

ARTICLE

Received 30 May 2013 | Accepted 10 Dec 2013 | Published 14 Jan 2014

DOI: 10.1038/ncomms4083

# Genome-scale metabolic modelling of hepatocytes reveals serine deficiency in patients with non-alcoholic fatty liver disease

Adil Mardinoglu<sup>1,\*</sup>, Rasmus Agren<sup>1,\*</sup>, Caroline Kampf<sup>2</sup>, Anna Asplund<sup>2</sup>, Mathias Uhlen<sup>3,4</sup> & Jens Nielsen<sup>1,4</sup>

Several liver disorders result from perturbations in the metabolism of hepatocytes, and their underlying mechanisms can be outlined through the use of genome-scale metabolic models (GEMs). Here we reconstruct a consensus GEM for hepatocytes, which we call *iHepatocytes2322*, that extends previous models by including an extensive description of lipid metabolism. We build *iHepatocytes2322* using Human Metabolic Reaction 2.0 database and proteomics data in Human Protein Atlas, which experimentally validates the incorporated reactions. The reconstruction process enables improved annotation of the proteomics data using the network centric view of *iHepatocytes2322*. We then use *iHepatocytes2322* to analyse transcriptomics data obtained from patients with non-alcoholic fatty liver disease. We show that blood concentrations of chondroitin and heparan sulphates are suitable for diagnosing non-alcoholic steatohepatitis and for the staging of non-alcoholic fatty liver disease. Furthermore, we observe serine deficiency in patients with NASH and identify *PSPH*, *SHMT1* and *BCAT1* as potential therapeutic targets for the treatment of non-alcoholic steatohepatitis.

<sup>1</sup>Department of Chemical and Biological Engineering, Chalmers University of Technology, Kamivangenvägen 10, Gothenburg SE-412 96, Sweden. <sup>2</sup>Department of Immunology, Genetics and Pathology, Science for Life Laboratory, Uppsala University, Uppsala SE-751 85, Sweden. <sup>3</sup>Department of Proteomics, KTH—Royal Institute of Technology, Stockholm SE-106 91, Sweden. <sup>4</sup>Science for Life Laboratory, KTH—Royal Institute of Technology, Stockholm SE-171 21, Sweden.

\* These authors contributed equally to this work. Correspondence and requests for materials should be addressed to J.N. (email: nielsenj@chalmers.se).

Hepatocytes have a wide range of physiological functions, including production of bile and hormones, removal of toxic substances, homeostatic regulation of the plasma constituents and synthesis of most plasma proteins<sup>1</sup>. The hepatocytes, the most metabolically active cell types in human, play a major role in overall human metabolism. Deficiency or alterations in the metabolism of hepatocytes can lead to complicated disorders such as hepatitis, non-alcoholic fatty liver disease (NAFLD), cirrhosis and liver cancer, which are serious threats to public health<sup>2</sup>. NAFLD is considered as the hepatic manifestation of obesity and metabolic syndrome, and encompasses a spectrum of pathological changes; ranging from simple fatty liver (FL; steatosis) to non-alcoholic steatohepatitis (NASH)<sup>3</sup>.

Even though it is well known that lipid accumulation in the liver is a hallmark of NAFLD<sup>4</sup>, the underlying mechanisms leading to steatosis and further transition to NASH still remain elusive. It is therefore difficult to track the onset and progression or to diagnose and design effective therapeutic techniques. The adverse outcomes of this pathology may possibly be prevented once the molecular mechanisms involved in the metabolism of hepatocytes are deciphered<sup>5</sup>. However, this requires understanding of the coordinated behaviour of a very large number of interconnected metabolic reactions and metabolites. Relating this behaviour with disease and patients have been a major focus in biomedicine<sup>6</sup>. A systems biology approach, based on the employment of genome-scale metabolic models (GEMs), can be used to extend our understanding of these molecular mechanisms, which in turn may enable future therapeutic discoveries<sup>7–10</sup>.

GEMs represent the current knowledge of metabolism generated through the integration of genetic and biochemical studies coupled with cellular, physiological and clinical data<sup>11</sup>. Several generic (non-cell type-specific) GEMs for human metabolism have been previously constructed<sup>12–16</sup>. However, neither of these generic networks contain extensive lipid metabolism, which is necessary to study the effect of lipids on the underlying molecular mechanism of NAFLD. Recently, a large-scale GEM for adipocytes, *iAdipocytes1809*, with a strong focus on lipid metabolism was presented<sup>17</sup>, and this model can provide a base for further integration of lipid metabolism into generic networks.

There is currently no efficient treatment for NASH<sup>18</sup> and new therapeutic approaches are in great demand. This study represents an attempt to rationally identify biomarkers and therapeutic targets using GEM modelling. To reconstruct a high-quality model for hepatocytes, we combine clinical, biochemical and genetic studies such as expression, localization and functional characteristics of the proteins. We first significantly expand the content of our Human Metabolic Reaction (HMR) database by including extensive lipid metabolism and generate HMR 2.0 database. This represents an important step forward, as lipids have major effect on the development of NAFLD and other metabolic diseases<sup>19</sup>. Second, we reconstruct a consensus GEM for hepatocytes, *iHepatocytes2322*, by using the HMR 2.0 database and large-scale proteomics data. We also merge previously published hepatocyte models to cover the entire known metabolic functions of hepatocytes and incorporated additional clinical data (for example, liver tissue and plasma fatty acid (FA) contents in lipid structures). During the reconstruction process, we reevaluate the hepatocyte proteomics data after identifying proteins that are included in the model to ensure network connectivity, but are assessed as absent in hepatocytes in the Human Protein Atlas (HPA, <http://www.proteinatlas.org>)<sup>20</sup>. Finally, we employ *iHepatocytes2322* for the analysis of differential gene expression data from liver tissues of subject groups with NAFLD. This lead

