

# Investigating the deregulation of metabolic tasks via Minimum Network Enrichment Analysis (MiNEA) as applied to nonalcoholic fatty liver disease using mouse and human omics data

---

Short title: Deregulated metabolic task in nonalcoholic fatty liver disease

Vikash Pandey<sup>1</sup>, Vassily Hatzimanikatis<sup>1\*</sup>

<sup>1</sup>Laboratory of Computational Systems Biotechnology, EPFL, Lausanne, Switzerland.

\*corresponding author: [vassily.hatzimanikatis@epfl.ch](mailto:vassily.hatzimanikatis@epfl.ch)

## Abstract

Nonalcoholic fatty liver disease (NAFLD) is associated with metabolic syndromes spanning a wide spectrum of diseases, from simple steatosis to the more complex nonalcoholic steatohepatitis. To identify the deregulation that occurs in metabolic processes at the molecular level that give rise to these various NAFLD phenotypes, algorithms such as pathway enrichment analysis (PEA) can be used. These analyses

20 require the use of predefined pathway maps, which are composed of reactions  
describing metabolic processes/subsystems. Unfortunately, the annotation of the  
metabolic subsystems can differ depending on the pathway database used, making  
these approaches subject to biases associated with different pathway annotations, and  
these methods cannot capture the balancing of cofactors and byproducts through the  
25 complex nature and interactions of genome-scale metabolic networks (GEMs). Here,  
we introduce a framework entitled **Minimum Network Enrichment Analysis**  
(MiNEA) that is applied to GEMs to generate all possible alternative minimal  
networks (MiNs), which are possible and feasible networks composed of all the  
reactions pertaining to various metabolic subsystems that can synthesize a target  
30 metabolite. We applied MiNEA to investigate deregulated MiNs and to identify key  
regulators in different NAFLD phenotypes, such as a fatty liver and liver  
inflammation, in both humans and mice by integrating condition-specific  
transcriptomics data from liver samples. We identified key deregulations in the  
synthesis of cholesteryl esters, cholesterol, and hexadecanoate in both humans and  
35 mice, and we found that key regulators of the hydrogen peroxide synthesis network  
were regulated differently in humans and mice. We further identified which MiNs  
demonstrate the general and specific characteristics of the different NAFLD  
phenotypes. MiNEA is applicable to any GEM and to any desired target metabolite,  
making MiNEA flexible enough to study condition-specific metabolism for any given  
40 disease or organism.

## Author Summary

This work aims to introduce a network-based enrichment analysis using metabolic networks and transcriptomics data. Previous pathways/subsystems enrichment methods use predefined gene annotations of metabolic processes and gene annotations  
45 can differ based on different resources and can produce bias in pathways definitions. Thus, we introduce a framework, Minimum Network Enrichment Analysis (MiNEA), which first finds all possible minimal-size networks for a given metabolic process/task and then identifies deregulated minimal networks using deregulated  
50 reactions that are overlapped across all possible minimal-size networks. We applied MiNEA to identify deregulated metabolic tasks and their synthesis networks in the steatosis and nonalcoholic steatohepatitis (NASH) disease using a metabolic network and transcriptomics data of mouse and human liver samples. We identified key regulators of NASH from the synthesis networks of hydrogen peroxide and ceramide  
55 in both humans and mice. We also identified opposite deregulation in NASH for the phosphatidylserine synthesis network between humans and mice. MiNEA finds synthesis networks for a given target metabolite and due to this it is flexible to study deregulation in different phenotypes. MiNEA can be widely applicable for studying context-specific metabolism for any organism because the metabolic networks and  
60 context-specific gene expression data are now available for many organisms.

## Introduction

Nonalcoholic fatty liver disease (NAFLD) is the most common cause of liver disease in western countries [1], affecting an estimated 25% to 45% of the general US

population, though this percentage is much higher among individuals suffering from  
65 obesity and diabetes [1, 2]. There are two categories of NAFLD: steatosis, which is an  
accumulation of fat in the liver, and nonalcoholic steatohepatitis (NASH), which  
consists of the additional presence of liver inflammation and hepatocellular injury  
with or without fibrosis [2].

Gene deregulation occurs when the cell no longer maintains precise control over  
70 certain genes, and therefore expressed proteins, and is associated with various  
diseases such as NAFLD. Determining which genes are deregulated in disease and  
how these genes are deregulated can increase the understanding of any disease as well  
as provide new pathways for therapeutic treatments. To study gene deregulation,  
pathway enrichment analysis (PEA) uses maps of metabolic processes and subsystem  
75 reactions to determine if a given list of genes (gene set) is associated with a certain  
biochemical pathway, shedding important conceptual insight into gene deregulation.  
Many PEA algorithms, such as ConsensusPathDB [3] and Piano [3, 4], have been  
developed to identify biological processes based on gene sets, while other algorithms,  
such as IMPaLA [5], are based on both gene and metabolite sets. While these  
80 algorithms can successfully provide insights about various disease phenotypes, they  
use predefined metabolic maps that are limited to our current knowledge of these  
hugely intricate pathways and can differ depending on the pathway database used,  
meaning they could miss information about the wide range of complex metabolic  
interactions. These approaches are also subject to biases associated with the different  
85 pathway annotations, and they cannot capture the balancing of cofactors and  
byproducts through the complex nature and interactions of genome-scale metabolic  
networks (GEMs).

To overcome this, a method has been developed based on elementary flux modes (EFMs), which are non-decomposable flux distributions in metabolic networks [6]. These flux distributions indicate a conceptual description of metabolic pathways. This method computes EFMs from a given GEM and uses tissue-specific transcriptomics to identify a subset of tissue-specific EFMs. However, the enumeration of all EFMs in a GEM can be computationally intractable, and EFMs are not necessarily specific to a target metabolite synthesis, meaning that targeted pathways are not always able to be elucidated using EFMs.

One could also use graph-based methods (GBM) that require network topologies and gene expression profiles as inputs to extract deregulated subnetworks that occur between two conditions [7-9]. These methods have been applied to protein-protein networks for prostate cancer study, and they are potentially applicable to metabolic networks. GBM methods use graph-theoretic properties on network topologies but miss information about additional constraints, such as mass balance [10] and thermodynamics [11]. Ideally, the results for these studies would include a set of mass balanced subnetworks that could be used to understand the carbon, energy, and redox flows from precursor metabolites to target metabolites and complex metabolic tasks. Here, we propose a method called **Minimal Network Enrichment Analysis** (MiNEA) that compares two conditions using transcriptomics, proteomics, and metabolomics to identify deregulated minimal networks (MiNs), which are reactions pertaining to various metabolic subsystems that can synthesize a target metabolite. The MiNEA algorithm works by formulating metabolic tasks (MTs) to mimic the various NAFLD phenotypes, such as lipid droplet formation, lipoapoptosis, liver inflammation, and oxidative stress. These MTs include the

known set of metabolites necessary to form the modeled phenotype. The cell can use different pathways to form the same metabolites and create the same phenotypes, and the purpose of MiNEA is to identify deregulated alternative routes for a given MT between two conditions, such as a control and treatment. First, MiNEA uses these MTs to compute alternative thermodynamically feasible MiNs applying thermodynamic constraints [11, 12] to a mouse GEM [13]. Then, it uses mouse and human liver sample expression data [1, 14] and identifies deregulated metabolic processes in mice and humans that potentially lead to the NAFLD phenotypes.

MiNEA identified an upregulation in the oxidative stress synthesis network in mice, but this network was found to be downregulated in humans. The cholesterol and triacylglycerol synthesis networks were deregulated in humans only, while the ceramide synthesis network was only deregulated in mice. We found downregulated reactions in the synthesis network for cholesteryl esters, cholesterol, and alanine in both humans and mice. We further identified downregulated reactions in the superoxide anion (SOA) synthesis network in humans, specifically in NASH as opposed to steatosis, while upregulation was found in the SOA synthesis network in mice. This perturbation through the SOA synthesis network in NASH suggested an unbalanced ceramide synthesis, and studies have shown that the ceramide is a key regulator of apoptosis and promotes fibrosis in the hepatic steatosis model [15, 16].

MiNEA can generate MiNs for any target metabolite, e.g. a metabolite produced under specific phenotypes or a biomass building block [17] that is needed for cell growth. Additionally, MiNEA can integrate condition- and context-specific omics data to understand the deregulated phenotypes associated with a set of differentially

expressed genes. These characteristics make MiNEA a versatile tool for exploring and understanding different metabolic phenotypes.

## Results and Discussions

### Experimental details for MiNEA

140 Nonalcoholic fatty liver disease (NAFLD) has been defined as a metabolic disease associated with insulin-resistance syndrome [18]. To study the differences seen in NAFLD phenotypes between mouse and human manifestations, we collected human expression data from the three diagnosis groups, normal (N), steatosis (S), and nonalcoholic steatohepatitis (NS), and mouse expression data from control and DDC-  
145 supplemented diet conditions for three genetically different mouse strains, AJ, B6, and PWD. DDC-supplemented diet reproduces steatosis and NASH phenotypes (see Materials and Methods for detail). We refer to these data as *human expression data* and *mouse expression data* throughout this section and integrate it into a mouse model iMM1415 that is constructed based on a human GEM Recon1 [13]. We analyzed both  
150 human and mouse MTs using iMM1415 with the assumption that both types of mammalian cells have a similar metabolism (See S1 text). In the following sections, we analyzed the deregulation of minimal networks in core reactions and enriched minimal networks based on the deregulation (up- or downregulated) genes.

### 155 Minimal networks

GEMs represent an entire cellular metabolic network in the form of mathematical constraints [19]. These network reconstructions have grown rapidly in the last decades, and now many GEMs for different organisms are available [20]. GEMs and condition-specific experimental data can be employed to generate and test hypotheses  
160 using the Minimal Network Enrichment analysis (MiNEA) framework that has been



developed in this study. As compared to a method that had been proposed for finding subnetworks from a GEM that optimizes only some part of metabolic networks to find subnetworks [21], MiNEA optimizes whole metabolic networks to find minimal networks.

