Computational Systems Biology Applied to Human Metabolism

Mathematical Modelling and Network Analysis

Ph.D. thesis of

Karla Misselbeck

Supervisor:
Prof. Corrado Priami

Co-Supervisors:
Dr. Silvia Parolo
Dr. Luca Marchetti

University of Trento
Department of Mathematics

March 14, 2019

Abstract

Human metabolism, an essential and highly organized process, which is required to run and maintain cellular processes and to respond to shifts in external and internal conditions, can be described as a complex and interconnected network of metabolic pathways. Computational systems biology provides a suitable framework to study the mechanisms and interactions of this network and to address questions that are difficult to reproduce *in vitro* or *in vivo*. This dissertation contributes to the development of computational strategies which help to investigate aspects of human metabolism and metabolic-related disorders.

In the first part, we introduce mathematical models of folate-mediated one-carbon metabolism in the cytoplasm and subsequently in the nucleus. A hybrid-stochastic framework is applied to investigate the behavior and stability of the complete metabolic network in response to genetic and nutritional factors. We analyse the effect of a common polymorphism of MTHFR, B_{12} and folate deficiency, as well as the role of the 5-formyltetrahydrofolate futile cycle on network dynamics. Furthermore, we study the impact of multienzyme complex formation and substrate channelling, which are key aspects related to nuclear folate-mediated one-carbon metabolism. Model simulations of the nuclear model highlight the importance of these two factors for normal functioning of the network and further identify folate status and enzyme levels as important influence factors for network dynamics.

In the second part, we focus on metabolic syndrome, a highly prevalent cluster of metabolic disorders. We develop a computational workflow based on network analysis to characterise underlying molecular mechanisms of the disorder and to explore possible novel therapeutic strategies by means of drug repurposing. To this end, genetic data, text mining results, drug expression profiles and drug target information are integrated in the setting of tissue-specific background networks and a proximity score based on topological distance and functional similarity measurements is defined to identify potential new therapeutic applications of already approved drugs. A filtering and prioritization analysis allow us to identify ibrutinib, an inhibitor of bruton tyrosine kinase, as the most promising repurposing candidate.

Publications included in this thesis

This thesis contains material which has already been published or has been submitted for publication:

- Misselbeck, K., Marchetti, L., Field, M.S., Scotti, M., Priami, C., Stover, P.S. (2017). A hybrid stochastic model of folate-mediated one-carbon metabolism: Effect of the common C677T MTHFR variant on de novo thymidylate biosynthesis. *Scientific Reports*, 7, 1-11
- Misselbeck, K., Marchetti, L., Priami, C., Stover, P.S., Field, M.S. (2019). The 5-formyltetrahydrofolate futile cycle reduces pathway stochasticity in an extended hybrid-stochastic model of folate-mediated one-carbon metabolism. *Scientific Reports*, 9, 1-10
- Palmer, A.M., Misselbeck, K., Marchetti, L., Finkelstein, J.M., Stabler, S.P., Brosnan, M.E., Brosnan, J.T., Priami, C., Field, M.S., Stover, P.S. (2018). Maternal vitamin B12 deficiency causes exencephaly in a mouse model of folate-responsive neural tube defects. *Manuscript in preparation*
- Misselbeck, K., Parolo, S., Leonardelli, L., Bora, P., Morine, M.J., Domenici, E., Priami, C. (2018). Metabolic syndrome: identification of deregulated pathways and drug effects by network analysis. *Submitted manuscript*

Contents

List of Figures

List of Tables

List of Abbreviations

5fTHF	5-formyl tetrahydrofolate
5mTHF	5-methyl tetrahydrofolate
10fTHF	10-formyltetrahydrofolate
AICAR	5-Aminoimidazole-4-carboxamide ribonucleotide
AICARFT	Phosphoribosylaminoimidazolecarboxamide formyltransferase
BHMT	Betaine-homocysteine methyltransferase
BTK	Bruton tyrosine kinase
CHF	5,10-methenyltetrahydrofolate
CH2F	5,10-methylenetetrahydrofolate
DHF	Dihydrofolate
DHFR	Dihydrofolate reductase
DNMT	DNA methyltransferase
dUMP	Deoxyuridine monophosphate
dTMP	Deoxythymidine monophosphate
FOCM	Folate-mediated one-carbon metabolism
FTD	10-formyltetrahydrofolate dehydrogenase
FTS	Formate-tetrahydrofolate ligase
GAR	Glycinamide ribonucleotide
GNMT	Glycine N-methyltransferase
HCY	Homocysteine
MAT-I	Methionine adenosyltransferase 1
MAT-III	Methionine adenosyltransferase 3
MET	Methionine
MetSyn	Metabolic syndrome
MTCH	Methenyltetrahydrofolate cyclohydrolase

MTD	Methylenetetrahydrofolate dehydrogenase
MTHFD	Methylenetetrahydrofolate dehydrogenase
MTHFR	Methylenetetrahydrofolate reductase
MTHFS	Methenyltetrahydrofolate synthetase
MTR	Methionine synthase
NADP+	Nicotinamide adenine dinucleotide phosphate
NADPH	Reduced form of Nicotinamide adenine dinucleotide phosphate
PGT	Phosphoribosylglycinamide formyltransferase
THF	Tetrahydrofolate
TYMS	Thymidylate synthase
SAH	S-adenosyl-homocysteine
SAHH	S-adenosylhomocysteine hydrolase
SAM	S-adenosyl-methionine
SHMT	Serine Hydroxymethyltransferase
SUMO	Small ubiquitin-like modifier

Preface

Systems biology has emerged as a powerful tool for studying complex biological systems at multiple scales from cell to organism level using a holistic perspective [131;192]. In contrast to approaches based on the concept of reductionism, which center around the study of individual, isolated components of a biological process, systems biology aims at understanding the organism as a whole and considers the single components and their interactions collectively [61;69]. Indeed, while understanding the functionality of single biological entities is still of great importance, obtaining a broad view of the interactions between them is of particular importance because it allows to elucidate underlying structures and dynamics [65;130].

A common way to represent biological processes in systems biology is given by networks, i.e. graphs in which nodes correspond to different biological entities such as genes, proteins and metabolites, and edges between them represent their interactions. Once the network structure has been defined, different analysis methods from other research areas may be applied to characterize and study the biological functions of the system at hand [284]. The static analysis of large-scale biological networks is based on the collation, integration, visualization and topological exploration of data sets spanning diverse layers of the physiology of interest or different stages of disease pathogenesis [69;284]. On the other hand, dynamic analysis of the system is provided by applying quantitative mathematical modeling and simulations, which allows to investigate adaptive changes in the molecular network in response to internal or external stimuli [183;192;203;284]. In this thesis we explore these two strategies and utilize them to study different aspects of human metabolism.

Human metabolism is defined by the totality of all life-sustaining physical or biochemical reactions occurring in the organism [191]. Two categories of reactions enable the organism to function: catabolic reactions break down complex compounds (e.g. carbohydrates, lipids, proteins, and nucleic acids) to obtain energy or macromolecule building blocks, while anabolic reactions synthesize new molecules from simpler and smaller entities [6]. This highly organized process is essential for organisms to run and maintain cellular processes, replicate and repair DNA, eliminate nitrogenous wastes, as well as to respond to shifting environmental conditions [6]. Regulation of these metabolic processes according to the cell's immediate demands and overall function is crucial to maintain a functioning system [182;191]. Based on the metabolites involved, biochemical reactions can be grouped in metabolic pathways, e.g. glycolysis is the set of reactions that transform glucose into biochemical energy. In consequence, we can understand metabolism as a complex and interconnected network of pathways, which must be monitored and coordinated at any given time.

Internal or external impairments or disruptions of the metabolic process introduced, e.g., by abnormal chemical reactions, nutritional deficiencies, genetic anomalies or organ-level diseases may cause dysfunction of these pathways and may lead to metabolic disorders [72]. The underlying

mechanisms and interactions of metabolic networks need still to be fully elucidated. Therefore, systems biology may play an important role in analyzing and understanding these complex processes.

This dissertation contributes to the development of computational frameworks, that help to investigate aspects of human metabolism and metabolic-related disorders. **Part I** (comprising Chapters 1 to 4) of this thesis presents a systems biology approach based on mathematical modelling to study the folate-mediated one-carbon metabolism, an important component of human metabolism connected to DNA replication. In **Part II** (consisting of Chapter 5) a network-based approach to study the underlying mechanisms of metabolic-related disorders and drug effects in a drug repurposing manner is introduced. This method is applied to metabolic syndrome, a highly prevelant cluster of metabolic disturbances.

Chapter 1 introduces a mathematical model of folate-mediated one-carbon metabolism (FOCM) in the cytoplasm. A hybrid-stochastic framework is applied to investigate the behavior and stability of this metabolic network in response to folate deficiency and the C667T polymorphism, a commen genetic variant of methylenetetrahydrofolate reductase (*MTHFR*), which has been associated with neural tube defects. The results indicate that *de novo* dTMP synthesis in the cytosol does not occur at rates sufficient for genome replication, and highlight therefore the need of nuclear FOCM to prevent impairment of dTMP synthesis, an aspect studied in Chapter 4. The content of this chapter has been published [168].

Chapter 2 is dedicated to the question of how vitamin B_{12} deficiency influences FOCM model dynamics with a special focus on *de novo* dTMP synthesis. For this purpose, deterministic simulations of the mathematical model introduced in Chapter 1 were used. The resulting simulations indicate that *de novo* dTMP synthesis is considerably more affected by vitamin B_{12} deficiency than either purine synthesis or homocysteine remethylation. The content of this chapter will be part of [194].

Chapter 3 provides an extension of the mathematical model presented in the preceding chapters. The main focus of this chapter is to study the effect of the 5fTHF futile cycle on overall network dynamics and stability. Model simulations indicate that the inclusion of the 5fTHF futile cycle stabilizes FOCM network in all considered scenarios regarding the *MTHFR* C667T polymorphism and folate deficiency. Furthermore, the model outcome shows that MTHFS plays a crucial role in preventing 5fTHF accumulation in mammalian cells, supporting experimental observations that *Mthfs* is an essential gene in mice. The content of this chapter has been published [169].

Chapter 4 is dedicated to the modeling efforts with respect to nuclear FOCM. Here the cytoplasmic model introduced in Chapter 3 is expanded to include the folate-dependent biochemical reactions occurring in the nuclear compartment. The primary focus of this study is to assess the effect of compartmentalization, enzymatic complex formation and substrate channelling on overall network dynamics and dTMP synthesis. The simulation results show that accounting for the kinetic effects of enzyme complex formation and substrate channelling is important to meet adequate dTMP synthesis rates.

Chapter 5 introduces a systems biology approach based on network analysis that integrates genomic data, text mining results and drug expression profiles to identify potential new therapeutic applications of approved drugs. A score that evaluates the similarity between drug and disease is defined by combining biological knowledge about gene functions and topological information derived from the network. This workflow has been applied to metabolic syndrome and allowed the identification of ibrutinib, an inhibitor of Bruton tyrosine kinase, as a candidate drug for the pharmacological treatment of obesity-related inflammation characterizing the disease. The content of this chapter has been submitted [170].

Introduction

In this dissertation, two computational systems biology strategies are introduced and utilized to study different aspects of human metabolism. The first method is based on mathematical modelling, which is applied to explore the dynamics of the folate-mediated one-carbon metabolism. Furthermore, a network-based approach to investigate the biological mechanisms of metabolic syndrome and to identify drug effects in a drug repurposing manner is presented. In the following, a short introduction to these topics is provided, starting with an overview of the folate-mediated one-carbon metabolism and a summary of previous mathematical models of this biological network. In the second part, an introduction to metabolic syndrome and the concept of drug repurposing is provided.

Folate-mediated one-carbon metabolism

Folate is an essential vitamin B, which occurs in diverse, but interconvertible chemical forms depending on the oxidative state of the one-carbon unit they carry. This family of enzymatic co-factors functions as one-carbon donors in a cellular network of biochemical reactions, by accepting and passing one-carbon units through enzymatic reactions, which compose the so-called folate-mediated one-carbon metabolism (FOCM). FOCM functions in the cellular compartments of the cytoplasm, mitochondria and nucleus and plays a crucial role for cell division, DNA repair and replication.

In the cytoplasm FOCM interconnects *de novo* purine synthesis, *de novo* thymidylate synthesis and the remethylation of homocysteine catalyzed by the B_{12}–dependent enzyme methionine synthase[1]. The primary functions of mitochondrial FOCM are to generate formate for the cytoplasm, to synthesize glycine from serine, to maintain mitochondrial dTMP synthesis, and to generate fMET-tRNA for mitochondrial protein synthesis initiation [11;227]. Recent biological evidence has demonstrated that FOCM mediated *de novo* dTMP biosynthesis also occurs in the nuclear compartment during S-phase or in response to DNA damage [12;13;42]. Nuclear dTMP synthesis enzymes form a multienzymatic complex at the replication fork, which may allow for substrate channelling between the active sites of the enzymes[2] [158].

In addition to its interconvertible chemical forms, folate derivatives also contain a polyglutamate peptide chain of varying length. Compared to the monoglutamate forms of folate, folate polyglutamates are the active form of folate cofactors in cells and have much higher affin-

[1] Folate is required to synthesis S-adenosyl methionine, a methyl donor used in hundreds of cellular reactions to promote methylation of DNA, histones, lipids, and neurotransmitters [227]

[2] Substrate channelling allows the direct transfer of co-factors between active sites of enzymes in the absence of diffusion.

ity for the respective folate-dependent enzymes and help to maintain intracellular folate levels [112;225;230;243;245].

Perturbations of the FOCM network can occur due to genetic or environmental/nutritional factors like folate deficiency, genetic polymorphisms in folate-dependent enzymes or vitamin deficiencies (i.e. vitamin B_6 or B_{12} deficiency). In particular, impaired FOCM is associated to severe health conditions and pathologies including neural tube defects (NTD) [25;175], certain types of cancer [29;31;204;292], and neurodegenerative diseases [134;219]. Several countries including the USA, Canada and Australia have established mandatory food fortification programs with folic acid, the synthesis form of folate to decrease the risk of folate-dependent NTDs and other pathologies. Even though these programmes seem to be sucessful [51], the causual pathways and underlying mechanisms have yet to be established. In addition to the associations to the physiological outcomes listed above, folate dependent enzymes are targeted by chemotherapeutic agents in cancer treatment, e.g. antifolate drugs like methotrexate (inhibitor of DHFR) [38] and 5-fluorouracil (inhibitior of TYMS) [152].

Mathematical Models of the folate-mediated one-carbon metabolism

Due to the key importance of FOCM, several mathematical models of FOCM have been introduced to address a wide range of questions including the effect of gene polymorphism and nutritional status of several vitamins on model outputs. Depending on the underlying research question, different aspects of the FOCM network are included and highlighted in the models. Mathematical models give a good framework to study FOCM, as the multiple interconnected reactions form a complex system that can not easily addressed by *in vivo/vitro* experiments as its complexity is increased by cellular compartmentalization and the role of genetic and environmental influence factors.

Starting from the 1970s, early models of FOCM have been developed [108;119;120;174;229;247]. The main focus of these studies was to quantify steady-state concentrations of folates and to predict how anti-cancer drugs like methotrexate and 5-fluorouracil affect the rates of *de novo* purine and thymidylate biosynthesis. In addition, similar models have been constructed to simulate and investigate the homocysteine remethylation cycle in liver [163;205;214]. These models considered FOCM reactions as occurring in a common cellular compartment, and described the reactions in terms of Michaelis-Menten kinetics.

More recent modeling efforts do not only include enhancements of FOCM modeling by integrating other pathways such as glutathione metabolism and polyamine synthesis, or by studying the influence of indirect effects beyond FOCM [185;215], but also investigate key regulative effects, long-range allosteric interactions, compartmentalization as well as genetic and environmental influence factors [185;187;188;212–215;261]

The published reaction-based models describe the continuous flux of metabolite concentrations in terms of ordinary differential equations (ODEs) with time as an independent variable. Sensitivity analysis is used to summarize the effects on metabolite concentrations and reaction velocities in response to changes in input variables or enzyme activities.

The initial work of Nijhout and Reed considers the folate cycle [187] and the methionine cycle [214] independently and standard biochemical kinetics were implemented to investigate key regulative

aspects like the influence of the methyl trap hypothesis induced by vitamin B_{12} deficiency, the inhibition of DHFR as used in cancer chemotherapy and enzyme or substrate deficiencies. By combining the two models a mathematical model of FOCM in the cytoplasm has been introduced and used to study a common polymorphism of MTHFR, as well as the effect of folate status, betaine, homocysteine and vitamin B_{12} on model dynamics [213]. In addition, simulations of this extended setting have been considered to investigate the effect of known and potential genetic polymorphisms of FOCM enzymes on biomarkers of cancer risk [261]. The same authors further extended the cytoplasmic FOCM model by including long-range allosteric interactions between the folate and methylation cycle [186] and by exploring mitochondrial compartimentalization of FOCM [188]. The *in silico* simulations of these last models matched experimental results regarding the variation in serine, glycine and SHMT expression and gave a better insight in the interplay between cytosol and mitochondria. The inclusion of glutathione metabolism and external in-/output from blood further extended this model and allowed to highlight the exchange with other cells to simulate, e.g., the metabolic profiles of Down syndrome and autism [215]. The effect of vitamin B_6 deficiency was considered in a further extension of the model [185]. A novelty of this publication is the direct modeling of the tight binding between 5mTHF and SHMT, for which the latter has been described as a model variable. This allowed to test the effect of different SHMT levels on model dynamics. More recently, a study on the regulatory properties of S-adenosylmethionine over long-range interactions and the binding of 5mTHF to GNMT considering only the cytosolic compartment was published [212]. The model predictions matched experimental results and provided further understanding of the underlying biological pathways.

Current mathematical models of FOCM have some limitations. First, these models are limited by adopting kinetic parameters determined using folate monoglutamate substrates, even if polyglutamate forms are crucial to keep folate in the cell [243]. Second, they are built upon data collected from different species and different tissues. Since the expression of folate-dependent enzymes and the regulation of their activity are known to vary across tissues and species this limits the interpretation of the results [227]. Third, the models do not incorporate 5fTHF, an intracellular storage form of folate, the role of which has not yet been fully understood. Fourth, so far published models focused only on the cytoplasmic and mitochondrial FOCM, excluding the nuclear compartment, which is important for *de novo* dTMP synthesis [158]. In **Part I** of this dissertation we will address these limitations.

Metabolic Syndrome

Metabolic Syndrome (MetSyn) is a common metabolic-related disorder defined as a complex cluster of co-occurring traits that increase the risk of cardiovascular diseases (CVD) and Type 2 Diabetes (T2D). Over the past decades the syndrome was known under alternative terms including Reaven syndrome, syndrome X, dysmetabolic syndrome X, CHAOS, plurimetabolic syndrome, the deadly quartet and insulin resistance syndrome [222]. Even though these terms came along with slightly varying definitions with respect to the relevant components and clinical criteria of the syndrome[3], they all describe a complex interplay of a number of similar risk factors such as

[3] Since 1998 several international organisations published statements on the clinical definition of MetSyn, among these are the World Health Organisation (WHO) [4], the European Group for the Study of Insulin resistance

hyperglycaemia, insulin resistance, obesity, dyslipidemia, hypertension and further components like chronic proinflammatory or pro-thrombotic states.

Following the commonly used definition of the joint interim statement from 2009 [3], the presence of three of the following five metabolic abnormalities qualify a patient for MetSyn: abdominal obesity, hyperglycaemia, hypertension, elevated levels of triglycerides and low levels of high-density lipoprotein (HDL) cholesterol. This definition highlights the equal interplay of the key features rather than centering MetSyn around one obligatory component [122].

As a direct consequence of the definitions of MetSyn, statistics on the prevalence of MetSyn differ depending on the selected clinical criteria and study cohort (based on age, sex, socio-economic status and ethnic background) [68;222]. However, a common observation of these studies is the increasing incidence of MetSyn in all western societies over the last decades [122]. A recently published analysis of the data from the National Health and Nutrition Examination Survey (NHANES) showed that among US adults the prevalence of MetSyn increased even up to 34.2% in 2012 [173].

These high numbers, and the fact that the underlying mechanisms of the complex metabolic dysregulation forming MetSyn are still not fully understood, implicate the need of further investigations. Especially with the elevated risk of developing T2D and CVDs – studies suggest a up to 5-times increased risk for MetSyn patients to develop T2D and a 2-times increased risk for CVDs, [122] – there is a rising need to develop new and efficient prevention and treatment strategies.

Because MetSyn is a complex, multifactorial disorder with genetic and environmental risk factors, and in view of the fact that the underlying etiology of MetSyn has only partially been elucidated, prevention strategies are complex. Initial treatment strategies always include a change in lifestyle based on physical activity and a balanced diet; while the consecutive step includes drug management, normally carried out as polydrug therapy according to the individual risk factors at hand [102]. However many patients are not able to fully reverse existing metabolic risk factors with lifestyle modification, and as risk factors worsen with advancing age, there is an increased need for efficacious and multi-functional drugs to manage particular risk factors [102].

Drug Repurposing

Drug repurposing (or repositioning) is the process of assigning a new indication to a previously approved drug. To name just two examples, methotrexate, an anti-folate drug originally used as chemotherapeutic agent, was later successfully approved for the treatment of rheumatoid arthritis [274]. Another example is the diabetes drug metformin, which is currently under investigation for novel applications as anti-cancer drug [39].

The benefits of drug repurposing arise from potential time, cost and risk reduction as approved drugs have already been optimized for safety and efficacy [19]. On average 10 to 15 years pass by until a drug is approved and introduced to the market, costing millions of dollars [263]. A recent

(EGIR) [22], the National Cholesterol Education Program (NCEP) Adult Treatment Panel III (ATPIII) [73], the American Association of Clinical Endocrinologists (AACE) [70] and the International Diabetes Federation (IDF) [5]. In 2009 a joint statement from the IDF, American Heart Association/National Heart, Lung and Blood Institute AHA/NHLBI, the World Heart Federation (WHF), the international Atherosclerosis Society, and the International Association for the Study of Obesity was published including a harmonized definition of MetSyn [3]. The dissimilarities of these definitions arise from the clinical criteria to manifest the included factors and from the emphasis on obligatory components like insulin-resistense or obesity. See [115;122;167;222] for a more detailed overview and the clinical criteria suggested in the single definitions.

study even estimated the overall costs above 1.3 billion dollar per drug [63;263]. When following a repurposing approach, the time-consuming and expensive trial phase for the novel application of the drug can often be abridged. Consequently, "approximately 50% of the cost may be reduced to get a drug approved for a new indication" [178].

In the past, one of the main sources for detecting repurposing candidates has been through serendipity, e.g., by observing side effects or by examining clinical or epidemological data related to a specific drug [146]. However, in consequence of the growing amount of available data for pharmacological processes and disease pathophysiology more systematic analysis approaches based on computational methods can be developed. Computational methods offer a meaningful framework to strategically investigate potential new treatment possibilities and accelerate therefore the drug development process. The methodologies used for these approaches include machine learning algorithms, modeling strategies, network analysis, text-mining or semantic inference [146]. Mining different data sources like genomic, transcriptomic, phenotypic or clinical data can provide a useful opportunity to identify potential new therapeutic uses by searching e.g. for similarities between drugs or disease profiles [114;176;231;232;290]. In addition also the identification of reversed/inversely correlated profiles of drugs and diseases may lead to a positive outcome [41]. Approaches based on network analysis successfully investigated new indications for existing drugs [44;104;153], predicted new potential anticancer treatments [46;267] and identified new promising targets [71;117;145].

Part I

Folate-Mediated One-Carbon Metabolism

Chapter 1

A hybrid stochastic model of folate-mediated one-carbon metabolism: Effect of the common C677T *MTHFR* variant on *de novo* thymidylate biosynthesis

Folate-mediated one-carbon metabolism (FOCM) is an interconnected network of metabolic pathways, including those required for the *de novo* synthesis of dTMP and purine nucleotides and for remethylation of homocysteine to methionine. Mouse models of folate-responsive neural tube defects (NTDs) indicate that impaired *de novo* thymidylate (dTMP) synthesis through changes in SHMT expression is causative in folate-responsive NTDs. In this chapter we present a hybrid computational model comprised of ordinary differential equations and stochastic simulation. We investigate whether the *de novo* dTMP synthesis pathway is sensitive to perturbations in FOCM that are known to be associated with human NTDs. This computational model shows that *de novo* dTMP synthesis is highly sensitive to the common *MTHFR* C677T polymorphism and that the effect of the polymorphism on FOCM is greater in folate deficiency. Computational simulations indicate that the *MTHFR* C677T polymorphism increases the stochastic behavior of the FOCM network, with the greatest instability observed for reactions catalyzed by serine hydroxymethyltransferase (SHMT). Furthermore, we show that *de novo* dTMP synthesis does not occur in the cytosol at rates sufficient for DNA replication, supporting empirical data indicating that impaired nuclear *de novo* dTMP synthesis results in uracil misincorporation into DNA.

1.1 Introduction

Perturbations in folate-mediated one-carbon metabolism (FOCM) are associated with numerous pathologies including neural tube defects (NTDs) [25], stroke [285], colorectal and other types of cancer [31;204;292]. Furthermore, enzymes in FOCM have been successful targets for the development of antineoplastic pharmaceutical agents including methotrexate and 5-flurouracil [265]. FOCM

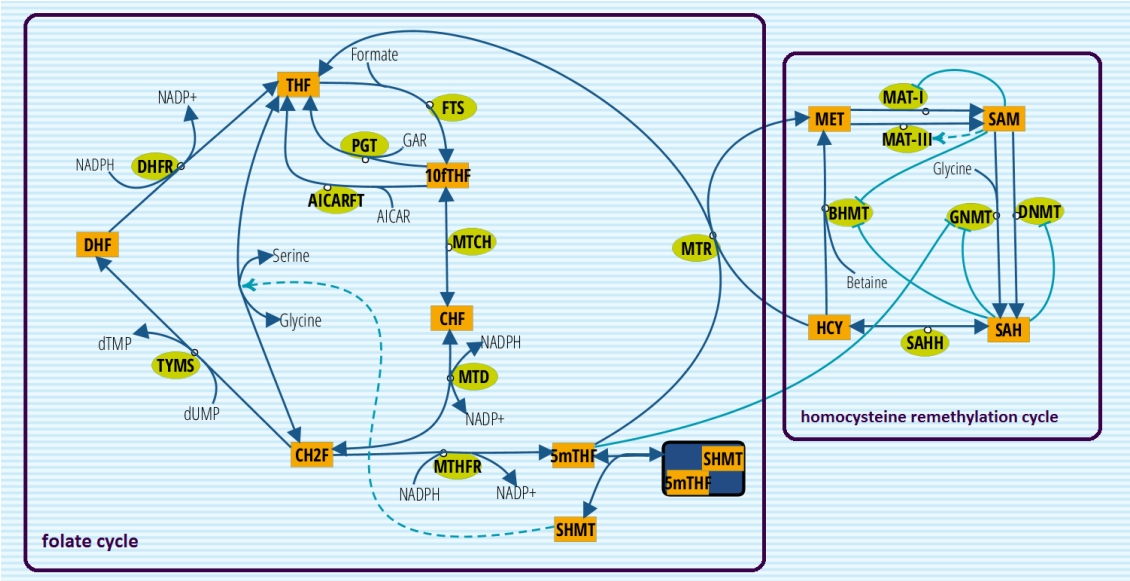

Figure 1.1: The reaction-based specification of the model according to the notation introduced in [94]. Rectangles identify model variables, non-boxed substrates are model constants, green circles identify enzymes, dark blue arcs identify matter transformation, and light blue arcs identify regulatory events (dotted lines indicate activations and solid lines indicate inhibitions). The purple boxes indicate reactions and variables associated with the folate cycle and the homocysteine remethylation cycle, respectively.

in the cytoplasm is composed of three interconnected biosynthetic pathways, which include *de novo* thymidylate (dTMP) synthesis, *de novo* purine synthesis and homocysteine remethylation to methionine (Figure 1.1). The FOCM network is sensitive to nutritional status for several vitamins that serve as enzyme cofactors (folate, riboflavin, vitamin B_6 and vitamin B_{12}) and genetic factors (coding and expression variants in folate-dependent enzymes) that can alter network outputs, including DNA synthesis, DNA repair and chromatin methylation [66;85;238]. Understanding the molecular basis of disease etiology has been limited by the ability to ascribe specific FOCM pathways and their biomarkers to clinical outcomes, because the pathways of FOCM are tightly interconnected [238]. FOCM complexity is manifest by: (a) competition among the pathways for a limiting pool of folate cofactors [110], (b) long-range and indirect regulatory processes, (c) formation of multi-enzyme complexes, (d) cellular compartmentalization, (e) interactions with other metabolic pathways, (f) nutritional status (g) penetrant genetic variants [85;227]. Mathematical models have been developed to assess this complexity and gain an understanding of the cause-and-effect relationships that regulate FOCM functioning in health and disease. The overall goal is to provide an understanding of function of the entire system *in silico* that can be used to accelerate discovery and guide the design of biological experimentation.

Here we present a hybrid stochastic model for simulating the FOCM dynamics where state-of-the art deterministic simulation (based on ODEs) has been coupled with exact stochastic simulation to assess metabolite variabilities in the FOCM network at steady state.

The deterministic approach used in isolation can only provide a rough estimate of FOCM model dynamics, because the deterministic approach is limited when enzyme substrates, such as folate cofactors, are present in the cells at low micromolar concentrations, and because reactions within

the network occur randomly at discrete time points. FOCM is expected to exhibit variability (i.e. stochasticity) in its behavior [93]. Capturing system stochasticity is essential when substrate concentrations are low and limiting, but requires consideration of molecules as discrete entities, rather than describing concentrations as continuous variables through ODEs [90]. Simulation strategies that combine both deterministic and stochastic approaches can give a more accurate and more detailed understanding of FOCM network functioning and stability. In contrast to approaches based solely on deterministic simulation, these studies can be used to assess the contributions of factors such as genetic variation and nutritional status on the stochastic behavior of individual pathways within the network, thereby aiding in establishing which system inputs (i.e. nutrition) and outputs (i.e. biomarkers) are most closely associated with human health outcomes.

Dysregulation of the partitioning of one-carbon units in the form of 5,10-methylenetetrahydrofolate (CH2F) cofactors between the *de novo* dTMP biosynthesis and homocysteine remethylation pathways is believed to underlie FOCM-associated pathologies including NTDs (Figure 1.1). A common variant of *MTHFR*, the C677T polymorphism, has been associated with numerous pathologies including birth defects, cancer, cardiovascular events, and other pathologies [226]. MTHFR catalyzes the FADH-dependent, irreversible conversion of CH2F to 5mTHF, which commits folate cofactors away from dTMP synthesis and towards homocysteine remethylation in the cytosol (Figure 1.1). The variant results from an alanine to valine substitution in the protein that decreases MTHFR activity by decreasing its affinity for the FADH cofactor. Such substitution affects enzyme stability and hence the partitioning of folates between dTMP synthesis and homocysteine remethylation [47;207]. Decreased MTHFR activity resulting from the polymorphism decreases 5mTHF synthesis, leading to impaired homocysteine remethylation and elevated serum homocysteine [226]. The 677T variant is also associated with a redistribution of cellular folate cofactors; 5mTHF is the predominate form of folate in red blood cells in *MTHFR* 677CC carriers, whereas 10fTHF is the predominate form of folate in *MTHFR* 677TT carriers [21;58;89]. 10fTHF is less chemically stable than 5mTHF, and the *MTHFR* 677TT variant is associated with lower folate status [244] and higher folate requirement [234]. Recent studies suggest that the contribution of the MTHFR variant to NTD risk is due to its impact on cellular folate status, rather than impaired homocysteine remethylation [257]. Likewise, mouse models of NTDs indicate that impaired dTMP synthesis, and not homocysteine remethylation, cause folate- responsive NTDs [23;24;162]. Reed et al. investigated the consequences of the *MTHFR* C677T polymorphism, assuming 70% enzyme activity for heterozygote and 30% enzyme activity for homozygote, in comparison to CC homozygotes (which was set to 100% activity) using parameters for folate monoglutamates, which are not the physiological form of folate cofactors in cells. Under these conditions, the variant allele decreased concentrations of 5mTHF and SAM and increased the concentrations of homocysteine, SAH, and rates of dTMP and purine biosynthesis [213]. The effect on the redistribution of folate cofactors towards 10fTHF that is associated with the 677T variant, or its impact on other pathways within the network, was not reported [213].

Here, we studied the partitioning of CH2F, a cofactor for both homocysteine remethylation and *de novo* dTMP biosynthesis [78;80;282], and the effects of known genetic and nutritional variables that impact movement of CH2F through the network.

Existing models are limited by adopting kinetic parameters determined from the use of folate monoglutamate substrates [185;215]. Folate polyglutamates are the functional form of folate

cofactors in cells and have much higher affinity for their respective FOCM enzymes than the corresponding monoglutamate forms of folate [112;225]. Therefore, we updated the parameters in the deterministic model to include the physiologically- relevant polyglutamate forms of folate cofactors and demonstrate that it faithfully recapitulates existing data in the literature. The decisions for selecting individual enzyme kinetic parameters for this model were driven by: 1) available data for physiologically relevant polyglutamate forms of the folate cofactors derived from characterization of mammalian enzymes, and 2) data from human models, specifically L1210 cells, because of the richness and quality of the data used to derive kinetic parameters. Given the high conservation of folate enzymes among mammals, our model could be applied to mammalian systems in general, even though we are not proposing a completely homogeneous model with respect to species.

We were able to identify key nodes in the network of Figure 1.1 that exhibit high degrees of stochastic behavior, including the influence of nutrient status and genetic variation on stochasticity through simulations. We explored the impact of the *MTHFR* C677T polymorphism and its interaction with folate status on partitioning of CH2F within the network, including its impact on *de novo* dTMP biosynthesis to understand the etiology of NTDs. The results of the computational model provide evidence that the rates of *de novo* dTMP synthesis as currently modeled in the cytosol are insufficient to support DNA synthesis in S-phase in mammals, accounting for uracil misincorporation into DNA that occurs in folate deficiency and in mouse models of NTDs.

1.2 Materials and Methods

1.2.1 Description of the Model and Simulation Techniques

The model was constructed as a closed system using the subset of reactions that describe the FOCM pathways and homocysteine remethylation in cytoplasm [215] (Figure 1.1). For the simulation, we employed a hybrid stochastic approach using deterministic simulation to compute the initial phase of the dynamics until a model steady state was reached, and then we assessed the stability of the achieved steady state by relying on the concept of total propensities arising from exact stochastic simulation. We adopted a hybrid approach, rather than one entirely based on exact stochastic simulation [161], because of the intensive computational effort introduced by the stiffness of the system during the simulation.

The deterministic simulation was based on ODEs, where reactions were described in terms of Michaelis-Menten equations consistent with the original model of Reed [215] and computed using the MATLAB integrator ode15s, whereas parameter estimates were derived from literature or calculated by nonlinear least squares optimization. The kinetic constants were obtained from folate polyglutamate cofactors and their interaction with enzymes purified from L1210 cells where possible and otherwise from other mammalian tissue.

The set of ODEs was further translated into a stochastic reaction-based model and a hybrid simulation approach was employed to quantify the level of stochasticity in the considered FOCM steady states. According to the seminal work of Gillespie [90], exact stochastic simulation allows simulating each reaction event explicitely when it's most likely to occur. In each step of the simulation algorithm a propensity function $a_j(x)$ for each modeled reaction R_j is calculated, where x is the current state of the system which provides the abundances of all modeled species at the

considered time. The propensity value of a reaction $a_j(x)$ has a direct link to the probability of its execution: reactions with higher propensities are more likely to be fired in the near future. To evaluate when the next reaction event will occur, the total sum of propensities $a_0(x) = \sum_j a_j(x)$ is computed, because this quantity is linked to the number of reaction events occurring in the next time unit. Indeed, with increasing total propensity the number of reaction events per unit of time also increases. In Table 1.8, the total propensities for four considered steady states are provided; to apply exact stochastic simulation we would need to generate up to $10^{15} - 10^{16}$ reaction events per unit of time. To circumvent this problem of stiffness, we relied on the concept of total propensity to evaluate the stability of the steady state, by assuming that a steady state is more stable than another, when it exhibits a lower total propensity $a_0(x)$. This means that on average this steady state leads to a lower number of reaction events that can perturb its equilibrium.

1.2.2 Technical description of the ODE model

In order to be consistent with previous literature [185;187;212–215;261], the model has been initially defined as a set of ordinary differential equations (ODEs) as introduced below and then translated into a stochastic model as explained in the following. The model can be considered as an extension of the previously published model [213] and a comparison to this model is provided in Section 1.2.4. The model consists of twelve variables:

- the different forms of folate: THF, 10fTHF, CHF, CH2F, DHF, and 5mTHF;
- the enzyme SHMT;
- the complex 5mTHF: SHMT formed by binding of 5mTHF to SHMT;
- the metabolites MET, HCY, SAM and SAH.

The other substrates (NADPH, NADP, dUMP, Serine, Glycine, GAR, AICAR, betaine and formate) are approximated to be constant over time in agreement with previous modeling literature [213]. In terms of reaction description, constant reactants are indicated above the arrow next to the enzyme:

$$\mathrm{S_1} \xrightarrow{\mathrm{enzyme,S_2}} \mathrm{P}$$

Due to the biological description of FOCM, our model can be divided into two connected modules, as indicated in Figure 1.1. The first one specifies the reactions linking the different forms of folate; we further refer to this module as the folate cycle. The second module is the homocysteine remethylation cycle, including the four metabolites HCY, MET, SAM and SAH and the six reactions associated with these. The connection between those two parts is formed by the biochemical reaction catalyzed by methionine synthase (MTR), which regenerates MET from HCY using 5mTHF as the donor of one methyl group. Following [213], most of the enzymatic reactions considered in the model have been translated in the set of ODEs by means of Michaelis-Menten kinetics, which consider one or two different substrates $\mathrm{S_1}$ and $\mathrm{S_2}$. The formula of the Michaelis-Menten kinetics with one substrate S is:

$$v(\mathrm{S}) = \frac{\mathrm{V_{max}} \cdot \mathrm{S}}{\mathrm{K_m} + \mathrm{S}},$$

where V_{max} indicates the maximum rate of the considered reaction and K_m is the Michaelis-Menten constant that specifies the concentration of the associated substrate for which the rate is half-maximum. In the same way, the formula can be extended to consider two substrates:

$$v(S_1, S_2) = \frac{V_{max} \cdot S_1 \cdot S_2}{(K_{m_1} + S)(K_{m_2} + S_2)},$$

where K_{m_1} and K_{m_2} are the Michaelis-Menten constants for the two substrates.

For a clear and unique description of the model, each reaction and each velocity is labeled with the name of the enzyme catalyzing it. For example, R_{DHFR} and v_{DHFR} define the reaction and the velocity of the transformation of DHF to THF, catalyzed by dihydrofolate reductase (DHFR). All variable initial concentrations, constant values, and parameter estimates for the folate cycle and homocysteine remethylation used in the model are listed in Appendix A.1 (Tables A.1, A.2, A.3, and A.4). All concentrations are expressed in µM, while time is expressed in hours.

For the sake of simplicity, hereafter we will present the mathematical model by focusing on reactions and their corresponding kinetic formulas. The differential equations of the model can then be derived by summing these formulas according to reaction stoichiometry. In fact, each arrow from the model visualization in Figure 1.1 connected to one variable corresponds to one term in the sum of the associated differential equation. For example, if we consider DHF we see that this variable is connected to one outgoing and one incoming arrow (promoted by DHFR and TYMS, respectively). Therefore, the corresponding differential equation is:

$$\frac{d[\text{DHF}]}{dt} = v_{\text{TYMS}}(\text{CH2F}, \text{dUMP}) - v_{\text{DHFR}}(\text{DHF}, \text{NADPH})$$

where v_{TYMS} and v_{DHFR} will be defined in the following.

The folate cycle

The majority of the reactions in the folate cycle are unidirectional with a time-variant and a constant reactant, like R_{DHFR}:

$$R_{\text{DHFR}}: \quad \text{DHF} \xrightarrow{\text{DHFR,NADPH}} \text{THF}$$

$$v_{\text{DHFR}}([\text{DHF}], \text{NADPH}) = \frac{V_{max} \cdot [\text{DHF}] \cdot \text{NADPH}}{(K_{\text{DHF}} + [\text{DHF}]) \cdot (K_{\text{NADPH}} + \text{NADPH})},$$

where squared brackets indicate the variable concentrations.

The same translation can be applied also to the following reactions:

$$R_{\text{TYMS}}: \quad \text{CH2F} \xrightarrow{\text{TYMS,dUMP}} \text{DHF}$$

$$R_{\text{AICARFT}}: \quad \text{10fTHF} \xrightarrow{\text{AICARFT,AICAR}} \text{THF},$$

$$R_{\text{PGT}}: \quad \text{10fTHF} \xrightarrow{\text{PGT,GAR}} \text{THF},$$

$$R_{\text{FTS}}: \quad \text{THF} \xrightarrow{\text{FTS,formate}} \text{10fTHF},$$

$$R_{\text{MTHFR}}: \quad \text{CH2F} \xrightarrow{\text{MTHFR,NADPH}} \text{5mTHF}$$

The only reaction with two non-constant substrates in the folate-cycle is the one catalyzed by methionine synthesis:

$$R_{\mathrm{MTR}}: \quad \mathrm{HCY} + 5\mathrm{mTHF} \xrightarrow{\mathrm{MTR}} \mathrm{THF} + \mathrm{MET}$$

$$v_{\mathrm{MTR}}([\mathrm{HCY}],[5\mathrm{mTHF}]) = \frac{\mathrm{V_{max}} \cdot [\mathrm{DHF}] \cdot [5\mathrm{mTHF}]}{(\mathrm{K_{HCY}} + [\mathrm{HCY}]) \cdot (\mathrm{K_{5mTHF}} + [5\mathrm{mTHF}])}.$$

The $\mathrm{V_{max}}$ estimate of this reaction is the only one that has not been directly taken from literature, but rather optimized in the range 0.024 - 500 µM/h from literature[164;215] to obtain the trends discussed above.

The next subset of reactions of the folate cycle contains the three bidirectional reactions R_{MTCH}, R_{MTD} and R_{SHMT}. R_{MTCH} links 10fTHF and CHF as a bidirectional reaction with one substrate:

$$R_{\mathrm{MTCH}}: \quad 10\mathrm{fTHF} \xleftrightarrow{\mathrm{MTCH}} \mathrm{CHF}$$

$$v_{\mathrm{MTCH}}([10\mathrm{fTHF}],[\mathrm{CHF}]) = \frac{\mathrm{V_{max,10fTHF}} \cdot [10\mathrm{fTHF}]}{\mathrm{K_{10fTHF}} + [10\mathrm{fTHF}]} - \frac{\mathrm{V_{max,CHF}} \cdot [\mathrm{CHF}]}{\mathrm{K_{CHF}} + [\mathrm{CHF}]}.$$

The subsequent reaction R_{MTD} is a bidirectional reaction with two substrates:

$$R_{\mathrm{MTD}}: \quad \mathrm{CHF} \underset{\mathrm{MTD,NADP}}{\overset{\mathrm{MTD,NADPH}}{\rightleftarrows}} \mathrm{CH2F}$$

$$v_{\mathrm{MTD}}([\mathrm{CHF}],\mathrm{NADPH},[\mathrm{CH2F}],\mathrm{NADP}) = \frac{\mathrm{V_{max,CHF}} \cdot [\mathrm{CHF}] \cdot \mathrm{NADPH}}{(\mathrm{K_{CHF}} + [\mathrm{CHF}]) \cdot (\mathrm{K_{NADPH}} + \mathrm{NADPH})}$$

$$- \frac{\mathrm{V_{max,CH2F}} \cdot [\mathrm{CH2F}] \cdot \mathrm{NADP}}{(\mathrm{K_{CH2F}} + [\mathrm{CH2F}]) \cdot (\mathrm{K_{NADP}} + \mathrm{NADP})}.$$

A slight change in terms of Michaelis-Menten kinetics can be found in the glycine and serine dependent reaction between THF and CH2F. As this reaction is catalyzed by the time-dependent enzyme SHMT, its change of concentration has to be taken into account[185]. Therefore, we used here the turnover number $\mathrm{k_{cat}}$, which describes the conversion of the Enzyme-Substrate complex to the product, where

$$\mathrm{V_{max}} = \mathrm{k_{cat}} \cdot [\mathrm{SHMT}].$$

The reaction is then defined in the following way:

$$R_{\mathrm{SHMT}}: \quad \mathrm{THF} \underset{\mathrm{SHMT,Glycine}}{\overset{\mathrm{SHMT,Serine}}{\rightleftarrows}} \mathrm{CH2F}$$

$$v_{\mathrm{SHMT}}([\mathrm{THF}],\mathrm{Serine},[\mathrm{CH2F}],\mathrm{Glycine}) = \frac{\mathrm{k_{cat,THF}} \cdot [\mathrm{SHMT}] \cdot [\mathrm{THF}] \cdot \mathrm{Serine}}{(\mathrm{K_{THF}} + [\mathrm{THF}]) \cdot (\mathrm{K_{Serine}} + \mathrm{Serine})}$$

$$- \frac{\mathrm{k_{cat,CH2F}} \cdot [\mathrm{SHMT}] \cdot [\mathrm{CH2F}] \cdot \mathrm{Glycine}}{(\mathrm{K_{CH2F}} + [\mathrm{CH2F}]) \cdot (\mathrm{K_{Glycine}} + \mathrm{Glycine})}.$$

The concentration of active SHMT enzyme changes over time because in the model we consider the tight binding of 5mTHF to SHMT[185]. The corresponding reactions are modeled by mass

action kinetics with rates $k_{binding}$ and $k_{unbinding}$ [239]:

$$\mathrm{SHMT} + \mathrm{5mTHF} \underset{k_{unbinding}}{\overset{k_{binding}}{\leftrightarrows}} \mathrm{SHMT : 5mTHF}$$

$$v_{binding} = k_{binding} \cdot [\mathrm{5mTHF}] \cdot [\mathrm{SHMT}]$$

$$v_{unbinding} = k_{unbinding} \cdot [\mathrm{SHMT : 5mTHF}].$$

Homocysteine Remethylation

The reactions responsible for homocysteine remethylation involve BHMT, DNMT, GNMT, MAT-I, MAT-III and SAHH. These reactions are entirely taken from literature [215].

A bidirectional Michaelis-Menten function is used to model the conversion between HCY and SAH.

$$R_{\mathrm{SAHH}}: \quad \mathrm{HCY} \overset{\mathrm{SAHH}}{\longleftrightarrow} \mathrm{SAH}$$

$$v_{\mathrm{SAHH}}([\mathrm{HCY}], [\mathrm{SAH}]) = \frac{V_{max,HCY} \cdot [\mathrm{HCY}]}{K_{HCY} + [\mathrm{HCY}]} - \frac{V_{max,SAH} \cdot [\mathrm{SAH}]}{K_{SAH} + [\mathrm{SAH}])}.$$

The betaine-dependent reaction of remethylation of HCY is presented as a two substrate Michaelis-Menten equation with an additional inhibition term dependent on SAM and SAH:

$$R_{\mathrm{BHMT}}: \quad \mathrm{HCY} \xrightarrow{\mathrm{BHMT,Betaine}} \mathrm{MET}$$

$$v_{\mathrm{BHMT}}([\mathrm{HCY}], \mathrm{Betaine}) = \frac{V_{max} \cdot [\mathrm{HCY}] \cdot \mathrm{Betaine}}{(K_{HCY} + [\mathrm{HCY}]) \cdot (K_{Betaine} + \mathrm{Betaine})} \cdot e^{-0.0021([\mathrm{SAM}]+[\mathrm{SAH}])} \cdot e^{0.0021 \cdot 102.6}.$$

For the conversion of MET to SAM, two reactions which are regulated by MAT-I and MAT-III, respectively, are considered. The first one is a first-order Michaelis-Menten function with a nonlinear inhibition term dependent on SAM. $R_{\mathrm{MAT-III}}$ was fitted to a Hill equation, including also an activation term by SAM:

$$R_{\mathrm{MAT-I}}: \quad \mathrm{MET} \xrightarrow{\mathrm{MAT-I}} \mathrm{SAM}$$

$$v_{\mathrm{MAT-I}}([\mathrm{MET}]) = \frac{V_{max} \cdot [\mathrm{MET}]}{K_{MET} + [\mathrm{MET}]} \cdot (0.23 + 0.8e^{-0.0026[\mathrm{SAM}]})$$

$$R_{\mathrm{MAT-III}}: \quad \mathrm{MET} \xrightarrow{\mathrm{MAT-III}} \mathrm{SAM}$$

$$v_{\mathrm{MAT-III}}([\mathrm{MET}]) = \frac{V_{max} \cdot [\mathrm{MET}]^{1.21}}{K_{MET} + [\mathrm{MET}]^{1.21}} \cdot \left(1 + \frac{7.2 \cdot [\mathrm{SAM}]^2}{K_a^2 + [\mathrm{SAM}]^2}\right)$$

Two methyltransferases DNMT and GNMT are included in our model; both are affected by the inhibition of SAH:

$$R_{\mathrm{DNMT}}: \quad \mathrm{SAM} \xrightarrow{\mathrm{DNMT}} \mathrm{SAH}$$

$$v_{\mathrm{DNMT}}([\mathrm{SAM}]) = \frac{\mathrm{V_{max}} \cdot [\mathrm{SAM}]}{\mathrm{K_{SAM}} \cdot (1 + \frac{[\mathrm{SAH}]}{\mathrm{K_i}}) + [\mathrm{SAM}]}$$

$$R_{\mathrm{GNMT}}: \quad \mathrm{SAM} \xrightarrow{\mathrm{GNMT}} \mathrm{SAH}$$

$$v_{\mathrm{GNMT}}([\mathrm{SAM}], \mathrm{Glycine}) = \frac{\mathrm{V_{max}} \cdot [\mathrm{MET}] \cdot \mathrm{Glycine}}{(\mathrm{K_{SAM}} + [\mathrm{SAM}]) \cdot (\mathrm{K_{Glycine}} + \mathrm{Glycine})} \cdot \frac{1}{1 + \frac{[\mathrm{SAH}]}{\mathrm{K_i}}}$$

1.2.3 Stochastic specification of the model

In addition to the ODE specification, the model has been translated to a stochastic reaction based formulation. This was achieved by scaling all metabolite and enzyme concentrations, as well as Michaelis-Menten constants, to number of molecules instead of concentrations. If we consider, for example, the concentration $[THF] = 0.12\,\mu\mathrm{M}$, the corresponding number of molecules $\#THF$ is

$$\begin{aligned}
\#\mathrm{THF} &= [\mathrm{THF}] \cdot \mathrm{N_A} \cdot \mathrm{K_{vol}} \cdot \mathrm{k_{cyt}} \cdot \mathrm{V_{L1210}} \\
&= 0.12\ \mu\mathrm{M} \cdot 6.022 \cdot 10^{23} \frac{1}{\mathrm{mol}} \cdot 10^{-6} \frac{\mathrm{M}}{\mu\mathrm{M}} \cdot 0.75 \cdot 0.63 \cdot 10^{-12}\mathrm{L} \\
&= 34145,
\end{aligned}$$

where $\mathrm{N_A} = 6.022 \cdot 10^{23} \frac{1}{\mathrm{mol}}$ is the Avogadro constant, $\mathrm{V_{L1210}} = 0.63 \cdot 10^{-12}\,\mathrm{L}$ is the average cell volume of the L1210 cell line[1] and $\mathrm{k_{vol}} = 10^{-6} \frac{\mathrm{M}}{\mu\mathrm{M}}$, $\mathrm{k_{cyt}} = 0.75$ are two scaling factors. The first is used to transform the concentration from µM to M; the second is introduced because only the reactions occurring in cytoplasm have been considered, referring to 75% of the total cell volume[155].

After the translation of concentrations and Michaelis-Menten constants to number of molecules, the propensities $a_j(x)$ for all reactions R_j were calculated. For the reactions formulated in terms of Michaelis-Menten or Hill kinetics, the propensities are computed by the same functions (see preceding sections) where parameters are expressed in terms of number of molecules, e.g. the propensity for the reaction catalyzed by DHFR is

$$a_{\mathrm{DHFR}}(\#\mathrm{DHF}, \#\mathrm{NADPH}) = \frac{\mathrm{V_{max}}^{\#} \cdot \#\mathrm{DHF} \cdot \#\mathrm{NADPH}}{(\mathrm{K}^{\#}_{\mathrm{DHF}} + \#\mathrm{DHF})(\mathrm{K}^{\#}_{\mathrm{NADPH}} + \#\mathrm{NADPH})},$$

where $\mathrm{K}^{\#}_{\mathrm{DHF}}$, $\mathrm{K}^{\#}_{\mathrm{NADPH}}$ and $\mathrm{V_{max}}^{\#}$ indicate the transformed Michaelis-Menten kinetic parameters. In the case of the two mass-action reactions modeling the binding/unbinding of 5mTHF and SHMT the propensities are:

$$a_{\mathrm{binding}}(\#\mathrm{5mTHF}, \#\mathrm{SHMT}) = \mathrm{k_{binding}}/\mathrm{V_{L1210}} \cdot \#\mathrm{5mTHF} \cdot \#\mathrm{SHMT},$$

$$a_{\mathrm{unbinding}}(\#\mathrm{5mTHF:SHMT}) = \mathrm{k_{unbinding}} \cdot \#\mathrm{5mTHF:SHMT}.$$

Finally,

$$a_0(x) = \sum_{R_j} a_j(x)$$

gives the sum of all propensities in the current state x.

1.2.4 Comparison with literature

The first mathematical models of FOCM were developed in the 1970s and 1980s [119;174;229] mainly focusing on the effect of anticancer drugs on the network. Starting from these first attempts, new models were introduced in literature to update/extend our understanding of the network based on new experimental evidence (refer to the Introduction or [189] for more details).

The folate and homocysteine remethylation cycles considered in our model have been initially modeled separately [187;214] and have later been merged in a single model [213]. The latter has been further extended by the same authors to study different aspects of FOCM (e.g. the interplay between mitochondrial and cytoplasmic FOCM or the inclusion of the glutathione metabolism) [185;212;215;261]. Following the same approach, the model herein proposed can be considered as another extension of this previously published model [213]. In the following, this model will be used as reference for the comparison with literature.

The model presented herein constitutes an update with respect to [213] according to two main aspects. The first and most important one relates to a more physiologically relevant selection of parameter estimates in modeling the folate cycle. The main improvement of our model with respect to literature is the selection of parameter estimates used in modeling the folate cycle. Parameter estimates were chosen according to two criteria. First, a homogeneous set of parameters was identified by referring, when possible, to L1210 cells. We chose this cell line because of the richness of the data available in for quantifying enzyme levels of enzymes and for their kinetic characterization of the enzyme using polyglutamate substrates in this cell line. The second criterion relates to the length of the glutamate chain attached to the folate. In previous models most of the parameters are estimated using the affinity of the enzymes for monoglutamates. However, polyglutamate forms are the physiologically relevant cofactors. They play a crucial role to sequester folates in the cell given their higher affinity with the enzymes if compared to monoglutamates [243]. Therefore, our preference was to select data from polyglutamate-derived coefficients. Detail on considered cell lines and length of glutamate chain are listed in Appendix A, Table A.3.

The second improvement to the structure of model equations, which have been modified to add new reactions and to update other reactions by including new biochemical interactions as described below. Figure 1.2 displays a point-to-point comparison between the structure of our model and the one in [213]. All modifications have been highlighted by coloring the corresponding part of the network (the green color identifies unmodified reactions, orange indicates updates/novelties, black indicates not included parts). In the following we will provide a detailed list of all the modifications.

1. Our model includes two new terms to highlight take into account that, according to [196], the rate of reaction R_{MTD} depends also on NADPH and NADP as second substrates.

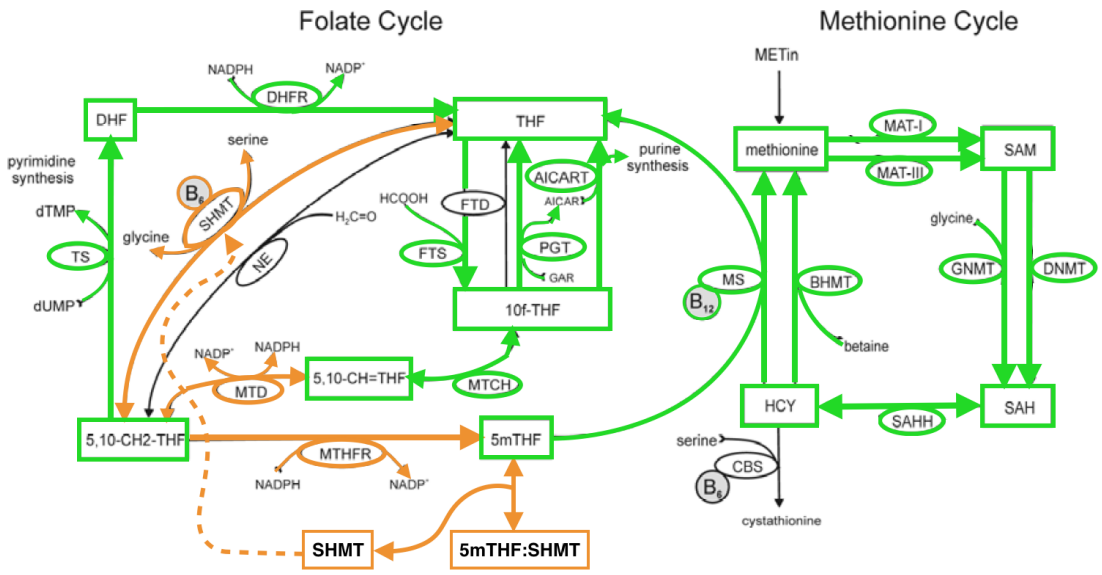

Figure 1.2: A point-to-point comparison between the herein proposed model and the one in [213]. Rectangles indicate variables of the model; enzymes are indicated in oval frames attached to the corresponding reaction represented by arcs; non-boxed substrates/products are model constants. Based on the model structure used in [213] the figure displays the applied modifications using different colors. Green colored reactions as well as green framed variables and enzymes indicate the adopted parts of the model. Orange colored reactions as well as orange framed variables and enzymes indicate parts of the network which have been modified or added. Black arcs and black framed variables and enzymes indicate parts of the network we did not consider. The comparison considers the structure of model equations, while parameter estimates have been updated with respect to [213]. The figure has been adapted from [213].

2. We further included the bidirectional reaction modeling the binding/unbinding of 5mTHF and SHMT [189;239] to study the effect of the MTHFR polymorphism on the availability of unbound SHMT. This update also affected on the definition of the reaction RSHMT, which has been updated to consider SHMT as model variable rather than a constant.

3. We also tried to build a parsimonious model in order to reduce the problem of overfitting of the system dynamics as much as possible. In particular, a closed model was built rather than considering external input/output pathways (see MET and cystathionine in Figure 1.2). Moreover,and the number of non-enzymatic reactions was reduced because of a high degree of uncertainty regarding their physiological significance (see the non-enzymatic reaction NE in Figure 1.2).

4. The inhibition of MTHFR by SAM initially proposed in [213] could not be integrated in our model, because it has been modeled by a nonlinear term whose parameters have been estimated under conditions that do not apply to our modeling scenario. In fact, the estimation of this term was based on the external methionine input (included in [213], but not in our closed model, see previous point) and on some kinetic parameters which differ by an order of magnitude with respect to estimates herein considered. Moreover, following [213],

the inhibition should only act when SAH < SAM, but this is not a common case in our simulations.

5. Finally, we decided not to include the reaction between 10fTHF and THF promoted by the enzyme FTD, because we observed that the strong inhibition of FTD by 10fTHF discussed in [127] makes the effect of this reaction in the considered steady states negligible.

1.2.5 Model validation

To validate the FOCM model, *in silico* experiments were performed to determine if the model could recapitulate empirical data generated in MCF-7 cells by Herbig et al. [110] focusing on the effect of glycine on FOCM (Table 1.1). Glycine is important because, as a second substrate, it has a direct influence on the reaction catalyzed by the enzymes GNMT as well as on the reversible reaction transforming CH2F to THF catalyzed by the enzyme SHMT. The purpose is to understand how the steady state of FOCM is affected by altering intracellular glycine concentrations. This was achieved by running several model simulations starting from different glycine concentrations and comparing the corresponding steady states with empirical data from Herbig et al. [110]. This study examined the effect of exogenous glycine at concentrations from 0 to 10 mM on the relative distribution of folate one-carbon forms as well as S-adenosylmethionine (SAM) and S-adenosylhomocysteine (SAH) levels. In summary, the empirical data revealed that as glycine concentrations increase intracellular 10fTHF levels increase at the expense of 5mTHF levels that decrease. Furthermore, as glycine concentrations increase SAM levels are depleted and SAH levels rise. These changes were interpreted by the effects of glycine concentration driving the reversible SHMT reaction in the direction of serine synthesis [110].

We simulated the effect of glycine on folate distribution, SAM, and SAH concentrations using the computational model for values of glycine ranging from 0 to 10 mM (Table 1.1). The trends obtained by the model simulations were in agreement with the literature, confirming the coherence between model outcomes and empirical data. We observed only one exception related to the total % of 5mTHF at 10 mM glycine. This discrepancy could be mainly due to two reasons: (1) 10 mM glycine is an extreme and non-physiological intracellular glycine concentration that could cause pharmacological effects, (2) the large magnitude of the experimental error in the Herbig et al. study at this glycine concentration.

Table 1.1: Steady states levels for five different values of glycine (folate cycle in % of total folate and the homocysteine remethylation cycle in µM). Where possible, a trend arrow is provided on the right to show the experimental outcome observed in [110]. All the trends are consistent with literature (green arrows) except for the total % of 5mTHF in the glycine range 5-10 mM (red arrow).

Glycine	**THF**	**10fTHF**	**CHF**	**CH2F**	**DHF**	**5mTHF**		
						free	**bound**	**total**
0 mM	0.19	36.17	7.25	2.17	0.04	10.31	48.87	54.18
1 mM	0.24	38.25	7.55	2.02	0.04	8.78	42.90	51.68
2 mM	0.74	39.37	7.70	1.91	0.04	7.97	42.26	50.23
5 mM	1.86	40.80	7.78	1.71	0.03	6.71	41.01	47.73
10 mM	5.32	40.40	7.71	1.52	0.03	5.56	39.48	45.04
Glycine	**HCY**	**MET**	**SAM**	**SAH**				
0 mM	2.72	30.58	99.99	17.23				
1 mM	3.44	42.09	54.44	50.55				
2 mM	3.48	42.48	50.96	53.60				
5 mM	3.52	42.83	46.49	57.67				
10 mM	3.57	43.04	42.40	61.51				

1.3 Results

1.3.1 The effect of folate status and the MTHFR polymorphism on pathways affecting NTD risk

In the current model, MTHFR activity was decreased to model the effect of the MTHFR C677T polymorphism. In addition, a two-fold increase in MTHFR activity was modeled to examine whether there was a dose-response relationship between MTHFR activity and various readouts of the network. This model shows that 5mTHF levels decrease as MTHFR activity decreases, reflecting the effects of the *MTHFR* C677T polymorphism (Table 1.2). The current model also recapitulates the biological observation that decreased MTHFR activity results in accumulation of 10fTHF and THF (Table 1.2). Inclusion of 10fTHF in the folate distribution is a strength of the current model in that it allows for estimation of the effect that perturbations to the system have on accumulation of this unstable form of folate, which likely accounts for the decreased folate status linked to NTDs in carriers of the polymorphism [244;257].

Table 1.2: Steady state distribution of folate (in percentage of total folate) for different levels of MTHFR activity (ranging from 2x to 0.3x of wild type). CC, CT and TT refer to the C677T polymorphism.

	THF	**10fTHF**	**CHF**	**CH2F**	**DHF**	**5mTHF**		
						free	**bound**	**total**
MTHFR × 2	0.25	14.43	2.86	0.79	0.01	33.45	48.20	81.65
MTHFR × 1 (CC)	0.70	39.24	7.69	1.93	0.04	8.07	42.34	50.42
MTHFR × 0.7 (CT)	1.59	48.76	9.45	2.18	0.04	3.37	34.61	37.98
MTHFR × 0.5	3.13	54.53	10.51	2.32	0.04	1.86	27.60	29.46
MTHFR × 0.3 (TT)	8.12	58.72	11.26	2.40	0.04	0.90	18.56	19.46

SAM and SAH levels vary markedly with changes in MTHFR activity (SAM levels decrease and SAH levels increase by more than 50% when comparing the "CC" model to the "TT" model), with a SAM/SAH ratio around one being achieved in CC homozygotes compared to 0.24 in TT homozygotes (Table 1.3). Although methylation potential, otherwise known as the SAM/SAH ratio, changes markedly with varying MTHFR activity, homocysteine concentrations appear relatively insensitive to changes in MTHFR in this model (Table 1.3), inconsistent with known effects of the *MTHFR* TT genotype in elevating serum homocysteine levels [149]. However, in this model the lack of elevation in homocysteine due to the *MTHFR* C677T polymorphism reflects that the model represents a closed system leading to intracellular conversion of cellular homocysteine to SAH as opposed to export of homocysteine into the circulation (Table 1.3).

The activity of each enzyme in the FOCM network as predicted by the current model indicates that accumulation of 10fTHF resulting from decreased MTHFR activity is due to two factors: (1) an increased flux through the 10fTHF synthetase activity leading to increased synthesis (FTS), and (2) an increased flux through the cyclohydrolase/dehydrogenase activity of MTHFD1 which converts CH2F to CHF to 10fTHF (MTCH, MTD activities, respectively, Table 1.4). Interestingly, the accumulation of 10fTHF does only mildly affect the flux through the enzymes that use 10fTHF as a co-factor for *de novo* purine synthesis (PGT, AICART; Table 1.4). This is consistent

Table 1.3: Steady state concentrations of model variables (in µM) for different levels of MTHFR activity (ranging from 2x to 0.3x of wild type). CC, CT and TT refer to the C667T polymorphism.

	THF	**10fTHF**	**CHF**	**CH2F**	**DHF**	**5mTHF**		
						free	**bound**	**total**
MTHFR × 2	0.04	2.58	0.51	0.14	0.00	5.97	8.60	14.57
MTHFR × 1 (CC)	0.12	7.00	1.37	0.34	0.01	1.44	7.56	9.00
MTHFR × 0.7 (CT)	0.28	8.70	1.69	0.39	0.01	0.60	6.18	6.78
MTHFR × 0.5	0.56	9.73	1.88	0.41	0.01	0.33	4.93	5.26
MTHFR × 0.3 (TT)	1.45	10.48	2.01	0.43	0.01	0.16	3.31	3.47
	SHMT	**HCY**	**MET**	**SAM**	**SAH**			
MTHFR × 2	0.40	3.10	37.99	78.41	31.01			
MTHFR × 1 (CC)	1.44	3.47	42.44	51.34	53.26			
MTHFR × 0.7 (CT)	2.82	3.67	43.00	34.87	68.98			
MTHFR × 0.5	4.07	3.78	42.59	26.72	77.42			
MTHFR × 0.3 (TT)	5.69	3.89	41.87	20.47	84.29			

with empirical experimental findings that increasing cellular levels of 10-formytetrahydrofolate dehydrogenase, which consumes 10fTHF, do not affect *de novo* purine synthesis [15]. Flux through the *de novo* dTMP synthesis pathway increases with decreasing MTHFR activity, consistent with empirical studies indicating these two pathways compete for CH2F (Table 1.4 , columns DHFR, TYMS and SHMT) [259]. These findings are in agreement with empirical data showing that the TT polymorphism results in an increase in CH2F available for dTMP synthesis as indicated by isotope tracer studies in humans [207].

5mTHF binds to and is a potent inhibitor of SHMT and glycine N-methyltransferase (GNMT). Decreased MTHFR activity, which lowers cellular 5mTHF levels (Table 1.3), increases the flux through the SHMT-catalyzed reaction in the direction of serine catabolism to glycine (Table 1.4). Methionine synthase (MTR) flux is highly sensitive to MTHFR genotype reflecting its dependence on availability of its substrate 5mTHF that is generated by MTHFR (Table 1.4).

The *MTHFR* C677T variant affects both CH2F partitioning (between homocysteine remethylation and dTMP biosynthesis) and intracellular folate concentrations; the 677T variant lowers intracellular folate levels. Therefore, the impact of this variant on FOCM was modeled at two different levels of folate (folate replete conditions, 18 µM and low-folate conditions, 9 µM) (Table 1.5,1.6, and 1.7) to understand how the variants function within the FOCM network as a function of folate cofactor availability. The results demonstrate that the changes in the percentage of 5mTHF (as well as other major one-carbon forms of folate) are more pronounced in the TT genotype than in the CC genotype when cellular folate levels are decreased (Table 1.5). The percentage of 5mTHF in CC homozygous does not change in the folate replete and deficiency states, whereas the accumulation of 5mTHF in TT homozygotes differs between the deficient and replete states (Table 1.5).

Fluxes through FOCM pathways are affected by both the *MTHFR* C677T polymorphism and folate levels. The most sensitive pathways to folate deficiency are the MTCH and MTD activities of MTHFD1, MTHFR, MTR, DNMT, DHFR, and TYMS (Table 1.7, row showing absolute flux

Table 1.4: Steady state fluxes of the model reactions (in µM/h) for different levels of MTHFR activity (ranging from 2x to 0.3x of wild type). CC, CT and TT refer to the C667T polymorphism. Reactions are indicated by the enzyme, which catalyzes them. For bidirectional reactions the direction is indicated behind the enzyme name (Enzyme: Substrate → Product).

	FTS	**MTCH**		**MTD**		**MTHFR**	**MTR**	
		10fTHF → CHF	**CHF → 10fTHF**	**CHF → CH2F**	**CH2F →CHF**			
MTHFR × 2	13,101.4	332,703.8	330,313.4	37,733.2	35,342.8	26.8	26.8	
MTHFR × 1 (CC)	23,582.5	756,270.3	744,706.4	89,938.7	78,374.9	21.6	21.6	
MTHFR × 0.7 (CT)	31,430.7	884,173.6	864,936.0	106,222.3	86,984.7	16.0	16.0	
MTHFR × 0.5	36,054.6	954,589.4	930,806.0	115,378.8	91,595.3	11.7	11.7	
MTHFR × 0.3 (TT)	39,755.7	1,002,680.1	975,243.2	121,653.1	94,216.2	7.1	7.1	
	SHMT		**SHMT & 5mTHF**		**DHFR**	**TYMS**	**PGT**	**AICARFT**
	CH2F →THF	**THF →CH2F**	**binding**	**unbinding**				
MTHFR × 2	2819.8	569.7	17,035.2	17,035.2	113.5	113.5	4,406.3	6,304.6
MTHFR × 1 (CC)	15,646.6	4,367.7	14,964.5	14,964.5	263.4	263.4	5,268.8	6,749.8
MTHFR × 0.7 (CT)	31,989.8	13,063.2	12,231.1	12,231.1	295.0	295.0	5,388.7	6,804.3
MTHFR × 0.5	47,127.3	23,667.8	9,754.6	9,754.6	312.3	312.3	5,442.7	6,828.4
MTHFR × 0.3 (TT)	66,523.2	39,415.5	6,559.7	6,559.7	322.1	322.1	5,475.7	6,843.0
	BHMT	**GNMT**	**DNMT**	**SAHH**		**MAT-I**	**MAT-III**	
				SAH → HCY	**HCY → SAH**			
MTHFR × 2	145.9	119.6	127.4	264.6	91.9	110.3	62.4	
MTHFR × 1 (CC)	161.0	370.9	87.2	285.2	102.6	123.0	59.6	
MTHFR × 0.7 (CT)	168.3	593.3	59.6	292.4	108.2	127.9	56.3	
MTHFR × 0.5	172.1	723.1	45.6	295.2	111.5	129.3	54.4	
MTHFR × 0.3 (TT)	175.5	830.4	34.7	297.1	114.5	129.8	52.7	

Table 1.5: Steady state distribution of folate (in percentage of total folate) for replete (18 µM) and low (9 µM) levels of total folate and for the CC and TT case of the C667T MTHFR polymorphism.

		THF	10fTHF	CHF	CH2F	DHF	5mTHF		
							free	bound	total
low folate	**CC**	1.27	38.48	7.46	1.65	0.03	4.98	46.13	51.11
	TT	4.24	49.75	9.45	1.68	0.03	0.93	33.91	34.84
replete folate	**CC**	0.70	39.24	7.69	1.93	0.04	8.07	42.34	50.42
	TT	8.12	58.72	11.26	2.40	0.04	0.90	18.56	19.46

Table 1.6: Steady state concentrations of model variables (in µM) for replete (18 µM) and low (9 µM) levels of total folate and for the CC and TT case of the C667T MTHFR polymorphism.

		THF	10fTHF	CHF	CH2F	DHF	5mTHF		
							free	bound	total
low folate	**CC**	0.11	3.43	0.67	0.15	0.00	0.44	4.12	4.56
	TT	0.38	4.44	0.84	0.15	0.00	0.08	3.03	3.11
replete folate	**CC**	0.12	7.00	1.37	0.34	0.01	1.44	7.56	9.00
	TT	1.45	10.48	2.01	0.43	0.01	0.16	3.31	3.47
		SHMT	**HCY**	**MET**	**SAM**	**SAH**			
low folate	**CC**	2.55	3.73	42.84	30.34	73.61			
	TT	10.00	3.95	41.35	17.43	87.79			
replete folate	**CC**	1.44	3.47	42.44	51.34	53.26			
	TT	5.69	3.89	41.87	20.47	84.29			

differences between folate levels). Flux through the dTMP synthesis pathway (DHFR and TYMS) is highly sensitive to folate status for both the MTHFR CC and TT genotypes (Table 1.7), with the TT homozygotes being the most sensitive. Flux through GNMT was also highly sensitive to folate status in the CC homozygotes, whereas GNMT flux in TT homozygotes was insensitive to folate status (Table 1.7). Similar but less pronounced effects were seen for flux through SHMT (Table 1.7).

To understand if *MTHFR* genotype affects the stability of the FOCM network at steady state, the deterministic simulation was coupled with stochastic simulation using a hybrid simulation strategy (see Materials and Methods). Model steady states were obtained under four different conditions that differed by *MTHFR* 677 genotype (CC and TT case) and intracellular folate levels (replete and low). Interestingly, the most stable steady state (the one with lowest total sum of reaction propensities $a_0(x)$, see Methods for details), was the CC case with folate replete concentrations, consistent with numerous epidemiological studies associating the *MTHFR* C677T genotype with folate-related pathologies (Table 1.8) [258]. The enzyme that exhibited the greatest level of stochasticity in response to folate levels and/or the *MTHFR* C677T polymorphism was SHMT1 (Appendix A.1, Table A.5).

Table 1.7: Steady state fluxes of the model reactions (in µM/h) for replete (18 µM) and low (9 µM) levels of total folate and for the CC and TT case of the C677T MTHFR polymorphism. Reactions are indicated by the enzyme, which catalyzes them. For bidirectional reactions the direction is indicated behind the enzyme name (Enzyme: Substrate → Product).

		FTS	MTCH		MTD		MTHFR	MTR	PGT	AICARFT	DHFR	TYMS
			10fTHF → CHF	CHF → 10fTHF	CHF → CH2F	CH2F →CHF						
CC	**replete folate**	23,582.5	756,270.3	744,706.4	89,938.7	78,374.9	21.6	21.6	5,268.8	6,749.8	263.4	263.4
	low folate	22,583.6	427,359.0	415,960.1	48,054.6	36,655.7	13.8	13.8	4,711.3	6,473.4	117.8	117.8
absolute difference (% of replete folate)		4.2	43.5	44.1	46.6	53.2	36.5	36.5	10.6	4.1	55.3	55.3
TT	**replete folate**	39,755.7	1,002,680.1	975,243.2	121,653.1	94,216.2	7.1	7.1	5,475.7	6,843.0	322.1	322.1
	low folate	33,618.2	529,771.1	507,690.2	59,371.2	37,290.2	4.2	4.2	4,943.9	6,593.4	120.0	120.0
absolute difference (% of replete folate)		15.4	47.2	47.9	51.2	60.4	41.2	41.2	9.7	3.6	62.8	62.8

		SHMT		SHMT & 5mTHF		BHMT	MAT-I	MAT-III	GNMT	DNMT	SAHH	
		CH2F →THF	THF →CH2F	binding	unbinding						SAH → HCY	HCY → SAH
CC	**replete folate**	15,646.6	4,367.7	14,964.5	14,964.5	161.0	123.0	59.6	370.9	87.2	285.2	102.6
	low folate	18,535.4	7,268.2	8150.7	8150.7	170.3	128.8	55.3	664.0	51.8	294.0	110.0
absolute difference (% of replete folate)		18.5	66.4	45.5	45.5	5.8	4.7	7.3	79.0	40.5	3.1	7.2
TT	**replete folate**	66,523.2	39,415.5	6,559.7	6,559.7	175.5	129.9	52.7	830.4	34.7	297.1	114.5
	low folate	73,586.0	51,629.1	5,992.1	5,992.1	177.4	129.8	51.8	884.4	29.4	297.9	116.3
absolute difference (% of replete folate)		10.6	31.0	8.7	8.7	1.1	0.0	1.8	6.5	15.2	0.3	1.6

Table 1.8: Total propensities obtained in four steady state conditions according to MTHFR polymorphism (CC and TT case) and total concentration of available folate (replete, 18 µM; low, 9 µM). To help comparisons, the differences between CC and TT (in % of CC) and between replete and low total folate (in % of replete folate) are indicated. A steady state that is less stable (or more noisy) than another one has higher total propensity.

$\mathbf{a_0(x)}$	**replete folate**	**low folate**	**Difference (% of replete folate)**
CC	$1.61 \cdot 10^{15}$	$2.10 \cdot 10^{15}$	30.62
TT	$8.59 \cdot 10^{15}$	$1.10 \cdot 10^{16}$	18.03
Difference (% of CC)	435.15	383.59	

1.3.2 The molecular basis of uracil misincorporation into DNA

Mouse models implicate SHMT and impaired *de novo* dTMP synthesis in NTD risk. Impaired *de novo* dTMP synthesis causes an increase in dUMP, which when converted to dUTP causes uracil misincorporation into DNA because DNA polymerases do not distinguish between dTTP and dUTP [29]. The dTMP biosynthesis pathway enzymes (MTHFD1, SHMT, TYMS, and DHFR) are present in both the cytosol and recently have been found to function in the nucleus. In the nucleus, they comprise a multi-enzyme complex at sites of DNA synthesis that may be critical to limit rates of uracil misincorporation into DNA, but regulatory mechanisms remain unknown [158]. These enzymes are modified by the Small Ubiquitin-like MOdifier (SUMO) protein at the G1/S boundary, which permits their nuclear translocation during S-phase of the cell cycle [13]. One study showed that when nuclear translocation of this complex is impaired in a mouse model over-expressing SHMT1, rates of uracil misincorporation into DNA increased several fold [158]. In this model, SHMT protein levels were elevated several fold in the liver, yet its localization was restricted to the cytoplasm and nuclear SHMT levels were depleted compared to wild-type mice [158]. Furthermore, nuclei isolated from SHMT overexpressing mice exhibited lower rates of *de novo* dTMP synthesis compared to nuclei isolated from wild-type mice [158]. This suggests that *de novo* dTMP synthesis occurs when the enzymes are present in the multi-enzyme complex within the nucleus in mammals. However, no definitive experiment has been performed that identifies the relative contribution of nuclear and cytosolic dTMP synthesis to overall dTMP synthesis. Interestingly, S. cerevesiae do not import the dTMP synthesis pathway into the nucleus [202].