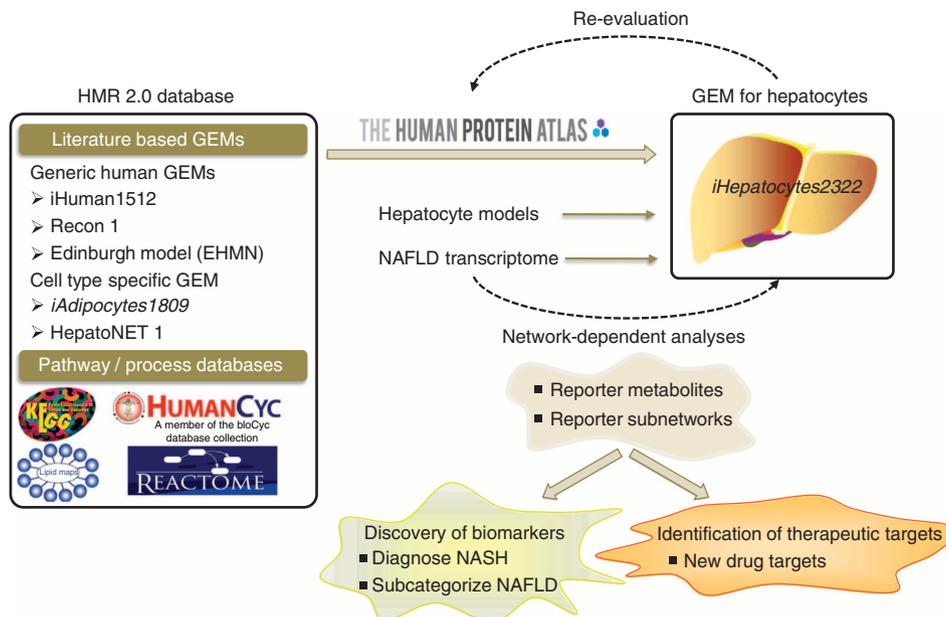
to new insights into the molecular mechanisms involved in NASH, which are used for the identification of potential metabolic biomarkers and therapeutic targets for treatment of NASH (Fig. 1).

## Results

**HMR 2.0 database.** To provide a resource for automated and semi-automated reconstruction of cell-type-specific GEMs, we previously constructed the HMR database<sup>15</sup>. This comprehensive database, together with the INIT (Integrative Network Inference for Tissues) algorithm, have been employed for automated generation of cell-type-specific GEMs<sup>15</sup>. These models form the basis for the Human Metabolic Atlas (<http://www.metabolicatlas.org>), which is a web-based resource for human metabolism. Here we expanded the HMR database by incorporating extensive lipid metabolism, which accounts for individual FAs rather than relying on generic FAs pool metabolite. The generated HMR 2.0 database is formulated using 59 FAs (Supplementary Table 1), which enables mapping and integration of lipidomics data. Integration of extensive lipid metabolism (for example, formation of lipid droplets (LDs) and lipoproteins) may allow not only for understanding the contribution of lipids to the development of diseases but also allow for study of the relationship between lipid metabolism and cellular molecular mechanisms<sup>17</sup>.

Reactions are included in the HMR 2.0 database depending on evidence from previously published models and databases (Supplementary Table 2) or on the availability of specific experimental evidence for the occurrence of the reaction. The reaction–gene associations of the generic human network were improved based on the publically available resources and literature review. The HMR 2.0 database is the largest biochemical reaction database for human metabolism in terms of number of reactions/genes/metabolites, as well as in terms of covering most parts of metabolism.

**Consensus GEM for hepatocytes and improved proteome annotation.** Cell-type-specific GEMs can be employed for the analysis of high-throughput patient (-omics) data, simulation of the metabolic differences under health and disease states, and eventually for predicting the cellular phenotype<sup>10</sup>. Previously, several GEMs for hepatocytes, including *HepatoNET 1* (ref. 1), *iLJ1046* (ref. 21), *iAB676* (ref. 22) and *iHepatocyte1154* (ref. 15) have been reconstructed. Here we generated a consensus GEM for hepatocytes, *iHepatocytes2322*, based on proteomics data (Supplementary Data 1) and the updated HMR 2.0 database. *iHepatocytes2322* contains all of the protein-coding genes and associated reactions in previously published liver models (Fig. 2a). In addition to the proteomics data and previous models, protein-coding genes are also included in *iHepatocytes2322* based on transcriptomics data and connectivity (Fig. 2b). Reactions and associated proteins were assigned into eight different compartments following our HMR database standard based on the subcellular localization of the proteins in HPA and Uniprot. The protein localization information in HPA and Uniprot were assigned to relevant compartments in the HMR 2.0 database (Supplementary Data 2). A confidence score for each protein was calculated based on the availability of knowledge in HPA and Uniprot (Supplementary Data 3). Furthermore, the connectivity in the model was checked carefully, such that all metabolites consumed in one reaction should be able to be produced by another reaction or they should be taken up from the plasma. Finally, additional clinical data for plasma and hepatocyte lipid concentrations for individual FAs were incorporated into the model (Supplementary Data 4).



**Figure 1 | Effective therapeutic approach through GEM modelling.** Schematic illustration of how a consensus GEM for hepatocytes, *iHepatocytes2322*, may contribute to the development of effective therapeutic approaches for NAFLD patients. The HMR 2.0 database was constructed through the use of previously published GEMs and pathway databases, including KEGG, HumanCyc, Reactome and LIPIDMAPS Lipidomics Gateway. Elements of lipid metabolism were included in the HMR database to understand the effect of the lipids and their interactions during the appearance of NAFLD. The HMR 2.0 database was used for reconstruction of *iHepatocytes2322* based on proteomics data in the HPA, transcriptomics data in NAFLD patients and previously published hepatocytes models. During the reconstruction process, *iHepatocytes2322* was employed for the improvement of proteomics data through identification and re-annotation of putative false-negative proteins. The resulting GEM was used for the analysis of clinical data obtained from NAFLD patients to investigate the alterations in their hepatocyte metabolism and eventually for the discovery of biomarkers and identifying therapeutic targets. Through our systems biology-based analysis, potential biomarkers for diagnosing NASH and for subcategorizing NAFLD patients were discovered. Furthermore, a list of candidate therapeutic targets was identified to develop efficient treatments for NASH patients.

*iHepatocytes2322* differs from previously published hepatocyte GEMs primarily in terms of coverage in lipid metabolism. Among the new lipid-related functions are uptake of the remnants of lipoproteins (chylomicrons, very-low-density lipoprotein (VLDL), low-density lipoproteins and high-density lipoproteins), the formation and degradation of LDs and secretion of synthesized lipoproteins (VLDL, low-density lipoprotein and high-density lipoprotein; Fig. 2c).

