (iii) the metabolite's exchange reaction can carry no flux (=blocked). Sampling distributions for each metabolite's exchange reaction in WT and MUT are shown on the right.

https://doi.org/10.1371/journal.pcbi.1011381.g008

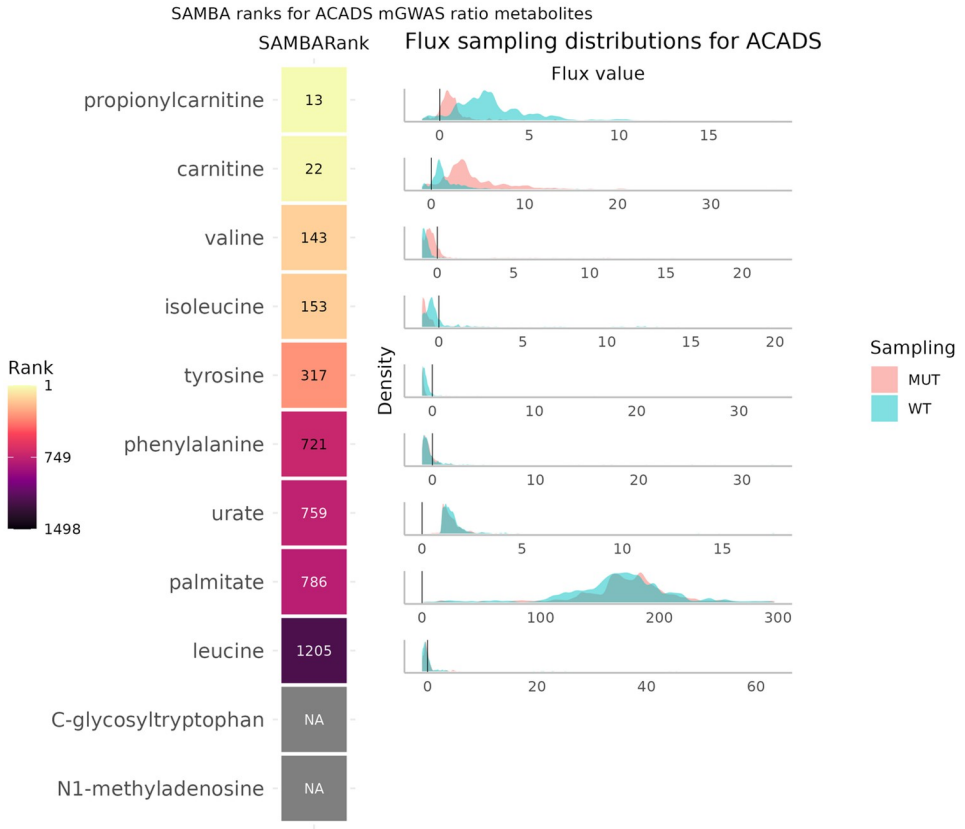

Fig 9. SAMBA ranks for the metabolites involved in significant ratios for the ACADS SNP from Suhre *et al.* 2011. [24] The SAMBARank column shows the predicted rank out of the 1498 metabolites in the network. The NAs represent metabolites for which SAMBA was unable to predict fluxes for one of the following reasons: (i) the metabolite is not in the network, (ii) the metabolite is in the network but has no exchange reaction, or (iii) the metabolite's exchange reaction can carry no flux (=blocked). Sampling distributions for each metabolite's exchange reaction in WT and MUT are shown on the right.

https://doi.org/10.1371/journal.pcbi.1011381.g009

The second most differentially abundant metabolite predicted by SAMBA for this condition is 10-heptadecenoate, which is present in at least one significant ratio in the mGWAS SCD dataset. In addition to this, there are 4 other highly ranked metabolites, all in the top 171 ranked metabolites out of 1497 (top 11%). The five metabolites ranked below the 50% mark have z-scores lower than 0.1. Interestingly, myristate is almost ranked last in the entire list of predictions. When taking a closer look at its flux distributions, the MUT distribution appears to be bimodal, meaning that while the flux seems to have shifted, the z-score was not able to detect this difference due to its reliance on the similar means.

A second example from the Suhre *et al.* [24] paper is the Acyl-CoA Dehydrogenase Short chain (ACADS) SNP. In the paper, it does not have any "single" metabolite trait associations, but has 11 significant ratio metabolite associations. The predictions using SAMBA for these metabolites involved in significant ratios are shown in Fig 9.

As in the previous example, there are some extremely well ranked metabolites while others are poorly ranked. The literature suggests that carnitine (rank 22) is intrinsically linked with CoA on multiple levels. Metabolically, the reactions catalysed by ACADS enzymes are two reactions away from propionycarnitine and L-carnitine, due to their interaction with propanoyl-CoA, their direct product. Carnitine also plays a role in the stabilisation of CoA and

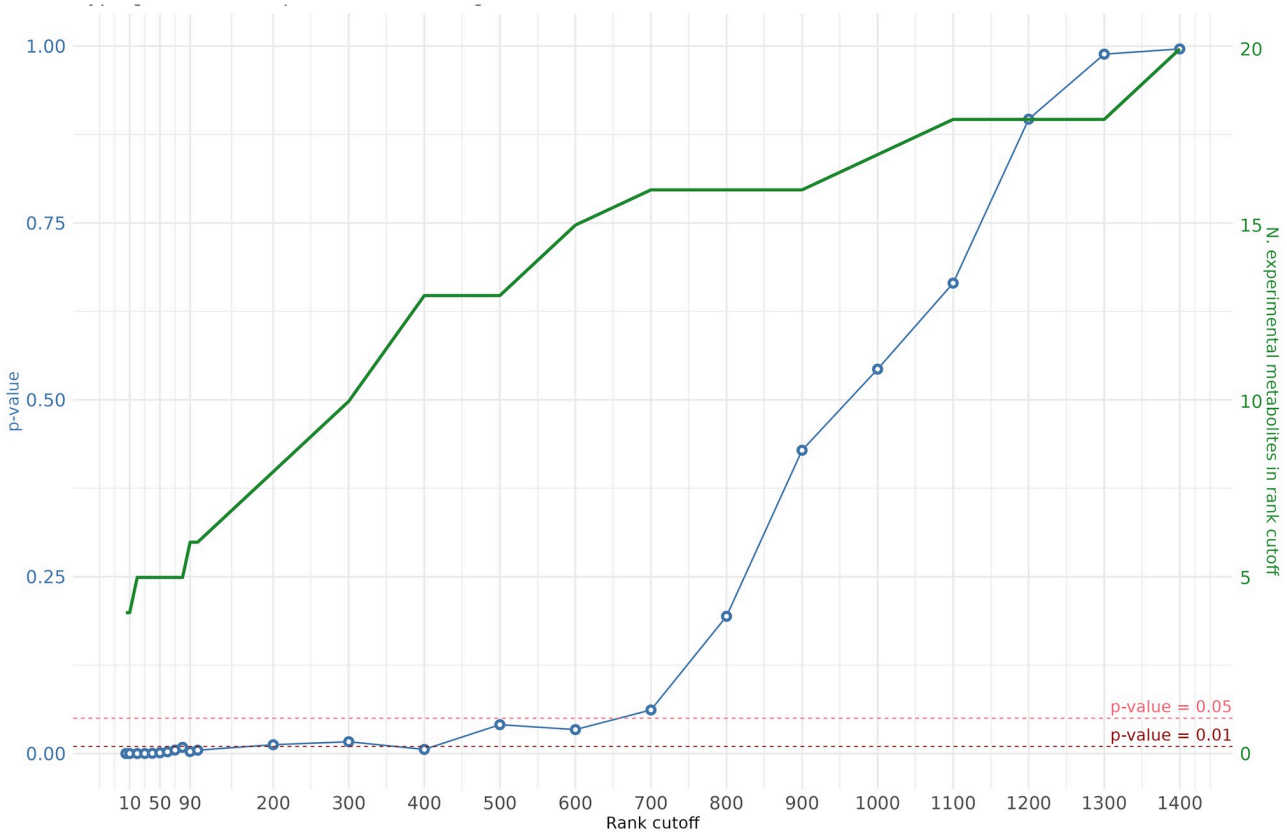

Fig 10. Hypergeometric test p-values for different rank cut-off values for SCD. Hypergeometric test p-values for different rank cut-off values for SCD. The left y axis (blue) shows the hypergeometric test p-values when using a given rank cut-off and the number of experimental metabolites predicted in that top ranking. The right y axis (green) shows the number of experimental metabolites predicted for each rank cut-off.

