npj Systems Biology and Applications https://doi.org/10.1038/s41540-026-00760-w

Article in Press

A mathematical model of folate-mediated one-carbon metabolism in Down syndrome

Received: 20 November 2025

Accepted: 28 May 2026

Cite this article as: Piovesan, A., Poluzzi, D., Ramacieri, G. *et al.* A mathematical model of folate-mediated one-carbon metabolism in Down syndrome. *npj Syst Biol Appl* (2026). https://doi.org/10.1038/s41540-026-00760-w

Allison Piovesan, Diletta Poluzzi, Giuseppe Ramacieri, Chiara Locatelli, Francesca Antonaros, Beatrice Vione, Maria Caracausi, Maria Chiara Pelleri & Luca Marchetti

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

ARTICLE IN PRESS

© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

ARTICLE IN PRESS

A mathematical model of folate-mediated one-carbon metabolism in Down syndrome

Allison Piovesan[1,§], Diletta Poluzzi[2,3,§], Giuseppe Ramacieri[4], Chiara Locatelli[5], Francesca Antonaros[1], Beatrice Vione[1,6], Maria Caracausi[1], Maria Chiara Pelleri[1], Luca Marchetti[3,7,*]

[1]Department of Biomedical and Neuromotor Sciences (DIBINEM), Unit of Histology, Embryology and Applied Biology, University of Bologna, Via Belmeloro 8, 40126 Bologna (BO), Italy

[2]Department of Mathematics, University of Trento, Via Sommarive 14, 38123 Povo (TN), Italy

[3]Fondazione The Microsoft Research, University of Trento, Centre for Computational and Systems Biology (COSBI), Piazza Manifattura, 1, 38068 Rovereto (TN), Italy

[4]Specialty School of Child Neuropsychiatry, University of Bologna, Via Massarenti 9, 40138 Bologna (BO), Italy

[5]Neonatology Unit, St. Orsola-Malpighi Polyclinic, Department of Medical and Surgical Sciences (DIMEC), University of Bologna, Via Massarenti 9, 40138 Bologna (BO), Italy

[6]Department of Medical and Surgical Sciences (DIMEC), Unit of Histology, Embryology and Applied Biology, University of Bologna, Via Belmeloro 8, 40126 Bologna (BO), Italy

[7]Department of Cellular, Computational and Integrative Biology (CIBIO), University of Trento, Via Sommarive 9, 38123 Povo (TN), Italy

[§]These authors contributed equally to this work

[*]To whom correspondence should be addressed: luca.marchetti@unitn.it marchetti@cosbi.eu

ARTICLE IN PRESS

ARTICLE IN PRESS

Abstract

Down syndrome (DS), the most frequent human genetic disorder marked by an extra copy of chromosome 21 (Hsa21) or a portion thereof, leads to physical and cognitive impairments. Following the Lejeune work, researchers focused on a potential anomaly within the folate-mediated one-carbon metabolism (FOCM). Here we present a FOCM model modified from a previous work with the incorporation of the enzyme CBS, whose encoding gene is located on Hsa21, coupled with the methionine input rate. Systematic perturbation of enzyme activity rates has been performed to explore possible *in silico* configurations to simulate the DS condition. The perturbed vs. unperturbed model-derived ratio concentrations of tetrahydrofolate, 5-formyl-tetrahydrofolate, 5-methyl-tetrahydrofolate, S-adenosyl-homocysteine, and S-adenosyl-methionine were compared with the known literature through various statistical approaches. After investigating public transcriptomic databases, the FTS (formate-tetrahydrofolate ligase) perturbation achieved the best overall score. Although the FTS encoding gene (*MTHFD1*) is not located on Hsa21, it was found to be over-expressed in the DS condition. In addition, an interesting correlation emerged with the PTG (phosphoribosylglycinamide formyltransferase) perturbation and the corresponding encoding gene, located on Hsa21 (*GART*) and notably over-expressed in the DS condition. The model thus identifies key enzyme activities that warrant further investigation.

ARTICLE IN PRESS

ARTICLE IN PRESS

Introduction

Down syndrome (DS) or trisomy 21[1] is the most frequent human genetic disorder in which an individual possesses an extra full or partial copy of human chromosome 21 (Hsa21)[2], with an incidence of 1 in 700–1,000 live births globally[3]. It is also the most frequent leading genetic cause of intellectual disability, which is of variable degree and always present together with other variable characteristic signs and symptoms[3-10]. While various interventions and therapies are available to address the physical manifestations linked to DS, there is currently a lack of specific treatments aimed directly at cognitive difficulties.

Lejeune hypothesized that a disturbance of the one-carbon cycle might occur in DS[11] and several other Authors reported the one-carbon cycle to be imbalanced in subjects with DS[12-18]. Folate-mediated one-carbon metabolism (FOCM) includes the folate cycle, which plays a vital role in delivering one-carbon units at different oxidation states to the homocysteine-methionine cycle for methylation reactions and biosynthesis of several metabolites. These include the biosynthesis of purines and thymidine, the maintenance of amino acid balance, the regulation of epigenetic functions, and the defence against oxidative stress. At the heart of this intricate process is methionine synthase (MS), the enzyme responsible for converting homocysteine (HCY) to methionine (MET) by using 5-methyl-tetrahydrofolate (5mTHF) as a methyl donor and vitamin B12 as an essential cofactor. HCY is converted to cysteine through the transsulfuration pathway mediated by cystathionine ß-synthase (CBS)[19,20]. MET is a precursor of S-adenosyl-methionine (SAM) which serves as a methyl donor for DNA methylation, protein methylation, and lipid and neurotransmitter synthesis[21].

Many of the biochemical reactions in the FOCM pathway depend on the concentrations of their substrates in a nonlinear way. Therefore, the behavior of the full cycle is relatively complex and cannot be easily deduced from the reaction diagram itself, making it a perfect candidate to be investigated through a mathematical model. Over the years, significant progress has been made in understanding FOCM, with numerous models developed to explain its pathways. The earliest models emerged in the 1970s and 1980s, focusing primarily on the impact of anticancer drugs on the system[22]. Notably, Michael C. Reed and Colleagues played a pivotal role in advancing these models. Firstly, they developed the homocysteine-methionine and folate cycle models separately[23,24] and later integrated them[25], thus contributing extensively to the understanding of FOCM across diverse scenarios and diseases. The integrated model simulates the interconnected biochemical reactions

ARTICLE IN PRESS

occurring in the cytoplasm and served as the foundation for the subsequent models developed and validated by Misselbeck *et al.*[26,27].

The aim of this work is to extend and leverage the mathematical model of the FOCM previously introduced in Misselbeck *et al.*[26,27] to explore possible *in silico* configurations of the FOCM that can be linked to the DS condition. The original model was extended through the incorporation of the reaction catalyzed by the enzyme CBS (of which the encoding gene is located on Hsa21), coupled with the MET input rate. The model behavior was newly evaluated under specific experimental conditions to validate the reliability of the extended model. Then, perturbations of enzyme activity rate were systematically performed to simulate DS imbalances, and the new steady state of the perturbed models was computed. The ratios of the model-predicted metabolite amounts in the perturbed vs. unperturbed model were compared to plasmatic DS vs. non-DS concentration ratios available in the literature for tetrahydrofolate (THF), 5mTHF, SAM, S-adenosyl-homocysteine (SAH) and 5-formyl-THF (5fTHF). This allowed us to identify a set of enzyme perturbations that, according to model simulation, can redistribute the folate forms within the FOCM pathway as reported in the data currently available for DS.

The analysis presented here will provide an *in silico* approach to investigate FOCM dysregulation in DS at a systemic level. By identifying potential regulatory components within the biological network, this work aims to enhance our understanding of the disease and the underlying mechanisms contributing to intellectual disability during development.

ARTICLE IN PRESS

Results

Extended FOCM model and perturbations

The starting published FOCM model has been extended in the present work through the incorporation of the reaction catalyzed by the enzyme CBS coupled with the MET input rate (Figure 1). The MATLAB code to simulate the extended mathematical model employed in this study is available at https://github.com/cosbi-research/FocmDownSyndrome. A validation procedure to ensure the reliability of the newly extended model was performed by a sequence of simulations regarding the effect of glycine on FOCM, the effect of different levels of MTHFR activity and the effect of different MTHFS and SHMT activity[26,27]. In all three considered

ARTICLE IN PRESS

validation cases, our extended model showed a good agreement with the experimental data (Supplementary Note 2 and Supplementary Data 1), comparable to the one computed by the previous version of the model, indicating that the behavior of the extended model remains consistent with the observations presented in the original papers [26,27].

Subsequently, the validated extended model was used and each of the 17 FOCM reactions was then perturbed by a reduction in enzyme activity rate by 2/3 or an enhancement by 3/2 to simulate possible cascade effects induced by the presence of the third copy of Hsa21 characteristic of the DS condition. As detailed in the Methods section, for each perturbation, a ratio was computed between the value reached at the steady state after enzyme activity rate perturbation representing DS condition and the initial steady state value (related to the unperturbed model as a non-DS control) for all 14 model variables. Computed value ratios for single enzyme activity perturbation by 2/3 or 3/2 for all variables are detailed in Supplementary Data 2. Notably, perturbation of MS and MTHFR activity rates by 2/3 and 3/2 induces changes in all 14 model variables, in some cases of substantial magnitude. On the other hand, other enzyme perturbations mostly changed their substrates or products, while BHMT, SAHH, GNMT, DNMT and CBS perturbations caused the greatest changes in all the homocysteine-methionine cycle variables, where they are involved. Table 1 shows the selection of computed value ratios for the five metabolites THF, 5mTHF, SAM, SAH and 5fTHF, which were successively compared with experimental data available from the literature through procedures based on relative error calculation and bootstrap technique.