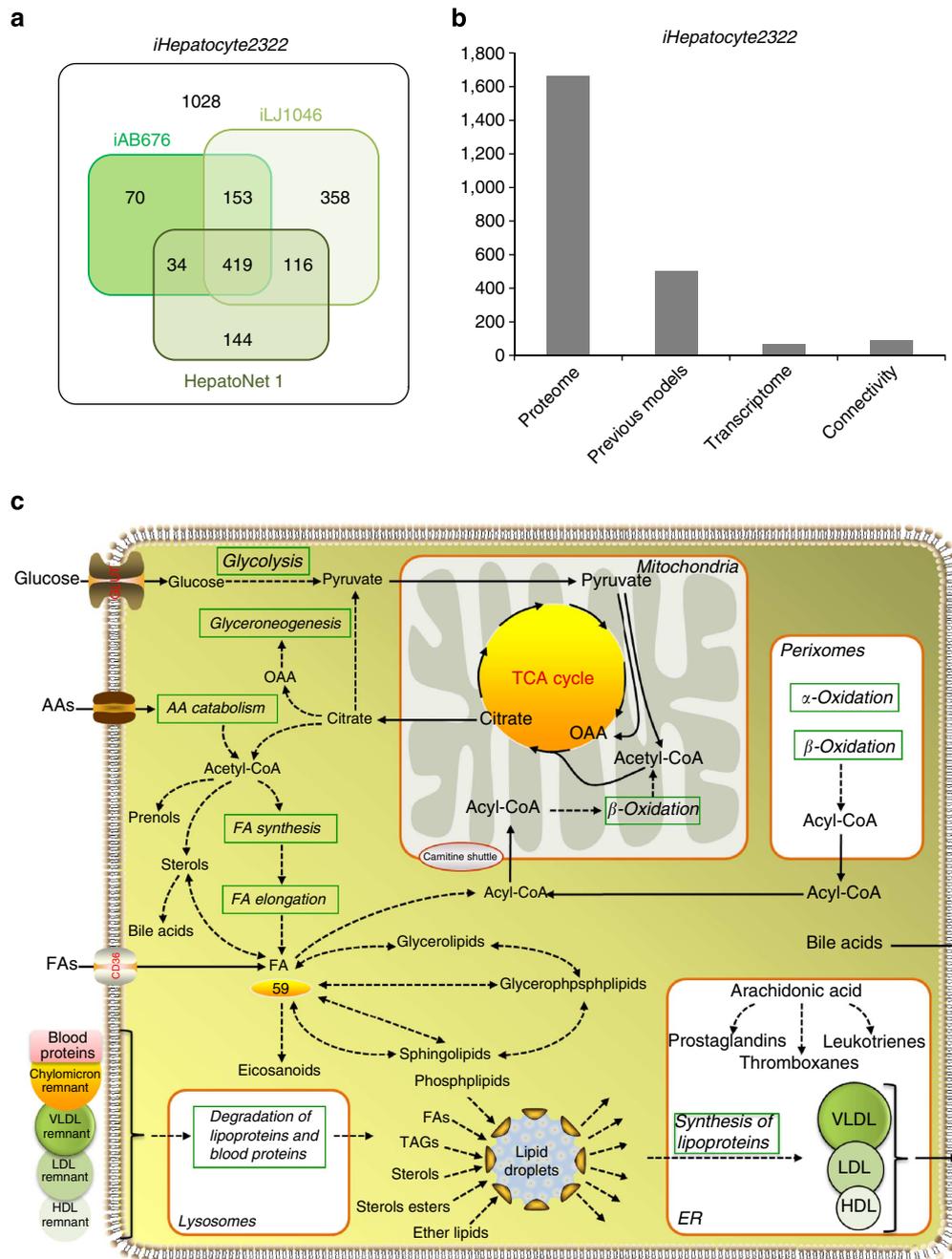
We tested *iHepatocytes2322* by simulating 256 different biologically defined metabolic functions (for example, the synthesis of FAs, amino acids, cholesterol and bile acids; Supplementary Data 5) that is known to occur in hepatocytes using the RAVEN Toolbox<sup>23</sup>. Furthermore, the ability of *iHepatocytes2322* for performing gluconeogenesis was demonstrated using experimentally measured secretion rates for glucose and albumin, and uptake rates for glycerol, lactate, amino acids and FAs in primary rat hepatocytes<sup>24</sup> (Supplementary Data 6).

The HPA covers the annotated expression of proteins and their subcellular localization in major human cell types, cancer and cell lines<sup>20</sup>. Relative abundance of proteins encoded by 15,155 genes in hepatocytes was analysed with 18,707 high-throughput-generated affinity-purified antibodies (Supplementary Data 1). The model reconstruction process was in excellent agreement with the protein profiling of hepatocytes in HPA. During the implementation of the metabolic tasks in *iHepatocytes2322*, merely 61 (~1.6%) out of 3,765 proteins in the HMR database and associated reactions had to be integrated into the model to maintain the functionality, even though they have been reported to be non-expressed in hepatocytes according to the HPA. We re-analysed the immunohistochemistry (IHC) data of these 61 proteins and found that 20 (0.5%) of these proteins actually

should display presence in the liver (Supplementary Data 7). Initial discordant data were due to the suboptimal titration of the antibody, misinterpretation of weak IHC staining or due to interference with other cell types besides hepatocytes present in the liver (for example, Kupffer cells and sinusoids). Nine (0.2%) of the investigated proteins showed more concordant results to the mathematic model when re-analysed using another antibody targeting the same protein. Fifteen proteins (0.4%) with negative IHC data were kept as negative in HPA data, as limited literature was available and/or concordant results were seen in subsets of the remaining panel of tissues included in the HPA high-throughput setup. The remaining 17 proteins (0.5%) were inaccurately assessed by IHC due to technical issues, such as antigen recognition due to antigen conformational changes, fixation or suboptimal antibody.

**Discovery of biomarkers for NASH.** NAFLD is progressively diagnosed worldwide<sup>25</sup>, is tightly associated with obesity, type 2 diabetes, insulin-resistance and hypertension, and represents a severe risk for the development of cirrhosis and hepatocellular carcinoma<sup>26</sup>. Despite its severe drawbacks, liver biopsy is still the most common procedure for diagnosing NASH<sup>18</sup>. Thus, there is a need for identifying non-invasive biomarkers to diagnose NASH and to subcategorize the NAFLD patients without taking biopsies.