https://doi.org/10.1371/journal.pcbi.1011381.g010

acetyl-CoA levels, as well as energy production by taking part in a rate controlling step in mitochondrial oxidation of long-chain fatty acids [27]. Medically, L-carnitine is used as treatment in some cases of ACADS deficiency (also known as SCAD deficiency (short chain acyl-CoA dehydrogenase)) [28]. Regarding the highly ranked amino acids, an adjacent enzyme Isobutyryl CoA Dehydrogenase (IBD), which is coded by ACAD8 and shares GPRs with ACADS, has been shown to be involved in valine metabolism [29, 30]. The ACADS gene is also involved GPRs in reactions in the "Valine, leucine, and isoleucine metabolism" pathway in the Human1 genome-scale metabolic network (GSMN).

Evaluating the statistical significance of SCD's predicted metabolic profile. Despite the problems that come with evaluating the false positives and negatives predicted by metabolite prediction methods, the statistical significance of the previous findings can be evaluated using a hypergeometric test. The test describes the statistical significance of predicting k number of metabolites correctly out of the top n predictions, when taking into account the total N number of predictions containing K number of experimentally significant metabolites.

Fig 10 shows the results of these tests for various rank cut-offs. For example, when looking at the top 300 (n) metabolites (x-axis), predicting 10 (k) experimentally significant metabolites (green y-axis) out of the 20 (K) total experimental metabolites for a total of 1497 (N) predictions, is significant (p-value < 0.05) (blue y-axis).

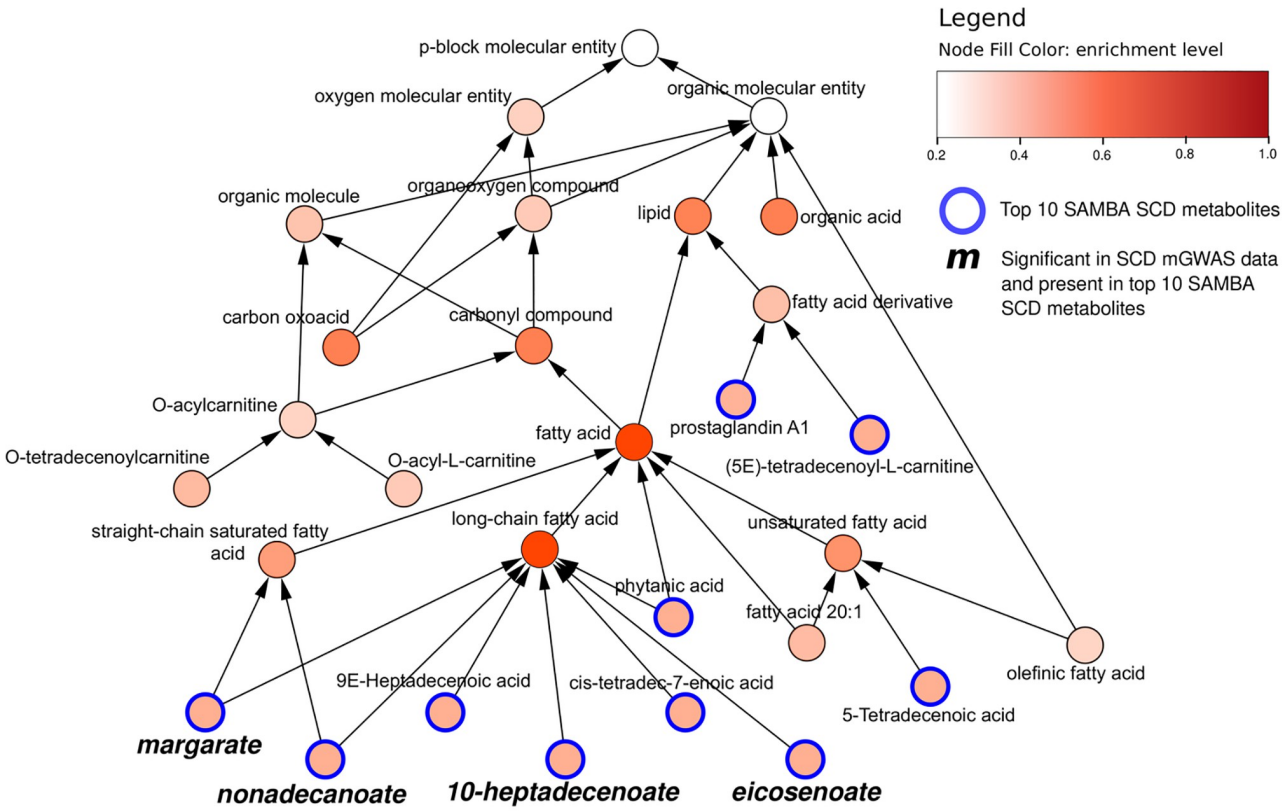

Fig 11. Hierarchical BiNChE CHEBI graph. Hierarchical CHEBI graph of the top 10 metabolites predicted to be differentially abundant (outlined in blue) predicted by SAMBA for SCD, extracted using BiNChE. The node colour corresponds to the BiNChE enrichment level. The metabolites in bold & italic were significant in the mGWAS dataset for SCD.

https://doi.org/10.1371/journal.pcbi.1011381.g011

The figure highlights the significance of finding theses numbers of expected metabolites in the top ranks of the SCD predictions. Until around the top 100, the test shows that predicting around 6 expected metabolites is extremely significant (p-value $\ll 0.01$) and remains significant (p-value < 0.05) until just below the halfway point of the ranked list.

Using SAMBA predicted metabolite lists enriches experimental knowledge and points to novel chemical classes of interest

The top 10 most differentially changed metabolites associated with SCD predicted using SAMBA can be used to form a list of new metabolites of interest for this condition. By examining the chemical class of each predicted highly differentially abundant metabolite, we can gather information on a general type of metabolite affected by the KO.

Fig 11 shows the CHEBI ontology extracted using these top 10 metabolites. This hierarchical graph was made using BiNChE [31], which creates and enriches a subnetwork using a list of CHEBI IDs and the CHEBI ontology.

All of the top 10 most changed metabolites are classed as lipids (outlined in blue in Fig 11), 8 of which are fatty acids, which is consistent with the functionality of the enzyme SCD. Indeed, the SNP rs603424 has been shown to be significantly associated with circulating phospholipid levels [32], as well as with low levels of palmitoleate [32]. SCD is a desaturase which leads to the formation of fatty acids, specifically monounsaturated fatty acids involved in membrane phospholipids [33].

Out of the top 10 metabolites, 4 were measured in the mGWAS study (margarate, 10-heptadecenoate, nonadecanoate, and eicosenoate), and they are all classified as saturated or long-chain fatty acids. This means that the other long-chain fatty acids could be potential metabolites of interest, such as 9-Heptadecenoic acid (rank 1) or cis-tetradec-7-enoic acid (rank 6), which weren't measured in the original mGWAS study.