ARTICLE IN PRESS

Effects of enzyme activity perturbation on FOCM model with fixed uncertainties

For the five metabolites THF, 5mTHF, SAM, SAH and 5fTHF, the ratio computed between the value reached at the steady state after perturbation representing DS condition and the initial steady state value (related to unperturbed model as a control) was compared to experimental DS/non-DS ratio values through a procedure based on relative error calculation. The first point of view to assess the extent of consistency between the ratios obtained through enzyme activity perturbations and the experimental data is to consider the total cumulative score ranging from 0 (complete lack of concordance) to 5 (concordance for all the five variables) for various

ARTICLE IN PRESS

predefined uncertainties for each enzyme (Supplementary Data 3). Table 2 shows a compact summary version with fixed uncertainties for which remarkable changes were evident.

When examining enzyme activity perturbation by 2/3, the score for each enzyme activity rate perturbation increases when uncertainty increases. The most unfavorable perturbation outcomes are associated with MTHFR, which reaches the highest score of 1 with a fixed uncertainty of 0.6, followed by MTHFS, which is the last to reach the score of 2 with a fixed uncertainty of 0.234. The first score of 2 is observed with a fixed uncertainty of 0.071 in the case of the GNMT enzyme perturbation, which also first reaches the score of 3 at a fixed uncertainty of 0.091. The score of 4 is first attained with a fixed uncertainty of 0.234 by the enzymes PGT and AICARFT. At a fixed uncertainty of 0.397, more than half of the perturbations (9) give a score of 4. The best score of 5 is reached by CBS perturbation, but only at 0.6 of uncertainty (Table 2).

Enzyme activity perturbation by 3/2 also shows an increase in score with increasing uncertainty. The enzyme MS produces the least favorable perturbation outcomes, with 1 as the maximum score at 0.58 uncertainty, followed by MTHFS, which is the last to reach the score of 2 with a fixed uncertainty of 0.213. The first score of 3 is observed with a fixed uncertainty of 0.213 in the case of the FTS enzyme perturbation. Overall, the best score is 4, reached with the first fixed uncertainty of 0.234 by the enzyme FTS. At a fixed uncertainty of 0.417, more than half of the perturbations (9) give a score of 4 (Table 2).

The second point of view to assess the extent of ratio concordance with experimental data is focusing on each of the five metabolites and their corresponding total scores. These range from 0 to 17, indicating that from no enzyme to all 17 enzymes' activity rates subject to perturbation gives a variable ratio value consistent with experimental reference data (Supplementary Data 4). These results highlight a general system perturbation at low levels of uncertainty and that enzyme activity perturbations almost never reach concordance with all five experimental ratio values because achieving consistency for the metabolite SAH seems to pose a greater challenge. Specifically, when perturbing enzymes by a factor of 2/3, concordance is achieved by none or one enzyme perturbation at a time from the fixed uncertainty of 0.417. Even more notable, when perturbing by a factor of 3/2, the score for SAH is always 0. Regarding the other metabolites, when uncertainty increases, the score increases as well, but not linearly. When perturbing by a factor of 2/3, the concordance for THF jumps from 6 (fixed uncertainty of 0.274) to 15 (fixed uncertainty of 0.295); the concordance for 5mTHF jumps from

ARTICLE IN PRESS

2 (fixed uncertainty of 0.091) to 10 (fixed uncertainty of 0.112); the concordance for SAM jumps from 2 (fixed uncertainty of 0.274) to 12 (fixed uncertainty of 0.295); the concordance for 5fTHF jumps from 2 (fixed uncertainty of 0.03) to 10 (fixed uncertainty of 0.051). When perturbing by a factor of 2/3, we observed a similar trend, with similar jumps at the same fixed uncertainty levels (Supplementary Data 4). While this method is useful to visually highlight how perturbations vary at different uncertainties, it is difficult to choose a fixed value for biological discussion. For this purpose, the comparison with the bootstrap approach described in the following section was performed.

Effects of enzyme activity perturbation on FOCM model with the bootstrap procedure

A bootstrap procedure was used to create *in silico* DS and non-DS datasets starting from the original experimental dataset for each of the five metabolites of interest. Due to practical limitations, the availability of metabolite concentration data varies for each patient, leading to non-uniform widths within the dataset (Supplementary Table 6 published in Vione *et al.*[18]). The real distribution of the experimental data populations was used to generate *in silico* datasets for the five distinct metabolites and the original and generated distributions are shown in Supplementary Figure 1. Therefore, we conducted an analysis for each of the five metabolites to determine whether the model computed variable ratio, taking into account the concentration under conditions with DS (where enzyme activity is perturbed) compared to the concentration in a control state (unperturbed model), falls within predefined intervals under the assumption of a general distribution (intervals 3 and 4 in Methods) derived from the bootstrap procedure with 1000 iterations (Table 3). A score of 1 was assigned if the criteria were met; otherwise, it was assigned as 0.

Considering the [First Quantile, Third Quantile] interval, the highest score obtained for enzymes perturbed by 2/3 was 1, and it was achieved by MTD. Meanwhile, enzymes perturbed by 3/2 resulted in a maximum score of 1 for three enzymes (MTHFR, GNMT, DNMT) (Supplementary Data 5).

Considering the [First Quantile - 1.5 · IQR, Third Quantile + 1.5 · IQR] interval, a score of 4 was achieved for two different enzymes (PGT and AICARFT) under a perturbation of 2/3 (Table 4). For enzymes perturbed by 3/2, only FTS obtained a score of 4 (Table 4). This analysis again highlights the difficulty in obtaining a computed variable ratio for SAH in agreement with experimental data. As in the previous interval,

ARTICLE IN PRESS

only the SAHH perturbation by 2/3 is in agreement for this metabolite (with a total score of 2 for this enzyme). Overall, these scores better resemble the score results obtained by 0.254 fixed uncertainty for enzyme activity rate perturbation by 2/3 and by 0.274 fixed uncertainty for perturbation by 3/2, with 13 out of 17 equal scores in both cases (Supplementary Data 6). Gradually expanding the intervals covering the entire population, the scores increase for each enzyme and for both types of activity rate perturbation. A closer agreement of the computed variable ratio for SAH with experimental data is shown with the interval [Minimum Value -0.5 · Minimum Value, Maximum Value + 0.5 · Maximum Value] which considers a variability of 50% for the minimum and maximum bounds of the interval, where the total score of 5 is obtained by 2/3 activity rate perturbation of SAHH and CBS (Supplementary Data 5). Results for other tested intervals did not highlight any other particular perturbations and are available upon request.

The overall score associated with the metabolite SAH remains consistently low across most of the analyzed intervals, thus the total score of 4 for each enzyme, accounting for consistent results with THF, 5mTHF, SAM, and 5fTHF, is truly noteworthy.

Gene expression analysis

Available quantitative transcriptome maps[28] were used to investigate ratio expression values between DS and non-DS conditions for genes encoding for FOCM enzymes (Table 5), considering enzyme scores at the predefined uncertainties of 0.254 (2/3 activity perturbation) and 0.274 (3/2 activity perturbation), which better resemble the bootstrap result procedures (Supplementary Data 7). Focusing on the brain transcriptome map, the highest scores for 2/3 enzyme activity perturbation concordant with gene expression ratio ≤0.76 are found only for *GNMT* (with an expression ratio value of 0.73, encoding for GNMT with a score of 3) and *MAT1A* (with an expression ratio value of 0.70, encoding for MAT-I and MAT-III with both a score of 3). The transcriptome map of all tissues did not show under-expressed genes. TissueNexus brain dataset highlighted, among the others, a connection of 0.8177 between *GNMT* and *CBS* (which emerged for high scores when its enzyme activity is perturbed by 2/3, in contrast with its over-expression), a connection of 0.8879 between *MAT1A* and *BHMT* (of which encoded enzymes never emerged in our perturbation analyses), and a connection of 0.844 between *MAT1A*

ARTICLE IN PRESS

and *MTHFD1* (Supplementary Data 8). Interestingly, *CBS*, *GNMT* and *MAT1A* show the lowest number of connections with other FOCM genes with edge weight >0.5 (Table 6).

The highest scores for 3/2 enzyme activity perturbation concordant with gene expression ratio ≥1.30 are found in the brain transcriptome map for *DHFR* (with an expression ratio value of 1.64, encoding for DHFR), *MTHFD1* (with a borderline expression ratio value of 1.28, encoding for FTS), *GART* (gene located on Hsa21 with an expression ratio of 1.43 and encoding for PTG), and *TYMS* (with an expression ratio value of 1.43, encoding for TYMS). The transcriptome map of all tissues confirmed the concordance between the highest scores for 3/2 enzyme activity perturbation and over-expression of *MTHFD1* (with an expression ratio value of 1.31) and *GART* (with an expression ratio of 1.31). Notably, the other Hsa21 gene (*CBS*, with an expression ratio value of 1.52 in the brain and 1.31 in all tissues' maps, encoding for CBS and located on Hsa21) resulted in being over-expressed in both transcriptome maps, but the encoded CBS enzyme is associated with low scores when 3/2 activity perturbation was performed (2 with fixed uncertainty or 1 with bootstrap procedure). TissueNexus brain dataset highlighted *TYMS* as the gene with the highest number of overall connections with edge weight >0.5, while *DHFR* and *MTHFD1* are the genes with the highest number of connections with other FOCM genes with edge >0.5 (Table 6). Among the others, *DHFR* shows high connections with *TYMS* (0.9454), *DNMT1* (0.9427), *MTHFD1* (0.9294), *SHMT1* (0.9081) and *GART* (0.873). *MTHFD1* is highly connected with *GART* (0.9396), *TYMS* (0.9343), *SHMT1* (0.9218), *AHCY* (0.9157) and *ATIC* (0.9036). *GART* is highly connected with *AHCY* (0.9314) and *TYMS* (0.9233) (Supplementary Data 8).