To date, there has been a number of studies aiming at finding non-invasive biomarkers for diagnosis and staging of NAFLD<sup>27,28</sup>. A routinely available biochemical marker for hepatocellular damage is alanine transferase, but this has only proven to have moderate diagnostic value. Other markers for metabolic syndrome (such as adiponectin or leptin),



**Figure 2 | Consensus GEM for hepatocytes, *iHepatocytes2322*.** (a) *iHepatocytes2322* is reconstructed through the use of the HMR 2.0 database and improved annotation of proteomics data in the HPA. This high-quality model includes the structure of major lipid metabolism in hepatocytes, as well as all of the reactions and associated genes to those reactions in previously published GEM for hepatocytes. The overlapping of the genes in *iHepatocytes2322* and previously published models are presented. One thousand and twenty-eight new protein-coding genes were included into *iHepatocytes2322* primarily based on proteomics evidence provided by HPA. (b) Genes and associated reactions in *iHepatocytes2322* are included into the model based on the high-quality proteome, transcriptome, previously published models, as well as the connectivity. The overall distribution is shown. (c) *iHepatocytes2322* contains extensive lipid metabolism that is known to exist in hepatocytes, in addition to other known metabolic pathways. In the model, 59 different individual FAs are used to allow the integration of high-quality lipidomics data rather than generic pool names. The model can uptake the remnants of chylomicrons, VLDL, low-density lipoproteins (LDLs) and high-density lipoproteins (HDLs) and can form and degrade LDs. Moreover, the model can synthesis VLDL, LDL and HDL, and secrete it to the blood. Some of the important elements of lipid metabolism are shown.

inflammation (such as tumour-necrosis factor- $\alpha$  or interleukin 6), oxidative stress (such as catalase, glutathione peroxidase and overall plasma ferric reducing ability), apoptosis and fibrosis have also been proposed as biomarkers for staging NAFLD<sup>29,30</sup>. Among these, the apoptosis markers are arguably the ones with the best predictive ability. One such marker is cytochrome 18

fragment, which is an intermediate filament expressed in single-layer epithelial tissues in patients with NAFLD. Its strong correlation with the occurrence of liver fibrosis and hepatic inflammation has been reported<sup>31,32</sup>.

In this study, we focused on predicting potential metabolite biomarkers rather than on proteins, as they can quickly and easily

be measured both in the plasma and urine<sup>33</sup>. We analysed the liver gene expression data obtained from NASH (severe stage of NAFLD) patients (Supplementary Fig. 1) to understand the multi-factorial nature of its appearance by using *iHepatocytes2322* as a scaffold for data analysis. The liver transcriptomics data include samples from 45 different subjects, and the samples were diagnosed as healthy ( $n = 19$ ), steatotic ( $n = 10$ ), NASH with FL ( $n = 9$ ) and NASH without FL ( $n = 7$ )<sup>34</sup>. Diagnosis of the liver samples was first established by a Liver Tissue Cell Distribution System medical pathologist and was confirmed by histological examination at the University of Arizona in a blinded fashion. Steatosis was diagnosed by  $>10\%$  fat deposition without inflammation or fibrosis. NASH with and without FL samples were characterized by  $>5\%$  fat deposition and  $<5\%$  fat deposition within hepatocytes, respectively, and both were accompanied by inflammation and fibrosis. The severity can be ordered as NASH without FL  $>$  NASH with FL  $>$  steatotic  $>$  healthy. Clinical information of these human liver samples has been described previously<sup>35</sup>. In brief, samples of frozen and formalin-fixed, paraffin embedded adult explant livers were obtained from the Liver Tissue Cell Distribution System at the University of Minnesota, Virginia Commonwealth University, and the University of Pennsylvania. Histological staining of progressive stages of NAFLD were provided for each human liver donor samples using a previously established scoring system<sup>36</sup>, and representative images of haematoxylin and eosin-stained livers from normal, steatotic, NASH with FL and NASH without FL have been provided<sup>35</sup>. The age, gender and disease state of the patients are included in Supplementary Data 8.

We identified metabolic differences by performing a pair-wise analysis of the gene expression of subjects with NASH with and without FL versus healthy (Fig. 3) and steatotic (Supplementary Fig. 2) samples using the Reporter Metabolite algorithm<sup>37</sup>. Reporter Metabolite analysis allows for the identification of metabolites in the network for which there is significant enrichment of associated gene expression changes. Such metabolites can therefore be used to discover key regions of the metabolic network, which are significantly perturbed between the compared conditions<sup>37</sup>. The analyses for two NASH patient groups were performed independently and a total of 60 statistically significant (Reporter Features,  $P$ -value  $< 0.05$ ) Reporter Metabolites for NASH with and without FL versus healthy samples were identified. The association of the Reporter Metabolites with up- and downregulated genes and their metabolic subsystems classified in the HMR database are presented (Fig. 3). NASH is associated with some well-studied major metabolic abnormalities. These include increased uptake of FAs, decreased  $\beta$ -oxidation and cholesterol synthesis, and irregular preparation and export of triacylglycerols and cholesterol in the form of VLDL particles<sup>38,39</sup>. Many of the metabolites associated with these functions were identified in our analysis, and literature evidence for their association with the appearance of NAFLD are included in Supplementary Data 9.

We focused here on the less well-studied metabolic subsystems involved in the progression of the NASH. In addition to subsystems previously implicated in NASH (for example, folate, vitamin B6, lipid, eicosanoid and amino acid metabolism<sup>40,41</sup>), new Reporter Metabolites involved in glycan metabolism and biosynthesis of chondroitin sulphate (CS), a proteoglycan (PG) were identified. Previously, the association of serum levels of hyaluronic acid (a non-sulphated glycosaminoglycan) with the fibrosis stage in chronic liver diseases, including NAFLD, was reported<sup>28,42</sup>. Hyaluronic acid is distributed throughout epithelial tissues and most of its disassembly takes place in endothelial liver cells. Fibrosis and cirrhosis lead to impaired clearance of hyaluronic fragments due to the lack of function in endothelial liver cells.

Kalsch *et al.*<sup>43</sup> re-evaluated the hyaluronic acid as biomarkers for NAFLD and fibrosis in a cohort of 127 patients, and compared these results with the histological diagnosis of NAFLD.