However, the ChEBI classification is limited by the annotation of each metabolite to the correct class. Upon manual inspection, both cis-tetradec-7-enoic acid and 5-tetradecenoic acid are C14:1 fatty acids, only differing by the position of the double bond, but they are classified separately in long-chain fatty acid and unsaturated fatty acid respectively. This indicates that 5-tetradecenoic acid could also be of interest for future studies. Furthermore, by looking at the chemical structures, 4 out of the top 10 are odd chain fatty acids which is interesting to highlight since they represent a very small percentage of the total human fatty acid plasma concentration [34].

Since BiNChE provides a view of the ChEBI ontology on a per-metabolite scale, using too many metabolites as input results in a large and difficult to read figure. Other methods can integrate more of the predicted metabolic profile (for example 50 metabolites).

As a step closer to using chemical structures as opposed to class annotations as well as using more of the metabolic profile, we ran a ChemRich [35] analysis using the top 50 metabolites predicted to be differentially abundant. It uses the chemical structure via SMILES, and the MeSH terms associated with PubChem IDs to highlight enriched chemical classes. Fig 12 represents the most enriched clusters from the top 50 metabolite set. The higher the -log(pvalue) (y axis), the more the group is enriched.

The ChemRich plot also shows that both saturated and unsaturated fatty acids are significantly enriched by this dataset. Fig 12 also highlights some other groups such as HETE (Hydroxyeicosatetraenoic acids (which are oxylipins)), cholestenes, and cholestadienols not detected using BiNChE. ChemRich serves as a complementary method to BiNChE for analysing predicted metabolic profiles, as highlighted in S2 Table. These are just two methods of enriching a metabolite set of interest. As well as looking into specific metabolites as potential biomarkers, using SAMBA could direct future research in SCD towards the general families of unsaturated fatty acids and straight-chain saturated fatty acids, with wider panels of measurements for better coverage.

Additionally, by using only the top 20 SAMBA predictions as input for ChemRich, the same classes as those obtained when using the list of experimentally significant metabolites are identified, shown in Fig 13. By going further down the list of ranked predictions, the information gained can be enriched using the simulated data. This figure clearly shows the gain of information as the list of original metabolites grows in length and is enriched.

The information gained by adding SAMBA predictions are the chemical classes mentioned previously: HETE, cholestenes and cholestadienols. HETE are eicosanoids produced from arachidonic acids, which are present in membrane phospholipids. The SCD enzyme is known to be a membrane-bound enzyme which contributes to the maintenance of the ratio of saturated to monounsaturated fatty acids for membrane fluidity, and is regulated by polyunsaturated fatty acids [36]. Furthermore, in Frainay *et al.* [3], the authors highlight the poor coverage of certain metabolic pathways in MS databases, one of which is Eicosanoid Metabolism. Enrichment methods using metabolomics data as input are highly dependent on the experimentally measured metabolites, and if a certain class of metabolites was not measured it cannot be enriched. In the mGWAS study, the most well-known eicosanoids were measured (adrenate, arachidonate, dihomo-linoleate, eicosenoate), but there are 81 metabolites in the Eicosanoids Metabolism pathway in Human1. The enrichment of the HETE chemical class appears to be relevant and could consist of a poorly known class to target for future analyses or experiments.

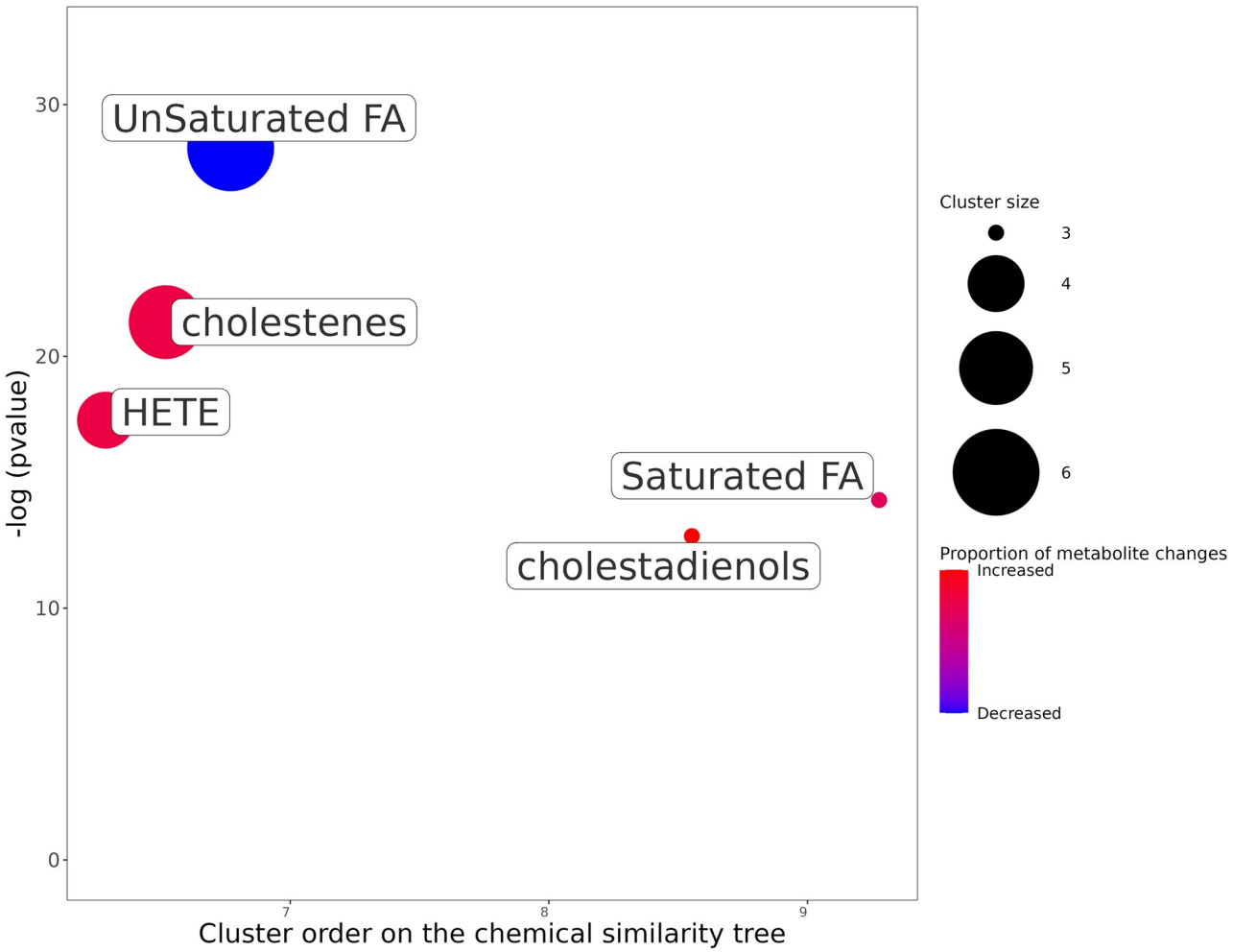

Fig 12. ChemRich enrichment of the top 50 most changed metabolites for SCD. The y-axis shows the most significantly altered clusters on the top. Each node reflects a significantly altered cluster of metabolites. Enrichment p-values are given by the Kolmogorov–Smirnov test. Node sizes represent the total number of metabolites in each cluster set. Cluster colours show the proportion of increased or decreased metabolites (red and blue respectively). The x axis represents a separation based on cluster order on the chemical similarity tree, and non-significant clusters are hidden.

https://doi.org/10.1371/journal.pcbi.1011381.g012

It is more specific compared to "Saturated FA" as an enrichment and therefore provides a more direct avenue of potential analysis.