To determine if nuclear import of the *de novo* dTMP pathway was required to meet cellular demands for dTTP during DNA replication, rates of dTMP synthesis were modeled for mammalian cells using standard Michaelis-Menten kinetics (Table 1.9, Table 1.4). Based on the number of A-T base pairs in the human genome and an 8-hour S-phase in embryonic stem cells (S-phase in L1210 cells is also 6-10 h [179;287]), the rate of dTMP synthesis required for faithful cell replication is 7.8 µM/min (Table 1.9, Appendix A, Section A.2) [26;221]. In the current model, which does not account for SHMT1/TYMS/DHFR/MTHFD1 nuclear localization nor complex formation, cytosolic dTMP synthesis rates are 4.4 µM/min (Table 1.4, DHFR and TYMS flux, 263.4 µM/h, assuming MTHFR 677CC genotype). This computational deficit between dTMP requirements and dTMP synthesis rates suggests that dTMP synthesis as currently modeled in the cytosol

Table 1.9: Cellular capacity for *de novo* dTMP synthesis in mammals and yeast at S-phase.

	Human	***S. cerevesiae***
Genome size	$3.0 \cdot 10^9$ bp	$1.2 \cdot 10^7$ bp
% AT	59 %	61.5 %
T bases needed for replication	$1.77 \cdot 10^9$ molecules	$7.5 \cdot 10^6$ molecules
Length of cell cycle	24 h	2.5 h
Length of S-phase	8 h	0.83 h
Cell volume	$8 \cdot 10^{-13}$ L (ES cell)	$5 \cdot 10^{-14}$ L
dTMP synthesis rate required to replicate genome	7.8 µM/min	0.5 µM/min
Measured dTMP synthesis rate (model outcomes)	4.4 µM/min	1.8 µM/min
Ratio of dTMP production relative to dTMP required for replication	0.6	3.6

where the enzymes are not present in a complex cannot meet cellular needs. Nuclear localization and complex formation of the *de novo* dTMP synthesis complex seem to be unique to mammalian cells. In S. cerevesiae, TYMS is not SUMOylated and localizes to the nuclear periphery [202]. The measured rate of dTMP synthesis in S. cerevesiae is 1.8 µM/min [100] (Table 1.9). The rate of dTMP synthesis required to replicate the S. Cerevesiae genome over the course of an S-phase (less than one hour [30;100]) is 0.5 µM/min, indicating that yeast synthesize dTMP at a rate that is more than 3-fold greater than necessary for adequate dTMP synthesis (Appendix A, Section A.2). Furthermore, in response to DNA damage, yeast increase dNTP concentrations 6-8 fold [37] and E. coli increase dNTP concentrations 1.8-3.7 fold [92], but dNTP concentrations do not increase after DNA damage in mammals [105;184].

1.4 Discussion

Understanding the dynamics of FOCM and its responsiveness to both genetic and environmental perturbations is the key to understanding the etiology of folate-related pathologies. Computational models and related simulations permit an identification of the most sensitive reactions within the network that exhibit the greatest degree of stochastic behavior leading to variability in network outputs. Furthermore, computational models allow an understanding of how both genetics and environmental factors can enhance or repress stochastic behavior at defined locations within the network, accelerating the development of diagnostics to identify those at risk for folate-related pathologies as well as lead to the development of targeted nutritional interventions for disease prevention.

The Shmt1 knockout mouse model (Shmt1+/-, Shmt1-/- embryos) exhibits impaired *de novo* dTMP synthesis in the absence of perturbations of homocysteine remethylation. It also recapitulates risk for NTDs in humans. Specifically, the mouse model exhibits folate-responsive NTDs that occur with minimal perturbation in FOCM, and exhibit low and variable penetrance [24;162]. In fact, most if not all, folate-related pathologies whose etiology involves interactions among genetic risk variants and nutrient exposures also exhibit low and/or variable clinical presentation. Under-

standing the stochastic behavior of the various reactions within FOCM that results in increased variability in FOCM network outputs is essential to understand which enzymes in the network contribute to folate-related pathologies.

Existing FOCM models rely on the limited quantity of kinetic data present in the literature, and the performance of the model will be dependent upon the kinetic parameters chosen to include in the model. Much of the available kinetic data for FOCM enzymes present in the literature was collected using the commercially available monoglutamate folate substrates, with few studies using the physiologically relevant polyglutamate forms of the cofactor. In the cell, newly transported monoglutamate folates are converted to folate polyglutamates, containing 3 to 7 polyglutamate moieties, though the action of folylpolyglutamate synthetase [243]. The polyglutamate chain (N=3 glutamate and higher) increases the affinity of folate cofactors for many folate-dependent enzymes by one to two orders of magnitude [112;225]. Models that include kinetic parameters derived from the use of folate monoglutamates can limit model reliability. Here we established a hierarchy of criteria to select a more homogeneous set of kinetic parameters (i.e. Km and Vmax) by referring, when possible, to L1210 cells because of the richness and quality of the data used to derive kinetic parameters. Furthermore, our preference was to select kinetic coefficients generated using folate polyglutamate cofactors and purified proteins from animal models closest to humans, as the variability in kinetic parameters among mammals is much less than the differences observed between folate monoglutamate and polyglutamate cofactor substrates.

The current model was validated by demonstrating that it recapitulates empirical observations regarding the impact of intracellular glycine on behavior of the FOCM network (Table 1.1). The validated model was then used to understand how the *MTHFR* C677T polymorphism, a known genetic risk factors for NTDs in humans, affects FOCM. This model shows that the lower levels of 5mTHF associated with the *MTHFR* 677T variant are accompanied by elevated levels of 10fTHF, which has been observed in animal models and in humans [21;58;89](Table 1.2). The model also indicates that 10fTHF accumulates in TT homozygotes as a result of increased flux through both the synthetase activity of MTHFD1 (FTS activity, Table 1.4), but also due to increased flux through MTHFD1 activity in the direction converting CH2F to 10fTHF (Table 1.4). Therefore, the model accurately predicts perturbations in FOCM that have been observed in human clinical and epidemiological studies. A recent study suggested that the risk of the *MTHFR* C677T polymorphism for NTDs was due to its known effect on lowering intracellular folate concentrations, rather than its role in providing 5mTHF for homocysteine remethylation [257]. This model demonstrates that the *MTHFR* C677T polymorphism elevates levels of 10fTHF, which is known to be a chemically unstable form of folate that is susceptible to oxidative degradation, providing a mechanism by which the *MTHFR* C677T polymorphism depletes intracellular folate levels. Importantly, the model reported here demonstrates that the *de novo* dTMP biosynthesis enzymes are the most sensitive to low intracellular folate concentrations, with both DHFR and TYMS activities being repressed by 63% (Table 1.7). This finding is consistent with the finding that mouse models with impaired *de novo* dTMP biosynthesis are susceptible to NTDs in folate deficiency [24]. The hybrid stochastic simulation also reveals that both folate deficiency and the *MTHFR* C677T polymorphism create overall instability in the network (Table 1.8), consistent with a vast body of literature demonstrating an association of both folate deficiency and the *MTHFR* C677T polymorphism with various pathologies [226;244]. Interestingly, SHMT exhibits the greatest increase in

stochastic behavior as a result of the *MTHFR* C677T polymorphism (Appendix A.1, Table A.5); the SHMT enzyme is the only FOCM enzyme that when disrupted results in folate-responsive NTDs[23].

The primary findings of this study are that the FOCM network is destabilized as a result of folate deficiency and the *MTHFR* C677T polymorphism, and that SHMT is the most sensitive enzyme within the network to this network instability. This finding nicely connects the MTHFR genetic variant, a known risk factor for human NTDs, and SHMT, the only folate enzyme whose disruption results in folate-responsive NTDs in mice. Furthermore, this model predicts that *de novo* dTMP synthesis rates in mammals are about half of what is required to meet DNA replication demands for dTMP (Table 1.9). Although mammals contain two pathways for dTMP synthesis, the folate-dependent *de novo* dTMP synthesis pathway described here and a salvage pathway catalyzed by thymidine kinase 1, the salvage pathway activity is insufficient to meet cellular needs based on observations that folate deficiency results in elevated uracil accumulation in DNA. In mammalian cells, the *de novo* dTMP synthesis enzymes form a multi-enzyme complex that interacts with DNA replication enzymes[13]. The discrepancy between *de novo* dTMP synthesis rates required to replicate the genome and the rate of dTMP synthesis currently predicted by the model indicates that the model should be extended to include multi-enzyme complex formation and substrate channelling in the nucleus to model more accurately determinants of FOCM and dTMP synthesis. The inclusion of the dTMP multi-enzyme metabolic complex in the model is expected to limit substrate diffusion and increase the rate of dTMP synthesis.

Author Contributions

The content of this chapter has been published in Scientific Reports [168].
K. Misselbeck (K.M.), L. Marchetti (L.M.), M. Scotti (M.S.), C. Priami (C.P.), M.S. Field (M.S.F.), P.J. Stover (P.J.S.) contributed to the experimental design. K.M., L.M., M.S. developed the computation model, while M.S.F., P.J.S., and M.S. provided all biological insight and extracted kinetic variables from the literature. M.S.F., P.J.S., and C.P. provided overall guidance of the project. All authors contributed to manuscript preparation and approved the final version.

Chapter 2

Effects of vitamin B_{12} deficiency on *de novo* dTMP synthesis

Folic acid intake in the periconceptional period reduces neural tube defect (NTD) occurrence by up to 70%, but not all NTDs are folate-responsive and thus further risk factors need to be identified. Observational studies demonstrate an association between maternal vitamin B_{12} status and NTD risk in both folate-fortified and non-fortified populations. Vitamin B_{12} is an essential co-factor of the folate-dependent enzyme methionine synthase. In this chapter we present a *in silico* study to explore the effect of vitamin B_{12} deficiency on the FOCM network in the cytoplasm. Simulations of the model introduced in Chapter 1 suggest that vitamin B_{12} deficiency leads to an accumulation of 5mTHF, which in turn disrupts the network functioning. Furthermore, model simulations indicate that *de novo* dTMP synthesis is considerably more sensitive to vitamin B_{12} deficiency than *de novo* purine synthesis.

2.1 Introduction

The connection between neural tube defects (NTDs), a class of severe malformations of the central nervous system occurring during early embryogenesis and leading to substantial mortatility, morbidity and long-term disability [28], and genes involved in the folate-mediated one-carbon metabolism as well as folate status is well recognized [8;99;172]. Indeed, folic acid supplementation at the beginning of pregnancy can decrease risk of NTDs and reduce the prevelence of these disorders [54;175]. However, not all NTDs are folate-responsive [111], and thus it is of interest to study which nutrition and environmental factors potentially play a role in the development and consequently in the treatment of NTDs.

The folate-mediated one-carbon metabolism (FOCM) network is sensitive to nutritional status of several vitamins that serve as essential enzyme cofactors, among which vitamin B_{12} (cobalamin) can be found. Vitamin B_{12} functions as a cofactor for the folate-dependent enzyme methionine synthase (MTR), which catalyses the conversion of homocysteine (HCY) to methionine(MET) and thereby regenerates THF for nucleotide synthesis (see Figure 2.1). Vitamin B_{12} is necassary for the maintanance of the normal function of the central nervous system. Diseases associated with B_{12} deficiency are reversible megaloblastic anemia, bone marrow failure, and demyelinating neurologic disease, while malabsorption following autoimmune gastritis is the most common cause

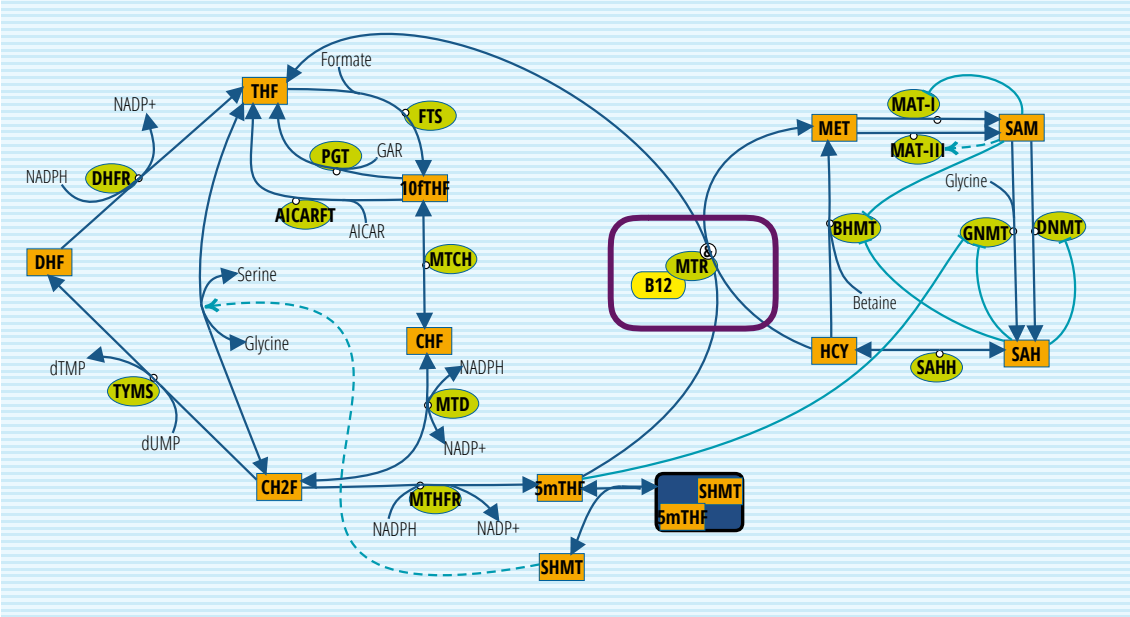

Figure 2.1: The reaction-based specification of the model according to the notation introduced in [94]. Rectangles identify model variables, non-boxed substrates are model constants, green circles identify enzymes, dark blue arcs identify matter transformation, and light blue arcs identify regulatory events (dotted lines indicate activations and solid lines indicate inhibitions). The purple box indicates the vitamin B_{12} dependend reaction catalyzed by MTR.

of vitamin B_{12} deficiency [236]. Various studies suggest that the maternal vitamin B_{12} status is also associated with the risk of NTD development [75;101;103;129;171;209;248;254;276;291]. Indeed, deficient vitamin B_{12} levels lead to impaired MTR activity, which in turn shifts the distribution of one-carbon folate forms towards 5-methylTHF (5mTHF) within the cell. [166;230;236]. Consequently, rates of *de novo* dTMP synthesis decrease and methylation reactions are impaired [166;230;236]. Evidence from epidemiologic and animal studies suggest a relation between folate-responsive NTDs and impaired dTMP biosynthesis [23;24;67;162].

To test the impact of vitamin B_{12} status on FOCM network dynamics with particular focus on dTMP production and to study propable underlying causative relations, the mathematical model introduced in Chapter 1 was used to run *in silico* experiments.

2.2 Materials and Methods

The mathematical model presented in Chapter 1 representing FOCM in the cytoplasm as depicted in Figures 1.1 and 2.1 was simulated to compare steady states and reaction velocities of the standard scenario and vitamin B_{12} deficient scenarios. The standard condition has been simulated by considering the model initial values and parameter estimates as described before (Appendix A.1, Tables A.1, A.2, A.3, and A.4). As vitamin B_{12} is an essential cofactor of methionine synthesis (MTR), we simulated vitamin B_{12} deficiency by decreasing MTR activity in a stepwise manner from the standard Vmax value of 30 µM/h to 10% of Vmax (3 µM/h). Two scenarios with MTR overexpression were also included, by considering the 2-fold and 5-fold increase of the standard

enzyme activity (60 and 150 µM/h). The effects of MTR activity on TYMS activity, and 5mTHF, methionine and homocysteine levels were simulated in the absence of BHMT expression, since BHMT is not present in most cells. All model simulations have been computed by implementing the mathematical model in MATLAB and by using the numerical ODE solver ode15s.

2.3 Results

To explore the effect of vitamin B_{12} deficiency on FOCM network dynamics, MTR activity was decreased step-wise and model outputs were compared across the selected scenarios. With decreasing MTR activity, the distribution of folate shifts towards the free and bound form of 5mTHF, reducing thereby the availability of other one-carbon substituted folate forms (Figure 2.2a, Appendix B, Tables B.2 and B.1). The formation of the so-called "5mTHF folate trap" [166;230;236] decreases the availability of nonmethylated functional folate substrates for FOCM enzymes other than MTR. Consequentially, in the state of increased 5mTHF levels, the fluxes through the reactions catalyzed by the trifunctional enzyme MTHFD1 (FTS, MTCH, MTD), and MTHFR (Figure 2.2b, Appendix B, Table B.3) are decreased, as well as the *de novo* dTMP synthesis activity (TYMS and DHFR, Figure 2.2c, Appendix B, Table B.3) and *de novo* purine synthesis activity (GART and AICART, Figure 2.2c, Appendix B, Table B.3). Interestingly, *de novo* dTMP synthesis was considerably more sensitive to vitamin B_{12} deficiency than was *de novo* purine synthesis, consistent with previous findings [193].

5mTHF is a tight-binding inhibitor of SHMT [239] and with decreased MTR activity the amount of 5mTHF binding to SHMT increased sharply (Figure 2.2d), resulting in decreased flux through SHMT (Figure 2.2d, Appendix B, Table B.3). The effects on SHMT were most pronounced at MTR activity levels in the direct surrounding of the standard Vmax (30 µM).

MET and HCY levels, as well as SAHH activity stayed relatively unaffected by vitamin B_{12} deficiency/reduced MTR activity (Figure 2.2e and f). This model also includes the vitamin B_{12} -independent conversion of HCY to MET, which is vatalyzed by betaine-homocysteine methyltransferase (BHMT). Inclusion of BHMT in the model and the fact that most intracellular HCY is converted to SAH may account for the lack of increase in HCY concentrations as a result of vitamin B_{12} deficiency in the model. It is also worth noting that free HCY is actively exported out of cells into the blood. SAM and SAH levels display a trade-off effect throughout the simulations (Figure 2.2f). A possible explanation for this behaviour is the decreased flux through GNMT (catalysing the transformation of SAM to SAH) caused by increased levels of 5mTHF, which is an inhibitor of GNMT. Like for SHMT, changes in SAM and SAH levels were most pronounced around the standard Vmax of MTR (30 µM).

Because most cells do not express BHMT, the simulation was repeated without BHMT activity. As anticipated, HCY and MET are more sensitive to MTR activity in the absence of BHMT (Figure 2.3, Appendix B, Table B.5). Importantly, TYMS activity, which is a proxy for dTMP synthesis, is more sensitive to MTR activity than 5mTHF levels (Figure 2.3, Appendix B, Tables B.4 and B.6), indicating that the inverse relationship between MTR and TYMS activity cannot be fully accounted for by the accumulation of 5mTHF, but that additional aspects of network dynamics are also operative.

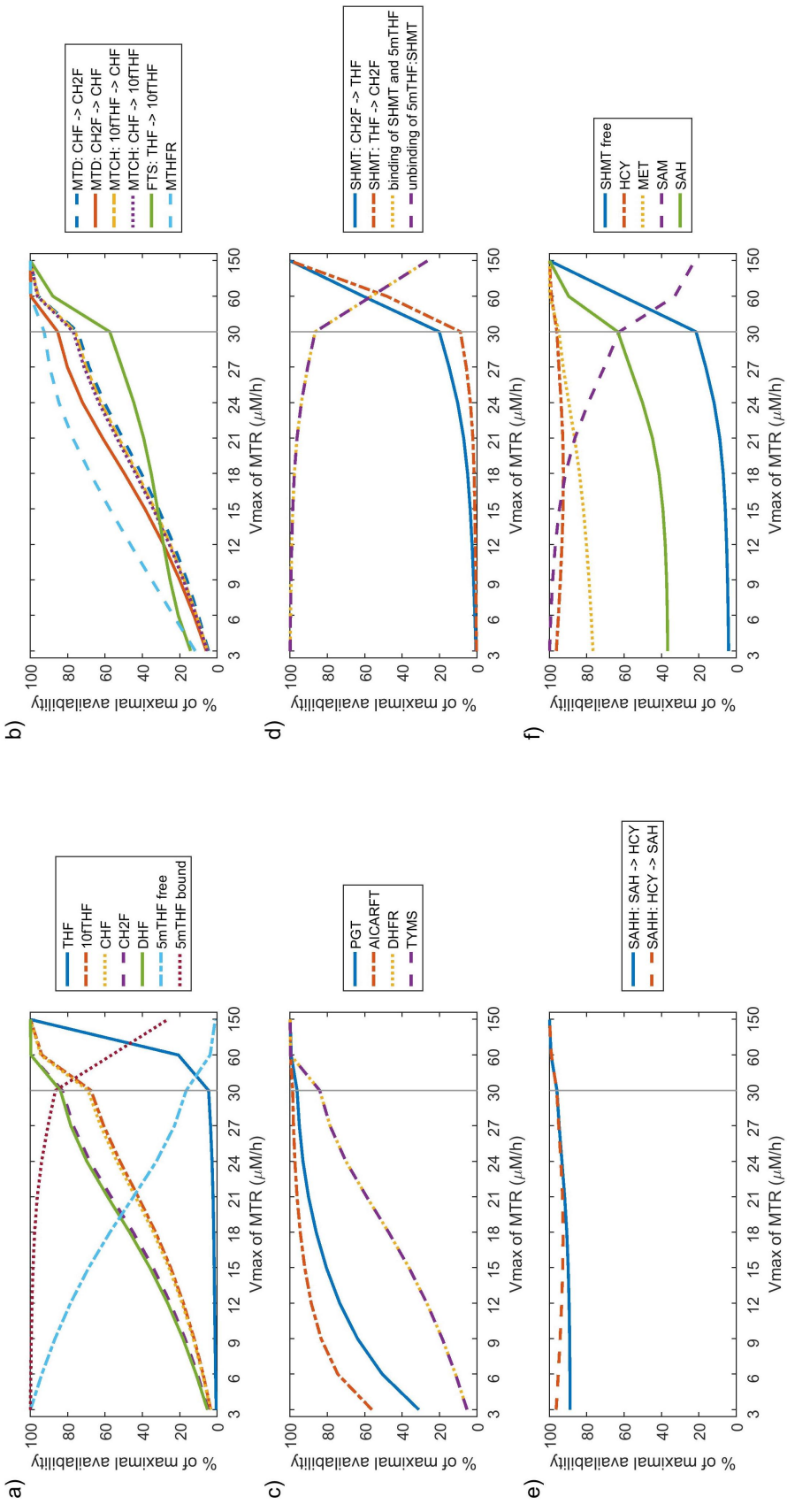

Figure 2.2: Effect of MTR activity on model variables (a, f) and fluxes (b–e). The provided values are computed by considering the model steady states achieved by varying the Vmax of MTR from 3 µM/h to 150 µM/hand are expressed in terms of percentage of the maximal activity over the considered range. The vertical gray line indicates the standard Vmax condition (30 µM/h); values lower than this threshold can be considered as scenarios with vitamin B_{12} deficiency.

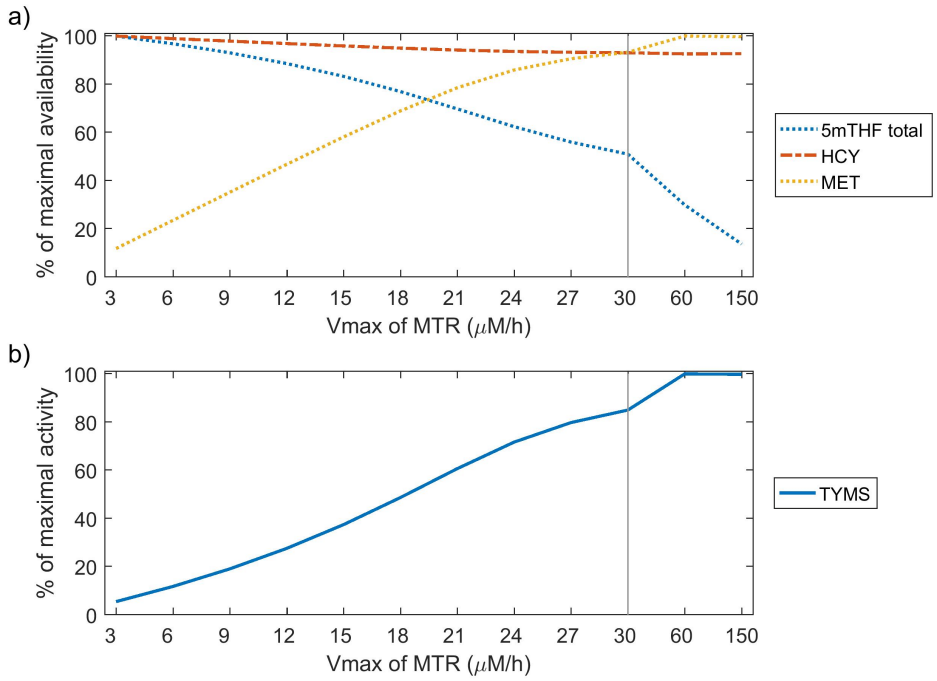

Figure 2.3: Effect of MTR activity on a) model variables and b) flux through TYMS for the model without BHMT activity. The provided values are computed by considering the model steady states by varying the V_{max} of MTR from 3 µM/h to 150 µM/h and are expressed in terms of percentage of the maximal availability/activity over the considered V_{max} range. The vertical gray line indicates the standard Vmax condition (30 µM/h); values lower than this threshold can be considered as scenarios with vitamin B_{12} deficiency.

2.4 Discussion

There is a need to identify additional risk factors for NTDs, as folic acid has not reduced the birth prevalence of NTDs in North America below an incidence of 7.5/10,000 in the post folic acid fortification era[28]. Maternal vitamin B_{12} deficiency has emerged as a candidate risk factor given its participation in folate metabolism as well as the findings from cross-sectional studies that demonstrate association between maternal deficiency and NTD risk[75;87;101;103;129;171;209;276;291]. Simulating vitamin B_{12} deficiency in the mathematical model of FOCM presented in Chapter 1 suggests that even relatively minor, physiologically relevant decreases in MTR activity resulting from vitamin B_{12} deficiency impact one-carbon distribution and flux through FOCM. With decreasing MTR activity, folate accumulates as 5mTHF derivative, which cannot be reintroduced to the folate pool via the MTHFR-catalyzed reaction, because this reaction is irreversible under physiological conditions[230]. The model further demonstrated that in metabolic systems, where BHMT is expressed, binding of 5mTHF to SHMT and inhibition of SHMT activity as well as *de novo* dTMP synthesis are considerably more affected by vitamin B_{12} deficiency than either *de novo* purine synthesis or homocysteine remethylation (Figure 2.2).

Author Contributions

The content of this chapter is part of a manuscript in preparation [194].
With respect to the computational results provided in this Chapter, K. Misselbeck (K.M.), L. Marchetti (L.M.), C. Priami (C.P.), M.S. Field (M.S.F.), and P.J. Stover (P.J.S.) contributed to the experimental design. K.M., carried out the computational simulations based on the model introduced in Chapter 1, while M.S.F., P.J.S., provided biological insight.

Chapter 3

An extended mathematical model of folate-mediated one-carbon metabolism: The role of the 5-formyltetrahydrofolate futile cycle

In folate-mediated one-carbon metabolism (FOCM), 5-formyltetrahydrofolate (5fTHF), a one-carbon substituted tetrahydrofolate (THF) vitamer, acts as an intracellular storage form of folate and as an inhibitor of the folate-dependent enzymes phosphoribosylaminoimidazole-carboxamide formyltransferase (AICARFT) and serine hydroxymethyltransferase (SHMT). Cellular levels of 5fTHF are regulated by a futile cycle comprising the enzymes SHMT and 5,10-methenyltetrahydrofolate synthetase (MTHFS). MTHFS is an essential gene in mice; however, the roles of both 5fTHF and MTHFS in mammalian FOCM remain to be fully elucidated. We present an extension of the hybrid-stochastic model of FOCM introduced in Chapter 1 by including the 5fTHF futile-cycle to explore its effect on the FOCM network. Simulations of this extended model indicate that MTHFS plays an essential role in preventing 5fTHF accumulation, which consequently averts inhibition of all other reactions in the metabolic network. Moreover, in silico experiments show that 10-formylTHF inhibition of MTHFS is critical for regulating purine synthesis. Model simulations also provide evidence that 5-methylTHF (and not 5fTHF) is the predominant physiological binder/inhibitor of SHMT. Finally, the model simulations indicate that the 5fTHF futile cycle dampens the stochastic noise in FOCM that results from both folate deficiency and a common variant in the methylenetetrahydrofolate reductase (*MTHFR*) gene.

3.1 Introduction

In the folate-mediated one-carbon metabolism (FOCM) tetrahydrofolates (THF) carry and chemically activate one-carbon moieties for biosynthetic reactions including *de novo* purine synthesis, *de novo* thymidylate (dTMP) synthesis, and remethylation of homocysteine to methionine [243].

There are five one-carbon substituted THF derivatives *in vivo*, and these derivatives carry one-carbon units at one of three oxidation levels ranging from formate to methanol, with each

substituted folate serving in unique one-carbon transfer reactions. 10-formylTHF (10fTHF), 5-formylTHF (5fTHF), and 5,10-methenylTHF (CHF) carry one-carbon units at the oxidation state of formate. The formyl group of 10fTHF, the folate co-factor used by phosphoribosylaminoimidazolecarboxamide formyltransferase (AICARFT) and phosphoribosylglycinamide formyltransferase (PGT), is incorporated into the #2 and #8 carbons of the purine ring. 10fTHF is formed from THF and ATP by the synthetase activity of the trifunctional enzyme methylenetetrahydrofolate dehydrogenase 1 (MTHFD1). 5fTHF is not used as a cofactor for folate-dependent biosynthetic reactions, rather it is thought to be an intracellular storage form of folate in dormant cells [239]. 5fTHF is regulated through a futile cycle catalyzed by serine hydroxymethyltransferase (SHMT) and methenyltetrahydrofolate synthetase (MTHFS). 5fTHF is generated from CHF in an irreversible reaction catalyzed by SHMT. 5fTHF is re-introduced to the folate cofactor pool by MTHFS, which converts 5fTHF to CHF in an irreversible, ATP-dependent reaction. CHF is not used directly as a co-factor for biosynthetic reactions. It is formed both enzymatically from MTHFS and the cyclohydrolase activity of MTHFD1 and non-enzymatically from both 5fTHF and 10fTHF [241].

MTHFS is an essential gene in mice [76], but its role in mammalian FOCM remains to be fully elucidated. Increased *MTHFS* expression in cultured cells shifts the distribution of folate cofactors toward 10fTHF at the expense of 5methylTHF (5mTHF) and also leads to increased rates of folate catabolism [16], presumably by shifting the folate distribution in favor of THF accumulation of more chemically unstable forms of folate. 10fTHF binds tightly to to and inhibits MTHFS *in vitro* [82;83], and increased *MTHFS* expression in cultured cells leads to increased rates of *de novo* purine biosynthesis [83]. *MTHFS* expression in cultured cells also decreases efficacy of anti-folate chemotherapeutic agents designed to target *de novo* purine synthesis [77]. MTHFS physically interacts with the "purinosome", a multi-enzyme complex that forms under purine-deficient conditions and consists of the six enzymes required for *de novo* purine synthesis. It has been suggested that MTHFS serves to channel 10fTHF cofactors to the purinosome [76].

The SHMT- and MTHFS-catalyzed "futile cycle" may serve regulatory functions by controlling 5fTHF concentrations. The primary metabolic function of SHMT is to reversibly interconvert serine and THF to glycine and 5,10-methyleneTHF (CH2F). 5fTHF is a feedback inhibitor of SHMT, and also binds to and inhibits AICARFT [27;91], but the purpose of the 5fTHF futile cycle in regulating SHMT and FOCM remains unresolved. This is due in part because 5mTHF, which is more abundant than 5fTHF, also serves as a potent inhibitor of SHMT.

The hybrid stochastic model of FOCM introduced in Chapter 1 showed that decreased 5mTHF binding to SHMT, as a result of an overall decrease in 5mTHF levels resulting from a methylenetetrahydrofolate reductase (*MTHFR*) polymorphism, led to increased flux through the reversible reactions catalyzed by SHMT and MTHFD1 (Chapter 1). The common *MTHFR* C677T polymorphism is known to lower total cellular MTHFR activity, leading to decreased 5mTHF production and altered one-carbon distribution [21;58;89;250]. Decreased 5mTHF levels also increase total reaction propensities, indicating a loss in overall FOCM network stability as a result of this common polymorphism. Here, we extended this hybrid stochastic model of FOCM to include the 5fTHF futile cycle according to Figure 3.1 in an effort to better understand: 1) the role of MTHFS and the 5fTHF futile cycle in FOCM, and 2) the relative contributions of 5fTHF and 5mTHF to SHMT activity and overall network stability.

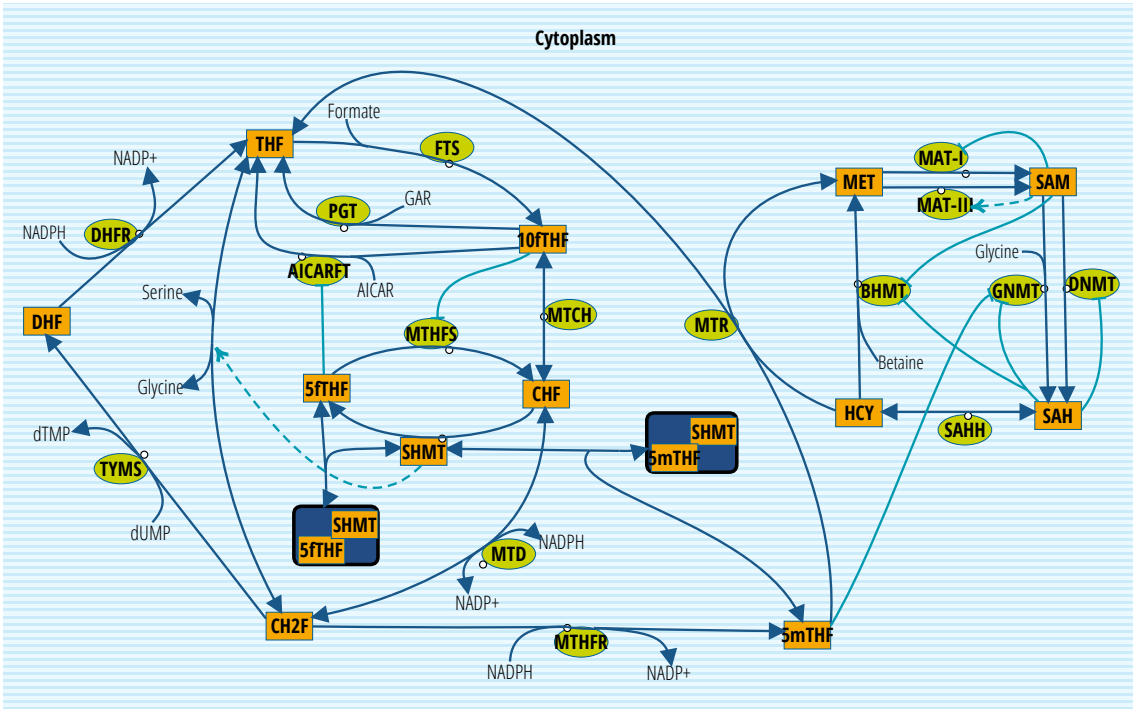

Figure 3.1: The reaction-based specification of the extended model according to the notation introduced in [94]. Rectangles identify model variables, non-boxed substrates are model constants, green circles identify enzymes, dark blue arcs identify matter transformation, and light blue arcs identify regulatory events (dotted lines indicate activations and solid lines indicate inhibitions).

3.2 Materials and Methods

3.2.1 Description of the model

The model presented in this paper is an extension of the mathematical model of FOCM introdced in Chapter 1. The model provides a description of FOCM in the cytoplasm including its regulation of key biological processes related to *de novo* dTMP synthesis, *de novo* purine synthesis and remethylation of homocysteine to methionine. With respect to the initial model, the model herein has been extended to include the folate form 5fTHF as well as the enzyme MTHFS and the relevant reactions involving these two molecules, according to the reaction network provided in Figure 3.1. The model is composed of 14 variables and 20 (reversible and irreversible) reactions, most of which have been parametrized by means of Michaelis-Menten kinetics with one or two substrates. Whenever possible, physiologically relevant forms of folate polyglutamate cofactors have been considered to derive kinetic coefficients, according to the modeling approach used in the original model.

MTHFS and 5fTHF have been included in the model using the following five reactions. In agreement with the FOCM modeling literature, ATP and ADP are not explicitly represented in reaction stoichiometry since they are assumed to be constantly present at their physiological level.

1. **Formation of 5fTHF catalyzed by MTHFS** (Michaelis-Menten kinetics with one sub-

strate)

$$R_{\mathrm{MTHFS}}: \quad 5\mathrm{fTHF} \xrightarrow{\mathrm{MTHFS}} \mathrm{CHF}$$

$$v_{\mathrm{MTHFS}}([5\mathrm{fTHF}]) = \frac{\mathrm{k_{cat}} \cdot \mathrm{MTHFS} \cdot [5\mathrm{fTHF}]}{(\mathrm{K_{5fTHF}} \cdot (1 + \frac{[10\mathrm{fTHF}]}{\mathrm{K_{i,10fTHF}}}) + [5\mathrm{fTHF}])}$$

The inhibition by 10fTHF has been encoded in the inhibition constant.

2. **Transformation of 5fTHF to CHF catalyzed by SHMT** (Michaelis-Menten kinetics with one substrate)

$$R_{\mathrm{SHMT}}: \quad \mathrm{CHF} \xrightarrow{\mathrm{SHMT}} 5\mathrm{fTHF}$$

$$v_{\mathrm{SHMT}}([\mathrm{CHF}]) = \frac{\mathrm{k_{cat}} \cdot \mathrm{SHMT} \cdot [\mathrm{CHF}]}{\mathrm{K_{CHF}} + \mathrm{CHF}}$$

3. **Tight binding of 5fTHF and SHMT** (mass-action kinetics)

$$\mathrm{SHMT} + 5\mathrm{mTHF} \xrightarrow{\mathrm{k_{binding}}} \mathrm{SHMT} : 5\mathrm{mTHF}$$

$$v_{\mathrm{binding}} = \mathrm{k_{binding}} \cdot [5\mathrm{mTHF}] \cdot [\mathrm{SHMT}]$$

4. **Unbinding of 5fTHF: SHMT** (mass-action kinetics)

$$\mathrm{SHMT} : 5\mathrm{mTHF} \xrightarrow{\mathrm{k_{unbinding}}} \mathrm{SHMT} + 5\mathrm{mTHF}$$

$$v_{\mathrm{unbinding}} = \mathrm{k_{unbinding}} \cdot [\mathrm{SHMT} : 5\mathrm{mTHF}]$$

5. **AICARFT mediated 10fTHF transformation to THF** (Michaelis-Menten kinetics with two substrates)

$$R_{\mathrm{AICARFT}}: \quad 10\mathrm{fTHF} \xrightarrow{\mathrm{AICARFT,AICAR}} \mathrm{THF}$$

$$v_{\mathrm{AICARFT}}([10\mathrm{fTHF}], \mathrm{AICAR}) = \frac{\mathrm{V_{max}} \cdot [10\mathrm{fTHF}] \cdot \mathrm{AICAR}}{(\mathrm{K_{10fTHF}} + [10\mathrm{fTHF}]) \cdot (\mathrm{K_{AICAR}} + \mathrm{AICAR})} \cdot \frac{1}{1 + \frac{[5\mathrm{fTHF}]}{\mathrm{K_{i,5fTHF}}}}$$

This reaction was already included in the initial model, but its rate formula has been extended to account for the inhibition by 5fTHF.

Regarding model initial values, the starting availability of 5fTHF has been set to 5% of total cytosolic folate, according to [91]. Furthermore, the initial value of SHMT has been refined with respect to the value provided in the original model to account for the fact that only two of the four enzyme sites are presumed to be active [250]. Because the additional components included in the model affect the behavior of the network, some of the reaction kinetics have been refined with respect to those previously provided to take into account the increased amount of information included in the extended network. All refined parameter estimates are listed in Appendix C, Table C.1.

3.2.2 Description of the simulation approach and computational environment

For the simulation of the model we followed the approach presented in Chapter 1 and we employed a hybrid-stochastic approach, for which the initial model and the extensions described above have first been translated to a set of ODEs and implemented in MATLAB. The initial part of the dynamics until reaching a model steady state was computed using deterministic simulation by means of the numerical ODE solver ode15s. By exploiting the hybrid stochastic framework, we then coupled the ODE-based model description with a stochastic reaction-based one following the same approach developed for the first version of the mathematical model introduced in Chapter 1. To assess the level of stochasticity of model steady states, we relied on the concept of total propensity, which is used to determine the next reaction event in stochastic simulation [90]. In more detail, the stochastic simulation algorithm computes a propensity function $a_j(x)$ for each model reaction R_j, where x is the current state of the system. The total propensity is then calculated as $a_0(x) = \sum_j a_j(x)$ and used to assess when the next reaction event will occur as is proportional to the number of reaction events occurring per unit of time.

3.2.3 *In silico* experiments

In the following a short techniqual description of the considered scenarios is provided.

The role of MTHFS

The influence of the addition of MTHFS on the FOCM network was studied by comparing steady states and reaction velocities of the standard scenario based on the parameters presented in Appendix C, Table C.1 with the following four scenarios: a) MTHFS activity scaled down to 50% of the standard case (0.04 µM), b) no availability of MTHFS (0 µM), c) no availability of MTHFS plus turning off of the reaction, d) no inhibition of MTHFS by 10fTHF.

Effect of glycine and MTHFS activity on 5mTHF binding to SHMT

Because 5fTHF inhibits SHMT activity, the interplay of glycine and MTHFS and their effect on steady state concentrations of 5mTHF (free, bound to SHMT and total) were compared for high and low levels of MTHFS (10% of standard activity and 5-fold increase in standard activity) as a function of glycine (considered levels: 1000 µM, 1850 µM, 2000 µM, 5000 µM and 10000 µM). These simulations were repeated for the case in which the binding of 5mTHF and SHMT was blocked, increasing therefore the availability of 5mTHF for other reactions.

Sensitivity analysis on MTHFS

The effect of MTHFS levels on the binding of 5fTHF and 5mTHF to SHMT was studied by comparing their steady state concentrations at different levels of MTHFS. Enzyme deficiency was simulated by decreasing MTHFS availability stepwise (90% to 10% of the standard activity, corresponding to 0.072 µM and 0.008 µM) whereas the 2-fold and 5-fold increase of the standard activity was considered in overexpression scenarios (corresponding to 0.16 µM and 0.4 µM).

Quantification of network stability

In the steady state comparisons herein presented, we exploited the concept of total propensity for comparing the level of stochasticity of the considered steady states as reported in section 3.2.2. The idea behind this approach is the following: intuitively, a steady state with lower total propensity can be interpreted as more stable because on average it will have a lower number of reaction events per unit of time that can perturb its equilibrium. We applied this intuition to compare different scenarios and assess the stability of the corresponding steady states. In particular, we investigated the effect of the common C677T MTHFR polymorphism, which reduces the enzymatic activity of MTHFR, and the effect of folate deficiency, to test the impact of these factors on the FOCM network with or without the 5fTHF futile cycle. Following Chapter 1, MTHFR activity was decreased to 30% of standard activity to model the effect of the MTHFR C677T polymorphism. We further considered two levels of folate availability: replete folate status (standard model parametrization) and low folate status (folate availability reduced to 50%). The resulting four scenarios were repeated for the standard parametrization of the model and for the one without the 5fTHF cycle (no availability of MTHFS and 5fTHF, and no activity of the SHMT-catalyzed reaction: CHF $\rightarrow$ 5fTHF). In all cases we used deterministic simulation to compute the steady states of the scenarios of interest and we then calculated the reaction propensities and the total propensities according to stochastic simulation.

3.3 Results

3.3.1 The role of MTHFS

The influence of the 5fTHF futile cycle, including the MTHFS-catalyzed synthesis of CHF from 5fTHF and the reverse reaction catalyzed by SHMT, on the FOCM network was studied by comparing steady states and reaction velocities of this updated model, which is described by the parameters presented in Table Appendix C, C.1. The outputs of the following four conditions were compared to describe the effects of the 5fTHF futile cycle: a) inclusion of the enzymatic reactions comprising the futile cycle (MTHFS and SHMT), b) MTHFS activity scaled down to 50% of the standard condition, c) no MTHFS activity or d) no activity of either MTHFS or its SHMT-catalyzed counterpart reaction R_{SHMT} : CHF $\rightarrow$ 5fTHF (elimination of the futile cycle).

In this model that includes the 5fTHF futile cycle, reducing MTHFS activity by 50% leads to increased 5fTHF levels, which come at the expense of 5mTHF levels, through 5mTHF remains the predominant form of intracellular folate under both conditions (Table 3.1 and 3.2, compare conditions A and B). Levels of other one-carbon substituted folate forms remain unchanged as a result of decreased MTHFS activity (Tables 3.1 and 3.2). 5fTHF is a known inhibitor of the folate-dependent enzymes AICARFT and SHMT [27]. We confirm that 5fTHF accumulation resulting of decreased MTHFS activity decreases the flux through the AICARFT-catalyzed reaction (Table 3.4). Flux through the other folate-dependent *de novo* purine synthesis enzyme, PGT, remains mostly unchanged when MTHFS activity is reduced by 50% (Table 3.4), which is also not unexpected because 5fTHF is not known to inhibit PGT. Overall, these results are consistent with the empirical observation that "cells with 50% reduced *Mthfs* expression have reduced *de novo* purine synthesis" [76]. Furthermore, the model also supports the empirical observation that

de novo dTMP synthesis is not affected in cells with reduced *Mthfs* expression, as flux through TYMS and DHFR is not affected by reduced MTHFS activity (Table 3.4) [76]. The 50% reduction in MTHFS activity moderately decreases fluxes through SHMT-catalyzed reactions (Table 3.3), and does not affect flux through the enzymes of the homocysteine remethylation cycle (Table 3.4).

The model also shows that MTHFS activity is necessary to prevent accumulation of cellular folate as 5fTHF. When FOCM was modeled without MTHFS activity, the distribution of folate shifted such that all the folate accumulated as 5fTHF (Table 3.1 and 3.2, comparing conditions A and C), leaving no folate co-factors available for *de novo* dTMP, purine synthesis (Table 3.4) or homocysteine remethylation (Table 3.3). Therefore, the model supports the experimental observation that *Mthfs* is an essential gene in mice [76] and suggests that *MTHFS* expression is necessary to prevent accumulation of 5fTHF and subsequent inhibition of FOCM.

If in addition to the deletion of MTHFS activity, the SHMT-catalyzed conversion of CHF to 5fTHF is also turned off thereby effectively modeling the absence of the 5fTHF futile cycle, the lethal pooling of folate as 5fTHF is prevented (Table 3.1 and 3.2, compare C to D). In addition, because there is no 5fTHF formed in the absence of the futile cycle, the steady state distribution of all remaining one-carbon substituted forms increases (Table 3.2, compare A to D). This does not appreciably affect the flux through other FOCM enzymes (MTHFD1, MTHFR, PGT, dihydrofolate reductase (DHFR), thymidylate synthase (TYMS), methionine synthase (MTR)) (Tables 3.3 and 3.4, compare A and D) or homocysteine remethylation enzymes (Table 3.4, compare A and D). As expected, the decrease in 5fTHF levels increases flux through AICARFT (Table 3.4, compare A and D) and SHMT (Table 3.3, compare A and D) due to loss of inhibition by 5fTHF. The model shows that for the latter, the increase in 5mTHF levels is compensating for the loss of 5fTHF inhibition. Indeed, the amount of SHMT bound by 5mTHF increases from 81% to 92% of total SHMT (Tables 3.1, compare A and D).

SAM and S-adenosylhomocysteine (SAH) levels are not affected by a 50% reduction in MTHFS activity (Table 3.1, compare conditions A and B). However, elimination of MTHFS activity causes SAM levels to decrease by more than 80% and SAH increases by almost 3-fold with respect to the standard condition. The SAM/SAH ratio, which is around 2.3 in the standard case, becomes 0.15 without MTHFS activity (Table 3.1, compare A and C). Even if 5mTHF is not available in scenario C, because of pooling all cofactors as 5fTHF, the remethylation cycle is active because the model includes the betaine-homocysteine transferase (BHMT)-catalyzed conversion of homocysteine to methionine. This reaction is folate-independent and is not directly affected by 5mTHF levels.

3.3.2 The role of 10fTHF inhibition of MTHFS

10fTHF is a tight-binding inhibitor of MTHFS [83] creating a mechanism for feedback inhibition of MTHFS such that 5fTHF levels are mobilized when 10fTHF levels are depleted. To understand the role of 10fTHF inhibition of MTHFS, the 10fTHF inhibition term was removed from the model (Tables 3.1, 3.2, 3.3, and 3.4, scenario E). Without inhibition of MTHFS by 10fTHF, increased MTHFS activity leads to depletion of 5fTHF (Tables 3.1 and 3.2). Interestingly, the steady-state distribution of folate forms and enzyme fluxes were nearly identical to those resulting from removal of the whole 5fTHF futile cycle (Tables 3.1, 3.2, 3.3, and 3.4, compare D and E). These data suggest that 10fTHF inhibition of MTHFS is necessary to maintain 5fTHF levels, which in

Table 3.1: Steady state concentrations of model variables (in µM) for different scenarios with respect to MTHFS and SHMT activity.

	THF	**10fTHF**	**CHF**	**CH2F**	**DHF**	**5mTHF**	
						free	**bound**
Standard MTHFS and regular SHMT activity (Condition A)	0.051	7.075	1.407	0.398	0.007	5.217	3.644
0.5x MTHFS and regular SHMT activity (Condition B)	0.047	6.984	1.390	0.394	0.007	4.972	3.249
0x MTHFS and regular SHMT activity (Condition C)	0.000	0.000	0.000	0.000	0.000	0.000	0.000
0x MTHFS and SHMT: CHF→5fTHF activity turned off (Condition D)	0.058	7.014	1.394	0.391	0.007	4.828	4.156
Standard MTHFS and regular SHMT activity without MTHFS inhibition by 10fTHF (Condition E)	0.056	7.180	1.428	0.403	0.007	5.517	4.194
	5fTHF		**SHMT**	**HCY**	**MET**	**SAM**	**SAH**
	free	**bound**	**free**				
Standard MTHFS and regular SHMT activity (Condition A)	0.413	0.577	0.279	3.206	38.116	76.071	33.127
0.5x MTHFS and regular SHMT activity (Condition B)	0.757	0.989	0.261	3.218	38.294	75.278	33.729
0x MTHFS and regular SHMT activity (Condition C)	14.351	4.438	0.062	4.039	40.573	14.010	91.898
0x MTHFS and SHMT: CHF→5fTHF activity turned off (Condition D)	0.000	0.000	0.344	3.226	38.402	74.790	34.102
Standard MTHFS and regular SHMT activity without MTHFS inhibition by 10fTHF (Condition E)	0.001	0.001	0.304	3.191	37.908	76.973	32.447

Table 3.2: Steady state distribution of folate (in percentage of total folate) for different scenarios with respect to MTHFS and SHMT activity.

	THF	10fTHF	CHF	CH2F	DHF	5mTHF			5fTHF		
						free	bound	total	free	bound	total
Standard MTHFS and regular SHMT activity (Condition A)	0.27	37.65	7.49	2.12	0.04	27.77	19.39	47.16	2.2	3.07	5.27
0.5x MTHFS and regular SHMT activity (Condition B)	0.25	37.17	7.40	2.10	0.04	26.46	17.29	43.75	4.03	5.27	9.29
0x MTHFS and regular SHMT activity (Condition C)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	76.38	23.62	100.00
0x MTHFS and SHMT: CHF→5fTHF activity turned off (Condition D)	0.33	39.30	7.81	2.19	0.04	27.05	23.28	50.33	0.00	0.00	0.00
Standard MTHFS and regular SHMT activity without MTHFS inhibition by 10fTHF (Condition E)	0.30	38.22	7.60	2.14	0.04	29.36	22.32	51.68	0.01	0.01	0.02

Table 3.3: Steady state fluxes of the reactions catalyzed by the enzymes FTS, MTCH, MTD, MTHFR, MTR, SHMT and the (un-)binding of 5mTHF/5fTHF and SHMT (in µM/h) for different scenarios with respect to MTHFS and SHMT activity.Reactions are indicated by the enzyme, which catalyzes them. For bidirectional reactions the direction is indicated behind the enzyme name (Enzyme: Substrate → Product).

	FTS	**MTCH**		**MTD**		**MTHFR**	**MTR**
		10fTHF → CHF	**CHF → 10fTHF**	**CHF → CH2F**	**CH2F →CHF**		
Standard MTHFS and regular SHMT activity (Condition A)	14281.6	761957.5	758886.0	91831.0	88759.5	23.0	23.0
0.5x MTHFS and regular SHMT activity (Condition B)	13563.0	754720.3	751813.8	90886.3	87979.8	22.9	22.9
0x MTHFS and regular SHMT activity (Condition C)	0.0	0.0	0.0	0.0	0.0	0.0	0.0
0x MTHFS and SHMT: CHF→5fTHF activity turned off (Condition D)	15638.3	757127.9	753509.5	91112.6	87494.2	22.9	22.9
Standard MTHFS and regular SHMT activity without MTHFS inhibition by 10fTHF (Condition E)	15326.6	770328.6	767040.0	92922.5	89634.0	23.1	23.1
	CH2F →THF	**SHMT THF →CH2F**	**CHF → 5fTHF**	**SHMT & 5mTHF binding**	**unbinding**	**SHMT & 5fTHF binding**	**unbinding**
Standard MTHFS and regular SHMT activity (Condition A)	3191.9	445.0	1.9	7214.3	7214.3	83.1	83.1
0.5x MTHFS and regular SHMT activity (Condition B)	2976.3	391.5	1.7	6433.6	6433.6	142.5	142.5
0x MTHFS and regular SHMT activity (Condition C)	0.0	0.0	0.0	0.0	0.0	639.1	639.1
0x MTHFS and SHMT: CHF→5fTHF activity turned off (Condition D)	3910.8	612.2	0.0	8228.4	8228.4	0.0	0.0
Standard MTHFS and regular SHMT activity without MTHFS inhibition by 10fTHF (Condition E)	3488.1	527.6	2.1	8305.0	8305.0	0.2	0.2

Table 3.4: Steady state fluxes of the reactions catalyzed by the enzymes MTHFS, PGT, AICARFT, DHFR, TYMS, BHMT, SAHH, MAT-I, MAT-III, GNMT, DNMT (in µM/h) for different scenarios with respect to MTHFS and SHMT activity.Reactions are indicated by the enzyme, which catalyzes them. For bidirectional reactions the direction is indicated behind the enzyme name (Enzyme: Substrate → Product).

	MTHFS	**PGT**	**AICARFT**	**DHFR**	**TYMS**	**BHMT**
Standard MTHFS and regular SHMT activity (Condition A)	1.9	5274.9	5935.2	301.7	301.7	149.7
0.5x MTHFS and regular SHMT activity (Condition B)	1.7	5267.2	5389.3	298.7	298.7	150.2
0x MTHFS and regular SHMT activity (Condition C)	0.0	0.0	0.0	0.0	0.0	180.1
0x MTHFS and SHMT: CHF→5fTHF activity turned off (Condition D)	0.0	5269.8	6750.2	296.9	296.9	150.5
Standard MTHFS and regular SHMT activity without MTHFS inhibition by 10fTHF (Condition E)	2.1	5283.7	6754.4	304.9	304.9	149.1
	SAHH		**MAT-I**	**MAT-III**	**GNMT**	**DNMT**
	SAH →HCY	**HCY →SAH**				
Standard MTHFS and regular SHMT activity (Condition A)	267.5	94.8	111.0	61.7	134.9	123.8
0.5x MTHFS and regular SHMT activity (Condition B)	268.3	95.2	111.5	61.7	140.7	122.7
0x MTHFS and regular SHMT activity (Condition C)	298.9	118.8	129.5	50.6	945.6	23.5
0x MTHFS and SHMT: CHF→5fTHF activity turned off (Condition D)	268.8	95.4	111.7	61.7	144.4	122.1
Standard MTHFS and regular SHMT activity without MTHFS inhibition by 10fTHF (Condition E)	266.6	94.4	110.5	61.7	128.3	125.0

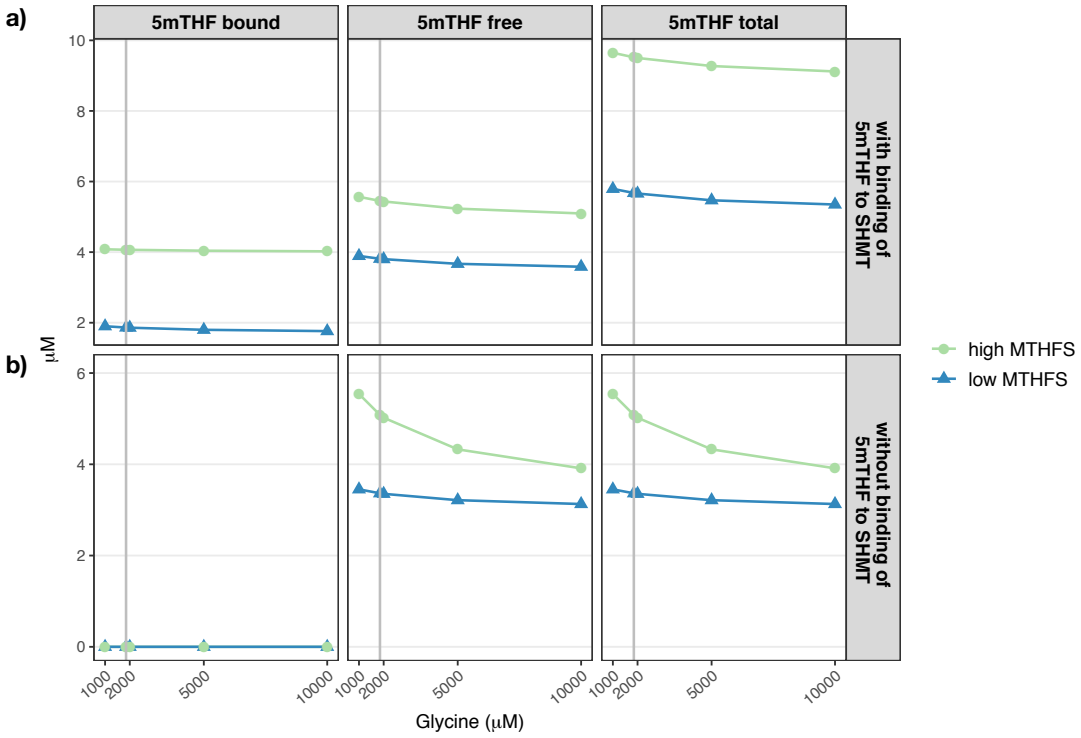

Figure 3.2: The effect of glycine on the availability of 5mTHF considering low and high levels of MTHFS. a) With the binding of 5mTHF and SHMT and b) Without the binding of 5mTHF and SHMT.

turn controls *de novo* purine synthesis by inhibiting AICARFT, as shown in cultured cells [27].

3.3.3 Effect of glycine and MTHFS activity on 5mTHF binding to SHMT

The SHMT-catalyzed interconversion of serine and glycine is reversible *in vitro* and *in vivo*, and increasing intracellular glycine concentrations has been shown to drive this reaction toward serine synthesis, which consumes CH2F otherwise available for dTMP or SAM synthesis [110]. CH2F exists at a branch point in FOCM and can be used for either synthesis of 5mTHF (through MTHFR) for SAM synthesis or for *de novo* synthesis of dTMP (through TYMS). Glycine has also been shown *in vivo* [110] and *in silico* (Chapter 1) to decrease 5mTHF levels. The relative contributions of 5fTHF and 5mTHF to regulating SHMT activity in response to changes in glycine concentration have not been determined. The interplay of glycine and MTHFS and their respective effects on 5mTHF levels was studied by comparing model steady-state concentrations of 5mTHF (free, bound to SHMT, and total) for high and low levels of MTHFS (10% of standard activity and 5-fold increase in standard activity) as a function of glycine concentration (considered levels: 1000 µM, 1850 µM, 2000 µM, 5000 µM and 10000 µM). These simulations were repeated for the case in which the binding of 5mTHF and SHMT was blocked, increasing therefore the availability of 5mTHF for other reactions.

When running the *in silico* experiment of FOCM, increasing MTHFS activity increased 5mTHF levels by making more folate available for conversion to 5mTHF, but it did not influence the effect

that glycine has on 5mTHF levels (Figure 3.2a). In this scenario, SHMT is inhibited by both 5fTHF and 5mTHF. Conversely, this is not the case when the binding of 5mTHF to SHMT is removed from the model (Figure 3.2b). In this case, 5mTHF levels become responsive to increasing glycine concentration and this decrease in 5mTHF is more pronounced with higher MTHFS activity (Figure 3.2b). This is consistent with the observation that a 50% reduction in MTHFS activity and concomitant 2-fold increase in 5fTHF levels (Table 3.2) do not affect flux through SHMT-catalyzed reactions (Table 3.3). In other words, intracellular 5fTHF is a meaningful inhibitor of SHMT activity (even when driven by increasing glycine concentrations) but only when 5mTHF levels are low, suggesting that 5mTHF is the predominant intracellular inhibitor of SHMT activity.

3.3.4 Model sensitivity analysis to unravel the effect of MTHFS activity on FOCM dynamics

Both 5mTHF and 5fTHF bind to SHMT and reduce levels of the free, active form of SHMT. The effect of MTHFS activity on the availability of 5fTHF, and the binding of 5fTHF and 5mTHF to SHMT was studied by comparing model steady-state concentrations of 5fTHF- and 5mTHF-bound SHMT at different levels of MTHFS activity. MTHFS deficiency was simulated by decreasing MTHFS availability stepwise from 90% to 10% of the standard activity whereas the 2-fold and 5-fold increase of the standard activity was considered in overexpression.

As MTHFS activity increases over a 50-fold range, 5fTHF levels decrease to less than 5% of the initial concentration (Appendix C, Table C.2). Furthermore, 5mTHF and THF levels increase by about 2-fold with increasing MTHFS activity, and the concentration/distribution of other folate forms is unchanged (Appendix C, Tables C.2 and C.3). Due to the relatively modest changes in levels of one-carbon substituted folates, flux through most FOCM enzymes remains unchanged as a result of elevated MTHFS activity (Appendix C, Table C.4). Increasing MTHFS activity decreases the amount of 5fTHF that is available to bind and inhibit AICARFT and SHMT (Figure 3.3, Appendix C, Tables C.4 and C.5). As expected, flux through both AICARFT and SHMT increases up to 2-fold as a result of decreased 5fTHF caused by elevated MTHFS activity (Appendix C, Tables C.4 and C.5). These data are consistent with observations in mammalian cells that MTHFS expression levels regulate *de novo* purine synthesis[76;77;83].

Both 5mTHF and 5fTHF are tight-binding inhibitors of SHMT, and interpreting inhibition of SHMT in this scenario where MTHFS activity is increased is complicated by the fact that 5fTHF and 5mTHF levels are affected in opposing directions (5mTHF increases and 5fTHF decreases). However, comparing binding of 5mTHF and 5fTHF as a function of MTHFS activity reveals that there is almost always more 5mTHF than 5fTHF bound to SHMT (Figure 3.3). 5fTHF binding to, and inhibition of, SHMT becomes predominant only at very high intracellular 5fTHF concentrations, as would result from an almost 90% reduction in MTHFS activity (Figure 3.3). Taken together, these data suggest that 5mTHF is the predominant physiological inhibitor of SHMT and are consistent with the observation that 5fTHF is a meaningful inhibitor of SHMT only when binding of 5mTHF to SHMT is not considered in the model (Figure 3.2)).

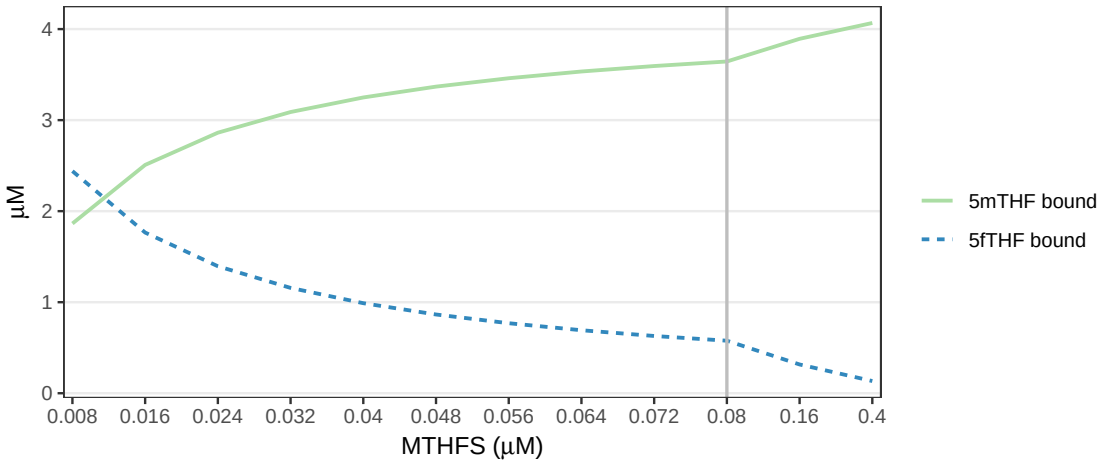

Figure 3.3: Sensitivity analysis: effect of MTHFS on 5fTHF and 5mTHF bound to SHMT.

3.3.5 Quantification of network stability

To understand the effect of the 5fTHF futile cycle on the stability of the FOCM network at steady state, we used the hybrid stochastic framework as described in Section 3.2.3. Model steady states were computed both with and without the 5fTHF futile cycle by considering four scenarios regarding two MTHFR genotype (CC, and TT) and two folate levels (replete and low folate), both of which are known to destabilize FOCM.

Interestingly, the addition of the 5fTHF futile cycle reduces the stochastic noise (indicated as lower total reaction propensities) throughout all considered conditions (Table 3.5, row 1 vs. row 2). Furthermore, the model suggests that the futile cycle mitigates the destabilizing effect that low folate has on the network for the MTHFR CC genotype (Table 3.5, CC columns). Overall, the SHMT-catalyzed reactions exhibited the greatest stochasticity in response to the absence of 5fTHF futile cycle, the MTHFR C677T polymorphism, and and/or low folate levels (Appendix C, Tables C.7, C.8, and C.9). In addition, AICARFT stochasticity also decreased in response to the addition of the futile cycle (Appendix C, Table C.7).

Table 3.5: Total propensities obtained in eight steady state conditions according to MTHFR polymorphism (CC and TT), folate status (replete, 19 uM; low, 9 uM) and presence of the 5fTHF futile cycle (FOCM network with and without futile cycle). Total propensities can be interpreted as a measure of stability when two steady states are compared since they provide an estimate of how many reactions events occur per unit of time. The steady state with lower total propensity can be consider more stable than the other.

a_0	**replete folate**		**low folate**	
	CC	**TT**	**CC**	**TT**
with 5fTHF futile cycle	$2.95 \cdot 10^{14}$	$5.74 \cdot 10^{14}$	$7.55 \cdot 10^{14}$	$9.56 \cdot 10^{14}$
without 5fTHF futile cycle	$3.67 \cdot 10^{14}$	$4.21 \cdot 10^{15}$	$1.24 \cdot 10^{15}$	$3.71 \cdot 10^{15}$

3.4 Discussion

Folate deficiency and/or genetic and environmental factors that impair FOCM are not only associated with development of pathology but often affect the function of the entire network, which makes it difficult to identify causal pathways associated with these pathologies [79;227]. Computational models advance our understanding of these complex interactions and the effects of perturbations on network function as a whole [227]. In this chapter, we focused on the role of the 5fTHF futile cycle on the FOCM network function and stability.

The 5fTHF futile cycle was shown to be critical to maintain the stability of the FOCM network, which is a novel finding. In Chapter 1, the initial model of FOCM identified both folate deficiency and the common *MTHFR* C677T polymorphism (a known human NTD risk factor through its effects on lowering folate levels [257]) as a source of network instability. The major source of the instability resulted from decreased 5mTHF levels leading to increasing flux through SHMT. Here we observed that inclusion of the 5fTHF futile cycle stabilized the FOCM network by introducing 5fTHF inhibition of both SHMT and AICARFT. This stabilization effect of the futile cycle was apparent both when folate deficiency and the MTHFR C677T polymorphism were introduced into the model (Table 3.5).

The metabolic roles of 5fTHF and MTHFS in mammalian FOCM have not been fully elucidated. 5fTHF does not serve as an enzyme cofactor, rather it acts as an inhibitor of folate-dependent enzymes AICARFT [27] and SHMT [239]. 5fTHF accumulates in seeds and spores which do not contain MTHFS activity, where 5fTHF is hypothesized to serve as a stable storage form of folate [135]. Conversely, 5fTHF is not known to accumulate in mammalian cells. *Mthfs* is an essential gene in mice [76], though this is not the case in some prokaryotes nor in Arabidopsis [96]. Humans with inborn errors of metabolism in MTHFS exhibit accumulation of 5fTHF in cultured fibroblasts and low cerebrospinal fluid (CSF) folate levels with accompanying neurological sequelae [220], which are common to many cerebral folate deficiency disorders [242].

Here, the updated model of mammalian FOCM indicates that *MTHFS* is essential in mammalian cells by preventing the pooling of cellular folate as 5fTHF (Tables 3.1 and 3.2). This depletion of folate cofactors forms induced by lack of MTHFS activity drives steady-state flux of all FOCM enzymes to zero (Tables 3.3 and 3.4). The model also shows that MTHFS is only necessary to prevent 5fTHF pooling when the 5fTHF synthesis activity of SHMT is included in the model (Tables 3.1 and 3.2), because when this activity is omitted, there is no 5fTHF formed. It is also worth noting that relatively low levels (only 10% of the standard modeled concentration) of MTHFS enzymatic activity are required to prevent this lethal 5fTHF pooling (Appendic C, Table C.3). 5fTHF and MTHFS levels affect *de novo* purine biosynthesis in cultured cells, although the relative contribution of MTHFS and 5fTHF to the regulation of *de novo* purine synthesis is unknown. The #2 and #8 carbons of the purine ring are formed through *de novo* purine synthesis in reactions catalyzed by the enzymes AICARFT and GARFT, respectively, and by the cofactor 10fTHF (Figure 3.1).

10fTHF tightly binds to and inhibits MTHFS [83], and increased MTHFS expression increases 10fTHF levels in cultured cells [27]. 5fTHF also affects *de novo* purine synthesis through its inhibition of AICARFT [27]. Increased *MTHFS* expression increased rates of *de novo* purine synthesis [83] and caused resistance to antifolates that specifically target *de novo* purine synthesis [77]. The po-

tential mechanisms underlying this observation were suggested to be either: 1) decreased 5fTHF levels and thereby less inhibition of AICARFT, and/or 2) the effect of MTHFS increasing cellular 10fTHF levels. When the 10fTHF inhibition of MTHFS term was removed from the model, folate cofactor distribution (Table 3.2) and steady-state reaction fluxes of almost all FOCM-dependent enzymes (Tables 3.3 and 3.4) were nearly identical to what was observed when the futile cycle was removed from the model. This indicates that 10fTHF inhibition of MTHFS is critical for controlling MTHFS activity and thereby maintaining cellular 5fTHF levels that limit AICARFT activity. It is worth noting that MTHFS has been shown to co-localize with the multi-enzyme *de novo* purine synthesis complex known as the "purinosome" [76]. Purinosomes form when mammalian cells are exposed to purine-deficient culture medium to increase rates of *de novo* purine synthesis [9;76], leading to the hypothesis that MTHFS delivers or "channels" 10fTHF to the purine synthesis enzymes. This physical interaction adds another layer of regulation among MTHFS activity, 10fTHF, and *de novo* purine synthesis, but the computation model does not yet account for these interactions.

The relative contribution of intracellular 5fTHF and 5mTHF in regulating SHMT, a key enzyme whose deficiency causes folate-responsive NTD risk in mice, has not been investigated to date. Both 5fTHF and 5mTHF tightly bind and inhibit SHMT [27], and SHMT activity is sensitive to 5mTHF accumulation, as has been demonstrated *in silico* and in cultured cells [110]. Comparing binding of both 5fTHF and 5mTHF to SHMT as a function of MTHFS activity indicates that 5mTHF is the predominant binder of SHMT (Figure 3.3). 5fTHF binding to SHMT becomes predominant only at high levels of intracellular 5fTHF induced by a 90% reduction in MTHFS activity. In other words, 5mTHF (and not 5fTHF) serves as the physiological inhibitor/regulator of SHMT. *De novo* purine biosynthesis is the primary pathway influenced by the 5fTHF futile cycle.

In summary, experimental investigation of the role of the 5fTHF futile cycle has been limited due to the embryonic lethality that occurs in MTHFS knock-out mice. Inclusion of the 5fTHF futile cycle in the *in silico* model has provided new insights into the metabolic functioning of FOCM, by allowing us to investigate conditions that are difficult to reproduce *in vitro* or *in vivo*. The model confirms that loss of MTHFS activity results in accumulation of folate as 5fTHF as occurs in seeds and spores, and replicates observations that the futile cycle impairs *de novo* purine biosynthesis. Importantly, the model provides new mechanistic evidence that the role of MTHFS in accelerating rates of *de novo* purine biosynthesis can be accounted for by its role in lowering 5fTHF levels and alleviating AICARFT inhibition. The model also indicates that the 5fTHF futile cycle plays an important role in limiting the increased stochastic behavior of SHMT1 introduced by the *MTHFR* C677T polymorphism and folate deficiency. Importantly, the updated model has identified a potential role of the 5fTHF futile cycle in the etiology of folic acid-responsive NTDs, as the SHMT1 knock-out mouse is the only folic acid-responsive mouse model of NTDs resulting from disruption of a folate-dependent enzyme [24;79].

Author Contributions

The content of this chapter has been submitted [169].
K. Misselbeck (K.M.), L. Marchetti (L.M.), C. Priami (C.P.), P.J. Stover (P.J.S.), and M.S. Field (M.S.F.) contributed to the experimental design. K.M., L.M. developed the computation model, while M.S.F. and P.J.S. provided all biological insight and extracted kinetic variables from the literature. C.P., P.J.S., and M.S.F. provided overall guidance of the project. All authors contributed to manuscript preparation and approved the final version.

Chapter 4

An extended mathematical model of folate-mediated one-carbon metabolism: Nuclear compartmentalization

In the preceding chapters a mathematical model and an extension of the initial model of folate-mediated one-carbon metabolism in the cytoplasm has been introduced. *In silico* simulations could successfully link genetic variants and nutritional factors to folate-related pathologies. Recent experimental evidence indicates that *de novo* dTMP synthesis must also occur in the nuclear compartment at sites of the replication machinery of the cell to prevent uracil misincorporation into DNA. Moreover, simulations of the initial model introduced in Chapter 1 indicate that *de novo* dTMP synthesis rates in the cytoplasm are insufficient to support DNA synthesis during S-phase (Table 1.9). In this chapter we present an additional extension of the cytoplasmic model presented in Chapter 3, which includes the folate-dependent reactions located in the nucleus. A key aspect of nuclear folate-mediated one-carbon metabolism is the formation of a multienzyme complex, which may allow for folate channelling between the active sites of the involved enzymes. In the model presented in this chapter nuclear complex formation and substrate channelling are taken into account and computational simulations highlight their importance for normal functioning of the network. *In silico* experiments also provide insight into the contribution of cytoplasmic and nuclear dTMP synthesis to overall dTMP synthesis. Moreover, model simulations provide evidence that the network is most sensitive to expression levels of TYMS and DHFR, while also folate partitioning between the cytoplasm and the nucleus plays an important role for dTMP activity.

4.1 Introduction

Folate-mediated one-carbon metabolism (FOCM) is required for the *de novo* synthesis of three of the four DNA bases and the remethylation of methionine (MET) to homocysteine (HCY). In the interconnected network folate serves as one-carbon carrier and donor for biochemical reactions associated to the anabolic pathways. FOCM plays an essential role for genome stability and methylation [282], and disruption of the network can be caused by genetic and/or nutritional factors, like genetic variants of the relevant enzymes or folate and vitamin deficiency. Impairment of FOCM

is associated with the pathogenesis of neural tube defects [175], neurodegenerative diseases [134] and cancer [7;29;48;128;199].

In the cytoplasm FOCM is required for the *de novo* nucleotide biosynthesis of purines and thymidylate, as well as for homocysteine remethylation to methionine, catalyzed by the vitamin B_{12} dependent enzyme methionine synthase (MTR). Recent studies demonstrate that FOCM also occurs in the nucleus at the site of DNA synthesis [12;13;42], where the *de novo* dTMP synthesis pathway functions through the four enzymes serine hydroxymethyltransferase (SHMT), dihydrofolate reductase (DHFR), methylenetetrahydrofolate dehydrogenase 1 (MTHFD1), and thymidylate synthase (TYMS). The required co-factor for dTMP synthesis, 5,10-methyleneTHF (CH2F), is independently generated by SHMT and MTHFD1 using serine or formate as one carbon source respectively. TYMS utilizes CH2F as folate co-factor to catalyze the conversion of dUMP to dTMP and dihydrofolate (DHF). DHFR reduces DHF to THF in a NADPH-dependent reaction.

A fraction of the enzymes involved in dTMP synthesis translocate to the nuclear compartment during S-phase of the cell cycle or in response to DNA damage after undergoing post-translational modification by the small ubiquitin like modifier (SUMO) protein. [12–14;78;282]. Nuclear enzymes involved in dTMP synthesis pathway form a multienzyme complex, which is associated with nuclear lamina and with the replication and epigenetic machinery [13]. Experimental observations show that SHMT acts as essential scaffold protein that anchors the multienzyme complex to nuclear lamina [13;81]. Formation of the nuclear enzyme complex appears to be essential for the functioning of *de novo* dTMP synthesis, as previous studies have demonstrated that dTMP synthesis activity in isolated nuclei is reduced following sonication [13]. Furthermore, nuclear metabolic complex formation may allow for channelling of folate polyglutamate cofactors among enzyme active sites, limiting substrate diffusion and accelerating enzymatic reaction rates. Interestingly, the formation of a multienzyme complex and resulting substrate channelling seems to be unique to mammalian cells [49]. Folate deficiency, anti-folate treatment or impaired assembly of the multienzyme complex responsible for nuclear dTMP synthesis can lead to depressed dTMP biosynthesis activity and consequently to elevated uracil incorporation in DNA [29;158;193]. During DNA replication, either dTTP or dUTP can be used to match an adenin nucleotide base on the template strand and uracil misincorporation into DNA will increase when thymidylate becomes limiting [49]. Therefore, adequate intracellular concentrations of folate and proper assembly of the dTMP synthesis pathway are important for genome stability.