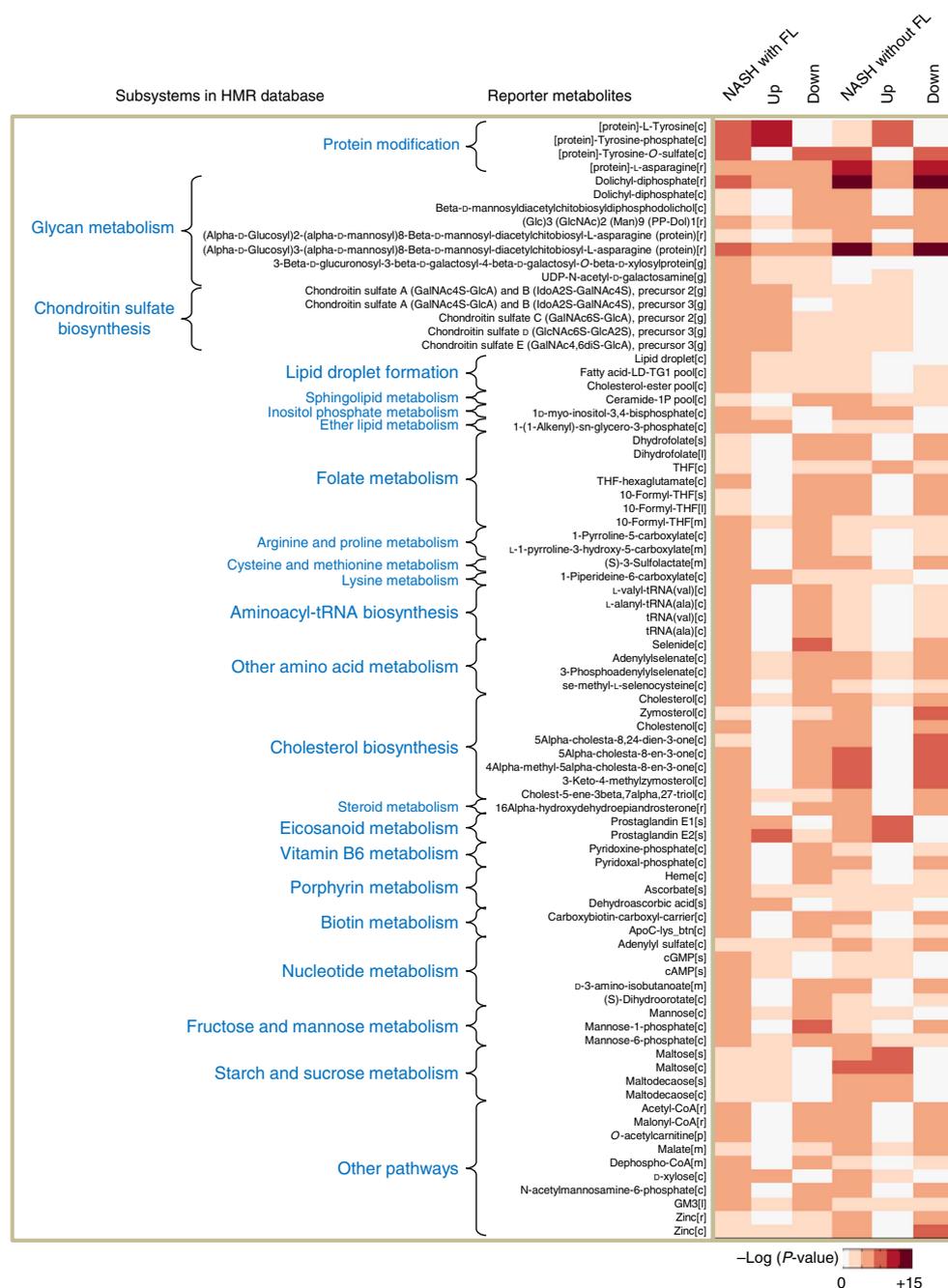
PGs are composed of glycosaminoglycans, including CS and heparan sulphate (HS) and core proteins. The biosynthesis of PGs starts with the xylosylation of serine residues in core proteins. To gain more knowledge about the metabolic differences around PGs, the detailed CS and HS biosynthesis pathway and the gene expression changes in NASH with and without FL patients versus healthy subjects are presented (Supplementary Fig. 3). It is observed that the expression of the genes involved in the CS biosynthesis are upregulated, whereas the expression of genes involved in the biosynthesis of HS are downregulated (Supplementary Table 3). It has been earlier reported that *CSPG2* gene is upregulated in biopsy-proven NASH patients versus obese controls<sup>44</sup>. CS and HS were also implicated in cancer progression<sup>45</sup>, one of the most severe outcomes of NASH. Therefore, we predicted that these changes in gene expression, in particular as it involves complete metabolic pathways, may correlate with a change in blood concentration of the pathway-associated metabolites. Hence, the blood level of CS and HS can be regarded as a potential biomarker for diagnosing NASH.

To validate the gene expression changes around CS and HS, we retrieved another microarray data set that includes liver samples from eight patients with morbid obesity and associated NASH, and seven control obese subjects from the Gene Expression Omnibus public repository under the accession number GSE37031. The control subjects had normal serum aminotransferase levels and liver histology. The Reporter Metabolites for this data set are presented (Supplementary Fig. 4). As can be seen, there are quite large differences in terms of overlap with the previously used data set. This can most likely be attributed to the small sample sizes in both data sets. Most importantly though, CS were found among the top-ranking metabolites for which there were transcriptional upregulation, and HS were found among the top-ranking metabolites for which there were transcriptional downregulation (Reporter Features,  $P$ -value  $< 0.05$ ).

**Potential therapeutic targets for NASH.** The Reporter Subnetwork algorithm identifies a set of metabolic reactions that exhibit transcriptional correlation after a perturbation (in this case NASH)<sup>37</sup>. We applied this algorithm to gain more insights into the molecular mechanisms involved in the appearance of NASH. After removing highly connected metabolites (for example, cofactors; Supplementary Table 4) in *iHepatocytes2322*, the involved subnetworks in either NASH with or without FL were identified and they are presented (Fig. 4a and Supplementary Fig. 5). The enzymes involved in the reactions are also represented (Fig. 4a), and related  $P$ -values and fold changes of their expression are shown (Supplementary Table 5).

The Reporter Subnetwork analysis showed that the non-essential amino acids serine, glycine, glutamate, glutamine, aspartate, asparagine and alanine, and the essential amino acid valine and methionine seem to be involved in the appearance of NASH. For reasons discussed later, serine, glycine and glutamate are of particular interest. Several metabolites involved in folate metabolism (for example, tetrahydrofolate (THF), 5-methyl-THF, 5-formyl-THF, 5,10-methenyl-THF and 5,10-methylene-THF) were also identified in the Reporter Subnetwork analysis, and these metabolites are involved in the interconversion of serine, glycine and glutamate. The metabolism around THF changed in NASH patients and this difference may be dependent on the uptake of 5-methyl-THF and 5-formyl-THF.