Finally, we ran ChemRich on the IEM examples from the Thiele *et al.* [14] study using the sampling predictions for some IEMs (the same as those seen in Fig A in S1 Text). As an example of one of these, Fig K in S1 Text shows the ChemRich figure for Fish-eye disease/ LCAT deficiency when using the top 54 sampling metabolites (z-scores) predicted using SAMBA for this condition as input. The significantly enriched compound classes here are oligopeptides and retinoids, going beyond the originally expected biomarker "cholesterol" from the IEM compendium [23]. Fish-eye disease/ LCAT deficiency, as the name suggests, affects eyesight, resulting in corneal opacifications. Retinoids are class of chemical compounds that are vitamers of vitamin A, which plays a vital role in maintaining a clear cornea [37]. Based on the literature, Fig K in S1 Text appears to be a more coherent enrichment than the previous predictions.

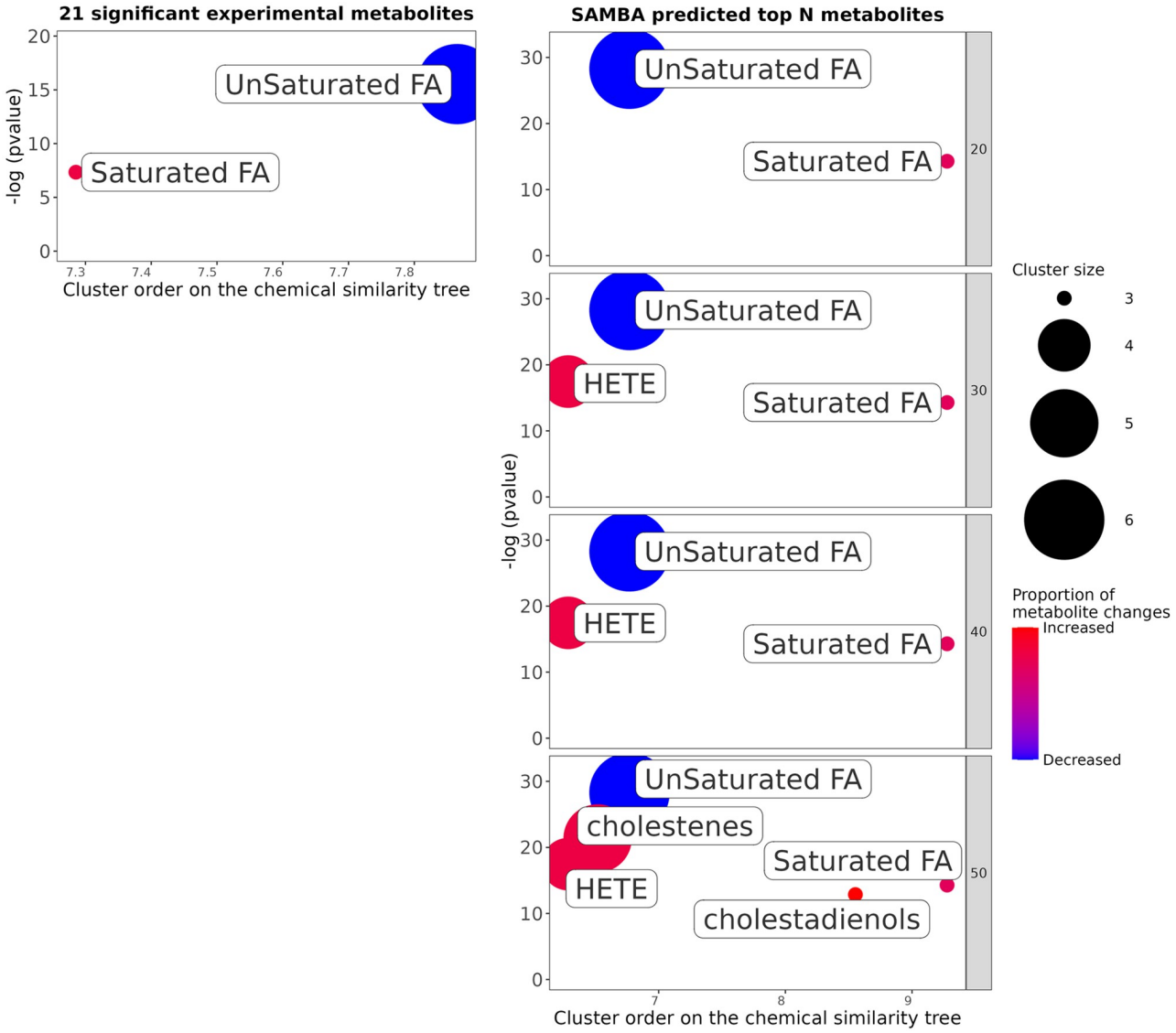

Fig 13. ChemRich enrichment of the top N most changed metabolites for SCD. ChemRich using only experimentally significant metabolites (left) and using increasing numbers of highly ranked SAMBA metabolites (right) for SCD.

https://doi.org/10.1371/journal.pcbi.1011381.g013

Discussion

The results presented in this study show that by using metabolism-simulating methods like SAMBA, we can predict metabolic profiles. For instance, for the SCD case study, the metabolites reported as associated with the SNP were highly ranked, especially when considering the total number of exchange metabolites in the whole human network.

Metabolites belonging to the predicted list but not in the original metabolomics fingerprint may be of interest to improve metabolic profiling. In fact, as it was shown in Frainay *et al.* [38], metabolites may be overlooked during the whole metabolomics pipeline. This can be for instance due to pre-processing steps since most peak picking methods [39] will define an intensity threshold to keep only intense peaks and, as a consequence, may discard peaks of interest that fall just below the threshold. Additionally, these metabolites are additional

metabolites of interest that could be future paths for analysis which could not be directly inferred from the affected reactions and scenario. This can be seen in Fig I in S1 Text and Fig J in S1 Text, which show that some highly ranked metabolites are not directly linked to the metabolic disruption. The far but highly ranked metabolites are metabolites we may not have thought of as potential metabolites of interest due to their distance from the disruption in the network, since they may seem unrelated at first glance.

Conversely, metabolites in the original experimental results but not predicted as highly ranked by SAMBA could be due to inconsistencies in the model, whether they are due to errors or unknowns, or an incorrectly simulated metabolic condition. Additionally, extra care should be taken when analysing low-ranking metabolites as their z-scores are very similar to each other. This means that their specific order does not indicate much information about the extent of how they were affected by the perturbation, only that they were affected very little.

The measurement of pure standards of metabolites is essential to obtain the highest level of confidence in metabolite identification (level 1 according to Metabolomics Standard Initiative [6]). Selecting which standards to measure is by itself a challenge, since samples can contain thousands of metabolites. Hence, SAMBA can be used by laboratories to select which standards to acquire in the context of the disease under study. More broadly, the top ranked list can also be used to identify families of metabolites to study as a whole, such as by extending the panel of measurable metabolites during a metabolomics experiment.

Although SAMBA is a predictive method, evaluating the predictions using traditional contingency tables, recall and precision is difficult due to the nature of metabolomics measurements and the available "truth" datasets. The model contains all known metabolites involved in metabolic reactions, but metabolomics methods are not able to detect and annotate all of them. This results in many cases where metabolites are predicted to be of interest while they are not detected by typical assays. In these cases, the predictions could be correct while being considered as a "false positive". Instead of using "false positive" to represent these predictions, we simply present the entire ranked prediction results in order to orient the user towards certain metabolites or metabolite classes. We then evaluate the method using true positive ranks and the list of the top most changed metabolites, some of which could be considered false positives, but could also be unmeasured metabolites. An additional method of evaluating the statistical validity of the results is by running a hypergeometric test, to test the significance of obtaining the number of correct predictions, for different rank cut-offs. This highlights that the number of experimental metabolites predicted in the top ranks is significant. Finally, for comparison reasons, we calculated the accuracy score used in a previous study [14] using our sampling predictions for the same dataset. Despite the slight improvement in accuracy, we believe that using this accuracy score does not represent the reality of the predictions. The score only takes into account predictions that match with observations, ignoring all predictions where no truth values were observed (false positives), which could be potential metabolites of interest if they were not measured in the patient dataset. It also ignores false negatives, where a prediction was expected to match an observation but did not, an essential part of evaluating the recall score of a method.