Our hybrid stochastic model of cytoplasmic FOCM could successfully link genetic predisposing factors and nutritional status to folate-dependent pathodologies (Chapter 1, 2, and 3). We have shown that low levels of folate, vitamin B_{12} and the common C677T *MTHFR* polymorphism affect cytoplasmic *de novo* dTMP synthesis (Chapters 1, and 2), whereas impairment of the 5fTHF futile cycle introduced by decreased MTHFS activity is mainly associated to loss in purine synthesis activity (Chapter 3). The simulations also indicate that *de novo* dTMP synthesis rates in the cytoplasm are insufficient to support DNA synthesis during S-phase (Chapter 1). Here, we extend the mathematical model of FOCM introduced in Chapter 3 by including its compartmentalization to the nucleus and by accounting for the kinetic effects of multienzyme complex formation. *In silico* experiments with the multi-compartmental model allow to better understand the impact of enzyme complex formation and substrate channelling on overall network outcomes with a special focus on *de novo* dTMP synthesis activity. Furthermore, sensitivity analysis for model enzymes

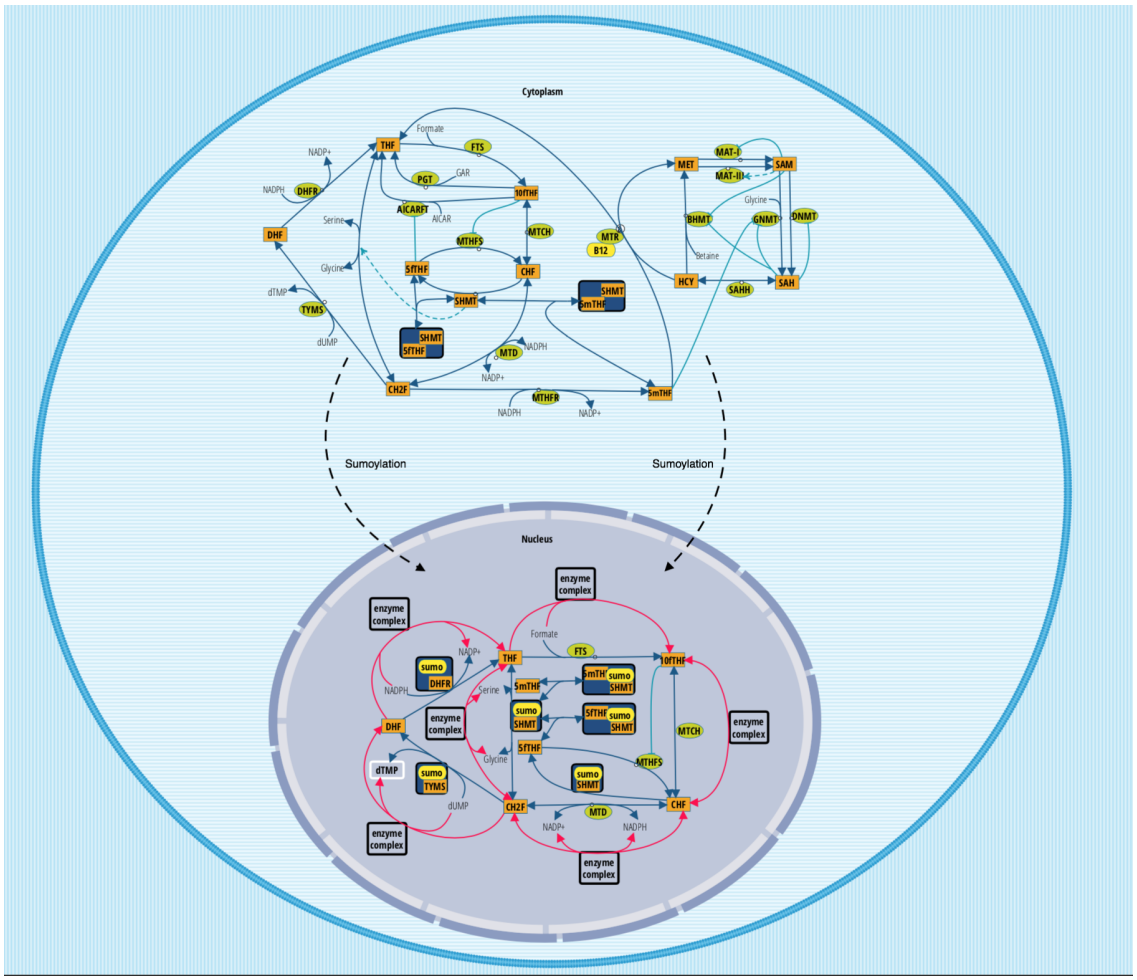

Figure 4.1: The reaction-based specification of the model including the cytoplasmic and nuclear compartment according to the notation introduced in [94]. Orange rectangles identify model variables, non-boxed substrates are model constants, green circles identify enzymes, dark blue arcs identify matter transformation, and light blue arcs identify regulatory events (dotted lines indicate activations and solid lines indicate inhibitions).

and variables is carried out to investigate their individual contribution on FOCM functioning.

4.2 Materials and Methods

4.2.1 Description of the model

The results presented in the following are based on an extension of the mathematical model of FOCM in the cytoplasm introduced in Chapter 3. This model provides a description of FOCM in the cytoplasm including its regulation of key biological processes related to *de novo* dTMP synthesis, *de novo* purine synthesis and remethylation of homocysteine to methionine. With respect to the initial model, the model herein employed has been extended to include the folate-mediated reactions occurring in the nuclear compartment, according to the graphical representation provided in Figure 4.1. The model consists of 36 reversible and irreversible reactions, most of which have

been parametrized by means of Michaelis-Menten kinetics with one or two substrates following Chapter 1 and 3. The parameter estimates for the folate cycle can be found in Appendix D, Table D.1, while the parameter estimates for the reactions related to the remethylation of methionine were unchanged with respect to Chapter 1 and are listed in Appendix A, Table A.4. Whenever possible, physiologically relevant forms of folate polyglutamate cofactors have been considered, in agreement with the cytoplasmic model. In the following we describe the technical procedure we applied to update the cytoplasmic model presented in the preceding sections by inlcuding the nuclear compartment.

Updating the cytoplasm

The cytoplasmic model introduced in Chapter 3 was extended to account for the fact that enzymes which translocate to the nucleus – SHMT, MTHFD1 (trifunctional enzyme encoded as FTS, MTD and MTCH), DHFR, TYMS and MTHFS – have been considered as variables that can change in time (some of them were encoded in the Michaelis-Menten constants in the previous model). Due to this, the Michaelis-Menten kinetics of the reactions catalyzed by the enzymes MTHFD1, DHFR and TYMS were updated to consider the turnover number and the enzyme concentration in place of the V_{max} value, considering the relation $V_{max} = k_{cat} \cdot [Enzyme]$.

Adding the nucleus

The inclusion of the nuclear compartment and its interplay with the cytosol is the main novelty of the model herein presented. The nucleus has been modeled to account for the folate-dependent reactions catalyzed by the enzymes MTHFD1, SHMT, DHFR, TYMS and MTHFS and consists of 17 variables and 16 (reversible and irreversible) reactions, as indicated in Figure 4.1. The nuclear stoichiometry and regulation of these reactions have been inherited from their counterparts in the cytoplasm, while model parameters have been derived as explained in the following. Furthermore, the model has been developed to account for observations that the four enzymes MTHFD1, SHMT, DHFR and TYMS form a multi-enzyme complex when they translocate to the nucleus [13]. This complex has the structure 1: TYMS, 1: DHFR, 1: MTHFD1, 2: SHMT and allows for substrate channelling, which was modeled as explained in the following.

Translocation to the nucleus

The translocation of enzymes and metabolites to the nucleus has been modeled by an abstract function representing the SUMO-dependent nuclear import process according to experimental results, e.g. [12]. The function derives the enzyme and folate availability in the nucleus by subtracting a given percentage of each folate form or enzyme from the steady state of the model restricted to the cytoplasm. In particular, the function follows the rules/steps described below.

1. **Nuclear kinetic parameter estimates.** The parameter estimates have been adapted from the values considered for the cytoplasm by scaling them according to the volume of the nucleus (Appendix D, Table D.1).

2. **Nuclear constant substrates.** The constant substrates (serine, glycine, formate, NADPH, NADP, and dUMP) were modeled at the same constant levels in both compartments (Appendix A, Table A.2).

3. **Nuclear folate forms.** Folate has been distributed between the two compartments by assuming that 10% of the total cytosolic folate mass is present in the nucleus, according to[12]. The initial nuclear folate distribution has been calculated by assigning 10% and 35% of total nuclear folate to 5mTHF and 5fTHF, respectively, according to[193]. The remaining 55% of nuclear folate have been assigned to THF, 10fTHF, CHF, CH2F and DHF by preserving the same proportions of their steady state distribution in the cytosol. The remaining 90% of total folate mass is redistributed in the cytoplasm according to its steady state distribution.

4. **Nuclear enzymes.** The availability of MTHFS in the nucleus has been estimated to preserve 35% of nuclear folate being 5fTHF at steady state according to[193], resulting in a concentration of 0.0039 µM (Table 4.1). The enzymes MTHFD1, SHMT, DHFR and TYMS have been translocated to the nucleus by assuming that all enzymes are in a complex of the structure 1:TYMS, 1:DHFR, 1:MTHFD1, 2:SHMT. Among the enzymes forming the complex, TYMS was the least abundant and therefore the availability of the complex has been based on a given percentage of TYMS availability in the cytoplasm. In particular, 75% of cytosolic TYMS was assumed to be translocated to the nucleus. An equal amount of DHFR and MTHFD1 has been translocated to the nucleus, whereas twice the amount of SHMT was imported to maintain stoichiometry (Table 4.1).

Table 4.1: Concentration of nuclear enzymes (in µM) before and after their translocation to the nuclear compartment. The values provided for the nucleus reflect the enzyme availability in their free form and after the multienzymatic complex formation.

Enzyme	**Cytoplasm**		**Nulcues**	
	before their translocation	**after their translocation**	**free form**	**after complex formation**
MTHFS	0.080	0.080	0.0039	0.0039
DHFR	0.617	0.335	1.202	-
MTHFD1	9.000	8.719	1.202	-
SHMT	4.500	3.937	2.404	-
TYMS	0.375	0.094	1.202	-
Complex	-	-	-	1.202

Reactions catalyzed by the enzymatic complex

The reactions catalyzed by the multi-enzyme complex are modeled considering replicates of the reactions catalyzed by the single, free enzymes. Since the enzyme complex allows for substrate channelling[227;243], the regulation of the reactions included in the channelling have been multiplied by an "acceleration" or scaling factor to enhance the reaction rates.

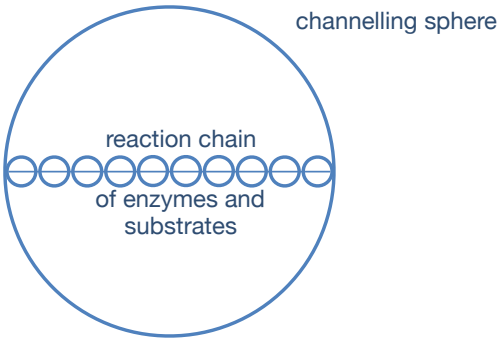

Figure 4.2: Schematic representation of the channelling reaction volume.

To derive the scaling factor, two approaches were considered:

1. **Approach based on literature.** According to the experimental evidence reported in [20] and [148], we considered a 20-fold scaling factor for the enzyme channelling, which refer empirical data reporting on the channelling of folates within the TYMS-DHFR bifunctional enzyme. Due to the lack of other available experimental data for the other enzymes SHMT and MTHFD1 forming the complete channelling network, this approach is based on the assumption that a similar scaling factor would also apply to the remaining enzymes.

2. **Approach based on the quantification of the channelling reaction volume.** As an alternative approach to the one described above, we derived the scaling factor based on the estimation of the reaction volume of the multi-enzyme complex. We interpret the channelling as a set of reactions working in close proximity in a smaller spherical reaction volume. A rough estimation of the volume diameter can be given by the length of the reaction chain that includes all the involved enzymes and substrates (Figure 4.2). The length of the reaction chain has been derived by summing up the diameters of all molecules, which have been approximated as spheres (Table 4.2). The diameter of each sphere representing a molecule has been computed by deriving the volume of the molecule according to the following formula

$$Volume = MW \cdot PSV,$$

where MW indicates the corresponding molecular weight (Table 4.2) and PSV indicates the average protein partial specific volume (0.72 cm^3/g in all cases, according to [107]). Once the total volume of the channelling including all complexes in the nucleus has been estimated, the final scaling factor has been computed by the ratio of the nuclear volume and the volume of the multi-enzyme complex. These calculations resulted in a scaling factor of 25 (see Tables 4.2 and 4.3).

Binding of 5fTHF and 5mTHF to SHMT within the enzymatic complex

The model has been developed according to the evidence that SHMT binds 5mTHF and 5fTHF when in the complex. Therefore, the model considered six different complex forms according to the binding state of the two SHMT monomers included in the complex (not shown in Figure 4.1,

Table 4.2: Diameters for the molecules involved in the channelling. The structure of the molecules has been approximated as sphere and its volume has beend calculated by considering the molecular weight (MW) and the average protein partial specifiv volume (PSV = 0.72 cm^3/g).

Enzyme	**MW (g/mol)**	**Volume** (L)	**radius** (dm)	**diameter** (dm)
MTHFD1	102000	$2.45{\cdot}10^{-19}$	$3.88{\cdot}10^{-8}$	$7.76{\cdot}10^{-8}$
SHMT	53000	$2.54{\cdot}10^{-19}$	$3.93{\cdot}10^{-8}$	$7.86{\cdot}10^{-8}$
TYMS	36000	$8.64{\cdot}10^{-23}$	$2.74{\cdot}10^{-8}$	$5.48{\cdot}10^{-8}$
DHFR	21000	$5.04{\cdot}10^{-23}$	$2.29{\cdot}10^{-8}$	$4.58{\cdot}10^{-7}$
Substrates & Products	**MW (g/mol)**	**Volume** (L)	**radius** (dm)	**diameter** (dm)
THF	445.43	$5.35{\cdot}10^{-25}$	$5.03{\cdot}10^{-9}$	$1.01{\cdot}10^{-8}$
10fTHF	473.44	$5.68{\cdot}10^{-25}$	$5.14{\cdot}10^{-9}$	$1.03{\cdot}10^{-8}$
CHF	582.53	$6.99{\cdot}10^{-25}$	$5.51{\cdot}10^{-9}$	$1.10{\cdot}10^{-8}$
CH2F	457.43	$5.49{\cdot}10^{-25}$	$5.08{\cdot}10^{-9}$	$1.02{\cdot}10^{-8}$
DHF	443.42	$5.32{\cdot}10^{-25}$	$5.03{\cdot}10^{-9}$	$1.01{\cdot}10^{-8}$
dUMP	308.18	$3.70{\cdot}10^{-25}$	$4.45{\cdot}10^{-9}$	$8.91{\cdot}10^{-9}$
dTMP	322.21	$3.87{\cdot}10^{-25}$	$4.52{\cdot}10^{-9}$	$9.04{\cdot}10^{-9}$
NADPH	833.35	$1.00{\cdot}10^{-24}$	$6.20{\cdot}10^{-9}$	$1.24{\cdot}10^{-8}$
NADP+	744.42	$8.93{\cdot}10^{-25}$	$5.97{\cdot}10^{-9}$	$1.19{\cdot}10^{-8}$
Formate	46.02	$5.52{\cdot}10^{-26}$	$2.36{\cdot}10^{-9}$	$4.72{\cdot}10^{-9}$
Serine	105.09	$1.26{\cdot}10^{-25}$	$3.11{\cdot}10^{-9}$	$6.22{\cdot}10^{-9}$
Glycine	75.07	$9.01{\cdot}10^{-26}$	$2.78{\cdot}10^{-9}$	$5.56{\cdot}10^{-9}$

Table 4.3: Calculation of the channelling speed-up factor. The diameter of the channelling sphere is estimated by the length of the reaction chain including all enzymes and substrates (diameter of involved enzymes, substrates and products, and diameter total). The standard equation for the volume of a sphere ($V = \frac{4}{3}r^3\pi$) is used to calculate the volume for one complex. The total channelling volume results from the multiplication of the number of complexes in the nucleus (Tabel 4.1).

Diameter of involved enzymes (dm)	$3.36{\cdot}10^{-7}$
Diameter of involved substrates and products (dm)	$1.35{\cdot}10^{-7}$
Diameter total (dm)	$4.70{\cdot}10^{-7}$
Volume of the reaction chain for one complex (L)	$5.45{\cdot}10^{-20}$
Number of complexes in the nucleus	$1.59{\cdot}10^{5}$
Volume of the reaction chain for all complexes in the nucleus (L)	$8.68{\cdot}10^{-15}$
Volume nucleus (L)	$2.20{\cdot}10^{-13}$
Scaling factor = Volume nuclues/Volume of the reaction chain for all complexes in the nucleus	**25**

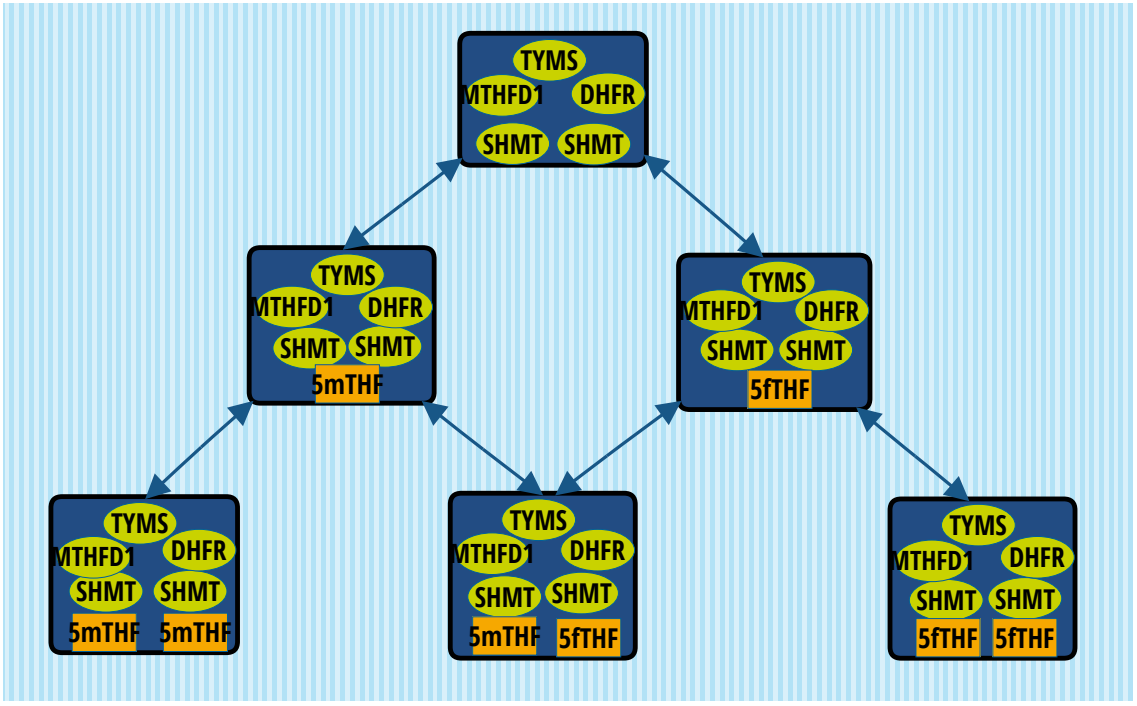

Figure 4.3: The six states of the nuclear enzymatic complex with respect to the binding of SHMT with 5mTHF and 5fTHF. Green cycles identify the enzymes of the complex, while orange rectangle boxes refer to the model variables 5mTHF and 5fTHF. The arrows between the complexes indicate the possibility to evolve from one state to the other.

where a generic enzyme complex has been included to simplify the network). According to Figure 4.3 the six complex configurations are: 1) both SHMT monomers are unliganded, 2) one SHMT is bound to 5mTHF, while the other is unliganded, 3) one SHMT is bound to 5fTHF, while the other is unliganded, 4) one SHMT monomer is bound to 5mTHF, while the other is bound to 5fTHF, 5) both SHMT monomers are bound to 5mTHF and 6) both SHMT monomers are bound to 5fTHF.

4.2.2 Description of the computational environment and simulation procedure

The mathematical model has been defined as a set of ODEs implemented as a MATLAB function. All model simulations have been computed with the numerical ODE solver ode15s.
All simulations for each scenario of interest were carried out considering the following steps: 1) the cytoplasmic compartment was simulated until reaching steady state. 2) The translocation was carried out and initial settings for the cytoplasm and nuclear compartment were updated according to the rules described above, and 3) the complete model including the cytoplasm and the nucleus was simulated until reaching steady state.

4.2.3 *In silico* experiments

In the following a short technical description of the considered simulation scenarios is provided. If not indicated differently, we used the 25x scaling factor to produce the *in silico* experiments.

Required dTMP molecules for cell replication

Based on the assumption that 59% of the human genome consists AT base pairs [17], the number of dTMP molecules needed for the replication of the human genome can be computed as

$$0.59 \cdot 3 \cdot 10^9 = 1.77 \cdot 10^9.$$

Model-derived production of dTMP molecules for cell replication

The predicted number of produced dTMP molecules has been derived by model simulation considering the final steady state of the system. The total number of dTMPs is given by the sum of the molecules produced in the cytosol and in the nucleus. For each one of the two compartments, the following formula has been used:

$$\#T = v_{TYMS} \cdot k_{vol} \cdot Sl \cdot N_A \cdot Vol,$$

where v_{TYMS} indicates the flux of the reaction catalyzed by TYMS in the corresponding compartment (cytoplasm or nucleus), $\mathrm{k_{vol}} = 10^{-6} \frac{\mathrm{M}}{\mu\mathrm{M}}$ indicates a scaling factor used used to transform the concentration from µM to M, Sl = 8h is the length of the S-phase, $N_A = 6.022 \cdot 10^{23}$ is the Avogadro constant and Vol is the volume of the respective compartment (Table D.1).

The role of enzymatic complex formation for sufficient dTMP production

The equation introduced above for estimating the total production of dTMP molecules has been computed for four different simulation scenarios to test the contribution of nuclear enzymatic complex formation and the effect of the substrate channelling on acceleration of the pathway. In particular, the following four scenarios have been considered: 1) restriction to the cytoplasmic model 2) entire network, without multi-enzyme complex formation (all nuclear enzymes are in free from), 3) entire network, with multi-enzyme complex formation and 20-fold substrate channelling acceleration (literature-based approach) and 4) entire network, with complex formation and 25-fold substrate channelling acceleration (volume-based approach).

The effect of folate partitioning on dTMP production

The impact of folate partitioning between the two compartments was studied by varying the percentgae of folate being translocated to the nuclear compartment. Starting from the standard scenario, in which 10% of cytoplasmic folate is translocated to the nucleus, the ratio of nuclear folate was increased up to 60% (considering the levels of 20%, 30%, 40% and 50%). *De novo* dTMP activity was measured for all the scenarios and expressed as number of predicted Ts, as well as percentage of dTMP synthesis rates required for genome replication. We further assessed with which folate partitioning the model would predict a sufficient dTMP activity for genome relication.

The role of nuclear MTHFS

The influence of the addition of MTHFS in the nuclear compartment was studied by comparing model steady states and the dTMP production in three different scenarios: 1) standard case with complex formation and 25-fold substrate channelling acceleration, 2) scenario 1, without MTHFS in the nucleus and 3) scenario 2 without the nuclear reaction R_{SHMT} : CHF $\rightarrow$ 10fTHF.

Sensitivity analysis of model enzymes

To study the impact of model enzymes on the network output variables a sensitivity analysis was carried out. The perturbation of the enzymes was calculated by a multiplicative scaling factor with the levels 0.25, 0.5, 1, 2 and 4. For each considered enzyme concentration we simulated the system to identify the steady state values of the following variables of interest: purine synthesis (measured as fluxes through PGT and AICARFT), thymidylate synthesis (measured as total dTMP production, and fluxes through TYMS in the cytoplasm and nucleus), and methionine synthesis (measured as flux through MTR). The influence of an enzyme on an output variable was measured by the coefficient of variation (CV):

$$\mathrm{CV} = \frac{\mathrm{sd}}{\mathrm{mean}} \cdot 100\%,$$

where a high value identifies those enzymes with a high influence on the output variable.

The contribution of NADPH and NADP to network dynamics

The contribution of NADPH and NADP to FOCM dynamics was studied by comparing model steady states of different scenarios corresponding, respectively, to sum variation and ratio variation of these two molecules. In the model NADPH and NADP are considered to be constant over time (NADPH: 58µM, NADP: 18µM for both the cytoplasm and nucleus). To assess the role of NADPH and NADP two *in silico* experiments were carried out. First, to understand the influence of the overall availability of NADPH and NADP, four scenarios with respect to the sum of NADPH and NADP were considered, while their ratio stayed fixed as NADPH:NADP = 70:30. Their overall availability was varied by means of a scaling factor with the levels of 0.5, 1, 1.5, and 2, corresponding to concentrations levels of 38µM, 76µM, 114µM, and 152µM, respectively (Table 4.4). Second, the impact of the partitioning of NADPH and NADP was assessed by considering a step-wise change in the ratio NADPH/NADP, while the overall availability of the two substrates remained constant (76 µM). This resulted in eleven scenarios indicated in Table 4.5.

Table 4.4: The scenarios considered to measure the effect of the overall availability of NADPH and NADP for a fixed ratio of the two variables NADPH/NADP = 70:30. The sum of NADPH and NADP was varied according to the scaling factors: 0.5x, 1x, 1.5x, and 2x.

Scaling factor	**NADPH+NADP**	**NADPH**		**NADP**	
	(µM)	(% of sum)	(µM)	(% of sum)	(µM)
0.5x	38	70	26.6	30	11.4
1x	76	70	53.2	30	22.8
1.5x	114	70	79.8	30	34.2
2x	152	70	106.4	30	45.6

Table 4.5: The scenarios considered to measure the effect of the partitioning of NADPH and NADP, while their overall sum stayed fixed at 76µM. NADPH and NADP are presented as percentage of their sum, as well as the responding concentration used for model simulations in the cytoplasm and nucleus.

Scenario	**NADPH**		**NADP**	
	(% of sum)	(µM)	(% of sum)	(µM)
1	0	0.0	100	76.0
2	10	7.6	90	68.4
3	20	15.2	80	60.8
4	30	22.8	70	53.2
5	40	30.4	60	45.6
6	50	38.0	50	38.0
7	60	45.6	40	30.4
8	70	53.2	30	22.8
9	80	60.8	20	15.2
10	90	68.4	10	7.6
11	100	76.0	0	0.0

The contribution of glycine and serine to network dynamics

The contribution of serine and glycine to FOCM dynamics was studied following the same approach as for NADPH and NADP described in the previous paragraph. For this model steady states of different scenarios corresponding, respectively, to sum variation and ratio variation of glycine and serine were compared. In the model serine and glycine are considered to be constant over time (serine: 468µM, glycine: 1850µM for both the cytoplasm and nucleus). To assess the role of glycine and serine two *in silico* experiments were carried out. First, to understand the influence of the overall availability of glycine and serine, four scenarios with respect to the sum of serine and glycine were considered, while their ratio stayed fixed as glycine:serine = 80:20. Their overall availability was varied by means of a scaling factor with the levels 0.001, 0.1, 0.5, 1, 1.5, 2, 5, and 10, corresponding to concentration levels of 23.2µM, 231.8µM, 1159µM, 2318µM, 3477µM, 11590µM, and 23180µM, respectively (Table 4.6). Second, the impact of the partitioning of glycine and serine was assessed by considering a step-wise change in the ratio glycine/serine, while the overall availability of the two substrates remained constant (2318 µM). This resulted in eleven scenarios indicated in Table 4.7.

Table 4.6: The scenarios considered to measure the effect of the overall availability of glycine and serine for a fixed ratio of the two variables glycine/serine = 80:20. The sum of glycine and serine was varied according to the scaling factors: 0.01x, 0.1x, 0.5x, 1x, 1.5x, 2x, 5x, and 10x.

Scaling factor	**Serine+Glycine**	**Serine**		**Glycine**	
	(µM)	(% of sum)	(µM)	(% of sum)	(µM)
0.01x	23.2	20	4.6	80	18.5
0.1x	231.8	20	46.4	80	185.4
0.5x	1159.0	20	231.8	80	927.2
1x	2318.0	20	463.6	80	1854.4
1.5x	3477.0	20	695.4	80	2781.6
2x	4636.0	20	927.2	80	3708.8
5x	11590.0	20	2318.0	80	9272.0
10x	23180.0	20	4636.0	80	18544.0

Table 4.7: The scenarios considered to measure the effect of the partitioning of glycine and serine, while their overall sum stayed fixed at 2318µM. Glycine and serine are presented as percentage of their sum, as well as the responding concentration used for model simulations in the cytoplasm and nucleus.

Scenario	**Serine**		**Glycine**	
	(% of sum)	(µM)	(% of sum)	(µM)
1	0	0.0	100	2318.0
2	10	231.8	90	2086.2
3	20	463.3	80	1854.4
4	30	695.4	70	1622.6
5	40	927.2	60	1390.8
6	50	1159.0	50	1159.0
7	60	1390.8	40	927.2
8	70	1622.6	30	695.4
9	80	1854.4	20	463.6
10	90	2086.2	10	231.8
11	100	2318.0	0	0.0

4.3 Results

4.3.1 The role of the enzymatic complex formation for sufficient dTMP production

To test the impact of the inclusion of the nuclear compartment and the contribution of nuclear enzymatic complex formation along with substrate channelling on *de novo* dTMP synthesis, model steady states have been compared across different experiments. In particular, the four following scenarios have been considered: 1) restriction to the cytoplasmic model 2) entire network, without multi-enzyme complex formation 3) entire network, with multi-enzyme complex formation and 20-fold substrate channelling acceleration (literature-based approach) and 4) entire network, with complex formation and 25-fold substrate channelling acceleration (volume-based approach).

The model considering only the cytoplasm leads to a dTMP production rate which produces 77% of required dTMP synthesis for genome replication (Table 4.8). However, based on the experimental evidence that the *de novo* dTMP pathway must localize to the nucleus to prevent uracil accumulation in DNA [158], this scenario can be considered as an artificial, fragmentary one. When considering the *in silico* experiments for the complete model including the translocation to the nucleus but not the multienzyme complex formation, cytoplasmic dTMP rates decrease by more than 75% as a response of the decreasing enzyme and substrate availability in the compartment (Table 4.8). Nuclear dTMP synthesis rates cannot compensate for this decline and only 21% of needed dTMP is manifactured (Table 4.8), indicating the lack of an important aspect.

Indeed, when metabolic channelling and multienzyme complex formation is considered, the nuclear dTMP synthesis pathway displays an enhanced activity and the number of produced Ts increases up to 25-fold (Table 4.8, row 3 and 4). While the model with the channelling accelaration factor of 20-fold predicts a dTMP rate which covers 72% of required Ts, the model with the scaling factor based on the volume of the sphere (25-fold) captures 85.5% of required synthesis, thus excelling the results of the model restricted to the cytoplasm.

Our observations indicate that dTMP synthesis functions mostly in the nucleus and when the respective enzymes SHMT, TYMS, DHFR and MTHFD1 form a multienzyme complex. Accounting for the kinetic effects of nuclear multienzyme complex formation on *de novo* dTMP synthesis is therefore essential to support genome stability.

The gap between predicted (85%) and required dTMP activity can be explained by two considerations: 1) The nuclear FOCM network is quite sensitive to folate availability and only a slight increase in nuclear folate levels rises *de novo* dTMP synthesis levels remarkly (section 4.3.2, Table 4.9). 2) Besides the *de novo* dTMP synthesis also the salvage pathway – catalyzed by TK1 – recovers dTMP from thymidine and contributes to the overall needed dTMP activity. However, this process is not yet included in the model.

Table 4.8: Predicted number of produced T during S-phase (8 hours) for the cytoplasm and nuclear compartment, as well as the ratio of required activity for human genome replication ($1.77 \cdot 10^9$) for different scenarios with respect to the compartmentalization, nuclear enzymatic complex formation and substrate channelling.

	Number of produced T in 8h			**% of required T**
	cytoplasm	**nucleus**	**total**	
only cytoplasm	$1.366 \cdot 10^9$	–	–	77.2
cytoplasm and nucleus without complex formation	$0.328 \cdot 10^9$	$0.047 \cdot 10^9$	$0.376 \cdot 10^9$	21.2
cytoplasm and nucleus with complex formation 20x scaling factor	$0.328 \cdot 10^9$	$0.948 \cdot 10^9$	$1.2768 \cdot 10^9$	72.1
cytoplasm and nucleus with complex formation 25x scaling factor	$0.328 \cdot 10^9$	$1.1858 \cdot 10^9$	$1.5138 \cdot 10^9$	85.5

4.3.2 The effect of folate partitioning on dTMP production

To assess the effect of folate partitioning between the two compartments on *de novo* dTMP synthesis, we evaluated seven scenarios. In particular, the percentage of folate translocated to the nuclear compartment was step-wise increased from 10 % (standard scenario) to 60%.

While cytoplasmic dTMP production rates decrease moderately in response to declining folate availability in the respective compartment, nuclear dTMP synthesis increases notably (Table 4.9); throughout the considered partitioning range the cytoplasmic proportion for dTMP synthesis diminishes by 60%, whereas nuclear production rate increases almost 10-fold.

We further assessed which folate partitioning allows to meet dTMP synthesis rates sufficient for genome replication. Interestingly, already the translocation of 11.6%, which is only 1.6% more than in the standard scenario, of cellular folate to the nucleus increases the flux through nuclear TYMS enough to meet adequate levels of dTMP synthesis. These results show that the network exhibits high sensitivity with respect to nuclear folate levels and further indicate the importance of the nuclear dTMP pathway as main contributor of thymidylate synthesis.

Table 4.9: Predicted number of produced Ts during S-phase (8 hours) and the ratio of required dTMP synthesis for human genome replication ($1.77{\cdot}10^9$) in response to the partitioning of folate between the nuclear and cytosol compartments.

Partitioning of folate nucleus/cytoplasm (%)	**Number of Ts produced in 8h**			**% of required T**		
	cytoplasm	**nucleus**	**total**	**cytoplasm**	**nucleus**	**total**
10/90	$3.28{\cdot}10^8$	$1.19{\cdot}10^9$	$1.51{\cdot}10^9$	18.54	66.96	85.50
11.6/88.4	$3.24{\cdot}10^8$	$1.45{\cdot}10^9$	$1.78{\cdot}10^9$	18.33	82.05	100.38
20/80	$3.02{\cdot}10^8$	$3.02{\cdot}10^9$	$3.32{\cdot}10^9$	17.06	170.48	187.54
30/70	$2.68{\cdot}10^8$	$5.01{\cdot}10^9$	$5.28{\cdot}10^9$	15.14	283.29	298.43
40/60	$2.27{\cdot}10^8$	$7.01{\cdot}10^9$	$7.24{\cdot}10^9$	12.83	396.17	409.00
50/50	$1.81{\cdot}10^8$	$8.97{\cdot}10^9$	$9.15{\cdot}10^9$	10.24	506.52	516.76
60/40	$1.33{\cdot}10^8$	$1.09{\cdot}10^{10}$	$1.10{\cdot}10^{10}$	7.54	613.59	621.13

4.3.3 The role of nuclear MTHFS activity

To study the impact of nuclear MTHFS activity model steady states and the dTMP production rates for three different scenarios were compared. In particular, we analyzed the following scenarios: 1) standard model with normal MTHFS activity; 2) no nuclear MTHFS availability and 3) no nuclear MTHFS availability and $SHMT : CHF \rightarrow 5fTHF$ activity turned off. The role of 5fTHF is not yet elucidated. It does not serve as a cofactor in folate-dependent biosynthetic reactions, but rather has been proposed to serve as an intracellular storage form of folate [239], and as an inhibitor of the folate dependent enzymes SHMT and AICARFT [27;91]. The model demonstrates that nuclear MTHFS activity is necessary to prevent the accumulation of nuclear folate as 5fTHF. This result is consistent with the observations made for cytoplasmic MTHFS activity presented in Chapter 3. When the nuclear compartment was modeled without MTHFS availability, nuclear folate pools irrevocably as 5fTHF (90% of total nuclear folate, Table 4.10), while the remaining 10% of nuclear folate is 5mTHF (the stoichiometry of the network implies no nuclear MTR nor

MTHFR activity, therefore initial 5mTHF levels remain unchanged at the predefined availability of 10% throughout all simulation scenarios (Figure 4.1). As a consequence no folate co-factors are available for nuclear dTMP synthesis and the overall production rate decrease to 18.5% of required TYMS activity (Table 4.10). The lethal pooling of nuclear folate as 5fTHF can be prevented, if in addition to the deletion of nuclear MTHFS activity the SHMT-catalyzed conversion of CHF to 5fTHF is also turned off.

Table 4.10: Predicted number of produced Ts during S-phase (8 hours), the ratio of required dTMP synthesis, and 5fTHF levels (as percentage of toal nuclear folate) for different scenarios with respect to nuclear MTHFS and SHMT activity.

	Number of produced Ts in 8h			**% of required T**	**% of nuclear 5fTHF**
	cytoplasm	**nucleus**	**total**		
Standard model	$0.328{\cdot}10^9$	$1.185{\cdot}10^9$	$1.513{\cdot}10^9$	85.5	35.3
No nuclear MTHFS	$0.328{\cdot}10^9$	0	$0.328{\cdot}10^9$	18.5	90.0
No nuclear MTHFS and SHMT: CHF → 5fTHF activity turned off	$0.328{\cdot}10^9$	$1.190{\cdot}10^9$	$1.518{\cdot}10^9$	85.8	35.0

4.3.4 Sensitivity analysis of model enzymes

We performed various computational experiments to summarize the effect of model enzymes on the biosynthetic reactions involved in FOCM. For each enzyme, five simulations were carried out to perturb the enzme availability according to a multiplicative scaling factor, and the impact of this variation was measured across the simulations by means of the coefficient of variation as presented in Section 4.2. A summary of the results, indicating the impact of each enzyme on network output variables of interest, is presented in Figure 4.4. The model predictions for all executed simulations are in Appendix D, Table D.4.

The steady state activity of purine synthesis depends highly on the magnitude of the two enzymes AICARFT and PGT, which is not unexpected, since AICARFT and PGT are the enzymes catalyzing the reactions of interest. Interestingly, this behaviour cannot be observed for the flux through MTR: here TYMS and DHFR are the enzymes with the highest impact, while MTR is the enzyme with the third highest influence. Indeed, TYMS and DHFR exhibit high values of variation for all model output variables of interest, including the relevant fluxes for purine synthesis, methionine synthase and thymidylate synthesis (Figure 4.4A). Changing TYMS levels affects the variables of interest in a linear way with the most pronounced effect in the scenario in which TYMS is upregulated by a 4-fold scaling factor (Appendix D, Table D.4). On the other hand, the high values of variation of DHFR is solely a consequence of the most extreme scenario considering the downregulation of DHFR by a factor of 0.25, while for the remaining DHFR levels, no variation in the variables of interest can be observed (Appendix D, Table D.4). The two extreme scenarios, 0.25×DHFR and 4×TYMS respectively, lead to an accumulation of DHF, which in turn leaves no folate co-factors available for *de novo* dTMP synthesis, purine synthesis or methionine synthase (Appendix D, Table D.4). These shifts in the steady state distribution of folate lead to

A

	V_AICARFT	V_PGT	V_MTR	#Ts total	V_TYMS cytoplasm	V_TYMS nucleus
AICARFT	98.43	0.11	0.09	0.05	0.21	0.00
BHMT	1.86	0.88	2.82	1.30	6.10	0.00
DHFR	55.90	55.90	55.90	24.36	55.90	16.85
DNMT	0.27	0.10	0.33	0.16	0.73	0.00
GNMT	0.12	0.04	0.14	0.07	0.31	0.00
MAT-I	0.29	0.11	0.35	0.17	0.80	0.00
MAT-III	0.10	0.04	0.12	0.06	0.27	0.00
MTHFD1	1.20	9.74	19.99	6.11	31.20	0.00
MTHFR	11.71	23.96	44.96	12.49	67.43	0.00
MTHFS	11.89	0.26	0.73	11,00	1.64	13.64
MTR	11.71	23.92	47.85	12.47	67.35	0.00
PGT	0.06	98.28	0.06	0,03	0.13	0.00
SAHH	1.62	0.76	2.43	1.12	5.27	0.00
SHMT	11.01	9.55	31.68	8.27	44.31	0.00
TYMS	55.92	55.90	55.91	51.48	106.81	49.07

B

V_AICARFT	V_PGT	V_MTR	#Ts total	V_TYMS cytoplasm	V_TYMS nucleus
AICARFT	PGT	TYMS	TYMS	TYMS	TYMS
TYMS	TYMS	DHFR	DHFR	MTHFR	DHFR
DHFR	DHFR	MTR	MTHFR	MTR	MTHFS
MTHFS	MTHFR	MTHFR	MTR	DHFR	MTHFD1
MTR	MTR	SHMT	MTHFS	SHMT	PGT
MTHFR	MTHFD1	MTHFD1	SHMT	MTHFD1	MTR
SHMT	SHMT	BHMT	MTHFD1	BHMT	BHMT
BHMT	BHMT	SAHH	BHMT	SAHH	MTHFR
SAHH	SAHH	MTHFS	SAHH	MTHFS	DNMT
MTHFD1	MTHFS	MAT-I	MAT-I	MAT-I	SHMT
MAT-I	MAT-I	DNMT	DNMT	DNMT	SAHH
DNMT	AICARFT	GNMT	GNMT	GNMT	MAT-III
GNMT	DNMT	MAT-III	MAT-III	MAT-III	MAT-I
MAT-III	GNMT	AICARFT	AICARFT	AICARFT	GNMT
PGT	MAT-III	PGT	PGT	PGT	AICARFT

Figure 4.4: The impact of model enzymes on network output variables of interest (purine synthesis: flux through PGT and AICARFT, methionine synthesis: flux through MTR and thymidylate synthesis: total dTMP synthesis by means of predicted number of Ts and fluxes through TYMS in the cytoplasm and nucleus). A) Coefficient of variation across enzyme levels for each enzyme-output combination. Colors encode the ranking of these values: red identifies the smallest CV relating to the enzyme with the least influence on the output variable, whereas green identifies the highest changes. B) Ranking of the enzyme influence on the output variables. For each variable of interest the enzymes are listed according to their influence on this variable, starting from the one with the highest influence. Colors refer to the values in subfigure A.

the high variation levels (Appendix D, Table D.4, compare e.g. the results for MTR with DHFR and TYMS).

Nuclear dTMP synthesis remains quite stable throughout all considered scenarios (Appendix D, Table D.4 and Figure 4.4); only changes in TYMS, DHFR or MTHFS activity affect the nuclear nucleotide synthesis. While the impact of TYMS levels on dTMP production is obvious, the simulations suggest that 5fTHF accumulation as a result of decreasing MTHFS levels lead in turn to a modest decrease of dTMP synthesis, because less folate co-factor is available (Section 4.3.3, Table 4.10). DHFR levels influence dTMP synthesis only in the restricting scenario, when DHFR levels are downregulated by the factor of 0.25 (Appendix D, Table D.4). In this scenario DHFR becomes the limiting enzyme in the cytoplasm and drives the translocation of the enzymes forming the nuclear complex. This reduces the availability of enzymatic complexes by 50% and in turn nuclear dTMP synthesis declines.

The effect of the enzymes of the remethylation pathway (MAT-I, MAT-III, GNMT, DNMT, BHMT, and SAHH) on network pathways is negligible (Figure 4.4). Indeed, only BHMT and SAHH slightly influence cytoplasmic dTMP activity due to diminishing HCY levels as response to downregulated SAHH levels and upregulated BHMT levels. In turn less flux goes through the MTR reaction allowing to accumulate 5mTHF levels, which on the other hand leaves less folate available as co-factor for cytoplasmic dTMP synthesis.

Overall the model suggests that there is substantial robustness in the FOCM network towards smaller enzymatic variation and identifies DHFR and TYMS as the main driving players for network dynamics.

4.3.5 The contribution of NADPH and NADP to network dynamics

To study the effect of changes in NADPH and NADP availability on FOCM dynamics, model steady states of different scenarios corresponding to sum and ratio variation of NADPH and NADP were considered, as described in Section 4.2.3.

Increasing overall availability of NADPH and NADP, while considering a fixed ratio (NADPH = 70%, NADP = 30%), leads to a shift in the steady state distribution of cytoplasmic folate with an accumulation of folate as 5mTHF (Appendix D, Table D.5). This comes at the expense of the remaining folate co-factors, including 5fTHF and 10fTHF, with the exception of THF, which decreases only moderately (Figure 4.5a). 10fTHF is the substrate of the purine synthesis reactions catalyzed by AICARFT and PGT. The simulations indicate that flux through the PGT-catalyzed reaction declines as response to the lack of substrate, whereas the flux through the second purine-relevant enzyme AICARFT exhibits the opposing behavior (Figure 4.5b). A possible explanation for this is the depletion of 5fTHF levels, following the increase of NADPH+NADP (Figure 4.5a). 5fTHF is a known inhibitor of AICARFT [27] and therefore lower concentration levels increase the flux through the AICARFT-catalyzed reaction (Figure 4.5b). These data suggest that the 5fTHF futile cycle and its link to purine synthesis could play an important role in controlling purine synthesis and in compensating for/preventing major damage due to downregulations of 10fTHF.

Even though 5fTHF depletion is reducing the amount of SHMT bound to 5fTHF, fluxes through SHMT-catalyzed reactions do not increase (Figure 4.5c and d). This is a consequence of increasing 5mTHF levels, which compensate for the loss of SHMT inhibition by 5fTHF by increasing the ratio of SHMT bound to 5mTHF from 16% to 94% of overall SHMT (Figure 4.5a,c and d, Appendix D, Table D.5).

The activity of *de novo* thymidylate synthesis depends on the substrate CH2F of the reaction catalyzed by TYMS. When NADPH and NADP levels are increasing, cytoplasmic CH2F levels decrease as response to accumulation of folate as 5mTHF and therefore cytoplasmic dTMP activity declines by more than 50% throughout the considered scenarios (Figure 4.5b, Table 4.11). However, nuclear CH2F levels increase alongside NADPH and NADP levels, because nuclear absense of MTHFR is averting 5mTHF accumulation in this compartment and because the flux through MTD (CHF $\rightarrow$ $CH2F$) accelerates (Figure 4.5e and f). Consequently, more folate cofactor is available for nuclear *de novo* dTMP production (Table 4.11, Figure 4.5e). The model suggests that the overall activity of dTMP synthesis is quite robust towards increases in NADPH and NADP, because nuclear productivity inceases to balance the decline in cytoplasmic TYMS activity (Table 4.11, rows 2 to 4). Lower disposability of NADPH and NADP affects total dTMP synthesis and simulations suggest that the ratio of required Ts reduces from around 80% to 70% (Table 4.11, rows 1 and 2).

Furthermore, we carried out a second set of *in silico* experiments, in which overall availability of NADPH and NADP has been considered constant and the ratio of NADPH and NADP has been varied according to Table 4.5. The model shows that NADPH is required to prevent impairment of dTMP synthesis: while nuclear dTMP activity enhances with increasing NADPH availability, cytoplasmic dTMP activity doesn't show a linear trend and reaches it's maximum with a NADPH:NADP ratio of 60:40 (Table 4.12). Moreover, for all considered scenarios nuclear dTMP synthesis remains the main contributor to overall dTMP activity. The depletion of NADPH

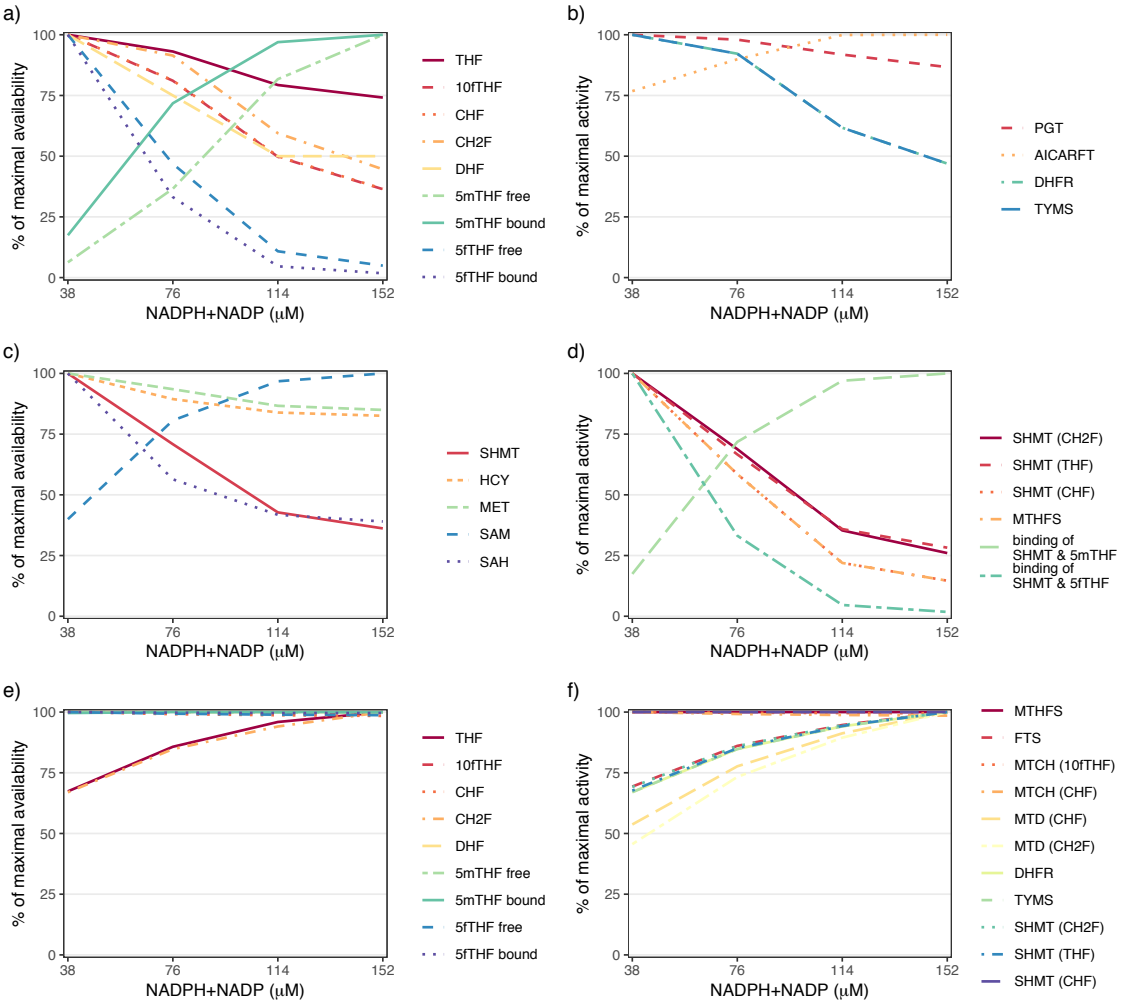

Figure 4.5: Effect of NADPH and NADP availability on model variables and fluxes. The provided values are computed by considering model steady states achieved by varying the total availability of NADPH+NADP from 38µM to 152 µM, according to Table 4.4. The results are expressed in terms of percentage of the maximal acitivity/availability over the considered range. Panel a) to d) display the results for cytoplasmic variables and fluxes, while panel e) and f) display those for the nuclear compartment. Reactions are indicated by the catalyzing enzyme, and for bidirectional reactions the direction is indicated behind the enzyme name: Enzyme (Substrate).

levels leads to an interconvertible pooling of folate as DHF and consequently stops dTMP synthesis (Table 4.12, Scenario 1). On the other hand, depletion of NADP increases nuclear dTMP synthesis notably (Table 4.12, Scenarios 1 and 11). This is a consequence of the accumulation of CH2F due to the loss of R_{MTD} : CH2F $\rightarrow$ CHF activity.

4.3.6 The contribution of gylcine and serine to network dynamics

To study the effect of changes in serine and glycine availability on FOCM dynamics, model steady states of different scenarios corresponding to sum and ratio variation were compared, as described in Section 4.2.

Table 4.11: Steady state activity of dTMP synthesis in response to the variation in cytoplasmic and nuclear overall availability of NADPH and NADP. The sum of NADPH and NADP was step-wise increased considering a scaling factor with the levels 0.5x, 1x, 2x and 4x, while the ratio NADPH:NADP stayed constant at 70:30. dTMP activity was measured in terms of steady state flux through the model reactions catalyzed by TYMS, the predicted number of produced Ts during S-phase (8h) and the ratio of required dTMP synthesis for human genome replication ($1.77{\cdot}10^9$)

NADPH+ NADP	v_{TYMS} (µM/h)		**Number of produced Ts**			**% of required Ts**		
	cyto-plasm	**nucleus**	**cyto-plasm**	**nucleus**	**total**	**cyto-plasm**	**nucleus**	**total**
0.5x	83.3	815.1	$3.77{\cdot}10^8$	$8.64{\cdot}10^8$	$1.24{\cdot}10^9$	21.31	48.81	70.12
1x	76.8	1030.7	$3.48{\cdot}10^8$	$1.09{\cdot}10^9$	$1.44{\cdot}10^9$	19.65	61.72	81.37
1.5x	51.4	1144.6	$2.33{\cdot}10^8$	$1.21{\cdot}10^9$	$1.45{\cdot}10^9$	13.15	68.54	81.69
2x	39.1	1216.0	$1.77{\cdot}10^8$	$1.29{\cdot}10^9$	$1.47{\cdot}10^9$	10.00	72.81	82.82

Table 4.12: Steady state activity of dTMP synthesis in response to the variation in partitioning of NADPH and NADP. The partitioning of NADPH and NADP was modeled by step-wise increasing the percentage of NADPH from 0% to 100% (Scenario 1 to 11, Table 4.5), while overall availability of NADPH and NADP stayed fixed at 76µM. dTMP activity was measured in terms of steady state flux through the model reactions catalyzed by TYMS, the predicted number of produced Ts during S-phase (8h) and the ratio of required dTMP synthesis for human genome replication ($1.77{\cdot}10^9$)

Scenario	v_{TYMS} (µM/h)		**Number of produced Ts**			**% of required Ts**		
	cyto-plasm	**nucleus**	**cyto-plasm**	**nucleus**	**total**	**cyto-plasm**	**nucleus**	**total**
1	0.0	0.0	0.00	0.00	0.00	0.00	0.00	0.00
2	41.8	242.7	$1.89{\cdot}10^8$	$2.57{\cdot}10^8$	$4.47{\cdot}10^8$	10.69	14.53	25.23
3	60.1	427.5	$2.72{\cdot}10^8$	$4.53{\cdot}10^8$	$7.25{\cdot}10^8$	15.38	25.60	40.98
4	70.2	575.5	$3.18{\cdot}10^8$	$6.10{\cdot}10^8$	$9.28{\cdot}10^8$	17.96	34.46	52.42
5	76.2	699.8	$3.45{\cdot}10^8$	$7.42{\cdot}10^8$	$1.09{\cdot}10^9$	19.50	41.90	61.40
6	79.5	810.4	$3.60{\cdot}10^8$	$8.59{\cdot}10^8$	$1.22{\cdot}10^9$	20.34	48.53	68.87
7	80.2	916.4	$3.63{\cdot}10^8$	$9.71{\cdot}10^8$	$1.33{\cdot}10^9$	20.52	54.87	75.39
8	76.8	1030.7	$3.48{\cdot}10^8$	$1.09{\cdot}10^9$	$1.44{\cdot}10^9$	19.65	61.72	81.37
9	69.6	1181.2	$3.15{\cdot}10^8$	$1.25{\cdot}10^9$	$1.57{\cdot}10^9$	17.81	70.73	88.54
10	62.1	1472.7	$2.81{\cdot}10^8$	$1.56{\cdot}10^9$	$1.84{\cdot}10^9$	15.89	88.18	104.07
11	57.5	3445.2	$2.60{\cdot}10^8$	$3.65{\cdot}10^9$	$3.91{\cdot}10^9$	14.71	206.30	221.01

Changing the overall availability of serine and glycine doesn't show significant effects on both cytoplasmic folate-dependent biosynthetic pathways (Figure 4.6a, AICARFT, PGT, and TYMS). Cytoplasmic dTMP activity remains mostly unaffected throughout the considered scenarios, because of quite constant steady state levels of CH2F (Figure 4.6b, Table 4.6). The main effect of changing serine+glycine can be observed in the model dynamics of SHMT-catalyzed reactions (Figure 4.6c and d). Not unexpectedly, flux through the glycine/serine-dependent bidirectional reaction R_{SHMT} : THF $\leftrightarrow$ CH2F increases with increasing availability of glycine and serine (Figure 4.6c).

For nuclear FOCM the model suggests a more pronounced trend as a consequence to changes in glycine and serine levels. Interestingly, a reduction (and not an augmentation) of overall availability of serine and glycine by 50% leads to a dTMP activity sufficient for the requirements of genome replication (Table 4.13). Indeed, increasing serine and glycine availability leads to an increase of nuclear THF concentration, while CH2F decreases notably (Figure 4.6e and f). In turn, this affects nuclear dTMP activity and the flux through TYMS decreases by 40% of initial activity (Figure 4.6f, Table 4.6). Consequently, the percentage of produced Ts required for genome replication decreases from 123% to 57% (Table 4.13).

Furthermore, we carried out a second set of *in silico* experiments, in which overall availability of serine and glycine has been considered constant and the ratio of serine and glyince has been varied according to Table 4.7. Simulations suggest that also varying the ratio of Glycine and Serine does not have a big effect on model outcomes related to cytoplasmic nucleotide synthesis (Table 4.14). Nuclear dTMP activity benefits from increasing serine availability and we observe a 2-fold increase of the flux through TYMS (Table 4.14). Consequently, overall dTMP activity increases and already at scenario 6 (glycine:serine = 50:50) we can observe dTMP rates sufficient for genome replication requirements.

Overall, the model suggests that nuclear FOCM is responsive to changes in serine and glycine availability, whereas cytoplasmic FOCM remains mostly unaffected. A possible explanation for this could be the substrate channelling, which occurs only in the nuclear compartment, and which leads to increased fluxes which in turn exhibit a higher sensitivity to changes of serine and glycine.

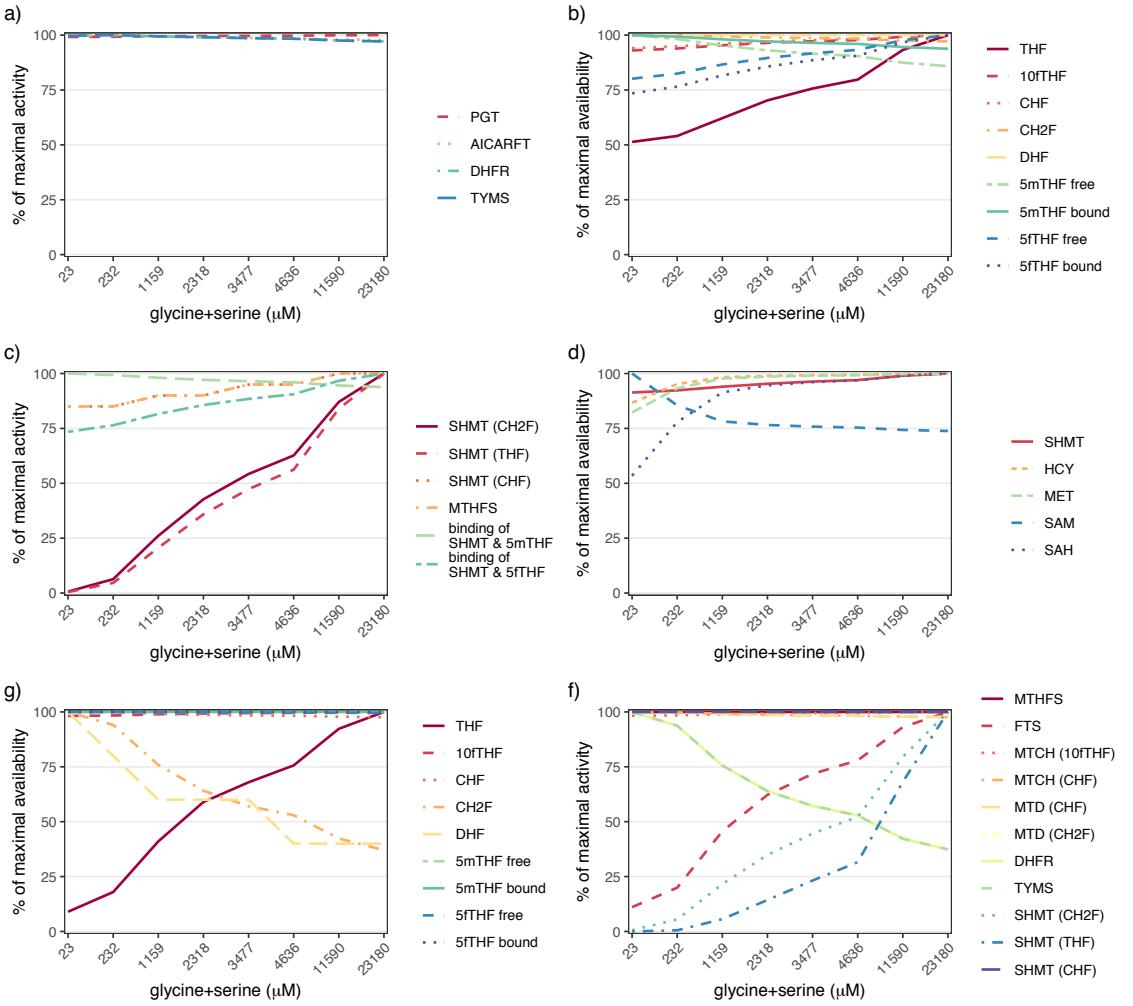

Figure 4.6: Effect of glycine and serine availability on model variables and fluxes. The provided values are computed by considering model steady states achieved by varying the total availability of glycine+serine from 23µM to 23180 µM, according to Table 4.6. They are expressed in terms of percentage of the maximal acitivity/availability over the considered range. Panel a) to d) display the results for cytoplasmic variables and fluxes, while panel e) and f) display those for the nuclear compartment. Reactions are indicated by the catalyzing enzyme, and for bidirectional reactions the direction is indicated behind the enzyme name: Enzyme (Substrate).

Table 4.13: Steady state activity of dTMP synthesis in response to the variation in cytoplasmic and nuclear overall availability of glycine and serine. The sum of glycine and serine was step-wise increased considering a scaling factor with the levels 0.01x, 0.1x, 0.5x, 1x, 1.5x, 2x and 5x, and 10x while the ratio glycine:serine stayed constant at 80:20. dTMP activity was measured in terms of steady state flux through the model reactions catalyzed by TYMS, the predicted number of produced Ts during S-phase (8h) and the ratio of required dTMP synthesis for human genome replication ($1.77 \cdot 10^9$)

glycine+ serine	v_{TYMS} (µM/h)		**Number of produced Ts**			**% of required Ts**		
	cyto-plasm	**nucleus**	**cyto-plasm**	**nucleus**	**total**	**cyto-plasm**	**nucleus**	**total**
0.01x	73.0	1747.2	$3.31 \cdot 10^8$	$1.85 \cdot 10^9$	$2.18 \cdot 10^9$	18.68	104.62	123.30
0.1x	73.2	1637.3	$3.31 \cdot 10^8$	$1.74 \cdot 10^9$	$2.07 \cdot 10^9$	18.73	98.04	116.77
0.5x	72.8	1320.9	$3.30 \cdot 10^8$	$1.40 \cdot 10^9$	$1.73 \cdot 10^9$	18.63	79.10	97.72
1x	72.5	1116.6	$3.28 \cdot 10^8$	$1.18 \cdot 10^9$	$1.51 \cdot 10^9$	18.55	66.86	85.41
1.5x	72.2	999.4	$3.27 \cdot 10^8$	$1.06 \cdot 10^9$	$1.39 \cdot 10^9$	18.47	59.84	78.32
2x	72.0	923.4	$3.26 \cdot 10^8$	$9.79 \cdot 10^8$	$1.30 \cdot 10^9$	18.42	55.29	73.71
5x	71.4	738.0	$3.23 \cdot 10^8$	$7.82 \cdot 10^8$	$1.11 \cdot 10^9$	18.27	44.19	62.46
10x	71.1	652.9	$3.22 \cdot 10^8$	$6.92 \cdot 10^8$	$1.01 \cdot 10^9$	18.19	39.10	57.29

Table 4.14: Steady state activity of dTMP synthesis in response to the variation in partitioning of glycine and serine. The partitioning of glycine and serine was modeled by step-wise increasing the percentage of serine from 0% to 100% (Scenario 1 to 11, Table 4.7, while overall availability of glycine and serine stayed fixed at 2318µM. dTMP activity was measured in terms of steady state flux through the model reactions catalyzed by TYMS, the predicted number of produced Ts during S-phase (8h) and the ratio of required dTMP synthesis for human genome replication ($1.77 \cdot 10^9$)

Scenario	v_{TYMS} (µM/h)		**Number of produced Ts**			**% of required Ts**		
	cyto-plasm	**nucleus**	**cyto-plasm**	**nucleus**	**total**	**cyto-plasm**	**nucleus**	**total**
1	72.1	921.6	$3.27 \cdot 10^8$	$9.77 \cdot 10^8$	$1.30 \cdot 10^9$	18.45	55.19	73.63
2	72.3	1025.3	$3.27 \cdot 10^8$	$1.09 \cdot 10^9$	$1.41 \cdot 10^9$	18.50	61.39	79.89
3	72.5	1116.6	$3.28 \cdot 10^8$	$1.18 \cdot 10^9$	$1.51 \cdot 10^9$	18.55	66.86	85.41
4	72.6	1200.8	$3.29 \cdot 10^8$	$1.27 \cdot 10^9$	$1.60 \cdot 10^9$	18.57	71.90	90.48
5	72.7	1281.5	$3.29 \cdot 10^8$	$1.36 \cdot 10^9$	$1.69 \cdot 10^9$	18.60	76.74	95.34
6	72.8	1361.2	$3.30 \cdot 10^8$	$1.44 \cdot 10^9$	$1.77 \cdot 10^9$	18.63	81.51	100.13
7	73.0	1442.0	$3.31 \cdot 10^8$	$1.53 \cdot 10^9$	$1.86 \cdot 10^9$	18.68	86.35	105.02
8	73.1	1525.7	$3.31 \cdot 10^8$	$1.62 \cdot 10^9$	$1.95 \cdot 10^9$	18.70	91.36	110.06
9	73.2	1613.9	$3.31 \cdot 10^8$	$1.71 \cdot 10^9$	$2.04 \cdot 10^9$	18.73	96.64	115.37
10	73.4	1708.4	$3.32 \cdot 10^8$	$1.81 \cdot 10^9$	$2.14 \cdot 10^9$	18.78	102.30	121.08
11	73.0	1811.0	$3.31 \cdot 10^8$	$1.92 \cdot 10^9$	$2.25 \cdot 10^9$	18.68	108.44	127.12

4.4 Discussion

In the preceding chapters we focused on cytoplasmic FOCM and simulations of our mathematical model contributed to the understanding of FOCM dynamics and its responsiveness to both genetic and environmental perturbations in this compartment. Recent experimental studies indicate that folate-dependent dTMP biosynthesis must also occur in the nucleus to limit genome instability [13;158]. An observation we could affirm with the initial model introduced in Chapter 1, where simulation results indicate that *de novo* dTMP synthesis rates in the cytoplasm are insufficient to support DNA synthesis during S-phase. Therefore, we extended our cytoplasmic FOCM model by building a multi-compartmental model that includes nuclear folate metabolism. A key aspect of the proposed model is that the enzymes constituting the *de novo* dTMP synthesis (SHMT, TYMS, DHFR, and MTHFD1) form a multi-enzyme complex in the nucleus [13]. These enzymes undergo SUMOylation during the G1/S phase, which leads to their nuclear import [13;84;282]. In the model, the translocation of matter to the nuclear compartment is encoded in functions and based on assumptions drawn from experimental data. A valuable extension would be to model the SUMOylation process and the following translocation expliclitly. However, for this purpose, more biological knowledge about this process has to be established.

The complex formation is proposed to accelerate the enzymatic reactions through folate substrate channelling (transfer of co-factors from the active site of an enzyme to another in the absence of diffusion), a process which, to the best of our knowledge, has not yet been modeled in FOCM literature. Experimental evidence indicates that *de novo* dTMP synthesis is effective only when the enzymes are present in the multi-enzyme complex within the nucleus and is therefore important to prevent uracil misincorporation into DNA and genome instability [158].

Mathematical models provide a feasible framework to study this aspect and *in silico* simulations enable rapid testing of assumptions related to this and other critical factors. Indeed, by including nuclear FOCM in the model we investigated factors that modify dTMP synthesis, including the effect of multienzyme complex formation, substrate channelling, nuclear MTHFS availability, enzyme expression levels, folate partitioning between the cytosol and nucleus, as well as availability of glycine, serine, NADPH and NADP.

We observed that modeling dTMP synthesis in the nucleus must include the occurrence of multienzyme complex formation and substrate channelling of folate substrates among the enzymes: only when these two aspects were taken into account, adequate levels of dTMP synthesis activity were reached. Model simulations of the standard scenario predicted that *de novo* dTMP synthesis may produce up to 85% of needed T. This portion can be increased when enhanced nuclear folate levels are considered, emphasizing the influence of folate partitioning between the two compartments on overall functionality of dTMP synthesis. Notably, already a very small increase of nuclear folate availability (from 10% to 11.6% of total folate) is sufficient to meet adequate levels of dTMP synthesis. The observed "gap" between *de novo* dTMP synthesis rates and T needs for replication can be explained by yet another aspect. There exist two distinct pathways for thymidine nucleotide synthesis and besides the *de novo* dTMP synthesis also a salvage pathway recovers thymidine nucleotides. The salvage pathway, which could be included in a possible extension of the proposed model, involves the conversion of thymidine to dTMP, and occurs in the cytoplasm/nucleus catalyzed by TK1 and in the mitochondria by TK2 [201]. Salvage pathway

synthesis of dTTP is not sufficient to sustain nuclear or mitochondrial DNA replication [293] and therefore requires folate-dependent *de novo* dTMP synthesis to maintain genomic integrity during cell division [85]. Presumably, most dTTP required for DNA replication is synthesized from dUMP by the *de novo* dTMP biosynthesis pathway [293], matching our predicted contribution of *de novo* dTMP synthesis around 85%.

An open question, which still needs definitive experiments, is related to the identification of the relative contribution of nuclear and cytosolic dTMP synthesis to overall dTMP synthesis. Our model clearly highlights the role of nuclear dTMP synthesis and predicts that during S-phase around 79% of overall Ts are synthesized in the nucleus. A possible explanation for this high contribution is given by the fact that the complex formation and substrate channelling are occurring solely in the nucleus. Without these two features, cytoplasmic *de novo* dTMP synthesis remains the main contributor to overall dTMP synthesis activity. Moreover, nuclear translocation uncouples *de novo* dTMP synthesis from *de novo* purine biosynthesis and homocysteine remethylation in the cytoplasm, thereby erasing competition for folate cofactors [142].

In Chapter 1, model simulations predicted that cytoplasmic *de novo* dTMP synthesis produces only 60% of required Ts and is therfore not sufficient for genome replication. In the extended model, we could observe that the inclusion of the nuclear compartment reduces the cytoplasmic contribution to overall dTMP synthesis even further by predicting that around 20% of required Ts are produced in this compartment. This highlights the importance of nuclear dTMP synthesis to prevent uracil misincorporation during DNA replication or repair.

The simulation results provide a computational indication that the translocation to the nuclear compartment is essential to prevent uracil incorporation in DNA by protecting biosynthesis from variability in the network, introduced e.g. by lack/increase of NADPH, NADP levels.

In summary, inclusion of the nuclear compartment in the mathematical model has provided new insights into the functioning of the FOCM network. The model confirms that accounting for the kinetic effects of nuclear multienzyme complex formation and substrate channelling is essential for the functioning of *de novo* dTMP synthesis. *In silico* simulations also indicate that the nuclear compartment plays an important role for regular cell replication and DNA repair.

Author Contributions

K. Misselbeck (K.M.), L. Marchetti (L.M.), C. Priami (C.P.), P.J. Stover (P.J.S.), and M.S. Field (M.S.F.) contributed to the experimental design. K.M., L.M. developed the computation model and carried out all simulations, while M.S.F. and P.J.S. provided biological insight and extracted kinetic variables from the literature. C.P., P.J.S., and M.S.F. provided overall guidance of the project.

Part II

Metabolic Syndrome

Chapter 5

Metabolic syndrome: identification of deregulated pathways and drug effects by network analysis.

Metabolic syndrome is a pathological condition characterized by obesity, hyperglycemia, hypertension, elevated levels of triglycerides and low levels of high-density lipoprotein cholesterol that increase the risk of cardiovascular diseases and type 2 diabetes. Even though numerous predisposing genetic risk factors have been identified for each component, the biological mechanisms underlying this complex phenotype are not fully elucidated. In this chapter we introduce a systems biology approach based on network analysis, integrating drug-related and metabolic-syndrome genes. Tissue-specific regulatory networks were constructed to investigate the biological processes deregulated in the disorder and subsequently identify new candidate treatments in a drug-repurposing manner. To this end, a proximity score describing the interaction between drugs and pathways was defined by combining topological and functional similarities. Our results highlight a prominent role of the immune system in metabolic syndrome and suggest a potential use of the BTK inhibitor ibrutinib as novel pharmacological treatment.

5.1 Introduction

Metabolic Syndrome (MetSyn) is a highly prevalent pathological condition defined by a complex clustering of comorbidities that increases the risk of cardiovascular diseases and type 2 diabetes mellitus. The risk factors commonly associated with MetSyn are abdominal obesity, hyperglycemia, hypertension, elevated levels of triglycerides and low levels of high-density lipoprotein (HDL) cholesterol. According to the criteria proposed by the main organizations involved in the study of MetSyn, this clinical condition can be diagnosed when three of these five metabolic abnormalities are present simultaneously [3]. Additional components such as chronic pro-inflammatory and pro-thrombotic states have been repeatedly implicated in MetSyn [5], highlighting the presence of numerous contributing risk factors.

While lifestyle changes are highly effective in an early phase of this metabolic disorder, pharmacological treatments are frequently required to control it in more advanced stages [102]. Currently,

the pharmacological interventions are mostly directed towards the single MetSyn components separately, raising the problem of polypharmacy [102]. Moreover, although an altered function of adipocytes is recognized as a pivotal driver of the observed metabolic dysregulation [18], most of the drugs approved for obesity act on the central nervous system while the pathways active in the unhealthy adipose tissue are less explored [121;137;235].

The increasing prevalence of MetSyn worldwide and the limited understanding of the pathophysiological mechanisms of MetSyn give rise to the need to study the underlying biological pathways and to develop more efficacious treatment strategies

An effective strategy to reduce time and cost of drug development is drug repurposing (or repositioning), which identifies new therapeutic applications of already approved drugs. For example, galantamine, an approved drug for Alzheimer's disease, was recently suggested as a candidate for MetSyn therapy [50]. The growing availability of high-resolution data allows researchers to establish new computational approaches to systematically investigate drug repurposing candidates [146;206]. For example, signature-based methods based on the Connectivity Map (CMap) [139;140] and LINCS data [249] allow to identify promising candidates by comparing the transcriptomic profiles of drugs and diseases [45;114;176;231;232;290]. Usually, algorithms based on this approach don't consider the interactions among the molecular components composing the expression profiles.

Network-based analysis is a method of choice to study *in silico* the complexity of biological systems and to evaluate the interactions among the different players involved, while serving as a powerful tool to link pharmacological and disease data [59]. Recent systems biology approaches based on network analysis successfully investigated new indications for existing drugs [44;104;153], predicted new potential anticancer treatments [46;267] and identified new promising targets [71;117;145].

In this chapter we present a systems biology approach based on network integration of genomic data, text mining results, drug expression profiles and drug target information to identify the underlying molecular mechanisms of MetSyn and to explore possible novel therapeutic strategies. This method identifies potential new therapeutic applications of already approved drugs using a proximity score, which integrates a network-based distance and a functional similarity measurement.

5.2 Results

5.2.1 Computational framework overview

To obtain a systems pharmacology view of MetSyn, we devised a network-based approach that identifies functional disease modules and connects them with drug targets and drug-perturbed genes. The analytical workflow consists of three interconnected parts as shown in Figure 5.1. The list of trait-associated genes was established starting from the results of published genome-wide association studies (GWAS) and from additional literature related to metabolic syndrome (Figure 5.1, step 1). Interconnections among disease genes were derived from existing biological networks. In particular, tissue-specific integrated networks for adipose tissue, liver and skeletal muscle were constructed by merging the HIPPIE protein-protein interaction (PPI) network [2] and the recently published Regulatory Circuits transcriptional regulatory network [160] (Figure 5.1, step 2). Mapping the gene set to the networks allowed us to identify tissue-specific trait-related modules

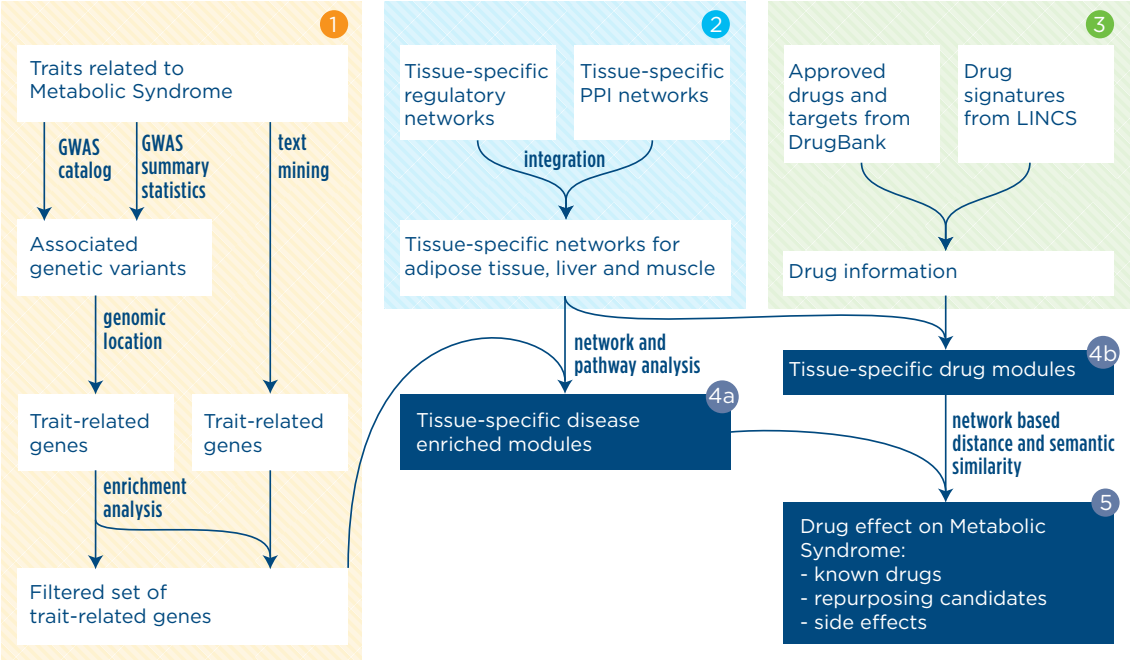

Figure 5.1: Schematic illustration of the computational framework. Step 1) MetSyn-related genes are identified by combining GWAS results and literature findings followed by a filtering step based on gene-set enrichment analysis. Step 2) Tissue-specific networks are constructed by integrating transcriptional regulatory networks from [160] and PPI networks from HIPPIE db [2]. Step 3) Drug information is retrieved from DrugBank [281] and LINCS database [249]. Step 4a) and b) Tissue-specific MetSyn and drug modules are established using network analysis. Step 5) To measure drug effects, a proximity score between drug and MetSyn modules is computed on the basis of network distance and semantic similarity.

(hereafter called MetSyn modules), for which pathway enrichment analysis provided insight into the associated biological processes (Figure 5.1, step 4a). The impact of existing drugs on MetSyn was studied by mapping drug targets and drug modulated genes on the networks in order to build drug modules (Figure 5.1, step 3 and 4b). Both drugs approved for MetSyn-related conditions and for other diseases were included in the study. This allowed us to gain insight into existing treatments and, at the same time, identify candidates for drug repositioning. To investigate the relationship between drug modules and MetSyn modules, we defined a proximity score that combines network-based distance with semantic similarity (Figure 5.1, step 5). A drug obtains a high score if its module is close to the MetSyn module and if the genes in the drug module and the genes in the MetSyn module have a similar biological function. A comparison of our approach with previously published network-based methods for drug repurposing can be found in Appendix E, Section E.1.

5.2.2 Identification of genes associated with MetSyn

A widely used approach to identify genetic variants associated with common traits and diseases is the use of GWAS that over the past ten years have been applied to hundreds of phenotypes. The increasing availability of GWAS summary data permits the development of methodologies aiming at understanding the biology of phenotypes of interest starting from association results [266].

Thus, we devised a multi-step procedure to identify genes associated with MetSyn starting from the results of GWAS of relevant traits. Since the genetic variants identified by GWAS do not directly yield specific gene targets or molecular mechanisms, our workflow includes a pathway enrichment step to identify the altered biological functions. Moreover, given the incompleteness of the currently available GWAS data in explaining the heritability of traits, an external source of information (i.e. literature-derived knowledge) was included in the study. Working along this line, to generate a list of MetSyn-related genes, we combined 3 different data sources. First, we queried the GWAS catalog for studies related to metabolic syndrome and retrieved those SNPs reaching genome-wide significance (p-value $< 5 \cdot 10^{-8}$).

Despite being located mainly in introns (Appendix E, Figure E.3a), the identified susceptibility variants showed a regulatory potential since they are enriched in SNPs located in genomic regions of epigenetic chromatin marks when compared with non-selected genome-wide common SNPs (Appendix E, Figure E.3b–d). To assign the association signal to a gene, we retrieved the genes located in the genomic region of the tagging SNP.