165 To examine the data that can be derived from MiNEA in terms of NAFLD, we first wanted to compute MiNs for each of the MTs in Table 7 (see Materials and Methods for the table) that were significant phenotypes in NAFLD. Instead of representing a main linear route between precursors and targets of MTs, a MiN is instead composed of many subsystems integrated into subnetworks that must be active, meaning  
170 reactions of subnetworks carries fluxes, to fulfill the target MT. The MiNEA algorithm facilitates enumeration of alternative MiNs and provides more flexibility to the analysis of different metabolic phenotypes and their respective environmental and genetic perturbations. The use of MT for analysis with MiNEA allows this method to be more easily generalized and applied to the study of other metabolic phenotypes and  
175 diseases.

A summary of the MiNs calculated for the MTs in Table 7 (see Materials and Methods for the table) are shown in Table 1. The shortest MiNs were for the synthesis of hydrogen peroxide ( $H_2O_2$ ) (network size = 37), and the longest were for the synthesis of cholesteryl ester (network size = 131; Table 1). We found the greatest  
180 number of possible alternatives for the hexadecanoate (HDCA) synthesis and the least number possible for the phosphatidylserine (PS) synthesis (Table 1), suggesting that the HDCA synthesis is more flexible and the PS synthesis less flexible towards alternative formation compared to the rest of the MTs. Reactions that overlapped between all alternative MiNs were called high-frequency reactions (HFRs), and the

185 percentage of HFRs shows how similar or divergent a MiN is compared to the other  
 MiNs. The percentage of HFRs was between 62% to 91% across the MTs (Table 1).  
 In the superoxide anion (SOA) and the hydrogen peroxide (H<sub>2</sub>O<sub>2</sub>) synthesis, 62% and  
 89% of the reactions were HFRs, respectively, which suggests that the alternative  
 MiNs for a SOA synthesis were more divergent compared to a H<sub>2</sub>O<sub>2</sub> synthesis.

190 Table 1. Summary of MiNs found for the target metabolites comprising the various NAFLD  
 phenotypes. Synthesis of a target metabolite represents a MT. For each MT, we enumerated alternative  
 MiNs. The number of reactions in each MiN represents the size of the MiN. “# Alt” and “# MT HFRs”  
 represent the number of alternatives and the number of overlapping reactions between all alternatives  
 of a given MT, respectively. The symbol “#Phenotypic HFRs” represents the number of overlapping  
 195 reactions form all alternatives of a set of MTs that are associated with a given phenotype.

Phenotypes	Metabolic Tasks (MTs)	Short Name	Min Size # rxns	# Alt	# MT HFRs	% MT HFRs	# Phenotypic HFRs
Oxidative Stress	Hydrogen peroxide	H <sub>2</sub> O <sub>2</sub>	37	126	33	89	16
	Superoxide anion	SOA	42	144	26	62	
Apoptosis and Inflammation	Alanine	ALA	98	39	84	86	36
	Ceramide	CRM	115	8	105	91	
	Glutamine	GLU	116	34	83	72	
Lipid droplet	Cholesterol	CHOL	121	60	108	89	27
	Cholesteryl ester	CHOL_ES	131	237	112	85	
	Diacylglycerol	DAG	82	232	67	82	
	Hexadecanoate	HDCA	75	266	61	81	
	Phosphatidic acid	PA	83	245	65	78	
	Phosphatidylinositol	PAIL	94	237	68	72	
	Phosphatidylethanolamine	PE	128	7	103	80	
	Phosphatidylserine	PS	125	6	98	78	
Triacylglycerol	TAG	84	257	62	74		

As an expanded example, a MiN from the H<sub>2</sub>O<sub>2</sub> synthesis comprises many reactions  
 from various metabolic subsystems (Fig. 1). Each MiN represents a group of active  
 reactions, which are reactions that carries fluxes, within multiple metabolic  
 200 subsystems/pathways that are required for a MT, while a pathway is a group of  
 annotated reactions. In the example in Figure 1, most of the active reactions for the  
 H<sub>2</sub>O<sub>2</sub> synthesis are from the pentose phosphate pathway (PPP) and the  
 glycolysis/gluconeogenesis pathway. PEA differs in that it identifies a marked

deregulation in a pathway as compared to MiNEA, which identifies deregulation in a  
205 MiN that could include multiple pathways.

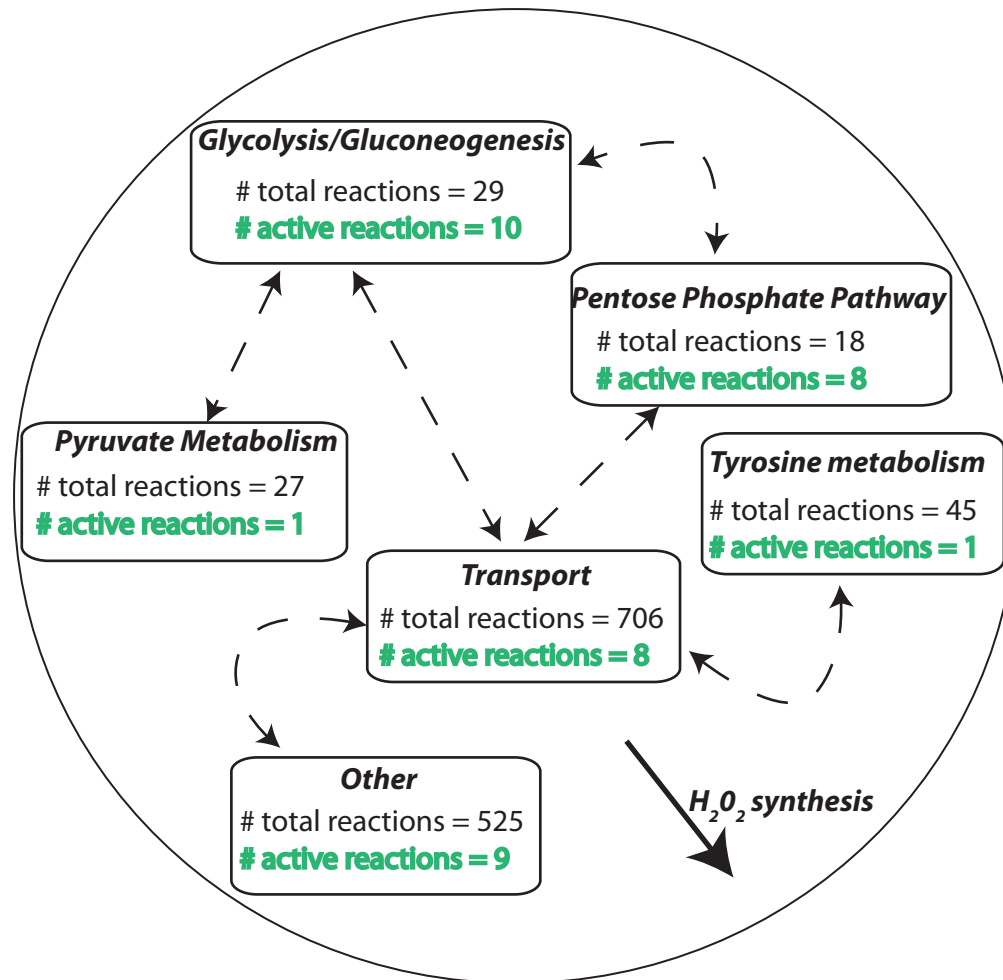


Figure 1. A minimal network (MiN) and high-frequency reactions (HFRs). A MiN  
of  $H_2O_2$  synthesis is illustrated. Each box represents a different metabolic  
210 subsystem/pathway. The number of total reactions listed in each box represents the  
annotated reactions for that subsystem. Active reactions carry non-zero flux, and the  
number of these present in each subsystem is shown in green text.

We further analyzed the HFRs within each phenotype and found that the most  
215 represented metabolic pathway in the apoptosis and inflammation (AI) phenotype was  
related to amino acid metabolism (# phenotypic HFRs=36; Table 1; Fig. 2).  
Interestingly, while the lipid droplet (LD) phenotype has many associated MTs, it  
shares 27 reactions in which 5 reactions are from the glycolysis/gluconeogenesis  
pathway (Fig. 2). These shared reactions can be constitutive candidates for the LD  
220 phenotypes. The HFRs forming the subsystems in Figure 2 are potential candidates  
for the key regulators of the NAFLD phenotypes.

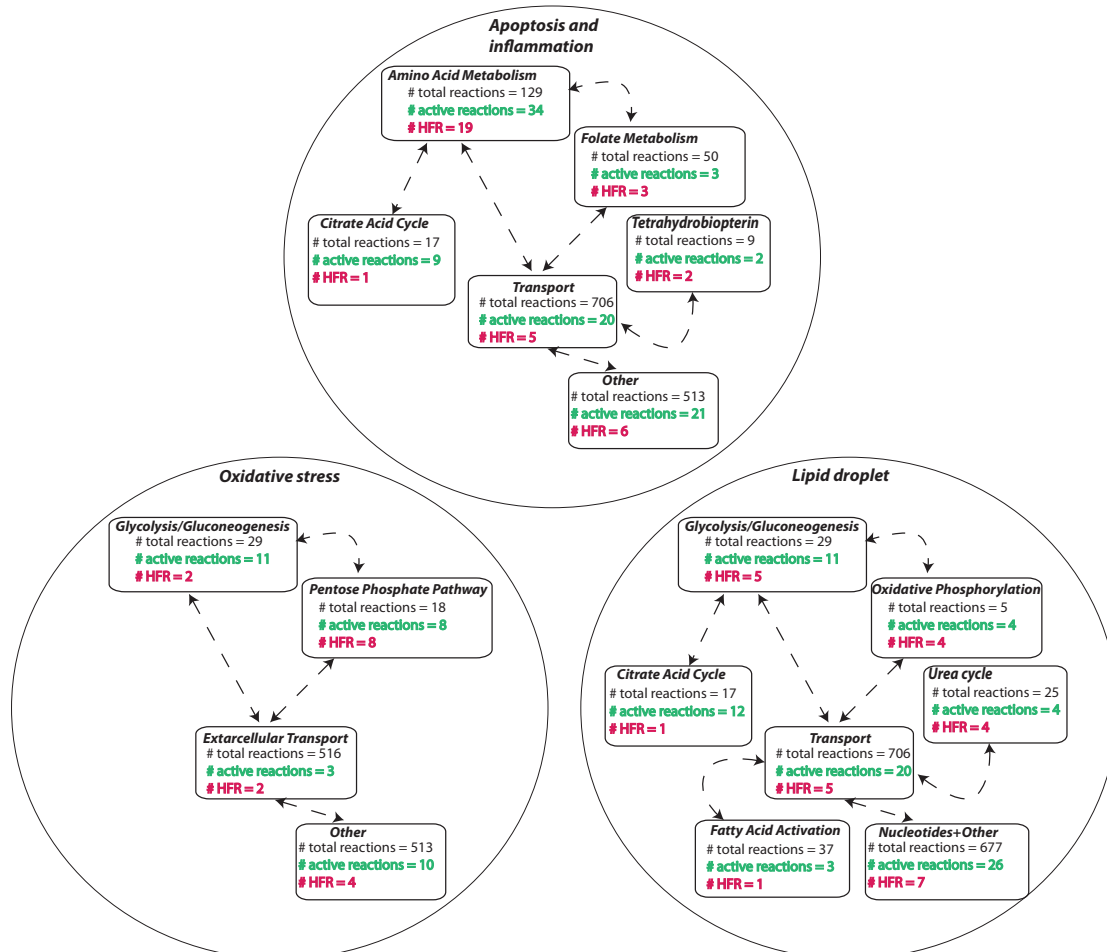


Figure 2. **HFRs for all the phenotypes.** The number of HFRs in the various pathways is shown for the oxidative stress, apoptosis and inflammation, and lipid droplets phenotypes. See Figure 1 caption for a detailed explanation of the chart.