SAMBA is based on ranking z-score absolute values, meaning that the metabolites whose exchange fluxes (and by extension concentrations) are more likely to change will be considered first. There are of course metabolites whose concentrations can change very little and have extreme consequences on the rest of the metabolism, such as via enzyme regulation, or if they are limiting substrates for example.

We demonstrated that sampling can add a layer of information to better improve metabolic profiling compared to FVA. Sampling provides a finer grained description of changes which helps order metabolites based on their likelihood to be affected by a perturbation. Compared

with FVA, sampling is more computationally intensive (CPU and memory) but recent strategies are reducing this computational burden [40–42]. Nevertheless, sampling is currently more than feasible on large networks such as Human1.

The FVA method used in previous work [13] to compare intervals calculates the greatest change between the two pairs of upper bounds and the two pairs of lower bounds. This comparison of boundary shifts is not always representative of the underlying changes and can mislead the interpretation of the intensity of these changes. Using other methods such as comparing the means of boundaries assumes a uniform or centred normal flux distribution within these bounds, which we have shown via sampling is rarely the case. Using the most frequent fluxes with sampling appears as a good approximation of the mix of metabolite exports that occurs in biofluids, but it should be noted that the most frequent flux value may not be the most frequently observed flux in reality. However, in some cases, the most frequently predicted flux value may not represent the biological reality of a cell, such as for cells in extreme conditions or fast-growing cancerous cells, for which fluxes might be more close to the extremes. To represent these extreme conditions in SAMBA, the initial parameters of the model could be adjusted (such as a higher minimal production of biomass) to force the model to operate within extreme (boundary) optimums, as opposed to more likely fluxes.

The boundary shifts evaluated by FVA are very sensitive to change, since a very low threshold (1e-6 tolerance and 0.01 factor) for change is used to report an increase or a decrease. Despite this, FVA is able to predict biomarkers, as shown in previous studies [13, 14], when aiming to predict specific biomarkers. We progressed from the calculation of a score to the ranking of these scores since ranking the change intensities via sampling means that the most changed metabolites can be highlighted, while still keeping information on the other subtle metabolite changes. Contrary to the binary change/no change method of reporting FVA results, sampling ranks provide information on a wider scale by taking into account relative changes between metabolites.

In order to continue to highlight the full benefits of using sampling distributions instead of FVA boundary values, further research for other applications and more validation data are required. For instance, experimentally-measured *in vivo* fluxomics data [43] could be matched to simulated import/export rates. A database of every unique KO could be simulated and compiled as a repository for comparison with real data to determine which metabolic perturbations are most likely to cause the condition tested by the experiment.

This ranking system bypasses the issues that come with using flux values directly, and especially helps in choosing which metabolites to focus on first. The comparison of metabolic profile recommendations between different scenario simulations can be achieved by considering the top most changed metabolites and their ranks, as opposed to the raw flux values.

Z-scores prove to be useful in that they reflect an intensity of change similar to fold changes, and are weighted by the standard deviation of the distributions, which helps the z-scores to remain flexible given the variable nature of these distributions. Initially, instead of using a z-score to compare sampling distributions, more widely used statistical metrics were tested, such as Kullback-Leibler Divergence, Kolmogorov-Smirnov, and Wasserstein. However, they did not prove to be informative in our use case since they lead to p-values being too sensitive, resulting in extremely significant p-values for very similar distributions. In addition to this, these tests provide scoring metrics which are unable to quantify or describe the differences in the way a z-score can. Z-scores efficiently capture both the intensity and extent of variation of flux distributions between conditions. Additionally, we assessed various other metrics in order to decipher their ability to capture relevant metabolite rankings (more detail can be found in Fig E in S1 Text).

The goal of this study is to simulate whole-body metabolic markers using a generic genome-scale model. From a physiological point of view these models may seem to be somewhat over simplified in that a single metabolic system is represented. However the examples used in this study are genetic diseases, therefore they affect the genome of all of the cells in the body. While gene expression can depend on organs and tissue regions, the hypothesis here is that experimentally observed metabolic profiles are a combination of metabolite exports from all tissues connected to biofluids, which is why they can be equated to metabolic profiles predicted using a genome-scale network. However, the modulation of a tissue-specific biomarker may be predicted incorrectly if it is normally (biologically) compensated by other tissues, which could result in false positives. In those cases, tissue-specific networks could be useful for analysing diseases that are known to affect a certain tissue, such as glycogen storage diseases. These diseases are a collection of genetic metabolic disorders, and the enzymes affected by the mutations are specific to the liver and muscle [44]. By using transcriptomics data to create a liver-specific model, the accuracy of metabolic simulations could be increased. This can be done using various integration methods such as iMAT [45] or DEXOM [46]. However, choosing any given model and tissue-specific conditions must be done with care as it will have a major impact on the resulting metabolite ranks. More broadly, the definition of constraints is key to adapting the model to the biological condition (e.g. availability of nutrients) and will impact predictions. These modelling steps can be performed upstream of SAMBA.

Furthermore, including the SAMBA approach in whole-body metabolic models [12] which combines the interactions of multiple human tissues is a potential path for future study. Since sampling algorithms are being continuously improved and iterated upon, and more CPU power is being added to computational clusters, running sampling on these larger models will become less of an issue. These models, with their different gene and reaction expressions per tissue, could reveal the different effects of genetic diseases or other metabolic disruptions on biofluid metabolites on a multi-tissular level.

Finally, while SAMBA was applied to KO scenarios in this paper, the method can be adapted to more complex constraints such as multiple gene KOs or even to simulate knock downs of reactions. Knock downs involve reducing the maximum flux capacity of affected reactions instead of blocking the flux completely and can be run directly using SAMBA by changing the input condition file (see Fig D in S1 Text for details). This can be particularly useful in the context of toxicology or drug development, where these subtle metabolic disruptions can lead to reduced enzyme activity. There are many potential applications for SAMBA recommendations, such as in predicting the effects of xenobiotics on human metabolism. In this paper we focused on simulating genetic diseases as the metabolic disruptions are simple to translate into the metabolic model, but the next challenge will be converting more complex metabolic perturbations into explicit reaction modulations. Effects like toxic environmental exposure can be simulated once the mechanism is narrowed down, while the effect of diet could be modelled by varying the input nutrients via the exchange reactions of the network.

Conclusion

Building upon constraint based modelling of metabolism through the use of random sampling of fluxes, we were able to predict large potential metabolic profiles and confirm measured metabolites both in targeted and untargeted assays. Ranking all metabolites becomes possible through the methodology's comparison of flux distributions between healthy and disease states. Metabolites revealed by this method are of potential interest to broaden the panel of targets for future metabolomics experiments, and can be identified as understudied metabolites, helping to develop our understanding of metabolic mechanisms. Furthermore, the rank of a

given metabolite can be compared between two different disruption scenarios, which provides information on the specificity of the disrupted metabolite to the scenario.

Although the methodology is designed to be used to predict external metabolite exchange fluxes, it can also be used to simulate the internal reaction fluxes, which can be useful for understanding internal metabolism along with external metabolites. Finally, simulated metabolic profiles can also be used to benchmark various analyses specific to metabolomics, such as pathway analysis, or other analyses which require lots of data like machine learning.