Second, we included the genes derived from summary statistics of 15 GWAS focused on MetSyn-related traits (Appendix E, Table E.6). Finally, the GWAS-derived genes were combined with a set of genes derived from a text mining analysis performed on PubMed abstracts searching MetSyn related terms co-occurring with gene names, as described in Section 5.4.

Given the heterogeneity of the data sources (GWAS catalog: top-scoring trait-associated SNPs and related genes, GWAS summary statistics: gene-level scores and text mining: genes co-occurring with MetSyn-terms) we devised a custom approach to combine and filter them. We performed a gene-set enrichment analysis of the GWAS genes selecting gene ontology biological processes and pathway databases available in EnrichR [136] to obtain pathway-level biological knowledge and we selected the genes belonging to at least one significant pathway (Appendix E, Table E.10). In total, we were able to identify 630 genes associated with MetSyn (Figure 5.2a, Appendix E, Table E.11, Figure E.4).

Interestingly, we identified pathway categories such as sugar metabolism, lipid metabolism and fat storage as significantly enriched (Figure 5.2b). This annotation supports the relevance of the selected genes, as they match the pathophysiological components of MetSyn, including hyperglycemia and dyslipidemia.

5.2.3 Tissue-specific disease modules

According to the pathological phenotypes associated with MetSyn, adipose, liver and skeletal muscle were selected as trait-relevant tissues and the corresponding tissue-specific background networks were generated by combining regulatory networks and PPI networks as described in Section 5.4 and displayed in Figure 5.3a. The tissue-specific regulatory networks were directly obtained from regulatory circuits [160], a resource that provides transcription factor–gene interactions inferred from the FANTOM5 data [151]. On the other hand, the tissue-specific PPI networks were created from HIPPIE interactions [2;224] by restricting to proteins expressed in the relevant tissue based on GTEx data [34;252].

The resulting network for adipose tissue contains 886 nodes and 9152 edges, the network for liver tissue 1544 nodes and 15846 edges, and the network for skeletal muscle tissue 1106 nodes and

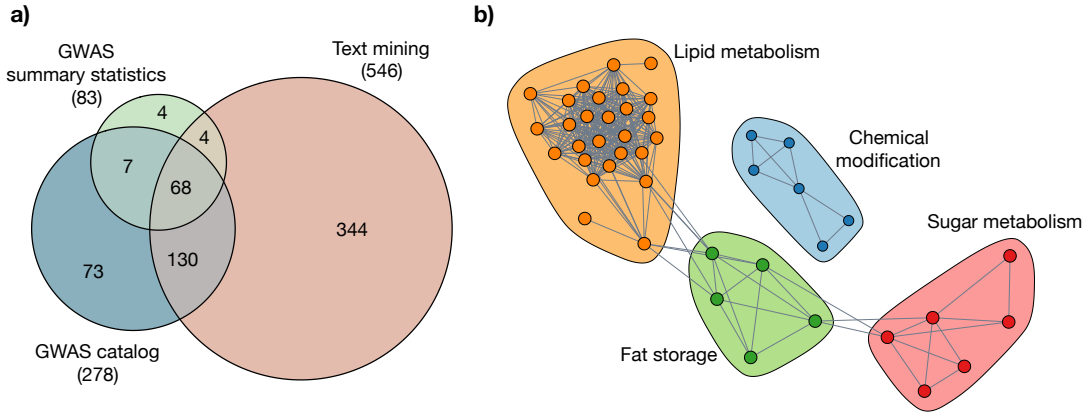

Figure 5.2: Identification of MetSyn-related genes. a) Venn diagram showing the overlap among MetSyn genes identified using GWAS catalog, GWAS summary statistics and text mining. b) Pathway enrichment map showing shared gene content among the pathways enriched in MetSyn genes. Each node corresponds to a pathway and edges between pathways indicate the presence of shared genes. Colors identify the membership to communities, detected using random walk clustering algorithm.

10555 edges. In total, these networks included 220 MetSyn genes, of which 81 were shared among the three networks (Figure 5.3b). The liver tissue has the highest number of MetSyn genes not present in the other networks (Figure 5.3b).

To identify trait relevant network subparts, we tested network modules for their overrepresentation in MetSyn genes. For both the liver and muscle network, three significant trait-related modules could be identified, whereas we found two significant modules for the adipose tissue. Reactome[74] pathway enrichment analysis of these network modules further allowed linking biological functions to the tissue-specific MetSyn modules.

For example, in the adipose tissue network, the most significantly enriched pathways of module 1 are related to cellular responses to external stimuli and immune function (Cellular responses to heat stress: adjusted p-value $9.65 \cdot 10^{-7}$; immune system: adjusted p-value $9.65 \cdot 10^{-7}$). Instead, pathways related to metabolism regulation resulted enriched in module 2 (PPARA activates gene expression: adjusted p-value $1.25 \cdot 10^{-11}$; regulation of lipid metabolism by Peroxisome proliferator-activated receptor alpha (PPARalpha): adjusted p-value $1.82 \cdot 10^{-11}$). The list of all significant pathways for the 3 tissue-specific networks can be found in Appendix E, Table E.12.

An overview of the module-related biological functions in all networks was obtained using the Top Level Pathways from the Reactome database[74] (Figure 5.4). Overall, the resulting pathways show an overlap across tissues, highlighting the overrepresentation of genes involved in signal transduction and gene expression. Moreover, our results suggest that the immune system plays a considerable role for MetSyn; across all three networks a module with a high number of immune-related genes and pathways was detected (Figure 5.4 and Appendix E, Table E.12), in agreement with previous reports[10;195]. The contribution of the different data sources (GWAS vs. text mining, PPI vs. regulatory networks) to the identification of MetSyn modules is described in Appendix E, Section E.2.

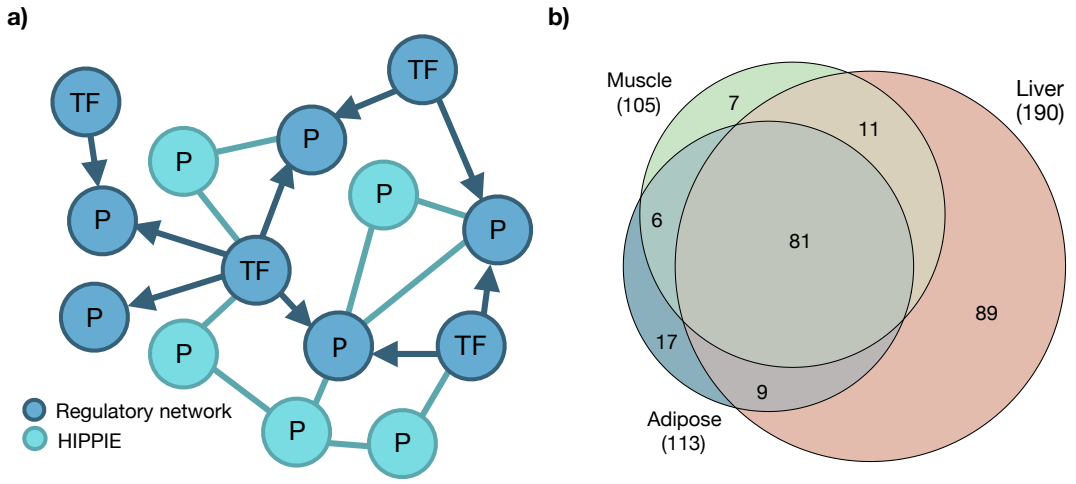

Figure 5.3: Network construction and disease module identification. a) Tissue-specific networks were constructed by integrating transcriptional regulatory networks consisting of interactions between transcription factors and genes (blue nodes) and PPI networks including interactions among proteins (turquoise nodes). For each tissue, the integrated network was built starting from the high-evidence associations from the regulatory network and extended with high-score interactions from the PPI network. b) Venn diagram of shared MetSyn genes among the three tissue-specific networks.

5.2.4 Drug Repurposing

To identify drugs potentially affecting MetSyn pathways, we selected approved drugs from DrugBank [280] which have a target in at least one of the three networks (Appendix E, Figure E.5). The interplay of MetSyn modules and 183 drugs was evaluated by computing the proximity score as shown in Figure 5.5 and described in Section 5.4. The score is based on topological properties of the network and functional similarity of the proteins. The contribution of these two components is described in Appendix E, Section E.2. Drugs with a significant score point toward a possible disease indication or drug side effect. For the adipose network, this analysis resulted in a list of 28 significant drugs, for the liver network 31 significant drugs were identified, while for the muscle network the analysis resulted in 50 significant drugs (Appendix E, Tables E.13, E.14, E.15).

To test the effectiveness of our approach, we checked if drugs with known indication for adiposity, which has a key role in leading the metabolic disturbances associated with MetSyn, were identified by our scoring system [18]. Bezafibrate, clofibrate, fenofibrate, gemfibrozil, mifepristone, pioglitazone were used for the evaluation process and 3 of them (pioglitazone, mifepristone and fenofibrate) were significance considering a threshold of 95 %. After lowering the significance threshold to 85%, all six drugs were significant (Appendix E, Section E.3.

After having obtained the preliminary list of significant predictions, we performed a filtering and prioritization analysis to identify the most promising repurposing candidates (Appendix E, Figure E.6. First, to exclude drugs with undesirable side effects, we evaluated information about contraindications from the DrugCentral platform [262] (Appendix E, Table E.16). For the adipose results we excluded 7 drugs, while for liver and muscle 12 and 24 drugs were excluded, restricting the list of repurposing candidates to 21, 19, and 26 drugs, respectively (Appendix E, Tables E.13,

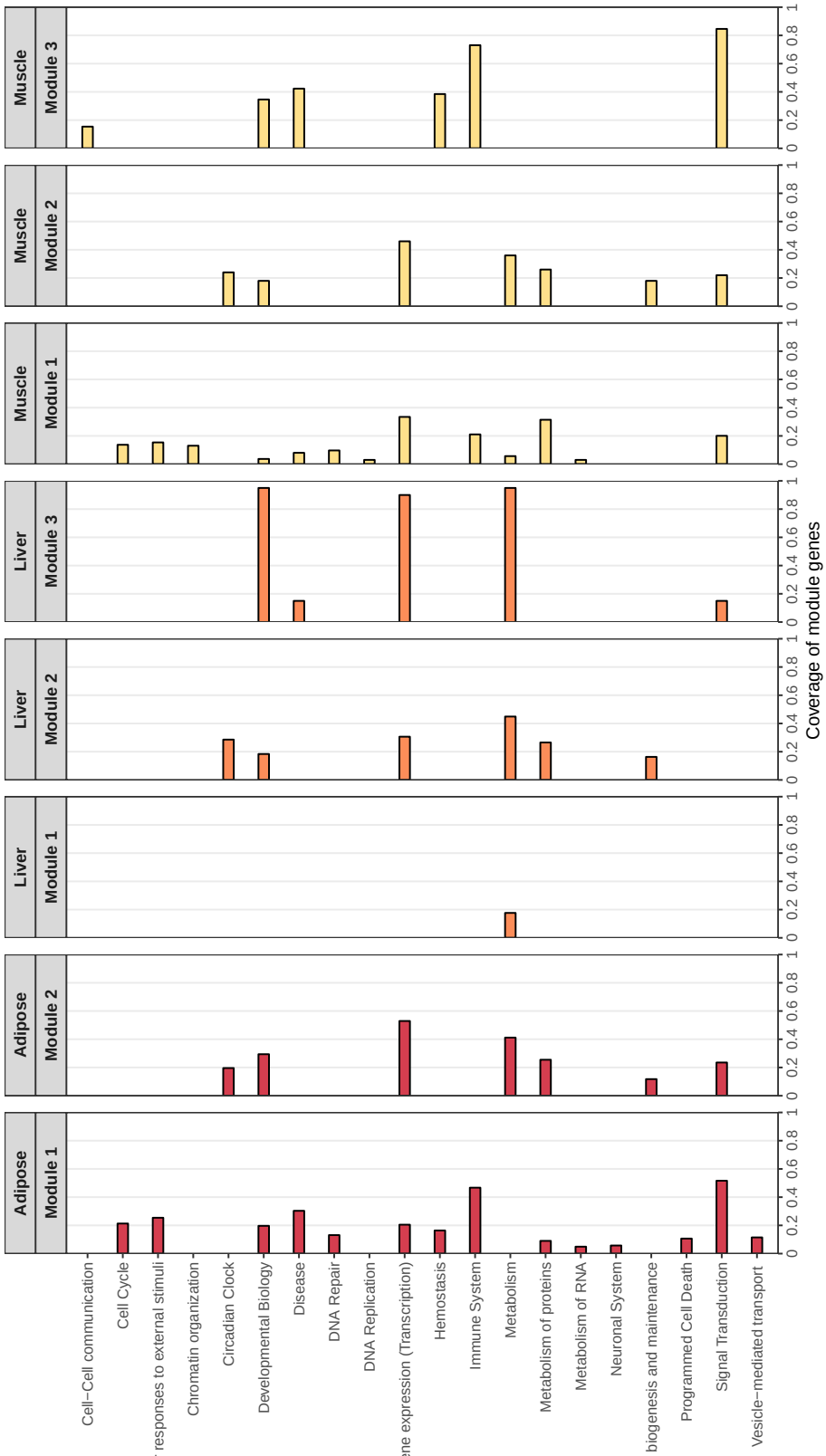

Figure 5.4: Functional annotation of the disease modules. For each network and disease module the coverage of module genes by significantly enriched Reactome pathways, grouped according to the TopLevel pathway classification, is presented.

E.14, E.15, column "SE"). Second, we filtered those drugs by focusing on their targets that were investigated using data from the OpenTargets platform [133]. For the adipose results, among the targets of the 21 drugs without known MetSyn-related side effects, 10 (AR, EGFR, HDAC6, IKBKB, NR3C1, PGR, PPARA, PPARG, RXRG, and VDR) have already been investigated for therapeutic interventions related to MetSyn and therefore we excluded them from further analyses (Figure 5.6). For liver and muscle, 6 (NR3C1, NR3C2, PPARA, PPARD, PPARG, RXRG) and 8 targets (ADRB2, AR, EGFR, ESR1, IKBKB, NR3C1, PPARA, RXRG) were removed, respectively (Appendix E, Figures E.7 and E.8, Tables E.13, E.14, E.15, column "OT").

After this filtering procedure, we identified the following drugs as having a potential novel therapeutic application for MetSyn: adapalene, afatinib, alitretinoin, belinostat, bosutinib, crizotinib, dequalinium, doconexent, erlotinib, ibrutinib, lapatinib, nintedanib, panobinostat, rucaparib, ruxolitinib, tamibarotene, tofacitinib (Table 5.1).

A final prioritization step was then carried out to evaluate if the tissue expression of the targets was concordant with the disease manifestations. Among the 18 targets of the drugs, only bruton tyrosine kinase (BTK), the target of Ibrutinib, and nuclear receptor subfamily 1 group I member 2 (NR1I2), the target of erlotinib, showed a tissue-specific expression relevant for MetSyn. According to the Human Protein Atlas [260], GTEx [252] and FANTOM5 [151] databases, *BTK* gene expression is consistently enhanced in immune-related tissues, and *NR1I2* expression is enriched in liver, while the other targets did not show any relevant tissue-specificity (Appendix E, Tables E.7, E.8, and E.9). NR1I2 is a nuclear receptor that regulates hepatic detoxification, and is involved in glucose and lipid metabolism. Recent studies indicate that an activation of the protein could contribute to the development of MetSyn and diabetes [106]. Since erlotinib is an agonist of NR1I2, we concluded that the significance of the proximity score in the liver network could be explained by this finding.

On the other hand, the BTK inhibitor ibrutinib is currently FDA-approved for the treatment of B cell cancers and the chronic graft-versus-host disease while ongoing clinical trials evaluate the use of BTK inhibitors in autoimmune diseases [52].

Given the important role of inflammation in the alteration of adipose tissue biology in obese patients, we investigated the relationship between BTK and the immune system in obesity using public datasets. According to ImmGen mouse RNAseq data [109], the immune cell populations expressing high levels of Bruton tyrosine kinase transcripts are B cells and myeloid lineage cells such as neutrophils and macrophages (Appendix E, Figure E.9). Interestingly, gene expression analysis of macrophages derived from adipose tissue of obese type II diabetic subjects [33] showed higher *BTK* expression as compared to macrophages of obese non diabetic subjects (t-test p-value 0.026) (Figure 5.7a). Expression data from a mouse model deficient in GRP120 [223], a receptor for long-chain free fatty acids involved in nutrient sensing and body weight regulation, showed higher *Btk* expression in white adipose tissues when fed with a high fat diet (HFD) compared with normal diet (Figure 5.7b). Accordingly, the estimated composition of the infiltrating immune cells in adipose tissue of HFD-fed GPR120-mutated mouse, computed with CIBERSORT [181], revealed a clear increase in macrophages (Figure 5.7d).

Since the macrophage-related inflammation in obese diabetic mice has been associated with inflammasome-dependent IL-1ß production, we also evaluated the levels of *Btk* mRNA in inflammasome-compromised mouse models [253]. In white adipose tissue from HFD-fed Caspase-1 null

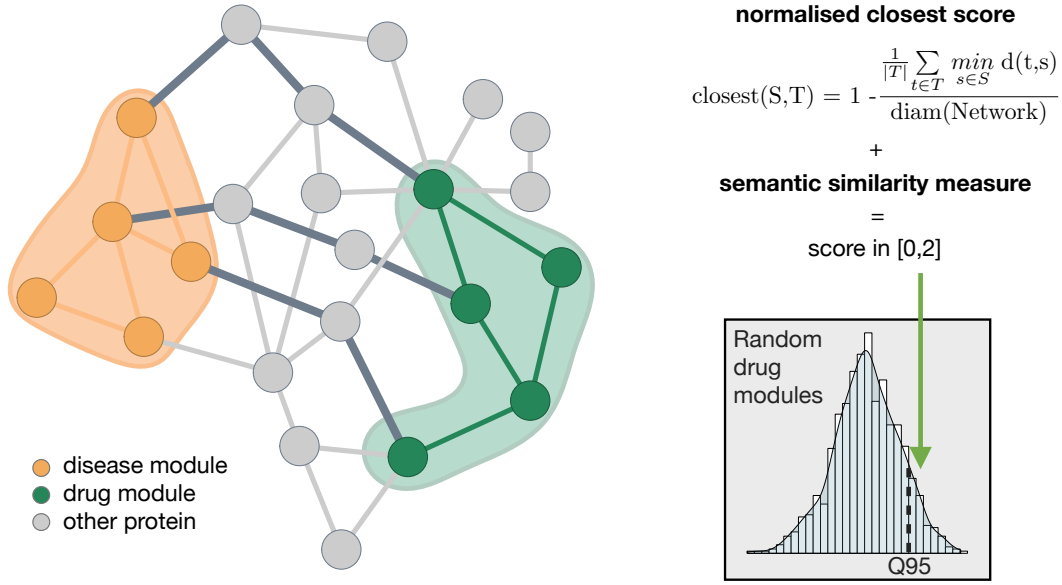

Figure 5.5: Illustration of the score calculation to evaluate the drug-disease interplay. The score combines network distances and functional similarity between proteins in the drug module (green nodes) and proteins in the disease module (orange nodes). The network score assesses the shortest path lengths connecting each protein of the drug module to the nearest protein in the disease module (dark gray edges) while the semantic similarity measure evaluates the functional similarity between the modules. To test the score significance, the distance between drug and disease modules is compared to a reference distribution of scores computed with drug modules randomly chosen from the network.

mouse, Bruton tyrosine kinase expression was lower as compared to wild type mice fed with the same diet (Figure 5.7c).

In addition, mice lacking the inflammasome adaptor protein ASC did not show a reduction in *Btk* expression. This is in agreement with the observations made by Stienstra et al., where the presence of macrophage infiltration in adipose tissue of ASC null mice was observed by immunohistochemistry [237]. In human samples, we observed the same pattern, although the differences are less pronounced (Appendix E, Figure E.10).

5.3 Discussion

The computational pipeline proposed in this Chapter depicts a systems biology approach to study the biological processes involved in MetSyn and to determine potential new pharmaceutical treatments in a drug repurposing manner. Given their high interconnectivity and the multifactorial aetiology, metabolic disorders related to MetSyn are particularly suited to a system-level analysis that integrates multiple data types [156]. Furthermore, the identification of new drug therapies is highly valuable to counteract serious complications, such as cardiovascular disease, that often follow MetSyn.

Although in this study we focused on MetSyn, the proposed method can be applied to other diseases. The only requirement is a list of disease genes that can be derived from many sources

Table 5.1: List of the drugs identified as possible repurposing candidates. For each drug, the DrugBank ID and name, the selected target to build the drug module, its action, the calculated score and the associated MetSyn module are listed.

Drug ID	Drug Name	Action	Target	Score	Module	Network
DB00210	Adapalene	agonist	RARG	1.6618	2	Adipose
			RXRB	1.6925	2	Adipose
DB00523	Alitretinoin	agonist	RARG	1.6661	2	Adipose
DB04209	Dequalinium	antagonist, inhibitor	XIAP	1.7909	1	Adipose
DB03756	Doconexent	activator	RXRA	1.7103	2	Adipose
			RXRB	1.6660	2	Adipose
DB09053	Ibrutinib	inhibitor	BTK	1.6147	1	Adipose
DB12332	Rucaparib	antagonist	PARP1	1.6678	2	Adipose
DB08877	Ruxolitinib	inhibitor	JAK1	1.7721	1	Adipose
DB04942	Tamibarotene	agonist	RARA	1.7222	2	Adipose
DB00210	Adapalene	agonist	RXRB	1,6871	2	Liver
DB03756	Doconexent	activator	RXRA	1,7047	2	Liver
			RXRA	1,4948	3	Liver
			RXRB	1,6819	2	Liver
			RXRB	1,5372	3	Liver
DB00530	Erlotinib	agonist	NR1I2	1,6684	2	Liver
DB09079	Nintedanib	inhibitor	FGFR3	1,4473	2	Liver
DB04942	Tamibarotene	agonist	RARA	1,6667	2	Liver
DB00210	Adapalene	agonist	RXRB	1,6759	2	Muscle
			RARG	1,6676	2	Muscle
DB08916	Afatinib	inhibitor	ERBB2	1,6765	3	Muscle
DB00523	Alitretinoin	agonist	RARG	1,7756	1	Muscle
			RARG	1,6643	2	Muscle
			RXRB	1,662	2	Muscle
			RARA	1,681	2	Muscle
DB05015	Belinostat	inhibitor	HDAC1	1,754	1	Muscle
			HDAC2	1,7615	1	Muscle
			HDAC4	1,8027	1	Muscle
DB06616	Bosutinib	inhibitor	ABL1	1,609	3	Muscle
DB08865	Crizotinib	inhibitor	MET	1,7883	1	Muscle
DB03756	Doconexent	activator	RXRA	1,7134	2	Muscle
DB01259	Lapatinib	antagonist	ERBB2	1,8184	1	Muscle
DB06603	Panobinostat	inhibitor	HDAC3	1,611	2	Muscle
DB08877	Ruxolitinib	inhibitor	JAK1	1,6745	3	Muscle
			JAK2	1,714	3	Muscle
DB04942	Tamibarotene	agonist	RARA	1,6708	2	Muscle
DB08895	Tofacitinib	antagonist	JAK1	1,71	3	Muscle
		inhibitor	JAK2	1,692	3	Muscle

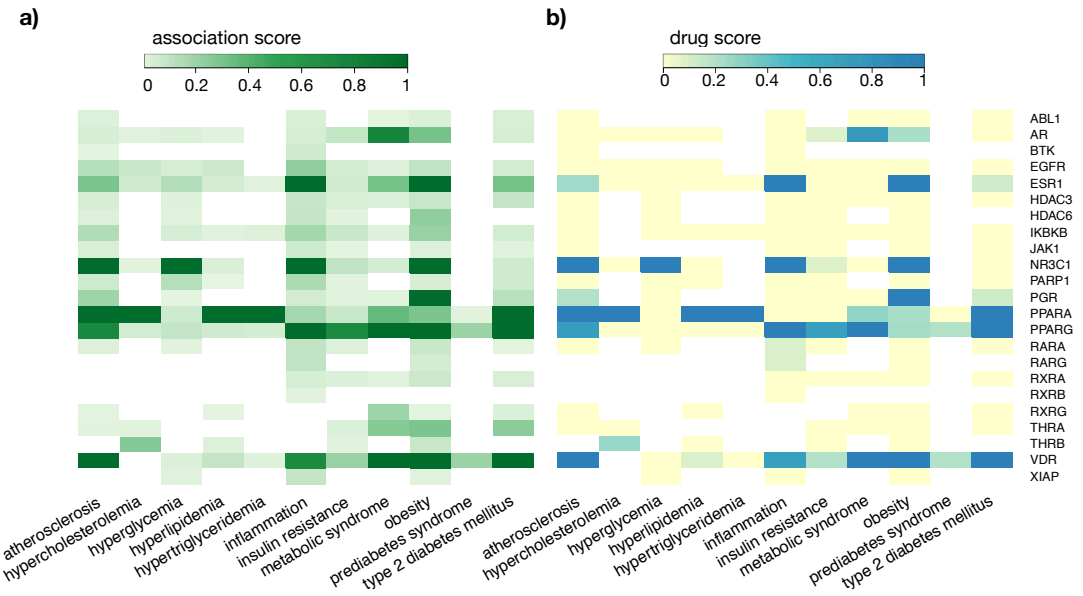

Figure 5.6: Association between the drug targets in the adipose network and MetSyn-related traits based on the scores provided by OpenTargets [133]. a) Heatmap of total association score and b) heatmap of association score based on ChEMBL information about drugs approved for marketing by FDA or under evaluation in clinical trials.

(e.g. from GWAS results or gene/protein expression profiles). Even in the case of Mendelian diseases, for which repurposing is an important opportunity to identify treatment strategies [62;95], the pipeline can be applied using information about affected pathways as input. Moreover, the network can be adapted for the specific requirements of the study. For example, if a disease-specific gene regulatory network is available, this can replace the GTEx-derived, tissue-specific networks selected here. On the other hand, the method can also provide useful insights for a specific drug of interest once network modules for various diseases are defined.

Our workflow relies on the integration of disease and drug data through network-based analysis, which has been shown as a promising methodology for drug repurposing [44;46;71;104;117;145;153;267]. To identify MetSyn genes, we exploited GWAS and text mining results. The use of genetic evidences has been demonstrated as a powerful resource to support drug discovery, both for pointing out disease-related pathways and for the identification of new drug targets [33]. Indeed, the targets of many drugs approved before the GWAS era are located in GWAS risk loci, thus supporting further exploration of this data [180;223]. Moreover, the integration of GWAS results with literature findings allowed us to take into account the knowledge derived from additional sources, such as functional studies, and thus to have a more comprehensive view of the pathways involved.

The tissue-specific integrated networks were generated by merging a protein-protein interaction network and a transcriptional regulatory network. While the first describes known associations between proteins, such as the formation of protein complexes or kinase-substrate interactions, regulatory networks are fundamental to keep into account the regulation exerted by transcription factors on gene expression. This is of particular importance when studying complex traits, because the regulation of gene expression in a tissue-specific manner is a key element in determining the

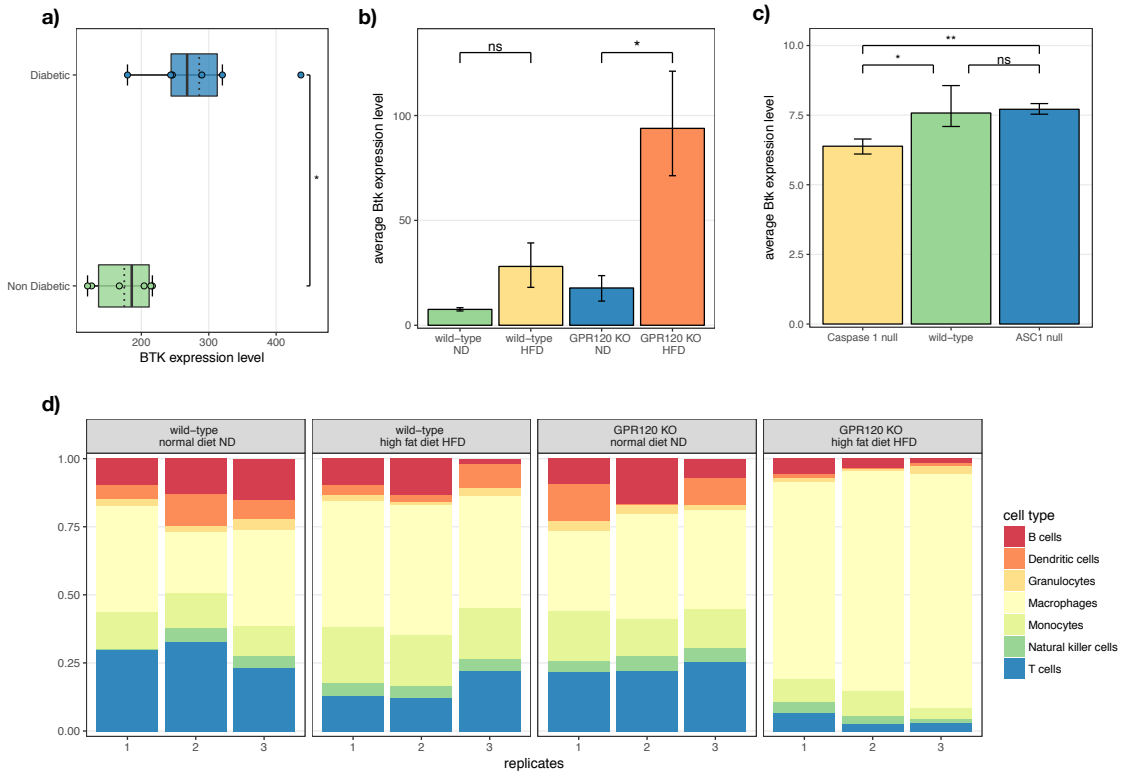

Figure 5.7: *BTK* expression in public datasets. a) Boxplots showing *BTK* gene expression in macrophages of diabetic and non-diabetic subjects. b) Barcharts showing the average gene expression level of *Btk* in adipose tissue for wild type and GPR120 KO mice fed with normal diet (ND) or high fat diet (HFD). c) Barcharts showing the average gene expression level of *Btk* in adipose tissue for wild-type, Caspase 1 null and ASC1 null mice. d) Estimated relative fraction of different immune cells in adipose tissue calculated via Cibersort for wild type and GPR120 KO mice fed with normal diet (ND) or high fat diet (HFD) in adipose tissue. In a)-c), statistical significance is denoted as follows: ns: not significant ($p > 0.05$), * : $0.01 < p \leq 0.05$, ** : $0.01 \leq p < 0.001$.

pathological phenotype[88;253]. Moreover, it has been shown that several GWAS traits show higher connectivity in regulatory networks than in other types of networks[160], further supporting our choice of including regulatory interactions in the background network.

A key point of the performed network analysis has been the definition of a proximity score that attempts to efficiently connect drug modules and MetSyn modules. In addition to the network-based distance between MetSyn and drug genes, our scoring system takes into account the similarity of biological gene functions. This approach increases the ability to identify potentially effective new therapeutic strategies because it adds direct biological knowledge to the topological information derived from the network.

The results obtained from the analysis of the adipose network remark the key role exerted by inflammation in obesity and suggest the adoption of anti-inflammatory therapies. This is in agreement with a growing body of literature supporting the use of anti-inflammatory agents to treat the chronic low-grade inflammation accompanying metabolic-related pathological conditions[126;216].

In particular, our results suggest Ibrutinib as the most promising drug repurposing candidate. ibrutinib is a small molecule that inhibits BTK, a protein well-known for its essential role in B cell development and maturation [270]. It has been FDA-approved for the treatment of mantle cell lymphoma, chronic lymphocytic leukemia, Waldenström's Macroglobulinemia and chronic graft versus host disease [147]. Moreover, BTK inhibitors are under investigation for the treatment of autoimmune diseases, such as multiple sclerosis (ClinicalTrials.gov Identifier: NCT02975349) and rheumatoid arthritis (ClinicalTrials.gov Identifier: NCT03233230) and showed promising effects in type 1 diabetes [124]. In addition to B cells, *BTK* is also expressed by macrophages, known key players in the development of the obesity-related chronic inflammation and insulin resistance. At the molecular level, BTK is involved in the regulation of macrophage Toll-like receptor-mediated immune response [97;150] and is essential for the activation of the NLRP3 inflammasome and IL-1ß production [118]. Importantly, NLRP3 activity has been linked to obesity and insulin resistance both in human and mouse studies [217]. In addition to macrophages, B cells themselves have been implicated in adipose tissue inflammation and insulin resistance [277;278], providing additional support for BTK involvement in obesity-related inflammation. However, experimental studies are needed to investigate the effect of BTK inhibitors such as ibrutinib, in obesity and the possible benefits for patients with metabolic syndrome.

Overall, this work describes an innovative methodology based on the integration of genetic and expression data with previous biological knowledge, which enables the identification of drug repurposing candidates for complex diseases. The application of the pipeline to MetSyn let us to identify the inhibition of BTK by ibrutinib as a promising repurposing strategy.

5.4 Materials and Methods

5.4.1 List of genes associated with MetSyn

To establish a list of genes associated with MetSyn, we used three different sources: GWAS catalog, GWAS summary statistics and text mining.

Data-mining of the GWAS catalog

Results of published genome-wide association studies were obtained from the NHGRI-EBI GWAS catalog (Ensembl release version E93, downloaded on October, 8 2018) [157;275]. MetSyn-related traits were manually selected among all those available and the results were filtered for SNPs with an association p-value $< 5 \cdot 10^{-08}$. The extracted SNPs were mapped to official gene symbols based on their genomic location. All genes located in the genomic interval were considered. As reference genes, we used the RefSeq genes, downloaded from the UCSC genome browser [125] using the table browser tool (human genome assembly: Dec 2013/HG38, downloaded on October, 11 2018 from: `http://genome.ucsc.edu/index.html`). Genes not assigned to chromosomes 1 to 22 were removed and the different transcriptional variants of one gene (isoforms) were merged by considering the minimal starting and the maximal ending position as new range of the gene. Moreover, the gene region was extended 110 kb upstream and 40 kb downstream of the transcript boundaries, following the approach used by MAGENTA [228]. The mapping procedure was executed with the R package GenomicRanges [143].

SNP functional annotation

The SNPs obtained from the GWAS catalog were annotated for their position within genes using the R package VariantAnnotation [190] and TxDb.Hsapiens.UCSC.hg19.knownGene [36] as annotation object. To evaluate the enrichment of the GWAS SNPs in regulatory regions, we used the chromatin state annotations from the NIH Roadmap Epigenomics project [218]. Specifically, the 18-state models for adipose nuclei (E063), liver (E066) and skeletal muscle female (E108) were downloaded from `https://egg2.wustl.edu/roadmap/web_portal/` and the overlap of the GWAS SNP locations with regulatory regions was computed using the R package GenomicRanges [143]. For comparison, we downloaded the full set of HapMap CEU SNPs from the UCSC website (`http://genome.ucsc.edu/`) using Table Browser and annotated them in the same way we did for the GWAS SNPs. Fisher's exact test was used to compare GWAS SNPs and HapMap SNPs.

GWAS summary statistics

A further resource for genetic factors associated with MetSyn is provided by GWAS summary statistics. Previously published results obtained by applying the PASCAL tool [141] were used. 15 GWASs connected to metabolic components were chosen (Appendix E, Table E.6) and genes with a p-value below the threshold of $5 \cdot 10^{-8}$ were selected.

Text mining

Additional MetSyn genes were identified using text mining of PubMed abstracts. The following MeSH terms were identified as being relevant to metabolic syndrome and its primary symptoms: metabolic syndrome x, hyperglycemia, insulin resistance, hyperinsulinism, glucose intolerance, hypertension, obesity, abdominal, hypertriglyceridemia, hypercholesterolemia, waist circumference, waist-hip ratio. The search was limited using the tags [Majr:NoExp], to restrict to articles having major focus on the searched MeSH terms (and no automatic inclusion of child terms of the searched term), and english[language] to restrict the search to English abstracts.

In addition, we complemented the MeSH search with a keyword search using the PubMed [TIAB] tag. The following search terms were used: metabolic syndrome, hyperglycem*, insulin resistan*, hyperinsulin*, glucose intoleran*, hypertension, abdominal obesity, central obesity, hypertriglyceridemia, high triglycerides, hypercholesterolemia, high cholesterol, waist circumference, waist-hip ratio, waist-to-hip ratio. To limit the results to the most relevant articles, the keyword search results were filtered as follows. First, we removed those articles already annotated with MeSH terms, because either the major topic of the article was not considered MetSyn related by the MeSH reviewers, or the articles were already captured by our MeSH search. The remaining articles were further reduced to those waiting for MeSH annotation according to the MedlineCitation Status in PubMed (In-Data-Review, In-Process, Publisher), to cover the recent literature not yet included in the MeSH indexing. Before performing the gene tagging, we removed those articles already present in our GWAS catalog results. The genes mentioned in the titles and abstracts of the selected articles of both search strategies were annotated using the PubTator gene annotation [271–273], and filtered for human genes using the R package org.Hs.eg.db [35]. PubMed was accessed on October, 24 2018 and PubTator annotation was downloaded on October, 25 2018.

Filtering and combining the data-driven and text mining approach

The approach we followed is based on genomic regions and thus includes also genes that are not causative risk factors for metabolic syndrome. Furthermore, the possibility of including false-positive results from the text mining approach cannot be discarded. To deal with these two limitations, gene set enrichment analysis for the genes identified by the data-driven approach was carried out and genes were prioritized based on their biological function. The enrichment analysis was performed using the enrichr tool[40;136], accessed through the RESTful API on October 12, 2018. We selected the following databases: GO biological processes, KEGG, WikiPathways, Reactome, Biocarta, Humancyc, NCI-Nature, and Panther. The analysis resulted in 47 significant gene sets (BH adj. pvalue < 0.05).

5.4.2 Constructing tissue-specific background networks

Integrated tissue-specific networks were constructed by combining two types of networks: transcriptional regulatory networks composed of interactions between transcription factors and the regulated genes, and human protein-protein interaction networks. We consider adipose tissue, liver tissue and skeletal muscle tissue as the three tissues mostly affected by MetSyn.

Regulatory networks were downloaded from `http://regulatorycircuits.org/download.html` as presented in[160]. These tissue-specific gene regulatory networks were inferred by combining transcription factor sequence motifs with activity data for promoters and enhancers from the FANTOM5 project[160]. Among the available individual networks, adipose_tissue_adult, liver_adult and skeletal_muscle_adult were selected. Based on the activity scores edge weights in the range of [0, 2] are provided. To filter for interactions with high evidence scores we choose a cut-off value for these edge weights of 0.4. This value is based on the considerations, that a) a threshold over 0.5 resulted in networks without/with only few nodes, b) a threshold beneath 0.1 rises the possibility of false positives steadily as the distribution of edge weights is highly skewed, and c) the threshold of 0.4 is the maximal value which secures that at least 25% of the TFs and PRs of the resulting network are also expressed in the tissue-specific corresponding network from HIPPIE as additional source for further protein-protein interactions.

Protein-protein interactions (PPI) were obtained from HIPPIE (v2.0)[2;224]. HIPPIE is a comprehensive database combining protein interactions from different sources. Furthermore, a confidence score for each interaction is provided. This score ranges in [0, 1] and reflects the reliability of the interaction based on the number and quality of the experimental technique, the number of studies the interaction is mentioned in and the number of non-human organisms in which the interaction was reproduced. To include only the most reliable interactions, the cut-off value of 0.73 was chosen considering that the curators of HIPPIE refer to this value for high evidence interactions. Following[2], tissue-specific PPIs were created using tissue expression RNA-Seq data from GTEx, while a gene was considered tissue-relevant if it showed an RPKM ≥ 1 in the given tissue. To generate the adipose tissue network we combined the data obtained from subcutaneous and visceral adipose tissue.

After this preprocessing analysis, the regulatory circuits were extended with high evidence interactions from the respective tissue-specific PPI. Relations for nodes in the gene regulatory networks as well as their first neighbors were included during this process.

5.4.3 Drug data

To evaluate the drug effect on MetSyn we combined drug target information and drug expression profiles.

Target information

Information about drugs, their indication, their stage of development (e.g. approved, experimental, withdrawn) and their target was obtained from the public database DrugBank [280] (version 5.1.1, release date: 2018-07-03, download date September, 11 2018). Drugs were retained if they were annotated to have a target gene and if they had an approved status. We further restricted our analysis to pharmacological active drug-target interactions. All target proteins were mapped and annotated using Entrez IDs and official gene symbols. In total, we obtained 3814 drug-target interactions for 1482 distinct drugs and 705 distinct targets (Appendix E, Figure E.5).

Drug expression profiles

The identified drug-target relations were extended by including knowledge about the gene expression profile related to the drug, obtained from the Library of Integrated Cellular Signatures (LINCS) [249], which provides gene expression profiles obtained by analyzing cellular responses (cellular signatures) across different cell-lines in response to a range of perturbations, including also single drug perturbations. We accessed the data using the RESTful API `https://clue.io/` in October 2018 and for each drug the 100 most up- and down-regulated genes were retrieved, relying on high quality signatures (is_gold = 1). If for a certain drug more than one signature was available, we selected the one with the highest signature strength parameter (distil_ss).

5.4.4 Network-based MetSyn modules

Trait-relevant network modules were detected using the walktrap algorithm [200] (implemented in the R package igraph [53]) in combination with an overrepresentation test. The walktrap algorithm identifies network communities based on the concept that random walks of a short length tend to stay in the same network area (identified as module). The algorithm was run using the default parameters. The enrichment in MetSyn genes was tested using Fisher's exact test (p-value $\leq$ 0.05). The communities significantly enriched in MetSyn genes were tested for their biological functionality using pathway enrichment analysis (R package reactomePA [289]).

5.4.5 Network-based drug modules

Network-based drug modules were generated by mapping drug profiles to the networks and connecting the mapped proteins. Starting from the drug target, the list of signature proteins was filtered to extract those having a medium to high semantic similarity with the target protein using the R package GoSemSim [288] (Wang method, cut-off value 0.5). The network-based drug module was then formed by the drug target, the selected subset of drug signature proteins and the shortest paths connecting them.

5.4.6 Proximity score

To quantify the interplay between a drug profile and a MetSyn module a proximity score was defined. This score combines the network-based distance between a drug- and a MetSyn-module, and the semantic similarity of the two modules. The network-based distance was calculated using the closest distance introduced in [104], where it has been shown to outperform other distance measurements. This measurement represents the average shortest path length between the drug module genes and their nearest disease proteins in the network [104]:

$$d_c = \frac{1}{|T|} \sum_{t \in T} \min_{s \in S} d(s,t),$$

where T is the drug module, S the disease module and $d(s,t)$ the shortest distance between two nodes s and t. Normalizing this measurement with the diameter of the network and considering the linear transformation $1 - d_{c,norm}$ defines a score in [0,1]. To include knowledge about the biological function of the drug and disease proteins their GO annotation restricted to biological processes was used to calculate a similarity score in [0,1]. Wang's method [269] combined with the Best-Match Average strategy was used as implemented in the R package GoSemSim [288]. Summing the two above measurements led to our final score in [0,2].

To assess the significance of the results, a reference score distribution corresponding to the expected scores for random sets of drug proteins was created. The construction of the random module follows the strategy to build drug modules described above by selecting first a target protein falling in the same degree bin as the original target, and by then selecting signature genes keeping the internal distances of the original module. Finally, we use the shortest paths between the target and signature genes to construct the random module. A drug resulting in a score higher than 95% of the reference distribution scores was considered significant.

5.4.7 Filtering and prioritization of candidate repurposing drugs

We retrieved data about the targets of the repurposing candidates using the Open Targets Platform [133] REST API (accessed November, 2018) to extract known associations between the target genes and a list of traits identified to be associated to MetSyn. Targets with at least one association score ≥ 0.2 were excluded from further considerations, because this indicates that the target has already been under investigation for therapeutic interventions related to MetSyn Furthermore, we extracted the known side effects of the candidate drugs using the DrugCentral platform [262], accessed via `http://drugcentral.org` in November 2018, and excluded those drugs with unwanted contraindications associated to MetSyn. A final prioritization step was carried out based on the tissue expression of the drug targets accessed using Human Protein Atlas [260], GTEx [252], and Fantom5 [151].

5.4.8 *BTK* gene expression analysis and immune cell component estimation

The mouse and human expression datasets described in this study are publicly available from NCBI GEO (`https://www.ncbi.nlm.nih.gov/geo/`) and EMBL-EBI ArrayExpress (`https://www.ebi.ac.uk/arrayexpress/`). From NCBI GEO we downloaded the Series_matrix files of the

following datasets: GSE54350[56], GSE32095[116], GSE25205[237], and GSE27951[123], while the processed data of the E-MTAB-54[64] dataset was downloaded from EMBL-EBI ArrayExpress.

The selected datasets were annotated using the R Bioconductor annotation packages corresponding to the microarray platform used in the respective study or the annotation file provided by NCBI and ArrayExpress. The probe signals were summarized at gene level considering the median. T-test was used to compare the mean of *BTK* transcript levels between different subgroups.

To estimate the abundances of immune cells in adipose tissue we used the online version of Cibersort[181](accessed via `https://cibersort.stanford.edu/`), run with default parameters. Cibersort is software based on a deconvolution algorithm that estimates the abundances of immune cells from gene-expression data on the basis of previous knowledge about immune cell gene expression (immune signature). For the human datasets, we used the immune signature provided by Cibersort that contains 22 immune cell types, while for the mouse datasets, we used the immune signature provided in[43] consisting of 25 immune cell types. For visualization purposes, the mouse immune cells were grouped in seven main classes: Granulocytes (Mast Cells, Neutrophil Cells, Eosinophil Cells), B cells (B Cells Memory, B Cells Naive, Plasma Cells), T cells (T Cells CD8 Actived, T Cells CD8 Naive, T Cells CD8 Memory, T Cells CD4 Memory, T Cells CD4 Naive, T Cells CD4 Follicular, Th1 Cells, Th17 Cells, Th2 Cells, GammaDelta T Cells), Macrophages (M0 Macrophage, M1 Macrophage, M2 Macrophage), Monocytes (Monocyte), Natural Killer cells (NK Resting, NK Actived), and Dendritic cells (DC Actived, DC Immature).

Author Contributions

The content of this chapter has been submitted [170].
S. Parolo (S.P.) and C. Priami (C.P.) conceived and designed the study. K. Misselbeck (K.M.) performed the computational analysis. S.P. performed the biological interpretation of the results. S.P., M. Morine (M.M.), L. Leonardelli (L.L.) and P. Bora (P.B.) contributed to the computational analysis. S.P., E. Domenici (E.D.) and C.P. supervised the study and provided guidance on data analysis. K.M. and S.P. wrote the manuscript, all authors reviewed the manuscript and approved the final version.

Conclusion

In this dissertation we applied computational systems biology methods to human metabolism, focusing on the folate-mediated one-carbon metabolism and metabolic syndrome. Our results highlight that computational methods provide a useful framework to gain biological insights by complementing experimental studies with computational evidence. In the following, further developments, possible extensions and potential limitations are discussed.

Folate-mediated one-carbon metabolism

The folate-mediated one-carbon metabolism is central for cell division, DNA repair and homeostasis because folate chemically activates, oxidates, and supplies one-carbon units for the synthesis of nucleotides and for remethylation of homocysteine. Impairment of FOCM, which disrupts cell replication and obstructs DNA repair, is associated with neural tube defects, cancer and other neurodegenerative disorders such as Alzheimer's and Parkinson's disease. Understanding the dynamics of FOCM and its responsiveness to both genetic and environmental perturbations is key to understanding the etiology of folate-related pathologies.

Mathematical models provide an excellent framework to obtain a thorough understanding not only of the individual molecules forming a biological system but also of their interactions, and help to identify underlying pathways and causative molecular mechanisms. The mathematical models of FOCM presented in Part I of this dissertation allowed us to develop a more comprehensive picture of the network and helped to identify how disruptions of the process may affect the overall behaviour of the system. Deterministic and stochastic simulation strategies were combined to systematically explore the effect of perturbations on model dynamics, quantified in terms of model steady states, and to assess their stochastic stability. *In silico* experiments of the model describing FOCM in the cytoplasm (Chapter 1 and Chapter 2) highlight the role of the common C677T MTHFR polymorphism, folate availability, as well as vitamin B_{12} disposability for the normal functioning of the network. The model connected disruptions of these three factors, quantified through sensitivity analysis, to impairments of *de novo* dTMP synthesis and to the pathological outcomes of NTDs. In addition, extending the inital model as presented in Chapter 3, allowed us to highlight the importance of 5fTHF and MTHFS to prevent network impairment by regulating purine synthesis.

Exploiting a hybrid-stochastic simulation approach allowed us to link both the MTHFR polymorphism and low levels of folate to a loss in overall network stability, while the presence of the 5fTHF futile cycle increased the stochastic stability. This observation highlights the advantage of simulation approach we conducted, as FOCM is expected to exhibit variability (i.e., stochasticity)

in its behavior [93], which can only be identified using a stochastic or hybrid-stochastic approach.

In addition to the aformentioned, our main contribution to the study of FOCM lies in the extension of the model with the nuclear compartment. By including ad-hoc functions to model the localization of folate and folate-dependent enzymes to the nucleus and by accounting for the kinetic effects of multienzyme complex formation and substrate channelling we explored the relative contribution of nuclear and cytosolic dTMP synthesis to overall dTMP synthesis. We quantified the effect of each enzyme in the pathway through sensitivity analysis, identifying the main regulatory reactions of the compartmentalized system.

However, the general limitations of mathematical models also apply to our hybrid-stochastic model of FOCM. Mathematical models are abstractions of the reality and are therefore necessarily incomplete by definition. In our specific scenario, the accuracy of model predictions could be improved by the inclusion of the mitochondrial FOCM-related reactions, the thymidylate salvage pathway or the addition of connected pathways. The functions encoding the translocation of matter to the nuclear compartment are based on assumptions drawn from experimental data, a particular worthwhile addition to the model would be the extension of these functions to model the SUMOylation process and the following translocation expliclitly. However, for this purpose, more biological knowledge about this process has to be established, as mathematical models are always restricted by the availability of experimental data and by the biological understanding of the system.

Metabolic Syndrome

Metabolic Syndrome (MetSyn), defined as the conincidence of multiple metabolic alterations leading to an increased risk of cardiovascular diseases and diabetes, is becoming increasingly prevalent in almost all countries. The mechanisms playing an etiologic role in the development of MetSyn are still elusive and therefore the description and interpretation of the underlying pathophysiology is essential to develop efficacious treatments. Standard treatment strategies are focused on lifestyle changes to manage excessive adipose tissue accumulation, normalize metabolic aberrances and reduce cardiovascular risk. However, responses to the lifestyle modifications alone are often not satisfying and an intensification of the treatment including drug management is necessary as the disease progresses.

In this dissertation, we presented a computational systems biology approach to identify biological processes involved in MetSyn and to detect potential new pharmaceutical treatments by means of drug repurposing. Adopting a methodology based on systems biology enabled us to explore the biological question more systematically by refraining from the single gene perspective. The decisive aspect of this study lies in the efficient integration of different experimental resources, which allowed us to improve data completeness and to accomplish a comprehensive view of the disease pathogenic mechanisms. For this purpose, we combined a) genetic evidence derived from GWAS with literature findings to establish a list of disorder-related genes, b) tissue-specific protein-protein interactions with transcriptional regulatory networks to build a background network, and c) drug target information with drug-modulated expression data to describe the drug's pharmacological effect. The prediction of potential drug - disease interactions was based on a proximity score, which relies on topological characteristics of the network and functional similarities between genes. We

further addressed the issue of identifying the most promising candidate among the results by including a filtering and prioritization analysis. This step is crucial to identify a limited and feasible number of drug candidates for the subsequent experimental validation.

Our main result, which has been obtained from the analysis of the adipose network, identified the key role exerted by inflammation in obesity and suggested the adoption of anti-inflammatory therapies. In particular, our results pointed out ibrutinib, an inhibitor of bruton tyrosine kinase (BTK), as the most promising candidate for drug repurposing.

While computational methods for drug repurposing can provide feasible and reliable results, there is no doubt that further experimental studies are crucial for the successful investigation of the identified drug candidates. This means that the next important step will be to perform *in vitro* and *in vivo* studies to experimentally validate our findings by exploring the immunological steps affected by BTK in obesity and by studying the effect of ibrutinib in metabolic inflammation and metabolic syndrome.

The proposed methodology was developed and introduced on the example of MetSyn, but can be applied to other diseases as well. The only requirement is a list of disease genes that can be derived from many sources (e.g. from GWAS results or gene/protein expression profiles). Furthermore, our method can provide useful insights for a specific drug of interest once network modules for various diseases have been defined.

A possible extension of our strategy could be the inclusion of drugs for which the mechanism of action is yet unknown, e.g. new compounds or neutraceuticals. Currently, the method assesses the therapeutic effect of a drug by building a network module centered around the drug target. Pharmaceutical compounds without available data on the mechanism of action could be included by detecting the respective network module solely using gene expression profiles (derived e.g. from RNA-Seq experiments). The drug module could be identified either by connecting the altered genes in the background network using shortest paths or by detecting network clusters enriched in drug-modulated genes, following the strategy used for the detection of disease modules. Moreover, the addition of further data relevant for the disease of interest could lead to an extension of the proposed methodology and improve the final prediction.

Bibliography

[1] Aherne, G. W., Ward, E., Lawrence, N., et al. (1998). Comparison of plasma and tissue levels of ZD1694 (Tomudex), a highly polyglutamatable quinazoline thymidylate synthase inhibitor, in preclinical models. *British Journal of Cancer*, 77, 221–226.

[2] Alanis-Lobato, G., Andrade-Navarro, M. A., & Schaefer, M. H. (2017). HIPPIE v2.0: Enhancing meaningfulness and reliability of protein-protein interaction networks. *Nucleic Acids Research*, 45(D1), D408–D414.

[3] Alberti, K. G. M. M., Eckel, R. H., Grundy, S. M., et al. (2009). Harmonizing the metabolic syndrome: A joint interim statement. *Circulation*, 120(16), 1640–1645.

[4] Alberti, K. G. M. M. & Zimmet, P. Z. (1998). Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus provisional report of a WHO consultation. *Diabetic medicine : a journal of the British Diabetic Association*, 15(7), 539–553.

[5] Alberti, K. G. M. M., Zimmet, P. Z., & Shaw, J. (2006). Metabolic syndrome - a new world-wide definition. A consensus statement from the international diabetes federation. *Diabetic Medicine*, 23(5), 469–480.

[6] Alberts, B., Bray, D., Hopkin, K., et al. (2014). *Essential Cell Biology*. New York: Garland Science.

[7] Ames, B. N. (2001). DNA damage from micronutrient deficiencies is likely to be a major cause of cancer. *Mutation Research*, 475(1–2), 7–20.

[8] Ami, N., Bernstein, M., Boucher, F., Rieder, M., & Parker, L. (2016). Folate and neural tube defects: The role of supplements and food fortification. *Paediatrics and Child Health*, 21(3), 145–149.

[9] An, S., Kumar, R., Sheets, E. D., & Benkovic, S. J. (2008). Reversible compartmentalization of de novo purine biosynthetic complexes in living cells. *Science*, 320(5872), 103–106.

[10] Andersen, C. J., Murphy, K. E., & Fernandez, M. L. (2016). Impact of Obesity and Metabolic Syndrome on Immunity. *Advances in Nutrition*, 7(1), 66–75.

[11] Anderson, D. D., Quintero, C. M., & Stover, P. (2011). Identification of a de novo thymidylate biosynthesis pathway in mammalian mitochondria. *PNAS*, 108(37), 15163–15168.

[12] Anderson, D. D. & Stover, P. J. (2009). SHMT1 and SHMT2 are functionally redundant in nuclear de novo thymidylate biosynthesis. *PLOS ONE*, 4(6).

[13] Anderson, D. D., Woeller, C. F., Chiang, E. P., Shane, B., & Stover, P. J. (2012). Serine hydroxymethyltransferase anchors de Novo thymidylate synthesis pathway to nuclear lamina for DNA synthesis. *Journal of Biological Chemistry*, 287(10), 7051–7062.

[14] Anderson, D. D., Woeller, C. F., & Stover, P. J. (2007). Small ubiquitin-like modifier-1 (SUMO-1) modification of thymidylate synthase and dihydrofolate reductase. *Clinical Chemistry and Laboratory Medicine*, 45(12), 1760–1763.

[15] Anguera, M. C., Field, M. S., Perry, C., et al. (2006). Regulation of Folate-mediated One-carbon Metabolism by 10-Formyltetrahydrofolate Dehydrogenase. *Journal of Biological Chemistry*, 281(27), 18335–18342.

[16] Anguera, M. C., Suh, J. R., Ghandour, H., et al. (2003). Methenyltetrahydrofolate synthetase regulates folate turnover and accumulation. *Journal of Biological Chemistry*, 278(32), 29856–29862.

[17] Antonarakis, E. (2010). Human Genome Sequence and Variation. In M. Speicher (Ed.), *Vogel and Motulsky's Human Genetics: Problems and Approaches* (pp. 31–53). Springer Berlin, Heidelberg.

[18] Armani, A., Berry, A., Cirulli, F., & Caprio, M. (2017). Molecular mechanisms underlying metabolic syndrome: the expanding role of the adipocyte. *The FASEB Journal*, 31(10), 4240–4255.

[19] Ashburn, T. T. & Thor, K. B. (2004). Drug repositioning: identifying and developing new uses for existing drugs. *Nature Reviews Drug Discovery*, 3, 673–683.

[20] Atreya, C. E. & Anderson, K. S. (2004). Kinetic characterization of bifunctional thymidylate synthase-dihydrofolate reductase (TS-DHFR) from Cryptosporidium hominis: A paradigm shift for TS activity and channeling behavior. *The Journal of Biological Chemistry*, 279(18), 18314–18322.

[21] Bagley, P. J. & Selhub, J. (1998). A common mutation in the methylenetetrahydrofolate reductase gene is associated with an accumulation of formylated tetrahydrofolates in red blood cells. *PNAS*, 95(22), 13217–13220.

[22] Balkau, B. & Charles, M. (1999). Comment on the provisional report from the WHO consultation. European Group for the Study of Insulin Resistance (EGIR). *Diabet Med*, 16, 442–443.

[23] Beaudin, A. E. (2012). Dietary folate, but not choline, modifies neural tube defect risk in Shmt1 knockout mice. *The American Journal of Clinical Nutrition*, 95, 109–114.

[24] Beaudin, A. E., Abarinov, E. V., Noden, D. M., et al. (2011). Shmt1 and de novo thymidylate biosynthesis underlie folate-responsive neural tube defects in mice. *American Journal of Clinical Nutrition*, 93, 789–798.

[25] Beaudin, A. E. & Stover, P. (2009). Insights into metabolic mechanisms underlying folate-responsive neural tube defects: a minireview. *Birth Defects Research Part A: Clinical and Molecular Teratology*, 85, 274–284.

[26] Becker, K. A., Ghule, P. N., Therrien, J. A., et al. (2006). Self-renewal of human embryonic stem cells is supported by a shortened G1 cell cycle phase. *Journal of Cellular Physiology*, 209(3), 883–893.

[27] Bertrand, R. & Jolivet, J. (1989). Methenyltetrahydrofolate synthetase prevents the inhibition of phosphoribosyl 5-aminoimidazole 4-carboxamide ribonucleotide formyltransferase by 5-formyltetrahydrofolate polyglutamates. *Journal of Biological Chemistry*, 264, 8843–8846.

[28] Blencowe, H., Kancherla, V., Moorthie, S., Darlison, M. W., & Modell, B. (2018). Estimates of global and regional prevalence of neural tube defects for 2015: a systematic analysis. *Annals of the New York Academy of Sciences*, 1414(1), 31–46.

[29] Blount, B. C., Mack, M. M., Wehr, C. M., et al. (1997). Folate deficiency causes uracil misincorporation into human DNA and chromosome breakage: implications for cancer and neuronal damage. *PNAS*, 94, 3290–3295.

[30] Brewer, B. J., Chlebowicz-Sledziewska, E., & Fangman, W. L. (1984). Cell cycle phases in the unequal mother/daughter cell cycles of Saccharomyces cerevisiae. *Molecular and Cellular Biology*, 4, 2529–2531.

[31] Burr, N. E., Hull, M. A., & Subramanian, V. (2017). Folic Acid Supplementation May Reduce Colorectal Cancer Risk in Patients With Inflammatory Bowel Disease: A Systematic Review and Meta-Analysis. *Journal of Clinical Gastroenterology*, 51(3), 247–253.

[32] Cai, J. J. & Petrov, D. A. (2010). Relaxed purifying selection and possibly high rate of adaptation in primate lineage-specific genes. *Genome Biology and Evolution*, 2, 393–409.

[33] Cardon, L. R. & Harris, T. (2016). Precision medicine, genomics and drug discovery. *Human Molecular Genetics*, 25(R2), R166–R172.

[34] Carithers, L. J., Ardlie, K., Barcus, M., et al. (2015). A Novel Approach to High-Quality Postmortem Tissue Procurement: The GTEx Project. *Biopreservation and Biobanking*, 13(5), 311–319.

[35] Carlson, M. (2018). org.Hs.eg.db: Genome wide annotation for Human, R package version 3.6.0.

[36] Carlson, M. & Maintainer, B. P. (2015). *TxDb.Hsapiens.UCSC.hg19.knownGene: Annotation package for TxDb object(s).*

[37] Chabes, A., Georgieva, B., Domkin, V., et al. (2003). Survival of DNA damage in yeast directly depends on increased dNTP levels allowed by relaxed feedback inhibition of ribonucleotide reductase. *Cell*, 112(3), 391–401.

[38] Chan, E. & Cronstein, B. (2013). Mechanisms of Action of Methotrexate. *Bulletin of the Hospital for Joint Diseases*, 7171(Suppl 1), 5–8.

[39] Chen, C., Wang, J.-L., Tsai, Y.-T., et al. (2017a). Repurposing Metformin for Lung Cancer Management. In M. Adonis (Ed.), *Prevention, Diagnosis, and Treatment of Lung Cancer* chapter 9. Intech.

[40] Chen, E. Y., Tan, C. M., Kou, Y., et al. (2013). Enrichr: Interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics*, 14.

[41] Chen, H.-R., Sherr, D. H., Hu, Z., & DeLisi, C. (2016). A network based approach to drug repositioning identifies plausible candidates for breast cancer and prostate cancer. *BMC Medical Genomics*, 9(51).

[42] Chen, Y.-L., Eriksson, S., & Chang, Z.-F. (2010). Regulation and functional contribution of thymidine kinase 1 in repair of DNA damage. *The Journal of Biological Chemistry*, 285(35), 27327–27335.

[43] Chen, Z., Huang, A., Sun, J., et al. (2017b). Inference of immune cell composition on the expression profiles of mouse tissue. *Scientific Reports*, 7(January), 1–11.

[44] Cheng, F., Desai, R. J., Handy, D. E., et al. (2018). Network-based approach to prediction and population-based validation of in silico drug repurposing. *Nature Communications*, 9.

[45] Cheng, F., Murray, J. L., Zhao, J., et al. (2016a). Systems Biology-Based Investigation of Cellular Antiviral Drug Targets Identified by Gene-Trap Insertional Mutagenesis. *PLOS Computational Biology*, 12(9), 1–27.

[46] Cheng, F., Zhao, J., Fooksa, M., & Zhao, Z. (2016b). A network-based drug repositioning infrastructure for precision cancer medicine through targeting significantly mutated genes in the human cancer genomes. *Journal of the American Medical Informatics Association*, 23(4), 681–691.

[47] Chiang, E. P., Wang, Y. C., & Tang, F. Y. (2007). Folate restriction and methylenetetrahydrofolate reductase 677T polymorphism decreases AdoMet synthesis via folate-dependent remethylation in human-transformed lymphoblasts. *Leukemia*, 21, 651–658.

[48] Choi, S.-W. & Mason, J. B. (2000). Folate and Carcinogenesis: An Integrated Scheme. *The Journal of Nutrition*, 130(2), 129–132.

[49] Chon, J., Stover, P. J., & Field, M. S. (2017). Targeting nuclear thymidylate biosynthesis. *Molecular Aspects of Medicine*, 53, 48–56.

[50] Consolim-Colombo, F. M., Sangaleti, C. T., Costa, F. O., et al. (2017). Galantamine alleviates inflammation and insulin resistance in patients with metabolic syndrome in a randomized trial. *JCI Insight*, 2(14), 1–13.

[51] Crider, K. S., Bailey, L. B., & Berry, R. J. (2011). Folic acid food fortification-its history, effect, concerns, and future directions. *Nutrients*, 3(3), 370–384.

[52] Crofford, L. J., Nyhoff, L. E., Sheehan, J. H., & Kendall, P. L. (2016). The role of Bruton's tyrosine kinase in autoimmunity and implications for therapy. *Expert Review of Clinical Immunology*, 12(7), 763–773.

[53] Csardi, G. & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal*, Complex Systems, 1–9.

[54] Czeizel, A. E. & Dudas, I. (1992). Prevention of the first occurrence of neural tube defects by periconceptional vitamin supplementation. *The New England Journal of Medicine*, 327(26), 1832–1835.

[55] D'Agati, V. D. (2008). The spectrum of focal segmental glomerulosclerosis: new insights. *Current Opinion in Nephrology and Hypertension*, 17(3), 271–281.

[56] Dalmas, E., Venteclef, N., Caer, C., et al. (2014). T cell-derived IL-22 amplifies IL-1β-driven inflammation in human adipose tissue: relevance to obesity and type 2 diabetes. *Diabetes*, 63(6), 1966–1977.

[57] Davies, M., Nowotka, M., Papadatos, G., et al. (2015). ChEMBL web services: streamlining access to drug discovery data and utilities. *Nucleic Acids Research*, 43(W1), W612–W620.

[58] Davis, S. R., Quinlivan, E. P., Shelnutt, K. P., et al. (2005). Homocysteine synthesis is elevated but total remethylation is unchanged by the methylenetetrahydrofolate reductase 677C->T polymorphism and by dietary folate restriction in young women. *The Journal of Nutrition*, 135, 1045–1050.

[59] de Anda-Jáuregui, G., Guo, K., McGregor, B. A., & Hur, J. (2018). Exploration of the anti-inflammatory drug space through network pharmacology: Applications for drug repurposing. *Frontiers in Physiology*, 9, 1–12.

[60] de Leeuw, C. A., Mooij, J. M., Heskes, T., & Posthuma, D. (2015). MAGMA: Generalized Gene-Set Analysis of GWAS Data. *PLOS Computational Biology*, 11(4), 1–19.

[61] De Lorenzo, V. & Galperin, M. (2009). Microbial systems biology: Bottom up and top down. *FEMS Microbiology Reviews*, 33(1), 1–2.

[62] Delavan, B., Roberts, R., Huang, R., et al. (2018). Computational drug repositioning for rare diseases in the era of precision medicine. *Drug Discovery Today*, 23(2), 382–394.

[63] DiMasi, J. A., Grabowski, H. G., & Hansen, R. W. (2016). Innovation in the pharmaceutical industry: New estimates of R&D costs. *Journal of Health Economics*, 47, 20–33.

[64] Drong, A. W., Nicholson, G., Hedman, Å. K., et al. (2013). The Presence of Methylation Quantitative Trait Loci Indicates a Direct Genetic Influence on the Level of DNA Methylation in Adipose Tissue. *PLOS ONE*, 8(2).

[65] Dubitzky, W. (2006). Understanding the computational methodologies of systems biology. *Briefings in Bioinformatics*, 7(4), 315–317.

[66] Ducker, G. S. & Rabinowitz, J. D. (2017). One-Carbon Metabolism in Health and Disease. *Cell Metabolism*, 25(1), 27–42.

[67] Dunlevy, L. P. E., Chitty, L. S., Burren, K. A., et al. (2007). Abnormal folate metabolism in foetuses affected by neural tube defects. *Brain*, 130(4), 1043–1049.

[68] Eckel, R. H., Grundy, Scott, M., & Zimmet, P. Z. (2005). The metabolic syndrome. *The Lancet*, 366(9468), 1415–1428.

[69] Edwards, L. M. (2017). Metabolic systems biology: a brief primer. *The Journal of Physiology*, 595(9), 2849–2855.

[70] Einhorn, D., Reaven, G., Cobin, R., et al. (2003). American College of Endocrinology position statement on the insulin resistance syndrome. *Endocrine Practice*, 9(2), 5–21.

[71] Emig, D., Ivliev, A., Pustovalova, O., et al. (2013). Drug Target Prediction and Repositioning Using an Integrated Network-Based Approach. *PLOS ONE*, 8(4).

[72] Enns, G. (2018). Metabolic disease. https://www.britannica.com/science/metabolic-disease.

[73] Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (2001). Executive Summary of the Third Report of the National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, and Treatment of High Blood Cholesterol in Adults (Adult Treatment Panel III). *JAMA: The Journal of the American Medical Association*, 285(19), 2486–2497.

[74] Fabregat, A., Jupe, S., Matthews, L., et al. (2018). The Reactome Pathway Knowledgebase. *Nucleic Acids Research*, 46, D649–D655.

[75] Felkner, M., Suarez, L., Canfield, M. A., Brender, J. D., & Sun, Q. (2009). Maternal serum homocysteine and risk for neural tube defects in a Texas-Mexico border population. *Birth Defects Research Part A - Clinical and Molecular Teratology*, 85(6), 574–581.

[76] Field, M. S., Anderson, D. D., & Stover, P. J. (2011). Mthfs is an essential gene in mice and a component of the purinosome. *Frontiers in Genetics*, 2(36), 1–13.

[77] Field, M. S., Anguera, M. C., Page, R., & Stover, P. J. (2009). 5,10-Methenyltetrahydrofolate synthetase activity is increased in tumors and modifies the efficacy of antipurine LY309887. *Archives of Biochemistry and Biophysics*, 481(2), 145–150.

[78] Field, M. S., Kamynina, E., Agunloye, O. C., et al. (2014). Nuclear enrichment of folate cofactors and methylenetetrahydrofolate dehydrogenase 1 (MTHFD1) protect de novo thymidylate biosynthesis during folate deficiency. *Journal of Biological Chemistry*, 289(43), 29642–29650.

[79] Field, M. S., Kamynina, E., Chon, J., & Stover, P. J. (2018). Nuclear Folate Metabolism. *Annual Review of Nutrition*, 38, 219–243.

[80] Field, M. S., Kamynina, E., Watkins, D., Rosenblatt, D. S., & Stover, P. (2015a). Human mutations in methylenetetrahydrofolate dehydrogenase 1 impair nuclear de novo thymidylate biosynthesis. *PNAS*, 112.

[81] Field, M. S., Kamynina, E., Watkins, D., Rosenblatt, D. S., & Stover, P. J. (2015b). New insights into the metabolic and nutritional determinants of severe combined immunodeficiency. *Rare Diseases*, 3(1), e1112479.

[82] Field, M. S., Szebenyi, D. M., Perry, C. A., & Stover, P. J. (2007). Inhibition of 5,10-Methenyltetrahydrofolate Synthetase. *Archives of Biochemistry and Biophysics*, 458(2), 194–201.

[83] Field, M. S., Szebenyi, D. M. E., & Stover, P. J. (2006). Regulation of de novo purine biosynthesis by methenyltetrahydrofolate synthetase in neuroblastoma. *Journal of Biological Chemistry*, 281(7), 4215–4221.

[84] Fox, J. T., Shin, W. K., Caudill, M. A., & Stover, P. J. (2009). A UV-responsive internal ribosome entry site enhances serine hydroxymethyltransferase 1 expression for DNA damage repair. *Journal of Biological Chemistry*, 284(45), 31097–31108.

[85] Fox, J. T. & Stover, P. J. (2008). Chapter 1 Folate-Mediated One-Carbon Metabolism. In *Folic Acid and Folates*, volume 79 of *Vitamins & Hormones* (pp. 1 – 44). Academic Press.

[86] Fujioka, A., Terai, K., Itoh, R. E., et al. (2006). Dynamics of the Ras/ERK MAPK cascade as monitored by fluorescent probes. *Journal of Biological Chemistry*, 281(13), 8917–8926.

[87] Gaber, K., Farang, M., Soliman, S., El-Bassyouni, H., & El-Kamah, G. (2007). Maternal vitamin B12 and the risk of fetal neural tube defects in Egyptian patients. *Clinical Laboratory*, 53(1-2), 69–75.

[88] Gamazon, E. R., Segrè, A. V., van de Bunt, M., Wen, X., & Xi, H. S. (2018). Using an atlas of gene regulation across 44 human tissues to inform complex disease- and trait-associated variation. *Nature Genetics*, 50, 956–967.

[89] Ghandour, H., Chen, Z., Selhub, J., & Rozen, R. (2004). Mice deficient in methylenetetrahydrofolate reductase exhibit tissue-specific distribution of folates. *The Journal of Nutrition*, 134, 2975–2978.

[90] Gillespie, D. T. (1977). Exact Stochastic Simulation of Coupled Chemical Reactions. *The Journal of Physical Chemistry*, 81(25), 2340–2361.

[91] Girgis, S., Suh, J. R., Jolivet, J., & Stover, P. J. (1997). 5-Formyltetrahydrofolate Regulates Homocysteine Remethylation in Human Neuroblastoma. *The Journal of Biological Chemistry*, 272, 4729–4734.

[92] Gon, S., Napolitano, R., Rocha, W., Coulon, S., & Fuchs, R. P. (2011). Increase in dNTP pool size during the DNA damage response plays a key role in spontaneous and induced-mutagenesis in Escherichia coli. *PNAS*, 108, 19311–19316.

[93] Goss, P. J. & Peccoud, J. (1998). Quantitative modeling of stochastic systems in molecular biology by using stochastic Petri nets. *PNAS*, 95(12), 6750–6755.

[94] Gostner, R., Baldacci, B., Morine, M. J., & Priami, C. (2015). Graphical Modeling Tools for Systems Biology. *ACM Computing Surveys*, 47(2), 1–21.

[95] Govindaraj, R. G., Naderi, M., Singha, M., Lemoine, J., & Brylinski, M. (2018). Large-scale computational drug repositioning to find treatments for rare diseases. *npj Systems Biology and Applications*, 4(1), 1–10.

[96] Goyer, A., Collakova, E., De La Garza, R. D., et al. (2005). 5-Formyltetrahydrofolate is an inhibitory but well tolerated metabolite in Arabidopsis leaves. *Journal of Biological Chemistry*, 280(28), 26137–26142.

[97] Gray, P., Dunne, A., Brikos, C., et al. (2006). MyD88 adapter-like (Mal) is phosphorylated by Bruton's tyrosine kinase during TLR2 and TLR4 signal transduction. *Journal of Biological Chemistry*, 281(15), 10489–10495.

[98] Greene, C. S., Krishnan, A., Wong, A. K., et al. (2015). Understanding multicellular function and disease with human tissue-specific networks. *Nature Genetics*, 47, 569–576.

[99] Greene, N. D. E., Stanier, P., & Copp, A. J. (2009). Genetics of human neural tube defects. *Human Molecular Genetics*, 18(R2), R113–R129.

[100] Greenwood, M. T., Calmels, E. M., & Storms, R. K. (1986). Growth-rate-dependent regulation of the expression and inactivation of thymidylate synthase in Saccharomyces cerevisiae. *Journal of Bacteriology*, 168(3), 1336–1342.

[101] Groenen, P. M. W., Van Rooij, I. A. L. M., Peer, P. G. M., et al. (2004). Marginal maternal vitamin B 12 status increases the risk of offspring with spina bifida. *American Journal of Obstetrics and Gynecology*, 191(1), 11–17.

[102] Grundy, S. M. (2006). Drug therapy of the metabolic syndrome: Minimizing the emerging crisis in polypharmacy. *Nature Reviews Drug Discovery*, 5, 295–309.

[103] Gu, Q., Li, Y., Cui, Z.-L., & Luo, X.-P. (2012). Homocysteine, folate, vitamin B12 and B6 in mothers of children with neural tube defects in Xinjiang, China. *Acta Paediatrica*, 101(11), e486–e490.

[104] Guney, E., Menche, J., Vidal, M., & Barabási, A.-L. (2016). Network-based in silico drug efficacy screening. *Nature Communications*, 7.

[105] Hakansson, P., Hofer, A., & Thelander, L. (2006). Regulation of mammalian ribonucleotide reduction and dNTP pools after DNA damage and in resting cells. *The Journal of Biological Chemistry*, 281, 7834–7841.

[106] Hakkola, J., Rysä, J., & Hukkanen, J. (2016). Regulation of hepatic energy metabolism by the nuclear receptor PXR. *Biochimica et Biophysica Acta - Gene Regulatory Mechanisms*, 1859(9), 1072–1082.

[107] Harpaz, Y., Gerstein, M., & Chothia, C. (1994). Volume changes on protein folding. *Structure*, 2(7), 641–649.

[108] Harvey, R. & Dev, I. (1975). Regulation in the folate pathway of Escherichia Coli. *Advances in Enzyme Regulation*, 13, 99–124.

[109] Heng, T. S. P. & Painter, M. W. (2008). The Immunological Genome Project: networks of gene expression in immune cells. *Nature Immunology*, 9, 1091–1094.

[110] Herbig, K., Chiang, E. P., Lee, L. R., et al. (2002). Cytoplasmic serine hydroxymethyltransferase mediates competition between folate-dependent deoxyribonucleotide and S-adenosylmethionine biosyntheses. *Journal of Biological Chemistry*, 277(41), 38381–38389.

[111] Heseker, H. B., Mason, J. B., Selhub, J., Rosenberg, I. H., & Jacques, P. F. (2009). Not all cases of neural-tube defect can be prevented by increasing the intake of folic acid. *British Journal of Nutrition*, 102(2), 173–180.

[112] Hilton, J. G., Cooper, B. A., & Rosenblatt, D. S. (1979). Folate polyglutamate synthesis and turnover in cultured human fibroblasts. *Journal of Biological Chemistry*, 254(17), 8398–8403.

[113] Hsu, F., Kent, J. W., Clawson, H., et al. (2006). The UCSC known genes. *Bioinformatics*, 22(9), 1036–1046.

[114] Hu, G. & Agarwal, P. (2009). Human disease-drug network based on genomic expression profiles. *PLOS ONE*, 4(8).

[115] Huang, P. L. (2009). A comprehensive definition for metabolic syndrome. *Disease Models & Mechanisms*, 2, 231–237.

[116] Ichimura, A., Hirasawa, A., Poulain-Godefroy, O., et al. (2012). Dysfunction of lipid sensor GPR120 leads to obesity in both mouse and human. *Nature*, 483, 350–354.

[117] Isik, Z., Baldow, C., Cannistraci, C. V., & Schroeder, M. (2015). Drug target prioritization by perturbed gene expression and network information. *Scientific Reports*, 5, 1–13.

[118] Ito, M., Shichita, T., Okada, M., et al. (2015). Bruton's tyrosine kinase is essential for NLRP3 inflammasome activation and contributes to ischaemic brain injury. *Nature Communications*, 6, 1–11.

[119] Jackson, R. C. & Harrap, K. R. (1973). Studies with a Mathematical of Folate Metabolism. *Archives of Biochemistry and Biophysics*, (16), 827–841.

[120] Jackson, R. C. & Harrap, K. R. (1979). Computer models of anticancer drug interaction. *Pharmacology & Therapeutics*, 4(2), 245–280.

[121] Jones, B. J. & Bloom, S. R. (2015). The New Era of Drug Therapy for Obesity: The Evidence and the Expectations. *Drugs*, 75(9), 935–945.

[122] Kassi, E., Pervanidou, P., Kaltsas, G., & Chrousos, G. (2011). Metabolic syndrome: Definitions and controversies. *BMC Medicine*, 9(48), 1–13.

[123] Keller, P., Gburcik, V., Petrovic, N., et al. (2011). Gene-chip studies of adipogenesis-regulated microRNAs in mouse primary adipocytes and human obesity. *BMC Endocrine Disorders*, 11(7).