### Deregulation in HFRs

The HFRs within each MT create a topology of reaction hubs in metabolic networks that can be similar to protein hubs in protein-protein interaction networks [22, 23], and the misregulation of these hubs in disease can explain the origin of disease phenotypes. Therefore, we performed an analysis of HFR deregulation to study the NAFLD-specific phenotypes. We identified the up- and downregulated reactions from mouse and human expression data (described in S1) and analyzed the HFRs for which the associated genes were found to be deregulated. In humans, the number of downregulated HFRs was higher compared to upregulated HFRs for the NS vs N and NS vs S across all phenotypes (Table 2). In mice, however, the NS vs N, NS vs S, and S vs N comparisons showed a higher number of upregulated HFRs compared to downregulated ones for the oxidative stress (OS) phenotype, but for the LD and AI phenotypes, we observed similar patterns of up- and downregulated HFRs to humans (Table 2).

Table 2. The deregulated HFRs across all comparisons. Numbers in the table indicate the number of up- and downregulated HFRs. In mice, N represents the control diet (N≈AJ, N≈PWD, or N≈B6 strain fed the control diet). Under the DDC-supplemented diet, AJ mice tended towards NASH phenotypes and PWD mice tended towards steatosis phenotypes. Thus, in mice, the symbols NS stand for AJ and S

stand for the PWD, all with DDC-supplemented diet (NS≈AJ with DDC-supplemented diet; S≈PWD DDC-supplemented diet).

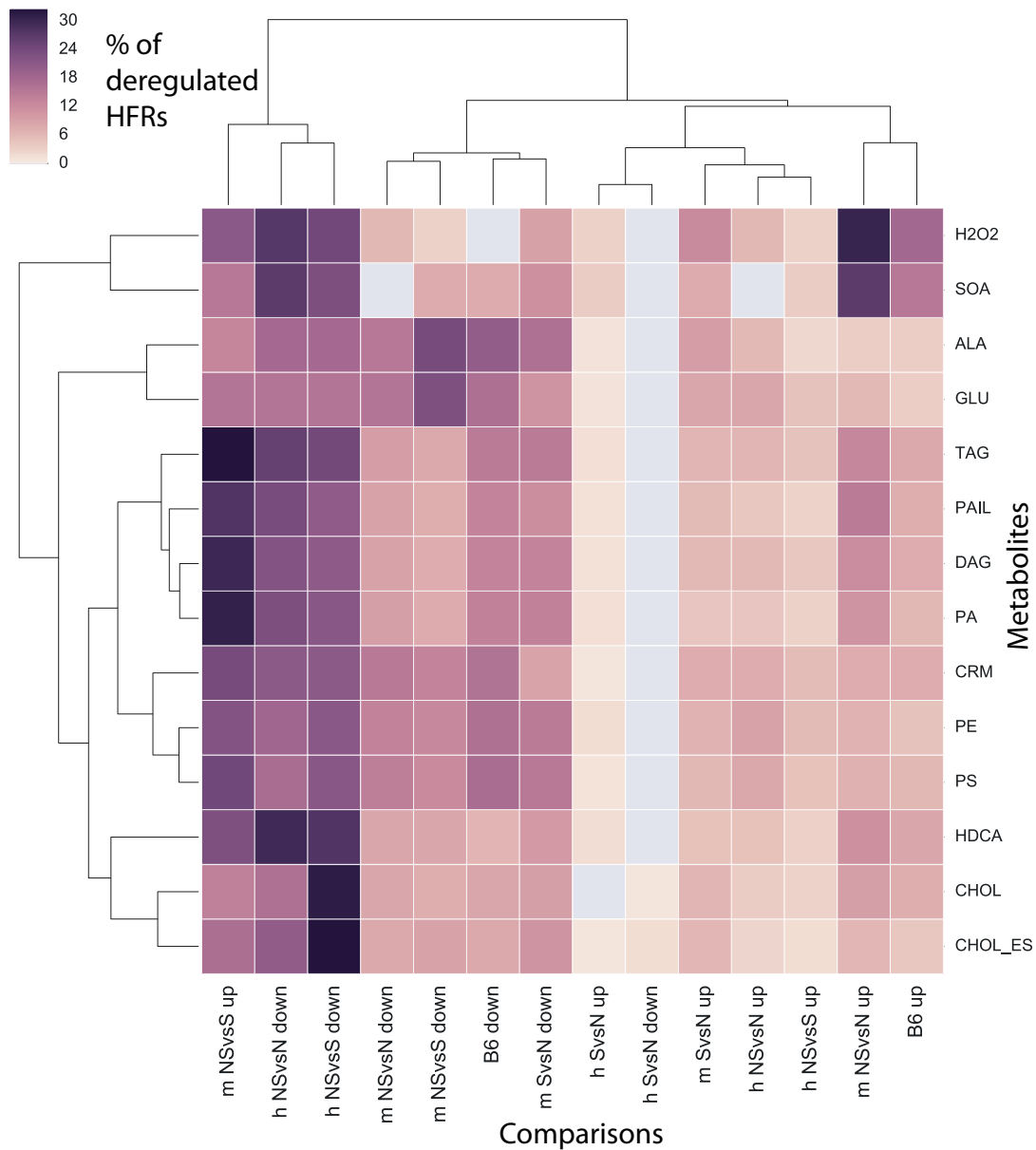
		Mouse							
		N		NS		S			
		up	down	up	down	up	down		
Human	N	OS	/		7		2		Mouse
		AI				9	3	5	
		LD			2	3	1	2	
	NS	OS		5	/		3	1	Mouse
		AI	3	5			6	10	
		LD	3	5			4	2	
	S	OS	1			5	/		Mouse
		AI			1	6			
		LD			1	5			
		N	NS		S		Human		

250

We further analyzed the deregulation in reaction hubs for each MT to identify the most perturbed MTs in NS and S and the differences between NS and S. In humans, the percentage of downregulated HFRs was higher than upregulated ones for the synthesis of all metabolites in the NS vs N and NS vs S states, while the percentage of upregulated HFRs was higher than downregulated ones for the S vs N state (Fig. 3). There was a higher percentage of downregulated HFRs for the cholesterol (CHOL) and cholesteryl ester (CHOL\_ES) synthesis networks in NS vs S and for the superoxide anion (SOA) and H<sub>2</sub>O<sub>2</sub> synthesis networks in NS vs S (Fig. 3). This means that downregulated HFRs have a higher impact on NASH phenotypes.

260

Since the mouse strains AJ and PWD show phenotype specific to NS and S, respectively, and the mouse strain B6 show low S and NS phenotypes after feeding with the DDC diet (Materials and Methods) and thus we analyzed B6 separately. For mice, in NS vs N (the comparison associated to AJ) and B6 DDC vs N comparisons, 265 we found that the percentage of downregulated HFRs in the SOA and H<sub>2</sub>O<sub>2</sub> synthesis networks was higher than for upregulated ones (Fig. 3). This indicates synthesis networks of oxidative stress were found to be perturbed for AJ and B6 mice. Additionally, for the NS vs S in mice, an elevated percentage of upregulated HFRs was identified for the synthesis of the triacylglycerol (TAG), phosphatidylinositol 270 (PAIL), diacylglycerol (DAG), and phosphatidic acid (PA) metabolites, and a high percentage of downregulated HFRs was found for the alanine (ALA) and glutamine (GLU) synthesis networks (Fig. 3). This is consistent with the observation of lipid droplet perturbation in NASH and steatosis [24], and it further suggests that the synthesis of TAG, PAIL, DAG, and PA are important for lipid droplet formation.



275

Figure 3. **Quantification of up- and downregulated HFRs.** Symbols ‘h’ and ‘m’ in comparison’s labels represent human and mouse, respectively, and ‘up’ and ‘down’ indicate up- and downregulation, respectively. See the legend of the Table 3 for the association of mouse strains to the symbols N, NS, and S. “B6 up” and “B6 down” represent the up- and downregulated HFRs for the B6 mouse strain comparing the DDC vs control diet. We quantified the percentage of up- and downregulated HFR using a colored heat map.

280



For the synthesis of ROS, such as  $H_2O_2$ , the human NS vs N case presented with a  
285 high percentage of downregulated HFRs, while the same case in mice presented with  
a high percentage of upregulated HFRs (Fig. 3). The reactions of the PPP associated  
with genes TKT1, TKT2, GND, and TALA were identified as downregulated HFRs  
in the NS vs N human. However, for the NS vs N mouse and B6 DDC vs N these  
reactions were identified as upregulated. PPP is known as a source of nicotinamide  
290 adenine dinucleotide phosphate (NADPH) that prevents oxidative stress. Thus, the  
different patterns of regulation observed for these reactions can unbalance both  
NADPH concentration and  $H_2O_2$  synthesis. Interestingly it has been reported that the  
unbalance in NADPH production through PPP by either an under or over-production  
of NADPH may induce oxidative stress [25, 26].