ARTICLE IN PRESS

Discussion

In the present work we have explored possible *in silico* configurations of the FOCM pathway to simulate imbalances in the DS condition. An original FOCM model has been extended here through the incorporation of the reaction catalyzed by the enzyme CBS (encoded by a gene located on Hsa21 that is found in an additional copy in DS), coupled with the MET input rate, leading to an open model more in agreement with the actual biological functioning of the pathway. The validation procedure performed on the newly extended model highlighted the consistency with previously reported findings described in the original papers[26,27], proving its reliability. The extended validated model has then been exploited to systematically perturb each enzyme activity

ARTICLE IN PRESS

rate, individually reducing or increasing each of the 17 model reaction activity rates, one at a time by a factor of 2/3 or 3/2, respectively, and computing each new steady state to simulate inhibitory or enhancing pathway imbalances potentially driven by the additional third copy of Hsa21 in DS condition. After each perturbation and simulation, the model variable values obtained at the new perturbed steady state (resembling possible hypotheses of DS condition) were compared with initial variable values at the steady state of the unperturbed model generating a computed model-derived DS/non-DS ratio. These ratios highlighted that the perturbed model configuration at steady state is affected in different ways by different enzyme activity rate perturbations with some of them more effectively altering than others, probably reflecting the relative enzymes that are more connected or more peripheral in the pathway.

The computed model-derived variable ratios were then in turn compared with experimental plasma concentration ratios measured in DS/non-DS children and young individuals and reported in the literature for five metabolites of interest: THF, 5mTHF, SAM, SAH and 5fTHF (see Figure 4 for a summary of the main results presented in this study). The aim was to identify the set of enzyme perturbations capable of redistributing the folate forms within the FOCM pathway, consistent with currently available data for DS at a young age. This assessment of the concordance between model-derived and experimental ratios was aided by utilizing a relative error calculation in combination with predetermined levels of uncertainty. As one might expect, overall, the concordance of each enzyme activity rate perturbation gives variable ratios that are more concordant with experimental ratios as the uncertainty increases. However, this approach also revealed that the concordance with all five experimental ratio values is almost never reached because achieving consistency for the metabolite SAH seems to pose a greater challenge. This may be attributed to the substantial difference in the number of data points between the DS and non-DS groups. Since the estimation of the best fixed uncertainty threshold δ to analyse the results could be difficult to identify, we also implemented a bootstrap procedure based on a randomized number generation approach to establish a data-driven approach for evaluating the concordance between model-derived variable ratios and experimental data. This procedure showed that the concordance increases for each enzyme and for both types of activity rate perturbation as the intervals are gradually expanded, covering the entire population. This approach highlights again that the overall score associated with the metabolite SAH remains consistently low across most of the analyzed intervals, thus the achievement of a total

ARTICLE IN PRESS

score of 4 for each enzyme, accounting for consistent results with THF, 5mTHF, SAM, and 5fTHF, is truly noteworthy. The scores obtained with the [First Quantile - 1.5 • IQR, Third Quantile + 1.5 • IQR] interval, which covers 98% of the data in a general distribution, better resemble the score results obtained by 0.254 fixed uncertainty for 2/3 enzyme activity perturbation and 0.274 fixed uncertainty for 3/2 enzyme activity perturbation. Thus, these results were taken into consideration when investigating the ratio expression values between DS and non-DS conditions for genes encoding FOCM enzymes. Analysing brain and all tissues' transcriptome maps available in the literature, concordance between gene under-expression ratio values and encoded enzyme 2/3 activity perturbation was found for GNMT, MAT-I, and MAT-III, which obtained a score of 3 in the comparison of model-derived with experimental metabolite ratios. Gene over-expression values concordance with encoded enzyme 3/2 activity perturbation was notably not favourable for CBS (encoded by *CBS* on Hsa21) as obtained model-derived vs. experimental metabolite ratio concordance scores of 2 with fixed uncertainty and 1 with bootstrap procedure. Gene over-expression value concordance was found instead for DHFR, PGT, SAHH and TYMS which obtained a score of 3 in the comparison of model-derived metabolite with experimental ratios and for FTS which notably obtained a score of 4 in both fixed uncertainties and bootstrap procedures, thereby strengthening our findings. Its encoding gene (*MTHFD1*), although located on chromosome 14, has one of the highest numbers of connections with other FOCM genes in TissueNexus database. This comes from the fact that FTS is an enzyme directly connected with THF, which has already been pointed out as the folate form most interconnected in the folate pathway and product of ten enzymatic reactions[18]. FTS holds a critical position as it facilitates the entry of formate and THF into FOCM through its three distinct enzymatic activities (5,10-methylenetetrahydrofolate dehydrogenase, 5,10-methenyltetrahydrofolate cyclohydrolase and 10-formyltetrahydrofolate synthetase), which catalyse the sequential interconversion of one-carbon derivatives of THF for methionine and thymidylate syntheses, and generate 10fTHF for *de novo* purine syntheses driven by PGT[29]. PGT also achieved good scores in 3/2 perturbation enzyme activity rate in line with the over-expression ratio of its encoding gene *GART*, which is located on Hsa21, also confirmed by proteome analysis in trisomy 21 versus euploid fibroblast cells[30]. Unfortunately, this and other proteome data available in the literature for the DS condition[31] did not allow us to confirm other results.

ARTICLE IN PRESS

The primary limitation of our study is the lack of transcriptomic, proteomic, and metabolomic data available for genes, metabolites, and enzymes of our interest and central to the FOCM pathway together for the same DS and non-DS sample groups. These data could help to better investigate our first assumption that the additional third copy of Hsa21 (in a ratio of 3:2 compared to the normal pair of Hsa21) may result in direct alteration of 3:2 or 2:3 of the FOCM enzyme activities. Another limitation concerns the quality and quantity of the literature data used to compare the computed model-derived variable ratios in our analysis for the five metabolites of our interest (THF, 5mTHF, SAM, SAH, and 5fTHF). In particular, the relatively low number of experimental data points for the SAM metabolite and the general difference in length between the DS and non-DS datasets in Vione *et al.*[18] should be highlighted as potential areas of improvement when more data points become available. While other datasets of plasma metabolomic analysis in trisomy 21 subjects compared with control samples are available[32,33], the measured abundance values refer to metabolites not central in the FOCM pathway and the average ages (mainly adults) is not comparable with the ones of our analysis. However, our tool is also made available for similar simulations (MATLAB code available at https://github.com/cosbi-research/FocmDownSyndromehttps://github.com/cosbi-research/FocmDownSyndrome), such as those involving other types of enzyme activity rate perturbations, when these data are made available in the literature. Finally, a promising model extension could be the inclusion of nuclear *de novo* thymidylate synthesis in the model, which is known to be affected in many pathologies resulting from impairments in one-carbon metabolism[34]. An ongoing research effort in this direction is currently underway within our research group.

In conclusion, we presented a FOCM model modified *ad hoc* to explore possible *in silico* configurations that can be linked to the DS condition and evaluate the concordance with experimental evidence. Our results, also supported by transcriptome evidence, strengthen the crucial role of the entrance of THF in the one-carbon cycle and that FTS activity needs to be investigated more in detail, as it had also been found to be over-expressed in fetuses with DS compared with control group (Cl-tetrahydrofolate synthase)[35].

Methods

Extension of FOCM model to simulate the Down syndrome imbalances

ARTICLE IN PRESS

The starting FOCM model used as the foundation of subsequent extensions and perturbations performed here to simulate DS imbalances derived from previous works[26,27]. Briefly, in 2017 Misselbeck *et al.*[26] developed a mathematical model to represent the FOCM pathways in cytoplasm, treating it as a closed system of reactions, expanding upon the fundamental work by Reed *et al.*[25]. The model included 12 variables that represented distinct forms of folate (THF, 10fTHF, 5,10-methenyl-THF, 5,10-methylen-THF, dihydrofolate, 5mTHF), the enzyme serine hydroxymethyltransferase (SHMT), the complex 5mTHF:SHMT, formed through the binding of 5mTHF and SHMT, and various metabolites (HCY, MET, SAM, and SAH). Certain substrates (nicotinamide adenine dinucleotide phosphate reduced form or NADPH, oxidated form or $NADP^+$, deoxyuridine monophosphate or dUMP, serine, glycine, glycinamide ribonucleotide or GAR, 5-aminoimidazole-4-carboxamide ribonucleotide or AICAR, betaine, and formate) were considered constants over time, based on observations from earlier work[25]. The model was further extended in the subsequent study[27] by incorporating the futile cycle, integrating 5fTHF and methenyltetrahydrofolate synthetase (MTHFS) along with the relevant reactions involving these variables. The extended model thus consisted of 14 variables and 17 reversible and irreversible enzyme reactions implemented as a set of Ordinary Differential Equations (ODEs), where reaction rates were specified by Michaelis-Menten formulas. Parameter estimates were obtained from literature or calculated through nonlinear least square optimization as detailed in Misselbeck *et al.*[26,27]. Kinetic constants were derived from polyglutamate cofactors and their interaction with enzymes, mainly purified from L1210 (murine lymphocytic leukemia cell line) cells or other mammalian tissue when possible[26].

In order to obtain a more suitable FOCM model for simulating the DS imbalances, the model publicly available in Misselbeck *et al.*[26,27] has been further extended in the present work (leading to an extended model as named from here on) through the incorporation of the reaction catalyzed by the enzyme CBS coupled with the MET input rate, as graphically represented in the model diagram provided in Figure 1.