Moreover, phosphatidylserine (PS), an essential component for formation of LDs, was identified through our analysis. LDs have

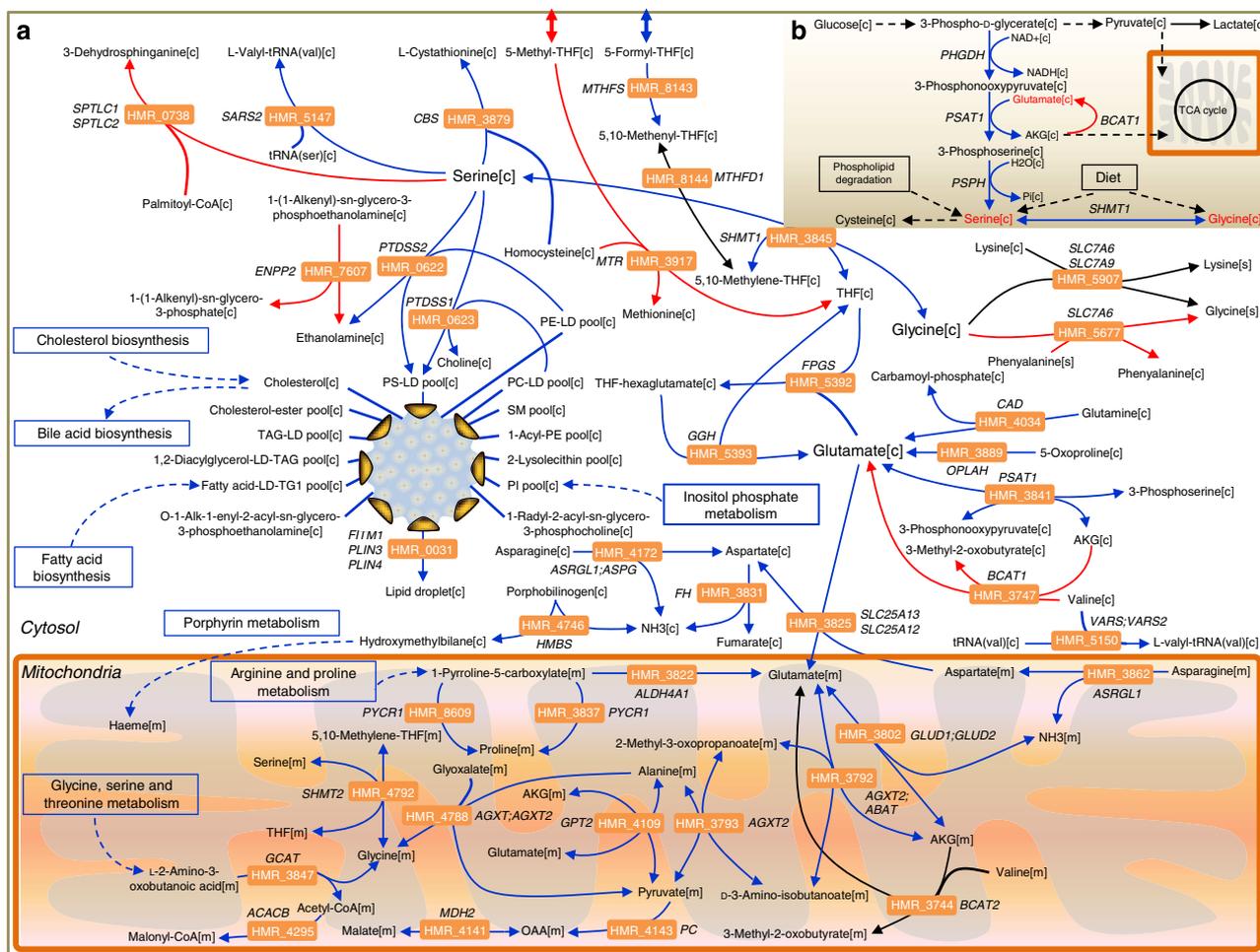


**Figure 3 | Reporter Metabolites through the global analysis of high-throughput data.** Metabolic differences between the liver gene expression profiles of subjects with and without FL, and healthy subjects were investigated through the employment of *iHepatoctyes2322*. In addition to the reporter metabolites associated with the known pathways involved in the appearance of NASH, reporter metabolites associated with the less well-studied subsystems, including Glycan metabolism and CS metabolism have been discovered. The blood level of CS, which can be secreted and taken up by the blood, is identified as a candidate biomarker for diagnosing NASH and staging NAFLD. *P*-values for each reporter metabolite were calculated for up- and downregulated genes, and minus logarithm of the *P*-values are presented.

diverse roles in the cell, such as serving as storage for triacylglycerols and CEs, or protecting the cell from excess lipids or lipophilic substances that may be toxic<sup>46</sup>. The enzymes PS synthases *PTDSS1* and *PTDSS2* that catalyse the production of PS by condensation of phosphatidylcholine and phosphatidylethanolamine, respectively, were significantly downregulated in NASH patients. The significant changes in the level of PS in cirrhotic (severe stage of NASH) livers was previously reported in a study on changes in lipid species in subjects with cirrhotic livers compared with healthy controls<sup>47</sup>. Given that PS is essential for

hepatocytes, we hypothesize that decreased activity of these enzymes may be associated with a decrease in the endogenous level of serine, which is the second most connected node in our identified Reporter Subnetworks (Fig. 4a).

Serine is endogenously biosynthesized from a glycolytic intermediate, 3-phospho-D-glycerate. This three-step process is catalysed by phosphoglycerate dehydrogenase (*PHGDH*), phosphoserine aminotransferase 1 (*PSAT1*) and phosphoserine phosphatase (*PSPH*), as shown in Fig. 4b. An alternative synthesis pathway is via the reversible interconversion with



**Figure 4 | Reporter Subnetworks through mapping of high-throughput data.** (a) Reporter Subnetworks were identified in *iHepatoctyes2322* through the pair-wise comparison of liver transcriptomics data obtained from NASH with and without FL versus healthy subjects. Amino acids, in particular serine, glutamate and glycine, have a role in the appearance of NASH, and the significant changes around these amino acids have been presented. Through our analysis, three different drug targets (*PSPH*, *SHMT1* and *BCAT1*) have been identified. Red arrows indicate overexpression of the associated genes, whereas blue arrows indicate underexpression. PS is presented as PS-LD pool in the subnetwork and the descriptions for other metabolites are included in the SBML (Systems Biology Mark-up Language) model file. (b) Serine is endogenously biosynthesized from 3-phospho-D-glycerate and glycine. It can also be derived from the diet and the degradation of protein and/or phospholipids.