295

The acetyl CoA C acetyltransferase (ACACT1m) and phosphoglycerate mutase  
reactions of the ceramide synthesis networks were identified as downregulated HFRs  
for the NS vs S and NS vs N in humans along with the NS vs N in mice. Thus,  
perturbation in the genes associated with these reactions can affect the ceramide level  
300 in NASH in both mice and humans. In line with this finding, a study suggested that  
deregulated ceramide production promotes liver injury and the development of NASH  
through disruption of endoplasmic calcium homeostasis, as well as through the  
inhibition of autophagy [27].

### 305 **Enriched MTs in human and mouse based on deregulated genes**

We want to identify deregulation at network level based on an enrichment analysis  
with deregulated enzymes or genes. A Minimal network (MiN) of a metabolic task

(MT) contains metabolites, active reactions and their associated enzymes in order to fulfill the MT. To identify deregulated MTs and their associated MiNs (see Materials and Methods section for definitions) in steatosis and NASH, we performed Minimal Network Enrichment Analysis (MiNEA) (see Materials and Methods section) using mouse and human expression data. We identified deregulated genes (S1 text; Table S1-S3), and for each set of deregulated genes, we estimated the  $p$  value.

Applying MiNEA to the NS vs N human case, we identified deregulated MiNs which corresponded to the synthesis of  $H_2O_2$ , PA, and TAG (Table 3; Table S4 & S5). We computed the percentage of significantly enriched MiNs from all generated alternatives of each MT and scored the deregulated MiNs using the **Alternative Minimal Network Frequency (AMiNF)**, which represents the percent of MiNs with enrichment in deregulated genes. For the NS vs N case in humans,  $H_2O_2$  synthesis had the highest AMiNF (AMiNF = 0.143) compared to all other deregulated MTs (Table 3), suggesting that the deregulation of  $H_2O_2$  metabolism and oxidative stress contributes to NASH. Indeed, a study has shown that a marked alteration of  $H_2O_2$  concentration can lead to different types of oxidative stress [25].

325

Table 3. The significantly deregulated MTs across all comparisons based on deregulated genes. The numbers in the table indicate Alternative Minimal Network Frequency (AMiNF). The mouse strains associated with N, NS, and S are described in the legend of Table 2. For the calculation of AMiNF  $p < 0.05$  was considered statistically significant.

		Mouse					
		N	NS	S	MT		
Human	N					Mouse	N
	NS			1	ALA		
				1	GLU		
				1	CRM		
				0.833	PS		
				0.714	PE		
					H2O2		
		0.143			TAG		
		0.016			PA		
		0.004					
S		0.7		CHOL		S	
		N	NS	S			
		Human					

330

When we applied MiNEA to NS vs S in humans, we found the most deregulated MiNs to be associated with CHOL synthesis (AMiNF = 0.7). Interestingly, TAG synthesis was also deregulated in NS vs N, suggesting that lipid droplet formation can  
 335 be perturbed in the NASH state but through different components.

In mice, we did not observe marked deregulation in NS vs N and S vs N, but we observed significant deregulation in NS vs S for the synthesis of ALA, GLU, ceramide (CRM), PS, and phosphatidylethanolamine (PE) (Table 3), suggesting that  
 340 these MTs are subject to different perturbations between the NASH and steatosis states.

In humans, MiNEA identified deregulation in the CHOL and TAG synthesis network, which are lipid droplet constituents. Similar deregulation was observed for the lipid  
 345 droplet constituents in mice through the PS and PE synthesis networks. Thus, because

MiNEA was applied with information on deregulated genes, we find that in both humans and mice, lipid droplet formation is perturbed but through different lipid constituents.

### 350 **Enriched MTs in humans and mice based on up- and downregulated reactions**

To identify whether a specific up- or downregulated MiN was associated with the NAFLD phenotypes, we performed MiNEA while enriching for up- and downregulated reactions separately (S1 Text).

355 In humans, we identified upregulated MiNs for the synthesis of PS and PE in N *vs* S (Table 4; Table S6 & S7). For the NS *vs* N and NS *vs* S cases, we identified downregulated MiNs associated with the synthesis of CHOL\_ES, CHOL, ALA, and GLU (Table 5; Table S8 & S9), but none were found to be upregulated. This implies that in NASH the synthesis of cholesterol and the glucogenic amino acids alanine and  
360 glutamine are downregulated. Interestingly, in the NS *vs* N case, only 3% of the cholesterol synthesis alternatives were significantly downregulated (Table 5). Although it is a relatively small frequency, it can provide important leads, such as deregulation in metabolic tasks and hypotheses that would have been missed by the commonly used pathway enrichment methods, such as [3, 5]. In general, these studies  
365 suggested that MiNEA can identify deregulated alternative MiNs under different conditions. For example, previous pathway analysis methods [3, 5] lack an enumeration of alternatives that would have been failed to identify cholesterol synthesis as downregulated in NS *vs* N (Table 5).

370 Table 4. The significantly deregulated MTs across all comparisons based on upregulated reactions. For a detailed understanding, see the legend of Table 2. Numbers of the table represents AMiNF and for the calculation of AMiNF  $p$  value  $< 0.005$  was used.

		Mouse					
		N	NS	S	MT		
Human	N			0.029	GLU	N	
	NS			1	PA	NS	
				1	TAG		
				1	DAG		
				0.979	PAIL		
				0.167	PS		
	S		0.429			PS	S
		0.167			PE		
		N	NS	S			
		Human					

375 In mice, the MiNs associated with the synthesis of ALA, CRM, PE, PS, CHOL\_ES, and GLU were identified as downregulated in NS vs N (Table 5). Interestingly, in S vs N, only the GLU synthesis network was upregulated (Table 4). For the NS vs S case, we found major deregulations. Specifically, we found five upregulated networks (PA, TAG, DAG, PAIL, and PS) and five downregulated (ALA, GLU, CRM, PS, and PE) synthesis networks. Within the alternative MiNs that are used for the synthesis of PS, we found some networks that were upregulated and some that were downregulated. Since there was a higher frequency of downregulated MiNs (AMiNF = 0.5) compared to upregulated ones (AMiNF = 0.167), we could hypothesize that the downregulation of PS synthesis is an important molecular mechanism specific to NASH. Here, if we use pathways enrichment methods [3, 5] then we cannot associate two different AMiNF score for a given MT due to the lack of alternative enumeration and would have been failed to give a higher confidence to downregulation of PS synthesis.

Table 5. The significantly deregulated MTs across all comparisons based on downregulated reactions.

390 For detail understanding see the legend of Table 4. For the AMiNF calculation  $p < 0.005$  was considered statistically significant.

		Mouse					
		N	NS	S	MT		
Human	N		1		ALA	N	
			1		CRM		
			0.857		PE		
			0.833		PS		
			0.62		CHOL ES		
			0.588		GLU		
	NS		0.034		CHOL ES	NS	
				1	ALA		
				1	GLU		
				1	CRM		
				0.5	PS		
			0.429	PE			
	S			1	CHOL ES	S	
				1	CHOL		
				0.128	ALA		
			0.004	GLU			
		N	NS	S			
		Human					

The cholesterol, glutamine, and alanine syntheses were found to be similarly  
 395 deregulated in humans and mice for the NS vs N case. In contrast, the PS synthesis  
 was differently deregulated between the two species (Table 4 and 5). Despite this  
 opposite response in the regulation of the PS synthesis between humans and mice, it  
 appears that the level of PS is perturbed in NASH. Overall this observation suggests  
 that different regulations in gene expression can translate to different effects on the  
 400 production of lipid droplet synthesis.

### **Integration of metabolite concentration data into a thermodynamically feasible metabolic model**

The integration of metabolomics can change the size and structure of the MiNs for  
405 MTs by eliminating thermodynamically infeasible reaction directionalities. In mice,  
there are available metabolic profiles for the N, NS, and B6 cases [1]. We integrated  
these profiles into the iMM1415 [11, 13] and found that, in this case, there are not  
marked changes in the MiNs and the result in the HFRs, which means we did not  
found a marked impact on the synthesis networks of NAFLD phenotypes. A further  
410 sensitivity analysis can guide future metabolomics studies to target metabolites that  
have a marked thermodynamic impact on a metabolic network [28]. Thus, the  
information content of such metabolite sets can be influential and more relevant to the  
phenotypes under study.

415 We applied MiNEA for the study of deregulated metabolic processes rather signaling  
and regulatory processed because metabolism is better characterized, as is shown with  
the increased availability of GEMS for many organisms. Despite the challenges in  
reconstructing constraint-based signaling networks, recently such reconstructions  
have started to become available. Signaling networks for the toll-like receptor (TLR)  
420 and epidermal growth factor receptor (EGFR) are now available, and MiNEA can be  
extended to include these networks [29, 30], meaning that MiNEA could easily be  
applied to the study of deregulation in signaling networks.

## Materials and Methods

425

### Microarray gene expression analysis of human liver samples

The microarray gene expression data pertaining to the three diagnostic groups (normal [N], Steatosis [S], and NASH [NS]) were collected from the ArrayExpress public repository for microarray data under the accession number E-MEXP-3291 [14]. We  
430 performed an analysis of the differentially expressed genes (DEGS) and a pairwise comparison between diagnostic groups. To control the false discovery rate at level of 0.05, multiple hypothesis testing was used [31].

### 435 Mice phenotypes after feeding a 3,5-diethoxycarbonyl-1,4-dihydrocollidine (DDC)-supplemented diet

NASH phenotypes can be reproduced in mouse models by treatment of chronic intoxication of with 3,5-diethoxycarbonyl-1,4-dihydrocollidine (DDC)-supplemented diet [1]. Three genetically different mouse strains AJ, B6, and PWD were fed with a  
440 DDC-supplemented diet. Under DDC-supplemented diet, the steatosis and NASH phenotypes were the most obvious in the PWD and AJ strains, respectively, while for the B6 strain, low levels of the steatosis and NASH phenotypes were observed [1].  
The B6 mouse strain showed less induction of the NASH phenotype than the AJ strain [1].