CBS, encoded by the *CBS* gene located on Hsa21, is a key player in the transsulfuration pathway. The intricacies of the CBS-catalyzed reaction's kinetics, activated by SAM, are clarified through the utilization of the following Michaelis-Menten formula (R: reaction, V: velocity), introduced in Reed *et al.*[25].

$$R_{CBS} : HCY \xrightarrow{CBS} cystathionine \quad (1)$$

ARTICLE IN PRESS

$$V_{CBS} = \left(\frac{V_{max}\,[HCY]}{K_m + [HCY]}\right)\left(\frac{(1.2)([SAM]+[SAH])^2}{30^2 + ([SAM]+[SAH])^2}\right) \quad (2)$$

where the Michaelis-Menten constant is *Km* = 1000 µM and the maximum rate of the reaction is *Vmax* = 90,000 µM/h.

The incorporation of the MET input rate was guided by the insights provided by Reed *et al.*[25] based on the work by Storch *et al.*[36], which considered both fasting and non-fasting states. As the context under consideration may involve scenarios both during fasting and non-fasting states, an average value of 100 µM for the MET input rate was adopted here, capturing a broader range of situations. The thus extended model, in agreement with Reed *et al.*[25], is no longer a closed model and should, therefore, be more in agreement with the actual biological functioning of the pathway.

Validation of the extended FOCM model

The modifications introduced in the newly extended model, as detailed above, might have changed the behaviour of the model itself. In order to ensure the reliability of the newly extended model and the consistency with previously reported findings, we performed a validation procedure by reproducing the experimental scenarios originally described in the original papers[26,27]. The validation step was carried out by performing a series of simulations using the MATLAB integrator ode15s and comparing the results with those obtained from the original model under the following three conditions. (1) The effect on FOCM of varying the amount of glycine, essential to compare the model's results with empirical data generated in MCF-7 cells, as observed by Herbig et al. (2002)[37]; simulations were carried out by altering glycine levels from 0 mM to 10 mM, while observing the variable concentration trends at the steady state of the model. (2) The effect on FOCM of different levels of methylenetetrahydrofolate reductase (MTHFR) activity, a critical enzyme involved in the folate cycle, to observe how the model responded to alterations due to C667T polymorphism[38]. Simulations were conducted at a steady state with diverse values of the parameter associated with the MTHFR-catalyzed reaction, spanning from 2 times to 0.3 times that of the wild type. (3) The effect on FOCM of different MTHFS and SHMT activity[26,27]. Simulations were executed under various scenarios for the activity of SHMT and MTHFS enzymes while considering model variables at a steady state. The specific conditions tested are the following: standard

ARTICLE IN PRESS

MTHFS and SHMT activity (parameters for MTHFS and SHMT were maintained at standard values); halved MTHFS activity and standard SHMT activity (SHMT activity remained unaffected while the MTHFS-catalyzed reaction parameter was reduced to 0.5 times its value); negligible MTHFS activity and standard SHMT activity (MTHFS-catalyzed reaction parameter was set to 0, while SHMT activity retained its standard value); negligible MTHFS and SHMT: CHF –> 5fTHF activity (the MTHFS parameter and initial value of 5fTHF were set to 0 and the SHMT activity was turned off by setting its parameter to 0 for the corresponding reaction); standard MTHFS and SHMT activity without MTHFS inhibition by 10fTHF (standard model parameterization where the part of the equation related to MTHFS inhibition by 10fTHF was omitted). The initial values considered for the simulations were those illustrated in the study conducted by Misselbeck *et al*. (2019)[27]. The trends of the concentration values and enzyme rates obtained by the simulations of the newly extended model (see Supplementary Note 2 and Supplementary Data 1) were compared with those available in the original papers (please refer to Figure 2 and Tables 1-6 published in Misselbeck *et al*. (2017)[26] and to Tables 1-5 published in Misselbeck *et al*. (2019)[27]) to evaluate the reliability and accuracy of the model modified here and its consistency with the previous findings.

Perturbation of enzyme activity rate

In order to simulate possible imbalances of FOCM pathways in DS condition, we proceeded to systematically perturb enzyme activity rates in the newly extended model. In concordance with the dosage hypothesis in DS condition[39,40], we assumed that the presence of the additional third copy of Hsa21 (in a ratio of 3:2 compared to the normal pair of Hsa21), may lead to an altered expression of genes located on Hsa21 itself and/or on other chromosomes by the expected 3:2 ratio or by a 2:3 ratio (the latter due to the inhibitory effect). This over- or under- expression may consequently result in alteration of 3:2 or 2:3 of the FOCM enzyme activities, finally leading to pathway imbalances. Therefore, perturbation was algorithmically imposed on the model by individually reducing or augmenting each of the 17 model reaction activity rates by a factor of 2/3 or 3/2, respectively. For bidirectional reactions catalyzed by the same enzyme (*e.g.*, methylene-tetrahydrofolate dehydrogenase or MTD, methenyl-tetrahydrofolate cyclohydrolase MTCH, and S-adenosyl-homocysteine hydrolase or SAHH), both directions were perturbed simultaneously. The resulting 34 perturbed models were

ARTICLE IN PRESS

then simulated to compute their new steady states, representing possible hypotheses of FOCM configurations in DS. A MATLAB function was developed to calculate, for each enzyme reaction rate perturbation, the ratio of perturbed to unperturbed concentrations for all 14 model variables, thereby representing a DS vs. non-DS ratio. Specifically, this computed ratio is calculated between the model variable concentration value (z) obtained at the steady state after perturbation for the DS condition and the model variable concentration value (y) in non-DS (represented by the unperturbed model at the steady state), following the approach depicted in Figure 2.

Experimental dataset

Computed model-derived variable ratios were then compared with experimental data available in the literature. In particular, a public dataset provided by a previous work[18] reported DS/non-DS concentration ratio for the five metabolites THF, 5mTHF, SAM, SAH and 5fTHF measured in plasma samples through the enzyme-linked immunosorbent assay (ELISA) and was used for this purpose. The global cohort was represented by 164 subjects with DS, with a mean age of 11.55 years and 54 subjects from the control group, with a mean age of 14.53 years[18]. We chose DS/non-DS mean concentration ratio values, excluding outliers: 0.72 for THF, 1.11 for 5mTHF, 1.28 for SAM, 3.45 for SAH, and 1.04 for 5fTHF (please refer to Supplementary Table 6 published in Vione *et al.*[18]).

In order to identify which enzyme reaction rate perturbation reaches the best concordance between the computed variable ratios and experimental data available in the literature, a procedure based on relative error calculation and a bootstrap technique was used as outlined in Figure 3 and detailed in the next paragraphs.

Comparison with experimental data through fixed uncertainties

For the five metabolites THF, 5mTHF, SAM, SAH and 5fTHF, the ratios computed by model simulation (variable steady state value after perturbation vs. unperturbed variable steady state value, Figure 2) were then compared with those experimentally provided by a previous work[18], as explained above, using the relative error calculation formula:

$$\frac{|v_r - v_p|}{v_r} \tag{3}$$

ARTICLE IN PRESS

where v_r denotes the real (experimental) value extracted from the literature, and v_p is the model-predicted value generated by the MATLAB function. In each case, for the five metabolites of interest, the MATLAB function allocated a score of 1 (positive match) when

$$\frac{|v_r - v_p|}{v_r} \leq \delta \tag{4}$$

according to a fixed uncertainty threshold δ, 0 (negative match) otherwise.

The fixed uncertainty thresholds δ were arbitrarily chosen and ranged from 0.01 to 0.6 (corresponding to 60% of the uncertainty). The MATLAB function was then extended to calculate, for each enzyme activity rate perturbation, two scores used to present the results. A total score for each perturbed enzyme, which ranges from 0 to 5, indicates the number of the five considered metabolites that have a computed ratio in agreement with the experimental data. In addition, for each metabolite of interest, an overall score which ranges from 0 to 17 indicates the extent of concordance between the computed ratios and those from the experimental data for each of the 17 enzyme reaction activity perturbations.

ARTICLE IN PRESS

Comparison with experimental data through a bootstrap procedure

A limitation of the analysis introduced above is that it requires an estimate of the fixed uncertainty threshold δ that could be difficult to identify. For this reason, to establish a data-driven approach for evaluating the concordance between model-derived variable ratios and experimental data, we also implemented a bootstrap procedure based on a randomized number generation approach. Specifically, a MATLAB function was developed to generate *in silico* datasets for both the DS and control conditions for each of the five metabolites of interest, ensuring that each dataset matched the length of the original experimental data. Within each *in silico* dataset, metabolite values were randomly generated to replicate the distribution observed in the corresponding experimental dataset.

Subsequently, metabolite mean values were computed for each condition and the DS/control ratio was calculated. This process was iteratively repeated for a predefined number of iterations (100, 300, 500, 800, and 1000), yielding a vector of ratio values. Since increasing the number of iterations from 800 to 1000 did not yield an appreciable difference in the computed values, we interpreted 1000 iterations as sufficient to achieve a good

ARTICLE IN PRESS

level of stability in our results. Therefore, we considered only the values computed with this number of iterations in the Results section.

From the computed vector of ratio values, key statistical measures were derived, including the mean, median, standard deviation, first quartile, and third quartile that were then used to assess the agreement between the model simulations and the experimental dataset[18]. For this reason, we employed a MATLAB function that assigns a score of 1 if the model variable ratio falls within a designated interval and 0 otherwise for each metabolite of interest. We examined the following intervals, unique to each metabolite, which have been derived from the outcomes of the bootstrap procedure:

1. the interval [Mean - Standard Deviation, Mean + Standard Deviation] which is based on the assumption of a normal distribution within the data population and encompasses 68.2% of the dataset.

2. The interval [Mean - 1.96 • Standard Deviation, Mean + 1.96 • Standard Deviation], assuming a normal distribution, which includes 95% of the data.

3. The interval [First Quantile, Third Quantile], accommodating the general distribution of data, which contains 50% of the data.