Materials and methods

Metabolic models

In this study, Human 1 v1.14 [47], containing 13 024 reactions, was used to carry out mGWAS analyses. https://github.com/SysBioChalmers/Human-GEM Recon 2 [14], containing 7 440 reactions, was used to carry out IEM analyses. https://github.com/opencobra/COBRA.papers/tree/master/2013_Recon2 Recon 2 was used for the IEM analyses as the idea was to compare results between FVA and sampling for the same set of conditions in the same model. This served as a proof of concept and we decided to publish the results using Recon 2 to show the comparison with previous work by Thiele *et al.* [14]. Human 1 was then used for the mGWAS analyses as we believe it is a more complete model, and it is in the community's best interest to use the latest model since it can then be improved by community efforts. It also highlights that SAMBA can scale to a larger model. Note that model choice will have an impact on any modelling approach and this selection step, out of the scope of this article, has to be taken with care. The top plot of Fig D in S1 Text shows the application of Xanthinuria Type I to Human 1 (simulated using Recon 2 in the main text).

WT and MUT states

When choosing which metabolic network to use, a decision must be made on whether to optimise the biomass reaction or not. This is done by providing the name of the reaction in the network (as they are not named uniformly between networks) as well as the fraction of biomass to optimise for, as SAMBA inputs. A list of genes or reactions to knock-out must also be provided to create the MUT state. Exchange reaction bounds are set to define the cellular medium by using a specific parameter X to set all bounds to [-X, 1000]. They can instead be set upstream of SAMBA by changing the model exchange bounds directly and then setting the config parameter to use the default bounds.

The WT state is created using the default network parameters (reaction bounds, biomass coefficients etc.). Then, the reaction(s) to be knocked out are forced to carry a non-zero flux. This is done by optimising for the reaction(s) to KO and then changing the minimum bound to 5% of the maximum flux value, or maximum bound to -5%, for forward and reverse reactions respectively. This avoids sampling and comparing two states where the fluxes are zero for the reactions of interest, and is also why each WT is specific to a MUT state.

In the case of a KO, the MUT state is created using the default network followed by setting the upper bound ub and lower bound lb for the reaction(s) to KO to zero: $ub_{new} = lb_{new} = 0$, resulting in [0, 0] for each reaction related to the KO gene.

Model parameters

Sampling and FVA were run using the same parameters as in Shlomi *et al.* and Thiele *et al.* [13, 14]: minimum fraction of optimum of the objective function (biomass) set to 0, and all exchange reaction bounds set to [-1, 1000].

Sampling method

Random sampling is done using Python code written for SAMBA, based on the cobrapy [48] Python package. The code uses the CPLEX 12.10 solver by default and uses the optGpsampler algorithm [49] to sample from the reaction flux solution space. optGpsampler begins with a warm-up phase to select starting points (by running a preliminary FVA on each reaction), followed by uniform sampling within this feasible solution space. Because each sample is selected from the solution space directly, there is no sample rejection since this would be extremely inefficient to do on genome-scale models. A thinning parameter of k (default k = 100) means that every k sample is saved and the rest is discarded in order to reduce intersample correlation. For large models such as Recon 2 and Human1, 100 000 samples with a thinning of 100 were used.

The number of samples can be changed, however, in order to sufficiently explore the solution space, a large number (at least 100 000) of samples must be used for larger networks such as Recon 2 or Human 1 which contain thousands of reactions. Determining a sufficient number of samples is discussed below in Sampling Convergence and supplementary data.

Parallel processing

Sampling can be run on a local computer for smaller models, but it needs a certain amount of resources to run correctly. More specifically, the amount of RAM required increases with the size of the model, and more CPUs will help generate the samples faster.

For this study, the larger metabolic models (Recon2 and Human1) were sampled using a computer cluster using 16 cores and 128GB of RAM for each job. The cluster we used is the Genotoul computational cluster which has about 3000 cores / 600 threads, 36 Tera Byte memory (3TB on a SMP machine), Infiniband interconnection (QDR/FDR), parallel file system (GPFS).

Sampling convergence

One of the main limitations of CBM is that it is impossible to fully describe such a large solution space. When using random sampling to explore the solution space, the number of samples to use must be provided, but choosing the ideal number for a given network is a challenge since by definition the structure of the solution space is unknown.

Therefore it is essential to know when to stop sampling: determining when the solution space has been sufficiently sampled. We ran convergence tests using various well-known sampling metrics: running means, traceplots, and shrink factor plots, to make sure that using 100 000 samples was enough for a network this large, for the goal of calculating z-scores on distributions. The results can be found in the supplementary data (Fig C in S1 Text).

Plots & libraries

The SAMBA pipeline is managed using Snakemake. Venn diagrams were made using the R library eulerr. Sampling distributions and mGWAS plot tables were made using ggplot2. BiNChE: the BiNChE plot was created by using a docker containing an old version of Firefox and Flash. The subnetwork was exported and then taken into Cytoscape [50] to change the layout, followed by Inkscape to edit the placement of labels. Code availability The code for the SAMBA project is freely available at https://forgemia.inra.fr/metexplore/cbm/samba-project or https://doi.org/10.5281/zenodo.8369624.

Data

Using the mGWAS SNP dataset, genes were mapped to the Human 1 network using the annotated ENSG IDs. Then, using the model's GPR relationships, reactions were automatically knocked out using SAMBA. In the case of SCD, the GPRs were manually checked. The SCD SNP only affects the SCD1 gene (known as SCD in the metabolic model), as SCD1 and SCD5 are two separate genes. SCD5 codes for the same enzymatic function as SCD1 but they are both expressed in different tissues, fat tissue for SCD1 and brain and pancreas for SCD5. However, Human1 is not tissue-specific and the reactions are not necessarily associated with the genes according to this tissue specificity, so in order to block the enzymatic function completely, both SCD1 and SCD5 were blocked.

Resulting SAMBA metabolites were manually mapped to the mGWAS significant metabolite names for SCD, with manual verification of metabolite synonyms as many lipids have multiple names and naming conventions.

Supporting information

S1 Text. Fig A: Reproduction of Fig 3 from Thiele *et al.* 2013 using FVA (original matlab code from the study) and the random sampling method used in SAMBA. Fig B: Venn diagrams of FVA (orange) and sampling (purple) predictions using Recon 2. Fig C: Running means, trace plots and PSRF plots for 3 random exchange reaction fluxes using 100, 10 000 and 100 000 samples, with 3 independent runs for each. Fig D: Flux bounds and distributions for urate and hypoxanthine exchange reactions in Human1 when affected by Xanthinuria Type I, for different flux range reduction values. Fig E: Z-score vs other metric-based rankings for SCD. Fig F: Heatmap based on Fig A in S1 Text, using predicted ranks instead of z-scores. Fig G: Distribution of metabolite z-scores for SCD. Fig H: Distributions of FVA bound differences. Fig I. Distances from the top 50 most changed metabolites to each of the reactions affected by the SCD SNP. Fig J: Undirected subnetwork showing paths between the reactions affected by SCD and the top 50 predicted most changed metabolites for this condition. Fig K: ChemRich using top 54 metabolite sampling predictions for Fish-eye disease/ LCAT deficiency. Table A: Table of the top 10 most differentially changed metabolites for the SCD gene KO using SAMBA. Table B: ChemRich provides a metabolite-level table with each metabolite assigned to a cluster, the top 10 of which are shown in S2 Table.
(PDF)

S1 Table. All SAMBA-predicted z-scores and ranks for SCD.
(XLSX)

S2 Table. Significant metabolites for SCD from Suhre *et al.* converted to Human1 identifiers.
(XLSX)

Acknowledgments

We would like to thank Jake Bundy for his participation in our discussions involving this work. The authors are grateful to the GenoToul bioinformatics platform (Toulouse, France) for hosting and providing their computer cluster, used for running all sampling runs.