### 445 Association of mouse strains with human diagnoses groups



The normal (N) feeding group is associated with mice fed the control diet. All mice strains, AJ, B6 and PWD, were fed this diet and are labeled as N $\approx$ AJ, N $\approx$ PWD or N $\approx$ B6 to indicate the control diet and mouse strain. After feeding the DDC-supplemented diet, AJ and PWD showed high NASH phenotypes and high steatosis phenotypes, respectively [1]. Thus, under the DDC-supplemented diet, the symbol for steatosis (S) is associated with the PWD mouse strain and the symbol for NASH (NS) is associated the AJ mouse strain (with DDC-supplemented diet: NS $\approx$ AJ, S $\approx$ PWD). Thus, NS *vs* N represents the AJ mouse strain treated with the DDC diet *vs* the AJ mouse strain fed the control diet, S *vs* N represents the PWD mouse strain treated with the DDC diet *vs* the PWD mouse strain fed the control diet, and NS *vs* S represents the AJ mouse strain treated with the DDC diet *vs* the PWD mouse strain fed the DDC diet.

#### 460 **Metabolite concentration and RNA-seq gene expression data from mouse liver samples**

Metabolite concentration data and RNA-seq gene expression data from mouse liver samples of the three mouse strains (AJ, B6, and PWD) measured under the control and DDC-supplemented diet conditions were collected from the work of Pandey et al. [1]. The R-package “edgeR” [32] was used to identify the differentially expressed genes (DEGs) using three biological replicates of each mouse strain for both the control and the DDC-treated conditions. Strain-wise identification of the DEGs between the DDC-treated and control states was performed separately for each mouse strain. We used the Benjamini-Hochberg procedure implemented in edgeR to control the false discovery rate to a level of 0.05 [31].

470

### Formulation of metabolic tasks (MTs) based on steatosis and steatohepatitis (NASH) phenotypes

Steatosis can occur due to the accumulation of lipid droplets in the liver, which can subsequently lead to NASH upon liver inflammation [33]. Lipid droplets are

475 composed of many lipid metabolites [34, 35], which are summarized in Table 7.

Reactive oxygen species (ROS), such as the superoxide anion, can also damage hepatic membranes and play an important role in the development of NASH [36]. To study these various phenotypes and their role in the various forms of liver disease, the metabolites associated with their various phenotypes were examined. Table 7

480 summarizes the key metabolites associated with the major liver disease phenotypes: lipid droplets, liver inflammation, apoptosis, and oxidative stress. We investigate the synthesis networks of these metabolites using MiNEA, where we called the synthesis of a metabolite as a metabolic task (MT).

485 Table 7: Metabolites associated with NAFLD phenotypes. This provides a summary of key metabolites associated with the four phenotypes and their corresponding references. Metabolites are taken from the mouse metabolic model of Sigurdsson and colleagues [13].

Metabolites	Phenotype	Reference
<b>Diacylglycerol, phosphatidylethanolamine, cholesteryl esters, triacylglycerol, hexadecanoate, cholesterol, phosphatidic acid, phosphatidylinositol, phosphatidylserine</b>	Lipid droplets	[34, 35, 37]
<b>Ceramide</b>	Inflammation	[27, 37, 38]
<b>Alanine, glutamine, ceramide</b>	Apoptosis	[39]
<b>Superoxide anion, hydrogen peroxide</b>	Oxidative stress	[36]

### Minimal network enrichment analysis (MiNEA) algorithm

490 The cell can use various different pathways depending on the current state and conditions of the environment to fulfill its immediate metabolic needs. In order for MiNEA to identify deregulated alternative routes for a given MT between two conditions (Fig. 4), the inputs required are a metabolic model, a list of MTs, and gene

or protein expression data. In this work, we used a mouse GEM (iMM1415)  
495 reconstructed by Sigurdsson and colleagues [13], a list of MTs which are the synthesis  
of the metabolites shown in Table 7, and gene expression data of mouse and human  
liver samples [1, 14].

The MiNEA calculations begin by applying thermodynamic constraints to the model  
500 as described previously [11] to eliminate thermodynamically infeasible reactions,  
meaning that reactions from the metabolic network can carry fluxes only if  
thermodynamics allows [12] (Fig. 4; step 1). Then, it enumerates all the  
thermodynamically feasible minimal-size networks (MiNs) that are active for the  
given list of provided MTs (Fig. 4; step 2), and all of these MiNs are composed of  
505 reactions that carry non-zero flux. For this step, the following mixed-integer linear  
programming (MILP) problem is applied with the objective of minimizing the number  
of reactions that carry flux or maximizing the number of reactions that cannot carry  
flux, while enforcing that the network should be able to synthesize the metabolites  
listed in Table 1:

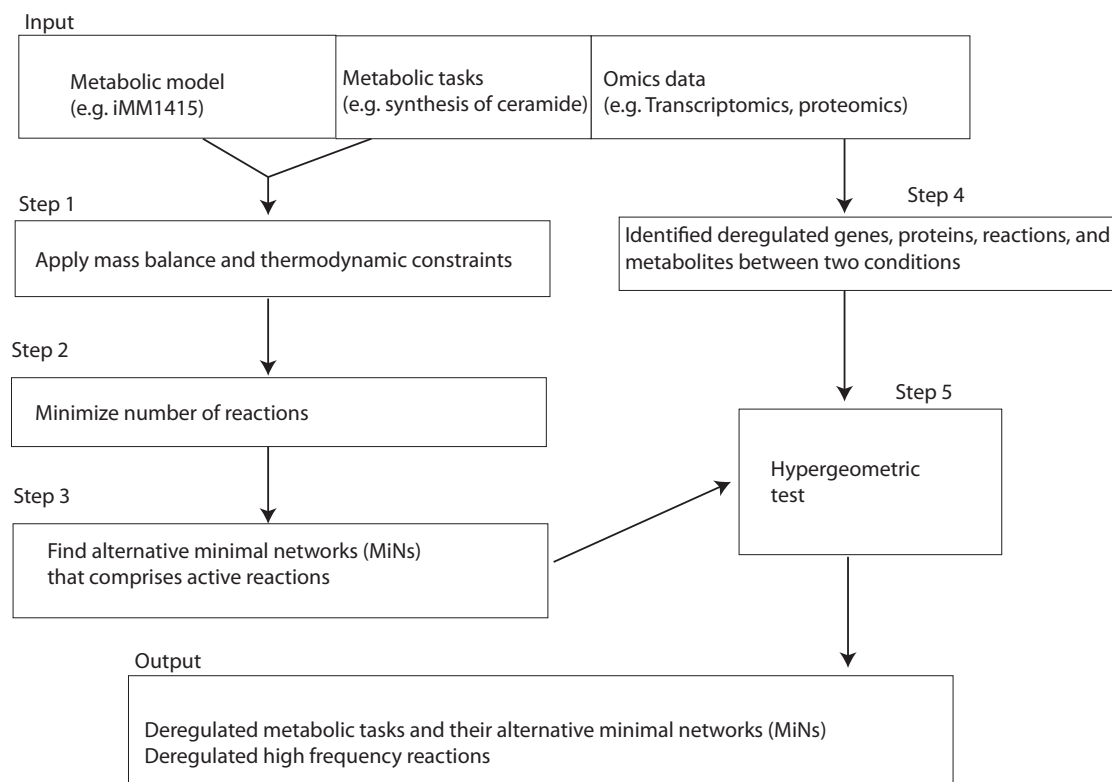
$$\begin{aligned} & \text{Maximize} \quad \sum_{i=1}^{R_{GEM}} z_{rxn,i} \\ & \text{S. } v = 0 \\ & 0 \leq v_i \leq v_{max,i} * Z_{rxn,i} \\ & v_{MT} \geq c * V_{MT,max} \end{aligned}$$

510  $R_{GEM}$  is the number of reactions in a GEM and  $z_{rxn,i}$  is a binary variable associated  
with reaction  $i$ . Reversible reactions were split into two forward reactions.  $V_{MT,max}$  is  
the maximum yield to produce a metabolite associated with a MT. The parameter  $c$  is

generates MiNs that allow for flexibility on the yield from the MT. We chose  $c = 0.8$  to allow for a yield of at least 80% of the maximum yield for the associated MT.

515

All alternative MiNs that are the minimum size (msize) for the synthesis of the metabolites in Table 1 are enumerated as described by Figueiredo and colleagues [40] (Fig. 4; step 3). One can enumerate many alternative networks larger than msize, such as msize+1 and msize+2. Each MiN is a subnetwork comprising a set of reactions, 520 metabolites, and reaction-associated genes through gene-protein-reaction (GPR) association. In step 4, the deregulated genes and proteins that differ between the two conditions under study are identified (Fig. 4). Finally, a hypergeometric test is performed on the sets of deregulated genes or deregulated reactions to identify the deregulated MiNs (Fig. 4; Step 5).



525

Figure 4. **Minimal network enrichment analysis (MiNEA) overview.** For each metabolic task (MT), MiNEA computationally enumerates MiNs that comprise active reactions for the MT. The inputs required for the MiNEA analysis are a genome-scale metabolic network (GEM), a given list of metabolic tasks (MTs), and transcriptomics data. Using this, alternative minimal networks (MiNs) are enumerated for MTs using a GEM. (steps 1–3). Transcriptomics data are used to identify differentially regulated genes between two conditions (step 4). To identify deregulated MTs and their associated MiNs, a hypergeometric test is performed with a set of deregulated genes (step 5).

535

#### Minimal network significance based on gene set and reaction set (Step 5)

The significance of the MiN based on the deregulated genes was calculated using the hypergeometric probability density function (P),

$$P(k,N,K,n) = \frac{\binom{K}{k} \binom{N-K}{n-k}}{\binom{N}{n}}, \quad (1)$$

540 where  $k$  and  $K$  are the numbers of deregulated genes and the number of total genes in a given MiN, respectively, and  $n$  and  $N$  are the total number of deregulated genes and the total number of genes of a metabolic network, respectively. In this context, the term “deregulated genes” is used for both up- or downregulated genes.