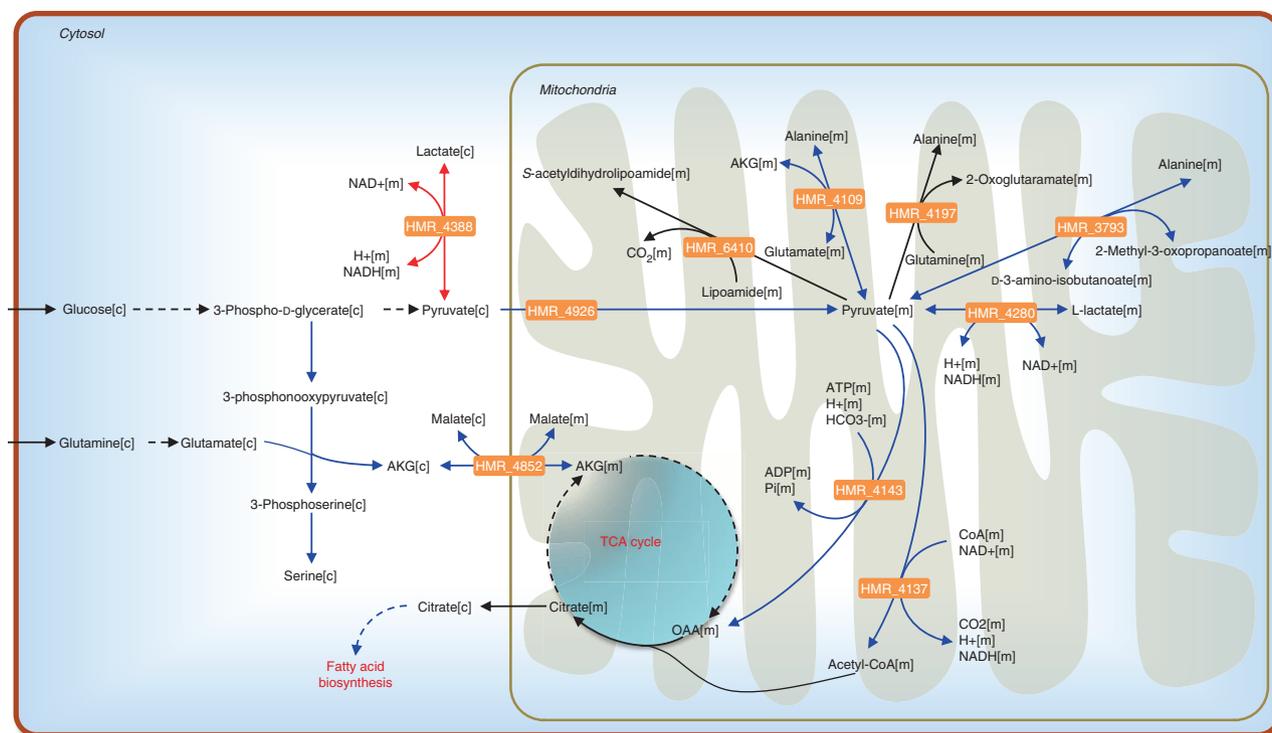
glycine through hydroxymethyltransferases *SHMT1* and *SHMT2*. Serine can also be derived from the diet and the degradation of protein and/or phospholipids. Serine plays a key role in the central metabolism, where it is involved in the formation of macromolecules, including lipids (sphingosine and PS), and other building blocks and cofactors, such as protein (glycine and cysteine), creatine, porphyrins, glutathione and nucleotides<sup>48</sup>.

Through differential analysis of transcriptomics data from the NASH patients, it was also observed that gene expression of several enzymes that either use serine as substrate or produce it as a product, including *CBS* (cysteine synthesis), *SARS2* (aminoacyl-tRNA biosynthesis), *SHMT1* and *SHMT2* (glycine synthesis) were significantly downregulated (*t*-test, *P*-value < 0.05), whereas *SPTLC1* and *SPTLC2* (sphingosine synthesis) were significantly upregulated (Supplementary Table 5). Downregulation of *CBS* that catalyses the conversion of serine and homocysteine to L-cystathionine and upregulation of *MTR* that condensates homocysteine to methionine through the use of 5-methyl-THF indicate that there are metabolic changes around homocysteine in NASH patients. Notably, it has been earlier reported that the plasma homocysteine level can be used for diagnosing NASH and

classifying steatosis and NASH patients<sup>49</sup>. It is not always straightforward to relate blood concentrations to gene expression levels of the involved enzymes, but our model-based analysis suggests a mechanistic explanation for this.

Taken together, the results suggest that the changes in the level of PS in the liver<sup>47</sup> as well as the relative increase in the homocysteine blood level<sup>49</sup> is caused by decreased level of endogenous serine. To test this hypothesis, we checked the expression level of enzymes that catalyse the biosynthesis of serine in the liver of NASH patients, and it was observed that the expression levels of *PHGDH*, *PSAT1*, *PSPH* in serine synthesis pathway (SSP), and *SHMT1* and *SHMT2* enzymes were significantly downregulated (Supplementary Table 6). Decreased levels of serine in NASH patients was supported by plasma profiling of amino acids, where the serine level in the plasma is relatively decreased (15% decrease, *P*-value = 0.0568) compared with healthy subjects<sup>50</sup>.

Equimolar amounts of serine and alpha-ketoglutarate (AKG) are synthesized in the SSP, and downregulation of reactions in SSP decrease the anaplerosis of AKG to the tricarboxylic acid (TCA) cycle in the form of glutamate<sup>51</sup>. Decreased level of serine



**Figure 5 | Mitochondrial dysfunction in NASH.** Mitochondrial reactions involving pyruvate as a reactant in *iHepatocytes2322* are presented. Arrows are coloured based on the direction of change in expression of the corresponding genes in NASH with and without FL samples versus healthy samples. Red arrows indicate overexpression of the associated genes, whereas blue arrows indicate underexpression. Not statistically significant changes (*t*-test, *P*-value > 0.05) associated to the reactions are indicated with black arrows.

also causes an accumulation of upstream glycolytic intermediates<sup>52</sup>, and a decreased flux of mitochondrial AKG is compensated by an increased flux of pyruvate to oxaloacetate in a healthy cell. To investigate the occurrence of this mechanism in NASH patients, we examined all mitochondrial reactions involving pyruvate as reactant in *iHepatocytes2322*. We found that the corresponding genes were downregulated for five out of seven such reactions (Fig. 5 and Supplementary Table 7). Furthermore, we investigated the expression of level of mitochondrial pyruvate carriers (*MPC1* and *MPC2*) and mitochondrial AKG/malate carrier (*SLC25A11*), and it was observed that their expression levels were downregulated in NASH patients. These indicate that the mitochondrial metabolic activity (TCA cycle) of hepatocytes is decreased in NASH patients in comparison with healthy subjects. This is in agreement with findings in our previous study wherein we investigated the metabolic changes in the case of fat accumulation in adipocytes in response to obesity<sup>17</sup>.