Author Contributions

Conceptualization: Juliette Cooke, Maxime Delmas, Cecilia Wieder, Pablo Rodríguez Mier, Clément Frainay, Timothy Ebbels, Nathalie Poupin, Fabien Jourdan.

Data curation: Juliette Cooke, Maxime Delmas, Clément Frainay.

Formal analysis: Juliette Cooke.

Funding acquisition: Fabien Jourdan.

Investigation: Juliette Cooke.

Methodology: Juliette Cooke, Fabien Jourdan.

Project administration: Timothy Ebbels, Fabien Jourdan.

Software: Juliette Cooke, Florence Vinson.

Supervision: Timothy Ebbels, Fabien Jourdan.

Validation: Juliette Cooke, Nathalie Poupin.

Visualization: Juliette Cooke.

Writing – original draft: Juliette Cooke, Nathalie Poupin, Fabien Jourdan.

Writing – review & editing: Juliette Cooke, Maxime Delmas, Cecilia Wieder, Pablo Rodríguez Mier, Clément Frainay, Florence Vinson, Timothy Ebbels, Nathalie Poupin, Fabien Jourdan.

References

1. Montgomery JE, Brown JR. Metabolic biomarkers for predicting cardiovascular disease. Vascular Health and Risk Management. 2013; 9:37–45. https://doi.org/10.2147/VHRM.S30378 PMID: 23386789
2. Nicholson JK, Lindon JC, Holmes E. 'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. Xenobiotica. 1999; 29(11):1181–1189. https://doi.org/10.1080/004982599238047 PMID: 10598751
3. Frainay C, Schymanski EL, Neumann S, Merlet B, Salek RM, Jourdan F, et al. Mind the Gap: Mapping Mass Spectral Databases in Genome-Scale Metabolic Networks Reveals Poorly Covered Areas. Metabolites. 2018; 8(3):E51. https://doi.org/10.3390/metabo8030051
4. Theodoridis G, Gika HG, Wilson ID. LC-MS-based methodology for global metabolite profiling in metabonomics/metabolomics. TrAC Trends in Analytical Chemistry. 2008; 27(3):251–260. https://doi.org/10.1016/j.trac.2008.01.008
5. Pan Z, Raftery D. Comparing and combining NMR spectroscopy and mass spectrometry in metabolomics. Analytical and Bioanalytical Chemistry. 2007; 387(2):525–527. https://doi.org/10.1007/s00216-006-0687-8 PMID: 16955259
6. Sumner LW, Amberg A, Barrett D, Beale MH, Beger R, Daykin CA, et al. Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). Metabolomics: Official journal of the Metabolomic Society. 2007; 3(3):211–221. https://doi.org/10.1007/s11306-007-0082-2 PMID: 24039616
7. Orth JD, Thiele I, Palsson BO. What is flux balance analysis? Nature Biotechnology. 2010; 28(3):245–248. https://doi.org/10.1038/nbt.1614 PMID: 20212490
8. Thiele I, Palsson BO. A protocol for generating a high-quality genome-scale metabolic reconstruction. Nature Protocols. 2010; 5(1):93–121. https://doi.org/10.1038/nprot.2009.203 PMID: 20057383
9. Sigurdsson MI, Jamshidi N, Steingrimsson E, Thiele I, Palsson BO. A detailed genome-wide reconstruction of mouse metabolism based on human Recon 1. BMC systems biology. 2010; 4:140. https://doi.org/10.1186/1752-0509-4-140 PMID: 20959003
10. Hädicke O, Klamt S. EColiCore2: a reference network model of the central metabolism of Escherichia coli and relationships to its genome-scale parent model. Scientific Reports. 2017; 7:39647. https://doi.org/10.1038/srep39647 PMID: 28045126
11. Bordbar A, Jamshidi N, Palsson BO. iAB-RBC-283: A proteomically derived knowledge-base of erythrocyte metabolism that can be used to simulate its physiological and patho-physiological states. BMC Systems Biology. 2011; 5:110. https://doi.org/10.1186/1752-0509-5-110 PMID: 21749716

12. Thiele I, Sahoo S, Heinken A, Hertel J, Heirendt L, Aurich MK, et al. Personalized whole-body models integrate metabolism, physiology, and the gut microbiome. Molecular Systems Biology. 2020; 16(5): e8982. https://doi.org/10.15252/msb.20198982 PMID: 32463598

13. Shlomi T, Cabili MN, Ruppin E. Predicting metabolic biomarkers of human inborn errors of metabolism. Molecular Systems Biology. 2009; 5:263. https://doi.org/10.1038/msb.2009.22 PMID: 19401675

14. Thiele I, Swainston N, Fleming RMT, Hoppe A, Sahoo S, Aurich MK, et al. A community-driven global reconstruction of human metabolism. Nature Biotechnology. 2013; 31(5):419–425. https://doi.org/10.1038/nbt.2488 PMID: 23455439

15. Mahadevan R, Schilling CH. The effects of alternate optimal solutions in constraint-based genome-scale metabolic models. Metabolic Engineering. 2003; 5(4):264–276. https://doi.org/10.1016/j.ymben.2003.09.002 PMID: 14642354

16. Smith RL. The hit-and-run sampler: a globally reaching Markov chain sampler for generating arbitrary multivariate distributions. In: Proceedings of the 28th conference on Winter simulation. WSC'96. USA: IEEE Computer Society; 1996. p. 260–264. Available from: https://dl.acm.org/doi/10.1145/256562.256619.

17. Price ND, Schellenberger J, Palsson BO. Uniform Sampling of Steady-State Flux Spaces: Means to Design Experiments and to Interpret Enzymopathies. Biophysical Journal. 2004; 87(4):2172–2186. https://doi.org/10.1529/biophysj.104.043000 PMID: 15454420

18. Fallahi S, Skaug HJ, Alendal G. A comparison of Monte Carlo sampling methods for metabolic network models. PLOS ONE. 2020; 15(7):e0235393. https://doi.org/10.1371/journal.pone.0235393 PMID: 32609776

19. Mo ML, Palsson BO, Herrgård MJ. Connecting extracellular metabolomic measurements to intracellular flux states in yeast. BMC Systems Biology. 2009; 3(1):37. https://doi.org/10.1186/1752-0509-3-37 PMID: 19321003

20. Dent CE, Philpot GR. Xanthinuria, an inborn error (or deviation) of metabolism. Lancet (London, England). 1954; 266(6804):182–185. https://doi.org/10.1016/S0140-6736(54)91257-X PMID: 13118765

21. Arikyants N, Sarkissian A, Hesse A, Eggermann T, Leumann E, Steinmann B. Xanthinuria type I: a rare cause of urolithiasis. Pediatric Nephrology. 2007; 22(2):310–314. https://doi.org/10.1007/s00467-006-0267-3 PMID: 17115198

22. Hamosh A, Scott AF, Amberger JS, Bocchini CA, McKusick VA. Online Mendelian Inheritance in Man (OMIM), a knowledgebase of human genes and genetic disorders. Nucleic Acids Research. 2005; 33 (Database Issue):D514–D517. https://doi.org/10.1093/nar/gki033 PMID: 15608251

23. Sahoo S, Franzson L, Jonsson JJ, Thiele I. A compendium of inborn errors of metabolism mapped onto the human metabolic network. Molecular BioSystems. 2012; 8(10):2545–2558. https://doi.org/10.1039/c2mb25075f PMID: 22699794