545 A reaction can have three states: upregulated, downregulated, or unregulated. The regulation of a reaction was determined based on its associated differentially expressed genes. According to this metric, a reaction was identified as upregulated or downregulated if the corresponding genes are only upregulated or only

downregulated, respectively. A reaction that is associated with a mixture of up- and  
550 downregulated genes is not characterized as regulated due to the inconsistency of  
gene expression. Up- and downregulated MiNs, which could contain various  
combination of up- or downregulated reactions, were identified in Step 4 of MiNEA  
(Fig. 1) based on the total number of up- or downregulated reactions, respectively. As  
expected, up- and downregulated MiNs comprise markedly high numbers of up- and  
555 downregulated reactions, respectively.

The significance of a MiN based on upregulated or downregulated reactions was  
calculated using multivariate Fisher's hypergeometric distribution. This method has  
been previously used for the selection of tissue-specific elementary modes using gene  
560 expression data [6]. To identify significantly upregulated MiNs in a given set of  
MiNs, we selected those MiNs that contained an elevated number of upregulated  
reactions and as few as possible downregulated ones, while for the identification of  
significantly downregulated MiNs, we selected MiNs with an elevated number of  
downregulated reactions and as few as possible upregulated ones.

565

Assume  $R$  and  $T$  are the total numbers of reactions in a GEM and a MiN, respectively,  
which can be decomposed as follows:

$$\begin{aligned} R &= R_{up} + R_{down} + R_{no} \\ T &= T_{up} + T_{down} + T_{no} \end{aligned} \quad (2)$$

570

$R_{up}$ ,  $R_{down}$ , and  $R_{no}$  represent the number of upregulated, downregulated, and  
unregulated reactions in a GEM, respectively. In the context of a MiN,  $T$  is the total

number of reactions, and  $T_{up}$ ,  $T_{down}$ , and  $T_{no}$  are the number of upregulated, downregulated, and unregulated reactions, respectively.

575 To consider a MiN as upregulated, we need to ensure that the pair  $(T_{up}, T_{down})$  in the MiN of  $T$  reactions cannot arise by chance in the context of the whole network. To obtain a better upregulation by chance, the  $p$  value was computed using equation (3). Note that an equal or better outcome than the pair  $(T_{up}, T_{down})$  satisfies two conditions: (i) the number of downregulated reactions is smaller than or equal to  $T_{down}$ , and (ii) 580 the number of upregulated reactions is greater than or equal to  $T_{up}$ , whereas the number of reactions in the MiN remains unchanged.

$$\text{p-value} = \sum_{i=T_{up}}^{\min(R_{up}, T)} \sum_{j=0}^{S_{down}} \frac{\binom{R_{up}}{i} \binom{R_{down}}{j} \binom{R_{no}}{S-i-j}}{\binom{R}{T}} \quad (3)$$

$i + j \leq T$

585 We can compute the  $p$  value for downregulated reactions the same way as for the upregulated reaction simply by changing **up** with **down** and **down** with **up** in the above equation.

### Deregulated percentage (DRP)

590 We calculated the deregulated percentage (DRP) for each MiN, which indicates the percentage of up- or downregulated genes in a MiN. For example, if a given MiN comprises 20 genes, and 5 genes of the MiN are deregulated, the then DRP is equal to 0.4. Since the active reactions in a MiN were classified as either up- or

downregulated, we calculated both upregulated percentage (UPR) and downregulated  
595 percentage (DnRP) for each MiN.

### **Alternative minimal network frequency (AMiNF)**

To extend the degree of confidence of the results from MiNEA, we introduced the  
alternative minimal network frequency (AMiNF) metric that determines the  
600 significantly deregulated percentage of MiNs for a given MT. If all the alternative  
MiNs for a MT are significantly deregulated, AMiNF takes the value 1. The AMiNF  
is 0 when none of the MiNs for a MT are significantly deregulated.

### **MiNEA implementation**

605 MiNEA was implemented using Matlab r2016a (The Mathworks, Natick, MA, USA),  
and MILP problems were solved using CPLEX solver (ILOG, Sunnyvale, CA, USA).  
The MILP gaps for all problems were converged to less than 0.05% in less than 2,400  
s. We also plan to make the MiNEA algorithm available as a tool for distribution to  
the community.

610



## Conclusions

In contrast to graph-based methods (GBM) and pathway enrichment analysis (PEA), the methodology introduced here (MiNEA) expands the notion of pathway into in a set of mass balanced subnetworks that can be used to understand the carbon, energy, and redox flows from precursor metabolites to target metabolites and complex metabolic tasks. One of the main advantages of MiNEA compared to PEA is that MiNEA attempts an enumeration of alternative minimal networks for each metabolic task (MT), which helps us understand and study the MT flexibility. Although a large number of alternative enumerations for a complex metabolic network can be time-consuming, once the enumeration is completed, MiNEA applies a statistical analysis that is fast and extracts additional information, such as pertaining to the deregulation of MTs or the deregulation in reaction hubs. As an example of the power of MiNEA, we identified deregulation in key metabolic network of the ceramide and hydrogen peroxide synthesis, of NASH in both humans and mice. We also identified similar deregulation in NASH for the cholesterol synthesis networks in humans and mice, and we found opposite deregulation for the phosphatidylserine synthesis network between humans and mice. MiNEA is highly applicable for the study of context- or condition-specific metabolism because using this one can identify synthesis networks for any given target metabolite and further can employ condition-specific transcriptomics, proteomics, and metabolomics data.

## 635 **Supplementary Tables**

Table S1. Differentially expressed genes for human and mouse expression data.

Table S2. Upregulated reactions for human and mouse data.

Table S3. Downregulated reactions for human and mouse data.

640 Table S4. Minimal network enrichment analysis based on differentially expressed genes. Description of the table heading is illustrated in ‘Symbol’ sheet of the excel file. Please see the Symbol sheet for the detail description for the tables: S5-S9.

Table S5. Minimal network enrichment analysis based on upregulated reactions.

Table S6. Minimal network enrichment analysis based on downregulated reactions.

Table S7. Marked deregulated minimal networks selected based on Table S4.

645 Table S8. Marked deregulated minimal networks selected based on Table S5.

Table S9. Marked deregulated minimal networks selected based on Table S6.

## **Acknowledgments**

650 VP and VH are supported by the RTD grant MalarX within SystemsX.ch, the Swiss Initiative for Systems Biology evaluated by the Swiss National Science Foundation. VP and VH are supported by the École polytechnique fédérale de Lausanne.

## **Author Contributions**

655 Conceptualization and methodology: VP and VH. Implementation of the method and data analysis: VP. Wrote the paper: VP and VH. Supervision: VH.

660

665

670

## Supporting Information

675

### Comparison of human and mouse genome-scale metabolic models

Mouse genome-scale metabolic model, iMM1415 [13] was reconstructed based on a  
680 human genome-scale metabolic model (Recon 1, [41]). Sigurdsson and colleagues  
found that the mammalian organism with the highest number of genes homologous to  
Recon 1 genes was the mouse (*Mus musculus*) (1,415 genes, 97%). We compared  
iMM1415 and Recon1 and found that the iMM1415 shares 98% of reactions with  
Recon1 and in the remaining 2 % reactions more than 1.5 % reactions were associated  
685 to the extracellular transport mechanism. This suggests that Recon1 and iMM1415  
have very similar metabolism.

### 690 Identification and analysis of deregulated genes and reactions in human and mouse liver samples

To understand the nonalcoholic fatty liver disease (NAFLD) and the difference of the  
disease in mouse and human we collected human expression data from the three  
diagnosis groups: normal (N), steatosis (S), and nonalcoholic steatohepatitis (NS), and  
695 mouse expression data from the control and DDC-supplement diet conditions for the  
three mouse strains: AJ, B6 and PWD (*Materials and Methods*). These data were  
referred as *human expression data* and *mouse expression data* throughout this section.

## 700 **Differential expressed genes in human liver samples**

To understand the NAFLD physiology we analyzed the differentially expressed genes (DEGs) in the *human expression data*. Out of the 1415 metabolic genes in iMM1415 [13], we identified 29, 484 and 363 DEGs in S vs N, NS vs S, and NS vs N, respectively (Fig. S1 upper panel and Table S1). Only 29 DEGs between S vs N may  
705 explain that the metabolic state was very similar between normal and steatosis. The total number of DEGs between NS vs N and NS vs S were much higher than S vs N, and thus suggest a more pronounced alteration of the metabolic state of nonalcoholic steatohepatitis compared to steatosis and normal. Furthermore, the number of downregulated genes is greater than the number of upregulated genes in exclusively  
710 NS vs N and NS vs S (Fig. S1), suggesting that, in human NASH, the perturbation leading to the metabolic state that characterizes it, is reached by downregulated genes.

## **Differential expressed genes in mouse liver samples**

We analyzed the *mouse expression data* for DEGs as described in the *material and*  
715 *methods* section. Out of the 1415 metabolic genes in iMM1415, the total number of DEGs between control and DDC-supplemented diet was similar across all strains with 247, 248 and 221 for AJ, B6 and PWD, respectively (Fig. S2 lower panel). Here, AJ and PWD are associated with NS vs N and S vs N, respectively (see materials and methods). Many up- and down-regulated genes were strain-specific. The number of  
720 up- and down-regulated strain-specific gene pairs for AJ, B6, and PWD were (48, 20), (33, 29), and (39, 62), respectively (Fig. S1 lower panel and Table S1). The number of

up-regulated genes was greater than the number of down-regulated genes for the AJ strain, while for the PWD strain the opposite was true. Numbers of up- and down-regulated genes were very similar for the B6 strain. Interestingly, the observed differences qualitatively correlate with the strains' phenotypes: steatohepatitis phenotypes were observed high, low, and unspecific for the AJ, B6, and PWD mouse strains, respectively [1].

We identified deregulated genes from the AJ under the DDC-supplemented diet vs PWD under DDC-supplemented diet and this is associated with NS vs S (see materials and methods). For this comparison we identified 191 and 100 up- and down-regulated genes, respectively (Table S1).

### **Up- and downregulated reactions in human and mouse**

A reaction is marked as down regulated if genes associated to the reaction is down-regulated and if genes associated to the reaction is upregulated then the reaction is called as upregulated. If a reaction is associated to genes with mix-regulation (up- and down-regulation) then the reaction is not marked with up- or down-regulated.

Regulations of reactions are computed for the *human expression data* and *mouse expression data* (Fig. S2).