4. The interval [First Quantile - 1.5 • interquartile range or IQR, Third Quantile + 1.5 • IQR], under the assumption of a general distribution, which covers 98% of the data.

In addition to the comparisons reported above, we conducted an additional analysis by defining an interval that encompassed the entire population and further expanded its minimum and maximum boundaries using the following formula [Minimum Value − γ • Minimum Value, Maximum Value + γ • Maximum Value], where $\gamma = 0.05 + k \bullet 0.05$ with $k = 0, 1, \ldots, 10, 11$. This test was designed to evaluate the level of uncertainty required to satisfy specific constraints that proved particularly challenging to fulfill in our simulations.

Gene expression meta-analysis

In order to explore if enzyme activity rate perturbations may be concordant with enzyme encoding gene expression values in DS (trisomy 21) compared with non-DS control (euploid) conditions, quantitative transcriptome maps of human trisomy 21 provided in the published study by Pelleri et al. (2018)[28] obtained through the use of the Transcriptome Mapper (TRAM) software[41] were used. Each FOCM enzyme name was

ARTICLE IN PRESS

searched in the KEGG (Kyoto Encyclopedia of Genes and Genomes) Enzyme database (https://www.genome.jp/kegg/annotation/enzyme.html)[42] and the corresponding enzyme commission number and encoding gene symbol for *Homo sapiens* were annotated. The list of gene symbols was used to export corresponding expression ratio values between DS (trisomy 21) and non-DS control (euploid) conditions in various available tissues focusing on brain transcriptome map from TRAM. Precisely, FOCM enzyme encoding genes with expression ratio ≤0.76 were considered in concordance with 2/3 enzyme activity rate perturbation, while genes with expression ratio ≥1.30 were considered in concordance with 3/2 enzyme activity perturbation. Genes for which under- or over-expression was considered in concordance with 2/3 or 3/2 enzyme activity rate perturbation statistically selected through fixed uncertainties and bootstrap procedure were searched in TissueNexus (https://diseaselinks.com/TissueNexus/)[43], a comprehensive database housing 49 functional gene networks constructed by integrating 1,341 RNA-seq datasets and 52,087 individual samples through artificial intelligence. This database was used to explore possible gene pairs, each associated with a Pearson correlation coefficient. Brain gene pair dataset with edge weight >0.5 was downloaded and searched for interesting genes, with particular attention to Hsa21.

ARTICLE IN PRESS

Data availability

This work is based on public datasets. All the computed data are provided in the Supplementary Data.

Code availability

The MATLAB code to simulate the mathematical model employed in this study is available at https://github.com/cosbi-research/FocmDownSyndrome.

Acknowledgements

The fellowships for BV and GR were funded by donations, as acknowledged below.

The fellowship for FA was founded by the Ministry of University and Research in the framework of PNC "DARE - Digital lifelong prevention project" (PNC0000002 – CUP B53C22006450001). The views and opinions expressed are solely those of the authors and do not necessarily reflect those of the European Union, nor can the European Union be held responsible for them.

This work has also been supported by a grant from Fondazione Cassa di Risparmio in Bologna (grant "Bando Alta Tecnologia 2021" to MCP).

ARTICLE IN PRESS

We are especially grateful to the "Fondazione Umano Progresso", Milano, for their valuable support to our research on trisomy 21, as well as to: the Mele Family; "Gruppo Arzdore di Dozza" and the Costa Family; "Associazione Vola con Martin oltre il 21 ODV", Mandello del Lario (LC); Gabriele's dad and mum; "Associazione Sette per te ventuno - Non abbiate paura" onlus; "Fondazione Rosa Prístina", Pisa; Paolo Ciaroni, main organizer of the car meetings "Top Down 2024", VII Edition, devoted to spider cars; "Associazione Amicorum", Cassano Magnago; the parents of Eleonora De Vita: Maurizio De Vita and Daniela Passerini, and the family of Giancarlo and Gabriella Bandera; "Unità Pastorale di Reda", "Parrocchie di S. Stefano in Corleto e S. Maria in Basiago", "Gruppo Dcuore"; "Monastero delle Clarisse di San Biagio, Forlì"; "Il popolo di Elisa Schira"; the Authors and the Editors of the book "Come fiori in un prato"; Jonatan Benvenuti and "Uniamoci per la Ricerca" initiative; Rotary Club Livorno Sud e Colline Pisano Livornesi; Rotary Club Cascina e Monte Pisano; Fondazione Mario Marianelli di Santa Croce sull'Arno (PI); "Associazione Capirsi Down ODV", Monza; Filippo Bono; Sichim Alfa Srl; Marco Bonafaccia, President of AMRP Onlus; A.MA.DOWN-Associazione Marchigiana Persone Down, Porto Potenza Picena (MC); "3x21 I sogni di Saveria Aps" and "Piramis Onlus"; "Amici di Simone ODV", Chiavari; Maria Aluigi; Maria Ilde Lomoro; "Associazione Cromosoma della Felicità APS", Chieti; "Gruppo Giovani Oratorio Narzole"; Ezio Meroni; Simone Bombardi and "Associazione Culturale Compagnia dell'Anello", Forlì; Maurizio Funazzi; Maria Grazia Verand-Bassani; Mimmo and Friends, dinner on 27/01/2024 at Toscanella di Dozza; Maurizio Poloni and Young Doctors in Rimini; Mario Gargantini, "Centro Culturale Marcello Candia", Melzo (MI); Giorgio Carpenito's Family.

We are deeply grateful for their generous bequest in favour of our research to: Mrs M.T.; to Mrs. Franca Troffei, in memory of Massimo Marchini.

We are very grateful to all the other people that have very kindly contributed through individual donations to support this research, in particular: Giuseppe Angelillo; Gaia Berlanda; Ornella Carciani; Carlo; Andrea Conti; Mariadele Conti; the family of Desirèe Donato; Simone Ferriani; the Grandparents of Carlotta and Maddalena; Anastasiya Koval; Domenico Lascala; Nicoletta Lastella; "Liceo scientifico A. Carrel" - "Fondazione V. Grossman"; Gilberto Lisi; the Mattia's Uncles; Nicolette; the Parents of Desiree; the Parents of Margherita; Domenico Ricci; Ernesto Salzano; Claudio Sandonà; Simona Serattini; Vincenzina; Alessandro Volta.

We are grateful to Kirsten Welter for her kind and expert revision of the manuscript.

The author would like to sincerely thank Prof. Pierluigi Strippoli for the valuable discussions.

ARTICLE IN PRESS

Author contributions

AP and DP performed the analyses; DP and LM implemented the computational extended model and the statistical significance procedures; AP, GR, CL, FA, BV, MC and MCP provided all biological insight; AP and LM supervised the project. All authors critically reviewed and approved the final manuscript.

Competing interests

The authors declare no competing financial or non-financial interests.

References

1. Lejeune, J., Gauthier, M. & Turpin, R. [Human chromosomes in tissue cultures]. *Comptes rendus hebdomadaires des seances de l'Academie des sciences* **248**, 602-603 (1959).
2. Pelleri, M. C. *et al.* Partial trisomy 21 map: Ten cases further supporting the highly restricted Down syndrome critical region (HR-DSCR) on human chromosome 21. *Mol. Genet. Genomic Med.* **7**, e797 (2019).