[124] Kendall, P. L., Moore, D. J., Hulbert, C., et al. (2009). Reduced diabetes in btk-deficient nonobese diabetic mice and restoration of diabetes with provision of an anti-insulin IgH chain transgene. *Journal of Immunology*, 183(10), 6403–6412.

[125] Kent, W. J., Sugnet, C. W., Furey, T. S., & Roskin, K. M. (2002). The Human Genome Browser at UCSC. *Genome Research*, 12(6), 996–1006.

[126] Kiechl, S., Wittmann, J., Giaccari, A., et al. (2013). Blockade of receptor activator of nuclear factor-κB (RANKL) signaling improves hepatic insulin resistance and prevents development of diabetes mellitus. *Nature Medicine*, 19, 358–363.

[127] Kim, D. W., Huang, T., Schirch, D., & Schirch, V. (1996). Properties of Tetrahydropteroylpentaglutamate Bound to 10-Formyltetrahydrofolate Dehydrogenase. *Biochemistry*, 35(49), 15772–15783.

[128] Kim, Y. I. (1999). Folate and cancer prevention: a new medical application of folate beyond hyperhomocysteinemia and neural tube defects. *Nutrition reviews*, 57(10), 314–21.

[129] Kirke, P. N., Molloy, A. M., Daly, L. E., et al. (1993). Maternal plasma folate and vitamin B12 are independent risk factors for neural tube defects. *The Quarterly Journal of Medicine*, 86(11), 703–8.

[130] Kitano, H. (2002). Systems biology: A brief overview. *Science*, 295(5560), 1662–1664.

[131] Kitano, H. (2015). Accelerating systems biology research and its real world deployment. *npj Systems Biology and Applications*, 1(15009).

[132] Kitsak, M., Sharma, A., Menche, J., et al. (2016). Tissue Specificity of Human Disease Module. *Scientific Reports*, 6, 1–12.

[133] Koscielny, G., An, P., Carvalho-Silva, D., et al. (2017). Open Targets: A platform for therapeutic target identification and validation. *Nucleic Acids Research*, 45(D1), D985–D994.

[134] Kronenberg, G., Colla, M., & Endres, M. (2009). Folic Acid, Neurodegenerative and Neuropsychiatric Disease. *Current Molecular Medicine*, 9(3), 315–323.

[135] Kruschwitz, H. L., McDonald, D., Cossins, E. A., & Schirch, V. (1994). 5-Formyltetrahydropteroylpolyglutamates are the major folate derivatives in Neurospora crassa conidiospores. *Journal of Biological Chemistry*, 269(46), 28757–28763.

[136] Kuleshov, M. V., Jones, M. R., Rouillard, A. D., et al. (2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Research*, 44(W1), W90–W97.

[137] Kusminski, C. M., Bickel, P. E., & Scherer, P. E. (2016). Targeting adipose tissue in the treatment of obesity-associated diabetes. *Nature Reviews Drug Discovery*, 15(9), 639–660.

[138] Lage, K., Hansen, N. T., Karlberg, E. O., et al. (2008). A large-scale analysis of tissue-specific pathology and gene expression of human disease genes and complexes. *PNAS*, 105(52), 1–6.

[139] Lamb, J. (2007). The Connectivity map: a new tool for biomedical research. *Nature Reviews Cancer*, 7, 54–60.

[140] Lamb, J., Crawford, E. D., Peck, D., et al. (2006). The Connectivity Map: Using gene-expression signatures to connect small molecules, genes, and disease. *Science*, 313(5795), 1929–1935.

[141] Lamparter, D., Marbach, D., Rueedi, R., Kutalik, Z., & Bergmann, S. (2016). Fast and Rigorous Computation of Gene and Pathway Scores from SNP-Based Summary Statistics. *PLOS Computational Biology*, 12(1), 1–20.

[142] Lan, X., Field, M. S., & Stover, P. J. (2018). Cell cycle regulation of folate-mediated one-carbon metabolism. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, (February), e1426.

[143] Lawrence, M., Huber, W., Pagès, H., et al. (2013). Software for Computing and Annotating Genomic Ranges. *PLOS Computational Biology*, 9(8).

[144] Lee, S., Zhang, C., Kilicarslan, M., et al. (2016). Integrated Network Analysis Reveals an Association between Plasma Mannose Levels and Insulin Resistance. *Cell Metabolism*, 24(1), 172–184.

[145] Lee, S., Zhang, C., Liu, Z., et al. (2017). Network analyses identify liver-specific targets for treating liver diseases. *Molecular Systems Biology*, 13(8), 938.

[146] Li, J., Zheng, S., Chen, B., et al. (2016). A survey of current trends in computational drug repositioning. *Briefings in Bioinformatics*, 17(1), 2–12.

[147] Liang, C., Tian, D., Ren, X., et al. (2018). The development of Bruton's tyrosine kinase (BTK) inhibitors from 2012 to 2017: A mini-review. *European Journal of Medicinal Chemistry*, 151, 315–326.

[148] Liang, P.-H. H. & Anderson, K. S. (1998). Substrate Channeling and Domain-Domain Interactions in Bifunctional Thymidylate Synthase-Dihydrofolate Reductase. *Biochemistry*, 37(35), 12195–12205.

[149] Liew, S. C. & Gupta, E. D. (2015). Methylenetetrahydrofolate reductase (MTHFR) C677T polymorphism: epidemiology, metabolism and the associated diseases. *European Journal of Medical Genetics*, 58, 1–10.

[150] Liljeroos, M., Vuolteenaho, R., Morath, S., et al. (2007). Bruton's tyrosine kinase together with PI 3-kinase are part of Toll-like receptor 2 multiprotein complex and mediate LTA induced Toll-like receptor 2 responses in macrophages. *Cellular Signalling*, 19(3), 625–633.

[151] Lizio, M., Harshbarger, J., Shimoji, H., et al. (2015). Gateways to the FANTOM5 promoter level mammalian expression atlas. *Genome Biology*, 16(1), 1–14.

[152] Longley, D. B., Harkin, D. P., & Johnston, P. G. (2003). 5-Fluorouracil: mechanisms of action and clinical strategies. *Nature Reviews Cancer*, 3(5), 330–338.

[153] Lotfi Shahreza, M., Ghadiri, N., Mousavi, S. R., Varshosaz, J., & Green, J. R. (2017). A review of network-based approaches to drug repositioning. *Briefings in Bioinformatics*, 19(5), 878–892.

[154] Lovelace, L. L., Johnson, S. R., Gibson, L. M., et al. (2009). Variants of human thymidylate synthase with loop 181-197 stabilized in the inactive conformation. *Protein Science*, 18(8), 1628–1636.

[155] Luby-Phelps, K. (2000). Cytoarchitecture and Physical properties of cytoplasm: Volume, Viscosity, Diffusion, Intracellular Surface Area. *Internal Review of Cytology*, 192, 189–221.

[156] Lusis, A. J., Attie, A. D., & Reue, K. (2008). Metabolic syndrome: from epidemiology to systems biology. *Nature Reviews Genetics*, 9, 819–830.

[157] MacArthur, J., Bowler, E., Cerezo, M., et al. (2017). The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Research*, 45(D1), D896–D901.

[158] MacFarlane, A. J., Anderson, D. D., Flodby, P., et al. (2011). Nuclear localization of de Novo thymidylate biosynthesis pathway is required to prevent uracil accumulation in DNA. *Journal of Biological Chemistry*, 286(51), 44015–44022.

[159] Manieri, W., Moore, M. E., Soeller, M. B., Tsang, P., & Caperelli, C. A. (2007). Human Glycinamide Ribonucleotide Transformylase: Active Site Mutants as Mechanistic Probes. *Biochemistry*, 46(1), 156 – 163.

[160] Marbach, D., Lamparter, D., Quon, G., et al. (2016). Tissue-specific regulatory circuits reveal variable modular perturbations across complex diseases. *Nature Methods*, 13, 366–370.

[161] Marchetti, L., Priami, C., & Thanh, V. H. (2016). HRSSA – Efficient hybrid stochastic simulation for spatially homogeneous biochemical reaction networks. *Journal of Computational Physics*, 317, 301–317.

[162] Martiniova, L., Field, M. S., Finkelstein, J. L., Perry, C. A., & Stover, P. (2015). Maternal dietary uridine causes, and deoxyuridine prevents, neural tube closure defects in a mouse model of folate-responsive neural tube defects. *The American Journal of Clinical Nutrition*, 101, 860–869.

[163] Martinov, M. V., Vitvitsky, V. M., Mosharov, E. V., Banerjee, R., & Ataullakhanov, F. I. (2000). A Substrate Switch: A New Mode of Regulation in the Methionine Metabolic Pathway. *Journal of Theoretical Biology*, 204(4), 521–532.

[164] Matthews, R. G., Ghose, C., Green, J. M., Matthews, K. D., & Bruce Dunlap, R. (1987). Folylpolyglutamates as substrates and inhibitors of folate-dependent enzymes. *Advances in Enzyme Regulation*, 26, 157–171.

[165] Matthews, R. G., Ross, J., Baugh, C. M., Cook, J. D., & Davis, L. (1982). Interactions of pig liver serine hydroxymethyltransferase with methyltetrahydropteroylpolyglutamate inhibitors and with tetrahydropteroylpolyglutamate substrates. *Biochemistry*, 21, 1230–1238.

[166] Metz, J., Kelly, A., Swett, V. C., Waxman, S., & Herbert, V. (1968). Deranged DNA synthesis by bone marrow from vitamin B-12-deficient humans. *British Journal of Haematology*, 14(6), 575–592.

[167] Milici, N. (2010). A Short History of the Metabolic Syndrome Definitions. *Proceedings of the Romanian Academy, Series B*, 1, 13–20.

[168] Misselbeck, K., Marchetti, L., Field, M. S., et al. (2017). A hybrid stochastic model of folate-mediated one-carbon metabolism: Effect of the common C677T MTHFR variant on de novo thymidylate biosynthesis. *Scientific Reports*, 7.

[169] Misselbeck, K., Marchetti, L., Priami, C., Stover, P. J., & Field, M. S. (2019). The 5-formyltetrahydrofolate futile cycle reduces pathway stochasticity in an extended hybrid-stochastic model of folate-mediated one-carbon metabolism. *Scientific Reports*, 9.

[170] Misselbeck, K., Parolo, S., Leonardelli, L., et al. (2018). Metabolic syndrome: identification of deregulated pathways and drug effects by network analysis. *Submitted manuscript*.

[171] Molloy, A. M., Kirke, P. N., Troendle, J. F., et al. (2009). Maternal vitamin B12 status and risk of neural tube defects in a population with high neural tube defect prevalence and no folic acid fortification. *Pediatrics*, 123(3), 917–923.

[172] Molloy, A. M., Mills, J. L., Kirke, P. N., Weir, D. G., & Scott, J. M. (1999). Folate status and neural tube defects. *Biofactors*, 10(2–3), 291–294.

[173] Moore, J., Chaudhary, N., & Akinyemiju, T. (2017). Metabolic syndrome prevalence by race/ethnicity and sex in the united states, national health and nutrition examination survey, 1988-2012. *Preventing Chronic Disease*, 14(E24), 1–16.

[174] Morrison, P. F. & Allegra, C. J. (1989). Folate cycle kinetics in human breast cancer cells. *The Journal of Biological Chemistry*, 264(18), 10552–10566.

[175] MRC Vitamin Study Research Group (1991). Prevention of neural tube defects: Results of the Medical Research Council Vitamin Study. *The Lancet*, 338(8760), 131–137.

[176] Nagaraj, A. B., Wang, Q. Q., Joseph, P., et al. (2018). Using a novel computational drug-repositioning approach (DrugPredict) to rapidly identify potent drug candidates for cancer treatment. *Oncogene*, 37, 403–414.

[177] Nakata, M., Nagasaka, S., Kusaka, I., et al. (2006). Effects of statins on the adipocyte maturation and expression of glucose transporter 4 (SLC2A4): Implications in glycaemic control. *Diabetologia*, 49(8), 1881–1892.

[178] Naveja, J. J., Dueñas-González, A., & Medina-Franco, J. L. (2016). Chapter 12 - drug repurposing for epigenetic targets guided by computational methods. In J. L. Medina-Franco (Ed.), *Epi-Informatics* (pp. 327 – 357). Boston: Academic Press.

[179] Neil, G. L. & Homan, E. R. (1973). The effect of dose interval on the survival of L1210 leukemic mice treated with DNA synthesis inhibitors. *Cancer Research*, 33, 895–901.

[180] Nelson, M. R., Tipney, H., Painter, J. L., et al. (2015). The support of human genetic evidence for approved drug indications. *Nature Genetics*, 47, 856–860.

[181] Newman, A. M., Liu, C. L., Green, M. R., et al. (2015). Robust enumeration of cell subsets from tissue expression profiles. *Nature Methods*, 12, 453–457.

[182] Nielsen, J. (2017). Systems Biology of Metabolism. *Annual Review of Biochemistry*, 86, 245–275.

[183] Nielsen, J. & Jewett, M. C. (2008). Impact of systems biology on metabolic engineering of Saccharomyces cerevisiae. *FEMS Yeast Research*, 8(1), 122–131.

[184] Niida, H., Shimada, M., Murakami, H., & Nakanishi, M. (2010). Mechanisms of dNTP supply that play an essential role in maintaining genome integrity in eukaryotic cells. *Cancer science*, 101, 2505–2509.

[185] Nijhout, H. F., Gregory, J. F., Fitzpatrick, C., et al. (2009). A Mathematical Model Gives Insight into the Effects of Vitamin B-6 Deficiency on 1-Carbon and Glutathione Metabolism. *The Journal of Nutrition*, 139(4), 784–791.

[186] Nijhout, H. F., Reed, M. C., Anderson, D. F., et al. (2006a). Long-range allosteric interactions between the folate and methionine cycles stabilize DNA methylation reaction rate. *Epigenetics*, 1(2), 81–87.

[187] Nijhout, H. F., Reed, M. C., Budu, P., & Ulrich, C. M. (2004). A mathematical model of the folate cycle. New insights into folate homeostasis. *Journal of Biological Chemistry*, 279(53), 55008–55016.

[188] Nijhout, H. F., Reed, M. C., Lam, S. L., et al. (2006b). In silico experimentation with a model of hepatic mitochondrial folate metabolism. *Theoretical Biology and Medical Modelling*, 3(40), 1–11.

[189] Nijhout, H. F., Reed, M. C., & Ulrich, C. M. (2008). Chapter 2 Mathematical Models of Metabolism. In *Folic Acid and Folates*, volume 79 of *Vitamins & Hormones* (pp. 45–82). Academic Press.

[190] Obenchain, V., Lawrence, M., Carey, V., et al. (2014). VariantAnnotation: a Bioconductor package for exploration and annotation of genetic variants. *Bioinformatics*, 30(14), 2076–2078.

[191] O'Connor, C. & Adams, J. (2010). *Essentials of cell biology.* Cambridge.

[192] Oresic, M. & Vidal-Puig, A., Eds. (2014). *A Systems Biology Approach to Study Metabolic Syndrome.*

[193] Palmer, A. M., Kamynina, E., Field, M. S., & Stover, P. J. (2017). Folate rescues vitamin B12 depletion-induced inhibition of nuclear thymidylate biosynthesis and genome instability. *PNAS*, 114(20), E4095–E4102.

[194] Palmer, A. M., Misselbeck, K., Marchetti, L., et al. (2018). Maternal vitamin B12 deficiency causes exencephaly in a mouse model of folate-responsive neural tube defects. *in prepartaion.*

[195] Paragh, G., Seres, I., Harangi, M., & Fulop, P. (2014). Dynamic interplay between metabolic syndrome and immunity. In J. Camps (Ed.), *Oxidative Stress and Inflammation in Non-communicable Diseases - Molecular Mechanisms and Perspectives in Therapeutics* (pp. 171–190). Cham: Springer International Publishing.

[196] Pawelek, P. D., Allaire, M., Cygler, M., & MacKenzie, R. E. (2000). Channeling efficiency in the bifunctional methylenetetrahydrofolate dehydrogenase/cyclohydrolase domain: The effects of site-directed mutagenesis of NADP binding residues. *Biochimica et Biophysica Acta - Protein Structure and Molecular Enzymology*, 1479(1–2), 59–68.

[197] Pengde, K., Fuxing, P., Bin, S., Jing, Y., & Jingqiu, C. (2008). Lovastatin inhibits adipogenesis and prevents osteonecrosis in steroid-treated rabbits. *Joint Bone Spine*, 75(6), 696–701.

[198] Pers, T. H., Karjalainen, J. M., Chan, Y., et al. (2015). Biological interpretation of genome-wide association studies using predicted gene functions. *Nature Communications*, 6.

[199] Pogribny, I. P., Basnakian, A. G., Miller, B. J., et al. (1995). Breaks in Genomic DNA and within the p53 Gene Are Associated with Hypomethylation in Livers of Folate/methyl-deficient Rats. *Cancer Research*, 55(9), 1894–1901.

[200] Pons, P. & Latapy, M. (2006). Computing Communities in Large Networks Using Random Walks. *Journal of Graph Algorithms and Applications*, 10(2), 191–218.

[201] Pontarin, G., Gallinaro, L., Ferraro, P., Reichard, P., & Bianchi, V. (2003). Origins of mitochondrial thymidine triphosphate : Dynamic relations to cytosolic pools. *PNAS*, 100(21), 12159–12164.

[202] Poon, P. P. & Storms, R. K. (1994). Thymidylate synthase is localized to the nuclear periphery in the yeast Saccharomyces cerevisiae. *The Journal of Biological Chemistry*, 269, 8341–8347.

[203] Priami, C. (2009). Algorithmic systems biology. *Communications of the ACM*, 52(5), 80–88.

[204] Price, A. J., Travis, R. C., Appleby, P. N., et al. (2016). Circulating Folate and Vitamin B12 and Risk of Prostate Cancer: A Collaborative Analysis of Individual Participant Data from Six Cohorts Including 6875 Cases and 8104 Controls. *European Urology*, 70(6), 941–951.

[205] Prudova, A., Martinov, M. V., Vitvitsky, V. M., Ataullakhanov, F. I., & Banerjee, R. (2005). Analysis of pathological defects in methionine metabolism using a simple mathematical model. *Biochimica et Biophysica Acta – Molecular Basis of Disease*, 1741(3), 331–338.

[206] Pushpakom, S., Iorio, F., Eyers, P. A., et al. (2018). Drug repurposing: progress, challenges and recommendations. *Nature Reviews Drug Discovery*, 18, 41–58.

[207] Quinlivan, E. P., Davis, S. R., Shelnutt, K. P., et al. (2005). Methylenetetrahydrofolate reductase 677C->T polymorphism and folate status affect one-carbon incorporation into human DNA deoxynucleosides. *The Journal of Nutrition*, 135, 389–396.

[208] Radparvar, S., Houghton, P. J., & Houghton, J. a. (1988). Characteristics of thymidylate synthase purified from a human colon adenocarcinoma. *Archives of Biochemistry and Biophysics*, 260(1), 342–350.

[209] Ray, J. G., Wyatt, P. R., Thompson, M. D., et al. (2007). Vitamin B12 and the risk of neural tube defects in a folic-acid-fortified population. *Epidemiology*, 18(3), 362–366.

[210] Raychaudhuri, S., Plenge, R. M., Rossin, E. J., et al. (2009). Identifying relationships among genomic disease regions: Predicting genes at pathogenic SNP associations and rare deletions. *PLOS Genetics*, 5(6).

[211] Rayl, E., Moroson, B., & Beardsley, P. (1996). The Human purH Gene Product, 5-Aminoimidazole-4-carboxamide Ribonucleotide Formyltransferase / IMP Cyclohydrolase. *Journal of Biological Chemistry*, 271(4), 2225–2233.

[212] Reed, M. C., Gamble, M. V., Hall, M. N., & Nijhout, H. F. (2015). Mathematical analysis of the regulation of competing methyltransferases. *BMC Systems Biology*, 9(69).

[213] Reed, M. C., Nijhout, H. F., Neuhouser, M. L., et al. (2006). A Mathematical Model Gives Insights into Nutritional and Genetic Aspects of Folate-Mediated One-Carbon Metabolism. *The Journal of Nutrition*, 136(10), 2653–2661.

[214] Reed, M. C., Nijhout, H. F., Sparks, R., & Ulrich, C. M. (2004). A mathematical model of the methionine cycle. *Journal of Theoretical Biology*, 226(1), 33–43.

[215] Reed, M. C., Thomas, R. L., Pavisic, J., et al. (2008). A mathematical model of glutathione metabolism. *Theoretical biology and Medical Modelling*, 5(8), 1–16.

[216] Reilly, S. M. & Saltiel, A. R. (2017). Adapting to obesity with adipose tissue inflammation. *Nature Reviews Endocrinology*, 13, 633–643.

[217] Rheinheimer, J., de Souza, B. M., Cardoso, N. S., Bauer, A. C., & Crispim, D. (2017). Current role of the NLRP3 inflammasome on obesity and insulin resistance: A systematic review. *Metabolism*, 74, 1–9.

[218] Roadmap Epigenomics Consortium, Kundaje, A., Meuleman, W., et al. (2015). Integrative analysis of 111 reference human epigenomes. *Nature*, 518(7539), 317–329.

[219] Robinson, N., Grabowski, P., & Rehman, I. (2018). Alzheimer's disease pathogenesis: Is there a role for folate? *Mechanisms of Ageing and Development*, 174, 86–94.

[220] Rodan, L. H., Qi, W., Ducker, G. S., et al. (2018). 5,10-Methenyltetrahydrofolate Synthetase Deficiency Causes a Neurometabolic Disorder Associated With Microcephaly, Epilepsy, and Cerebral Hypomyelination. *Molecular Genetics and Metabolism*, 125(1-2), 118–126.

[221] Sagi, I., Chia, G., Golan-Lev, T., et al. (2016). Derivation and differentiation of haploid human embryonic stem cells. *Nature*, 532, 107–111.

[222] Samson, S. L. & Garber, A. J. (2014). Metabolic syndrome. *Endocrinology and Metabolism Clinics*, 43(1), 1–23.

[223] Sanseau, P., Agarwal, P., Barnes, M. R., et al. (2012). Use of genome-wide association studies for drug repositioning. *Nature Biotechnology*, 30, 317–320.

[224] Schaefer, M. H., Fontaine, J. F., Vinayagam, A., et al. (2012). Hippie: Integrating protein interaction networks with experiment based quality scores. *PLOS ONE*, 7(2), 1–8.

[225] Schirch, V. & Strong, W. B. (1989). Interaction of folylpolyglutamates with enzymes in one-carbon metabolism. *Archives of Biochemistry and Biophysics*, 269, 371–380.

[226] Schwahn, B. & Rozen, R. (2001). Polymorphisms in the methylenetetrahydrofolate reductase gene: clinical consequences. *American Journal of Pharmacogenomics*, 1(3), 189–201.

[227] Scotti, M., Stella, L., Shearer, E. J., & Stover, P. J. (2013). Modeling cellular compartmentation in one-carbon metabolism. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, 5(3), 343–365.

[228] Segrë, A. V., Groop, L., Mootha, V. K., Daly, M. J., & Altshuler, D. (2010). Common inherited variation in mitochondrial genes is not enriched for associations with type 2 diabetes or related glycemic traits. *PLOS Genetics*, 6(8).

[229] Seither, R. L., Trent, D. F., Mikulecky, D. C., Rape, T. J., & Goldman, I. D. (1989). Folate-pool interconversions and inhibition of biosynthetic processes after exposure of L1210 leukemia cells to antifolates. *Journal of Biological Chemistry*, 264(29), 17016–17023.

[230] Shane, B. & Stokstad, E. L. R. (1985). Vitamin B12-Folate Interrelationships. *Annual Review of Nutrition*, 5, 115–41.

[231] Sirota, M., Dudley, J. T., Kim, J., et al. (2011). Discovery and Preclinical Validation of Drug Indications Using Compendia of Public Gene Expression Data. *Science Translational Medicine*, 3(96).

[232] So, H. C., Chau, C. K. L., Chiu, W. T., et al. (2017). Analysis of genome-wide association data highlights candidates for drug repositioning in psychiatry. *Nature Neuroscience*, 20, 1342–1349.

[233] Sola, D., Rossi, L., Schianca, G. P. C., et al. (2015). Sulfonylureas and their use in clinical practice. *Archives of Medical Science*, 11(4), 840–848.

[234] Solis, C. (2008). Folate intake at RDA levels is inadequate for Mexican American men with the methylenetetrahydrofolate reductase 677TT genotype. *The Journal of Nutrition*, 138, 67–72.

[235] Srivastava, G. & Apovian, C. M. (2018). Current pharmacotherapy for obesity. *Nature Reviews Endocrinology*, 14, 12–24.

[236] Stabler, S. P. (2013). Clinical practice: Vitamin B12 Deficiency. *New England Journal of Medicine*, 368(2), 149–160.

[237] Stienstra, R., van Diepen, J. A., Tack, C. J., et al. (2011). Inflammasome is a central player in the induction of obesity and insulin resistance. *PNAS*, 108(37), 15324–15329.

[238] Stover, P. (2004). Physiology of folate and vitamin B12 in health and disease. *Nutrition Reviews*, 62, S3–S12.

[239] Stover, P. & Schirch, V. (1991). 5-Formyltetrahydrofolate Polyglutamates Are Slow Tight Binding Inhibitors of Serine Hydroxymethyltransferase. *Journal of Biological Chemistry*, 266(3), 1543–1550.

[240] Stover, P. & Schirch, V. (1992). Enzymatic mechanism for the hydrolysis of 5,10- methenyltetrahydropteroylglutamate to 5-formyltetrahydropteroylglutamate by serine hydroxymethyltransferase. *Biochemistry*, 31(7), 2155–2164.

[241] Stover, P. & Schirch, V. (1993). The metabolic role of leucovorin. *Trends in Biochemical Sciences*, 18(3), 102–106.

[242] Stover, P. J., Durga, J., & Field, M. S. (2017). Folate nutrition and blood–brain barrier dysfunction. *Current Opinion in Biotechnology*, 44, 146–152.

[243] Stover, P. J. & Field, M. S. (2011). Trafficking of Intracellular Folates. *Advances in Nutrition*, 2, 325–331.

[244] Stover, P. J., MacFarlane, A. J., & Field, M. S. (2015). Bringing clarity to the role of MTHFR variants in neural tube defect prevention. *The American Journal of Clinical Nutrition*, 101(6), 1111–1112.

[245] Strong, W. B., Cook, R., & Schirch, V. (1989). Interaction of Tetrahydropteroylpolyglutamates with two Enzymes from Mitochondria. *Biochemistry*, 28(1), 106–114.

[246] Strong, W. B. & Schirch, V. (1989). In vitro conversion of formate to serine: effect of tetrahydropteroylpolyglutamates and serine hydroxymethyltransferase on the rate of 10-formyltetrahydrofolate synthetase. *Biochemistry*, 28(24), 9430–9439.

[247] Strong, W. B., Tendlers, S. J., Seither, R. L., Goldman, D., & Schirch, V. (1990). Purification and properties of serine hydroxymethyltransferase and C1-tetrahydrofolate synthase from L1210 cells. *The Journal of Biological Chemistry*, 265(21), 12149–12155.

[248] Suarez, L., Hendricks, K., Felkner, M., & Gunter, E. (2003). Maternal serum b12 levels and risk for neural tube defects in a texas-mexico border population. *Annals of Epidemiology*, 13(2), 81 – 88.

[249] Subramanian, A., Narayan, R., Corsello, S. M., et al. (2017). A Next Generation Connectivity Map: L1000 Platform and the First 1,000,000 Profiles. *Cell*, 171(6), 1437–1452.

[250] Szebenyi, D. M. E., Liu, X., Kriksunov, I. A., Stover, P. J., & Thiel, D. J. (2000). Structure of a murine cytoplasmic serine hydroxymethyltransferase quinonoid ternary complex: Evidence for asymmetric obligate dimers. *Biochemistry*, 39(44), 13313–13323.

[251] Tan, F., Yang, R., Xu, X., et al. (2014). Drug repositioning by applying 'expression profiles' generated by integrating chemical structure similarity and gene semantic similarity. *Molecular BioSystems*, 10(5), 1126–1138.

[252] The GTEx Consortium (2015). The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. *Science*, 348(6235), 648–660.

[253] The GTEx Consortium (2017). Genetic effects on gene expression across human tissues. *Nature*, 550, 204–213.

[254] Thompson, M. D., Cole, D. E. C., & Ray, J. G. (2009). Vitamin B-12 and neural tube defects: the Canadian experience. *American Journal of Clinical Nutrition*, 89(2), 697S–701S.

[255] Thorndike, J., Gaumont, Y., Kisliuk, R. L., et al. (1989). Inhibition of glycinamide ribonucleotide formyltransferase and other folate enzymes by homofolate polyglutamates in human lymphoma and murine leukemia cell extracts. *Cancer Research*, 49, 158–163.

[256] Thulé, P. M. & Umpierrez, G. (2014). Sulfonylureas: a new look at old therapy. *Current Diabetes Reports*, 14(4), 473.

[257] Tsang, B. L., Devine, O. J., Cordero, A. M., et al. (2015). Assessing the association between the methylenetetrahydrofolate reductase (MTHFR) 677C > T polymorphism and blood folate concentrations: a systematic review and meta-analysis of trials and observational studies. *The American Journal of Clinical Nutrition*, 101, 1286–1294.

[258] Ueland, P. M., Hustad, S., Schneede, J., Refsum, H., & Vollset, S. E. (2001a). Biological and clinical implications of the MTHFR C677T polymorphism. *TRENDS in Pharmacological Sciences*, 22, 195–201.

[259] Ueland, P. M., Nygård, O., Vollset, S. E., & Refsum, H. (2001b). The hordaland homocysteine studies. *Lipids*, 36(1), S33–S39.

[260] Uhlen, M., Fagerberg, L., Hallstrom, B. M., et al. (2015). Tissue-based map of the human proteome. *Science*, 347(6220), 1–9.

[261] Ulrich, C. M., Neuhouser, M., Liu, A. Y., et al. (2008). Mathematical modeling of folate metabolism: predicted effects of genetic polymorphisms on mechanisms and biomarkers relevant to carcinogenesis. *Cancer Epidemiology Biomarkers and Prevention*, 17(7), 1822–1831.

[262] Ursu, O., Holmes, J., Knockel, J., et al. (2017). DrugCentral: Online drug compendium. *Nucleic Acids Research*, 45(D1), D932–D939.

[263] Van Norman, G. A. (2016). Drugs, Devices, and the FDA: Part 1: An Overview of Approval Processes for Drugs. *JACC: Basic to Translational Science*, 1(3), 170–179.

[264] Vella, D., Marini, S., Vitali, F., et al. (2018). MTGO: PPI Network Analysis Via Topological and Functional Module Identification. *Scientific Reports*, 8, 1–13.

[265] Visentin, M., Zhao, R., & Goldman, I. D. (2012). The Antifolates. *Hematology/Oncology Clinics of North America*, 26(3), 629–648.

[266] Visscher, P. M., Wray, N. R., Zhang, Q., et al. (2017). 10 years of gwas discovery: Biology, function, and translation. *The American Journal of Human Genetics*, 101, 5–22.

[267] Vitali, F., Cohen, L. D., Demartini, A., et al. (2016). A Network-Based Data Integration Approach to Support Drug Repurposing and Multi-Target Therapies in Triple Negative Breast Cancer. *PLOS ONE*, 11(9).

[268] Volpato, J. P., Fossati, E., & Pelletier, J. N. (2007). Increasing Methotrexate Resistance by Combination of Active-site Mutations in Human Dihydrofolate Reductase. *Journal of Molecular Biology*, 373(3), 599–611.

[269] Wang, J. Z., Du, Z., Payattakool, R., Yu, P. S., & Chen, C. F. (2007). A new method to measure the semantic similarity of GO terms. *Bioinformatics*, 23(10), 1274–1281.

[270] Weber, A. N. R., Bittner, Z., Liu, X., et al. (2017). Bruton's Tyrosine Kinase: An Emerging Key Player in Innate Immunity. *Frontiers in Immunology*, 8, 1–6.

[271] Wei, S. T., Sun, Y. H., & Zong, S. H. (2017). A novel method to identify hub pathways of rheumatoid arthritis based on differential pathway networks. *Molecular Medicine Reports*, 16(3), 3187–3193.

[272] Wei, W. Q., Cronin, R. M., Xu, H., et al. (2013a). Development and evaluation of an ensemble resource linking medications to their indications. *Journal of the American Medical Informatics Association*, 20(5), 954–961.

[273] Wei, W.-Q., Mosley, J. D., Bastarache, L., & Denny, J. C. (2013b). Validation and enhancement of a computable medication indication resource (MEDI) using a large practice-based dataset. *AMIA Annual Symposium proceedings*, 2013, 1448–56.

[274] Weinblatt, M. E. (2013). Methotrexate in rheumatoid arthritis: a quarter century of development. *Transactions of the American Clinical and Climatological Association*, 124, 16–25.

[275] Welter, D., MacArthur, J., Morales, J., et al. (2014). The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. *Nucleic Acids Research*, 42(D1), 1001–1006.

[276] Wilson, A., Platt, R., Wu, Q., et al. (1999). A common variant in methionine synthase reductase combined with low cobalamin (Vitamin B12) increases risk for spina bifida. *Molecular Genetics and Metabolism*, 67(4), 317–323.

[277] Winer, D. A., Winer, S., Chng, M. H. Y., Shen, L., & Engleman, E. G. (2014). B Lymphocytes in obesity-related adipose tissue inflammation and insulin resistance. *Cellular and Molecular Life Sciences*, 71(6), 1033–1043.

[278] Winer, D. A., Winer, S., Shen, L., et al. (2011). B cells promote insulin resistance through modulation of T cells and production of pathogenic IgG antibodies. *Nature Medicine*, 17, 610–617.

[279] Winter, E. E., Goodstadt, L., & Ponting, C. P. (2004). Elevated rates of protein secretion, evolution, and disease among tissue-specific genes. *Genome Research*, 14, 54–61.

[280] Wishart, D. S., Feunang, Y. D., Guo, A. C., et al. (2018). DrugBank 5.0: A major update to the DrugBank database for 2018. *Nucleic Acids Research*, 46(D1), D1074–D1082.

[281] Wishart, D. S., Knox, C., Guo, A. C., et al. (2008). DrugBank: A knowledgebase for drugs, drug actions and drug targets. *Nucleic Acids Research*, 36(SUPPL. 1), 901–906.

[282] Woeller, C. F., Anderson, D. D., Szebenyi, D. M., & Stover, P. (2007). Evidence for small ubiquitin-like modifier-dependent nuclear import of the thymidylate biosynthesis pathway. *The Journal of Biological Chemistry*, 282, 17623–17631.

[283] Wu, Z., Li, W., Liu, G., & Tang, Y. (2018). Network-Based Methods for Prediction of Drug-Target Interactions. *Frontiers in Pharmacology*, 9(October), 1–14.

[284] Xia, T. (2010). *Network modeling in systems biology*. PhD thesis, Iowa State University.

[285] Xu, X., Qin, X., Li, Y., et al. (2016). Efficacy of Folic Acid Therapy on the Progression of Chronic Kidney Disease: The Renal Substudy of the China Stroke Primary Prevention Trial. *JAMA Internal Medicine*, 176(10), 1443–1450.

[286] Yamada, Y., Takeuchi, S., Yoneda, M., et al. (2017). Atorvastatin reduces cardiac and adipose tissue inflammation in rats with metabolic syndrome. *International Journal of Cardiology*, 240, 332–338.

[287] Young, R. C. & DeVita, V. T. (1970). The effect of chemotherapy on the growth characteristics and cellular kinetics of leukemia L1210. *Cancer Research*, 30, 1789–1794.

[288] Yu, G., Li, F., Qin, Y., et al. (2010). GOSemSim: An R package for measuring semantic similarity among GO terms and gene products. *Bioinformatics*, 26(7), 976–978.

[289] Yu, L., Ma, X., Zhang, L., Zhang, J., & Gao, L. (2016). Prediction of new drug indications based on clinical data and network modularity. *Scientific Reports*, 6.

[290] Zhang, M., Luo, H., Xi, Z., & Rogaeva, E. (2015). Drug repositioning for diabetes based on 'Omics' data mining. *PLOS ONE*, 10(5), 1–13.

[291] Zhang, T., Xin, R., Gu, X., et al. (2009). Maternal serum vitamin B12, folate and homocysteine and the risk of neural tube defects in the offspring in a high-risk area of China. *Public Health Nutrition*, 12(5), 680–686.

[292] Zhao, M., Li, X., Xing, C., & Zhou, B. (2013). Association of methylenetetrahydrofolate reductase C677T and A1298C polymorphisms with colorectal cancer risk: A meta-analysis. *Biomedical Reports*, 1, 781–791.

[293] Zhou, X., Solaroli, N., Bjerke, M., et al. (2008). Progressive loss of mitochondrial DNA in thymidine kinase 2-deficient mice. *Human Molecular Genetics*, 17(15), 2329–2335.

Appendix A

Additional material for Chapter 1

A.1 Supplementary Tables

Table A.1: Initial concentrations of the twelve model variables.

Substrate/Enzyme	µM	Reference
THF	4.61	[215]
10fTHF	3.41	[215]
CHF	0.28	[215]
CH2H	0.51	[215]
DHF	0.039	[215]
5mTHF free	4.5	[215]
5mTHF:SHMT	4.5	[215]
SHMT free	4.5	[215]
HCY	1.12	[215]
MET	49.2	[215]
SAM	81.1	[215]
SAH	19.1	[215]

Table A.2: Concentrations of the constant substrates included in the model.

Constant Substrate	µM	Cell line	Reference
NADPH	58.0	L1210	[229]
NADP+	18.0	L1210	[229]
dUMP	20.0	L1210	[229]
Serine	468.0	L1210	[229]
Glycine	1850.0	L1210	[229]
GAR	10.0	L1210	[229]
AICAR	2.1	L1210	[229]
Formate	200.0	L1210	[229]
Betaine	50.0	–	[215]

Table A.3: Model parameter estimates for the folate cycle grouped by reactions. All concentrations are expressed in µM, while time is expressed in hours. For each value the reference cell line and the length of the glutamate chain of the associated folate are indicated when available in literature.

Parameter	Metabolite	Value	Length of glutamate chain	Cell line	Reference
$R_{AICARFT}$: 10fTHF → THF					
V_{max}		63350		MCF-7	[174]
K_m	10fTHF	0.3	4-6	Human leukemia	[255]
K_m	AICAR	16.8		Human purH	[211]
R_{DHFR} : DHF → THF					
V_{max}		22200		L1210	[229]
K_m	DHF	0.5		L1210	[229]
K_m	NADPH	4.3		L1210	[229]
R_{FTS} : THF → 10fTHF					
V_{max}		45900	6	L1210	[247]
K_m	THF	0.1	5	L1210	[247]
K_m	formate	16	5	L1210	[247]
R_{MTCH} : CHF → 10fTHF					
V_{max}		2916000	5	L1210	[247]
K_m	CHF	4	5	L1210	[247]
R_{MTCH} : 10fTHF → CHF					
V_{max}		2916000	5	L1201	[247]
K_m	10fTHF	20		L1210	[247]
R_{MTD} : CHF → CH2F					
V_{max}		594000	5/6	L1210	[247]
K_m	CHF	6.3		Human DC301	[196]
K_m	NADPH	10.5		Human DC301	[196]
R_{MTD} : CH2F → CHF					
V_{max}		594000	5/6	L1210	[247]
K_m	CH2F	2	5	L1210	[247]
K_m	NADP+	2	5	L1201	[247]
R_{MTHFR} : CH2F → 5mTHF					
V_{max}		120	5	Pig liver	[165]
K_m	CH2F	0.26	5	Pig liver	[165]
K_m	NADPH	125	5	Pig liver	[165]

to continue on next page ...

Table A.3 continued

Parameter	Metabolite	Value	Length of glutamate chain	Cell line	Reference
R_{MTR} : 5mTHF + HCY → THF + MET					
V_{max}		30			Estimated in the range 0.024 [164] - 50 [215] μM/h
K_m	5mTHF	0.5	6	Pig liver	[164]
K_m	HCY	0.1			[213]
R_{PGT} : 10fTHF → THF					
V_{max}		6600			[213]
K_m	10fTHF	0.9		human	[159]
K_m	GAR	1.1		human	[159]
R_{SHMT} : THF → CH2F					
K_m	Serine	600	5	L1210	[247]
K_m	THF	0.2	5/6	L1210	[247]
k_{cat}		18000	5/6	L1210	[247]
R_{SHMT} : CH2F → THF					
K_m	Glycine	3000		L1210	[247]
K_m	CH2F	0.2		L1210	[247]
k_{cat}		45000	4	Rabbit liver	[246]
R_{TYMS} : CH2F → DHF					
V_{max}		4200		L1210	[229]
K_m	CH2F	4.3	1	Human colon	[208]
K_m	dUMP	3.6	1	Human colon	[208]
(un-)binding of 5mTHF and SHMT					
$k_{unbinding}$		1980	3	Rabbit liver	[239]
$k_{binding}$		7200	3	Rabbit liver	[239]

Table A.4: Model parameter estimates for the homocysteine remethylation cycle grouped by reactions. All concentrations are expressed in µM, while time is expressed in hours.

Parameter	Metabolite	Value	Reference
R_{BHMT} : HCY → MET			
V_{max}		2160	[215]
K_m	HCY	12	[215]
K_m	Betaine	100	[215]
R_{DNMT} : SAM → SAH			
V_{max}		180	[215]
K_m	SAM	1.4	[215]
K_i	Inhibition by SAH	1.4	[215]
R_{GNMT} : SAM → SAH			
V_{max}		245	[215]
K_m	SAM	32	[215]
K_m	Glycine	130	[215]
K_i	Inhibition by SAH	18	[215]
R_{MAT-I} : MET → SAM			
V_{max}		260	[215]
K_m	MET	41	[215]
$R_{MAT-III}$: MET → SAM			
V_{max}		220	[215]
K_m	MET	300	[215]
K_a	Activation by SAM	360	[215]
R_{SAHH} : SAH → HCY			
V_{max}		320	[215]
K_m	SAH	6.5	[215]
R_{SAHH} : HCY → SAH			
V_{max}		4530	[215]
K_m	HCY	150	[215]

Table A.5: Stochastic propensities for all model reactions. Reactions are indicated by the enzyme catalyzing them. The propensities are calculated in the model steady state by considering the CC and TT case of the MTHFR polymorphism. Comparison between the two scenarios is provided by their difference expressed in % of CC.

	FTS	**MTCH**		**MTD**	
		10fTHF → CHF	**CHF → 10fTHF**	**CHF → CH2F**	**CH2F → CHF**
CC	6,596,297,718	215,116,909,117	211,683,279,282	25,562,411,625	22,102,671,431
TT	11,313,005,226	285,290,029,076	277,499,278,768	34,615,583,310	26,918,068,257
Difference (% of CC)	71.5	32.6	31.1	35.4	21.8
	AICARFT	**PGT**	**MTHFR**	**SHMT & 5mTHF**	
				Binding	**Unbinding**
CC	1,920,578,414	1,499,148,516	6,132,514	4,248,251,645	4,259,314,801
TT	1,947,149,894	1,558,090,684	2,023,259	1,865,166,039	18,648,587,298
Difference (% of CC)	1.4	3.9	-67	-56.1	-56.2
	MTR	**DHFR**	**TYMS**	**SHMT**	
				CH2F → THF	**THF → CH2F**
CC	6,158,799	115,312,264	74,213,168	1,260,072,867,375,030	344,862,817,955,790
TT	2,017,561	115,312,264	92,071,952	5,397,444,475,851,730	3,193,367,766,042,920
Difference (% of CC)	-67.2	0	24.1	328.3	826
	BHMT	**MAT-I**	**MAT-III**	**GNMT**	**DNMT**
CC	45,761,513	34,996,581	58,188,162	107,600,905	24,807,058
TT	49,927,635	36,950,816	51,905,890	239,168,069	9,876,025
Difference (% of CC)	9.1	5.7	-10.8	122.3	-60.2
	SAHH				
	SAH → HCY	**HCY → SAH**			
CC	81,150,893	29,144,522			
TT	84,535,820	32,582,932			
Difference (% of CC)	4.2	11.8			

A.2 Calculations of dTMP synthesis capacity in mammals and yeast

The capacity of mammalian cells and yeast cells to synthesize sufficient levels of dTMP for DNA replication during S-phase was calculated using results from the model and other values from the literature, as discussed in the preceding sections (see also Table 1.9). Detailed calculations are listed in the following.

Rate of dTMP synthesis based on the computational model (Table 1.4, CC case)

$$263.4\,\frac{\mu\mathrm{M}}{\mathrm{h}} = 4.38\,\frac{\mu\mathrm{M}}{\mathrm{min}}$$

dTMP synthesis required for replication
Based on 59% of bp in human genome being AT [17]

$$3\cdot 10^9\,\mathrm{bp}\cdot 0.59 = 1.77\cdot 10^9\ \mathrm{T\ molecules} = 2.94\cdot 10^{-15}\,\mathrm{mol\,T}$$

are required. Assuming furthermore 8 hour replication time in human ES cells or L1210 cells,

$$3.67\cdot 10^{-16}\,\frac{\mathrm{mol}}{\mathrm{h}} = 1.02\cdot 10^{-19}\,\frac{\mathrm{mol}}{\mathrm{s}}$$

synthesis is required. Assuming ES cell volume of $800\,\mu\mathrm{m}^3 = 8\cdot 10^{-13}\,\mathrm{L}$ the rate of dTMP synthesis required for faithful cell replication is

$$\frac{1.02\cdot 10^{-19}\,\frac{\mathrm{mol}}{\mathrm{s}}}{8\cdot 10^{-13}\,\mathrm{L}} = 1.28\cdot 10^{-7}\,\frac{\mathrm{M}}{\mathrm{s}} = 0.13\,\frac{\mu\mathrm{M}}{\mathrm{s}} = 7.8\,\frac{\mu\mathrm{M}}{\mathrm{min}}$$

Rate of dTMP synthesis in S. cerevisiae based on [100]
dTMP production as measured by ^{3}H-thymidine incorporation.

$$\frac{44.8\,\mu\mathrm{U\ TS\ activity}}{10^8\ \mathrm{haploid\ cells}} = \frac{44.8\,\frac{\mathrm{pmol}}{\mathrm{min}}}{5\cdot 10^7\mathrm{diploid\ cell}} = 8.96\cdot 10^{-7}\,\frac{\frac{\mathrm{pmol}}{\mathrm{min}}}{\mathrm{diploid\ cell}}$$

dTMP synthesis required for replication in S. cerevisiae
By considering the diploid genome with 61.5% AT base pairs

$$12,156,677\,\mathrm{bp}\cdot 0.615 = 7.5\cdot 10^6\ \mathrm{T\ molecules}$$

are required. We assume 150 minute generation time [100]

$$\frac{7.5\cdot 10^6}{150}\,\frac{\frac{\mathrm{T\ molecules}}{\mathrm{min}}}{\mathrm{diploid\ cell}}\cdot\frac{1}{6.02\cdot 10^{23}}\,\frac{\mathrm{mol}}{\mathrm{molecules}} = 8.3\cdot 10^{-20}\,\frac{\frac{\mathrm{mol}}{\mathrm{min}}}{\mathrm{diploid\ cell}} = 8.3\cdot 10^{-8}\,\frac{\frac{\mathrm{pmol}}{\mathrm{min}}}{\mathrm{diploid\ cell}}$$

If we assume that S-phase is 1/3 of cell cycle (50 min), then

$$8.3\cdot 10^{-8}\,\frac{\frac{\mathrm{pmol}}{\mathrm{min}}}{\mathrm{diploid\ cell}}\cdot 3 = 2.5\cdot 10^{-7}\,\frac{\frac{\mathrm{pmol}}{\mathrm{min}}}{\mathrm{diploid\ cell}}$$

is needed to replicate the genome.

Appendix B

Additional material for Chapter 2

Table B.1: Steady state distribution of folate (in percentage of total folate) for different levels of MTR activity (ranging from 3 µM/hto 150 µM/h).The results for the standard MRT activity (30 µM/h) are highlighted in light gray.

Vmax of MTR (µM/h)	**THF**	**10fTHF**	**CHF**	**CH2F**	**DHF**	**5mTHF free**	**5mTHF bound**	**total**
3	0.09	2.03	0.40	0.11	0.00	48.50	48.87	97.37
6	0.14	4.41	0.88	0.25	0.00	45.55	48.77	94.33
9	0.18	7.20	1.43	0.40	0.01	42.13	48.64	90.78
12	0.22	10.52	2.09	0.58	0.01	38.12	48.46	86.58
15	0.25	14.48	2.87	0.80	0.02	33.39	48.20	81.59
18	0.29	19.17	3.80	1.04	0.02	27.89	47.78	75.67
21	0.34	24.58	4.86	1.32	0.03	21.78	47.09	68.87
24	0.42	30.23	5.97	1.58	0.03	15.82	45.95	61.77
27	0.54	35.27	6.94	1.79	0.03	11.13	44.29	55.43
30	0.70	39.24	7.69	1.92	0.04	8.07	42.34	50.42
60	3.13	54.51	10.50	2.32	0.04	1.87	27.63	29.50
150	15.00	58.02	11.10	2.32	0.04	0.53	12.97	13.50

Table B.2: Steady state concentrations of model variables (in µM) for different levels of MTR activity (ranging from 3 µM/hto 150 µM/h). The results for the standard MRT activity (30 µM/h) are highlighted in light gray.

Vmax of MTR (µM/h)	**THF**	**10fTHF**	**CHF**	**CH2F**	**DHF**	**5mTHF free**	**5mTHF bound**	**5mTHF total**
3	0.02	0.36	0.07	0.02	0.00	8.66	8.72	17.38
6	0.03	0.79	0.16	0.04	0.00	8.13	8.71	16.84
9	0.03	1.29	0.26	0.07	0.00	7.52	8.68	16.20
12	0.04	1.88	0.37	0.10	0.00	6.80	8.65	15.45
15	0.05	2.58	0.51	0.14	0.00	5.96	8.60	14.56
18	0.05	3.42	0.68	0.19	0.00	4.98	8.53	13.51
21	0.06	4.39	0.87	0.24	0.01	3.89	8.41	12.29
24	0.08	5.40	1.07	0.28	0.01	2.82	8.20	11.03
27	0.10	6.30	1.24	0.32	0.01	1.99	7.91	9.89
30	0.13	7.00	1.37	0.34	0.01	1.44	7.56	9.00
60	0.56	9.73	1.88	0.41	0.01	0.33	4.93	5.27
150	2.68	10.36	1.98	0.41	0.01	0.10	2.32	2.41

Vmax of MTR (µM/h)	**SHMT free**	**HCY**	**MET**	**SAM**	**SAH**
3	0.28	3.49	34.27	81.89	30.87
6	0.29	3.45	34.73	81.33	31.02
9	0.32	3.41	35.25	80.52	31.35
12	0.35	3.38	35.85	79.33	31.96
15	0.40	3.36	36.58	77.58	33.01
18	0.47	3.35	37.50	74.90	34.77
21	0.60	3.36	38.68	70.79	37.70
24	0.80	3.39	40.07	64.89	42.77
27	1.09	3.44	41.41	57.91	47.77
30	1.44	3.48	42.44	51.34	53.26
60	4.07	3.58	44.59	27.03	75.32
150	6.68	3.62	44.68	18.30	83.93

Table B.3: Steady state fluxes of all model reactions (in µM/h) for different levels of MTR activity (ranging from 3 µM/hto 150 µM/h). Reactions are indicated by the enzyme, which catalyzes them.

Vmax MTR (µM/h)	FTS	MTCH		MTD		MTHFR	MTR	
		10fTHF → CHF	CHF → 10fTHF	CHF → CH2F	CH2F →CHF			
3	5847.5	51775.1	51478.3	5674.0	5377.2	2.8	2.8	
6	8558.4	110304.2	109613.3	12170.8	11479.9	5.5	5.5	
9	10358.2	176142.4	174989.1	19592.4	18439.1	8.2	8.2	
12	11790.9	250284.1	248582.0	28096.8	26394.6	10.9	10.9	
15	13103.0	333619.4	331232.9	37842.8	35456.3	13.4	13.4	
18	14482.1	426086.1	422783.6	48886.9	45584.4	15.9	15.9	
21	16142.4	524546.8	519926.5	60901.8	56281.4	18.1	18.1	
24	18301.7	619536.7	612999.0	72710.1	66172.4	19.8	19.8	
27	20917.1	698125.7	689129.5	82592.9	73596.7	21.0	21.0	
30	23582.5	756270.1	744706.2	89938.7	78374.9	21.7	21.7	
60	36044.3	954342.8	930569.4	115345.6	91572.2	23.3	23.3	
150	40970.8	995200.6	966541.4	120418.2	91759.0	23.4	23.4	

Vmax MTR (µM/h)	SHMT		SHMT & binding	5mTHF unbinding	DHFR	TYMS	PGT	AICARFT
	CH2F →THF	THF →CH2F						
3	438.7	161.5	17271.3	17271.3	16.7	16.7	1704.0	3846.8
6	909.5	260.0	17237.0	17237.0	36.0	36.0	2772.5	5095.0
9	1434.5	347.5	17191.4	17191.3	58.2	58.2	3497.7	5707.3
12	2051.5	444.2	17127.7	17127.7	84.0	84.0	4019.5	6069.3
15	2829.8	570.6	17034.1	17034.0	113.8	113.8	4409.9	6306.6
18	3902.0	763.3	16887.0	16887.0	147.9	147.9	4707.9	6471.6
21	5517.0	1099.4	16642.8	16642.8	184.7	184.7	4933.7	6588.3
24	8027.0	1728.6	16238.6	16238.6	219.4	219.4	5095.9	6668.1
27	11545.6	2816.5	15654.0	15654.0	246.0	246.0	5202.2	6718.7
30	15646.6	4367.7	14964.5	14964.5	263.4	263.4	5268.8	6749.8
60	47065.2	23627.3	9764.4	9764.4	312.2	312.2	5442.5	6828.4
150	77378.2	49055.3	4586.7	4586.7	312.9	312.9	5470.8	6840.8

Vmax MTR (µM/h)	BHMT	GNMT	DNMT	SAHH		MAT-I	MAT-III	
				SAH → HCY	HCY → SAH			
3	158.7	85.0	129.1	264.3	102.9	103.8	57.7	
6	157.4	90.1	128.7	264.6	101.7	104.6	58.2	
9	156.2	96.8	128.0	265.0	100.6	105.6	58.8	
12	155.3	106.1	126.7	265.9	99.8	106.8	59.3	
15	154.8	119.7	124.7	267.4	99.2	108.4	59.8	
18	154.8	140.5	121.4	269.6	98.9	110.4	60.3	
21	155.6	174.0	116.0	272.9	99.3	113.0	60.6	
24	157.2	226.5	107.7	277.3	100.2	116.4	60.6	
27	159.3	296.6	97.3	281.7	101.5	120.0	60.3	
30	161.0	370.9	87.2	285.2	102.6	123.0	59.6	
60	165.6	726.7	46.9	294.6	105.6	132.2	56.8	
150	166.9	887.1	31.8	297.0	106.7	134.6	55.7	

Table B.4: Steady state distribution of folate (in percentage of total folate) for different levels of MTR activity (ranging from 3 µM/h to 150 µM/h) for the model without BHMT activity.The results for the standard MRT activity (30 µM/h) are highlighted in light gray.

Vmax of MTR (µM/h)	**THF**	**10fTHF**	**CHF**	**CH2F**	**DHF**	**5mTHF free**	**5mTHF bound**	**total**
3	0.09	2.07	0.41	0.12	0.00	48.44	48.87	97.31
6	0.09	2.07	0.41	0.12	0.00	48.44	48.87	97.31
9	0.18	7.38	1.47	0.41	0.01	41.91	48.64	90.55
12	0.22	10.81	2.15	0.60	0.01	37.77	48.45	86.22
15	0.25	14.91	2.96	0.82	0.02	32.88	48.17	81.04
18	0.30	19.79	3.92	1.08	0.02	27.18	47.72	74.90
21	0.35	25.36	5.02	1.36	0.03	20.93	46.97	67.89
24	0.44	31.07	6.13	1.62	0.03	14.99	45.72	60.72
27	0.57	36.02	7.08	1.82	0.03	10.51	43.98	54.48
30	0.73	39.86	7.80	1.94	0.04	7.65	41.97	49.62
60	3.24	54.75	10.55	2.32	0.04	1.82	27.29	29.11
150	15.3	57.96	11.09	2.32	0.04	0.52	12.77	13.29

Table B.5: Steady state concentrations of model variables (in µM) for different levels of MTR activity (ranging from 3 µM/h to 150 µM/h) for the model without BHMT activity. The results for the standard MRT activity (30 µM/h) are highlighted in light gray.

Vmax of MTR (µM/h)	**THF**	**10fTHF**	**CHF**	**CH2F**	**DHF**	**5mTHF free**	**5mTHF bound**	**5mTHF total**
3	0.02	0.37	0.07	0.02	0.00	8.65	8.72	17.37
6	0.03	0.80	0.16	0.04	0.00	8.11	8.70	16.81
9	0.03	1.32	0.26	0.07	0.00	7.48	8.68	16.16
12	0.04	1.93	0.38	0.11	0.00	6.74	8.65	15.39
15	0.05	2.66	0.53	0.15	0.00	5.87	8.60	14.47
18	0.05	3.53	0.70	0.19	0.00	4.85	8.52	13.37
21	0.06	4.53	0.90	0.24	0.00	3.74	8.38	12.12
24	0.08	5.55	1.09	0.29	0.01	2.68	8.16	10.84
27	0.10	6.43	1.26	0.32	0.01	1.88	7.85	9.72
30	0.13	7.12	1.39	0.35	0.01	1.37	7.49	8.86
60	0.58	9.77	1.88	0.41	0.01	0.32	4.87	5.19
150	2.73	10.34	1.98	0.41	0.01	0.09	2.28	2.37

Vmax of MTR (µM/h)	**SHMT free**	**HCY**	**MET**	**SAM**	**SAH**
3	0.28	10.74	0.40	1.65	137.72
6	0.30	10.63	0.79	3.27	135.83
9	0.32	10.52	1.19	4.82	134.00
12	0.35	10.41	1.58	6.25	132.29
15	0.40	10.30	1.96	7.48	130.78
18	0.48	10.20	2.33	8.38	129.61
21	0.62	10.12	2.65	8.72	129.03
24	0.84	10.06	2.90	8.36	129.21
27	1.15	10.02	3.06	7.49	129.95
30	1.51	10.00	3.15	6.58	130.80
60	4.13	9.95	3.37	3.46	133.73
150	6.72	9.96	3.36	2.38	134.82

Table B.6: Steady state fluxes of all model reactions (in µM/h) for different levels of MTR activity (ranging from 3 µM/h to 150 µM/h) for the model without BHMT activity.

Vmax MTR (µM/h)	FTS	MTCH		MTD		MTHFR	MTR	
		10fTHF → CHF	CHF → 10fTHF	CHF → CH2F	CH2F →CHF			
3	5915.6	52835.1	52531.4	5790.8	5487.2	2.8	2.8	
6	8640.0	112722.5	112014.9	12441.3	11733.7	5.6	5.6	
9	10449.9	180275.2	179092.2	20062.3	18879.4	8.4	8.4	
12	11896.8	256548.3	254797.7	28822.5	27071.9	11.1	11.1	
15	13234.1	342430.7	339965.6	38884.9	36419.7	13.7	13.7	
18	14660.5	437620.7	434186.3	50281.2	46846.9	16.2	16.2	
21	16406.3	538119.5	533274.2	62577.1	57731.8	18.3	18.3	
24	18682.8	632985.9	626096.4	74395.8	67506.2	20.0	20.0	
27	21378.8	709363.3	699925.5	84011.1	74573.2	21.1	21.1	
30	24051.0	765162.0	753143.3	91063.7	79045.1	21.7	21.7	
60	36226.3	956974.3	933021.6	115689.7	91737.0	23.4	23.4	
150	40998.6	994097.0	965409.0	120257.7	91569.7	23.3	23.3	

Vmax MTR (µM/h)	SHMT		SHMT & 5mTHF		DHFR	TYMS	PGT	AICARFT
	CH2F →THF	THF →CH2F	binding	unbinding				
3	447.4	163.7	17270.7	17270.7	17.1	17.1	1729.2	3882.8
6	928.7	263.5	17235.5	17235.5	36.8	36.8	2805.9	5126.5
9	1467.8	352.8	17188.2	17188.2	59.6	59.6	3533.2	5733.7
12	2106.1	452.8	17121.6	17121.6	86.2	86.2	4054.6	6091.7
15	2920.4	586.0	17022.3	17022.3	117.0	117.0	4443.2	6325.8
18	4059.6	793.6	16864.2	16864.2	152.2	152.2	4738.4	6487.8
21	5801.5	1164.3	16598.0	16598.0	189.7	189.7	4959.7	6601.3
24	8509.0	1863.7	16159.3	16159.3	224.2	224.2	5115.7	6677.6
27	12217.5	3050.3	15541.3	15541.3	249.6	249.6	5215.8	6725.1
30	16426.4	4695.2	14833.2	14833.2	265.8	265.8	5278.3	6754.1
60	47813.7	24197.2	9642.1	9642.1	312.8	312.8	5444.4	6829.2
150	77744.4	49391.9	4513.3	4513.3	312.2	312.2	5470.1	6840.5

Vmax MTR (µM/h)	BHMT	GNMT	DNMT	SAHH		MAT-I	MAT-III	
				SAH → HCY	HCY → SAH			
3	0.0	6.0	2.1	305.6	302.8	2.6	0.2	
6	0.0	12.0	4.2	305.4	299.8	5.0	0.6	
9	0.0	18.2	6.2	305.2	296.8	7.5	0.9	
12	0.0	25.1	8.0	305.0	293.9	9.8	1.3	
15	0.0	33.2	9.6	304.8	291.2	12.0	1.6	
18	0.0	43.5	10.8	304.7	288.6	14.1	2.0	
21	0.0	57.2	11.3	304.7	286.3	16.0	2.4	
24	0.0	74.6	10.8	304.7	284.7	17.4	2.6	
27	0.0	92.9	9.7	304.8	283.7	18.3	2.8	
30	0.0	108.4	8.5	304.9	283.1	18.8	2.9	
60	0.0	157.9	4.5	305.2	281.8	20.2	3.2	
150	0.0	170.7	3.1	305.3	281.9	20.2	3.1	

Appendix C

Additional material for Chapter 3

Table C.1: Model parameter estimates (concentrations are expressed in µM, time is expressed in hours). Parameter estimated not included in the table have been set to the values provided in Tables A.3 and A.4.

Parameter	Metabolite	Value	Length of glutamate chain	Cell line	Reference
R_{MTHFS} : 5fTHF → CHF					
MTHFS		0.08			Estimated to preserve 5fTHF being 5% of total folate [91]
k_{cat}		5400			[83]
K_m	5fTHF	0.2	3		[83]*
K_i)	10fTHF	0.015	3		[83]*
R_{SHMT} : CHF → 5fTHF					
k_{cat}		198			[240]
K_m	CHF	40	4		[240]
(un-)binding of 5fTHF and SHMT					
K_D		0.2		Rabbit liver	[239]
$k_{unbinding}$		144	3	Rabbit liver	[239]
$k_{binding}$		720	3	Rabbit liver	$k_{binding} = k_{unbinding}/K_D$
$R_{AICARFT}$: 10fTHF → THF					
K_i	5fTHF	3	5	MCF-7	[27]
(un-)binding of 5mTHF and SHMT					
K_D		0.4		Rabbit liver	[239]
$k_{unbinding}$		1980	3	Rabbit liver	[239]
$k_{binding}$		4950	3	Rabbit liver	$k_{binding} = k_{unbinding}/K_D$
R_{MTR} : 5mTHF + HCY → THF + MET					
V_{max}		26			Estimated in the range 0.024 [164] -50 [215] µM/h

*The reference value was halved, because the measurements were made with racemic folates, only half of which are physiological substrates.

Table C.2: Steady state concentrations of model variables (in µM) for different levels of MTHFS (ranging from 0.008 µMto 0.4 µM). The results for the standard MTHFS concentration (0.08 µM) are highlighted in light grey.

MTHFS	THF	10fTHF	CHF	CH2F	DHF	5mTHF	
						free	bound
0.008	0.036	6.477	1.290	0.369	0.007	3.812	1.864
0.016	0.041	6.762	1.346	0.383	0.007	4.421	2.508
0.024	0.043	6.878	1.369	0.389	0.007	4.700	2.862
0.032	0.045	6.943	1.381	0.392	0.007	4.864	3.089
0.040	0.047	6.984	1.390	0.394	0.007	4.972	3.249
0.048	0.048	7.013	1.395	0.395	0.007	5.049	3.368
0.056	0.049	7.034	1.399	0.396	0.007	5.107	3.460
0.064	0.050	7.051	1.403	0.397	0.007	5.152	3.534
0.072	0.050	7.064	1.405	0.398	0.007	5.188	3.594
0.080	0.051	7.075	1.407	0.398	0.007	5.217	3.644
0.160	0.053	7.125	1.417	0.400	0.007	5.359	3.893
0.400	0.055	7.158	1.423	0.402	0.007	5.452	4.067

MTHFS	5fTHF		SHMT	HCY	MET	SAM	SAH
	free	bound	free				
0.008	2.495	2.440	0.196	3.287	39.263	70.648	37.322
0.016	1.556	1.765	0.227	3.249	38.727	73.282	35.261
0.024	1.146	1.395	0.244	3.233	38.502	74.335	34.450
0.032	0.910	1.157	0.254	3.224	38.376	74.912	34.009
0.040	0.757	0.989	0.261	3.218	38.294	75.278	33.729
0.048	0.648	0.865	0.267	3.214	38.237	75.533	33.535
0.056	0.567	0.769	0.271	3.211	38.195	75.721	33.393
0.064	0.504	0.692	0.274	3.209	38.163	75.865	33.284
0.072	0.454	0.629	0.277	3.207	38.137	75.979	33.197
0.080	0.413	0.577	0.279	3.206	38.116	76.071	33.127
0.160	0.217	0.316	0.291	3.199	38.016	76.506	32.799
0.400	0.090	0.134	0.298	3.194	37.952	76.783	32.590

Table C.3: Steady state distribution of folate (in percentage of total folate) for different levels of MTHFS (ranging from 0.008 µMto 0.4 µM). The results for the standard MTHFS concentration (0.08 µM) are highlighted in light grey.

MTHFS	THF	10fTHF	CHF	CH2F	DHF	5mTHF			5fTHF		
						free	bound	total	free	bound	total
0.008	0.19	34.47	6.87	1.96	0.04	20.29	9.92	30.21	13.28	12.99	26.26
0.016	0.22	35.99	7.16	2.04	0.04	23.53	13.35	36.87	8.28	9.40	17.68
0.024	0.23	36.61	7.29	2.07	0.04	25.02	15.23	40.25	6.10	7.42	13.52
0.032	0.24	36.95	7.35	2.09	0.04	25.89	16.44	42.33	4.85	6.16	11.00
0.040	0.25	37.17	7.40	2.10	0.04	26.46	17.29	43.75	4.03	5.27	9.29
0.048	0.26	37.32	7.43	2.10	0.04	26.87	17.93	44.80	3.45	4.60	8.05
0.056	0.26	37.44	7.45	2.11	0.04	27.18	18.42	45.59	3.02	4.09	7.11
0.064	0.26	37.53	7.46	2.11	0.04	27.42	18.81	46.23	2.68	3.68	6.37
0.072	0.27	37.60	7.48	2.12	0.04	27.61	19.13	46.74	2.42	3.35	5.77
0.080	0.27	37.65	7.49	2.12	0.04	27.77	19.39	47.16	2.20	3.07	5.27
0.160	0.28	37.92	7.54	2.13	0.04	28.52	20.72	49.24	1.16	1.68	2.84
0.400	0.29	38.10	7.58	2.14	0.04	29.02	21.65	50.66	0.48	0.71	1.19

Table C.4: Steady state fluxes of model reactions catalyzed by the enzymes FTS, MTCH, MTD, MTHFR, SHMT and the (un-)binding of SHMT and 5mTHF/5fTHF (in µM/h) for different levels of MTHFS (ranging from 0.008 µM to 0.4 µM). Reactions are indicated by the enzyme, which catalyzes them. For bidirectional reactions the direction is indicated behind the enzyme name (Enzyme: Substrate → Product). The results for the standard MTHFS concentration (0.08 µM) are highlighted in light grey.

	FTS	**MTCH**		**MTD**		**MTHFR**	**MTR**
		10fTHF → CHF	**CHF → 10fTHF**	**CHF → CH2F**	**CH2F →CHF**		
0.008	11141.5	713316.8	711068.9	85479.3	83231.4	22.3	22.3
0.016	12257.1	736798.2	734226.5	88544.9	85973.3	22.7	22.7
0.024	12874.6	746211.2	743475.4	89774.8	87039.0	22.8	22.8
0.032	13277.1	751402.1	748565.0	90452.9	87615.8	22.9	22.9
0.040	13563.0	754720.3	751813.9	90886.3	87979.8	22.9	22.9
0.048	13777.5	757034.3	754077.1	91188.4	88231.3	22.9	22.9
0.056	13944.8	758743.9	755748.1	91411.6	88415.8	23.0	23.0
0.064	14079.1	760060.3	757033.9	91583.4	88557.1	23.0	23.0
0.072	14189.4	761106.1	758055.0	91719.9	88668.8	23.0	23.0
0.080	14281.6	761957.5	758886.0	91831.0	88759.5	23.0	23.0
0.160	14749.0	765971.0	762799.2	92354.5	89182.6	23.1	23.1
0.400	15080.7	768546.7	765306.9	92690.3	89450.5	23.1	23.1

	SHMT			**SHMT & 5mTHF**		**SHMT & 5fTHF**	
	CH2F →THF	**THF →CH2F**	**CHF →5fTHF**	**binding**	**un-binding**	**binding**	**un-binding**
0.008	2177.2	232.8	1.2	3691.5	3691.5	351.4	351.4
0.016	2559.3	301.6	1.5	4965.2	4965.2	254.2	254.2
0.024	2760.6	342.9	1.6	5665.9	5665.9	200.9	200.9
0.032	2887.8	371.0	1.7	6116.9	6116.9	166.6	166.6
0.040	2976.3	391.5	1.7	6433.6	6433.6	142.5	142.5
0.048	3041.7	407.2	1.8	6669.1	6669.1	124.6	124.6
0.056	3092.0	419.6	1.8	6851.3	6851.3	110.7	110.7
0.064	3132.1	429.6	1.8	6996.7	6996.7	99.6	99.6
0.072	3164.7	438.0	1.9	7115.4	7115.4	90.6	90.6
0.080	3191.9	445.0	1.9	7214.3	7214.3	83.1	83.1
0.160	3326.8	481.3	2.0	7709.0	7709.0	45.5	45.5
0.400	3420.2	507.7	2.0	8053.3	8053.3	19.3	19.3

Table C.5: Steady state fluxes of model reactions catalyzed by the enzymes PGT, AICARFT, TYMS, DHFR, MTHFS, BHMT, GNMT, DNMT, SAHH, MAT-I and MAT-III (in µM/h) for different levels of MTHFS (ranging from 0.008 µM to 0.4 µM). Reactions are indicated by the enzyme, which catalyzes them. For bidirectional reactions the direction is indicated behind the enzyme name (Enzyme: Substrate → Product). The results for the standard MTHFS concentration (0.08 µM) are highlighted in light grey.

	PGT	**AICARFT**	**DHFR**	**TYMS**	**MTHFS**	**BHMT**
0.008	5220.5	3673.1	281.2	281.2	1.2	153.1
0.016	5247.5	4437.9	291.3	291.3	1.5	151.5
0.024	5258.0	4880.9	295.2	295.2	1.6	150.8
0.032	5263.6	5176.4	297.4	297.4	1.7	150.5
0.040	5267.2	5389.3	298.7	298.7	1.7	150.2
0.048	5269.7	5550.7	299.7	299.7	1.8	150.1
0.056	5271.5	5677.5	300.4	300.4	1.8	149.9
0.064	5272.9	5779.9	300.9	300.9	1.8	149.8
0.072	5274.0	5864.3	301.3	301.3	1.9	149.8
0.080	5274.9	5935.2	301.7	301.7	1.9	149.7
0.160	5279.1	6298.0	303.2	303.2	2.0	149.4
0.400	5281.8	6559.0	304.2	304.2	2.0	149.2

	GNMT	**DNMT**	**SAHH**		**MAT-I**	**MAT-III**
			SAH → HCY	**HCY → SAH**		
0.008	177.0	116.3	272.5	97.1	113.9	61.5
0.016	155.9	120.0	270.2	96.0	112.6	61.6
0.024	147.8	121.4	269.2	95.6	112.0	61.6
0.032	143.5	122.2	268.7	95.3	111.7	61.7
0.040	140.7	122.7	268.3	95.2	111.5	61.7
0.048	138.8	123.1	268.0	95.0	111.3	61.7
0.056	137.4	123.3	267.9	95.0	111.2	61.7
0.064	136.4	123.5	267.7	94.9	111.2	61.7
0.072	135.5	123.7	267.6	94.8	111.1	61.7
0.080	134.9	123.8	267.5	94.8	111.0	61.7
0.160	131.7	124.4	267.1	94.6	110.8	61.7
0.400	129.7	124.8	266.8	94.5	110.6	61.7

Table C.6: Stochastic propensities for all model reactions, calculated for model steady states considering the MTHFR polymorphism (CC and TT), folate status (replete, 19 uM; low, 9 uM) and the presence of the 5fTHF futile cycle (FOCM network with and without futile cycle).

		FTS	**MTCH:**		**MTD:**		**MTHFR**	**PGT**	**AICARFT**	**DHFR**
			10fTHF → CHF	CHF → 10fTHF	CHF → CH2F	CH2F → CHF				
Replete folate, with 5fTHF futile cycle	**CC**	$4.06{\cdot}10^{9}$	$2.17{\cdot}10^{11}$	$2.16{\cdot}10^{11}$	$2.61{\cdot}10^{10}$	$2.53{\cdot}10^{10}$	$6.55{\cdot}10^{6}$	$1.50{\cdot}10^{9}$	$1.69{\cdot}10^{9}$	$8.58{\cdot}10^{7}$
	TT	$4.48{\cdot}10^{9}$	$2.83{\cdot}10^{11}$	$2.81{\cdot}10^{11}$	$3.51{\cdot}10^{10}$	$3.35{\cdot}10^{10}$	$2.22{\cdot}10^{6}$	$1.56{\cdot}10^{9}$	$1.29{\cdot}10^{9}$	$1.18{\cdot}10^{8}$
Replete folate, without 5fTHF futile cycle	**CC**	$4.45{\cdot}10^{9}$	$2.15{\cdot}10^{11}$	$2.14{\cdot}10^{11}$	$2.59{\cdot}10^{10}$	$2.49{\cdot}10^{10}$	$6.50{\cdot}10^{6}$	$1.50{\cdot}10^{9}$	$1.92{\cdot}10^{9}$	$8.45{\cdot}10^{7}$
	TT	$9.73{\cdot}10^{9}$	$3.20{\cdot}10^{11}$	$3.14{\cdot}10^{11}$	$3.99{\cdot}10^{10}$	$3.37{\cdot}10^{10}$	$2.23{\cdot}10^{6}$	$1.58{\cdot}10^{9}$	$1.96{\cdot}10^{9}$	$1.19{\cdot}10^{8}$
Low folate, with 5fTHF futile cycle	**CC**	$4.71{\cdot}10^{9}$	$1.28{\cdot}10^{11}$	$1.26{\cdot}10^{11}$	$1.46{\cdot}10^{10}$	$1.29{\cdot}10^{10}$	$4.51{\cdot}10^{6}$	$1.36{\cdot}10^{9}$	$1.65{\cdot}10^{9}$	$4.19{\cdot}10^{7}$
	TT	$5.08{\cdot}10^{9}$	$1.43{\cdot}10^{11}$	$1.41{\cdot}10^{11}$	$1.65{\cdot}10^{10}$	$1.44{\cdot}10^{10}$	$1.45{\cdot}10^{6}$	$1.39{\cdot}10^{9}$	$1.57{\cdot}10^{9}$	$4.70{\cdot}10^{7}$
Low folate, without 5fTHF futile cycle	**CC**	$5.69{\cdot}10^{9}$	$1.36{\cdot}10^{11}$	$1.34{\cdot}10^{11}$	$1.56{\cdot}10^{10}$	$1.31{\cdot}10^{10}$	$4.56{\cdot}10^{6}$	$1.38{\cdot}10^{9}$	$1.86{\cdot}10^{9}$	$4.27{\cdot}10^{7}$
	TT	$8.46{\cdot}10^{9}$	$1.91{\cdot}10^{11}$	$1.85{\cdot}10^{11}$	$2.21{\cdot}10^{10}$	$1.70{\cdot}10^{10}$	$1.60{\cdot}10^{6}$	$1.47{\cdot}10^{9}$	$1.91{\cdot}10^{9}$	$5.61{\cdot}10^{7}$

		TYMS	**SHMT:**		**SHMT:5mTHF**		**SHMT:5fTHF**		**MTHFS**
			CH2F → THF	THF → CH2F	CHF → 5fTHF	**binding**	**unbinding**	**binding**	**unbinding**	
Replete folate, with 5fTHF futile cycle	**CC**	$8.58{\cdot}10^{7}$	$2.58{\cdot}10^{14}$	$3.60{\cdot}10^{13}$	$5.35{\cdot}10^{5}$	$2.05{\cdot}10^{9}$	$2.05{\cdot}10^{9}$	$2.36{\cdot}10^{7}$	$2.36{\cdot}10^{7}$	$5.35{\cdot}10^{5}$
	TT	$1.18{\cdot}10^{8}$	$5.02{\cdot}10^{14}$	$7.11{\cdot}10^{13}$	$1.34{\cdot}10^{6}$	$1.54{\cdot}10^{8}$	$1.54{\cdot}10^{8}$	$1.53{\cdot}10^{8}$	$1.53{\cdot}10^{8}$	$1.34{\cdot}10^{6}$
Replete folate, without 5fTHF futile cycle	**CC**	$8.45{\cdot}10^{7}$	$3.17{\cdot}10^{14}$	$4.96{\cdot}10^{13}$	0	$2.34{\cdot}10^{9}$	$2.34{\cdot}10^{9}$	0	0	0
	TT	$1.19{\cdot}10^{8}$	$2.97{\cdot}10^{15}$	$1.24{\cdot}10^{15}$	0	$9.09{\cdot}10^{8}$	$9.09{\cdot}10^{8}$	0	0	0
Low folate, with 5fTHF futile cycle	**CC**	$4.19{\cdot}10^{7}$	$6.13{\cdot}10^{14}$	$1.42{\cdot}10^{14}$	$9.09{\cdot}10^{5}$	$1.08{\cdot}10^{9}$	$1.08{\cdot}10^{9}$	$6.82{\cdot}10^{7}$	$6.82{\cdot}10^{7}$	$9.09{\cdot}10^{5}$
	TT	$4.70{\cdot}10^{7}$	$7.71{\cdot}10^{14}$	$1.84{\cdot}10^{14}$	$1.23{\cdot}10^{6}$	$1.92{\cdot}10^{8}$	$1.92{\cdot}10^{8}$	$1.26{\cdot}10^{8}$	$1.26{\cdot}10^{8}$	$1.23{\cdot}10^{6}$
Low folate, without 5fTHF futile cycle	**CC**	$4.27{\cdot}10^{7}$	$9.63{\cdot}10^{14}$	$2.80{\cdot}10^{14}$	0	$1.73{\cdot}10^{9}$	$1.73{\cdot}10^{9}$	0	0	0
	TT	$5.61{\cdot}10^{7}$	$2.57{\cdot}10^{15}$	$1.14{\cdot}10^{15}$	0	$6.65{\cdot}10^{8}$	$6.65{\cdot}10^{8}$	0	0	0

		MTR	**SAHH:**		**BHMT**	**MAT-I**	**MAT-III**	**GNMT**	**DNMT**
			SAH → HCY	HCY → SAH					
Replete folate, with 5fTHF futile cycle	**CC**	$6.55{\cdot}10^{6}$	$7.61{\cdot}10^{7}$	$2.70{\cdot}10^{7}$	$4.26{\cdot}10^{7}$	$3.16{\cdot}10^{7}$	$6.49{\cdot}10^{7}$	$3.95{\cdot}10^{7}$	$3.52{\cdot}10^{7}$
	TT	$2.22{\cdot}10^{6}$	$8.44{\cdot}10^{7}$	$3.24{\cdot}10^{7}$	$4.97{\cdot}10^{7}$	$3.68{\cdot}10^{7}$	$5.22{\cdot}10^{7}$	$2.27{\cdot}10^{8}$	$1.10{\cdot}10^{7}$
Replete folate, without 5fTHF futile cycle	**CC**	$6.50{\cdot}10^{6}$	$7.65{\cdot}10^{7}$	$2.71{\cdot}10^{7}$	$4.28{\cdot}10^{7}$	$3.18{\cdot}10^{7}$	$6.45{\cdot}10^{7}$	$4.23{\cdot}10^{7}$	$3.47{\cdot}10^{7}$
	TT	$2.23{\cdot}10^{6}$	$8.44{\cdot}10^{7}$	$3.24{\cdot}10^{7}$	$4.97{\cdot}10^{7}$	$3.68{\cdot}10^{7}$	$5.22{\cdot}10^{7}$	$2.27{\cdot}10^{8}$	$1.10{\cdot}10^{7}$
Low folate, with 5fTHF futile cycle	**CC**	$4.51{\cdot}10^{6}$	$8.26{\cdot}10^{7}$	$3.06{\cdot}10^{7}$	$4.76{\cdot}10^{7}$	$3.58{\cdot}10^{7}$	$5.55{\cdot}10^{7}$	$1.47{\cdot}10^{8}$	$1.96{\cdot}10^{7}$
	TT	$1.45{\cdot}10^{6}$	$8.47{\cdot}10^{7}$	$3.29{\cdot}10^{7}$	$5.03{\cdot}10^{7}$	$3.69{\cdot}10^{7}$	$5.17{\cdot}10^{7}$	$2.46{\cdot}10^{8}$	$9.19{\cdot}10^{6}$
Low folate, without 5fTHF futile cycle	**CC**	$4.56{\cdot}10^{6}$	$8.26{\cdot}10^{7}$	$3.05{\cdot}10^{7}$	$4.75{\cdot}10^{7}$	$3.58{\cdot}10^{7}$	$5.56{\cdot}10^{7}$	$1.45{\cdot}10^{8}$	$1.99{\cdot}10^{7}$
	TT	$1.60{\cdot}10^{6}$	$8.46{\cdot}10^{7}$	$3.28{\cdot}10^{7}$	$5.02{\cdot}10^{7}$	$3.69{\cdot}10^{7}$	$5.18{\cdot}10^{7}$	$2.42{\cdot}10^{8}$	$9.51{\cdot}10^{6}$	

Table C.7: Comparison between the propensities provided in Table C.6. with respect to the presence/absence of the 5fTHF futile cycle. The difference propensity_without_futile_cycle - propensity_with_futile_cycle is expressed in % of the scenario without futile cycle.