24. Suhre K, Shin SY, Petersen AK, Mohney RP, Meredith D, Wägele B, et al. Human metabolic individuality in biomedical and pharmaceutical research. Nature. 2011; 477(7362):54–60. https://doi.org/10.1038/nature10354 PMID: 21886157

25. Illig T, Gieger C, Zhai G, Römisch-Margl W, Wang-Sattler R, Prehn C, et al. A genome-wide perspective of genetic variation in human metabolism. Nature Genetics. 2010; 42(2):137–141. https://doi.org/10.1038/ng.507 PMID: 20037589

26. McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GRS, Thormann A, et al. The Ensembl Variant Effect Predictor. Genome Biology. 2016; 17(1):122. https://doi.org/10.1186/s13059-016-0974-4 PMID: 27268795

27. Longo N, Frigeni M, Pasquali M. CARNITINE TRANSPORT AND FATTY ACID OXIDATION. Biochimica et biophysica acta. 2016; 1863(10):2422–2435. https://doi.org/10.1016/j.bbamcr.2016.01.023 PMID: 26828774

28. Reddy GS, Sujatha M. A Rare Case of Short-Chain Acyl-COA Dehydrogenase Deficiency: The Apparent Rarity of the Disorder Results in Under Diagnosis. Indian Journal of Clinical Biochemistry. 2011; 26 (3):312–315. https://doi.org/10.1007/s12291-011-0139-x PMID: 22754199

29. Rani N, Hazra S, Singh A, Surolia A. Functional annotation of putative fadE9 of Mycobacterium tuberculosis as isobutyryl-CoA dehydrogenase involved in valine catabolism. International Journal of Biological Macromolecules. 2019; 122:45–57. https://doi.org/10.1016/j.ijbiomac.2018.10.040 PMID: 30316772

30. Roe CR, Cederbaum SD, Roe DS, Mardach R, Galindo A, Sweetman L. Isolated isobutyryl-CoA dehydrogenase deficiency: an unrecognized defect in human valine metabolism. Molecular Genetics and Metabolism. 1998; 65(4):264–271. https://doi.org/10.1006/mgme.1998.2758 PMID: 9889013

31. Moreno P, Beisken S, Harsha B, Muthukrishnan V, Tudose I, Dekker A, et al. BiNChE: A web tool and library for chemical enrichment analysis based on the ChEBI ontology. BMC Bioinformatics. 2015; 16 (1):56. https://doi.org/10.1186/s12859-015-0486-3 PMID: 25879798

32. Demirkan A, Duijn CMv, Ugocsai P, Isaacs A, Pramstaller PP, Liebisch G, et al. Genome-Wide Association Study Identifies Novel Loci Associated with Circulating Phospho- and Sphingolipid Concentrations. PLoS Genetics. 2012; 8(2). https://doi.org/10.1371/journal.pgen.1002490 PMID: 22359512

33. Paton CM, Ntambi JM. Biochemical and physiological function of stearoyl-CoA desaturase. American Journal of Physiology—Endocrinology and Metabolism. 2009; 297(1):E28–E37. https://doi.org/10.1152/ajpendo.90897.2008 PMID: 19066317

34. Jenkins B, West JA, Koulman A. A Review of Odd-Chain Fatty Acid Metabolism and the Role of Pentadecanoic Acid (C15:0) and Heptadecanoic Acid (C17:0) in Health and Disease. Molecules. 2015; 20 (2):2425–2444. https://doi.org/10.3390/molecules20022425 PMID: 25647578

35. Barupal DK, Fiehn O. Chemical Similarity Enrichment Analysis (ChemRICH) as alternative to biochemical pathway mapping for metabolomic datasets. Scientific Reports. 2017; 7(1):14567. https://doi.org/10.1038/s41598-017-15231-w PMID: 29109515

36. Ntambi JM. Regulation of stearoyl-CoA desaturase by polyunsaturated fatty acids and cholesterol. Journal of Lipid Research. 1999; 40(9):1549–1558. https://doi.org/10.1016/S0022-2275(20)33401-5 PMID: 10484602

37. Sajovic J, Meglič A, Glavač D, Markelj S, Hawlina M, Fakin A. The Role of Vitamin A in Retinal Diseases. International Journal of Molecular Sciences. 2022; 23(3):1014. https://doi.org/10.3390/ijms23031014 PMID: 35162940

38. Frainay C, Aros S, Chazalviel M, Garcia T, Vinson F, Weiss N, et al. MetaboRank: network-based recommendation system to interpret and enrich metabolomics results. Bioinformatics (Oxford, England). 2019; 35(2):274–283. https://doi.org/10.1093/bioinformatics/bty577 PMID: 29982278

39. Smith CA, Want EJ, O'Maille G, Abagyan R, Siuzdak G. XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. Analytical Chemistry. 2006; 78(3):779–787. https://doi.org/10.1021/ac051437y PMID: 16448051

40. Haraldsdóttir HS, Cousins B, Thiele I, Fleming RMT, Vempala S. CHRR: coordinate hit-and-run with rounding for uniform sampling of constraint-based models. Bioinformatics. 2017; 33(11):1741–1743. https://doi.org/10.1093/bioinformatics/btx052 PMID: 28158334

41. Jadebeck JF, Wiechert W, Nöh K. CHRRT: boosting coordinate hit-and-run with rounding by thinning; 2022. Available from: https://www.biorxiv.org/content/10.1101/2022.11.17.516802v1.

42. Theorell A, Jadebeck JF, Nöh K, Stelling J. PolyRound: polytope rounding for random sampling in metabolic networks. Bioinformatics. 2022; 38(2):566–567. https://doi.org/10.1093/bioinformatics/btab552 PMID: 34329395

43. Hui S, Cowan AJ, Zeng X, Yang L, TeSlaa T, Li X, et al. Quantitative Fluxomics of Circulating Metabolites. Cell Metabolism. 2020; 32(4):676–688.e4. https://doi.org/10.1016/j.cmet.2020.07.013 PMID: 32791100

44. Özen H. Glycogen storage diseases: New perspectives. World Journal of Gastroenterology: WJG. 2007; 13(18):2541–2553. https://doi.org/10.3748/wjg.v13.i18.2541 PMID: 17552001

45. Shlomi T, Cabili MN, Herrgård MJ, Palsson BO, Ruppin E. Network-based prediction of human tissue-specific metabolism. Nature Biotechnology. 2008; 26(9):1003–1010. https://doi.org/10.1038/nbt.1487 PMID: 18711341

46. Rodríguez-Mier P, Poupin N, Blasio Cd, Cam LL, Jourdan F. DEXOM: Diversity-based enumeration of optimal context-specific metabolic networks; 2020. Available from: https://www.biorxiv.org/content/10.1101/2020.07.17.208918v1.

47. Robinson JL, Kocabaş P, Wang H, Cholley PE, Cook D, Nilsson A, et al. An atlas of human metabolism. Science Signaling. 2020; 13(624):eaaz1482. https://doi.org/10.1126/scisignal.aaz1482 PMID: 32209698

48. Ebrahim A, Lerman JA, Palsson BO, Hyduke DR. COBRApy: COnstraints-Based Reconstruction and Analysis for Python. BMC Systems Biology. 2013; 7(1):74. https://doi.org/10.1186/1752-0509-7-74 PMID: 23927696

49. Megchelenbrink W, Huynen M, Marchiori E. optGpSampler: An Improved Tool for Uniformly Sampling the Solution-Space of Genome-Scale Metabolic Networks. PLOS ONE. 2014; 9(2):e86587. https://doi.org/10.1371/journal.pone.0086587 PMID: 24551039

50. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Research. 2003; 13 (11):2498–2504. https://doi.org/10.1101/gr.1239303 PMID: 14597658