ARTICLE IN PRESS

3. Bull, M. J. Down Syndrome. *N. Engl. J. Med.* **382**, 2344-2352 (2020).
4. Anneren, G., Gustavson, K. H., Sara, V. R. & Tuvemo, T. Growth retardation in Down syndrome in relation to insulin-like growth factors and growth hormone. *Am. J. Med. Genet. Suppl.* **7**, 59-62 (1990).
5. Gardiner, K. *et al.* Down syndrome: from understanding the neurobiology to therapy. *J. Neurosci.* **30**, 14943-14945 (2010).
6. Roizen, N. J. & Patterson, D. Down's syndrome. *Lancet* **361**, 1281-1289 (2003).
7. Strippoli, P. *et al.* An integrated route to identifying new pathogenesis-based therapeutic approaches for trisomy 21 (Down Syndrome) following the thought of Jérôme Lejeune. *Science Postprint* **1**, e00010 (2013).
8. Pelleri, M. C. *et al.* Genotype-phenotype correlation for congenital heart disease in Down syndrome through analysis of partial trisomy 21 cases. *Genomics* **109**, 391-400 (2017).
9. Strippoli, P. *et al.* in *State of the Art of Research on Down Syndrome* Vol. 56 1-39 (Academic Press, 2019).
10. Locatelli, C. *et al.* A reassessment of Jackson's checklist and identification of two Down syndrome sub-phenotypes. *Sci. Rep.* **12**, 3104 (2022).
11. Lejeune, J. [Biochemical investigations and trisomy 21 (author's transl)]. *Ann. Genet.* **22**, 67-75 (1979).
12. Rosenblatt, D. S. *et al.* Differential effects of folinic acid and glycine, adenosine, and thymidine as rescue agents in methotrexate-treated human cells in relation to the accumulation of methotrexate polyglutamates. *Mol. Pharmacol.* **21**, 718-722 (1982).
13. Peeters, M. A., Rethore, M. O. & Lejeune, J. In vivo folic acid supplementation partially corrects in vitro methotrexate toxicity in patients with Down syndrome. *Br. J. Haematol.* **89**, 678-680 (1995).
14. Song, C. *et al.* Effect of the onecarbon unit cycle on overall DNA methylation in children with Down's syndrome. *Mol Med Rep* **12**, 8209-8214 (2015).
15. Funk, R. S., Talib, N. J., Zimmerman, K. O., van Haandel, L. & Becker, M. L. Altered Folate Homeostasis in Children with Down Syndrome: A Potential Basis for Enhanced Methotrexate Toxicity. *J. Pediatr.* **221**, 235-239 (2020).
16. Antonaros, F. *et al.* Plasma metabolome and cognitive skills in Down syndrome. *Sci. Rep.* **10**, 10491 (2020).
17. Antonaros, F. *et al.* One-carbon pathway and cognitive skills in children with Down syndrome. *Sci. Rep.* **11**, 4225 (2021).
18. Vione, B. *et al.* One-carbon pathway metabolites are altered in the plasma of subjects with Down syndrome: Relation to chromosomal dosage. *Front Med (Lausanne)* **9**, 1006891 (2022).
19. Lyon, P., Strippoli, V., Fang, B. & Cimmino, L. B Vitamins and One-Carbon Metabolism: Implications in Human Health and Disease. *Nutrients* **12** (2020).
20. Caracausi, M. *et al.* The functional roles of S-adenosyl-methionine and S-adenosyl-homocysteine and their involvement in trisomy 21. *Biofactors* **50**, 709-724 (2024).
21. Ducker, G. S. & Rabinowitz, J. D. One-Carbon Metabolism in Health and Disease. *Cell Metab.* **25**, 27-42 (2017).
22. Jackson, R. C. & Harrap, K. R. Studies with a mathematical model of folate metabolism. *Arch. Biochem. Biophys.* **158**, 827-841 (1973).
23. Nijhout, H. F., Reed, M. C., Budu, P. & Ulrich, C. M. A mathematical model of the folate cycle: new insights into folate homeostasis. *J. Biol. Chem.* **279**, 55008-55016 (2004).
24. Reed, M. C., Nijhout, H. F., Sparks, R. & Ulrich, C. M. A mathematical model of the methionine cycle. *J. Theor. Biol.* **226**, 33-43 (2004).
25. Reed, M. C. *et al.* A mathematical model gives insights into nutritional and genetic aspects of folate-mediated one-carbon metabolism. *J. Nutr.* **136**, 2653-2661 (2006).
26. Misselbeck, K. *et al.* A hybrid stochastic model of folate-mediated one-carbon metabolism: Effect of the common C677T MTHFR variant on de novo thymidylate biosynthesis. *Sci. Rep.* **7**, 797 (2017).
27. Misselbeck, K., Marchetti, L., Priami, C., Stover, P. J. & Field, M. S. The 5-formyltetrahydrofolate futile cycle reduces pathway stochasticity in an extended hybrid-stochastic model of folate-mediated one-carbon metabolism. *Sci. Rep.* **9**, 4322 (2019).
28. Pelleri, M. C. *et al.* Integrated Quantitative Transcriptome Maps of Human Trisomy 21 Tissues and Cells. *Front. Genet.* **9**, 125 (2018).
29. Coppedè, F. The complex relationship between folate/homocysteine metabolism and risk of Down syndrome. *Mutation Research/Reviews in Mutation Research* **682**, 54-70 (2009).

ARTICLE IN PRESS

ARTICLE IN PRESS

30. Liu, Y. *et al.* Systematic proteome and proteostasis profiling in human Trisomy 21 fibroblast cells. *Nat. Commun.* **8**, 1212 (2017).
31. Sullivan, K. D. *et al.* Trisomy 21 causes changes in the circulating proteome indicative of chronic autoinflammation. *Sci. Rep.* **7**, 14818 (2017).
32. Dunn, L. N. *et al.* Altered hepatic metabolism in Down syndrome. *Cell reports* **45**, 116835 (2026).
33. Powers, R. K. *et al.* Trisomy 21 activates the kynurenine pathway via increased dosage of interferon receptors. *Nat. Commun.* **10**, 4766 (2019).
34. Field, M. S., Kamynina, E., Chon, J. & Stover, P. J. Nuclear Folate Metabolism. *Annu. Rev. Nutr.* **38**, 219-243 (2018).
35. Fountoulakis, M., Gulesserian, T. & Lubec, G. in *Advances in Down Syndrome Research* (ed G. Lubec) 85-93 (Springer Vienna, 2003).
36. Storch, K. J., Wagner, D. A., Burke, J. F. & Young, V. R. Quantitative study in vivo of methionine cycle in humans using [methyl-2H3]- and [1-13C]methionine. *Am. J. Physiol.* **255**, E322-331 (1988).
37. Herbig, K. *et al.* Cytoplasmic serine hydroxymethyltransferase mediates competition between folate-dependent deoxyribonucleotide and S-adenosylmethionine biosyntheses. *J. Biol. Chem.* **277**, 38381-38389 (2002).
38. Frosst, P. *et al.* A candidate genetic risk factor for vascular disease: a common mutation in methylenetetrahydrofolate reductase. *Nat. Genet.* **10**, 111-113 (1995).
39. Hunter, S., Hendrix, J., Freeman, J., Dowell, R. D. & Allen, M. A. Transcription dosage compensation does not occur in Down syndrome. *BMC Biol.* **21**, 228 (2023).
40. Liu, Y. *et al.* Systematic proteome and proteostasis profiling in human Trisomy 21 fibroblast cells. *Nat. Commun.* **8**, 1212 (2017).
41. Lenzi, L. *et al.* TRAM (Transcriptome Mapper): database-driven creation and analysis of transcriptome maps from multiple sources. *BMC Genomics* **12**, 121 (2011).
42. Kanehisa, M. Enzyme Annotation and Metabolic Reconstruction Using KEGG. *Methods Mol. Biol.* **1611**, 135-145 (2017).
43. Lin, C.-X., Li, H.-D., Deng, C., Guan, Y. & Wang, J. TissueNexus: a database of human tissue functional gene networks built with a large compendium of curated RNA-seq data. *Nucleic Acids Res.* **50**, D710-D718 (2021).

ARTICLE IN PRESS

Additional information

Supplementary information Supplementary notes, figures and data are available with the online version of this article.

Figures

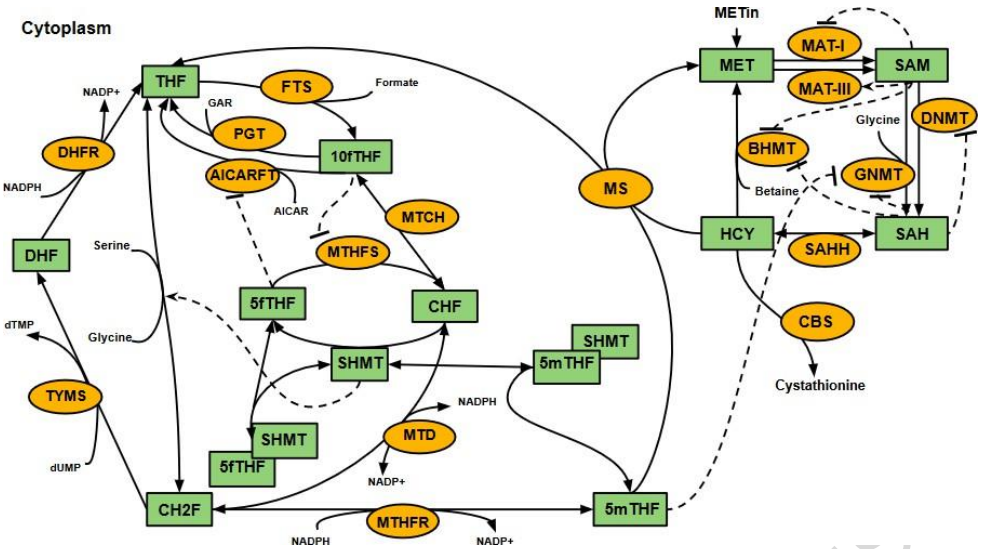


Figure 1. Model diagram of the folate-mediated one-carbon metabolism (FOCM), derived from the model described in Misselbeck *et al.*[27] extended with the addition of the methionine input rate (METin, displayed in the top right area of the diagram) and the reaction catalyzed by cystathionine ß-synthase (CBS, displayed in the bottom right area of the diagram). Green rectangles identify model variables, orange circles represent enzymes (except for serine hydroxymethyltransferase or SHMT, which is considered a model variable in the original model), non-boxed substrates indicate model constants, solid black arcs identify matter transformations, dashed black arcs represent regulatory events ("T" shaped arrows identify inhibitions and standard arrows identify activations). The list of abbreviations is available in Supplementary Note 1.

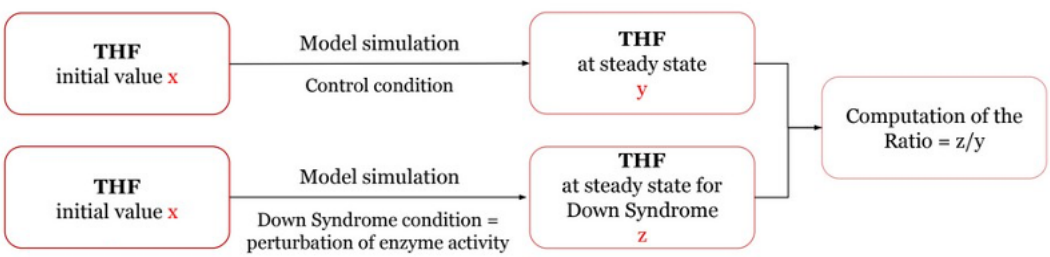


Figure 2. Diagram describing the workflow for the computation of the model variable ratios in the example case of tetrahydrofolate (THF). The final output is the ratio of the THF value at steady state simulated by our model

ARTICLE IN PRESS

under a perturbation relative to the Down syndrome (DS) condition (z) to the THF value at steady state relative to the non-DS control condition (y) from the unperturbed model. This procedure was repeated for each enzyme reaction rate perturbation and for each of the 14 model variables.