		FTS	**MTCH:**		**MTD:**		**MTHFR**	**PGT**	**AICARFT**	**DHFR**
			10fTHF → CHF	CHF → 10fTHF	CHF → CH2F	CH2F → CHF				
Replete folate	**CC**	8.68	-0.64	-0.71	-0.79	-1.45	-0.68	-0.10	12.07	-1.59
	TT	53.94	11.72	10.5	11.99	0.67	0.27	1.42	34.07	0.76
Low folate	**CC**	17.16	6.33	5.89	6.24	1.74	1.11	1.48	11.31	1.82
	TT	40.02	24.68	23.75	25.33	15.42	9.44	5.29	17.56	16.24
		TYMS	**SHMT:**			**SHMT:5mTHF**		**SHMT:5fTHF**		**MTHFS**
			CH2F → THF	THF → CH2F	CHF → 5fTHF	**binding**	**unbinding**	**binding**	**unbinding**	
Replete folate	**CC**	-1.59	18.38	27.32	NaN	12.32	12.32	NaN	NaN	NaN
	TT	0.76	83.08	94.27	NaN	83.11	83.11	NaN	NaN	NaN
Low folate	**CC**	1.82	36.34	49.41	NaN	37.62	37.62	NaN	NaN	NaN
	TT	16.24	70.01	83.86	NaN	71.15	71.15	NaN	NaN	NaN
		MTR	**SAHH:**		**BHMT**	**MAT-I**	**MAT-III**	**GNMT**	**DNMT**	
			SAH → HCY	HCY → SAH						
Replete folate	**CC**	-0.68	0.47	0.61	0.56	0.63	-0.57	6.51	-1.43	
	TT	0.27	0.00	-0.01	-0.01	0.00	0.01	-0.07	0.14	
Low folate	**CC**	1.11	-0.08	-0.17	-0.14	-0.13	0.22	-1.54	1.34	
	TT	9.44	-0.06	-0.29	-0.21	-0.02	0.19	-1.38	3.40	

Table C.8: Comparison between the propensities provided in Table C.6 regarding the MTHFR polymorphism. The difference propensity_CC - propensity_TT is expressed in % of CC.

		FTS	**MTCH:**		**MTD:**		**MTHFR**	**PGT**	**AICARFT**	**DHFR**
			10fTHF → CHF	CHF → 10fTHF	CHF → CH2F	CH2F → CHF				
Replete folate	**with futile cycle**	10.28	30.45	30.22	34.49	32.67	-66.04	3.70	-23.63	37.03
	without futile cycle	118.66	48.72	46.55	54.01	35.50	-65.71	5.30	1.85	40.28
Low folate	**with futile cycle**	-7.75	-12.27	-12.12	-12.92	-11.47	67.89	-2.63	4.75	-12.08
	without futile cycle	-48.83	-39.62	-38.38	-41.79	-29.51	64.94	-6.75	-2.46	-31.38
		TYMS	**SHMT:**			**SHMT:5mTHF**		**SHMT:5fTHF**		**MTHFS**
			CH2F → THF	THF → CH2F	CHF → 5fTHF	**binding**	**un-binding**	**binding**	**un-binding**	
Replete folate	**with futile cycle**	37.03	94.39	97.31	151.39	-92.52	-92.52	547.76	547.76	151.39
	without futile cycle	40.28	837.57	2402.38	NaN	-61.16	-61.16	NaN	NaN	NaN
Low folate	**with futile cycle**	-12.08	-25.78	-30.06	-35.39	82.24	82.24	-84.55	-84.55	-35.39
	without futile cycle	-31.38	-167.01	-307.64	NaN	61.6	61.6	NaN	NaN	NaN
		MTR	**SAHH:**		**BHMT**	**MAT-I**	**MAT-III**	**GNMT**	**DNMT**	
			SAH → HCY	HCY → SAH						
Replete folate	**with futile cycle**	-66.04	10.82	20.14	16.73	16.57	-19.48	474.5	-68.74	
	without futile cycle	-65.71	10.30	19.40	16.07	15.83	-19.02	436.72	-68.25	
Low folate	**with futile cycle**	67.89	-2.45	-7.68	-5.76	-3.01	6.96	-66.96	53.23	
	without futile cycle	64.94	-2.47	-7.54	-5.68	-3.13	6.99	-67.23	52.23	

Table C.9: Comparison between the propensities provided in Table C.6 with respect to folate status. The difference propensity_replete_folate – propensity_low_folate is expressed in % of replete folate.

		FTS	**MTCH:**		**MTD:**		**MTHFR**	**PGT**	**AICARFT**	**DHFR**
			10fTHF → CHF	CHF → 10fTHF	CHF → CH2F	CH2F → CHF				
with futile cycle	**CC**	-15.93	41.05	41.6	44.05	48.86	31.16	9.62	2.27	51.18
	TT	-13.26	49.27	49.72	53.02	57.03	34.92	10.55	-21.9	60.07
without futile cycle	**CC**	-27.8	36.67	37.5	39.85	47.2	29.91	8.17	3.11	49.49
	TT	13.02	40.54	40.99	44.63	49.54	28.33	6.90	2.52	52.69
		TYMS	**SHMT:**			**SHMT:5mTHF**		**SHMT:5fTHF**		**MTHFS**
			CH2F → THF	THF → CH2F	CHF → 5fTHF	**binding**	**un-binding**	**binding**	**un-binding**	
with futile cycle	**CC**	51.18	-137.29	-293.08	-70.00	47.37	47.37	-188.62	-188.62	-70.00
	TT	60.07	-53.53	-159.11	8.44	-24.89	-24.89	17.77	17.77	8.44
without futile cycle	**CC**	49.49	-204.21	-464.71	NaN	26.03	26.03	NaN	NaN	NaN
	TT	52.69	13.36	8.01	NaN	26.86	26.86	NaN	NaN	NaN
		MTR	**SAHH:**		**BHMT**	**MAT-I**	**MAT-III**	**GNMT**	**DNMT**	
			SAH → HCY	HCY → SAH						
with futile cycle	**CC**	31.16	-8.55	-13.33	-11.63	-13.37	14.4	-272.39	44.24	
	TT	34.92	-0.35	-1.58	-1.13	-0.18	1.09	-8.22	16.58	
without futile cycle	**CC**	29.91	-7.96	-12.45	-10.86	-12.51	13.72	-242.85	42.67	
	TT	28.33	-0.30	-1.29	-0.93	-0.17	0.91	-6.82	13.76	

Appendix D

Additional material for Chapter 4

Table D.1: Parameter estimates for the extended model including cytoplasmic and nuclear FOCM grouped by reactions. All concentrations are expressed in µM, while time is expressed in hours. The parameter estimates have been calculated by scaling the respective cytoplasmic values to the nuclear volume. The estimates for the reactions forming the remethylation of homocysteine cycle are listed in Appendix A, Table A.4

Parameter	**Metabolite**	**Value**		**Cell line**	**Reference**
		cytoplasm	**nucleus**		
Volume		$9.4 \cdot 10^{13}$	$2.2 \cdot 10^{13}$	HeLa	[86]
$R_{AICARFT}$: 10fTHF → THF					
V_{max}		63350		MCF-7	[174]
K_m	10fTHF	0.3		Human leukemia	[255]
K_m	AICAR	16.8		Human purH	[211]
K_i	5fTHF	3			[27]
R_{DHFR} : DHF → THF					
k_{cat}		36000	36000	Human	[268]
K_m	DHF	0.5	2.14	L1210	[229]
K_m	NADPH	4.3	18.37	L1210	[229]
R_{FTS} : THF → 10fTHF					
k_{cat}		5100	5100	L1210	[247]
K_m	THF	0.1	0.43	L1210	[247]
K_m	formate	16	68.36	L1210	[247]
R_{MTCH} : CHF → 10fTHF					
k_{cat}		324000	324000	L1210	[247]
K_m	CHF	4	17.09	L1210	[247]
R_{MTCH} : 10fTHF → CHF					

to continue on next page ...

Table D.1 continued

Parameter	Metabolite	Value		Cell line	Reference
		cytoplasm	nucleus		
k_{cat}		324000	324000	L1210	[247]
K_m	10fTHF	20	85.45	L1210	[247]
R_{MTD} : CHF → CH2F					
k_{cat}		66000	66000	L1210	[247]
K_m	CHF	6.3	26.92	Human DC301	[196]
K_m	NADPH	10.5	44.86	Human DC301	[196]
R_{MTD} : CH2F → CHF					
k_{cat}		66000	66000	L1210	[247]
K_m	CH2F	2	8.55	L1210	[247]
K_m	NADP+	2	8.55	L1201	[247]
R_{MTHFR} : CH2F → 5mTHF					
V_{max}		120		Pig liver	[164]
K_m	CH2F	0.26		Pig liver	[164]
K_m	NADPH	125		Pig liver	[164]
R_{MTHFS} : 5fTHF → CHF					
k_{cat}		5400	5400		[83]
K_m	5fTHF	0.2	0.85		[83]
K_i	10fTHF	0.015	0.64		[83]
R_{MTR} : 5mTHF + HCY → THF + MET					
V_{max}		26			Estimated in the range 0.024[164] - 50[215] μM/h
K_m	5mTHF	0.5		Pig liver	[164]
K_m	HCY	0.1			[213]
R_{PGT} : 10fTHF → THF					
V_{max}		6600			[213]
K_m	10fTHF	0.9		human	[159]
K_m	GAR	1.1		human	[159]
R_{SHMT} : THF → CH2F					
k_{cat}		18000	18000	L1210	[247]
K_m	Serine	600	2563.6	L1210	[247]
K_m	THF	0.2	0.85	L1210	[247]
R_{SHMT} : CH2F → THF					

to continue on next page ...

Table D.1 continued

Parameter	Metabolite	Value		Cell line	Reference
		cytoplasm	nucleus		
k_{cat}		45000	45000	Rabbit liver	[246]
K_m	Glycine	3000	12818	L1210	[247]
K_m	CH2F	0.2	0.85	L1210	[247]
R_{SHMT} : CHF $\rightarrow$ 5fTHF					
k_{cat}		198	198		[240]
K_m	CHF	40	170.9		[240]
R_{TYMS} : CH2F $\rightarrow$ DHF					
k_{cat}		11196	11196	Human	[154]
K_m	CH2F	4.3	18.37	Human colon	[208]
K_m	dUMP	3.6	15.38	Human colon	[208]
(un-)binding of 5mTHF and SHMT					
$k_{unbinding}$		1980	1980	Rabbit liver	[239]
k_D		0.4	1.71	Rabbit liver	[239]
$k_{binding}$		4950	1158	Rabbit liver	$k_{binding} = \frac{k_{unbinding}}{K_D}$
(un-)binding of 5fTHF and SHMT					
$k_{unbinding}$		144	144	Rabbit liver	[239]
k_D		0.2	0.85	Rabbit liver	[239]
$k_{binding}$		720	169	Rabbit liver	$k_{binding} = \frac{k_{unbinding}}{K_D}$

Table D.2: Steady state concentration (in µM) and steady state distribution of folate (in percentage of total folate per compartment) of the cytoplasm and nucleus.

	cytoplasm		**nucleus**	
	uM	**%**	**uM**	**%**
THF	0.052	0.31	0.045	0.57
10fTHF	6.765	40.01	3.544	44.15
CHF	1.346	7.96	0.688	8.57
CH2F	0.381	2.25	0.109	1.35
DHF	0.003	0.02	0.003	0.04
5mTHF free	4.324	25.57	0.549	6.84
5mTHF bound	3.096	18.31	0.254	3.16
5mTHF total	7.420	43.88	0.803	10.00
5fTHF free	0.388	2.29	1.475	18.37
5fTHF bound	0.555	3.28	1.361	16.96
5fTHF total	0.943	5.58	2.836	35.33
SHMT	0.286			
HCY	3.255			
MET	38.808			
SAM	72.899			
SAH	35.559			

Table D.3: Steady state fluxes for all model reactions (in µM/h).

	cytoplasm	**nucleus**
FTS	14069.6	10980.7
MTCH (10fTHF → CHF)	713991.6	387726.9
MTCH (CHF → 10fTHF)	711138.2	376746.2
MTD (CHF → CH2F)	85758.9	27868.4
MTD (CH2F → CHF)	82905.5	16887.7
DHFR	72.5	1118.3
TYMS	72.5	1118.3
SHMT (CH2F → THF)	3223.8	12630.2
SHMT (THF → CH2F)	465.5	2767.7
SHMT (CHF → 5fTHF)	1.8	0.6
MTHFS	1.8	0.6
MTHFR	22.6	
MTR	22.6	
PGT	5247.8	
AICARFT	5968.4	
BHMT	151.7	
MAT_I	112.8	
MAT_III	61.6	
GNMT	158.9	
DNMT	119.4	
SAHH_SAH	270.5	
SAHH_HCY	96.2	
5mTHF:SHMT (unbinding)	6129.4	
5mTHF:SHMT (binding)	6129.4	
5fTHF:SHMT (unbinding)	80.0	
5fTHF:SHMT (binding)	80.0	

Table D.4: The impact of variation in model enzymes activity on network output variables. Enzyme levels were modified according to a scaling factor (0.5x, 1x, 2x, and 4x) and for each enzyme the resulting model predictions for the flux through the enzymes AIACRFT, PGT, MTR, TYMS as well as the total number of produced Ts is provided. The variation across scenarios was assessed using the coefficient of variation (CV).

	v_{AICARFT}	v_{PGT}	v_{MTR}	#T total	v_{TYMS}	
					cytoplasm	nucleus
AICARFT						
0.25x	1491.5	5249.3	22.6	$1.51{\cdot}10^9$	72.5	1118.2
0.5x	2983.5	5248.8	22.6	$1.51{\cdot}10^9$	72.5	1118.2
1x	5968.4	5247.8	22.6	$1.51{\cdot}10^9$	72.5	1118.3
2x	11943.1	5245.2	22.6	$1.51{\cdot}10^9$	72.5	1118.2
4x	23901.3	5235.8	22.6	$1.51{\cdot}10^9$	72.1	1118.2
CV (%)	98.43	0.11	0.09	0.05	0.21	0.00
BHMT						
0.25x	5913.7	5264.1	22.8	$1.52{\cdot}10^9$	74.2	1118.2
0.5x	5929.6	5259.5	22.8	$1.52{\cdot}10^9$	73.7	1118.2
1x	5968.4	5247.8	22.6	$1.51{\cdot}10^9$	72.5	1118.3
2x	6050.6	5218.9	22.2	$1.50{\cdot}10^9$	69.6	1118.2
4x	6186.3	5152.1	21.3	$1.47{\cdot}10^9$	63.7	1118.3
CV (%)	1.86	0.88	2.82	1.30	6.10	0.00
DHFR						
0.25x	0.0	0.0	0.0	$0.77{\cdot}10^9$	0.0	726.5
0.5x	5966.7	5246.9	22.6	$1.51{\cdot}10^9$	72.4	1118.2
1x	5968.4	5247.8	22.6	$1.51{\cdot}10^9$	72.5	1118.3
2x	5968.5	5247.8	22.6	$1.51{\cdot}10^9$	72.5	1118.3
4x	5968.5	5247.8	22.6	$1.51{\cdot}10^9$	72.5	1118.2
CV (%)	55.90	55.90	55.90	24.36	55.90	16.85
DNMT						
0.25x	5997.9	5238.1	22.5	$1.51{\cdot}10^9$	71.5	1118.2
0.5x	5982.3	5243.3	22.5	$1.51{\cdot}10^9$	72.0	1118.3
1x	5968.4	5247.8	22.6	$1.51{\cdot}10^9$	72.5	1118.3
2x	5961.6	5249.9	22.6	$1.51{\cdot}10^9$	72.7	1118.3
4x	5959.1	5250.7	22.7	$1.51{\cdot}10^9$	72.8	1118.2
CV (%)	0.27	0.1	0.33	0.16	0.73	0.00

to continue on next page ...

Table D.4 continued

	$v_{AICARFT}$	v_{PGT}	v_{MTR}	#T total	v_{TYMS}	
					cytoplasm	nucleus
GNMT						
0.25x	5978.2	5244.7	22.6	$1.51{\cdot}10^9$	72.2	1118.3
0.5x	5973.4	5246.2	22.6	$1.51{\cdot}10^9$	72.3	1118.3
1x	5968.4	5247.8	22.6	$1.51{\cdot}10^9$	72.5	1118.3
2x	5964.1	5249.1	22.6	$1.51{\cdot}10^9$	72.6	1118.3
4x	5961.0	5250.1	22.6	$1.51{\cdot}10^9$	72.7	1118.3
CV (%)	0.12	0.04	0.14	0.07	0.31	0.00
MAT-I						
0.25x	6003.5	5236.2	22.4	$1.51{\cdot}10^9$	71.3	1118.3
0.5x	5980.2	5244.0	22.6	$1.51{\cdot}10^9$	72.1	1118.3
1x	5968.4	5247.8	22.6	$1.51{\cdot}10^9$	72.5	1118.3
2x	5963.4	5249.4	22.6	$1.51{\cdot}10^9$	72.6	1118.3
4x	5961.1	5250.1	22.6	$1.51{\cdot}10^9$	72.7	1118.3
CV (%)	0.29	0.11	0.35	0.17	0.80	0.00
MAT-III						
0.25x	5977.7	5244.8	22.6	$1.51{\cdot}10^9$	72.2	1118.3
0.5x	5973.0	5246.3	22.6	$1.51{\cdot}10^9$	72.3	1118.2
1x	5968.4	5247.8	22.6	$1.51{\cdot}10^9$	72.5	1118.3
2x	5965.0	5248.9	22.6	$1.51{\cdot}10^9$	72.6	1118.3
4x	5962.8	5249.6	22.6	$1.51{\cdot}10^9$	72.7	1118.3
CV (%)	0.10	0.04	0.12	0.06	0.27	0.00
MTHFD1						
0.25x	5831.6	4150.1	13.3	$1.31{\cdot}10^9$	28.2	1118.2
0.5x	5928.8	5255.6	22.5	$1.51{\cdot}10^9$	71.4	1118.2
1x	5968.4	5247.8	22.6	$1.51{\cdot}10^9$	72.5	1118.3
2x	5994.8	5239.4	22.7	$1.52{\cdot}10^9$	73.0	1118.2
4x	6008.8	5234.5	22.7	$1.52{\cdot}10^9$	73.3	1118.3
CV (%)	1.2	9.74	19.99	6.11	31.2	0
MTHFR						
0.25x	4749.5	5432.2	6.3	$1.62{\cdot}10^9$	96.1	1118.2
0.5x	4860.4	5422.3	12.6	$1.61{\cdot}10^9$	94.6	1118.2
1x	5968.4	5247.8	22.6	$1.51{\cdot}10^9$	72.5	1118.3

to continue on next page ...

Table D.4 continued

	$v_{AICARFT}$	v_{PGT}	v_{MTR}	#T total	v_{TYMS}	
					cytoplasm	**nucleus**
2x	6113.6	4151.5	24.0	$1.29 \cdot 10^9$	24.2	1118.2
4x	5196.7	2886.8	24.2	$1.23 \cdot 10^9$	10.0	1118.2
CV (%)	11.71	23.96	44.96	12.49	67.43	0.00
MTHFS						
0.25x	4801.0	5222.5	22.3	$1.26 \cdot 10^9$	70.3	886.8
0.5x	5458.6	5238.4	22.5	$1.40 \cdot 10^9$	71.7	1012.8
1x	5968.4	5247.8	22.6	$1.51 \cdot 10^9$	72.5	1118.3
2x	6309.5	5253.0	22.7	$1.60 \cdot 10^9$	72.9	1200.5
4x	6512.7	5255.8	22.7	$1.67 \cdot 10^9$	73.2	1261.0
CV (%)	11.89	0.26	0.73	11.00	1.64	13.64
MTR						
0.25x	5200.0	2890.6	6.1	$1.23 \cdot 10^9$	10.1	1118.2
0.5x	6115.6	4155.1	12.0	$1.29 \cdot 10^9$	24.2	1118.2
1x	5968.4	5247.8	22.6	$1.51 \cdot 10^9$	72.5	1118.3
2x	4860.9	5422.3	25.2	$1.61 \cdot 10^9$	94.5	1118.2
4x	4749.7	5432.2	25.4	$1.62 \cdot 10^9$	96.1	1118.3
CV (%)	11.71	23.92	47.85	12.47	67.35	0.00
PGT						
0.25x	5966.4	1312.3	22.6	$1.51 \cdot 10^9$	72.5	1118.2
0.5x	5967.1	2624.4	22.6	$1.51 \cdot 10^9$	72.5	1118.2
1x	5968.4	5247.8	22.6	$1.51 \cdot 10^9$	72.5	1118.3
2x	5971.2	10491.1	22.6	$1.51 \cdot 10^9$	72.5	1118.3
4x	5975.4	20953.5	22.6	$1.51 \cdot 10^9$	72.3	1118.2
CV (%)	0.06	98.28	0.06	0.03	0.13	0.00
SAHH						
0.25x	6169.5	5162.4	21.4	$1.48 \cdot 10^9$	64.5	1118.2
0.5x	6031.3	5226.3	22.3	$1.50 \cdot 10^9$	70.3	1118.3
1x	5968.4	5247.8	22.6	$1.51 \cdot 10^9$	72.5	1118.3
2x	5943.6	5255.4	22.7	$1.52 \cdot 10^9$	73.3	1118.3
4x	5934.6	5258.1	22.8	$1.52 \cdot 10^9$	73.5	1118.2
CV (%)	1.62	0.76	2.43	1.12	5.27	0
SHMT						

to continue on next page ...

Table D.4 continued

	$v_{AICARFT}$	v_{PGT}	v_{MTR}	#T total	v_{TYMS}	
					cytoplasm	**nucleus**
0.25x	6658.1	5287.2	23.6	$1.55 \cdot 10^9$	79.7	1118.3
0.5x	6439.4	5277.7	23.3	$1.54 \cdot 10^9$	77.8	1118.2
1x	5968.4	5247.8	22.6	$1.51 \cdot 10^9$	72.5	1118.3
2x	5283.1	5098.1	19.7	$1.43 \cdot 10^9$	54.2	1118.2
4x	5232.0	4170.1	8.8	$1.26 \cdot 10^9$	16.0	1118.2
CV (%)	11.01	9.55	31.68	8.27	44.31	0.00
TYMS						
0.25x	5884.2	5239.1	22.4	$0.47 \cdot 10^9$	17.8	365.6
0.5x	5912.0	5242.1	22.5	$0.88 \cdot 10^9$	35.8	674.1
1x	5968.4	5247.8	22.6	$1.51 \cdot 10^9$	72.5	1118.3
2x	6082.4	5257.2	22.8	$2.26 \cdot 10^9$	147.8	1497.0
4x	0.0	0.0	0.0	$1.61 \cdot 10^9$	0.0	1516.2
CV (%)	55.92	55.90	55.91	51.48	106.81	49.07

Table D.5: Steady state concentrations of model variables (in µM) for different levels of NADPH+NADP (ranging from 0.5x to 2x of the standard concentration of 76 µM).

cytoplasm									
NADPH+ NADP	**THF**	**10fTHF**	**CHF**	**CH2F**	**DHF**	**5mTHF**		**5fTHF**	
						free	**bound**	**free**	**bound**
0.5x	0.058	9.334	1.852	0.444	0.004	0.551	0.648	1.200	2.820
1x	0.054	7.567	1.504	0.406	0.003	3.209	2.669	0.562	0.935
1.5x	0.046	4.643	0.925	0.264	0.002	7.164	3.605	0.130	0.131
2x	0.043	3.398	0.677	0.198	0.002	8.767	3.718	0.059	0.050
	SHMT	**HCY**	**MET**	**SAM**	**SAH**				
0.5x	0.470	3.725	42.631	33.416	70.748				
1x	0.333	3.331	39.863	67.433	39.893				
1.5x	0.201	3.125	36.945	80.901	29.549				
2x	0.170	3.075	36.224	83.612	27.610				
nucleus									
NADPH+ NADP	**THF**	**10fTHF**	**CHF**	**CH2F**	**DHF**	**5mTHF**		**5fTHF**	
						free	**bound**	**free**	**bound**
0.5x	0.033	3.558	0.696	0.079	0.003	0.550	0.252	1.489	1.476
1x	0.042	3.548	0.690	0.100	0.003	0.550	0.253	1.479	1.472
1.5x	0.047	3.543	0.687	0.111	0.003	0.549	0.253	1.473	1.470
2x	0.049	3.540	0.685	0.118	0.003	0.549	0.253	1.470	1.470

Table D.6: Steady state concentrations of model variables (in µM) for different levels of glycine+serine (ranging from 0.01x to 10x of the standard concentration of 2318 µM).

cytoplasm									
glycine+ serine	**THF**	**10fTHF**	**CHF**	**CH2F**	**DHF**	**5mTHF**		**5fTHF**	
						free	**bound**	**free**	**bound**
0.01x	0.038	6.523	1.304	0.384	0.003	4.647	3.187	0.347	0.476
0.1x	0.040	6.584	1.315	0.385	0.003	4.565	3.164	0.357	0.496
0.5x	0.046	6.688	1.333	0.383	0.003	4.427	3.126	0.375	0.529
1x	0.052	6.765	1.346	0.381	0.003	4.324	3.095	0.388	0.555
1.5x	0.056	6.817	1.354	0.380	0.003	4.254	3.074	0.397	0.574
2x	0.059	6.855	1.360	0.379	0.003	4.204	3.059	0.404	0.587
5x	0.069	6.959	1.377	0.375	0.003	4.064	3.014	0.423	0.627
10x	0.074	7.012	1.386	0.374	0.003	3.991	2.989	0.433	0.648
	SHMT	**HCY**	**MET**	**SAM**	**SAH**				
0.01x	0.274	2.851	32.379	95.210	20.081				
0.1x	0.277	3.131	36.709	81.414	29.265				
0.5x	0.282	3.234	38.454	74.467	34.365				
1x	0.286	3.255	38.809	72.893	35.563				
1.5x	0.289	3.264	38.962	72.190	36.105				
2x	0.291	3.270	39.052	71.763	36.435				
5x	0.297	3.283	39.262	70.749	37.226				
10x	0.300	3.289	39.357	70.281	37.593				
nucleus									
glycine+ serine	**THF**	**10fTHF**	**CHF**	**CH2F**	**DHF**	**5mTHF**		**5fTHF**	
						free	**bound**	**free**	**bound**
0.01x	0.007	3.505	0.697	0.170	0.005	0.549	0.253	1.478	1.363
0.1x	0.014	3.512	0.696	0.160	0.004	0.549	0.253	1.477	1.363
0.5x	0.032	3.532	0.691	0.129	0.003	0.549	0.253	1.476	1.363
1x	0.046	3.544	0.688	0.109	0.003	0.549	0.253	1.475	1.363
1.5x	0.053	3.551	0.686	0.097	0.003	0.549	0.253	1.474	1.361
2x	0.059	3.555	0.685	0.090	0.002	0.549	0.253	1.473	1.361
5x	0.072	3.566	0.682	0.072	0.002	0.549	0.253	1.472	1.361
10x	0.078	3.570	0.680	0.063	0.002	0.549	0.253	1.471	1.361

Appendix E

Additional material for Chapter 5

E.1 Comparison with related work

In recent years, numerous computational drug repurposing methods have been developed. In particular, network-based *in silico* repositioning strategies have received great attention given the suitability of network data structures to represent biological data in an efficient way. The most commonly adopted network-based computational strategies have been described in recent review articles[153;283].

The available methods can be categorized according to the type of input data they require as well as the type of network used and the computational strategy applied to identify repurposing drug candidates.

With regard to the input data, common strategies are based on gene/protein expression profiles of drugs and diseases, disease symptoms, chemical structures of drugs, reported drug side effects, known drug targets, known drug indications, and/or clinical data related to the drug of interest (from clinical trials and electronic health records)[206].

To describe the interactions among the players of interest, different types of networks can be considered. Protein-protein interaction networks, gene regulatory networks and metabolic networks are commonly used to represent the interactions among the molecular players that constitute a biological system (protein-protein, transcription factor-gene, and metabolite-metabolite interactions). For example, a recent study by Cheng and collaborators identified new drug repurposing candidates for cardiovascular disease mapping both disease genes and drug targets on a protein-protein interaction network[44]. The transcriptomic profiles of drugs and diseases can also be analyzed using network-based approaches. For example, the drug-gene signatures from Connectivity Map have been integrated with genes frequently mutated in cancer to identify new drugable targets and anticancer indications for existing drugs[46]. In addition, also information present in genome-scale metabolic models have been successfully integrated in a network-based setting to identify drug targets for the treatment of metabolic disorders[144;145]. Heterogeneous networks constituted by different types of nodes, such as drugs, targets or diseases, can be generated. The analysis of these kinds of networks is commonly used for the prediction of new drug-target interactions[153]. In this case networks with at least two distinct types of nodes are constructed on the basis of known associations and different algorithms are used to predict new interactions.

The prediction of new drug applications or drug-disease associations can be based solely on the

network characteristics or exploit other sources of information such as drug structure similarity, target similarity, disease similarity [283].

Inspired by these network-based repurposing studies, we developed a new computational strategy that tries to combine different input data, background networks and similarity measures in an innovative framework. The background network was built by integrating protein-protein and transcription factor-gene interactions; the latter was demonstrated as particularly suitable to cluster GWAS derived disease genes. A further strength of our study is the tissue-specificity of the generated networks, as it was extensively demonstrated that gene function is frequently related to the tissue and disease considered [32;55;98;132;138;279].

To obtain a detailed description of the disease pathways, we integrated genetic information derived from GWAS with literature-derived knowledge. GWAS data has already been exploited for computational repurposing [223]. However, to address the problem of distinguishing causal genes from genes located in the same genomic region but not involved in the disease predisposition, we performed a preliminary filtering step based on pathway enrichment analysis, similarly to previous GWAS prioritization studies [60;141;198;210;228]. Moreover, the GWAS results were complemented with literature that captures also non-GWAS findings, such as the results of functional studies.

For the drugs, we combined gene expression profiles with information about the drug target. The network-based setting allowed us integrating this two distinct information using a novel strategy to identify drug modules. Indeed, the module detection is based on the idea that the gene expression perturbation observed in the drug signature profiles is a cascade starting with the drug-target binding. Moreover, the regulatory-network component of our background network supports this reasoning because it describes the transcriptional regulation exerted by transcription factors.

Another aspect that separates our strategy from previously published methods is the combined use of topological network information and semantic similarities to identify repurposing candidates. The two approaches have been used separately by others to identify new disease-drug interactions [104;251]. While the idea to combine network characteristics and functional similarities has already been applied for the detection of network modules [264], the integration of semantic similarities and topological properties between network modules has not yet been applied to identify novel disease-drug relations.

E.2 Contribution of the different data sources to the final predictions

In the pipeline we developed, several data sources are integrated by means of network analysis. Tissue-specific networks, generated by merging protein-protein interaction networks (PPI) and transcriptional regulatory networks (RegNet), provided the scaffold of the analysis. GWAS results and text mining results were combined to detect disease relevant network modules (MetSyn modules). Gene set enrichment analysis of GWAS genes identified significant pathways that were the basis for filtering both the GWAS and text mining genes. Specifically, only genes that are part of at least one significant pathway were selected for subsequent analyses. The identification of drug repurposing candidates was based on the definition of a proximity score that connects drug modules and MetSyn modules combining network-based distance and functional similarity. In the following section, the contribution of the individual system components to the final prediction will be described (GWAS vs. text mining, PPI vs. regulatory network, network-based distance vs. GO functional similarity).

GWAS vs. text mining

With our workflow that combines genes derived from GWAS and text mining results, two MetSyn modules were identified in the adipose network, and three MetSyn modules were detected for liver and muscle. If we consider only the genes derived from GWAS results (286 genes, Appendix E, Table E.11), one out of two MetSyn module in the adipose network, one out of three MetSyn module in the muscle network, and three out of three MetSyn modules in the liver network can be detected. On the other hand, taking into account only genes derived from text mining results (546 genes, Appendix E, Table E.11), leads to the detection of two out of two MetSyn module in the adipose network, three out of three MetSyn modules in the muscle network, and two out of three MetSyn modules in the liver network. However, the final selection of text mining derived genes is guided by the GWAS results through pathway analysis and thus the comparison of the individual contribution of text mining and GWAS is not straightforward and should be interpreted in the context of the entire methodological framework. Overall, we can observe that the addition of text mining genes allows a more detailed disease characterisation. For example, the inflammation-related MetSyn module in the adipose network would not have been identified if the analysis had been limited to the genes derived solely from GWAS. Moreover, the source (GWAS and text mining) of the MetSyn genes included in the modules was evaluated and is summarised in Table E.1. Overall, the text mining genes play a major role for the disease enrichment analysis. In all except two cases (Liver Modules 1 and 3), text mining genes accounted for the significance of the modules.

PPI vs. regulatory network

To evaluate the contribution of the protein-protein and regulatory network interactions, we first evaluated the composition of the complete networks (Table E.2) and afterwards the composition of the MetSyn modules for both edge and node source (Tables E.3 and E.4). Overall, we can conclude that the edges and nodes derived from the PPI networks contribute more to the overall network structure. On the level of MetSyn modules, only one module (Liver Module 1) is solely constructed

Table E.1: Contribution of GWAS vs. text mining results to the MetSyn modules.

Network	Module	Module size	Number of disease genes			
			Total	GWAS	Text mining	GWAS & text mining
Adipose	1	122	23	0	19	4
	2	51	19	1	10	8
Liver	1	17	8	3	1	4
	2	49	21	0	10	11
	3	20	6	3	1	2
Muscle	1	299	39	2	31	6
	2	50	17	1	9	7
	3	26	7	0	5	2

based on edges and nodes from the regulatory network. This relatively small module is related to the modification of chemicals (Reactome pathway: Phase II - Conjugation of compounds).

Table E.2: Contribution of regulatory and protein-protein interactions (RegNet vs. PPI) to the final networks.

Network	Number of edges			Number of nodes			
	Total	RegNet	PPI	Total	RegNet	PPI	RegNet and PPI
Adipose	9152	255	8900	886	124	661	101
Liver	15846	1520	14326	1544	395	983	166
Muscle	10555	603	9952	1106	232	741	133

Table E.3: Contribution of regulatory and protein-protein interactions (network edges) to the edges of the MetSyn modules.

Network	Module	Number of edges		
		Total	RegNet	PPI
Adipose	1	872	0	872
	2	324	0	324
Liver	1	29	29	0
	2	220	8	212
	3	238	0	238
Muscle	1	3512	2	3510
	2	248	0	248
	3	202	0	202

network-based distance vs. GO functional similarity

Finally, we evaluated the contribution of the closest network distance and the GO similarity to the final significant results by testing how many of the significant results remain significant if only the topological aspect or the semantic similarity was considered. With the integrated proximity score,

Table E.4: Contribution of regulatory and protein-protein interactions (network nodes) to the nodes of the MetSyn modules.

Network	Module	Number of nodes			
		Total	**RegNet**	**PPI**	**RegNet & PPI**
Adipose	1	122	0	119	3
	2	51	0	44	7
Liver	1	17	17	0	0
	2	49	3	37	9
	3	20	0	19	1
Muscle	1	299	0	265	34
	2	50	0	38	12
	3	26	0	25	1

we identified 28, 31 and 50 significant drugs in the adipose, liver and muscle network, respectively. Overall, for all MetSyn modules the number of significant results increases using the integrated score (Figure E.1a). The evaluation of the relative contribution of the two aspects showed that the GO similarity score alone identified more significant drugs than the topological score alone (Figure E.1b). The only module for which the topological score has a greater impact is Module 2 in the muscle network (Figure E.1).

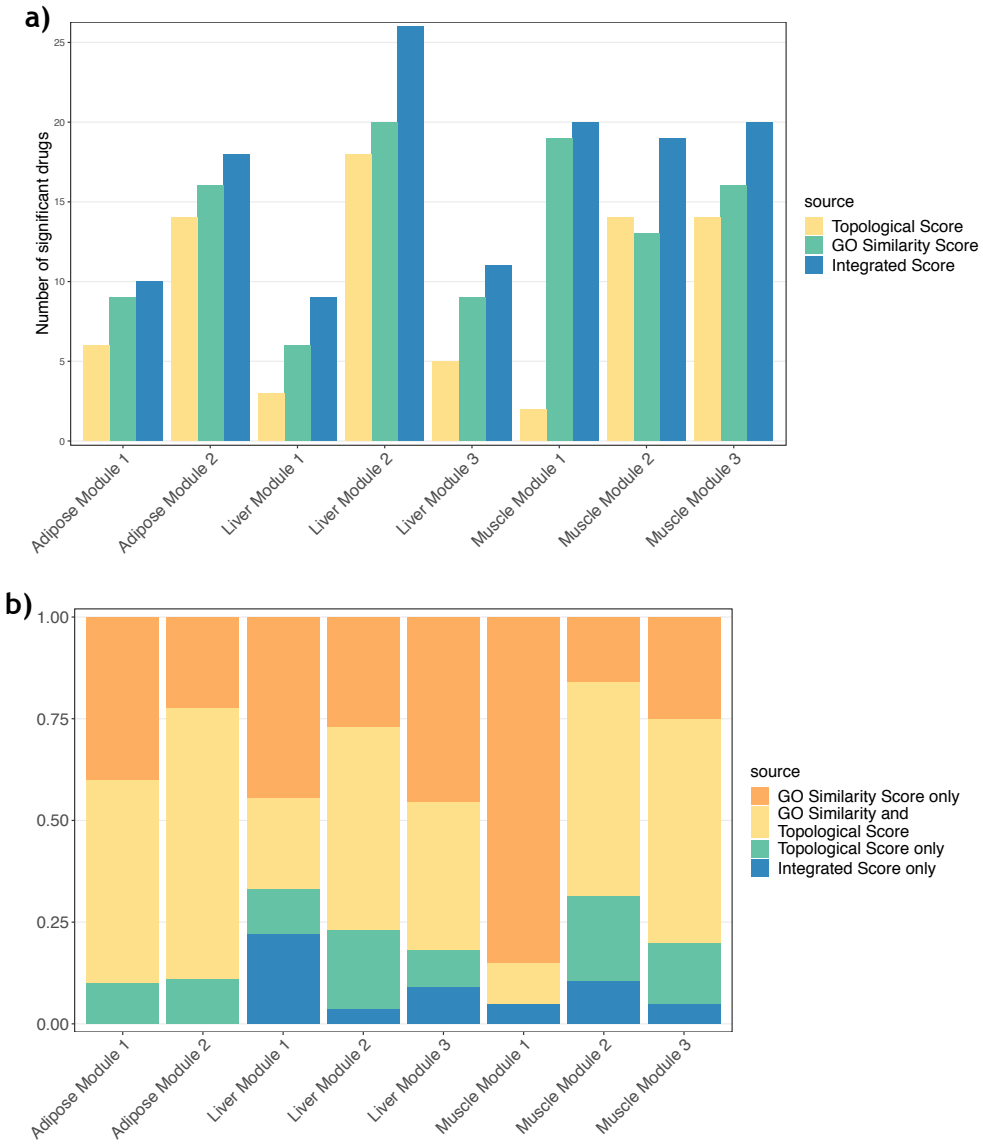

Figure E.1: Contribution of the topological score and the functional similarity to the final significant results. a) Number of significant drugs detected considering only the topological score, only the similarity score, or the integrated score. b) Relative contribution of the two components of the integrated score to the final predictions. The blue bar indicates drugs only significant when the integrated score was applied.

E.3 Evaluation of score effectiveness

To test the effectiveness of our approach in identifying drugs affecting MetSyn-pathways, we looked at drugs already approved for MetSyn-related traits. We identified them by parsing the indications provided by DrugBank [280] (version 5.1.1) and filtering for the following search-terms:

obesity, hyperglycemia, glycemicontrol, dyslipid(a)emia, high cholesterol, high triglyceride, hyperlipid(a)emia, hypertension, high blood pressure, NIDDM, cholesterol-lowering, antilipemic, statins, lipid-lowering, hypercholesterolemia, and diabetes.

For the diabetes search-term, a further refinement was carried out limiting the results to those referring to type 2 diabetes. Moreover, for the adipose tissue, the resulting list was trimmed to those drugs affecting adipose tissue and those for which we were able to generate a drug module in the adipose network (dependent on availability of drug target and signature) (Table E.5).

Based on the information provided in DrugBank [280] and ChEMBL [57], six of the resulting ten drugs have an obvious impact on adiposity (bezafibrate, clofibrate, fenofibrate, gemfibrozil, mifepristone, and pioglitazone), while four drugs (atorvastatin, glipizide, lovastatin and repaglinide) did not have a direct connection to adiposity. However, for Atrovastatin and Lovastatin recent publications suggest their impact also on adiposity [177;197;286]. The two sulfonylureas drugs Glipizide and Repaglinide were also kept because of a described effect on adipose tissue via the inhibition of lipolysis [233;256]. For this reason we executed two analyses, one with all the 6 top MetSyn drugs (Analysis A: bezafibrate, clofibrate, fenofibrate, gemfibrozil, mifepristone, and pioglitazone), and an additional one including all ten MetSyn drugs (Analysis B: bezafibrate, clofibrate, fenofibrate, gemfibrozil, mifepristone, pioglitazone, atorvastatin, glipizide, lovastatin and repaglinide).

To evaluate the effectiveness of our scoring system, we tested how many MetSyn drugs our method could identify at different significance thresholds (Figure E.2).
With our selected significance threshold of 95% we are able to detect half of the MetSyn drugs in both Analysis (Analysis A: pioglitazone, mifepristone, and fenofibrate; Analysis B: pioglitazone, mifepristone, fenofibrate, glipizide and lovastatin). This still holds for the threshold of 99% for Analysis A, while for Analysis B the ratio of detected MetSyn drugs decreases because Glipizide and Lovastatin are no longer significant. Lowering the threshold to 85% we could identify all six top MetSyn drugs, while Atorvastatin and Repaglinide are only identified at a threshold level of 75% and 50%, respectively.

Table E.5: Drugs with indications for MetSyn-related traits affecting adipose tissue.

Drug Name	Indication in DrugBank	Search-term	MoA (ChEMBL)
Atorvastatin	May be used as primary prevention in individuals with multiple risk factors for coronary heart disease (CHD) and as secondary prevention in individuals with CHD to reduce the risk of myocardial infarction (MI), stroke, angina, and revascularization procedures. May be used to reduce the risk of cardiovascular events in patients with acute coronary syndrome (ACS). May be used in the treatment of primary hypercholesterolemia and mixed dyslipidemia, homozygous familial hypercholesterolemia, primary dysbetalipoproteinemia, and/or hypertriglyeridemia as an adjunct to dietary therapy to decrease serum total and low-density lipoprotein cholesterol (LDL-C), apolipoprotein B (apoB), and triglyceride concentrations, while increasing high-density lipoprotein cholesterol (HDL-C) levels.	hypercholesterolemia	HMG-CoA reductase inhibitor
Bezafibrate	For the treatment of primary hyperlipidaemia types IIa, IIb, III, IV and V (Fredrickson classification) corresponding to groups I, II and III of the European Atherosclerosis Society guidelines - when diet alone or improvements in lifestyle such as increased exercise or weight reduction do not lead to an adequate response. Also for the treatment of secondary hyperlipidaemias, e.g. severe hypertriglyceridemias, when sufficient improvement does not occur after correction of the underlying disorder (e.g. diabetes mellitus).	diabetes	Peroxisome proliferator-activated receptor agonist
Clofibrate	For Primary Dysbetalipoproteinemia (Type III hyperlipidemia) that does not respond adequately to diet. This helps control high cholesterol and high triglyceride levels.	hyperlipidemia, high cholesterol	Peroxisome proliferator-activated receptor alpha agonist
Fenofibrate	For use as adjunctive therapy to diet to reduce elevated LDL-C, Total-C,Triglycerides and Apo B, and to increase HDL-C in adult patients with primary hypercholesterolemia or mixed dyslipidemia (Fredrickson Types IIa and IIb)	hypercholesterolemia	Peroxisome proliferator-activated receptor alpha agonist

Table E.5 continued

Drug Name	Indication in DrugBank	Search-term	MoA (ChEMBL)
Gemfibrozil	For treatment of adult patients with very high elevations of serum triglyceride levels (types IV and V hyperlipidemia) who are at risk of developing pancreatitis (inflammation of the pancreas) and who do not respond adequately to a strict diet.	hyperlipidemia	Peroxisome proliferator-activated receptor alpha agonist
Glipizide	For use as an adjunct to diet for the control of hyperglycemia and its associated symptomatology in patients with non-insulin-dependent diabetes mellitus (NIDDM; type II), formerly known as maturity-onset diabetes, after an adequate trial of dietary therapy has proved unsatisfactory.	diabetes, hyperglycemia, NIDDM	Sulfonylurea receptor 1, Kir6.2 blocker
Lovastatin	For management as an adjunct to diet to reduce elevated total-C, LDL-C, apo B, and TG levels in patients with primary hypercholesterolemia and mixed dyslipidemia. For primary prevention of coronary heart disease and to slow progression of coronary atherosclerosis in patients with coronary heart disease.	hypercholesterolemia	HMG-CoA reductase inhibitor
Mifepristone	For the medical termination of intrauterine pregnancy through 49 days' pregnancy. Also indicated to control hyperglycemia secondary to hypercortisolism in adult patients with endogenous Cushing's syndrome who have type 2 diabetes mellitus or glucose intolerance and are not candidates for surgery or have had unsuccessful surgery.	Diabetes, hyperglycemia	Glucocorticoid receptor antagonist - Progesterone receptor antagonist
Pioglitazone	Treatment of Type II diabetes mellitus	diabetes	Peroxisome proliferator activated receptor gamma agonist
Repaglinide	As an adjunct to diet and exercise to improve glycemic control in adults with type 2 diabetes mellitus.	diabetes	Sulfonylurea receptor 1, Kir6.2 blocker

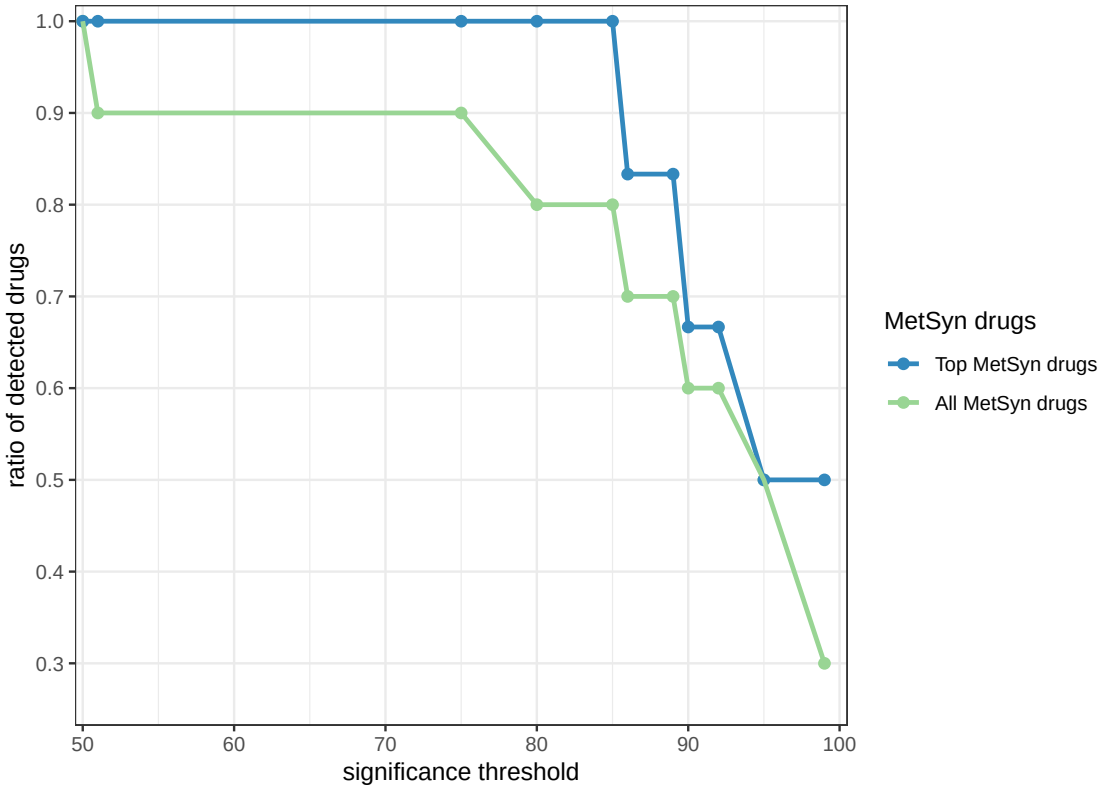

Figure E.2: Ratio of detected MetSyn drugs at different significance thresholds.

E.4 Supplementary Figures

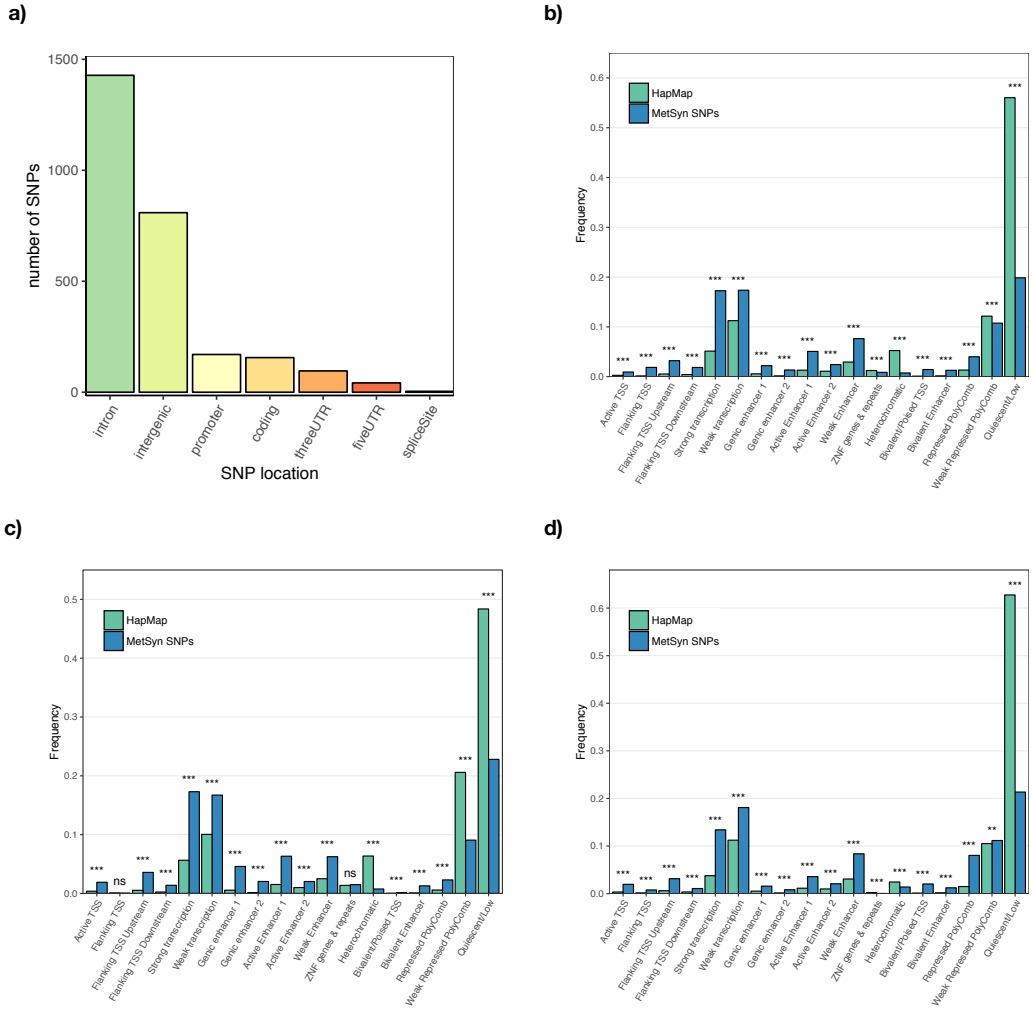

Figure E.3: Functional annotation of SNPs from GWAS catalog. a) Functional annotation of the SNPs extracted from GWAS catalog. SNPs were categorized in one or more of the indicated categories according to their genomic location. The location of the analyzed gene elements was retrieved from the University of California Santa Cruz (UCSC) Known Genes dataset [113]. b)-d) Annotation of SNPs according to their position relative to regulatory regions. The positions of MetSyn-related GWAS SNP and the 18-chromatin-state annotation of a) adipose tissue, b) liver tissue and c) skeletal muscle tissue obtained from the NIH Roadmap Epigenomics project [218] were overlapped to evaluate the regulatory potential of the identified SNPs. The full set SNP of HapMap SNPs was analyzed in the same way for comparison. Fisher exact test was used to test the over-/underrepresentation of GWAS SNPs in regulatory regions. Statistical significance is denoted as follows: ns: not significant ($p > 0.05$), * : $.01 < p \leq 0.05$, ** : $0.01 < p \leq 0.001$, *** : $p < 0.001$.

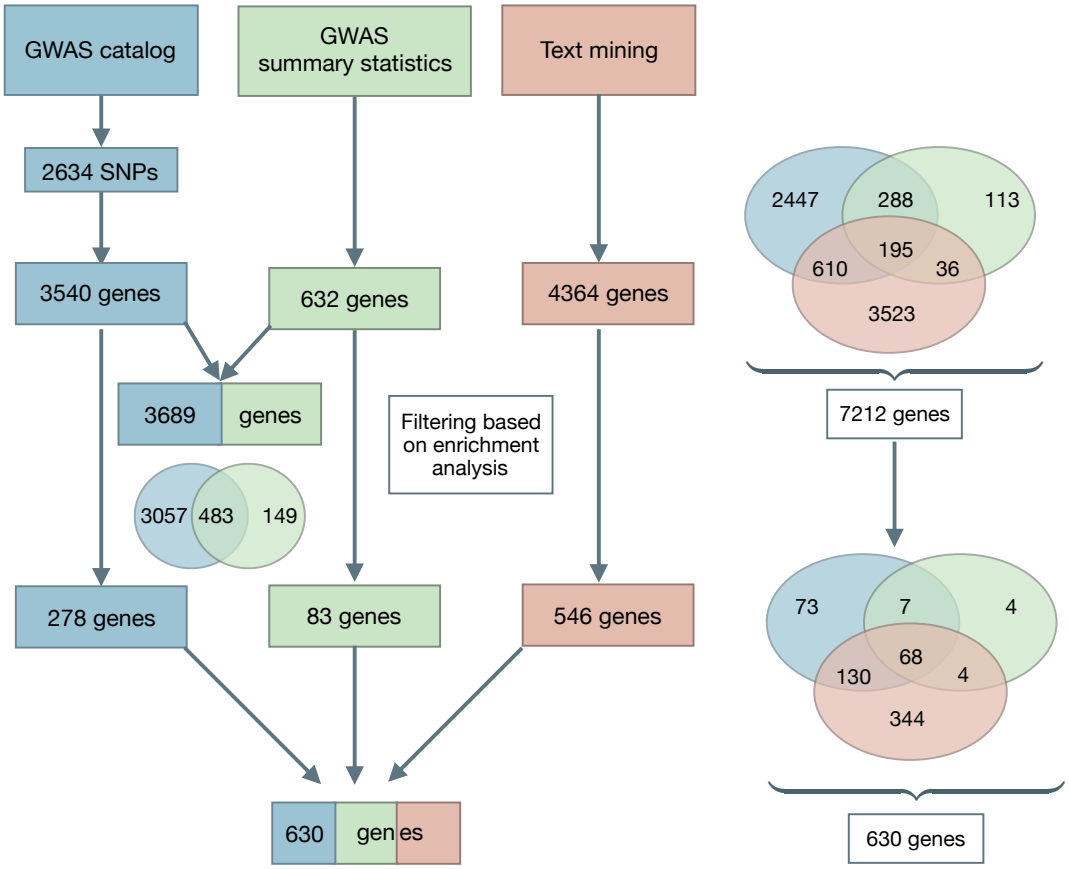

Figure E.4: Overview of MetSyn genes identification. For each of the three resources (GWAS catalog, GWAS summary statistics and text mining) the number of genes and their intersection is shown.

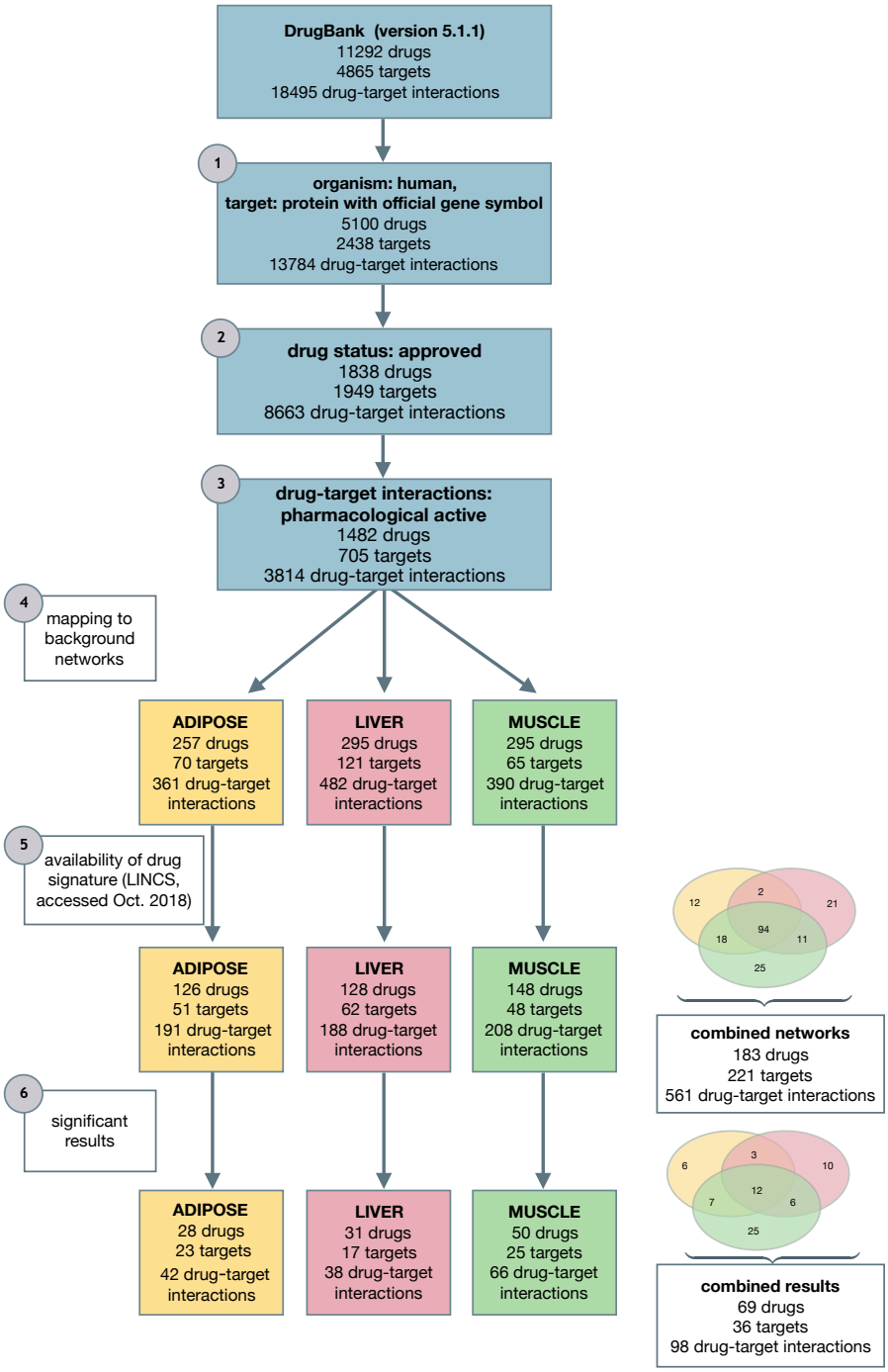

Figure E.5: Overview of drug selection. The drugs available in DrugBank version 5.1.1 were filtered according to the following steps: We kept 1) drugs having human as the organisms for which the drug is most effective and whose target(s) could be mapped to official gene symbol(s). 2) approved drugs. 3) targets with a pharmacological active status. 4) targets that could be mapped to at least one of our background networks. 5) drugs for which a drug signature could be retrieved from LINCS. After the proximity score calculation, we identified the drugs with a significant score (step 6). For step 5 and 6, venn diagrams indicate the overlap of selected drugs across the three networks.

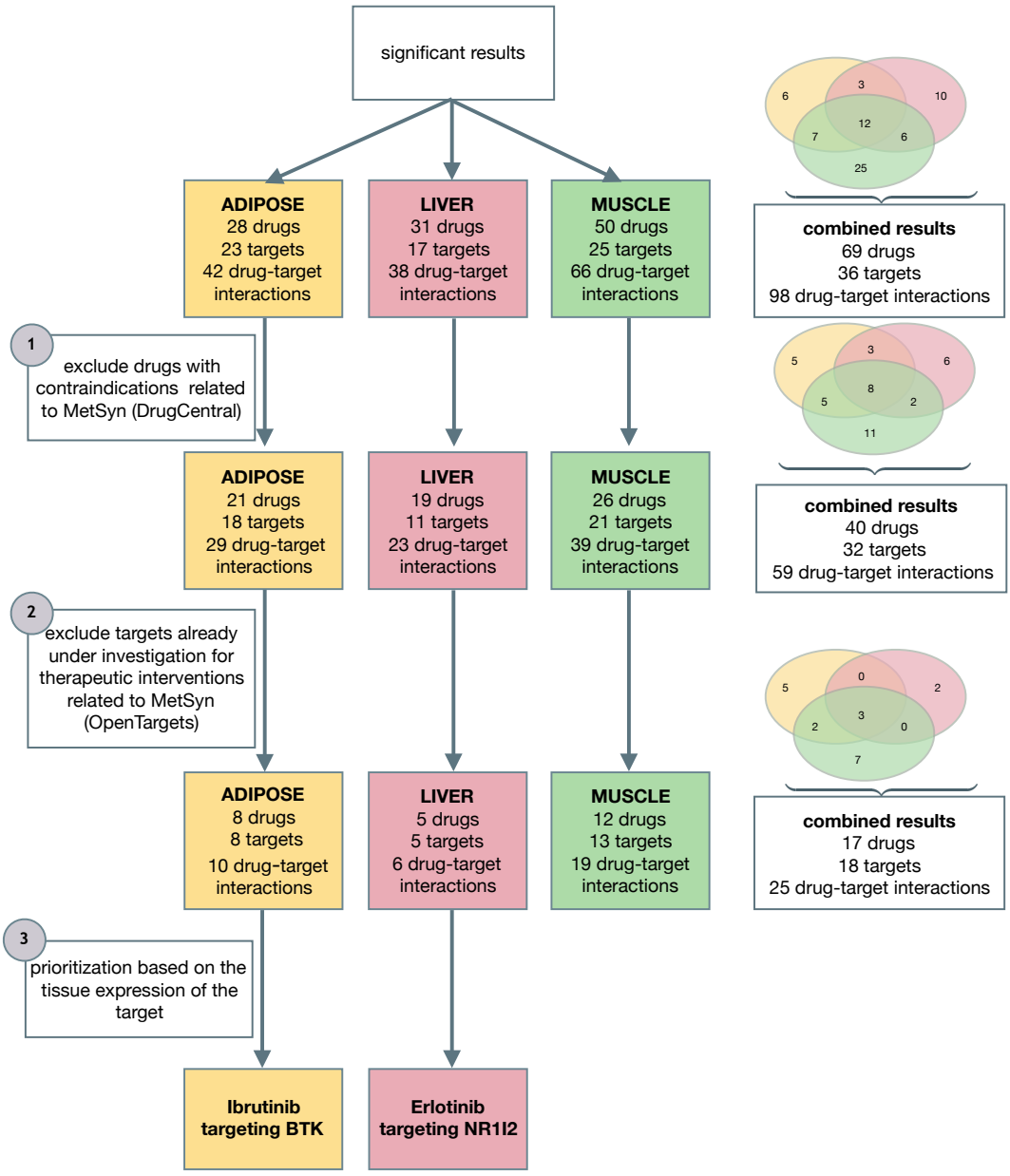

Figure E.6: Overview of the filtering and prioritization of the significant results. The drugs with a significant score were filtered according to the following steps: 1) drugs with a contraindication related to MetSyn were excluded, 2) drugs with a target that is already under investigation for therapeutic interventions related to MetSyn were excluded, 3) the remaining results were filtered according to the tissue specific expression of the targets. Only the targets of drug-targets interactions with a tissue-specific expression in agreement with MetSyn pathophysiology were kept.

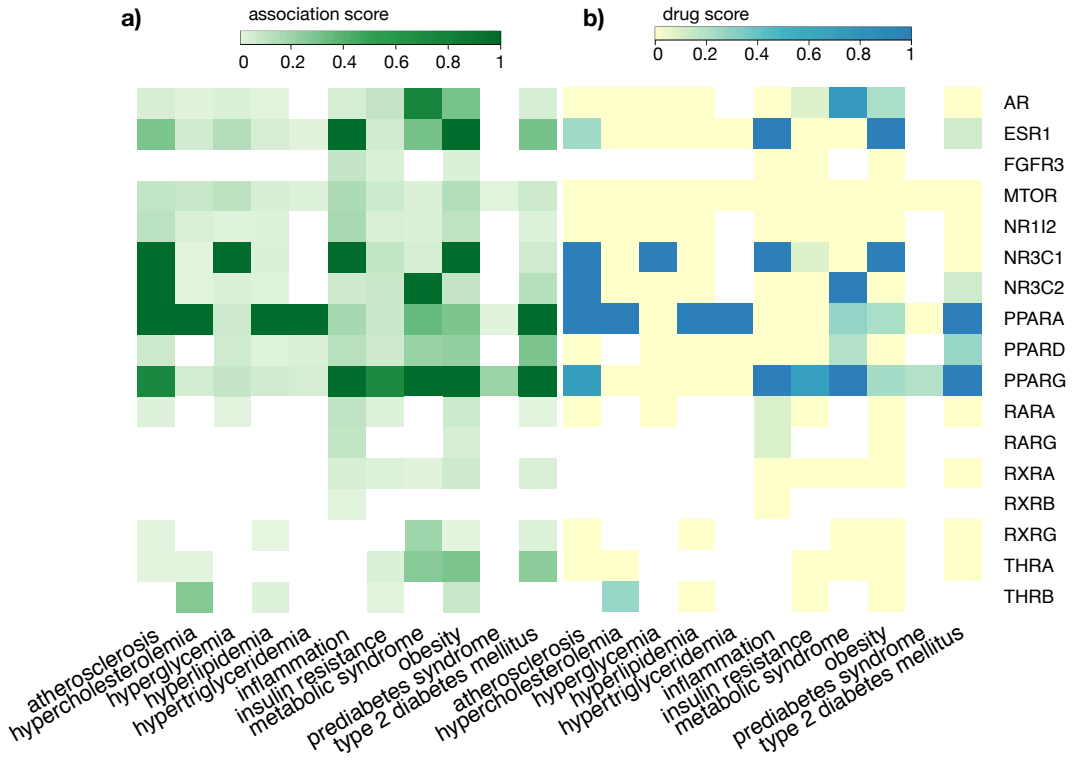

Figure E.7: Association between the identified drug targets in the liver network and traits selected to be relevant for MetSyn-related traits based on the association scores provided in by the OpenTargets platform [133].; a) Heatmap of total association score and b) heatmap of association score based on ChEMBL information about drugs approved for marketing by FDA or under evaluation in clinical trials.

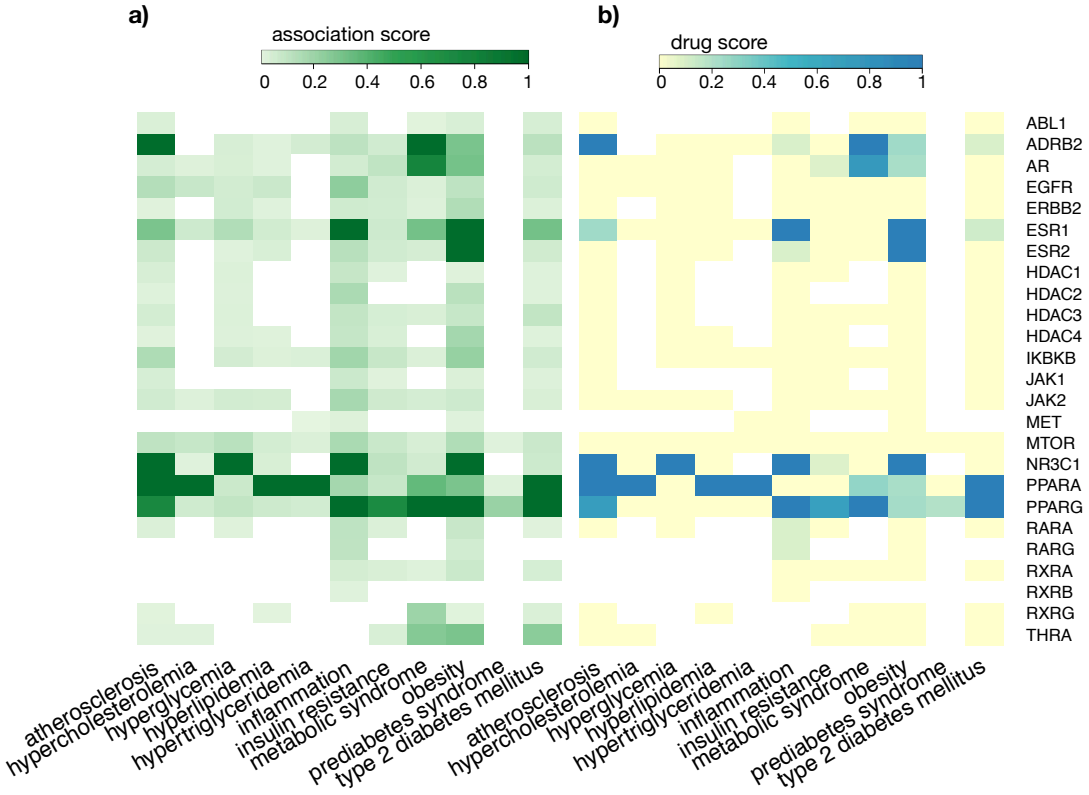

Figure E.8: Association between the identified drug targets in the muscle network and traits selected to be relevant for MetSyn-related traits based on the association scores provided in by the OpenTargets platform [133].; a) Heatmap of total association score and b) heatmap of association score based on ChEMBL information about drugs approved for marketing by FDA or under evaluation in clinical trials.

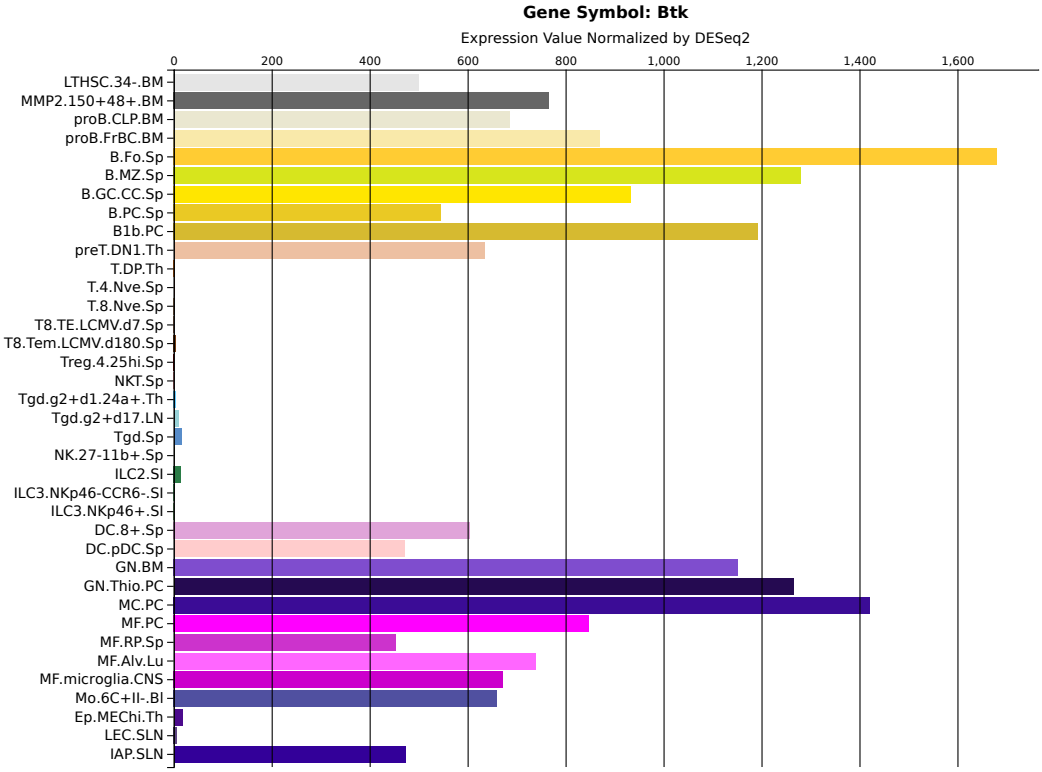

Figure E.9: *Btk* expression in immune related cell from ImmGen database. The bar chart shows *Btk* levels in derived from RNASeq experiments performed by ImmGen [109].

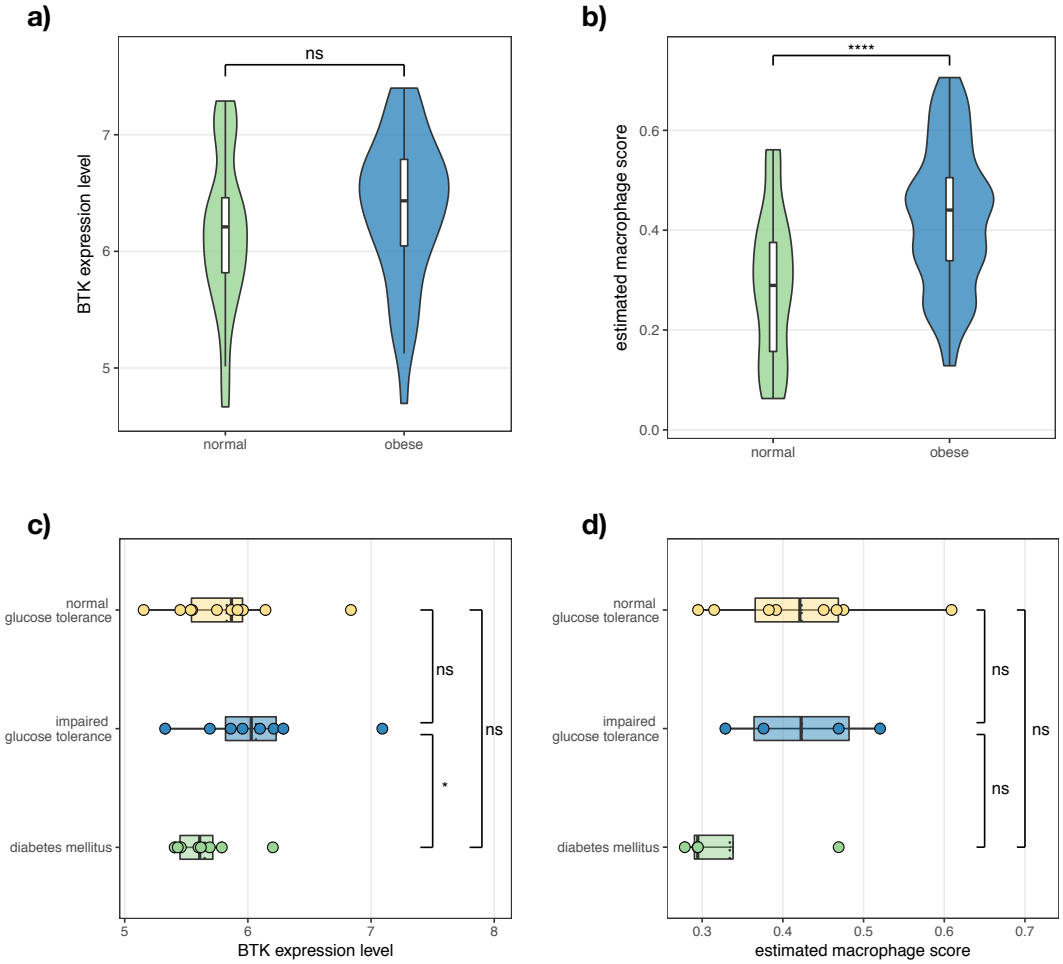

Figure E.10: *BTK* expression profiles and macrophages enrichment in human adipose tissue. a) *BTK* expression profiles obtained from the dataset E-MTAB-54 [64] in abdominal adipose tissue of obese and normal subjects. b) Estimation of macrophages enrichment obtained from the dataset E-MTAB-54 [64] in abdominal adipose tissue of obese and normal subjects using CIBERSORT [181]. c) *BTK* expression profile obtained from the dataset GSE27951 [123] comparing overweight subjects with normal glucose tolerance, impaired glucose tolerance and diabetes mellitus. d) Estimation of macrophages enrichment obtained from the dataset GSE27951 [123] comparing overweight subjects with normal glucose tolerance, impaired glucose tolerance and diabetes mellitus using CIBERSORT [181]. In a)-d), statistical significance is denoted as follows: ns: not significant ($p > 0.05$), * : $0.01 < p \leq 0.05$, ** : $0.01 \leq p < 0.001$, ***: $p \leq 0.001$.