Only for NS vs S, the number of down-regulated reactions was higher than the number of up-regulated reactions (Fig. S2 upper panel), and a similar trend was observed for the number of up- and down-regulated genes (Fig. S1 upper panel and Tables S2-S3). The number of reactions up- and down-regulated was similar for NS vs N, but the number of down-regulated genes was higher than the number of

upregulated genes (Fig. S1 and S2 upper panels). For S vs N, the number of down-regulated reactions was higher than the number of up-regulated reactions; however, the numbers of up- and down-regulated genes for these conditions were similar (Fig. S1 and S2 upper panels). This observation indicates that the reaction regulation (RR) is not always in agreement with the gene regulation, which results from the dependency on the gene association set rather than a single gene.

For the AJ and B6 mouse strains, we found that the number of upregulated reactions was higher than the number of downregulated reactions, whereas for the PWD mouse strain the inverse was observed (Fig. S2 lower panel). For AJ and PWD, the gene regulation followed the same trend for reaction regulation, whereas for B6 the numbers of up- and down-regulated genes and reactions remained comparable (Fig. S1).

To represent as the human NS vs S we compared mice that have shown high nonalcoholic steatohepatitis (NS) phenotypes (AJ mice fed with the DDC-supplemented diet) to mice that have shown high steatosis (S) phenotypes (PWD mice fed with DDC supplemented diet). For the NS vs S in mice (see materials and methods; AJ DDC vs PWD DDC) we identified 459 and 191 up- and down-regulated reactions, respectively (Table S 2-S3).

765

### Supplementary Figures

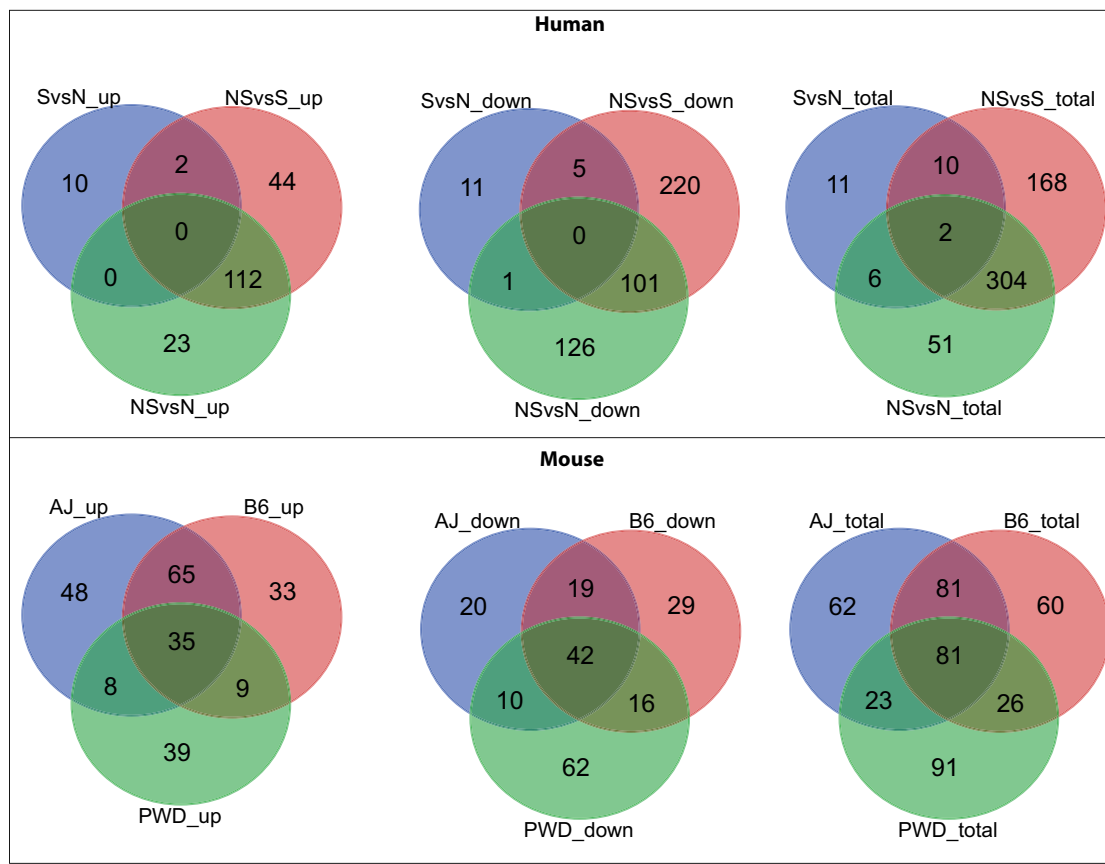
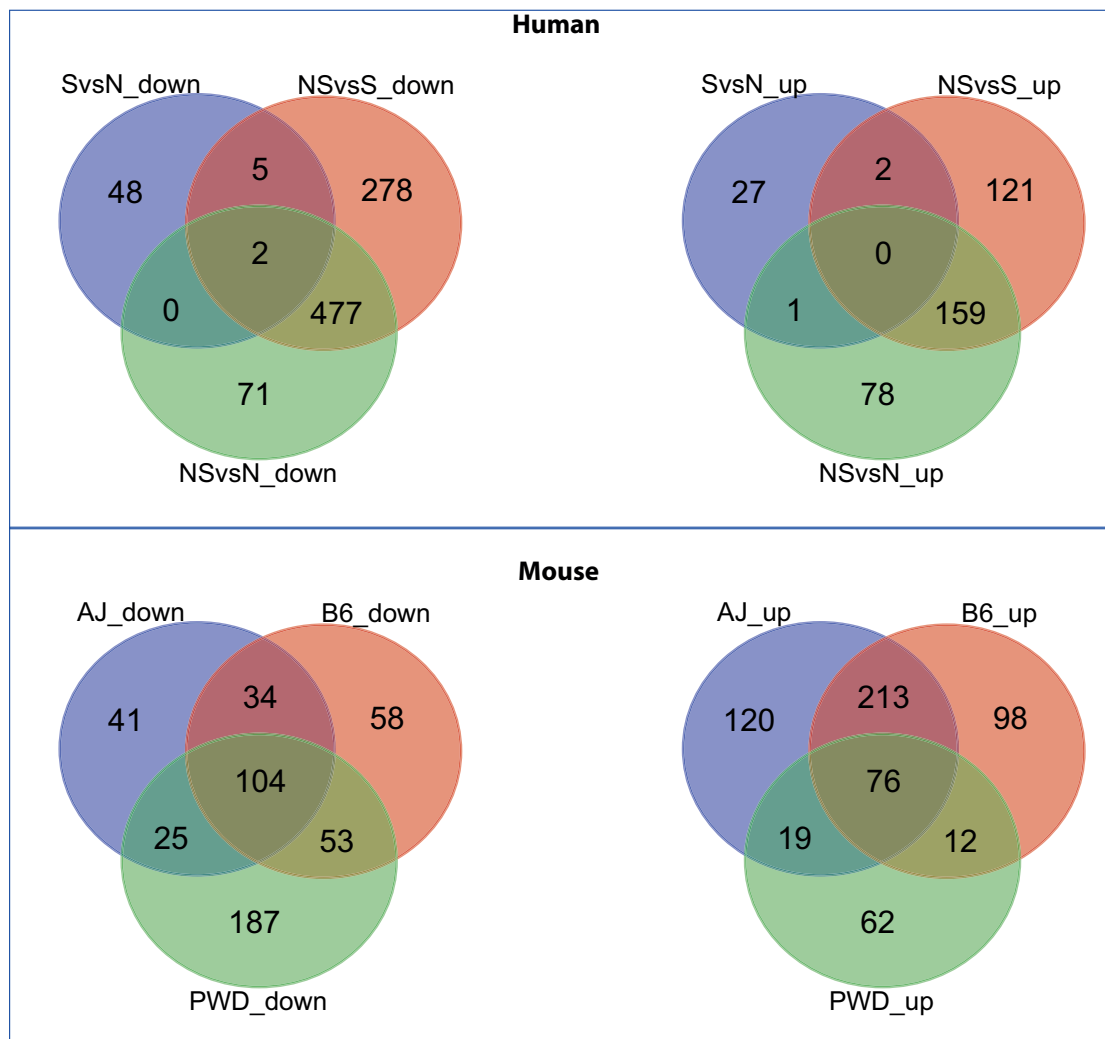


Figure S1. Venn diagram of differentially expressed genes of human and mouse liver samples. Upper and lower panels represent human and mouse, respectively.



770

Figure S2. Venn diagram of up- and downregulated reactions of the iMM1415 in human and mouse liver samples. The reaction regulation metric was used to identify up- and downregulated reactions.



## References

775

1. Pandey V, Sultan M, Kashofer K, Ralser M, Amstislavskiy V, Starmann J, et al. Comparative Analysis and Modeling of the Severity of Steatohepatitis in DDC-Treated Mouse Strains. *PLoS ONE*. 2014;9:e111006. doi: 10.1371/journal.pone.0111006.

780

2. ME R. Nonalcoholic fatty liver disease: A systematic review. *JAMA*. 2015;313:2263-73. doi: 10.1001/jama.2015.5370.

3. Kamburov A, Pentchev K, Galicka H, Wierling C, Lehrach H, Herwig R. ConsensusPathDB: toward a more complete picture of cell biology. *Nucleic Acids Research*. 2011;39:D712-D7. doi: 10.1093/nar/gkq1156.

785

4. Våremo L, Nielsen J, Nookaew I. Enriching the gene set analysis of genome-wide data by incorporating directionality of gene expression and combining statistical hypotheses and methods. *Nucleic Acids Research*. 2013;41:4378-91. doi: 10.1093/nar/gkt111.

790

5. Kamburov A, Cavill R, Ebbels TMD, Herwig R, Keun HC. Integrated pathway-level analysis of transcriptomics and metabolomics data with IMPaLA. *Bioinformatics*. 2011;27:2917-8. doi: 10.1093/bioinformatics/btr499.

6. Rezola A, Pey J, de Figueiredo LF, Podhorski A, Schuster S, Rubio A, et al. Selection of human tissue-specific elementary flux modes using gene expression data. *Bioinformatics (Oxford, England)*. 2013;29:2009-16. doi:

795

10.1093/bioinformatics/btt328.