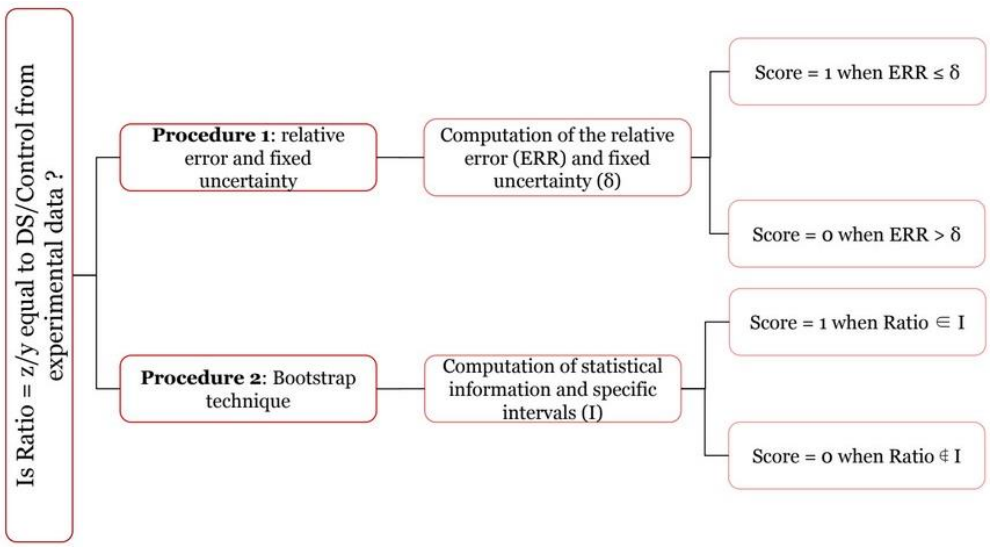


Figure 3. Diagram describing two different approaches to assess the agreement between computed model-derived variable ratios ("Ratio z/y") and experimental data for the example case shown in Figure 2, and applied for the five metabolites of interest (tetrahydrofolate (THF), 5-methyl-tetrahydrofolate (5mTHF), S-adenosyl-methionine (SAM), S-adenosyl-homocysteine (SAH) and 5-formyl-tetrahydrofolate (5fTHF)). The first procedure is based on the computation of a relative error and a fixed uncertainty, while the second is based on a bootstrap technique. Details are provided in the Methods section.

A

Enzyme	Fixed uncertainties δ						Bootstrap	Gene	Chr	Expression ratio	
	0.051	0.112	0.234	0.376	0.417	0.600				Brain	All tissues
DHFR	1	2	3	3	4	4	3	*DHFR*	5	1.64	1.29
TYMS	1	2	3	4	4	4	3	*TYMS*	18	1.43	1.17
PGT	1	2	4	4	4	4	4	*GART*	21	1.43	1.31
AICARFT	1	2	4	4	4	4	4	*ATIC*	2	1.17	0.96
MTHFR	0	0	0	0	0	1	0	*MTHFR*	1	1.15	1.11
MS	0	1	2	2	2	3	2	*MTR*	1	1.40	0.85
MTHFS	0	0	2	3	4	4	2	*MTHFS*	15	0.86	1.01
FTS	1	2	3	3	3	3	3	*MTHFD1*	14	1.28	1.31
MTD	1	1	2	4	4	4	3	*MTHFD2*	2	1.58	1.02
MTCH	1	1	2	4	4	4	3	*MTHFD2*	2	1.58	1.02
BHMT	1	1	2	2	3	4	1	*BHMT*	5	1.13	0.96
SAHH	0	1	2	3	4	4	2	*AHCY*	20	1.42	0.97
MAT-III	1	2	3	3	4	4	3	*MAT1A*	10	0.70	1.13
MAT-I	1	2	3	3	4	4	3	*MAT1A*	10	0.70	1.13
GNMT	1	3	3	3	4	4	3	*GNMT*	6	0.73	1.12
DNMT	0	2	2	3	3	4	2	*DNMT1*	19	1.36	0.90
CBS	0	1	2	2	4	5	1	CBS	21	1.52	1.31

B

Enzyme	Fixed uncertainties δ						Bootstrap	Gene	Chr	Expression ratio	
	0.051	0.112	0.234	0.376	0.437	0.600				Brain	All tissues
DHFR	1	2	3	3	4	4	3	*DHFR*	5	1.64	1.29
TYMS	1	2	3	3	4	4	3	*TYMS*	18	1.43	1.17
PGT	1	2	3	3	3	3	3	*GART*	21	1.43	1.31
AICARFT	1	2	3	3	3	3	3	*ATIC*	2	1.17	0.96
MTHFR	1	1	2	2	2	3	2	*MTHFR*	1	1.15	1.11
MS	0	0	0	0	0	1	0	*MTR*	1	1.40	0.85
MTHFS	0	1	2	3	4	4	2	*MTHFS*	15	0.86	1.01
FTS	1	2	4	4	4	4	4	*MTHFD1*	14	1.28	1.31
MTD	0	2	2	4	4	4	3	*MTHFD2*	2	1.58	1.02
MTCH	1	2	2	4	4	4	3	*MTHFD2*	2	1.58	1.02
BHMT	0	2	2	3	4	4	2	*BHMT*	5	1.13	0.96
SAHH	1	1	2	3	3	4	2	*AHCY*	20	1.42	0.97
MAT-III	1	2	3	3	4	4	3	*MAT1A*	10	0.70	1.13
MAT-I	1	2	3	3	4	4	3	*MAT1A*	10	0.70	1.13
GNMT	1	2	2	2	4	4	3	*GNMT*	6	0.73	1.12
DNMT	1	1	2	2	3	4	1	*DNMT1*	19	1.36	0.90
CBS	1	1	2	3	3	4	1	*CBS*	21	1.52	1.31

ARTICLE IN PRESS

Figure 4. Summary diagram of the main results obtained following folate-mediated one-carbon metabolism (FOCM) enzyme activity perturbations by 2/3 (A) or by 3/2 (B) and comparison of the computed ratio between the steady-state value after perturbation (representing the Down syndrome, DS, condition) and the initial steady-state value of the unperturbed model (non-DS) with literature data for the five metabolites of interest (tetrahydrofolate (THF), 5-methyl-tetrahydrofolate (5mTHF), S-adenosyl-methionine (SAM), S-adenosyl-homocysteine (SAH) and 5-formyl-tetrahydrofolate (5fTHF)). The first column reports the perturbed enzymes. For enzymes catalyzing bidirectional reactions (*e.g.*, MTD, MTCH, and SAHH), both reaction directions were perturbed simultaneously. For each enzyme are shown: the concordance score obtained using a relative error-based procedure for a selection of fixed uncertainties showing notable changes; the score determined with the bootstrap procedure for the interval [First Quantile - 1.5 • IQR, Third Quantile + 1.5 • IQR]; the relative encoding gene and its chromosome (Chr) location; the DS versus non-DS gene expression ratio derived from transcriptome maps available in Pelleri *et al.* 2018[28]. Enzymes with the highest concordance scores whose perturbation was consistent with relative encoding gene under- or over-expression are highlighted with a rectangle. Complete results are provided in the Supplementary Data, and additional methodological details are described in the Methods section.

ARTICLE IN PRESS

Tables

Table 1. Computed ratio values for perturbation of enzyme activity rates by 2/3 or by 3/2.

	THF	5mTHF	SAM	SAH	5fTHF		THF	5mTHF	SAM	SAH	5fTHF
Exp. Ref.	0.72	1.11	1.28	3.45	1.04	**Exp. Ref.**	0.72	1.11	1.28	3.45	1.04
2/3						**3/2**					
DHFR	1.0001	0.9995	0.9997	1.0001	1.0003	**DHFR**	1	1.0003	1.0002	0.9999	0.9998
TYMS	0.9904	1.0020	1.0011	0.9996	0.9964	**TYMS**	1.0146	0.9970	0.9984	1.0006	1.0053
PGT	0.8307	1.0002	1.0001	1	1.0014	**PGT**	1.2944	0.9996	0.9998	1	0.9979
AICARF	0.8035	1.0002	1.0001	1	1.0016	**AICARF**	1.3509	0.9994	0.9997	1	0.9975
MTHFR	1.1646	0.1920	0.3988	1.4442	4.0865	**MTHFR**	0.8118	1.7181	1.3328	0.8998	0.1451
MS	0.8122	1.7151	1.1531	0.8623	0.1468	**MS**	1.1651	0.1939	0.4363	1.5158	4.0759
MTHFS	0.9687	0.9815	0.9899	1.0040	1.4588	**MTHFS**	1.0236	1.0130	1.0070	0.9972	0.6799
FTS	1.9216	0.9981	0.9990	1.0004	0.9944	**FTS**	0.5804	1.0003	1.0002	0.9999	1.0038

ARTICLE IN PRESS

MTD	1.002	0.9781	0.9881	1.0048	1.0481	**MTD**	0.9987	1.0143	1.0077	0.997	0.9695
MTCH	1.0004	0.9977	0.9987	1.0005	1.0036	**MTCH**	0.9998	1.0015	1.0008	0.9997	0.9976
BHMT	1.0063	0.9689	0.6848	0.8822	1.0666	**BHMT**	0.9943	1.0280	1.7442	1.2048	0.9428
SAHH	0.9916	1.041	2.6112	3.0359	0.9171	**SAHH**	1.0078	0.9616	0.7351	0.5393	1.0827
MAT-III	1	1	1	1	1	**MAT-III**	1	1	1	1	1
MAT-I	1	1	1	1	1	**MAT-I**	1	1	1	1	1
GNMT	0.9977	1.0114	1.2474	0.9139	0.9764	**GNMT**	1.0026	0.9874	0.7932	1.1117	1.0266
DNMT	0.9942	1.0282	1.8832	0.7973	0.9423	**DNMT**	1.0053	0.9739	0.6099	1.2610	1.0556
CBS	1.0113	0.9442	1.7941	1.4363	1.1218	**CBS**	0.9883	1.0572	0.6806	0.7717	0.8859

List of ratio values computed between the value reached at the steady state after each enzyme activity perturbation by 2/3 (on the left) or by 3/2 (on the right) representing Down syndrome condition and the initial steady state value (related to unperturbed model as a control) for tetrahydrofolate (THF), 5-methyl-tetrahydrofolate (5mTHF), S-adenosyl-methionine (SAM), S-adenosyl-homocysteine (SAH) and 5-formyl-tetrahydrofolate (5fTHF). The first column reports the perturbed enzymes. For enzymes catalyzing bidirectional reactions (*e.g.*, MTD, MTCH, and SAHH), both reaction directions were perturbed simultaneously. All the 14 variable ratio values are available in Supplementary Data 2. The complete list of abbreviations is available in Supplementary Note 1. In the first row (Exp.Ref.) the mean Down syndrome/non-DS metabolite concentration ratio values measured in plasma samples by Vione *et al*.[18] are reported.