E.5 Supplementary Tables

Table E.6: The 15 selected studies used as resource for GWAS summary statistics, obtained from the results of Marbach et al. applying the PASCAL tool[141] (downloaded from `http://regulatorycircuits.org/download.html`)

Trait	File	Pmid
Coronary artery disease	23_coronary_artery_disease.txt	21378990
Blood pressure (systolic)	22_blood_pressure_systolic.txt	21909115
HDL cholesterol	18_hdl_cholesterol.txt	20686565
LDL cholesterol	19_ldl_cholesterol.txt	20686565
Total cholesterol	20_total_cholesterol.txt	20686565
Triglycerides	21_triglycerides.txt	20686565
Type 2 diabetes	26_type_2_diabetes.txt	22885922
Insulin secretion (CIR)	29_insulin_secretion.txt	24699409
Glucose tolerance	27_2hr_glucose.txt	22885924
Fasting glucose	24_fasting_glucose.txt	22885924
Fasting insulin (BMI-adj.)	32_fasting_insulin.txt	22885924
Glycated hemoglobin	25_glycated_hemoglobin.txt	20858683
Beta-cell function	31_beta-cell_function.txt	20081858
Insulin resistance	30_insulin_resistance.txt	20081858
Fasting proinsulin	28_fasting_proinsulin.txt	21873549

Table E.7: Tissue expression for the filtered significant targets in the adipose network.

Drug Target	Tissue specificity		
	Human Protein Atlas	**GTEx**	**FANTOM5**
BTK	tissue enhanced (tonsil, lymph node)	tissue enhanced (spleen)	tissue enhanced (appendix, spleen, tonsil)
JAK1	expressed in all	expressed in all	expressed in all
PARP1	expressed in all	expressed in all	expressed in all
RARA	expressed in all	expressed in all	expressed in all
RARG	tissue enhanced (skin)	expressed in all	expressed in all
RXRA	mixed	expressed in all	expressed in all
RXRB	mixed	expressed in all	expressed in all
XIAP	expressed in all	expressed in all	expressed in all

Table E.8: Tissue expression for the filtered significant targets in the liver network.

Drug Target	**Tissue specificity**		
	Human Protein Atlas	**GTEx**	**FANTOM5**
FGFR3	tissue enriched (skin)	tissue enhanced (skin)	tissue enhanced (caudate, hippocampus)
NR1I2	group enriched (colon, duodenum, gallbladder, liver, rectum, small intestine)	group enriched (colon, liver, small intestine)	tissue enhanced (colon, small intestine)
RARA	expressed in all	expressed in all	expressed in all
RXRA	mixed	expressed in all	expressed in all
RXRB	mixed	expressed in all	expressed in all

Table E.9: Tissue expression for the filtered significant targets in the muscle network.

Drug Target	**Tissue specificity**		
	Human Protein Atlas	**GTEx**	**FANTOM5**
ABL1	expressed in all	expressed in all	expressed in all
ERBB2	expressed in all	expressed in all	expressed in all
HDAC1	expressed in all	expressed in all	expressed in all
HDAC2	expressed in all	expressed in all	expressed in all
HDAC3	expressed in all	expressed in all	expressed in all
HDAC4	mixed	expressed in all	expressed in all
JAK1	expressed in all	expressed in all	expressed in all
JAK2	mixed	expressed in all	expressed in all
MET	expressed in all	mixed	expressed in all
RARA	expressed in all	expressed in all	expressed in all
RARG	tissue enhanced (skin)	expressed in all	expressed in all
RXRA	mixed	expressed in all	expressed in all
RXRB	mixed	expressed in all	expressed in all

Table E.10: Pathways and GO categories identified as enriched in MetSyn genes.

NAME	ID	OVER-LAP	PVAL	ADJ. PVAL	DATABASE	MODULE NAME
carbohydrate homeostasis	GO:0033500	34/64	$4.47 \cdot 10^{-10}$	$6.86 \cdot 10^{-7}$	GO (BP)	Sugar metabolism
glucose homeostasis	GO:0042593	37/75	$1.21 \cdot 10^{-9}$	$1.39 \cdot 10^{-6}$	GO (BP)	Sugar metabolism
regulation of insulin secretion	GO:0050796	34/99	$1.20 \cdot 10^{-4}$	$2.75 \cdot 10^{-2}$	GO (BP)	Sugar metabolism
regulation of peptide hormone secretion	GO:0090276	27/74	$1.88 \cdot 10^{-4}$	$3.93 \cdot 10^{-2}$	GO (BP)	Sugar metabolism
Type II diabetes mellitus	hsa04930	19/48	$5.06 \cdot 10^{-4}$	$3.70 \cdot 10^{-2}$	KEGG	Sugar metabolism
Type II diabetes mellitus	WP1584	12/22	$1.55 \cdot 10^{-4}$	$9.31 \cdot 10^{-3}$	WikiPathways	Sugar metabolism
acylglycerol homeostasis	GO:0055090	20/23	$1.98 \cdot 10^{-12}$	$9.13 \cdot 10^{-9}$	GO (BP)	Lipid metabolism
cholesterol efflux	GO:0033344	14/24	$1.56 \cdot 10^{-5}$	$5.12 \cdot 10^{-3}$	GO (BP)	Lipid metabolism
cholesterol homeostasis	GO:0042632	27/59	$1.38 \cdot 10^{-6}$	$6.33 \cdot 10^{-4}$	GO (BP)	Lipid metabolism
cholesterol transport	GO:0030301	25/45	$2.75 \cdot 10^{-8}$	$2.11 \cdot 10^{-5}$	GO (BP)	Lipid metabolism
chylomicron remnant clearance	GO:0034382	7/8	$4.85 \cdot 10^{-5}$	$1.31 \cdot 10^{-2}$	GO (BP)	Lipid metabolism
Chylomicron-mediated lipid transport	R-HSA-174800	13/17	$3.17 \cdot 10^{-7}$	$2.22 \cdot 10^{-4}$	Reactome	Lipid metabolism
Composition of Lipid Particles	WP3601	8/9	$1.00 \cdot 10^{-5}$	$1.40 \cdot 10^{-3}$	WikiPathways	Lipid metabolism
Estrogen Receptor Pathway	WP2881	8/13	$7.05 \cdot 10^{-4}$	$2.68 \cdot 10^{-2}$	WikiPathways	Lipid metabolism
high-density lipoprotein particle remodeling	GO:0034375	13/18	$9.46 \cdot 10^{-7}$	$5.45 \cdot 10^{-4}$	GO (BP)	Lipid metabolism
LDL-mediated lipid transport	R-HSA-171052	6/6	$3.92 \cdot 10^{-5}$	$1.10 \cdot 10^{-2}$	Reactome	Lipid metabolism
Lipid digestion, mobilization, and transport	R-HSA-73923	28/71	$2.87 \cdot 10^{-5}$	$1.10 \cdot 10^{-2}$	Reactome	Lipid metabolism
lipid homeostasis	GO:0055088	21/50	$9.60 \cdot 10^{-5}$	$2.33 \cdot 10^{-2}$	GO (BP)	Lipid metabolism
Lipoprotein metabolism	R-HSA-174824	21/34	$2.79 \cdot 10^{-8}$	$3.91 \cdot 10^{-5}$	Reactome	Lipid metabolism
negative regulation of lipid metabolic process	GO:0045833	11/20	$2.67 \cdot 10^{-4}$	$4.56 \cdot 10^{-2}$	GO (BP)	Lipid metabolism
positive regulation of lipid biosynthetic process	GO:0046889	14/29	$2.43 \cdot 10^{-4}$	$4.35 \cdot 10^{-2}$	GO (BP)	Lipid metabolism
positive regulation of lipid metabolic process	GO:0045834	12/22	$1.55 \cdot 10^{-4}$	$3.41 \cdot 10^{-2}$	GO (BP)	Lipid metabolism
PPAR signaling pathway	hsa03320	26/69	$1.34 \cdot 10^{-4}$	$2.55 \cdot 10^{-2}$	KEGG	Lipid metabolism
regulation of lipoprotein lipase activity	GO:0051004	16/21	$1.38 \cdot 10^{-8}$	$1.27 \cdot 10^{-5}$	GO (BP)	Lipid metabolism
reverse cholesterol transport	GO:0043691	13/18	$9.46 \cdot 10^{-7}$	$5.45 \cdot 10^{-4}$	GO (BP)	Lipid metabolism
secondary alcohol metabolic process	GO:1902652	15/32	$2.21 \cdot 10^{-4}$	$4.24 \cdot 10^{-2}$	GO (BP)	Lipid metabolism
Statin Pathway	WP430	23/31	$2.07 \cdot 10^{-11}$	$8.69 \cdot 10^{-9}$	WikiPathways	Lipid metabolism
sterol homeostasis	GO:0055092	27/59	$1.38 \cdot 10^{-6}$	$6.33 \cdot 10^{-4}$	GO (BP)	Lipid metabolism
sterol metabolic process	GO:0016125	25/60	$2.52 \cdot 10^{-5}$	$7.75 \cdot 10^{-3}$	GO (BP)	Lipid metabolism
triglyceride homeostasis	GO:0070328	22/29	$2.72 \cdot 10^{-11}$	$6.26 \cdot 10^{-8}$	GO (BP)	Lipid metabolism
triglyceride-rich lipoprotein particle remodeling	GO:0034370	11/14	$1.74 \cdot 10^{-6}$	$7.24 \cdot 10^{-4}$	GO (BP)	Lipid metabolism
very-low-density lipoprotein particle	GO:0034361	11/16	$1.44 \cdot 10^{-5}$	$5.75 \cdot 10^{-3}$	GO (CC)	Lipid metabolism
very-low-density lipoprotein particle remodeling	GO:0034372	9/10	$2.05 \cdot 10^{-6}$	$7.24 \cdot 10^{-4}$	GO (BP)	Lipid metabolism
Adipocytokine signaling pathway	hsa04920	26/70	$1.76 \cdot 10^{-4}$	$2.55 \cdot 10^{-2}$	KEGG	Fat Storage
Adipogenesis	WP236	42/130	$1.01 \cdot 10^{-4}$	$7.08 \cdot 10^{-3}$	WikiPathways	Fat Storage
negative regulation of lipid storage	GO:0010888	10/17	$2.45 \cdot 10^{-4}$	$4.35 \cdot 10^{-2}$	GO (BP)	Fat Storage
negative regulation of macrophage derived foam cell differentiation	GO:0010745	9/14	$1.99 \cdot 10^{-4}$	$3.99 \cdot 10^{-2}$	GO (BP)	Fat Storage
Transcriptional regulation of white adipocyte differentiation	R-HSA-381340	29/79	$9.74 \cdot 10^{-5}$	$1.95 \cdot 10^{-2}$	Reactome	Fat Storage
Arylamine metabolism	WP694	6/6	$3.92 \cdot 10^{-5}$	$3.29 \cdot 10^{-3}$	WikiPathways	Chemical Modifications
Chemical carcinogenesis	hsa05204	28/82	$5.11 \cdot 10^{-4}$	$3.70 \cdot 10^{-2}$	KEGG	Chemical Modifications
Ethanol oxidation	R-HSA-71384	9/12	$3.14 \cdot 10^{-5}$	$1.10 \cdot 10^{-2}$	Reactome	Chemical Modifications
Fatty Acid Omega Oxidation	WP206	9/16	$7.96 \cdot 10^{-4}$	$2.78 \cdot 10^{-2}$	WikiPathways	Chemical Modifications
flavonoid glucuronidation	GO:0052696	9/10	$2.05 \cdot 10^{-6}$	$7.24 \cdot 10^{-4}$	GO (BP)	Chemical Modifications
xenobiotic glucuronidation	GO:0052697	9/12	$3.14 \cdot 10^{-5}$	$9.02 \cdot 10^{-3}$	GO (BP)	Chemical Modifications
Peptide hormone metabolism	R-HSA-2980736	30/81	$6.14 \cdot 10^{-5}$	$1.43 \cdot 10^{-2}$	Reactome	
renal system development	GO:0072001	24/60	$8.01 \cdot 10^{-5}$	$2.05 \cdot 10^{-2}$	GO (BP)	
Wnt Signaling Pathway	WP428	25/67	$2.14 \cdot 10^{-4}$	$1.12 \cdot 10^{-2}$	WikiPathways	

Table E.11: Final list of genes identified associated with MetSyn (MetSyn genes). MetSyn genes are identified by the official gene symbol and entrezid. The source indicates how the genes were identified (TM = Text Mining, GC = GWAS catalog, GS = GWAS summary statistics).

GENE SYMBOL	GENE ID	SOURCE	GENE SYMBOL	GENE ID	SOURCE	GENE SYMBOL	GENE ID	SOURCE
A2M	2	TM	FFAR1	2864	TM	ORMDL3	94103	GS
ABCA1	19	GC,GS,TM	FFAR2	2867	TM	OSBPL8	114882	TM
ABCB4	5244	TM	FFAR3	2865	TM	OSM	5008	TM
ABCC8	6833	GC,GS,TM	FFAR4	338557	TM	P4HB	5034	TM
ABCG1	9619	TM	FGA	2243	TM	PARK7	11315	TM
ABCG2	9429	TM	FGB	2244	TM	PAX2	5076	GC
ABCG5	64240	GC,GS,TM	FGFR4	2264	GC,TM	PAX6	5080	GC,TM
ABCG8	64241	GC,GS,TM	FGG	2266	TM	PAX8	7849	GC
ABHD5	51099	TM	FITM2	128486	GC,TM	PCK1	5105	GC,TM
ACAA1	30	TM	FOSL1	8061	TM	PCK2	5106	TM
ACACB	32	TM	FOXA2	3170	GC,GS,TM	PCSK1	5122	GC,GS,TM
ACE	1636	GC,TM	FOXC2	2303	TM	PCSK2	5126	TM
ACE2	59272	TM	FOXO1	2308	TM	PCSK5	5125	GC,TM
ACHE	43	GC,TM	FRAT1	10023	GC	PCSK9	255738	GC,GS,TM
ACOX1	51	GC,TM	FRZB	2487	TM	PDE3B	5140	GC,TM
ACOX2	8309	GC	FURIN	5045	GC,TM	PDIA2	64714	GC,TM
ACSBG1	23205	TM	FZD1	8321	TM	PDK2	5164	TM
ACSL1	2180	GC,TM	FZD2	2535	TM	PDK4	5166	GC,TM
ACSL3	2181	TM	FZD5	7855	GC	PDX1	3651	GC,GS,TM
ACSL4	2182	TM	FZD8	8325	TM	PER2	8864	GC,TM
ACSL5	51703	GC,TM	FZD9	8326	GC,GS	PHKA2	5256	TM
ACSM1	116285	TM	G6PC	2538	TM	PIK3CA	5290	TM
ACSM2A	123876	GC	G6PC2	57818	GC,GS,TM	PIK3CB	5291	TM
ACSM3	6296	TM	GADD45A	1647	TM	PIK3CD	5293	TM
ACSS2	55902	GC,TM	GATA2	2624	TM	PIK3CG	5294	GC,TM
ADCY5	111	GC,GS,TM	GATA3	2625	TM	PIK3R1	5295	TM
ADH1A	124	GC,TM	GATA4	2626	GC,TM	PIK3R2	5296	TM
ADH1B	125	GC,TM	GCG	2641	TM	PIK3R3	8503	GC
ADH1C	126	GC,TM	GCGR	2642	TM	PIK3R5	23533	TM
ADH4	127	GC,TM	GCK	2645	GC,GS,TM	PKD1	5310	TM
ADH5	128	GC	GCKR	2646	GC,GS,TM	PKHD1	5314	GC,TM
ADH6	130	GC	GDF10	2662	TM	PKM	5315	TM
ADH7	131	GC	GH1	2688	GC,TM	PLA2G10	8399	TM
ADIPOQ	9370	GC,TM	GHRL	51738	TM	PLA2G7	7941	GC,TM
ADIPOR1	51094	TM	GIP	2695	GC,TM	PLAU	5328	TM
ADIPOR2	79602	TM	GIPR	2696	GC,GS,TM	PLIN1	5346	TM
ADRA1B	147	TM	GK	2710	TM	PLIN2	123	TM
ADRA2A	150	TM	GLI2	2736	TM	PLIN5	440503	TM
ADRA2C	152	TM	GLP1R	2740	GC,TM	PLTP	5360	GC,GS,TM
AGPAT2	10555	TM	GLUD1	2746	TM	PNLIP	5406	TM
AGRP	181	GC,TM	GNAT3	346562	TM	PNLIPRP1	5407	GC
AGT	183	GC,TM	GNB3	2784	TM	PNLIPRP2	5408	GC
AGTR1	185	TM	GPAM	57678	GC,GS,TM	PNPLA2	57104	TM
AHR	196	TM	GPER1	2852	GC,TM	PNPLA3	80339	GC,TM
AKR1C2	1646	TM	GPIHBP1	338328	TM	PNPLA4	8228	TM
AKT1	207	GC,TM	GPLD1	2822	TM	PNPLA5	150379	TM
AKT2	208	GC,TM	GPR119	139760	TM	POMC	5443	GC,TM
AKT3	10000	TM	GRP	2922	GC,TM	PPARA	5465	GC,TM
ALB	213	TM	GSK3B	2932	GC,TM	PPARD	5467	TM
ALDH1A3	220	TM	GSTA1	2938	TM	PPARG	5468	GC,GS,TM
ALDH1B1	219	TM	GSTK1	373156	TM	PPARGC1A	10891	TM
ALDH2	217	GC,GS,TM	GSTM1	2944	GC,TM	PPP1CA	5499	TM
AMN	81693	GC	GSTM2	2946	GC,TM	PPP1CB	5500	TM
ANGPTL3	27329	GC,GS,TM	GSTM3	2947	TM	PPP1CC	5501	TM
ANGPTL4	51129	GC,TM	GSTM4	2948	GC,TM	PPP2R5E	5529	GC
ANGPTL8	55908	GC,TM	GSTM5	2949	GC,TM	PPP3CB	5532	TM
ANPEP	290	GC,TM	GSTO1	9446	TM	PRKAA1	5562	TM

APOA1	335	GC,GS,TM
APOA2	336	TM
APOA4	337	GC,GS,TM
APOA5	116519	GC,GS,TM
APOB	338	GC,GS,TM
APOC1	341	GC,GS,TM
APOC2	344	GC,GS,TM
APOC3	345	GC,GS,TM
APOC4	346	GC,GS,TM
APOD	347	TM
APOE	348	GC,GS,TM
APOF	319	GC,TM
APOH	350	GC,TM
APOL1	8542	TM
APOM	55937	GC,TM
AQP7	364	TM
ARL3	403	GC
ARNT	405	GC,TM
ARNTL	406	GC,TM
ATP6AP2	10159	TM
AXIN1	8312	GC,TM
BAD	572	TM
BAG6	7917	GC,TM
BCHE	590	TM
BLK	640	GC,TM
BMP2	650	TM
BMP4	652	TM
BMP8A	353500	GC,GS
BSCL2	26580	GC,TM
C1QTNF3	114899	TM
CACNA1A	773	TM
CACNA1C	775	TM
CACNA1D	776	GC,TM
CACNA1E	777	TM
CACNA2D2	9254	GC,TM
CAMKK2	10645	GC,TM
CAPN10	11132	TM
CARM1	10498	GS
CARTPT	9607	TM
CAV1	857	TM
CAV3	859	TM
CBR1	873	TM
CCL5	6352	TM
CCND1	595	TM
CCND2	894	GC,TM
CCND3	896	TM
CD36	948	GC,TM
CDK4	1019	TM
CDK8	1024	GC
CDKN1A	1026	TM
CDX2	1045	GC,GS,TM
CEACAM1	634	TM
CEBPA	1050	GC,TM
CEBPB	1051	TM
CEBPD	1052	TM
CEL	1056	TM
CELF1	10658	GC,GS,TM
CETP	1071	GC,GS,TM
CFD	1675	TM
CGA	1081	TM
CH25H	9023	TM
CHD7	55636	TM
CHUK	1147	TM
CIDEA	1149	TM
CLN8	2055	GC,TM
CLOCK	9575	TM
CLU	1191	TM
GSTP1	2950	TM
GSTT1	2952	TM
GTF3A	2971	GC
GZMH	2999	TM
HAS2	3037	TM
HCAR2	338442	GC,TM
HDAC3	8841	TM
HECTD4	283450	GC,GS,TM
HFE	3077	GC,TM
HIF1A	3091	TM
HK1	3098	GC,GS,TM
HK2	3099	TM
HKDC1	80201	GC,TM
HMGA1	3159	GC,TM
HMGCR	3156	GC,GS,TM
HMGCS2	3158	GC,TM
HNF1A	6927	GC,GS,TM
HNF1B	6928	GC,GS,TM
HNF4A	3172	GC,TM
HPGDS	27306	TM
HSD11B1	3290	TM
HSD17B13	345275	GC,TM
HSPG2	3339	TM
HTR2C	3358	TM
IFNG	3458	TM
IGF1	3479	GC,TM
IGF2	3481	GC,TM
IKBKB	3551	TM
IKBKG	8517	TM
IL18	3606	TM
IL1B	3553	TM
IL4	3565	TM
IL6	3569	GC,TM
IL6ST	3572	TM
ILK	3611	TM
INHBB	3625	TM
INHBC	3626	GC,GS
INHBE	83729	GC,GS,TM
INPP5K	51763	GC
INS	3630	GC,TM
INS-IGF2	723961	GC
INSIG1	3638	TM
INSR	3643	GC,TM
IRS1	3667	GC,TM
IRS2	8660	TM
IRS4	8471	TM
ISL1	3670	TM
ITGAV	3685	TM
ITGB3	3690	GC,TM
ITPR2	3709	GC
ITPR3	3710	TM
JAK2	3717	TM
JUN	3725	TM
KCNB1	3745	TM
KCNJ11	3767	GC,GS,TM
KCTD11	147040	GC
KIF5A	3798	GC
KLF15	28999	TM
KLF5	688	TM
KLF6	1316	TM
KREMEN1	83999	GC,TM
LCAT	3931	GC,GS,TM
LDLR	3949	GC,GS,TM
LDLRAP1	26119	GC,TM
LEP	3952	GC,TM
LEPR	3953	GC,TM
LIF	3976	TM
PRKAA2	5563	TM
PRKAB1	5564	TM
PRKAG1	5571	GC
PRKAG2	51422	TM
PRKAG3	53632	GC
PRKCA	5578	GC,TM
PRKCB	5579	TM
PRKCD	5580	GC,TM
PRKCE	5581	TM
PRKCQ	5588	TM
PRKCZ	5590	GC,TM
PRKD1	5587	TM
PRKN	5071	GC
PRLR	5618	TM
PROX1	5629	GC,TM
PSMD9	5715	TM
PTCH1	5727	GC,TM
PTGIS	5740	TM
PTGS2	5743	TM
PTPN11	5781	GC,GS,TM
PTPN2	5771	GC,TM
PYGO1	26108	TM
RAC1	5879	GC,TM
RAP1A	5906	TM
RAPGEF4	11069	GC
RARA	5914	GC,TM
RB1	5925	TM
RBL2	5934	GC,TM
RELA	5970	GC,TM
REN	5972	TM
RETN	56729	TM
RFX6	222546	GC,TM
RHOA	387	TM
RORA	6095	GC,TM
RPS6	6194	TM
RXRA	6256	TM
RXRB	6257	GC
RXRG	6258	TM
RYK	6259	GC
SALL1	6299	GC
SAR1B	51128	GC
SCARB1	949	GC,GS,TM
SCD	6319	GC,TM
SCP2	6342	TM
SDC1	6382	GC
SEC11A	23478	GC
SELENOS	55829	TM
SELENOT	51714	TM
SERAC1	84947	GC
SERPINE1	5054	TM
SFRP4	6424	TM
SHH	6469	TM
SIRT1	23411	TM
SIRT4	23409	TM
SIX1	6495	TM
SIX2	10736	GC
SLC25A4	291	TM
SLC25A5	292	TM
SLC25A6	293	TM
SLC27A1	376497	TM
SLC27A2	11001	TM
SLC27A4	10999	GC,TM
SLC27A5	10998	TM
SLC27A6	28965	TM
SLC2A1	6513	TM
SLC2A2	6514	GC,GS,TM
SLC2A4	6517	GC,TM

CMA1	1215	GC,TM	LIPC	3990	GC,GS,TM	SLC30A7	148867	TM
CNTFR	1271	TM	LIPE	3991	TM	SLC30A8	169026	GC,GS,TM
CPB1	1360	TM	LIPF	8513	TM	SLC8B1	80024	GC
CPB2	1361	TM	LIPG	9388	GC,GS,TM	SMAD3	4088	GC,TM
CPE	1363	GC	LIPI	149998	TM	SMO	6608	TM
CPN1	1369	TM	LIPN	643418	TM	SNAP25	6616	TM
CPT1A	1374	GC,TM	LMF1	64788	TM	SOAT1	6646	TM
CPT1B	1375	TM	LMNA	4000	TM	SOAT2	8435	GC
CPT1C	126129	TM	LPA	4018	GC,GS,TM	SOCS1	8651	TM
CPT2	1376	TM	LPIN1	23175	TM	SOCS2	8835	TM
CREB1	1385	TM	LPIN3	64900	GS	SOCS3	9021	GC,TM
CREBBP	1387	TM	LPL	4023	GC,GS,TM	SORBS1	10580	TM
CRHBP	1393	TM	LRP1	4035	GS,TM	SORT1	6272	GC,GS,TM
CRP	1401	GC,TM	LRP4	4038	GS	SOX17	64321	TM
CRTC2	200186	TM	LRP5	4041	TM	SP1	6667	TM
CRY1	1407	TM	LRP6	4040	TM	SPCS1	28972	GC
CRY2	1408	GC,GS,TM	MAFA	389692	TM	SQLE	6713	TM
CSNK1E	1454	GC,TM	MALL	7851	TM	SREBF1	6720	GC,TM
CTDNEP1	23399	GC	MAPK1	5594	TM	SREBF2	6721	TM
CTNNB1	1499	TM	MAPK10	5602	TM	SSTR5	6755	TM
CTSD	1509	TM	MAPK3	5595	TM	STARD3	10948	GC,GS,TM
CTSG	1511	TM	MAPK8	5599	TM	STAT1	6772	TM
CTSZ	1522	GC	MAPK9	5601	TM	STAT2	6773	GC,TM
CYP11A1	1583	TM	MARCKS	4082	TM	STAT3	6774	TM
CYP11B1	1584	GC,TM	MBNL1	4154	TM	STAT5A	6776	TM
CYP11B2	1585	GC,TM	MCU	90550	GC,TM	STAT6	6778	TM
CYP17A1	1586	GC,GS,TM	MED1	5469	GS,TM	STK11	6794	TM
CYP19A1	1588	TM	MED11	400569	GC	STX1A	6804	GC,TM
CYP1A1	1543	GS,TM	MED16	10025	TM	STXBP4	252983	TM
CYP1A2	1544	GC,GS,TM	MED17	9440	GC	SULF1	23213	TM
CYP1B1	1545	TM	MED22	6837	GC,GS	SULF2	55959	TM
CYP21A2	1589	GC,GS,TM	MED28	80306	TM	SULT1A1	6817	GC,TM
CYP26A1	1592	GC	MED30	90390	GC	SULT1A2	6799	GC
CYP26C1	340665	GC	MED7	9443	GC	SULT1A3	6818	GC,TM
CYP27A1	1593	GC,TM	MED9	55090	GC	SULT1A4	445329	GC
CYP2A6	1548	TM	MEF2A	4205	TM	SULT2A1	6822	TM
CYP2C19	1557	TM	MEF2C	4208	GC	SURF1	6834	GC,GS
CYP2C8	1558	TM	MGLL	11343	TM	SYT7	9066	GC
CYP2C9	1559	TM	MIF	4282	TM	TBL1XR1	79718	TM
CYP2D6	1565	TM	MIR33A	407039	TM	TBX18	9096	TM
CYP2E1	1571	TM	MIR33B	693120	GC,TM	TCF21	6943	GS,TM
CYP3A4	1576	TM	MIXL1	83881	TM	TCF7L2	6934	GC,GS,TM
CYP3A5	1577	TM	MLXIPL	51085	GC,GS,TM	TFAP2A	7020	TM
CYP46A1	10858	TM	MME	4311	GC,TM	TFAP2B	7021	GC,TM
CYP4A11	1579	TM	MMP1	4312	TM	TFR2	7036	TM
CYP4V2	285440	GC	MSMO1	6307	GC	TGFB1	7040	TM
CYP51A1	1595	TM	MSR1	4481	TM	TGFB2	7042	GC,TM
CYP7A1	1581	GC,GS,TM	MTNR1B	4544	GC,GS,TM	TGFBR1	7046	TM
CYP8B1	1582	TM	MTOR	2475	TM	THADA	63892	GC,GS,TM
DBI	1622	TM	MTTP	4547	TM	TNF	7124	GC,GS,TM
DDIT3	1649	GC,TM	MVK	4598	GC,GS	TNFRSF1A	7132	TM
DGAT1	8694	TM	MYC	4609	TM	TNFRSF1B	7133	TM
DGAT2	84649	GC,TM	MYO5A	4644	TM	TRADD	8717	GC,GS,TM
DHCR7	1717	TM	NAMPT	10135	TM	TRAF2	7186	TM
DLK1	8788	TM	NAT1	9	TM	TRH	7200	TM
DPP4	1803	TM	NAT2	10	GC,GS,TM	TRIB3	57761	TM
DVL1	1855	TM	NCOA1	8648	GC,TM	TWIST1	7291	TM
DVL2	1856	GC,TM	NCOA3	8202	TM	UBC	7316	GC,GS,TM
DVL3	1857	TM	NCOA6	23054	GC	UCP1	7350	TM
E2F1	1869	TM	NCOR1	9611	TM	UGT1A1	54658	GC,TM
E2F4	1874	GC,GS,TM	NCOR2	9612	GC,TM	UGT1A10	54575	GC
EBF1	1879	GC,TM	NDN	4692	TM	UGT1A3	54659	GC
EBP	10682	TM	NEUROD1	4760	TM	UGT1A4	54657	GC
EEF1A2	1917	TM	NFKB1	4790	TM	UGT1A5	54579	GC
EGR2	1959	GC	NFKBIA	4792	GC,TM	UGT1A6	54578	GC
EHD1	10938	TM	NFKBIB	4793	TM	UGT1A7	54577	GC
EHHADH	1962	TM	NGFR	4804	TM	UGT1A8	54576	GC

ENPEP	2028	GC,TM	NOV	4856	GC	UGT1A9	54600	GC
EP300	2033	TM	NPC1	4864	GC,TM	USF1	7391	TM
EPAS1	2034	TM	NPC1L1	29881	GC,GS,TM	VAMP2	6844	TM
EPHX1	2052	TM	NPY	4852	TM	VDR	7421	TM
EPHX2	2053	TM	NR0B2	8431	GC,TM	VEGFA	7422	GC,TM
ESR1	2099	TM	NR1D1	9572	TM	WFS1	7466	GC,GS,TM
EXOC2	55770	TM	NR1D2	9975	TM	WNT1	7471	GC,TM
EXOC6	54536	GC	NR1H2	7376	TM	WNT10A	80326	GC
F2	2147	GC,GS,TM	NR1H3	10062	GC,GS,TM	WNT10B	7480	GC,TM
FABP1	2168	TM	NR1H4	9971	TM	WNT2	7472	GC
FABP2	2169	TM	NR1I2	8856	GC,TM	WNT2B	7482	GC
FABP3	2170	TM	NR1I3	9970	TM	WNT3	7473	TM
FABP4	2167	TM	NR2F2	7026	TM	WNT3A	89780	GC,TM
FABP5	2171	TM	NR3C1	2908	TM	WNT4	54361	GC
FABP6	2172	GC,TM	NRIP1	8204	TM	WNT5A	7474	TM
FADS2	9415	GC,GS,TM	NUCKS1	64710	GC	WNT6	7475	GC
FAS	355	TM	NUP93	9688	GC,GS	WT1	7490	TM
FBXW2	26190	GC	OAS1	4938	GC	XBP1	7494	TM
FDFT1	2222	TM	OLR1	4973	TM	ZMPSTE24	10269	TM

Table E.12: Tissue-specific pathways for each MetSyn module.

ID	NAME	PVAL	ADJ. PVAL	TOP LEVEL PATHWAY	NETWORK MODULE
		Adipose Network			
R-HSA-3371556	Cellular response to heat stress	$3.32 \cdot 10^{-9}$	$9.65 \cdot 10^{-7}$	Cellular responses to external stimuli	Adipose_Mod_1
R-HSA-168256	Immune System	$6.52 \cdot 10^{-9}$	$9.65 \cdot 10^{-7}$	Immune System	Adipose_Mod_1
R-HSA-3371453	Regulation of HSF1-mediated heat shock response	$2.72 \cdot 10^{-8}$	$2.69 \cdot 10^{-6}$	Cellular responses to external stimuli	Adipose_Mod_1
R-HSA-2262752	Cellular responses to stress	$1.65 \cdot 10^{-7}$	$1.22 \cdot 10^{-5}$	Cellular responses to external stimuli	Adipose_Mod_1
R-HSA-8953897	Cellular responses to external stimuli	$2.39 \cdot 10^{-7}$	$1.42 \cdot 10^{-5}$	Cellular responses to external stimuli	Adipose_Mod_1
R-HSA-168249	Innate Immune System	$2.93 \cdot 10^{-7}$	$1.45 \cdot 10^{-5}$	Immune System	Adipose_Mod_1
R-HSA-422475	Axon guidance	$3.61 \cdot 10^{-7}$	$1.53 \cdot 10^{-5}$	Developmental Biology	Adipose_Mod_1
R-HSA-73887	Death Receptor Signalling	$9.44 \cdot 10^{-7}$	$3.49 \cdot 10^{-5}$	Signal Transduction	Adipose_Mod_1
R-HSA-1280215	Cytokine Signaling in Immune system	$1.15 \cdot 10^{-6}$	$3.77 \cdot 10^{-5}$	Immune System	Adipose_Mod_1
R-HSA-168164	Toll Like Receptor 3 (TLR3) Cascade	$1.36 \cdot 10^{-6}$	$4.02 \cdot 10^{-5}$	Immune System	Adipose_Mod_1
R-HSA-449147	Signaling by Interleukins	$1.60 \cdot 10^{-6}$	$4.31 \cdot 10^{-5}$	Immune System	Adipose_Mod_1
R-HSA-166166	MyD88-independent TLR4 cascade	$2.22 \cdot 10^{-6}$	$5.06 \cdot 10^{-5}$	Immune System	Adipose_Mod_1
R-HSA-937061	TRIF(TICAM1)-mediated TLR4 signaling	$2.22 \cdot 10^{-6}$	$5.06 \cdot 10^{-5}$	Immune System	Adipose_Mod_1
R-HSA-166016	Toll Like Receptor 4 (TLR4) Cascade	$8.63 \cdot 10^{-6}$	$1.83 \cdot 10^{-4}$	Immune System	Adipose_Mod_1
R-HSA-448424	Interleukin-17 signaling	$1.22 \cdot 10^{-5}$	$2.26 \cdot 10^{-4}$	Immune System	Adipose_Mod_1
R-HSA-450294	MAP kinase activation	$1.22 \cdot 10^{-5}$	$2.26 \cdot 10^{-4}$	Immune System	Adipose_Mod_1
R-HSA-168898	Toll-like Receptor Cascades	$1.31 \cdot 10^{-5}$	$2.28 \cdot 10^{-4}$	Immune System	Adipose_Mod_1
R-HSA-162582	Signal Transduction	$2.07 \cdot 10^{-5}$	$3.41 \cdot 10^{-4}$	Signal Transduction	Adipose_Mod_1
R-HSA-373760	L1CAM interactions	$2.28 \cdot 10^{-5}$	$3.55 \cdot 10^{-4}$	Developmental Biology	Adipose_Mod_1
R-HSA-76002	Platelet activation, signaling and aggregation	$2.94 \cdot 10^{-5}$	$4.07 \cdot 10^{-4}$	Hemostasis	Adipose_Mod_1
R-HSA-9006934	Signaling by Receptor Tyrosine Kinases	$3.38 \cdot 10^{-5}$	$4.07 \cdot 10^{-4}$	Signal Transduction	Adipose_Mod_1
R-HSA-168254	Influenza Infection	$3.94 \cdot 10^{-5}$	$4.07 \cdot 10^{-4}$	Disease	Adipose_Mod_1
R-HSA-168255	Influenza Life Cycle	$3.94 \cdot 10^{-5}$	$4.07 \cdot 10^{-4}$	Disease	Adipose_Mod_1
R-HSA-168138	Toll Like Receptor 9 (TLR9) Cascade	$4.12 \cdot 10^{-5}$	$4.07 \cdot 10^{-4}$	Immune System	Adipose_Mod_1
R-HSA-168142	Toll Like Receptor 10 (TLR10) Cascade	$4.12 \cdot 10^{-5}$	$4.07 \cdot 10^{-4}$	Immune System	Adipose_Mod_1
R-HSA-168176	Toll Like Receptor 5 (TLR5) Cascade	$4.12 \cdot 10^{-5}$	$4.07 \cdot 10^{-4}$	Immune System	Adipose_Mod_1
R-HSA-168181	Toll Like Receptor 7/8 (TLR7/8) Cascade	$4.12 \cdot 10^{-5}$	$4.07 \cdot 10^{-4}$	Immune System	Adipose_Mod_1
R-HSA-975138	TRAF6 mediated induction of NFkB and MAP kinases upon TLR7/8 or 9 activation	$4.12 \cdot 10^{-5}$	$4.07 \cdot 10^{-4}$	Immune System	Adipose_Mod_1
R-HSA-975155	MyD88 dependent cascade initiated on endosome	$4.12 \cdot 10^{-5}$	$4.07 \cdot 10^{-4}$	Immune System	Adipose_Mod_1
R-HSA-975871	MyD88 cascade initiated on plasma membrane	$4.12 \cdot 10^{-5}$	$4.07 \cdot 10^{-4}$	Immune System	Adipose_Mod_1
R-HSA-168643	Nucleotide-binding domain, leucine rich repeat containing receptor (NLR) signaling pathways	$6.57 \cdot 10^{-5}$	$6.27 \cdot 10^{-4}$	Immune System	Adipose_Mod_1
R-HSA-5653656	Vesicle-mediated transport	$7.64 \cdot 10^{-5}$	$7.07 \cdot 10^{-4}$	Vesicle-mediated transport	Adipose_Mod_1
R-HSA-166058	MyD88:MAL(TIRAP) cascade initiated on plasma membrane	$1.36 \cdot 10^{-4}$	$1.12 \cdot 10^{-3}$	Immune System	Adipose_Mod_1
R-HSA-168179	Toll Like Receptor TLR1:TLR2 Cascade	$1.36 \cdot 10^{-4}$	$1.12 \cdot 10^{-3}$	Immune System	Adipose_Mod_1
R-HSA-168188	Toll Like Receptor TLR6:TLR2 Cascade	$1.36 \cdot 10^{-4}$	$1.12 \cdot 10^{-3}$	Immune System	Adipose_Mod_1
R-HSA-181438	Toll Like Receptor 2 (TLR2) Cascade	$1.36 \cdot 10^{-4}$	$1.12 \cdot 10^{-3}$	Immune System	Adipose_Mod_1
R-HSA-446652	Interleukin-1 family signaling	$1.47 \cdot 10^{-4}$	$1.18 \cdot 10^{-3}$	Immune System	Adipose_Mod_1
R-HSA-194138	Signaling by VEGF	$2.07 \cdot 10^{-4}$	$1.57 \cdot 10^{-3}$	Signal Transduction	Adipose_Mod_1
R-HSA-4420097	VEGFA-VEGFR2 Pathway	$2.07 \cdot 10^{-4}$	$1.57 \cdot 10^{-3}$	Signal Transduction	Adipose_Mod_1
R-HSA-109581	Apoptosis	$2.35 \cdot 10^{-4}$	$1.70 \cdot 10^{-3}$	Programmed Cell Death	Adipose_Mod_1
R-HSA-5357801	Programmed Cell Death	$2.35 \cdot 10^{-4}$	$1.70 \cdot 10^{-3}$	Programmed Cell Death	Adipose_Mod_1
R-HSA-3371497	HSP90 chaperone cycle for steroid hormone receptors (SHR)	$2.93 \cdot 10^{-4}$	$2.07 \cdot 10^{-3}$	Cellular responses to external stimuli	Adipose_Mod_1
R-HSA-109606	Intrinsic Pathway for Apoptosis	$3.76 \cdot 10^{-4}$	$2.47 \cdot 10^{-3}$	Programmed Cell Death	Adipose_Mod_1
R-HSA-6804757	Regulation of TP53 Degradation	$3.76 \cdot 10^{-4}$	$2.47 \cdot 10^{-3}$	Gene expression (Transcription)	Adipose_Mod_1
R-HSA-6806003	Regulation of TP53 Expression and Degradation	$3.76 \cdot 10^{-4}$	$2.47 \cdot 10^{-3}$	Gene expression (Transcription)	Adipose_Mod_1
R-HSA-1640170	Cell Cycle	$4.28 \cdot 10^{-4}$	$2.75 \cdot 10^{-3}$	Cell Cycle	Adipose_Mod_1
R-HSA-199991	Membrane Trafficking	$4.50 \cdot 10^{-4}$	$2.84 \cdot 10^{-3}$	Vesicle-mediated transport	Adipose_Mod_1
R-HSA-5621481	C-type lectin receptors (CLRs)	$5.49 \cdot 10^{-4}$	$3.39 \cdot 10^{-3}$	Immune System	Adipose_Mod_1
R-HSA-5693532	DNA Double-Strand Break Repair	$6.93 \cdot 10^{-4}$	$4.19 \cdot 10^{-3}$	DNA Repair	Adipose_Mod_1

R-HSA-5607764	CLEC7A (Dectin-1) signaling	$8.60 \cdot 10^{-4}$	$4.99 \cdot 10^{-3}$	Immune System	Adipose_Mod_1
R-HSA-8951664	Neddylation	$8.60 \cdot 10^{-4}$	$4.99 \cdot 10^{-3}$	Metabolism of proteins	Adipose_Mod_1
R-HSA-983169	Class I MHC mediated antigen processing & presentation	$1.24 \cdot 10^{-3}$	$7.07 \cdot 10^{-3}$	Immune System	Adipose_Mod_1
R-HSA-109582	Hemostasis	$1.41 \cdot 10^{-3}$	$7.53 \cdot 10^{-3}$	Hemostasis	Adipose_Mod_1
R-HSA-112315	Transmission across Chemical Synapses	$1.42 \cdot 10^{-3}$	$7.53 \cdot 10^{-3}$	Neuronal System	Adipose_Mod_1
R-HSA-416476	G alpha (q) signalling events	$1.42 \cdot 10^{-3}$	$7.53 \cdot 10^{-3}$	Signal Transduction	Adipose_Mod_1
R-HSA-8856828	Clathrin-mediated endocytosis	$1.42 \cdot 10^{-3}$	$7.53 \cdot 10^{-3}$	Vesicle-mediated transport	Adipose_Mod_1
R-HSA-9020702	Interleukin-1 signaling	$1.64 \cdot 10^{-3}$	$8.50 \cdot 10^{-3}$	Immune System	Adipose_Mod_1
R-HSA-69278	Cell Cycle, Mitotic	$1.77 \cdot 10^{-3}$	$9.05 \cdot 10^{-3}$	Cell Cycle	Adipose_Mod_1
R-HSA-112316	Neuronal System	$2.65 \cdot 10^{-3}$	$1.33 \cdot 10^{-2}$	Neuronal System	Adipose_Mod_1
R-HSA-5693567	HDR through Homologous Recombination (HRR) or Single Strand Annealing (SSA)	$3.11 \cdot 10^{-3}$	$1.53 \cdot 10^{-2}$	DNA Repair	Adipose_Mod_1
R-HSA-168928	DDX58/IFIH1-mediated induction of interferon-alpha/beta	$3.82 \cdot 10^{-3}$	$1.82 \cdot 10^{-2}$	Immune System	Adipose_Mod_1
R-HSA-376176	Signaling by ROBO receptors	$3.82 \cdot 10^{-3}$	$1.82 \cdot 10^{-2}$	Developmental Biology	Adipose_Mod_1
R-HSA-372790	Signaling by GPCR	$3.95 \cdot 10^{-3}$	$1.83 \cdot 10^{-2}$	Signal Transduction	Adipose_Mod_1
R-HSA-388396	GPCR downstream signalling	$3.95 \cdot 10^{-3}$	$1.83 \cdot 10^{-2}$	Signal Transduction	Adipose_Mod_1
R-HSA-2454202	Fc epsilon receptor (FCERI) signaling	$4.10 \cdot 10^{-3}$	$1.84 \cdot 10^{-2}$	Immune System	Adipose_Mod_1
R-HSA-69620	Cell Cycle Checkpoints	$4.10 \cdot 10^{-3}$	$1.84 \cdot 10^{-2}$	Cell Cycle	Adipose_Mod_1
R-HSA-1643685	Disease	$4.21 \cdot 10^{-3}$	$1.86 \cdot 10^{-2}$	Disease	Adipose_Mod_1
R-HSA-1227986	Signaling by ERBB2	$4.54 \cdot 10^{-3}$	$1.97 \cdot 10^{-2}$	Signal Transduction	Adipose_Mod_1
R-HSA-2219528	PI3K/AKT Signaling in Cancer	$4.66 \cdot 10^{-3}$	$1.97 \cdot 10^{-2}$	Disease	Adipose_Mod_1
R-HSA-5693538	Homology Directed Repair	$4.66 \cdot 10^{-3}$	$1.97 \cdot 10^{-2}$	DNA Repair	Adipose_Mod_1
R-HSA-5663202	Diseases of signal transduction	$4.96 \cdot 10^{-3}$	$1.99 \cdot 10^{-2}$	Disease	Adipose_Mod_1
R-HSA-111885	Opioid Signalling	$5.10 \cdot 10^{-3}$	$1.99 \cdot 10^{-2}$	Signal Transduction	Adipose_Mod_1
R-HSA-2565942	Regulation of PLK1 Activity at G2/M Transition	$5.10 \cdot 10^{-3}$	$1.99 \cdot 10^{-2}$	Cell Cycle	Adipose_Mod_1
R-HSA-69563	p53-Dependent G1 DNA Damage Response	$5.10 \cdot 10^{-3}$	$1.99 \cdot 10^{-2}$	Cell Cycle	Adipose_Mod_1
R-HSA-69580	p53-Dependent G1/S DNA damage checkpoint	$5.10 \cdot 10^{-3}$	$1.99 \cdot 10^{-2}$	Cell Cycle	Adipose_Mod_1
R-HSA-69615	G1/S DNA Damage Checkpoints	$5.10 \cdot 10^{-3}$	$1.99 \cdot 10^{-2}$	Cell Cycle	Adipose_Mod_1
R-HSA-8878166	Transcriptional regulation by RUNX2	$5.25 \cdot 10^{-3}$	$2.02 \cdot 10^{-2}$	Gene expression (Transcription)	Adipose_Mod_1
R-HSA-5693607	Processing of DNA double-strand break ends	$5.89 \cdot 10^{-3}$	$2.24 \cdot 10^{-2}$	DNA Repair	Adipose_Mod_1
R-HSA-193704	p75 NTR receptor-mediated signalling	$7.31 \cdot 10^{-3}$	$2.64 \cdot 10^{-2}$	Signal Transduction	Adipose_Mod_1
R-HSA-2871837	FCERI mediated NF-kB activation	$7.31 \cdot 10^{-3}$	$2.64 \cdot 10^{-2}$	Immune System	Adipose_Mod_1
R-HSA-69473	G2/M DNA damage checkpoint	$7.31 \cdot 10^{-3}$	$2.64 \cdot 10^{-2}$	Cell Cycle	Adipose_Mod_1
R-HSA-73894	DNA Repair	$7.36 \cdot 10^{-3}$	$2.64 \cdot 10^{-2}$	DNA Repair	Adipose_Mod_1
R-HSA-1280218	Adaptive Immune System	$7.41 \cdot 10^{-3}$	$2.64 \cdot 10^{-2}$	Immune System	Adipose_Mod_1
R-HSA-2559583	Cellular Senescence	$8.52 \cdot 10^{-3}$	$2.98 \cdot 10^{-2}$	Cellular responses to external stimuli	Adipose_Mod_1
R-HSA-174143	APC/C-mediated degradation of cell cycle proteins	$8.77 \cdot 10^{-3}$	$2.98 \cdot 10^{-2}$	Cell Cycle	Adipose_Mod_1
R-HSA-450531	Regulation of mRNA stability by proteins that bind AU-rich elements	$8.77 \cdot 10^{-3}$	$2.98 \cdot 10^{-2}$	Metabolism of RNA	Adipose_Mod_1
R-HSA-453276	Regulation of mitotic cell cycle	$8.77 \cdot 10^{-3}$	$2.98 \cdot 10^{-2}$	Cell Cycle	Adipose_Mod_1
R-HSA-9006925	Intracellular signaling by second messengers	$9.14 \cdot 10^{-3}$	$3.07 \cdot 10^{-2}$	Signal Transduction	Adipose_Mod_1
R-HSA-69275	G2/M Transition	$9.78 \cdot 10^{-3}$	$3.25 \cdot 10^{-2}$	Cell Cycle	Adipose_Mod_1
R-HSA-2029480	Fcgamma receptor (FCGR) dependent phagocytosis	$1.12 \cdot 10^{-2}$	$3.63 \cdot 10^{-2}$	Immune System	Adipose_Mod_1
R-HSA-418594	G alpha (i) signalling events	$1.12 \cdot 10^{-2}$	$3.63 \cdot 10^{-2}$	Signal Transduction	Adipose_Mod_1
R-HSA-2559580	Oxidative Stress Induced Senescence	$1.25 \cdot 10^{-2}$	$3.97 \cdot 10^{-2}$	Cellular responses to external stimuli	Adipose_Mod_1
R-HSA-450282	MAPK targets/ Nuclear events mediated by MAP kinases	$1.25 \cdot 10^{-2}$	$3.97 \cdot 10^{-2}$	Immune System	Adipose_Mod_1
R-HSA-453274	Mitotic G2-G2/M phases	$1.30 \cdot 10^{-2}$	$4.10 \cdot 10^{-2}$	Cell Cycle	Adipose_Mod_1
R-HSA-6804756	Regulation of TP53 Activity through Phosphorylation	$1.63 \cdot 10^{-2}$	$4.97 \cdot 10^{-2}$	Gene expression (Transcription)	Adipose_Mod_1
R-HSA-202403	TCR signaling	$1.63 \cdot 10^{-2}$	$4.97 \cdot 10^{-2}$	Immune System	Adipose_Mod_1
R-HSA-202424	Downstream TCR signaling	$1.63 \cdot 10^{-2}$	$4.97 \cdot 10^{-2}$	Immune System	Adipose_Mod_1
R-HSA-1989781	PPARA activates gene expression	$1.07 \cdot 10^{-13}$	$1.25 \cdot 10^{-11}$	Metabolism	Adipose_Mod_2
R-HSA-400206	Regulation of lipid metabolism by Peroxisome proliferator-activated receptor alpha (PPARalpha)	$3.13 \cdot 10^{-13}$	$1.82 \cdot 10^{-11}$	Metabolism	Adipose_Mod_2
R-HSA-556833	Metabolism of lipids	$5.16 \cdot 10^{-12}$	$1.99 \cdot 10^{-10}$	Metabolism	Adipose_Mod_2
R-HSA-383280	Nuclear Receptor transcription pathway	$1.43 \cdot 10^{-11}$	$4.16 \cdot 10^{-10}$	Gene expression (Transcription)	Adipose_Mod_2
R-HSA-1430728	Metabolism	$9.28 \cdot 10^{-8}$	$2.15 \cdot 10^{-6}$	Metabolism	Adipose_Mod_2
R-HSA-381340	Transcriptional regulation of white adipocyte differentiation	$4.53 \cdot 10^{-7}$	$8.75 \cdot 10^{-6}$	Developmental Biology	Adipose_Mod_2

R-HSA-1368108	BMAL1:CLOCK,NPAS2 activates circadian gene expression	$4.57 \cdot 10^{-6}$	$6.92 \cdot 10^{-5}$	Circadian Clock	Adipose_Mod_2
R-HSA-400253	Circadian Clock	$4.77 \cdot 10^{-6}$	$6.92 \cdot 10^{-5}$	Circadian Clock	Adipose_Mod_2
R-HSA-9006931	Signaling by Nuclear Receptors	$3.20 \cdot 10^{-5}$	$4.12 \cdot 10^{-4}$	Signal Transduction	Adipose_Mod_2
R-HSA-2990846	SUMOylation	$6.64 \cdot 10^{-5}$	$7.01 \cdot 10^{-4}$	Metabolism of proteins	Adipose_Mod_2
R-HSA-3108232	SUMO E3 ligases SUMOylate target proteins	$6.64 \cdot 10^{-5}$	$7.01 \cdot 10^{-4}$	Metabolism of proteins	Adipose_Mod_2
R-HSA-3108232	SUMO E3 ligases SUMOylate target proteins	$6.64 \cdot 10^{-5}$	$7.01 \cdot 10^{-4}$	Metabolism of proteins	Adipose_Mod_2
R-HSA-4090294	SUMOylation of intracellular receptors	$9.37 \cdot 10^{-5}$	$9.06 \cdot 10^{-4}$	NA	Adipose_Mod_2
R-HSA-212436	Generic Transcription Pathway	$5.69 \cdot 10^{-4}$	$4.82 \cdot 10^{-3}$	Gene expression (Transcription)	Adipose_Mod_2
R-HSA-1368082	RORA activates gene expression	$5.82 \cdot 10^{-4}$	$4.82 \cdot 10^{-3}$	Circadian Clock	Adipose_Mod_2
R-HSA-73857	RNA Polymerase II Transcription	$1.23 \cdot 10^{-3}$	$9.53 \cdot 10^{-3}$	Gene expression (Transcription)	Adipose_Mod_2
R-HSA-2151201	Transcriptional activation of mitochondrial biogenesis	$1.57 \cdot 10^{-3}$	$1.14 \cdot 10^{-2}$	Organelle biogenesis and maintenance	Adipose_Mod_2
R-HSA-74160	Gene expression (Transcription)	$2.63 \cdot 10^{-3}$	$1.70 \cdot 10^{-2}$	Gene expression (Transcription)	Adipose_Mod_2
R-HSA-1592230	Mitochondrial biogenesis	$2.64 \cdot 10^{-3}$	$1.70 \cdot 10^{-2}$	Organelle biogenesis and maintenance	Adipose_Mod_2
R-HSA-8957322	Metabolism of steroids	$3.35 \cdot 10^{-3}$	$2.04 \cdot 10^{-2}$	Metabolism	Adipose_Mod_2
R-HSA-2426168	Activation of gene expression by SREBF (SREBP)	$3.53 \cdot 10^{-3}$	$2.05 \cdot 10^{-2}$	Metabolism	Adipose_Mod_2
R-HSA-1655829	Regulation of cholesterol biosynthesis by SREBP (SREBF)	$4.65 \cdot 10^{-3}$	$2.57 \cdot 10^{-2}$	Metabolism	Adipose_Mod_2
R-HSA-5617472	Activation of anterior HOX genes in hindbrain development during early embryogenesis	$5.17 \cdot 10^{-3}$	$2.61 \cdot 10^{-2}$	Developmental Biology	Adipose_Mod_2
R-HSA-5619507	Activation of HOX genes during differentiation	$5.17 \cdot 10^{-3}$	$2.61 \cdot 10^{-2}$	Developmental Biology	Adipose_Mod_2
Liver Network					
R-HSA-156580	Phase II - Conjugation of compounds	$2.00 \cdot 10^{-4}$	$8.99 \cdot 10^{-3}$	Metabolism	Liver_Mod_1
R-HSA-1989781	PPARA activates gene expression	$1.10 \cdot 10^{-12}$	$6.72 \cdot 10^{-11}$	Metabolism	Liver_Mod_2
R-HSA-383280	Nuclear Receptor transcription pathway	$1.16 \cdot 10^{-12}$	$6.72 \cdot 10^{-11}$	Gene expression (Transcription)	Liver_Mod_2
R-HSA-400206	Regulation of lipid metabolism by Peroxisome proliferator-activated receptor alpha (PPARalpha)	$1.99 \cdot 10^{-12}$	$7.70 \cdot 10^{-11}$	Metabolism	Liver_Mod_2
R-HSA-400253	Circadian Clock	$2.37 \cdot 10^{-11}$	$6.86 \cdot 10^{-10}$	Circadian Clock	Liver_Mod_2
R-HSA-1368108	BMAL1:CLOCK,NPAS2 activates circadian gene expression	$9.61 \cdot 10^{-11}$	$2.23 \cdot 10^{-9}$	Circadian Clock	Liver_Mod_2
R-HSA-556833	Metabolism of lipids	$2.12 \cdot 10^{-9}$	$4.11 \cdot 10^{-8}$	Metabolism	Liver_Mod_2
R-HSA-556833	Metabolism of lipids	$2.12 \cdot 10^{-9}$	$4.11 \cdot 10^{-8}$	Metabolism	Liver_Mod_2
R-HSA-4090294	SUMOylation of intracellular receptors	$2.17 \cdot 10^{-8}$	$3.60 \cdot 10^{-7}$		Liver_Mod_2
R-HSA-1368082	RORA activates gene expression	$1.19 \cdot 10^{-7}$	$1.73 \cdot 10^{-6}$	Circadian Clock	Liver_Mod_2
R-HSA-3108232	SUMO E3 ligases SUMOylate target proteins	$3.20 \cdot 10^{-6}$	$4.13 \cdot 10^{-5}$	Metabolism of proteins	Liver_Mod_2
R-HSA-2990846	SUMOylation	$3.72 \cdot 10^{-6}$	$4.32 \cdot 10^{-5}$	Metabolism of proteins	Liver_Mod_2
R-HSA-2151201	Transcriptional activation of mitochondrial biogenesis	$1.74 \cdot 10^{-5}$	$1.84 \cdot 10^{-4}$	Organelle biogenesis and maintenance	Liver_Mod_2
R-HSA-2426168	Activation of gene expression by SREBF (SREBP)	$3.20 \cdot 10^{-5}$	$3.09 \cdot 10^{-4}$	Metabolism	Liver_Mod_2
R-HSA-381340	Transcriptional regulation of white adipocyte differentiation	$4.00 \cdot 10^{-5}$	$3.57 \cdot 10^{-4}$	Developmental Biology	Liver_Mod_2
R-HSA-1430728	Metabolism	$6.10 \cdot 10^{-5}$	$5.05 \cdot 10^{-4}$	Metabolism	Liver_Mod_2
R-HSA-1592230	Mitochondrial biogenesis	$8.59 \cdot 10^{-5}$	$6.64 \cdot 10^{-4}$	Organelle biogenesis and maintenance	Liver_Mod_2
R-HSA-1655829	Regulation of cholesterol biosynthesis by SREBP (SREBF)	$1.14 \cdot 10^{-4}$	$8.28 \cdot 10^{-4}$	Metabolism	Liver_Mod_2
R-HSA-211945	Phase I - Functionalization of compounds	$1.50 \cdot 10^{-4}$	$1.02 \cdot 10^{-3}$	Metabolism	Liver_Mod_2
R-HSA-1852241	Organelle biogenesis and maintenance	$4.30 \cdot 10^{-4}$	$2.77 \cdot 10^{-3}$	Organelle biogenesis and maintenance	Liver_Mod_2
R-HSA-211897	Cytochrome P450 - arranged by substrate type	$5.46 \cdot 10^{-4}$	$3.33 \cdot 10^{-3}$	Metabolism	Liver_Mod_2
R-HSA-211859	Biological oxidations	$1.03 \cdot 10^{-3}$	$5.99 \cdot 10^{-3}$	Metabolism	Liver_Mod_2
R-HSA-8957322	Metabolism of steroids	$2.29 \cdot 10^{-3}$	$1.26 \cdot 10^{-2}$	Metabolism	Liver_Mod_2
R-HSA-194068	Bile acid and bile salt metabolism	$9.42 \cdot 10^{-3}$	$4.97 \cdot 10^{-2}$	Metabolism	Liver_Mod_2
R-HSA-381340	Transcriptional regulation of white adipocyte differentiation	$1.16 \cdot 10^{-27}$	$5.82 \cdot 10^{-26}$	Developmental Biology	Liver_Mod_3
R-HSA-1989781	PPARA activates gene expression	$1.44 \cdot 10^{-24}$	$3.59 \cdot 10^{-23}$	Metabolism	Liver_Mod_3
R-HSA-400206	Regulation of lipid metabolism by Peroxisome proliferator-activated receptor alpha (PPARalpha)	$2.84 \cdot 10^{-24}$	$4.74 \cdot 10^{-23}$	Metabolism	Liver_Mod_3
R-HSA-556833	Metabolism of lipids	$1.28 \cdot 10^{-18}$	$1.60 \cdot 10^{-17}$	Metabolism	Liver_Mod_3
R-HSA-1266738	Developmental Biology	$9.24 \cdot 10^{-13}$	$9.24 \cdot 10^{-12}$	Developmental Biology	Liver_Mod_3
R-HSA-1430728	Metabolism	$4.39 \cdot 10^{-12}$	$3.66 \cdot 10^{-11}$	Metabolism	Liver_Mod_3

R-HSA-212436	Generic Transcription Pathway	$5.97 \cdot 10^{-10}$	$4.27 \cdot 10^{-9}$	Gene expression (Transcription)	Liver_Mod_3
R-HSA-73857	RNA Polymerase II Transcription	$1.05 \cdot 10^{-9}$	$6.59 \cdot 10^{-9}$	Gene expression (Transcription)	Liver_Mod_3
R-HSA-74160	Gene expression (Transcription)	$5.82 \cdot 10^{-9}$	$3.23 \cdot 10^{-8}$	Gene expression (Transcription)	Liver_Mod_3
R-HSA-2644602	Signaling by NOTCH1 PEST Domain Mutants in Cancer	$4.80 \cdot 10^{-3}$	$1.71 \cdot 10^{-2}$	Disease	Liver_Mod_3
R-HSA-2644603	Signaling by NOTCH1 in Cancer	$4.80 \cdot 10^{-3}$	$1.71 \cdot 10^{-2}$	Disease	Liver_Mod_3
R-HSA-2644606	Constitutive Signaling by NOTCH1 PEST Domain Mutants	$4.80 \cdot 10^{-3}$	$1.71 \cdot 10^{-2}$	Disease	Liver_Mod_3
R-HSA-2894858	Signaling by NOTCH1 HD+PEST Domain Mutants in Cancer	$4.80 \cdot 10^{-3}$	$1.71 \cdot 10^{-2}$	Disease	Liver_Mod_3
R-HSA-2894862	Constitutive Signaling by NOTCH1 HD+PEST Domain Mutants	$4.80 \cdot 10^{-3}$	$1.71 \cdot 10^{-2}$	Disease	Liver_Mod_3
R-HSA-2122947	NOTCH1 Intracellular Domain Regulates Transcription	$5.50 \cdot 10^{-3}$	$1.83 \cdot 10^{-2}$	Signal Transduction	Liver_Mod_3
R-HSA-1980143	Signaling by NOTCH1	$7.94 \cdot 10^{-3}$	$2.48 \cdot 10^{-2}$	Signal Transduction	Liver_Mod_3
	Muscle Network				
R-HSA-2990846	SUMOylation	$1.92 \cdot 10^{-9}$	$3.45 \cdot 10^{-7}$	Metabolism of proteins	Muscle_Mod_1
R-HSA-3108232	SUMO E3 ligases SUMOylate target proteins	$1.92 \cdot 10^{-9}$	$3.45 \cdot 10^{-7}$	Metabolism of proteins	Muscle_Mod_1
R-HSA-2559583	Cellular Senescence	$3.78 \cdot 10^{-9}$	$4.54 \cdot 10^{-7}$	Cellular responses to external stimuli	Muscle_Mod_1
R-HSA-6807070	PTEN Regulation	$2.57 \cdot 10^{-8}$	$2.32 \cdot 10^{-6}$	Signal Transduction	Muscle_Mod_1
R-HSA-3214815	HDACs deacetylate histones	$4.87 \cdot 10^{-8}$	$2.92 \cdot 10^{-6}$	Chromatin organization	Muscle_Mod_1
R-HSA-3232118	SUMOylation of transcription factors	$4.87 \cdot 10^{-8}$	$2.92 \cdot 10^{-6}$	Metabolism of proteins	Muscle_Mod_1
R-HSA-8943724	Regulation of PTEN gene transcription	$8.78 \cdot 10^{-8}$	$4.51 \cdot 10^{-6}$	Signal Transduction	Muscle_Mod_1
R-HSA-392499	Metabolism of proteins	$1.98 \cdot 10^{-7}$	$8.91 \cdot 10^{-6}$	Metabolism of proteins	Muscle_Mod_1
R-HSA-3247509	Chromatin modifying enzymes	$5.10 \cdot 10^{-7}$	$1.84 \cdot 10^{-5}$	Chromatin organization	Muscle_Mod_1
R-HSA-4839726	Chromatin organization	$5.10 \cdot 10^{-7}$	$1.84 \cdot 10^{-5}$	Chromatin organization	Muscle_Mod_1
R-HSA-5693532	DNA Double-Strand Break Repair	$2.66 \cdot 10^{-6}$	$8.70 \cdot 10^{-5}$	DNA Repair	Muscle_Mod_1
R-HSA-2408522	Selenoamino acid metabolism	$5.29 \cdot 10^{-6}$	$1.47 \cdot 10^{-4}$	Metabolism	Muscle_Mod_1
R-HSA-4551638	SUMOylation of chromatin organization proteins	$5.29 \cdot 10^{-6}$	$1.47 \cdot 10^{-4}$	Metabolism of proteins	Muscle_Mod_1
R-HSA-2262752	Cellular responses to stress	$8.40 \cdot 10^{-6}$	$2.16 \cdot 10^{-4}$	Cellular responses to external stimuli	Muscle_Mod_1
R-HSA-8939243	RUNX1 interacts with co-factors whose precise effect on RUNX1 targets is not known	$9.17 \cdot 10^{-6}$	$2.20 \cdot 10^{-4}$	Gene expression (Transcription)	Muscle_Mod_1
R-HSA-597592	Post-translational protein modification	$1.11 \cdot 10^{-5}$	$2.50 \cdot 10^{-4}$	Metabolism of proteins	Muscle_Mod_1
R-HSA-3899300	SUMOylation of transcription cofactors	$1.80 \cdot 10^{-5}$	$3.13 \cdot 10^{-4}$	Metabolism of proteins	Muscle_Mod_1
R-HSA-168164	Toll Like Receptor 3 (TLR3) Cascade	$1.90 \cdot 10^{-5}$	$3.13 \cdot 10^{-4}$	Immune System	Muscle_Mod_1
R-HSA-166058	MyD88:MAL(TIRAP) cascade initiated on plasma membrane	$2.53 \cdot 10^{-5}$	$3.13 \cdot 10^{-4}$	Immune System	Muscle_Mod_1
R-HSA-168138	Toll Like Receptor 9 (TLR9) Cascade	$2.53 \cdot 10^{-5}$	$3.13 \cdot 10^{-4}$	Immune System	Muscle_Mod_1
R-HSA-168142	Toll Like Receptor 10 (TLR10) Cascade	$2.53 \cdot 10^{-5}$	$3.13 \cdot 10^{-4}$	Immune System	Muscle_Mod_1
R-HSA-168176	Toll Like Receptor 5 (TLR5) Cascade	$2.53 \cdot 10^{-5}$	$3.13 \cdot 10^{-4}$	Immune System	Muscle_Mod_1
R-HSA-168179	Toll Like Receptor TLR1:TLR2 Cascade	$2.53 \cdot 10^{-5}$	$3.13 \cdot 10^{-4}$	Immune System	Muscle_Mod_1
R-HSA-168181	Toll Like Receptor 7/8 (TLR7/8) Cascade	$2.53 \cdot 10^{-5}$	$3.13 \cdot 10^{-4}$	Immune System	Muscle_Mod_1
R-HSA-168188	Toll Like Receptor TLR6:TLR2 Cascade	$2.53 \cdot 10^{-5}$	$3.13 \cdot 10^{-4}$	Immune System	Muscle_Mod_1
R-HSA-181438	Toll Like Receptor 2 (TLR2) Cascade	$2.53 \cdot 10^{-5}$	$3.13 \cdot 10^{-4}$	Immune System	Muscle_Mod_1
R-HSA-975138	TRAF6 mediated induction of NFkB and MAP kinases upon TLR7/8 or 9 activation	$2.53 \cdot 10^{-5}$	$3.13 \cdot 10^{-4}$	Immune System	Muscle_Mod_1
R-HSA-975155	MyD88 dependent cascade initiated on endosome	$2.53 \cdot 10^{-5}$	$3.13 \cdot 10^{-4}$	Immune System	Muscle_Mod_1
R-HSA-975871	MyD88 cascade initiated on plasma membrane	$2.53 \cdot 10^{-5}$	$3.13 \cdot 10^{-4}$	Immune System	Muscle_Mod_1
R-HSA-3108214	SUMOylation of DNA damage response and repair proteins	$2.70 \cdot 10^{-5}$	$3.13 \cdot 10^{-4}$	Metabolism of proteins	Muscle_Mod_1
R-HSA-5693607	Processing of DNA double-strand break ends	$2.70 \cdot 10^{-5}$	$3.13 \cdot 10^{-4}$	DNA Repair	Muscle_Mod_1
R-HSA-166166	MyD88-independent TLR4 cascade	$3.54 \cdot 10^{-5}$	$3.77 \cdot 10^{-4}$	Immune System	Muscle_Mod_1
R-HSA-937061	TRIF(TICAM1)-mediated TLR4 signaling	$3.54 \cdot 10^{-5}$	$3.77 \cdot 10^{-4}$	Immune System	Muscle_Mod_1
R-HSA-5693567	HDR through Homologous Recombination (HRR) or Single Strand Annealing (SSA)	$3.56 \cdot 10^{-5}$	$3.77 \cdot 10^{-4}$	DNA Repair	Muscle_Mod_1
R-HSA-1257604	PIP3 activates AKT signaling	$3.95 \cdot 10^{-5}$	$4.06 \cdot 10^{-4}$	Signal Transduction	Muscle_Mod_1
R-HSA-5693538	Homology Directed Repair	$4.15 \cdot 10^{-5}$	$4.15 \cdot 10^{-4}$	DNA Repair	Muscle_Mod_1
R-HSA-71291	Metabolism of amino acids and derivatives	$4.47 \cdot 10^{-5}$	$4.35 \cdot 10^{-4}$	Metabolism	Muscle_Mod_1
R-HSA-1280215	Cytokine Signaling in Immune system	$5.05 \cdot 10^{-5}$	$4.77 \cdot 10^{-4}$	Immune System	Muscle_Mod_1
R-HSA-8878171	Transcriptional regulation by RUNX1	$5.16 \cdot 10^{-5}$	$4.77 \cdot 10^{-4}$	Gene expression (Transcription)	Muscle_Mod_1
R-HSA-166016	Toll Like Receptor 4 (TLR4) Cascade	$6.33 \cdot 10^{-5}$	$5.70 \cdot 10^{-4}$	Immune System	Muscle_Mod_1

R-HSA-2559582	Senescence-Associated Secretory Phenotype (SASP)	$9.81 \cdot 10^{-5}$	$8.62 \cdot 10^{-4}$	Cellular responses to external stimuli	Muscle_Mod_1
R-HSA-168898	Toll-like Receptor Cascades	$1.10 \cdot 10^{-4}$	$9.40 \cdot 10^{-4}$	Immune System	Muscle_Mod_1
R-HSA-1834949	Cytosolic sensors of pathogen-associated DNA	$1.15 \cdot 10^{-4}$	$9.59 \cdot 10^{-4}$	Immune System	Muscle_Mod_1
R-HSA-9006925	Intracellular signaling by second messengers	$1.39 \cdot 10^{-4}$	$1.14 \cdot 10^{-3}$	Signal Transduction	Muscle_Mod_1
R-HSA-2559586	DNA Damage/Telomere Stress Induced Senescence	$1.56 \cdot 10^{-4}$	$1.25 \cdot 10^{-3}$	Cellular responses to external stimuli	Muscle_Mod_1
R-HSA-69563	p53-Dependent G1 DNA Damage Response	$2.24 \cdot 10^{-4}$	$1.68 \cdot 10^{-3}$	Cell Cycle	Muscle_Mod_1
R-HSA-69580	p53-Dependent G1/S DNA damage checkpoint	$2.24 \cdot 10^{-4}$	$1.68 \cdot 10^{-3}$	Cell Cycle	Muscle_Mod_1
R-HSA-69615	G1/S DNA Damage Checkpoints	$2.24 \cdot 10^{-4}$	$1.68 \cdot 10^{-3}$	Cell Cycle	Muscle_Mod_1
R-HSA-449147	Signaling by Interleukins	$2.62 \cdot 10^{-4}$	$1.93 \cdot 10^{-3}$	Immune System	Muscle_Mod_1
R-HSA-2122947	NOTCH1 Intracellular Domain Regulates Transcription	$2.77 \cdot 10^{-4}$	$1.96 \cdot 10^{-3}$	Signal Transduction	Muscle_Mod_1
R-HSA-2559580	Oxidative Stress Induced Senescence	$2.77 \cdot 10^{-4}$	$1.96 \cdot 10^{-3}$	Cellular responses to external stimuli	Muscle_Mod_1
R-HSA-8953897	Cellular responses to external stimuli	$2.84 \cdot 10^{-4}$	$1.97 \cdot 10^{-3}$	Cellular responses to external stimuli	Muscle_Mod_1
R-HSA-2173793	Transcriptional activity of SMAD2/SMAD3:SMAD4 heterotrimer	$5.42 \cdot 10^{-4}$	$3.51 \cdot 10^{-3}$	Gene expression (Transcription)	Muscle_Mod_1
R-HSA-1640170	Cell Cycle	$6.07 \cdot 10^{-4}$	$3.51 \cdot 10^{-3}$	Cell Cycle	Muscle_Mod_1
R-HSA-6791226	Major pathway of rRNA processing in the nucleolus and cytosol	$6.28 \cdot 10^{-4}$	$3.51 \cdot 10^{-3}$	Metabolism of RNA	Muscle_Mod_1
R-HSA-6804758	Regulation of TP53 Activity through Acetylation	$6.28 \cdot 10^{-4}$	$3.51 \cdot 10^{-3}$	Gene expression (Transcription)	Muscle_Mod_1
R-HSA-72312	rRNA processing	$6.28 \cdot 10^{-4}$	$3.51 \cdot 10^{-3}$	Metabolism of RNA	Muscle_Mod_1
R-HSA-8868773	rRNA processing in the nucleus and cytosol	$6.28 \cdot 10^{-4}$	$3.51 \cdot 10^{-3}$	Metabolism of RNA	Muscle_Mod_1
R-HSA-8868773	rRNA processing in the nucleus and cytosol	$6.28 \cdot 10^{-4}$	$3.51 \cdot 10^{-3}$	Metabolism of RNA	Muscle_Mod_1
R-HSA-9022692	Regulation of MECP2 expression and activity	$6.28 \cdot 10^{-4}$	$3.51 \cdot 10^{-3}$	NA	Muscle_Mod_1
R-HSA-448424	Interleukin-17 signaling	$6.52 \cdot 10^{-4}$	$3.51 \cdot 10^{-3}$	Immune System	Muscle_Mod_1
R-HSA-450294	MAP kinase activation	$6.52 \cdot 10^{-4}$	$3.51 \cdot 10^{-3}$	Immune System	Muscle_Mod_1
R-HSA-2644602	Signaling by NOTCH1 PEST Domain Mutants in Cancer	$6.63 \cdot 10^{-4}$	$3.51 \cdot 10^{-3}$	Disease	Muscle_Mod_1
R-HSA-2644603	Signaling by NOTCH1 in Cancer	$6.63 \cdot 10^{-4}$	$3.51 \cdot 10^{-3}$	Disease	Muscle_Mod_1
R-HSA-2644606	Constitutive Signaling by NOTCH1 PEST Domain Mutants	$6.63 \cdot 10^{-4}$	$3.51 \cdot 10^{-3}$	Disease	Muscle_Mod_1
R-HSA-2894858	Signaling by NOTCH1 HD+PEST Domain Mutants in Cancer	$6.63 \cdot 10^{-4}$	$3.51 \cdot 10^{-3}$	Disease	Muscle_Mod_1
R-HSA-2894862	Constitutive Signaling by NOTCH1 HD+PEST Domain Mutants	$6.63 \cdot 10^{-4}$	$3.51 \cdot 10^{-3}$	Disease	Muscle_Mod_1
R-HSA-450282	MAPK targets/ Nuclear events mediated by MAP kinases	$6.63 \cdot 10^{-4}$	$3.51 \cdot 10^{-3}$	Immune System	Muscle_Mod_1
R-HSA-72766	Translation	$6.63 \cdot 10^{-4}$	$3.51 \cdot 10^{-3}$	Metabolism of proteins	Muscle_Mod_1
R-HSA-73894	DNA Repair	$7.23 \cdot 10^{-4}$	$3.77 \cdot 10^{-3}$	DNA Repair	Muscle_Mod_1
R-HSA-157118	Signaling by NOTCH	$7.38 \cdot 10^{-4}$	$3.80 \cdot 10^{-3}$	Signal Transduction	Muscle_Mod_1
R-HSA-5689880	Ub-specific processing proteases	$1.28 \cdot 10^{-3}$	$6.47 \cdot 10^{-3}$	Metabolism of proteins	Muscle_Mod_1
R-HSA-5688426	Deubiquitination	$1.65 \cdot 10^{-3}$	$8.26 \cdot 10^{-3}$	Metabolism of proteins	Muscle_Mod_1
R-HSA-1606322	ZBP1(DAI) mediated induction of type I IFNs	$1.73 \cdot 10^{-3}$	$8.31 \cdot 10^{-3}$	Immune System	Muscle_Mod_1
R-HSA-5693571	Nonhomologous End-Joining (NHEJ)	$1.73 \cdot 10^{-3}$	$8.31 \cdot 10^{-3}$	DNA Repair	Muscle_Mod_1
R-HSA-69541	Stabilization of p53	$1.73 \cdot 10^{-3}$	$8.31 \cdot 10^{-3}$	Cell Cycle	Muscle_Mod_1
R-HSA-2173795	Downregulation of SMAD2/3:SMAD4 transcriptional activity	$1.84 \cdot 10^{-3}$	$8.50 \cdot 10^{-3}$	Gene expression (Transcription)	Muscle_Mod_1
R-HSA-5607761	Dectin-1 mediated noncanonical NF-kB signaling	$1.84 \cdot 10^{-3}$	$8.50 \cdot 10^{-3}$	Immune System	Muscle_Mod_1
R-HSA-69239	Synthesis of DNA	$1.84 \cdot 10^{-3}$	$8.50 \cdot 10^{-3}$	DNA Replication	Muscle_Mod_1
R-HSA-69481	G2/M Checkpoints	$2.25 \cdot 10^{-3}$	$1.03 \cdot 10^{-2}$	Cell Cycle	Muscle_Mod_1
R-HSA-212436	Generic Transcription Pathway	$2.28 \cdot 10^{-3}$	$1.03 \cdot 10^{-2}$	Gene expression (Transcription)	Muscle_Mod_1
R-HSA-69620	Cell Cycle Checkpoints	$2.45 \cdot 10^{-3}$	$1.09 \cdot 10^{-2}$	Cell Cycle	Muscle_Mod_1
R-HSA-170834	Signaling by TGF-beta Receptor Complex	$2.73 \cdot 10^{-3}$	$1.20 \cdot 10^{-2}$	Signal Transduction	Muscle_Mod_1
R-HSA-5633007	Regulation of TP53 Activity	$2.77 \cdot 10^{-3}$	$1.20 \cdot 10^{-2}$	Gene expression (Transcription)	Muscle_Mod_1
R-HSA-8878159	Transcriptional regulation by RUNX3	$3.64 \cdot 10^{-3}$	$1.54 \cdot 10^{-2}$	Gene expression (Transcription)	Muscle_Mod_1
R-HSA-9006936	Signaling by TGF-beta family members	$3.64 \cdot 10^{-3}$	$1.54 \cdot 10^{-2}$	Signal Transduction	Muscle_Mod_1
R-HSA-1980143	Signaling by NOTCH1	$3.93 \cdot 10^{-3}$	$1.65 \cdot 10^{-2}$	Signal Transduction	Muscle_Mod_1
R-HSA-3214841	PKMTs methylate histone lysines	$3.98 \cdot 10^{-3}$	$1.65 \cdot 10^{-2}$	Chromatin organization	Muscle_Mod_1
R-HSA-3214858	RMTs methylate histone arginines	$4.39 \cdot 10^{-3}$	$1.79 \cdot 10^{-2}$	Chromatin organization	Muscle_Mod_1
R-HSA-446652	Interleukin-1 family signaling	$4.42 \cdot 10^{-3}$	$1.79 \cdot 10^{-2}$	Immune System	Muscle_Mod_1
R-HSA-5668541	TNFR2 non-canonical NF-kB pathway	$4.67 \cdot 10^{-3}$	$1.81 \cdot 10^{-2}$	Immune System	Muscle_Mod_1