7. He H, Lin DD, Zhang JG, Wang YP, Deng HW. Comparison of statistical methods for subnetwork detection in the integration of gene expression and protein interaction network. *Bmc Bioinformatics*. 2017;18:6. doi: 10.1186/s12859-017-1567-2. PubMed PMID: WOS:000397507800004.

800

8. Glaab E, Baudot A, Krasnogor N, Schneider R, Valencia A. EnrichNet: network-based gene set enrichment analysis. *Bioinformatics*. 2012;28:i451-i7. doi: 10.1093/bioinformatics/bts389.

9. Liu L, Wei JM, Ruan JH. Pathway Enrichment Analysis with Networks. *Genes*. 2017;8(10):12. doi: 10.3390/genes8100246. PubMed PMID:

805

WOS:000414862300011.

10. Orth JD, Thiele I, Palsson BØ. What is flux balance analysis? *Nature Biotechnology*. 2010;28:245-8. doi: 10.1038/nbt.1614.

11. Henry CS, Broadbelt LJ, Hatzimanikatis V. Thermodynamics-Based Metabolic Flux Analysis. *Biophysical Journal*. 2007;92:1792-805.

810

12. Ataman M, Hatzimanikatis V. Heading in the right direction: thermodynamics-based network analysis and pathway engineering. *Current opinion in biotechnology*. 2015;36:176-82. doi: 10.1016/j.copbio.2015.08.021.

13. Sigurdsson MI, Jamshidi N, Steingrimsson E, Thiele I, Palsson BØ. A detailed genome-wide reconstruction of mouse metabolism based on human

815

Recon 1. *BMC systems biology*. 2010;4:140. doi: 10.1186/1752-0509-4-140.

14. Lake AD, Novak P, Fisher CD, Jackson JP, Hardwick RN, Billheimer DD, et al. Analysis of global and absorption, distribution, metabolism, and elimination

- gene expression in the progressive stages of human nonalcoholic fatty liver disease. *Drug Metabolism and Disposition: The Biological Fate of Chemicals*. 2011;39:1954-60. doi: 10.1124/dmd.111.040592.
- 820 15. Ordonez R, Fernandez A, Prieto-Dominguez N, Martinez L, Garcia-Ruiz C, Fernandez-Checa JC, et al. Ceramide metabolism regulates autophagy and apoptotic cell death induced by melatonin in liver cancer cells. *J Pineal Res*. 2015;59(2):178-89. doi: 10.1111/jpi.12249. PubMed PMID: 825 WOS:000358958000004.
16. Lee AY, Lee JW, Kim JE, Mock HJ, Park S, Kim S, et al. Dihydroceramide is a key metabolite that regulates autophagy and promotes fibrosis in hepatic steatosis model. *Biochemical and Biophysical Research Communications*. 2017;494(3-4):460-9. doi: 10.1016/j.bbrc.2017.10.110. PubMed PMID: 830 WOS:000416393300006.
17. Smith CA. *Physiology of the bacterial cell. A molecular approach*. By F C Neidhardt, J L Ingraham and M Schaechter. pp 507. Sinauer associates, Sunderland, MA. 1990. \$43.95 ISBN 0-87893-608-4. *Biochemical Education*. 20:124-5. doi: 10.1016/0307-4412(92)90139-D.
- 835 18. Marchesini G, Bugianesi E, Forlani G, Cerrelli F, Lenzi M, Manini R, et al. Nonalcoholic fatty liver, steatohepatitis, and the metabolic syndrome. *Hepatology (Baltimore, Md)*. 2003;37:917-23. doi: 10.1053/jhep.2003.50161.
19. Terzer M, Maynard ND, Covert MW, Stelling J. Genome-scale metabolic networks. *Wiley Interdiscip Rev-Syst Biol*. 2009;1(3):285-97. doi: 840 10.1002/wsbm.37. PubMed PMID: WOS:000283710600002.
20. King ZA, Lu J, Drager A, Miller P, Federowicz S, Lerman JA, et al. BiGG Models: A platform for integrating, standardizing and sharing genome-scale models. *Nucleic Acids Research*. 2016;44(D1):D515-D22. doi: 10.1093/nar/gkv1049. PubMed PMID: WOS:000371261700071.
- 845 21. Ataman M, Hatzimanikatis V. lumpGEM: Systematic generation of subnetworks and elementally balanced lumped reactions for the biosynthesis of target metabolites. *Plos Computational Biology*. 2017;13(7):21. doi: 10.1371/journal.pcbi.1005513. PubMed PMID: WOS:000406619800006.
22. Ekman D, Light S, Bjorklund AK, Elofsson A. What properties characterize the hub proteins of the protein-protein interaction network of *Saccharomyces cerevisiae*? *Genome Biology*. 2006;7(6):13. doi: 10.1186/gb-2006-7-6-r45. 850 PubMed PMID: WOS:000241243800010.
23. Hu G, Wu ZH, Uversky VN, Kurgan L. Functional Analysis of Human Hub Proteins and Their Interactors Involved in the Intrinsic Disorder-Enriched Interactions. *Int J Mol Sci*. 2017;18(12):40. doi: 10.3390/ijms18122761. PubMed 855 PMID: WOS:000418896700261.
24. Gluchowski NL, Becuwe M, Walther TC, Farese RV. Lipid droplets and liver disease: from basic biology to clinical implications. *Nat Rev Gastroenterol Hepatol*. 2017;14(6):343-55. doi: 10.1038/nrgastro.2017.32. PubMed PMID: 860 WOS:000402116200006.
25. Sies H. Hydrogen peroxide as a central redox signaling molecule in physiological oxidative stress: Oxidative eustress. *Redox Biol*. 2017;11:613-9. doi: 10.1016/j.redox.2016.12.035. PubMed PMID: WOS:000398212000060.

26. Christodoulou D, Link H, Fuhrer T, Kochanowski K, Gerosa L, Sauer U.  
865 Reserve Flux Capacity in the Pentose Phosphate Pathway Enables *Escherichia coli*'s Rapid Response to Oxidative Stress. *Cell Systems*. 2018;0. doi: 10.1016/j.cels.2018.04.009.
27. Musso G, Cassader M, Gambino R. Non-alcoholic steatohepatitis: emerging molecular targets and therapeutic strategies. *Nature Reviews Drug Discovery*.  
870 2016;advance online publication. doi: 10.1038/nrd.2015.3.
28. Kiparissides A, Hatzimanikatis V. Thermodynamics-based Metabolite Sensitivity Analysis in metabolic networks. *Metabolic Engineering*. 2017;39:117-27. doi: 10.1016/j.ymben.2016.11.006. PubMed PMID: WOS:000392565200013.
29. Li F, Thiele I, Jamshidi N, Palsson BØ. Identification of Potential Pathway  
875 Mediation Targets in Toll-like Receptor Signaling. *PLOS Comput Biol*. 2009;5:e1000292. doi: 10.1371/journal.pcbi.1000292.
30. Choudhary KS, Rohatgi N, Halldorsson S, Briem E, Gudjonsson T, Gudmundsson S, et al. EGFR Signal-Network Reconstruction Demonstrates Metabolic Crosstalk in EMT. *PLoS computational biology*. 2016;12:e1004924.  
880 doi: 10.1371/journal.pcbi.1004924.
31. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B (Methodological)*. 1995;57:289-300.
32. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for  
885 differential expression analysis of digital gene expression data. *Bioinformatics* (Oxford, England). 2010;26:139-40. doi: 10.1093/bioinformatics/btp616.
33. Day CP, James OFW. Steatohepatitis: A tale of two "hits"? *Gastroenterology*. 1998;114(4):842-5. doi: 10.1016/s0016-5085(98)70599-2. PubMed PMID: WOS:000072770100030.
- 890 34. Guo Y, Cordes KR, Farese RV, Walther TC. Lipid droplets at a glance. *Journal of Cell Science*. 2009;122:749-52. doi: 10.1242/jcs.037630.
35. Bartz R, Li W-H, Venables B, Zehmer JK, Roth MR, Welti R, et al. Lipidomics reveals that adiposomes store ether lipids and mediate phospholipid traffic. *Journal of Lipid Research*. 2007;48:837-47. doi: 10.1194/jlr.M600413-  
895 JLR200.
36. Koek GH, Liedorp PR, Bast A. The role of oxidative stress in non-alcoholic steatohepatitis. *Clin Chim Acta*. 2011;412(15-16):1297-305. doi: 10.1016/j.cca.2011.04.013. PubMed PMID: WOS:000291917500002.
37. Yetukuri L, Katajamaa M, Medina-Gomez G, Seppanen-Laakso T, Vidal-Puig A, Oresic M. Bioinformatics strategies for lipidomics analysis: characterization of obesity related hepatic steatosis. *Bmc Systems Biology*. 2007;1:15. doi: 10.1186/1752-0509-1-12. PubMed PMID: WOS:000250001800001.
- 900 38. Srivastava S, Chan C. Hydrogen peroxide and hydroxyl radicals mediate palmitate-induced cytotoxicity to hepatoma cells: Relation to mitochondrial permeability transition. *Free Radic Res*. 2007;41(1):38-49. doi: 10.1080/10715760600943900. PubMed PMID: WOS:000242799200005.
39. Halama A, Riesen N, Moller G, de Angelis MH, Adamski J. Identification of biomarkers for apoptosis in cancer cell lines using metabolomics: tools for

- 910 individualized medicine. *J Intern Med.* 2013;274(5):425-39. doi:  
10.1111/joim.12117. PubMed PMID: WOS:000328284700004.
40. de Figueiredo LF, Podhorski A, Rubio A, Kaleta C, Beasley JE, Schuster S, et al. Computing the shortest elementary flux modes in genome-scale metabolic networks. *Bioinformatics (Oxford, England).* 2009;25:3158-65. doi:  
915 10.1093/bioinformatics/btp564.
41. Duarte NC, Becker SA, Jamshidi N, Thiele I, Mo ML, Vo TD, et al. Global reconstruction of the human metabolic network based on genomic and bibliomic data. *Proceedings of the National Academy of Sciences of the United States of America.* 2007;104(6):1777-82. doi: 10.1073/pnas.0610772104. PubMed PMID:  
920 WOS:000244127900011.