ARTICLE IN PRESS

Table 2. Total scores considering fixed uncertainties δ for each enzyme when its activity was decreased using the factor 2/3 (left) or increased 3/2 (right).

Enzyme	**Fixed uncertainties perturbation 2/3**						**Fixed uncertainties perturbation 3/2**					
δ	0.051	0.112	0.234	0.376	0.417	0.600	0.051	0.112	0.234	0.376	0.437	0.600
DHFR	1	2	3	3	4	4	1	2	3	3	4	4
TYMS	1	2	3	4	4	4	1	2	3	3	4	4
PGT	1	2	4	4	4	4	1	2	3	3	3	3
AICARFT	1	2	4	4	4	4	1	2	3	3	3	3
MTHFR	0	0	0	0	0	1	1	1	2	2	2	3
MS	0	1	2	2	2	3	0	0	0	0	0	1
MTHFS	0	0	2	3	4	4	0	1	2	3	4	4
FTS	1	2	3	3	3	3	1	2	4	4	4	4
MTD	1	1	2	4	4	4	0	2	2	4	4	4
MTCH	1	1	2	4	4	4	1	2	2	4	4	4

ARTICLE IN PRESS

BHMT	1	1	2	2	3	4	0	2	2	3	4	4
SAHH	0	1	2	3	4	4	1	1	2	3	3	4
MAT-III	1	2	3	3	4	4	1	2	3	3	4	4
MAT-I	1	2	3	3	4	4	1	2	3	3	4	4
GNMT	1	3	3	3	4	4	1	2	2	2	4	4
DNMT	0	2	2	3	3	4	1	1	2	2	3	4
CBS	0	1	2	2	4	5	1	1	2	3	3	4

For each perturbation of enzyme activity (reduced to 2/3, left, or increased to 3/2, right), a score was assigned by comparing the computed ratio—defined as the steady-state value after perturbation (representing the Down syndrome, DS, condition) relative to the initial steady-state value of the unperturbed model (non-DS)—with the corresponding experimental ratio. This comparison was performed for five key metabolites: tetrahydrofolate (THF), 5-methyl-tetrahydrofolate (5mTHF), S-adenosyl-methionine (SAM), S-adenosyl-homocysteine (SAH), and 5-formyl-tetrahydrofolate (5fTHF). The concordance score was determined using a relative error-based procedure, ranging from 0 (indicating a complete lack of agreement) to 5 (indicating concordance across all five metabolites). The first column reports the perturbed enzymes. For enzymes catalyzing bidirectional reactions (*e.g.,* MTD, MTCH, and SAHH), both reaction directions were perturbed simultaneously. A selection of fixed uncertainties for which remarkable changes were evident was made, while complete results for all considered fixed uncertainties (δ) are available in the Supplementary Data 3. The list of abbreviations is available in Supplementary Note 1.

ARTICLE IN PRESS

Table 3. Intervals for the five metabolites dedicated to the analysis of enzyme activity perturbation with the bootstrap procedure.

Metabolite	**First Quantile**	**Third quantile**
THF	0.684	0.792
5mTHF	1.085	1.157
SAM	1.323	1.720
SAH	3.291	3.899
5fTHF	1.023	1.061
Metabolite	**First Quantile - 1.5 · IQR**	**First Quantile + 1.5 · IQR**
THF	0.521	0.955
5mTHF	0.978	1.265
SAM	0.727	2.316
SAH	2.380	4.810

ARTICLE IN PRESS

5fTHF	0.966	1.118

The two intervals assume a general distribution derived from the bootstrap procedure with 1000 iterations, with which to compare the model computed variable ratios for tetrahydrofolate (THF), 5-methyl-tetrahydrofolate (5mTHF), S-adenosyl-methionine (SAM), S-adenosyl-homocysteine (SAH) and 5-formyl-tetrahydrofolate (5fTHF). IQR: interquartile range.

Table 4. Results of enzyme activity perturbations (2/3, top; 3/2, bottom) evaluated within the interval [First Quantile - 1.5 · IQR, Third Quantile + 1.5 · IQR].

Perturb. 2/3	**THF**	**5mTHF**	**SAM**	**SAH**	**5fTHF**	**Score**
DHFR	0	1	1	0	1	3
TYMS	0	1	1	0	1	3
PGT	1	1	1	0	1	4
AICARFT	1	1	1	0	1	4
MTHFR	0	0	0	0	0	0
MS	1	0	1	0	0	2
MTHFS	0	1	1	0	0	2
FTS	0	1	1	0	1	3
MTD	0	1	1	0	1	3
MTCH	0	1	1	0	1	3
BHMT	0	0	0	0	1	1
SAHH	0	1	0	1	0	2
MAT-III	0	1	1	0	1	3
MAT-I	0	1	1	0	1	3
GNMT	0	1	1	0	1	3
DNMT	0	1	1	0	0	2
CBS	0	0	1	0	0	1
Total score	**3**	**13**	**14**	**1**	**11**	

Perturb. 3/2	**THF**	**5mTHF**	**SAM**	**SAH**	**5fTHF**	**Score**
DHFR	0	1	1	0	1	3
TYMS	0	1	1	0	1	3
PGT	0	1	1	0	1	3
AICARFT	0	1	1	0	1	3
MTHFR	1	0	1	0	0	2
MS	0	0	0	0	0	0
MTHFS	0	1	1	0	0	2
FTS	1	1	1	0	1	4
MTD	0	1	1	0	1	3

ARTICLE IN PRESS

MTCH	0	1	1	0	1	3
BHMT	0	1	1	0	0	2
SAHH	0	0	1	0	1	2
MAT-III	0	1	1	0	1	3
MAT-I	0	1	1	0	1	3
GNMT	0	1	1	0	1	3
DNMT	0	0	0	0	1	1
CBS	0	1	0	0	0	1
Total score	**2**	**13**	**14**	**0**	**12**	

For each enzyme activity perturbation (2/3, top; 3/2, bottom), the five metabolites—tetrahydrofolate (THF), 5-methyl-tetrahydrofolate (5mTHF), S-adenosyl-methionine (SAM), S-adenosyl-homocysteine (SAH), and 5-formyl-tetrahydrofolate (5fTHF)—were evaluated. For each metabolite, the computed ratio (steady-state value after perturbation, representing the DS condition, relative to the initial steady-state value of the unperturbed model used as control) was compared with the interval [First Quantile - 1.5 • IQR, Third Quantile + 1.5 • IQR] obtained from a bootstrap procedure with 1000 iterations (Table 3). A score of 1 was assigned if the ratio fell within this interval; otherwise, a score of 0 was assigned. The first column reports the perturbed enzymes. For enzymes catalyzing bidirectional reactions (*e.g.,* MTD, MTCH, and SAHH), both reaction directions were perturbed simultaneously. The list of abbreviations is available in Supplementary Note 1.

Table 5. Folate-mediated one-carbon metabolism enzymes and corresponding encoding genes.

Enzyme	**Enzyme commission number**	**Gene**	**Chromosome**
AICARFT	2.1.2.3	*ATIC*	2
BHMT	2.1.1.5	*BHMT*	5
CBS	4.2.1.22	*CBS*	21
DHFR	1.5.1.3	*DHFR*	5
DNMT	2.1.1.37	*DNMT1*	19
FTS	6.3.4.3	*MTHFD1*	14
GNMT	2.1.1.20	*GNMT*	6
MAT-I	2.5.1.6	*MAT1A*	10

ARTICLE IN PRESS

MAT-III	2.5.1.6	*MAT1A*	10
MS	2.1.1.13	*MTR*	1
MTCH	3.5.4.9	*MTHFD2*	2
MTD	1.5.1.15	*MTHFD2*	2
MTHFR	1.5.1.53	*MTHFR*	1
MTHFS	6.3.3.2	*MTHFS*	15
PGT	2.1.2.2	*GART*	21
SAHH	3.13.2.1	*AHCY*	20
SHMT	2.1.2.1	*SHMT1*	17
TYMS	2.1.1.45	*TYMS*	18

For each enzyme considered in the folate-mediated one-carbon metabolism **(**FOCM) model, relative commission number, encoding gene and its chromosomal location as retrieved in the KEGG (Kyoto Encyclopedia of Genes and Genomes) Enzyme database are reported.

Table 6. TissueNexus correlations

Gene	**Total correlations**	**FOCM correlations**
AHCY	1198	10
CBS	1162	8
DHFR	1780	14
GART	1385	10
GNMT	831	8
MAT1A	477	8
MTHFD1	1239	13
TYMS	1928	10

For each enzyme achieving the highest scores in the perturbation analyses, the corresponding encoding gene was queried in the TissueNexus database. Correlations with an edge weight > 0.5 in brain tissue—both with other genes in general and specifically with other folate-mediated one-carbon metabolism (FOCM) genes—were then identified and reported. The correlations of these genes with other FOCM genes are reported in Supplementary Data 8.