R-HSA-5676590	NIK–¿noncanonical NF-kB signaling	$4.67{\cdot}10^{-3}$	$1.81{\cdot}10^{-2}$	Immune System	Muscle_Mod_1
R-HSA-5693565	Recruitment and ATM-mediated phosphorylation of repair and signaling proteins at DNA double strand breaks	$4.67{\cdot}10^{-3}$	$1.81{\cdot}10^{-2}$	DNA Repair	Muscle_Mod_1
R-HSA-5693606	DNA Double Strand Break Response	$4.67{\cdot}10^{-3}$	$1.81{\cdot}10^{-2}$	DNA Repair	Muscle_Mod_1
R-HSA-162909	Host Interactions of HIV factors	$6.62{\cdot}10^{-3}$	$2.48{\cdot}10^{-2}$	Disease	Muscle_Mod_1
R-HSA-168928	DDX58/IFIH1-mediated induction of interferon-alpha/beta	$6.62{\cdot}10^{-3}$	$2.48{\cdot}10^{-2}$	Immune System	Muscle_Mod_1
R-HSA-9010553	Regulation of expression of SLITs and ROBOs	$6.62{\cdot}10^{-3}$	$2.48{\cdot}10^{-2}$	Developmental Biology	Muscle_Mod_1
R-HSA-3700989	Transcriptional Regulation by TP53	$7.41{\cdot}10^{-3}$	$2.75{\cdot}10^{-2}$	Gene expression (Transcription)	Muscle_Mod_1
R-HSA-1169091	Activation of NF-kappaB in B cells	$7.78{\cdot}10^{-3}$	$2.86{\cdot}10^{-2}$	Immune System	Muscle_Mod_1
R-HSA-9020702	Interleukin-1 signaling	$7.89{\cdot}10^{-3}$	$2.87{\cdot}10^{-2}$	Immune System	Muscle_Mod_1
R-HSA-73857	RNA Polymerase II Transcription	$8.97{\cdot}10^{-3}$	$3.23{\cdot}10^{-2}$	Gene expression (Transcription)	Muscle_Mod_1
R-HSA-74160	Gene expression (Transcription)	$1.03{\cdot}10^{-2}$	$3.66{\cdot}10^{-2}$	Gene expression (Transcription)	Muscle_Mod_1
R-HSA-212165	Epigenetic regulation of gene expression	$1.04{\cdot}10^{-2}$	$3.66{\cdot}10^{-2}$	Gene expression (Transcription)	Muscle_Mod_1
R-HSA-174143	APC/C-mediated degradation of cell cycle proteins	$1.15{\cdot}10^{-2}$	$3.90{\cdot}10^{-2}$	Cell Cycle	Muscle_Mod_1
R-HSA-176408	Regulation of APC/C activators between G1/S and early anaphase	$1.15{\cdot}10^{-2}$	$3.90{\cdot}10^{-2}$	Cell Cycle	Muscle_Mod_1
R-HSA-453276	Regulation of mitotic cell cycle	$1.15{\cdot}10^{-2}$	$3.90{\cdot}10^{-2}$	Cell Cycle	Muscle_Mod_1
R-HSA-69052	Switching of origins to a post-replicative state	$1.15{\cdot}10^{-2}$	$3.90{\cdot}10^{-2}$	DNA Replication	Muscle_Mod_1
R-HSA-168249	Innate Immune System	$1.26{\cdot}10^{-2}$	$4.23{\cdot}10^{-2}$	Immune System	Muscle_Mod_1
R-HSA-69242	S Phase	$1.27{\cdot}10^{-2}$	$4.23{\cdot}10^{-2}$	Cell Cycle	Muscle_Mod_1
R-HSA-1234174	Regulation of Hypoxia-inducible Factor (HIF) by oxygen	$1.39{\cdot}10^{-2}$	$4.50{\cdot}10^{-2}$	Cellular responses to external stimuli	Muscle_Mod_1
R-HSA-2173796	SMAD2/SMAD3:SMAD4 heterotrimer regulates transcription	$1.39{\cdot}10^{-2}$	$4.50{\cdot}10^{-2}$	Gene expression (Transcription)	Muscle_Mod_1
R-HSA-2262749	Cellular response to hypoxia	$1.39{\cdot}10^{-2}$	$4.50{\cdot}10^{-2}$	Cellular responses to external stimuli	Muscle_Mod_1
R-HSA-383280	Nuclear Receptor transcription pathway	$9.78{\cdot}10^{-13}$	$1.10{\cdot}10^{-10}$	Gene expression (Transcription)	Muscle_Mod_2
R-HSA-1989781	PPARA activates gene expression	$7.28{\cdot}10^{-11}$	$4.11{\cdot}10^{-9}$	Metabolism	Muscle_Mod_2
R-HSA-400206	Regulation of lipid metabolism by Peroxisome proliferator-activated receptor alpha (PPARalpha)	$1.61{\cdot}10^{-10}$	$6.07{\cdot}10^{-9}$	Metabolism	Muscle_Mod_2
R-HSA-1368108	BMAL1:CLOCK,NPAS2 activates circadian gene expression	$3.79{\cdot}10^{-9}$	$1.07{\cdot}10^{-7}$	Circadian Clock	Muscle_Mod_2
R-HSA-400253	Circadian Clock	$1.30{\cdot}10^{-8}$	$2.95{\cdot}10^{-7}$	Circadian Clock	Muscle_Mod_2
R-HSA-556833	Metabolism of lipids	$1.67{\cdot}10^{-8}$	$3.15{\cdot}10^{-7}$	Metabolism	Muscle_Mod_2
R-HSA-1368082	RORA activates gene expression	$1.41{\cdot}10^{-6}$	$2.28{\cdot}10^{-5}$	Circadian Clock	Muscle_Mod_2
R-HSA-1368082	RORA activates gene expression	$1.41{\cdot}10^{-6}$	$2.28{\cdot}10^{-5}$	Circadian Clock	Muscle_Mod_2
R-HSA-4090294	SUMOylation of intracellular receptors	$3.58{\cdot}10^{-6}$	$5.06{\cdot}10^{-5}$	NA	Muscle_Mod_2
R-HSA-2990846	SUMOylation	$9.96{\cdot}10^{-6}$	$1.13{\cdot}10^{-4}$	Metabolism of proteins	Muscle_Mod_2
R-HSA-3108232	SUMO E3 ligases SUMOylate target proteins	$9.96{\cdot}10^{-6}$	$1.13{\cdot}10^{-4}$	Metabolism of proteins	Muscle_Mod_2
R-HSA-1430728	Metabolism	$1.48{\cdot}10^{-5}$	$1.52{\cdot}10^{-4}$	Metabolism	Muscle_Mod_2
R-HSA-2151201	Transcriptional activation of mitochondrial biogenesis	$2.66{\cdot}10^{-5}$	$2.50{\cdot}10^{-4}$	Organelle biogenesis and maintenance	Muscle_Mod_2
R-HSA-381340	Transcriptional regulation of white adipocyte differentiation	$5.55{\cdot}10^{-5}$	$4.83{\cdot}10^{-4}$	Developmental Biology	Muscle_Mod_2
R-HSA-9006931	Signaling by Nuclear Receptors	$7.67{\cdot}10^{-5}$	$6.19{\cdot}10^{-4}$	Signal Transduction	Muscle_Mod_2
R-HSA-1592230	Mitochondrial biogenesis	$8.90{\cdot}10^{-5}$	$6.71{\cdot}10^{-4}$	Organelle biogenesis and maintenance	Muscle_Mod_2
R-HSA-2426168	Activation of gene expression by SREBF (SREBP)	$1.69{\cdot}10^{-4}$	$1.19{\cdot}10^{-3}$	Metabolism	Muscle_Mod_2
R-HSA-1655829	Regulation of cholesterol biosynthesis by SREBP (SREBF)	$2.43{\cdot}10^{-4}$	$1.61{\cdot}10^{-3}$	Metabolism	Muscle_Mod_2
R-HSA-8957322	Metabolism of steroids	$8.22{\cdot}10^{-4}$	$5.16{\cdot}10^{-3}$	Metabolism	Muscle_Mod_2
R-HSA-1852241	Organelle biogenesis and maintenance	$1.04{\cdot}10^{-3}$	$6.21{\cdot}10^{-3}$	Organelle biogenesis and maintenance	Muscle_Mod_2
R-HSA-212436	Generic Transcription Pathway	$1.91{\cdot}10^{-3}$	$1.08{\cdot}10^{-2}$	Gene expression (Transcription)	Muscle_Mod_2
R-HSA-73857	RNA Polymerase II Transcription	$3.93{\cdot}10^{-3}$	$2.12{\cdot}10^{-2}$	Gene expression (Transcription)	Muscle_Mod_2
R-HSA-512988	Interleukin-3, Interleukin-5 and GM-CSF signaling	$8.78{\cdot}10^{-15}$	$1.08{\cdot}10^{-12}$	Immune System	Muscle_Mod_3
R-HSA-9006934	Signaling by Receptor Tyrosine Kinases	$3.46{\cdot}10^{-14}$	$2.13{\cdot}10^{-12}$	Signal Transduction	Muscle_Mod_3
R-HSA-6806834	Signaling by MET	$5.23{\cdot}10^{-12}$	$2.15{\cdot}10^{-10}$	Signal Transduction	Muscle_Mod_3
R-HSA-451927	Interleukin-2 family signaling	$1.33{\cdot}10^{-11}$	$3.27{\cdot}10^{-10}$	Immune System	Muscle_Mod_3

R-HSA-451927	Interleukin-2 family signaling	$1.33 \cdot 10^{-11}$	$3.27 \cdot 10^{-10}$	Immune System	Muscle_Mod_3
R-HSA-9006335	Signaling by Erythropoietin	$1.33 \cdot 10^{-11}$	$3.27 \cdot 10^{-10}$	NA	Muscle_Mod_3
R-HSA-5684996	MAPK1/MAPK3 signaling	$2.65 \cdot 10^{-11}$	$5.42 \cdot 10^{-10}$	Signal Transduction	Muscle_Mod_3
R-HSA-5683057	MAPK family signaling cascades	$6.26 \cdot 10^{-10}$	$1.10 \cdot 10^{-8}$	Signal Transduction	Muscle_Mod_3
R-HSA-1433557	Signaling by SCF-KIT	$8.20 \cdot 10^{-10}$	$1.26 \cdot 10^{-8}$	Signal Transduction	Muscle_Mod_3
R-HSA-1280215	Cytokine Signaling in Immune system	$3.45 \cdot 10^{-9}$	$4.53 \cdot 10^{-8}$	Immune System	Muscle_Mod_3
R-HSA-186763	Downstream signal transduction	$4.05 \cdot 10^{-9}$	$4.53 \cdot 10^{-8}$	Signal Transduction	Muscle_Mod_3
R-HSA-186797	Signaling by PDGF	$4.05 \cdot 10^{-9}$	$4.53 \cdot 10^{-8}$	Signal Transduction	Muscle_Mod_3
R-HSA-449147	Signaling by Interleukins	$2.86 \cdot 10^{-8}$	$2.93 \cdot 10^{-7}$	Immune System	Muscle_Mod_3
R-HSA-168256	Immune System	$8.24 \cdot 10^{-8}$	$7.60 \cdot 10^{-7}$	Immune System	Muscle_Mod_3
R-HSA-162582	Signal Transduction	$8.65 \cdot 10^{-8}$	$7.60 \cdot 10^{-7}$	Signal Transduction	Muscle_Mod_3
R-HSA-177929	Signaling by EGFR	$1.06 \cdot 10^{-7}$	$8.66 \cdot 10^{-7}$	Signal Transduction	Muscle_Mod_3
R-HSA-2404192	Signaling by Type 1 Insulin-like Growth Factor 1 Receptor (IGF1R)	$2.27 \cdot 10^{-7}$	$1.65 \cdot 10^{-6}$	Signal Transduction	Muscle_Mod_3
R-HSA-2428924	IGF1R signaling cascade	$2.27 \cdot 10^{-7}$	$1.65 \cdot 10^{-6}$	Signal Transduction	Muscle_Mod_3
R-HSA-76002	Platelet activation, signaling and aggregation	$6.43 \cdot 10^{-7}$	$4.39 \cdot 10^{-6}$	Hemostasis	Muscle_Mod_3
R-HSA-74751	Insulin receptor signalling cascade	$8.10 \cdot 10^{-7}$	$4.98 \cdot 10^{-6}$	Signal Transduction	Muscle_Mod_3
R-HSA-74752	Signaling by Insulin receptor	$8.10 \cdot 10^{-7}$	$4.98 \cdot 10^{-6}$	Signal Transduction	Muscle_Mod_3
R-HSA-5673001	RAF/MAP kinase cascade	$1.02 \cdot 10^{-6}$	$5.96 \cdot 10^{-6}$	Signal Transduction	Muscle_Mod_3
R-HSA-109582	Hemostasis	$1.78 \cdot 10^{-6}$	$9.97 \cdot 10^{-6}$	Hemostasis	Muscle_Mod_3
R-HSA-2428928	IRS-related events triggered by IGF1R	$4.84 \cdot 10^{-6}$	$2.59 \cdot 10^{-5}$	Signal Transduction	Muscle_Mod_3
R-HSA-2219530	Constitutive Signaling by Aberrant PI3K in Cancer	$8.68 \cdot 10^{-6}$	$4.45 \cdot 10^{-5}$	Disease	Muscle_Mod_3
R-HSA-5654741	Signaling by FGFR3	$1.46 \cdot 10^{-5}$	$6.89 \cdot 10^{-5}$	Signal Transduction	Muscle_Mod_3
R-HSA-5655302	Signaling by FGFR1 in disease	$1.46 \cdot 10^{-5}$	$6.89 \cdot 10^{-5}$	Disease	Muscle_Mod_3
R-HSA-5654743	Signaling by FGFR4	$2.32 \cdot 10^{-5}$	$1.06 \cdot 10^{-4}$	Signal Transduction	Muscle_Mod_3
R-HSA-5663202	Diseases of signal transduction	$3.49 \cdot 10^{-5}$	$1.45 \cdot 10^{-4}$	Disease	Muscle_Mod_3
R-HSA-5654736	Signaling by FGFR1	$3.53 \cdot 10^{-5}$	$1.45 \cdot 10^{-4}$	Signal Transduction	Muscle_Mod_3
R-HSA-877300	Interferon gamma signaling	$3.53 \cdot 10^{-5}$	$1.45 \cdot 10^{-4}$	Immune System	Muscle_Mod_3
R-HSA-422475	Axon guidance	$4.92 \cdot 10^{-5}$	$1.93 \cdot 10^{-4}$	Developmental Biology	Muscle_Mod_3
R-HSA-913531	Interferon Signaling	$5.03 \cdot 10^{-5}$	$1.93 \cdot 10^{-4}$	Immune System	Muscle_Mod_3
R-HSA-1227986	Signaling by ERBB2	$5.18 \cdot 10^{-5}$	$1.93 \cdot 10^{-4}$	Signal Transduction	Muscle_Mod_3
R-HSA-1236394	Signaling by ERBB4	$7.37 \cdot 10^{-5}$	$2.67 \cdot 10^{-4}$	Signal Transduction	Muscle_Mod_3
R-HSA-1226099	Signaling by FGFR in disease	$1.02 \cdot 10^{-4}$	$3.59 \cdot 10^{-4}$	Disease	Muscle_Mod_3
R-HSA-6811558	PI5P, PP2A and IER3 Regulate PI3K/AKT Signaling	$1.38 \cdot 10^{-4}$	$4.73 \cdot 10^{-4}$	Signal Transduction	Muscle_Mod_3
R-HSA-909733	Interferon alpha/beta signaling	$1.47 \cdot 10^{-4}$	$4.89 \cdot 10^{-4}$	Immune System	Muscle_Mod_3
R-HSA-166520	Signaling by NTRKs	$1.59 \cdot 10^{-4}$	$5.16 \cdot 10^{-4}$	Signal Transduction	Muscle_Mod_3
R-HSA-1500931	Cell-Cell communication	$2.26 \cdot 10^{-4}$	$6.95 \cdot 10^{-4}$	Cell-Cell communication	Muscle_Mod_3
R-HSA-388841	Costimulation by the CD28 family	$2.26 \cdot 10^{-4}$	$6.95 \cdot 10^{-4}$	Immune System	Muscle_Mod_3
R-HSA-199418	Negative regulation of the PI3K/AKT network	$2.40 \cdot 10^{-4}$	$7.02 \cdot 10^{-4}$	Signal Transduction	Muscle_Mod_3
R-HSA-5654738	Signaling by FGFR2	$2.40 \cdot 10^{-4}$	$7.02 \cdot 10^{-4}$	Signal Transduction	Muscle_Mod_3
R-HSA-190236	Signaling by FGFR	$3.08 \cdot 10^{-4}$	$8.81 \cdot 10^{-4}$	Signal Transduction	Muscle_Mod_3
R-HSA-187687	Signalling to ERKs	$3.32 \cdot 10^{-4}$	$9.28 \cdot 10^{-4}$	Signal Transduction	Muscle_Mod_3
R-HSA-6785807	Interleukin-4 and Interleukin-13 signaling	$6.03 \cdot 10^{-4}$	$1.65 \cdot 10^{-3}$	Immune System	Muscle_Mod_3
R-HSA-447115	Interleukin-12 family signaling	$6.43 \cdot 10^{-4}$	$1.72 \cdot 10^{-3}$	Immune System	Muscle_Mod_3
R-HSA-2219528	PI3K/AKT Signaling in Cancer	$7.37 \cdot 10^{-4}$	$1.93 \cdot 10^{-3}$	Disease	Muscle_Mod_3
R-HSA-187037	Signaling by NTRK1 (TRKA)	$8.93 \cdot 10^{-4}$	$2.29 \cdot 10^{-3}$	Signal Transduction	Muscle_Mod_3
R-HSA-5655253	Signaling by FGFR2 in disease	$3.96 \cdot 10^{-3}$	$9.73 \cdot 10^{-3}$	Disease	Muscle_Mod_3
R-HSA-9020591	Interleukin-12 signaling	$3.96 \cdot 10^{-3}$	$9.73 \cdot 10^{-3}$	Immune System	Muscle_Mod_3
R-HSA-2029480	Fcgamma receptor (FCGR) dependent phagocytosis	$5.16 \cdot 10^{-3}$	$1.22 \cdot 10^{-2}$	Immune System	Muscle_Mod_3
R-HSA-2871796	FCERI mediated MAPK activation	$5.16 \cdot 10^{-3}$	$1.22 \cdot 10^{-2}$	Immune System	Muscle_Mod_3
R-HSA-375165	NCAM signaling for neurite out-growth	$6.57 \cdot 10^{-3}$	$1.50 \cdot 10^{-2}$	Developmental Biology	Muscle_Mod_3
R-HSA-416476	G alpha (q) signalling events	$6.57 \cdot 10^{-3}$	$1.50 \cdot 10^{-2}$	Signal Transduction	Muscle_Mod_3
R-HSA-1643685	Disease	$6.78 \cdot 10^{-3}$	$1.52 \cdot 10^{-2}$	Disease	Muscle_Mod_3
R-HSA-194138	Signaling by VEGF	$2.26 \cdot 10^{-2}$	$4.88 \cdot 10^{-2}$	Signal Transduction	Muscle_Mod_3
R-HSA-4420097	VEGFA-VEGFR2 Pathway	$2.26 \cdot 10^{-2}$	$4.88 \cdot 10^{-2}$	Signal Transduction	Muscle_Mod_3
R-HSA-2454202	Fc epsilon receptor (FCERI) signaling	$2.34 \cdot 10^{-2}$	$4.97 \cdot 10^{-2}$	Immune System	Muscle_Mod_3

Table E.13: Repurposing candidates obtained from the analysis of the adipose network. The columns SE indicates if the drug has an undesired side effect related to MetSyn (DrugCentral platform [262]), while the column OT indicates if the target has been/is under investigation for theurapeutic interventions related to MetSyn (OpenTargets platform [133])

DRUG ID	DRUG NAME	TARGET	ACTION	SCORE	MODULE	SE	OT
DB00459	Acitretin	RARA	agonist	1.678	Adipose_Mod_2	1	0
DB00459	Acitretin	RXRA	agonist	1.721	Adipose_Mod_2	1	0
DB00459	Acitretin	RXRB	agonist	1.673	Adipose_Mod_2	1	0
DB00459	Acitretin	RXRG	agonist	1.652	Adipose_Mod_2	1	1
DB00210	Adapalene	RARG	agonist	1.662	Adipose_Mod_2	0	0
DB00210	Adapalene	RXRB	agonist	1.692	Adipose_Mod_2	0	0
DB00210	Adapalene	RXRG	agonist	1.680	Adipose_Mod_2	0	1
DB00523	Alitretinoin	RARG	agonist	1.666	Adipose_Mod_2	0	0
DB00523	Alitretinoin	RXRG	agonist	1.704	Adipose_Mod_2	0	1
DB00995	Auranofin	IKBKB	inhibitor	1.797	Adipose_Mod_1	NA	1
DB00136	Calcitriol	VDR	antagonist	1.666	Adipose_Mod_2	0	1
DB04209	Dequalinium	XIAP	antagonist, inhibitor	1.791	Adipose_Mod_1	NA	0
DB03756	Doconexent	RXRA	activator	1.710	Adipose_Mod_2	NA	0
DB03756	Doconexent	RXRB	activator	1.666	Adipose_Mod_2	NA	0
DB03756	Doconexent	PPARA	ligand	1.716	Adipose_Mod_2	NA	1
DB03756	Doconexent	PPARG	ligand	1.709	Adipose_Mod_2	NA	1
DB03756	Doconexent	RXRG	activator	1.665	Adipose_Mod_2	NA	1
DB00530	Erlotinib	EGFR	anotagonist	1.802	Adipose_Mod_1	0	1
DB01039	Fenofibrate	PPARA	agonist	1.741	Adipose_Mod_2	0	1
DB00317	Gefitinib	EGFR	antagonist	1.668	Adipose_Mod_1	0	1
DB09053	Ibrutinib	BTK	inhibition	1.615	Adipose_Mod_1	0	0
DB00451	Levothyroxine	THRA	agonist	1.667	Adipose_Mod_2	1	1
DB00279	Liothyronine	THRA	agonist	1.687	Adipose_Mod_2	1	1
DB00279	Liothyronine	THRB	agonist	1.706	Adipose_Mod_2	1	1
DB00253	Medrysone	NR3C1	agonist	1.659	Adipose_Mod_2	NA	1
DB00244	Mesalazine	PPARG	agonist	1.722	Adipose_Mod_2	0	1
DB00834	Mifepristone	NR3C1	antagonist	1.708	Adipose_Mod_2	0	1
DB00834	Mifepristone	PGR	antagonist	1.714	Adipose_Mod_2	0	1
DB11828	Neratinib	EGFR	inhibitor	1.811	Adipose_Mod_1	NA	1
DB04868	Nilotinib	ABL1	inhibitor	1.726	Adipose_Mod_1	1	0

DB00665	Nilutamide	AR	antagonist	1.664	Adipose_Mod_1	0	1
DB06603	Panobinostat	HDAC6	inhibitor	1.508	Adipose_Mod_1	NA	1
DB01132	Pioglitazone	PPARG	agonist	1.698	Adipose_Mod_2	0	1
DB00396	Progesterone	ESR1	agonist	1.639	Adipose_Mod_2	1	1
DB12332	Rucaparib	PARP1	antagonist	1.668	Adipose_Mod_2	NA	0
DB08877	Ruxolitinib	JAK1	inhibitor	1.772	Adipose_Mod_1	NA	0
DB00795	Sulfasalazine	PPARG	agonist	1.684	Adipose_Mod_2	0	1
DB04942	Tamibarotene	RARA	agonist	1.722	Adipose_Mod_2	NA	0
DB00755	Tretinoin	RARG	agonist	1.695	Adipose_Mod_2	1	0
DB00755	Tretinoin	RXRB	agonist	1.668	Adipose_Mod_2	1	0
DB00755	Tretinoin	RXRG	agonist	1.656	Adipose_Mod_2	1	1
DB02546	Vorinostat	HDAC3	inhibitor	1.671	Adipose_Mod_2	1	0

Table E.14: Repurposing candidates obtained from the analysis of the liver network.The columns SE indicates if the drug has an undesired side effect related to MetSyn (DrugCentral platform [262]), while the column OT indicates if the target has been/is under investigation for theurapeutic interventions related to MetSyn (OpenTargets platform [133])

DRUG ID	DRUG NAME	TARGET	ACTION	SCORE	MODULE	SE	OT
DB00459	Acitretin	RXRA	agonist	1.671	Liver_Mod_2	1	0
DB00459	Acitretin	RARG	agonist	1.407	Liver_Mod_3	1	0
DB00210	Adapalene	RXRB	agonist	1.687	Liver_Mod_2	0	0
DB00523	Alitretinoin	RXRG	agonist	1.390	Liver_Mod_3	0	1
DB01014	Balsalazide	PPARG	agonist	1.642	Liver_Mod_2	0	1
DB01393	Bezafibrate	PPARD	agonist	1.299	Liver_Mod_1	0	1
DB01393	Bezafibrate	PPARG	agonist	1.636	Liver_Mod_2	0	1
DB01128	Bicalutamide	AR	antagonist	1.314	Liver_Mod_1	1	1
DB00636	Clofibrate	PPARA	agonist	1.655	Liver_Mod_2	NA	1
DB03756	Doconexent	PPARA	ligand	1.325	Liver_Mod_1	NA	1
DB03756	Doconexent	PPARA	ligand	1.739	Liver_Mod_2	NA	1
DB03756	Doconexent	PPARG	ligand	1.668	Liver_Mod_2	NA	1
DB03756	Doconexent	RXRA	activator	1.705	Liver_Mod_2	NA	0
DB03756	Doconexent	RXRA	activator	1.495	Liver_Mod_3	NA	0
DB03756	Doconexent	PPARG	ligand	1.486	Liver_Mod_3	NA	1
DB03756	Doconexent	RXRB	activator	1.682	Liver_Mod_2	NA	0
DB03756	Doconexent	RXRB	activator	1.537	Liver_Mod_3	NA	0
DB00530	Erlotinib	NR1I2	agonist	1.668	Liver_Mod_2	0	0
DB00783	Estradiol	ESR1	agonist	1.679	Liver_Mod_2	1	1
DB00655	Estrone	ESR1	agonist	1.671	Liver_Mod_2	1	1
DB01039	Fenofibrate	PPARA	agonist	1.681	Liver_Mod_2	0	1
DB01039	Fenofibrate	PPARA	agonist	1.614	Liver_Mod_3	0	1
DB00687	Fludrocortisone	NR3C2	agonist	1.647	Liver_Mod_2	NA	1
DB00324	Fluorometholone	NR3C1	agonist	1.288	Liver_Mod_1	1	1
DB01241	Gemfibrozil	PPARA	agonist	1.694	Liver_Mod_2	0	1
DB00159	Icosapent	PPARG	agonist	1.312	Liver_Mod_1	NA	1
DB00159	Icosapent	PPARG	agonist	1.596	Liver_Mod_2	NA	1
DB00451	Levothyroxine	THRB	agonist	1.300	Liver_Mod_1	1	1
DB00451	Levothyroxine	THRA	agonist	1.623	Liver_Mod_2	1	1
DB00451	Levothyroxine	THRB	agonist	1.633	Liver_Mod_2	1	1
DB00451	Levothyroxine	THRA	agonist	1.430	Liver_Mod_3	1	1

DB00451	Levothyroxine	THRB	agonist	1.542	Liver_Mod_3	1	1
DB00279	Liothyronine	THRA	agonist	1.658	Liver_Mod_2	1	1
DB00279	Liothyronine	THRB	agonist	1.696	Liver_Mod_2	1	1
DB00279	Liothyronine	THRA	agonist	1.435	Liver_Mod_3	1	1
DB00253	Medrysone	NR3C1	agonist	1.634	Liver_Mod_2	NA	1
DB00244	Mesalazine	PPARG	agonist	1.330	Liver_Mod_1	0	1
DB00244	Mesalazine	PPARG	agonist	1.686	Liver_Mod_2	0	1
DB00244	Mesalazine	PPARG	agonist	1.512	Liver_Mod_3	0	1
DB00834	Mifepristone	NR3C1	antagonist	1.645	Liver_Mod_2	0	1
DB00834	Mifepristone	NR3C1	antagonist	1.459	Liver_Mod_3	0	1
DB09079	Nintedanib	FGFR3	inhibitor	1.447	Liver_Mod_2	NA	0
DB01132	Pioglitazone	PPARG	agonist	1.312	Liver_Mod_1	0	1
DB01132	Pioglitazone	PPARG	agonist	1.619	Liver_Mod_2	0	1
DB00481	Raloxifene	ESR1	agonist	1.649	Liver_Mod_2	1	1
DB00421	Spironolactone	NR3C2	antagonist	1.644	Liver_Mod_2	1	1
DB00421	Spironolactone	NR3C2	antagonist	1.476	Liver_Mod_3	1	1
DB00795	Sulfasalazine	PPARG	agonist	1.672	Liver_Mod_2	0	1
DB04942	Tamibarotene	RARA	agonist	1.667	Liver_Mod_2	NA	0
DB00966	Telmisartan	PPARG	partial agonist	1.478	Liver_Mod_3	1	1
DB06287	Temsirolimus	MTOR	inhibitor	1.581	Liver_Mod_2	1	0
DB00374	Treprostinil	PPARD	agonist	1.321	Liver_Mod_1	0	1
DB00755	Tretinoin	RARG	agonist	1.658	Liver_Mod_2	1	0
DB00755	Tretinoin	RARG	agonist	1.430	Liver_Mod_3	1	0

Table E.15: Repurposing candidates obtained from the analysis of the muscle network.The columns SE indicates if the drug has an undesired side effect related to MetSyn (DrugCentral platform [262]), while the column OT indicates if the target has been/is under investigation for theurapeutic interventions related to MetSyn (OpenTargets platform [133])

DRUG ID	DRUG NAME	TARGET	ACTION	SCORE	MODULE	SE	OT
DB00459	Acitretin	RARA	agonist	1.814	Muscle_Mod_1	1	0
DB00459	Acitretin	RARA	agonist	1.686	Muscle_Mod_2	1	0
DB00459	Acitretin	RXRA	agonist	1.718	Muscle_Mod_2	1	0
DB00459	Acitretin	RXRG	agonist	1.658	Muscle_Mod_2	1	1
DB00210	Adapalene	RXRG	agonist	1.826	Muscle_Mod_1	0	1
DB00210	Adapalene	RXRG	agonist	1.682	Muscle_Mod_2	0	1
DB00210	Adapalene	RXRB	agonist	1.676	Muscle_Mod_2	0	0
DB00210	Adapalene	RARG	agonist	1.668	Muscle_Mod_2	0	0
DB08916	Afatinib	ERBB2	inhibitor	1.595	Muscle_Mod_3	NA	1
DB00523	Alitretinoin	RARG	agonist	1.776	Muscle_Mod_1	0	0
DB00523	Alitretinoin	RXRG	agonist	1.836	Muscle_Mod_1	0	1
DB00523	Alitretinoin	RARG	agonist	1.664	Muscle_Mod_2	0	0
DB00523	Alitretinoin	RXRG	agonist	1.676	Muscle_Mod_2	0	1
DB00523	Alitretinoin	RXRB	agonist	1.662	Muscle_Mod_2	0	0
DB00523	Alitretinoin	RARA	agonist	1.681	Muscle_Mod_2	0	0
DB00288	Amcinonide	NR3C1	agonist	1.737	Muscle_Mod_1	1	1
DB00995	Auranofin	IKBKB	inhibitor	1.682	Muscle_Mod_3	NA	1
DB05015	Belinostat	HDAC1	inhibitor	1.754	Muscle_Mod_1	NA	0
DB05015	Belinostat	HDAC2	inhibitor	1.761	Muscle_Mod_1	NA	0
DB05015	Belinostat	HDAC4	inhibitor	1.803	Muscle_Mod_1	NA	0
DB06616	Bosutinib	ABL1	inhibitor	1.609	Muscle_Mod_3	NA	0
DB01222	Budesonide	NR3C1	antagonist	1.649	Muscle_Mod_2	1	1
DB00636	Clofibrate	PPARA	agonist	1.804	Muscle_Mod_1	NA	1
DB00636	Clofibrate	PPARA	agonist	1.658	Muscle_Mod_2	NA	1
DB08865	Crizotinib	MET	inhibitor	1.788	Muscle_Mod_1	NA	0
DB01406	Danazol	ESR1	agonist	1.698	Muscle_Mod_3	1	1
DB03756	Doconexent	PPARA	ligand	1.688	Muscle_Mod_2	NA	1
DB03756	Doconexent	RXRA	activator	1.713	Muscle_Mod_2	NA	0
DB03756	Doconexent	RXRG	activator	1.675	Muscle_Mod_2	NA	1
DB08899	Enzalutamide	AR	inhibitor	1.708	Muscle_Mod_1	NA	1
DB08899	Enzalutamide	AR	inhibitor	1.674	Muscle_Mod_2	NA	1

DB00530	Erlotinib	EGFR	antagonist	1.714	Muscle_Mod_3	0	1
DB00783	Estradiol	ESR1	agonist	1.825	Muscle_Mod_1	1	1
DB01590	Everolimus	MTOR	inhibitor	1.798	Muscle_Mod_1	1	0
DB01039	Fenofibrate	PPARA	agonist	1.717	Muscle_Mod_2	0	1
DB01288	Fenoterol	ADRB2	agonist	1.631	Muscle_Mod_3	NA	1
DB00180	Flunisolide	NR3C1	agonist	1.797	Muscle_Mod_1	0	1
DB00180	Flunisolide	NR3C1	agonist	1.610	Muscle_Mod_3	0	1
DB00499	Flutamide	AR	antagonist	1.636	Muscle_Mod_2	0	1
DB00947	Fulvestrant	ESR1	antagonist	1.811	Muscle_Mod_1	0	1
DB00317	Gefitinib	EGFR	antagonist	1.627	Muscle_Mod_3	0	1
DB01241	Gemfibrozil	PPARA	agonist	1.647	Muscle_Mod_2	0	1
DB01259	Lapatinib	ERBB2	antagonist	1.818	Muscle_Mod_1	0	0
DB01259	Lapatinib	EGFR	antagonist	1.642	Muscle_Mod_3	0	1
DB00451	Levothyroxine	THRA	agonist	1.653	Muscle_Mod_2	1	1
DB00279	Liothyronine	THRA	agonist	1.667	Muscle_Mod_2	1	1
DB00253	Medrysone	NR3C1	agonist	1.730	Muscle_Mod_1	NA	1
DB00253	Medrysone	NR3C1	agonist	1.632	Muscle_Mod_2	NA	1
DB01357	Mestranol	ESR1	agonist	1.652	Muscle_Mod_3	1	1
DB00834	Mifepristone	NR3C1	antagonist	1.637	Muscle_Mod_2	0	1
DB11828	Neratinib	EGFR	inhibitor	1.676	Muscle_Mod_3	NA	1
DB04868	Nilotinib	ABL1	inhibitor	1.661	Muscle_Mod_3	1	0
DB00621	Oxandrolone	AR	agonist	1.815	Muscle_Mod_1	1	1
DB06412	Oxymetholone	AR	agonist, activator	1.800	Muscle_Mod_1	1	1
DB06603	Panobinostat	HDAC3	inhibitor	1.611	Muscle_Mod_2	NA	0
DB00860	Prednisolone	NR3C1	agonist	1.629	Muscle_Mod_3	1	1
DB00635	Prednisone	NR3C1	agonist	1.594	Muscle_Mod_3	1	1
DB00481	Raloxifene	ESR2	agonist	1.620	Muscle_Mod_2	1	1
DB00867	Ritodrine	ADRB2	agonist	1.599	Muscle_Mod_3	1	1
DB08877	Ruxolitinib	JAK2	inhibitor	1.714	Muscle_Mod_3	NA	0
DB08877	Ruxolitinib	JAK1	inhibitor	1.674	Muscle_Mod_3	NA	0
DB01001	Salbutamol	ADRB2	agonist	1.599	Muscle_Mod_3	1	1
DB04942	Tamibarotene	RARA	agonist	1.671	Muscle_Mod_2	NA	0
DB00966	Telmisartan	PPARG	partial agonist	1.733	Muscle_Mod_2	1	1
DB06287	Temsirolimus	MTOR	inhibitor	1.693	Muscle_Mod_1	1	0
DB06287	Temsirolimus	MTOR	inhibitor	1.613	Muscle_Mod_3	1	0
DB00624	Testosterone	AR	agonist	1.793	Muscle_Mod_1	1	1
DB00277	Theophylline	HDAC2	activator	1.721	Muscle_Mod_1	1	0
DB00373	Timolol	ADRB2	antagonist	1.619	Muscle_Mod_3	1	1

DB08895	Tofacitinib	JAK2	antagonist, inhibitor	1.692	Muscle_Mod_3	NA	0
DB08895	Tofacitinib	JAK1	inhibitor	1.710	Muscle_Mod_3	NA	0
DB00755	Tretinoin	RARG	agonist	1.665	Muscle_Mod_2	1	0
DB00755	Tretinoin	RXRG	agonist	1.702	Muscle_Mod_2	1	1
DB02546	Vorinostat	HDAC1	inhibitor	1.793	Muscle_Mod_1	1	0

Table E.16: Side effects of the repurposing candidate drugs extracted from DrugCentral platform.

DRUG NAME	SIDE EFFECTS (SNOMED)	METSYN RELATED
Acitretin	Acute nephropathy (58574008), Acute pancreatitis (197456007), Alcoholism (7200002), Benign intracranial hypertension (68267002), Breastfeeding (mother) (413712001), Cardiovascular event risk (395112001), Depressive disorder (35489007), Diabetes mellitus (73211009), Drug-induced hepatitis (235876009), Hepatic failure (59927004), Hypercholesterolemia (13644009), Hypertriglyceridemia (302870006), Hypertrophy of bone (203514008), Hypoalphalipoproteinemia (190785000), Liver function tests abnormal (166603001), Night blindness (65194006), Obesity (414916001), Pregnancy, function (289908002), Premature epiphyseal closure (89493005), Psychiatric Disturbance (NA), Visual impairment (397540003)	1
Adapalene	Erythroderma (399992009), Inflammatory dermatosis (703938007), Photosensitivity (90128006)	0
Alitretinoin	Breastfeeding (mother) (413712001), Mycosis fungoides (118618005), Pregnancy, function (289908002)	0
Amcinonide	Atrophoderma (399979006), Bilateral cataracts (95722004), Diabetes mellitus type 1 (46635009), Diabetes mellitus type 2 (44054006), Glaucoma (23986001), Peripheral vascular disease (400047006), Tuberculosis of skin (66986005)	1
Balsalazide	Kidney disease (90708001), Pyloric obstruction (244815007)	0
Bezafibrate	End stage renal disease (46177005)	0
Bicalutamide	Anemia (271737000), Diabetes mellitus (73211009), Disease of liver (235856003), Drug-induced hepatitis (235876009), Hyperglycemia (80394007), Interstitial pneumonia (64667001), Liver function tests abnormal (166603001), Pregnancy, function (289908002)	1
Budesonide	Acute tuberculosis (25629007), Adrenal cortical hypofunction (386584007), Anastomosis of intestine (235407009), Arginase deficiency (23501004), Arginosuccinate Lyase Deficiency (NA), Avascular necrosis of bone (397758007), Bacterial infectious disease (87628006), Bilateral cataracts (95722004), Cerebral malaria (53622003), Cerebral trauma (275382005), Chronic heart failure (48447003), Cirrhosis of liver (19943007), Citrullinemia (398680004), Congenital hyperammonemia, type I (62522004), Diabetes mellitus (73211009), Disease caused by parasite (17322007), Disease of liver (235856003), Disorder of muscle (129565002), Diverticulitis of gastrointestinal tract (271366000), Edema (267038008), Epistaxis (12441001), Exposure to varicella (444453009), Gastritis (4556007), Glaucoma (23986001), Herpes simplex (88594005), Hypercholesterolemia (13644009), Hypercortisolism (47270006), Hyperglycemia (80394007), Hypertensive disorder (38341003), Hypokalemia (43339004), Hypopituitarism (74728003), Hypothyroidism (40930008), Immunosuppression (38013005), Inactive tuberculosis (11999007), Infection by Strongyloides (1214006), Infectious disease (40733004), Measles (14189004), Muscle atrophy (88092000), Mycosis (3218000), Nasal Candidiasis (NA), Nasal Septal Ulcers (NA), Nasal Trauma (NA), Ocular hypertension (4210003), Open-angle glaucoma (84494001), Operation on nose (88733004), Ophthalmic herpes simplex (186542001), Ornithine Carbamyltransferase Deficiency (NA), Oropharyngeal Candidiasis (NA), Osteopenia (312894000), Osteoporosis (64859006), Pathological fracture (268029009), Peptic ulcer (13200003), Perforation of nasal septum (80142000), Psychotic disorder (69322001), Pulmonary tuberculosis (154283005), Seizure disorder (128613002), Traumatic rupture of tendon (415749005), Tuberculosis (56717001), Uncontrolled Bacterial Infections (NA), Untreated Fungal Infection (NA), Varicella-zoster virus infection (309465005), Viral disease (34014006)	1
Calcitriol	Arteriosclerosis obliterans (361133006), Hypercalcemia (66931009), Hypercalcemia associated with Sarcoidosis (NA), Hyperphosphatemia (20165001), Hypervitaminosis D (27712000), Kidney disease (90708001), Kidney stone (95570007), Sarcoidosis (31541009)	0
Danazol	Angina pectoris (194828000), Breastfeeding (mother) (413712001), Carcinoma of female breast (447782002), Chronic heart failure (48447003), Disease of liver (235856003), Epilepsy (84757009), Intermenstrual bleeding - irregular (64996003), Kidney disease (90708001), Migraine (37796009), Mixed hyperlipidemia (267434003), Porphyria (418470004), Pregnancy, function (289908002), Thromboembolic disorder (371039008)	1
Erlotinib	Adult respiratory distress syndrome (67782005), Bleeding (131148009), Blood coagulation disorder (64779008), Breastfeeding (mother) (413712001), Cerebrovascular accident (230690007), Corneal ulcer (91514001), Dehydration (34095006), Disease of liver (235856003), Fibrosis of lung (51615001), Gastrointestinal perforation (51875005), Hemolytic anemia (61261009), Hepatorenal syndrome (51292008), Hyperbilirubinemia (14783006), Hypokalemia (43339004), Interstitial lung disease (233703007), Interstitial pneumonia (64667001), Kidney disease (90708001), Liver function tests abnormal (166603001), Malignant neoplasm of liver (93870000), Myocardial infarction (22298006), Myocardial ischemia (414795007), Obstructive Bronchiolitis (NA), Pregnancy, function (289908002), Smokes tobacco daily (449868002), Thrombocytopenic disorder (302215000), Thrombotic thrombocytopenic purpura (78129009)	0

Estradiol	Acute nephropathy (58574008), Acute Thromboembolic Stroke (NA), Adrenal cortical hypofunction (386584007), Alcoholism (7200002), Angina pectoris (194828000), Anorexia nervosa (56882008), Asthma (195967001), Bed-ridden (160685001), Benign Hepatic Cell Adenoma (NA), Benign mammary dysplasia (57993004), Benign prostatic hyperplasia (266569009), Body fluid retention (43498006), Breast Carcinoma in Males (NA), Breast lump (89164003), Breastfeeding (mother) (413712001), Carcinoma of female breast (447782002), Cardiovascular event risk (395112001), Cerebrovascular accident (230690007), Cerebrovascular disease (62914000), Chloasma (36209000), Cholestasis of pregnancy (235888006), Chorea (271700006), Chronic heart failure (48447003), Chronic lung disease (413839001), Deep venous thrombosis (128053003), Dementia (52448006), Depressive disorder (35489007), Diabetes mellitus (73211009), Diabetes with Vascular Disease Complication (NA), Diplopia (24982008), Disease of liver (235856003), Disorder of cardiovascular system (49601007), Disorder of coronary artery (414024009), Disorder of gallbladder (39621005), Edema (267038008), Endometrial carcinoma (254878006), Endometriosis (129103003), Epilepsy (84757009), Estrogen receptor positive tumor (416053008), Family history of malignant neoplasm of breast (429740004), Fibroadenosis of breast (23260002), Functional visual loss (313165001), Gynecomastia (4754008), Heart disease (56265001), Heart failure (84114007), Heart valve disorder (368009), Hepatic failure (59927004), Hepatic porphyria (55056006), Humoral hypercalcemia of malignancy (47709007), Hypercalcemia (66931009), Hypercholesterolemia (13644009), Hyperglycemia (80394007), Hyperkalemia (14140009), Hyperlipidemia (55822004), Hyperlipoproteinemia (3744001), Hypertensive disorder (38341003), Hypertensive urgency (443482000), Hypertriglyceridemia (302870006), Hypocalcemia (5291005), Hypothyroidism (40930008), Intermenstrual bleeding - irregular (64996003), Jaundice (18165001), Kidney disease (90708001), Liver function tests abnormal (166603001), Major Surgery with Prolonged Post-Operative Immobilization (NA), Malignant neoplasm of liver (93870000), Malignant tumor of cervix (363354003), Malignant tumor of ovary (363443007), Mammography abnormal (168750009), Metabolic syndrome X (237602007), Migraine (37796009), Myocardial infarction (22298006), Neoplasm of female genital organ (126907002), Neoplasm of prostate (126906006), Non-Q wave myocardial infarction (314207007), Nonspecific Abnormal Papanicolaou Smear of Cervix (NA), Obesity (414916001), Obstructive hyperbilirubinemia (59848001), Optic disc edema (423341008), Osteopenia (312894000), Osteoporosis (64859006), Porphyria (418470004), Pregnancy, function (289908002), Pulmonary thromboembolism (233935004), Resistance to activated protein C due to Factor V Leiden (421527008), Retinal hemorrhage (28998008), Seizure disorder (128613002), Sleep apnea (73430006), Smokes tobacco daily (449868002), Systemic lupus erythematosus (55464009), Thromboembolic disorder (371039008), Thrombophilia (234467004), Thrombophlebitis (64156001), Thrombosis of retinal vein (46085004), Toxic shock syndrome (18504008), Uterine leiomyoma (95315005), Weight gain finding (8943002), Worsening Headache Disorder (NA)	1
Estrone	Angina pectoris (194828000), Asthma (195967001), Benign mammary dysplasia (57993004), Body fluid retention (43498006), Breast lump (89164003), Carcinoma of female breast (447782002), Cardiovascular event risk (395112001), Cerebrovascular accident (230690007), Chloasma (36209000), Cholestasis of pregnancy (235888006), Chorea (271700006), Deep venous thrombosis (128053003), Dementia (52448006), Depressive disorder (35489007), Diabetes mellitus (73211009), Disease of liver (235856003), Disorder of coronary artery (414024009), Disorder of gallbladder (39621005), Endometrial carcinoma (254878006), Endometriosis (129103003), Epilepsy (84757009), Estrogen receptor positive tumor (416053008), Extrapyramidal disease (76349003), Family history of malignant neoplasm of breast (429740004), Functional visual loss (313165001), Heart valve disorder (368009), Hepatic failure (59927004), Hepatic porphyria (55056006), Humoral hypercalcemia of malignancy (47709007), Hypercholesterolemia (13644009), Hyperglycemia (80394007), Hyperlipidemia (55822004), Hyperlipoproteinemia (3744001), Hypertensive disorder (38341003), Hypertensive urgency (443482000), Hypertriglyceridemia (302870006), Hypocalcemia (5291005), Hypothyroidism (40930008), Intermenstrual bleeding - irregular (64996003), Liver function tests abnormal (166603001), Major Surgery with Prolonged Post-Operative Immobilization (NA), Malignant neoplasm of liver (93870000), Malignant tumor of cervix (363354003), Malignant tumor of ovary (363443007), Mammography abnormal (168750009), Migraine (37796009), Myocardial infarction (22298006), Nonspecific Abnormal Papanicolaou Smear of Cervix (NA), Obesity (414916001), Obstructive hyperbilirubinemia (59848001), Porphyria (418470004), Pregnancy, function (289908002), Pulmonary thromboembolism (233935004), Smokes tobacco daily (449868002), Systemic lupus erythematosus (55464009), Thromboembolic disorder (371039008), Thrombophilia (234467004), Thrombophlebitis (64156001), Thrombosis of retinal vein (46085004), Uterine leiomyoma (95315005)	1
Everolimus	Acute infectious disease (63171007), Acute nephropathy (58574008), Anemia (271737000), Ascites (389026000), BK Polyomavirus Reactivation Nephropathy (NA), Breastfeeding (mother) (413712001), Diabetes mellitus (73211009), Disease of liver (235856003), Edema (267038008), Hemolytic uremic syndrome (111407006), Hepatic failure (59927004), Hypercholesterolemia (13644009), Hyperglycemia (80394007), Hyperlipidemia (55822004), Hypertriglyceridemia (302870006), Impaired wound healing (271618001), Interstitial pneumonia (64667001), Kidney disease (90708001), Leukopenia (84828003), Malignant lymphoma (118600007), Neutropenic disorder (303011007), Pericardial effusion (373945007), Pleural Effusions (NA), Pregnancy, function (289908002), Primary malignant neoplasm (372087000), Proteinuria (29738008), Stomatitis (61170000), Thrombocytopenic disorder (302215000), Thrombosis of renal artery (95579008), Thrombotic thrombocytopenic purpura (78129009)	1
Fenofibrate	Acute nephropathy (58574008), Acute pancreatitis (197456007), Agranulocytosis (17182001), Anemia (271737000), Breastfeeding (mother) (413712001), Calculus in biliary tract (266474003), Disease of liver (235856003), Disorder of gallbladder (39621005), Disorder of muscle (129565002), Hepatic failure (59927004), Impaired renal function disorder (197663003), Leukopenia (84828003), Liver function tests abnormal (166603001), Myositis (26889001), Primary biliary cirrhosis (31712002), Rhabdomyolysis (240131006), Thrombocytopenic disorder (302215000), Thromboembolic disorder (371039008)	0

Flunisolide	Acute tuberculosis (25629007), Adrenal cortical hypofunction (386584007), Bacterial infectious disease (87628006), Bilateral cataracts (95722004), Disease caused by parasite (17322007), Disease of liver (235856003), Epistaxis (12441001), Exposure to varicella (444453009), Glaucoma (23986001), Inactive tuberculosis (11999007), Measles (14189004), Mycosis (3218000), Nasal Candidiasis (NA), Nasal Septal Ulcers (NA), Nasal Trauma (NA), Operation on nose (88733004), Ophthalmic herpes simplex (186542001), Oropharyngeal Candidiasis (NA), Perforation of nasal septum (80142000), Uncontrolled Bacterial Infections (NA), Untreated Fungal Infection (NA), Varicella-zoster virus infection (309465005)	0
Fluorometholone	Anemia due to enzyme deficiency (111577008), Bacterial infection of eye (128984004), Deficiency of glucose-6-phosphate dehydrogenase (124134002), Diabetes mellitus (73211009), Eye infection (128351009), Fungal infection of eye (31194008), Herpes simplex dendritic keratitis (29943008), Herpes simplex keratitis (9389005), Krukenberg spindle (85430004), Ocular hypertension (4210003), Open-angle glaucoma (84494001), Porphyria (418470004), Severe myopia (34187009), Tuberculosis of eye (49107007), Vaccinia keratitis (397552005), Viral eye infection (312132001)	1
Flutamide	Anemia due to enzyme deficiency (111577008), Breastfeeding (mother) (413712001), Deficiency of cytochrome-b¿5¡ reductase (124184009), Deficiency of glucose-6-phosphate dehydrogenase (124134002), Disease of liver (235856003), Hepatic failure (59927004), Hyperbilirubinemia (14783006), Methemoglobinemia (38959009), Pregnancy, function (289908002)	0
Fulvestrant	Blood coagulation disorder (64779008), Breastfeeding (mother) (413712001), Disease of liver (235856003), Pregnancy, function (289908002), Thrombocytopenic disorder (302215000)	0
Gefitinib	Acute nephropathy (58574008), Breastfeeding (mother) (413712001), Disease of liver (235856003), Fibrosis of lung (51615001), Interstitial pneumonia (64667001), Non-small cell lung cancer, negative for epidermal growth factor receptor expression (427038005), Pregnancy, function (289908002)	0
Gemfibrozil	Acute nephropathy (58574008), Anemia (271737000), Breastfeeding (mother) (413712001), Calculus in biliary tract (266474003), Disease of liver (235856003), Disorder of gallbladder (39621005), Impaired renal function disorder (197663003), Leukopenia (84828003), Primary biliary cirrhosis (31712002), Thrombocytopenic disorder (302215000)	0
Ibrutinib	Pregnancy, function (289908002)	0
Lapatinib	Breastfeeding (mother) (413712001), Congenital long QT syndrome (442917000), Drug-induced hepatitis (235876009), Hepatic failure (59927004), Hypokalemia (43339004), Hypomagnesemia (190855004), Interstitial pneumonia (64667001), Left heart failure (85232009), Liver function tests abnormal (166603001), Pregnancy, function (289908002), Prolonged QT interval (111975006), Severe diarrhea (409587002), Torsades de pointes (31722008)	0
Levothyroxine	Angina pectoris (194828000), Conduction disorder of the heart (44808001), Diabetes mellitus (73211009), Disorder of cardiovascular system (49601007), Disorder of coronary artery (414024009), Hypertensive disorder (38341003), Hyperthyroidism (34486009), Hypopituitarism (74728003), Myocardial infarction (22298006), Osteopenia (312894000), Osteoporosis (64859006), Primary adrenocortical insufficiency (373662000), Thyrotoxic crisis (29028009)	1
Liothyronine	Angina pectoris (194828000), Conduction disorder of the heart (44808001), Diabetes mellitus (73211009), Disorder of cardiovascular system (49601007), Disorder of coronary artery (414024009), Hypertensive disorder (38341003), Hyperthyroidism (34486009), Hypopituitarism (74728003), Myocardial infarction (22298006), Osteopenia (312894000), Osteoporosis (64859006), Primary adrenocortical insufficiency (373662000), Thyrotoxic crisis (29028009)	1
Mesalazine	Anemia (271737000), Disease of liver (235856003), Interstitial nephritis (28689008), Kidney disease (90708001), Minimal change disease (44785005), Neutropenic disorder (303011007), Pyloric obstruction (244815007)	0
Mestranol	Angina pectoris (194828000), Benign mammary dysplasia (57993004), Body fluid retention (43498006), Carcinoma of female breast (447782002), Cardiovascular event risk (395112001), Cerebrovascular accident (230690007), Cholestasis of pregnancy (235888006), Deep venous thrombosis (128053003), Depressive disorder (35489007), Diabetes mellitus (73211009), Disease of liver (235856003), Disorder of coronary artery (414024009), Disorder of gallbladder (39621005), Endometrial carcinoma (254878006), Estrogen receptor positive tumor (416053008), Heart valve disorder (368009), Hypercholesterolemia (13644009), Hyperglycemia (80394007), Hyperlipidemia (55822004), Hypertensive disorder (38341003), Hypertensive urgency (443482000), Hypertriglyceridemia (302870006), Intermenstrual bleeding - irregular (64996003), Liver function tests abnormal (166603001), Major Surgery with Prolonged Post-Operative Immobilization (NA), Malignant neoplasm of liver (93870000), Malignant tumor of cervix (363354003), Migraine (37796009), Myocardial infarction (22298006), Nonspecific Abnormal Papanicolaou Smear of Cervix (NA), Obesity (414916001), Obstructive hyperbilirubinemia (59848001), Porphyria (418470004), Pregnancy, function (289908002), Pulmonary thromboembolism (233935004), Smokes tobacco daily (449868002), Thromboembolic disorder (371039008), Thrombophlebitis (64156001), Thrombosis of retinal vein (46085004)	1
Mifepristone	Acute intermittent porphyria (234422006), Adrenal cortical hypofunction (386584007), Bacterial infectious disease (87628006), Blood coagulation disorder (64779008), Ectopic pregnancy (34801009), Erythropoietic protoporphyria (51022005), Hepatic porphyria (55056006), Incomplete miscarriage (156072005), Intermenstrual bleeding - irregular (64996003), Leukocytosis (111583006), Miscarriage with sepsis (67465009), Porphyria cutanea tarda (61860000), Postabortal Hemorrhage (NA), Severe adrenal insufficiency (24867002), Uterine Adnexal Mass (NA), Variegate porphyria (58275005)	0

Nilotinib	Anemia (271737000), Breastfeeding (mother) (413712001), Congenital long QT syndrome (442917000), Diabetes mellitus (73211009), Disease of liver (235856003), Elevated Serum Lipase (NA), Gastrectomy (53442002), Hyperbilirubinemia (14783006), Hypercholesterolemia (13644009), Hyperglycemia (80394007), Hyperkalemia (14140009), Hypertensive disorder (38341003), Hypocalcemia (5291005), Hypokalemia (43339004), Hypomagnesemia (190855004), Hyponatremia (89627008), Hypophosphatemia (4996001), Liver function tests abnormal (166603001), Neutropenic disorder (303011007), Pancreatitis (75694006), Pregnancy, function (289908002), Prolonged QT interval (111975006), Thrombocytopenic disorder (302215000), Torsades de pointes (31722008)	1
Nilutamide	Decreased respiratory function (80954004), Disorder of lung (19829001), Hepatic failure (59927004), Incomplete Testicular Development (NA), Interstitial pneumonia (64667001)	0
Oxandrolone	Arteriosclerotic vascular disease (72092001), Benign prostatic hyperplasia (266569009), Blood coagulation disorder (64779008), Breast Carcinoma in Males (NA), Breastfeeding (mother) (413712001), Cholestatic hepatitis (95556007), Disease of liver (235856003), Disorder of coronary artery (414024009), Edema (267038008), Heart failure (84114007), Hypercalcemia (66931009), Hypercholesterolemia (13644009), Hypoalphalipoproteinemia (190785000), Kidney disease (90708001), Neoplasm of liver (126851005), Neoplasm of prostate (126906006), Nephrotic syndrome (52254009), Peliosis hepatis (58008004), Pregnancy, function (289908002)	1
Oxymetholone	Arteriosclerotic vascular disease (72092001), Benign prostatic hyperplasia (266569009), Blood coagulation disorder (64779008), Breast Carcinoma in Males (NA), Breastfeeding (mother) (413712001), Carcinoma of female breast (447782002), Chronic heart failure (48447003), Diabetes mellitus (73211009), Disease of liver (235856003), Disorder of coronary artery (414024009), Edema (267038008), Heart disease (56265001), Hepatic failure (59927004), Humoral hypercalcemia of malignancy (47709007), Hypercholesterolemia (13644009), Hypoalphalipoproteinemia (190785000), Kidney disease (90708001), Liver function tests abnormal (166603001), Neoplasm of prostate (126906006), Nephrotic syndrome (52254009), Peliosis hepatis (58008004), Pregnancy, function (289908002)	1
Pioglitazone	Acute vomiting (23971007), Adrenal cortical hypofunction (386584007), Alcohol intoxication (25702006), Alcoholism (7200002), Asthenia (13791008), Autonomic dysreflexia (129618003), Body fluid retention (43498006), Breastfeeding (mother) (413712001), Cardiogenic shock (89138009), Chronic heart failure (48447003), Cobalamin deficiency (190634004), Decompensated cardiac failure (195111005), Dehydration (34095006), Disease of liver (235856003), Edema (267038008), Fever (386661006), Fever greater than 100.4 Fahrenheit (4.26e+08), Fracture of bone (125605004), Glucose-6-phosphate dehydrogenase deficiency anemia (62403005), Hemolytic anemia (61261009), Hepatic porphyria (55056006), Hypoglycemic disorder (237630007), Hypopituitarism (74728003), Infectious disease (40733004), Ketoacidosis (56051008), Kidney disease (90708001), Lactic acidosis (91273001), Liver function tests abnormal (166603001), Macular retinal edema (37231002), Malignant tumor of urinary bladder (399326009), Metabolic acidosis (59455009), Myocardial infarction (22298006), Primary adrenocortical insufficiency (373662000), Radiography with IV Iodinated Contrast Agent (NA), Sepsis syndrome (238150007), Severe diarrhea (409587002), Severe Hypoxemia (NA), Shock (27942005), Surgical procedure (387713003), Traumatic injury (417746004)	0
Prednisolone	Anastomosis of intestine (235407009), Anemia due to enzyme deficiency (111577008), Arginase deficiency (23501004), Arginosuccinate Lyase Deficiency (NA), Atrophoderma (399979006), Avascular necrosis of bone (397758007), Bacterial infection of eye (128984004), Bilateral cataracts (95722004), Cerebral malaria (53622003), Cerebral trauma (275382005), Chronic heart failure (48447003), Cirrhosis of liver (19943007), Citrullinemia (398680004), Congenital hyperammonemia, type I (62522004), Deficiency of glucose-6-phosphate dehydrogenase (124134002), Diabetes mellitus (73211009), Disorder of muscle (129565002), Diverticulitis of gastrointestinal tract (271366000), Edema (267038008), Exposure to varicella (444453009), Eye infection (128351009), Fungal infection of eye (31194008), Gastritis (4556007), Herpes simplex (88594005), Herpes simplex dendritic keratitis (29943008), Herpes simplex keratitis (9389005), Hypercholesterolemia (13644009), Hyperglycemia (80394007), Hypertensive disorder (38341003), Hypokalemia (43339004), Hypopituitarism (74728003), Hypothyroidism (40930008), Immunosuppression (38013005), Inactive tuberculosis (11999007), Infection by Strongyloides (1214006), Infectious disease (40733004), Injury of eye region (282752000), Krukenberg spindle (85430004), Measles (14189004), Muscle atrophy (88092000), Mycosis (3218000), Ocular hypertension (4210003), Open-angle glaucoma (84494001), Ophthalmic herpes simplex (186542001), Ornithine Carbamyltransferase Deficiency (NA), Osteopenia (312894000), Osteoporosis (64859006), Pathological fracture (268029009), Peptic ulcer (13200003), Porphyria (418470004), Pregnancy, function (289908002), Psychotic disorder (69322001), Seizure disorder (128613002), Severe myopia (34187009), Skin striae (201066002), Telangiectasia disorder (247479008), Traumatic rupture of tendon (415749005), Tuberculosis (56717001), Tuberculosis of eye (49107007), Vaccinia keratitis (397552005), Varicella-zoster virus infection (309465005), Viral eye infection (312132001)	1
Prednisone	Anastomosis of intestine (235407009), Arginase deficiency (23501004), Arginosuccinate Lyase Deficiency (NA), Avascular necrosis of bone (397758007), Bilateral cataracts (95722004), Cerebral malaria (53622003), Cerebral trauma (275382005), Chronic heart failure (48447003), Cirrhosis of liver (19943007), Citrullinemia (398680004), Congenital hyperammonemia, type I (62522004), Diabetes mellitus (73211009), Disorder of muscle (129565002), Diverticulitis of gastrointestinal tract (271366000), Edema (267038008), Exposure to varicella (444453009), Gastritis (4556007), Herpes simplex (88594005), Hypercholesterolemia (13644009), Hyperglycemia (80394007), Hypertensive disorder (38341003), Hypokalemia (43339004), Hypopituitarism (74728003), Hypothyroidism (40930008), Immunosuppression (38013005), Inactive tuberculosis (11999007), Infection by Strongyloides (1214006), Infectious disease (40733004), Measles (14189004), Muscle atrophy (88092000), Mycosis (3218000), Open-angle glaucoma (84494001), Ophthalmic herpes simplex (186542001), Ornithine Carbamyltransferase Deficiency (NA), Osteopenia (312894000), Osteoporosis (64859006), Pathological fracture (268029009), Peptic ulcer (13200003), Pregnancy, function (289908002), Psychotic disorder (69322001), Seizure disorder (128613002), Traumatic rupture of tendon (415749005), Tuberculosis (56717001), Varicella-zoster virus infection (309465005)	1

Progesterone	Alcoholism (7200002), Anorexia nervosa (56882008), Asthma (195967001), Bed-ridden (160685001), Body fluid retention (43498006), Breast lump (89164003), Carcinoma of female breast (447782002), Cardiovascular event risk (395112001), Cerebrovascular accident (230690007), Cerebrovascular disease (62914000), Chloasma (36209000), Chorea (271700006), Chronic heart failure (48447003), Deep venous thrombosis (128053003), Dementia (52448006), Depressive disorder (35489007), Diabetes mellitus (73211009), Diplopia (24982008), Disease of liver (235856003), Disorder of coronary artery (414024009), Disorder of gallbladder (39621005), Ectopic pregnancy (34801009), Epilepsy (84757009), Estrogen receptor positive tumor (416053008), Family history of malignant neoplasm of breast (429740004), Fibroadenosis of breast (23260002), Functional visual loss (313165001), Hepatic failure (59927004), Hepatic porphyria (55056006), Humoral hypercalcemia of malignancy (47709007), Hypercholesterolemia (13644009), Hyperglycemia (80394007), Hyperlipidemia (55822004), Hyperlipoproteinemia (3744001), Hypertensive disorder (38341003), Hypertensive urgency (443482000), Hypertriglyceridemia (302870006), Hypocalcemia (5291005), Hypothyroidism (40930008), Impaired glucose tolerance (9414007), Incomplete miscarriage (156072005), Intermenstrual bleeding - irregular (64996003), Jaundice (18165001), Kidney disease (90708001), Malignant neoplasm of liver (93870000), Malignant tumor of ovary (363443007), Mammography abnormal (168750009), Migraine (37796009), Mild pre-eclampsia (41114007), Myocardial infarction (22298006), Neoplasm of liver (126851005), Obesity (414916001), Optic disc edema (423341008), Osteopenia (312894000), Osteoporosis (64859006), Porphyria (418470004), Predisposition to Thrombosis (NA), Pregnancy, function (289908002), Pulmonary thromboembolism (233935004), Retinal hemorrhage (28998008), Seizure disorder (128613002), Smokes tobacco daily (449868002), Systemic lupus erythematosus (55464009), Thromboembolic disorder (371039008), Thrombophilia (234467004), Thrombophlebitis (64156001), Thrombosis of retinal vein (46085004), Uterine leiomyoma (95315005), Weight gain finding (8943002)	1
Raloxifene	Atrial fibrillation (49436004), Breastfeeding (mother) (413712001), Cardiovascular event risk (395112001), Cerebrovascular accident (230690007), Chronic heart failure (48447003), Deep venous thrombosis (128053003), Disease of liver (235856003), Disorder of coronary artery (414024009), Hypertensive disorder (38341003), Hypertriglyceridemia (302870006), Impaired renal function disorder (197663003), Lupus anticoagulant disorder (19267009), Pregnancy, function (289908002), Pulmonary thromboembolism (233935004), Smokes tobacco daily (449868002), Thrombophilia (234467004), Thrombophlebitis (64156001), Thrombosis of retinal vein (46085004), Transient ischemic attack (266257000)	1
Ritodrine	Antepartum hemorrhage (34842007), Chorioamnionitis (11612004), Conduction disorder of the heart (44808001), Dehydration (34095006), Eclampsia in pregnancy (198992004), Fetal death (276507005), Gestational diabetes mellitus (11687002), Hypertensive disorder (38341003), Hypertensive urgency (443482000), Hyperthyroidism (34486009), Hypovolemia (28560003), Migraine (37796009), Mild pre-eclampsia (41114007), Pulmonary hypertension (70995007), Severe pre-eclampsia (46764007)	1
Salbutamol	Angle-closure glaucoma (392291006), Benign prostatic hyperplasia (266569009), Bladder outflow obstruction (236645006), Chronic myocardial ischemia (413844008), Conduction disorder of the heart (44808001), Congenital long QT syndrome (442917000), Diabetes mellitus (73211009), Hypertensive disorder (38341003), Hyperthyroidism (34486009), Hypokalemia (43339004), Ketoacidosis (56051008), Metabolic acidosis (59455009), Myocardial ischemia (414795007), Ocular hypertension (4210003), Paradoxical bronchospasm (102578005), Prolonged QT interval (111975006), Retention of urine (267064002), Seizure disorder (128613002)	1
Spironolactone	Acute nephropathy (58574008), Acute pancreatitis (197456007), Anuria (2472002), Azotemia (445009001), Dehydration (34095006), Diabetes mellitus (73211009), Gout (90560007), Hepatic coma (72836002), Hypercalcemia (66931009), Hypercholesterolemia (13644009), Hyperkalemia (14140009), Hyperparathyroidism (66999008), Hyperuricemia (35885006), Hypochloremic alkalosis (70134007), Hypokalemia (43339004), Hypomagnesemia (190855004), Hyponatremia (89627008), Hypovolemia (28560003), Kidney disease (90708001), Metabolic acidosis, normal anion gap, acidifying salts (18104000), Neonatal hyperbilirubinemia (281610001), Secondary angle-closure glaucoma (21571006), Sympathectomy (57071006), Systemic lupus erythematosus (55464009)	1
Sulfasalazine	Anemia due to enzyme deficiency (111577008), Aplastic anemia (306058006), Asthma (195967001), Deficiency of glucose-6-phosphate dehydrogenase (124134002), Disease of liver (235856003), Gastrointestinal obstruction (126765001), Kidney disease (90708001), Neutropenic disorder (303011007), Porphyria (418470004), Slow acetylator due to N-acetyltransferase enzyme variant (425079005), Urinary tract obstruction (7163005)	0
Telmisartan	Acute pancreatitis (197456007), Anuria (2472002), Azotemia (445009001), Chronic idiopathic constipation (82934008), Dehydration (34095006), Diabetes mellitus (73211009), Disease of liver (235856003), Gout (90560007), Hepatic coma (72836002), Hepatic failure (59927004), Hypercalcemia (66931009), Hypercholesterolemia (13644009), Hyperkalemia (14140009), Hyperparathyroidism (66999008), Hyperuricemia (35885006), Hypochloremic alkalosis (70134007), Hypokalemia (43339004), Hypomagnesemia (190855004), Hyponatremia (89627008), Hypovolemia (28560003), Kidney disease (90708001), Low blood pressure (45007003), Neonatal hyperbilirubinemia (281610001), Obstruction of bile duct (30144000), Pregnancy, function (289908002), Renal artery stenosis (302233006), Secondary angle-closure glaucoma (21571006), Severe Aortic Valve Stenosis (NA), Sympathectomy (57071006), Systemic lupus erythematosus (55464009)	1
Temsirolimus	Acute infectious disease (63171007), Breastfeeding (mother) (413712001), Cerebrovascular accident (230690007), Diabetes mellitus (73211009), Disease of liver (235856003), Gastrointestinal perforation (51875005), Hyperbilirubinemia (14783006), Hyperglycemia (80394007), Hyperlipidemia (55822004), Impaired wound healing (271618001), Interstitial pneumonia (64667001), Kidney disease (90708001), Malignant neoplasm of brain (428061005), Perioperative care (133897009), Pregnancy, function (289908002), Secondary malignant neoplasm of cerebrum (94248000), Surgical procedure (387713003)	1

Testosterone	Arteriosclerotic vascular disease (72092001), Asthma (195967001), Benign prostatic hyperplasia (266569009), Body fluid retention (43498006), Breast Carcinoma in Males (NA), Breast lump (89164003), Breastfeeding (mother) (413712001), Carcinoma of female breast (447782002), Cardiovascular event risk (395112001), Cerebrovascular accident (230690007), Chloasma (36209000), Chorea (271700006), Chronic heart failure (48447003), Chronic lung disease (413839001), Deep venous thrombosis (128053003), Dementia (52448006), Diabetes mellitus (73211009), Disease of liver (235856003), Disorder of coronary artery (414024009), Disorder of gallbladder (39621005), Edema (267038008), Endometrial carcinoma (254878006), Endometriosis (129103003), Epilepsy (84757009), Estrogen receptor positive tumor (416053008), Family history of malignant neoplasm of breast (429740004), Functional visual loss (313165001), Gynecomastia (4754008), Heart disease (56265001), Heart failure (84114007), Hepatic porphyria (55056006), Humoral hypercalcemia of malignancy (47709007), Hypercalcemia (66931009), Hypercholesterolemia (13644009), Hyperlipoproteinemia (3744001), Hypertensive disorder (38341003), Hypertensive urgency (443482000), Hypertriglyceridemia (302870006), Hypocalcemia (5291005), Hypothyroidism (40930008), Intermenstrual bleeding - irregular (64996003), Kidney disease (90708001), Magnetic resonance imaging (113091000), Malignant tumor of ovary (363443007), Mammography abnormal (168750009), Migraine (37796009), Myocardial infarction (22298006), Neoplasm of liver (126851005), Neoplasm of prostate (126906006), Obesity (414916001), Pregnancy, function (289908002), Pulmonary thromboembolism (233935004), Sleep apnea (73430006), Smokes tobacco daily (449868002), Systemic lupus erythematosus (55464009), Thromboembolic disorder (371039008), Thrombophilia (234467004), Thrombophlebitis (64156001), Thrombosis of retinal vein (46085004), Uterine leiomyoma (95315005)	1
Theophylline	Acute hepatitis (37871000), Acute tuberculosis (25629007), Alcoholism (7200002), Anemia (271737000), Angina pectoris (194828000), Angle-closure glaucoma (392291006), Benign prostatic hyperplasia (266569009), Bladder outflow obstruction (236645006), Breastfeeding (mother) (413712001), Chronic heart failure (48447003), Chronic idiopathic constipation (82934008), Chronic myocardial ischemia (413844008), Conduction disorder of the heart (44808001), Continuous fever (271751000), Cor pulmonale (83291003), Coronary arteriosclerosis (53741008), Cystic fibrosis (190905008), Depressive disorder (35489007), Dermatitis herpetiformis (111196000), Diabetes mellitus (73211009), Disease of liver (235856003), Disorder of cardiovascular system (49601007), Disorder of coronary artery (414024009), Dyspnea (267036007), Goiter (3716002), Hashimoto thyroiditis (21983002), Hepatic coma (72836002), Hepatic encephalopathy (13920009), Hepatic failure (59927004), Hyperammonemia (9360008), Hyperkalemia (14140009), Hypertensive disorder (38341003), Hypertensive urgency (443482000), Hyperthyroidism (34486009), Hypocomplementemic urticarial vasculitis (239945009), Hypothyroidism (40930008), Kidney disease (90708001), Multiple organ failure (57653000), Myocardial infarction (22298006), Myocardial ischemia (414795007), Neutropenic disorder (303011007), Open-angle glaucoma (84494001), Peptic ulcer (13200003), Poisoning by phenobarbital (64921004), Porphyria (418470004), Pregnancy, function (289908002), Psychotic disorder (69322001), Pulmonary edema (19242006), Retention of urine (267064002), Seizure disorder (128613002), Sepsis syndrome (238150007), Severe Hypoxemia (NA), Shock (27942005), Sleep apnea (73430006), Smokes tobacco daily (449868002), Smoking cessation assistance (384742004), Substance abuse (66214007), Suicidal thoughts (6471006), Tachyarrhythmia (6285003), Theophylline Toxicity (NA), Third trimester pregnancy (41587001), Thrombocytopenic disorder (302215000)	1
Timolol	Acute cerebrovascular insufficiency (29322000), Acute disease of cardiovascular system (128487001), Acute nephropathy (58574008), Acute pancreatitis (197456007), Anaphylaxis (39579001), Anuria (2472002), Azotemia (445009001), Bronchospasm (4386001), Cardiogenic shock (89138009), Cerebrovascular disease (62914000), Complete atrioventricular block (27885002), Decompensated cardiac failure (195111005), Dehydration (34095006), Depressive disorder (35489007), Diabetes mellitus (73211009), Disease of liver (235856003), General anesthesia (50697003), Gout (90560007), Hepatic coma (72836002), Hypercalcemia (66931009), Hypercholesterolemia (13644009), Hyperparathyroidism (66999008), Hyperthyroidism (34486009), Hyperuricemia (35885006), Hypochloremic alkalosis (70134007), Hypoglycemic disorder (237630007), Hypokalemia (43339004), Hypomagnesemia (190855004), Hyponatremia (89627008), Hypovolemia (28560003), Kidney disease (90708001), Myasthenia gravis (91637004), Neonatal hyperbilirubinemia (281610001), Orthostatic hypotension (28651003), Partial atrioventricular block (195039008), Pregnancy, function (289908002), Pulmonary emphysema (87433001), Raynaud's phenomenon (266261006), Right ventricular failure (367363000), Secondary angle-closure glaucoma (21571006), Severe chronic obstructive pulmonary disease (313299006), Sinus bradycardia (49710005), Sympathectomy (57071006), Systemic lupus erythematosus (55464009), Thromboangiitis obliterans (52403007)	1
Treprostinil	Blood coagulation disorder (64779008), Disease of liver (235856003), Disorder of lung (19829001), Kidney disease (90708001), Low blood pressure (45007003)	0

Tretinoin	Acute pancreatitis (197456007), Adrenal cortical hypofunction (386584007), Agranulocytosis (17182001), Alcoholism (7200002), Anorexia nervosa (56882008), Atopic dermatitis (24079001), Atrophoderma (399979006), Benign intracranial hypertension (68267002), Bleeding (131148009), Breastfeeding (mother) (413712001), Congenital neutropenia (89655007), Corneal opacity (64634000), Crohn's disease (34000006), Denuded skin (418242004), Depressive disorder (35489007), Diabetes mellitus (73211009), Disease of liver (235856003), Disorder of bone (76069003), Disorder of cardiovascular system (49601007), Dyspnea (267036007), Eczema (43116000), Edema (267038008), Fever (386661006), Hearing loss (15188001), Hypercholesterolemia (13644009), Hypertriglyceridemia (302870006), Inflammatory bowel disease (24526004), Inflammatory disease of liver (128241005), Keratoconjunctivitis sicca (302896008), Kidney disease (90708001), Leukocytosis (111583006), Liver function tests abnormal (166603001), Mycosis fungoides (118618005), Myocardial infarction (22298006), Neutropenic disorder (303011007), Obesity (414916001), Osteomalacia (4598005), Osteoporosis (64859006), Pericardial effusion (373945007), Peripheral vascular disease (400047006), Pleural Effusions (NA), Predisposition To Hypertriglyceridemia (NA), Pregnancy, function (289908002), Pseudomembranous enterocolitis (397683000), Psychotic disorder (69322001), Pulmonary Infiltrates (NA), Rectal hemorrhage (12063002), Retinoic acid syndrome (450887006), Severe diarrhea (409587002), Suicidal thoughts (6471006), Sunburn (23346002), Telangiectasia disorder (247479008), Tinnitus (60862001), Ulcerative colitis (64766004), Venous thrombosis (111293003), Vitiligo (56727007)	1
Vorinostat	Anemia (271737000), Breastfeeding (mother) (413712001), Congenital long QT syndrome (442917000), Deep venous thrombosis (128053003), Dehydration (34095006), Diabetes mellitus (73211009), Diarrhea (62315008), Disease of liver (235856003), Hyperglycemia (80394007), Hypocalcemia (5291005), Hypokalemia (43339004), Hypomagnesemia (190855004), Pregnancy, function (289908002), Prolonged QT interval (111975006), Pulmonary thromboembolism (233935004), Thrombocytopenic disorder (302215000), Thromboembolic disorder (371039008), Vomiting (422400008)	1