In the appearance of NASH, glutamate, which is the most connected node in our Reporter Subnetwork, plays a significant role as well. All enzymes linked to glutamate, except the branched chain amino-acid transaminase 1 (*BCAT1*) that convert AKG and valine to glutamate in the cytosol, are downregulated. In NASH patients, upregulation of *BCAT1* is previously reported<sup>40</sup> and our study identifies its mechanism in the appearance of NASH. The simultaneous upregulation of *BCAT1* and downregulation of *PSPH* could point to an imbalance in intercellular level of AKG and glutamate. This could possibly result in an accumulation of intracellular glutamate, which would then be compensated for by reduced uptake from/increased export to the blood. Notably, a previous study detected a significantly higher level of glutamate (60% increase,  $P = 9.808E - 09$ ) in the plasma profiling of the amino acids in NASH patients<sup>50</sup>. At the same time, this may

result in a higher demand for valine in hepatocytes, and given that valine is an essential amino acid this would arguably correlate with an increased uptake. Indeed, the same study<sup>50</sup> reported that the plasma valine concentration displayed significant changes (10% increase,  $P$ -value = 0.012).

## Discussion

To gain insights into the underlying molecular mechanisms of NASH, we reconstructed a consensus GEM for hepatocytes, which allowed for study of the interactions between lipid metabolism and other cellular metabolic functions. The reconstruction process enabled re-evaluation and improved annotation of the proteomics data, and our analysis clearly showed the power of using reconstructed metabolic networks for improving the annotation of experimental expression data. This highly curated GEM reconstructed through the use of the HMR 2.0 database enables interpretation of systemic effects, provides deeper insight into omics data for better understanding of the genotype-phenotype relationship in NASH subjects and allows for application of constraint-based modelling techniques to distinguish the NASH-specific metabolic features. On the basis of liver transcriptomics data of NASH patients and systems biology-based approaches, we proposed potential biomarkers and identified candidate therapeutic targets for NASH. Throughout our analysis, the gene expression levels served only as cues for the changes in the metabolic flux of its associated reaction. It is known that the correlation between these measurements is known to be limited, but several methods for inferring flux rates using gene expression data have been successfully applied<sup>53,54</sup>.

On the basis of our analysis, increasing the serine level in hepatocytes through the uptake of serine as a dietary supplement could be beneficial for NASH patients. Similarly, activity loss of

*PHGDH* in SSP in the brain, which causes low serine and glycine levels, and affects neuronal function, is reversed by serine supplementation<sup>55</sup>. The toxicity and the dosage of serine during its uptake through diet have been previously studied. Furthermore, long-term serine treatment decreased the homocysteine level in animal studies<sup>56</sup> and in humans, in a single-dose situation<sup>57</sup>.

One other possible way to increase the serine level to offer the possibility for therapeutic interventions is activation of the enzymes in SSP or *SHMT1* and *SHMT2* that converts glycine to serine. Three different enzymes constitute the SSP and it is earlier reported that *PSPH* is the rate-controlling enzyme for SSP in the liver<sup>58</sup>. Activation of the SSP through the amplification of *PSPH* may also decrease the flux through pyruvate and lactate formation in cytosol, as increased pyruvate and lactate levels were previously reported in NASH patients<sup>50</sup>. Recently, Frayling *et al.*<sup>59</sup> performed a genome-wide association study using 1,004 non-diabetic individuals and identified 8 common genetic variants relevant to insulin sensitivity and type 2 diabetes that are strongly linked to NASH phenotype. Their findings are in agreement with our results and they reported that three variants were associated with serine levels, out of which one is in the *PHGDH* gene and the other two are independently in the *PSPH* gene.

Boosting the serine level through SSP will also increase the flux on unregulated anaplerotic reactions that drive glutamine-derived carbon (via glutamate) into the TCA cycle through increased level of AKG and result in increased TCA cycle flux<sup>51</sup>. TCA cycle intermediates, besides being involved in driving energy production, are also used for biosynthesis of lipids (citrate), porphyrin (succinyl-CoA) and amino acids (AKG and oxaloacetate). AKG concentration in the cytosol can also be increased by inhibition of *BCAT1*, which converts AKG and valine to glutamate. Increasing the AKG level by either overexpressing *PSPH* or inhibiting *BCAT1* may change the NASH-specific patterns to healthy patterns, and it may potentially be used to develop an effective treatment for NASH.

In conclusion, we reconstructed a consensus GEM for hepatocytes, showed how reconstructed metabolic networks can be used for improving the annotation of experimental data and employed the model to gain more insight into the metabolic transformations associated with the development of NASH. Our analysis suggests that it is possible to diagnose NASH through identified metabolic biomarkers such as CS and HS levels in the blood. Furthermore, the development of therapeutic techniques based on the enhancement of endogenous serine and AKG levels may correct the underlying aetiology of NASH. This could be achieved by activation (or elevated expression) of *PSPH* and *SHMT1*, and inhibition of *BCAT1*. This study demonstrates that a deeper understanding of the metabolic changes obtained through GEM modelling may allow for elucidating the unknown aetiology of NASH, discovery of novel biomarkers, identification of drug targets and, eventually, development of efficient treatment strategies.

## Methods

**Expansion of HMR database.** The HMR database was constructed<sup>15</sup> by integrating the elements of stoichiometric networks of human metabolism, Recon 1 (ref. 12) and EHMN<sup>13,14</sup>, as well as the Kyoto Encyclopedia of Genes and Genomes (KEGG) database<sup>60</sup>. To generate a generic network for studying the effect of the lipids on the cellular metabolism, we expanded the coverage of the HMR database. We first merged metabolism of lipids and lipoproteins in Reactome, a manually curated and peer-reviewed pathway database<sup>61</sup>, and the literature-based GEMs including Recon 1 (ref. 12), EHMN<sup>13,14</sup> and HepatoNet 1, a manually reconstructed GEM for hepatocytes<sup>1</sup>. Extensive lipid metabolism involving 59 different FAs in the comprehensive database for lipid biology for mammalian cells, Lipidomics Gateway<sup>62</sup>, were included in the network and the gaps in the resulting network were filled using public databases such as KEGG<sup>60</sup> and HumanCyc<sup>63</sup>.

The resulting generic human GEM is called HMR 2.0 database, and to ensure the standardization of the HMR database all model components were extensively annotated with database identifiers. HMDB, Lipid Map<sup>62</sup>, KEGG and ChEBI identifiers were assigned for each metabolite, and KEGG ids and enzyme commission numbers were assigned for each reaction. Alternative genes associated to reactions were assigned using UniProt<sup>64</sup> and Lipid Map proteome database<sup>65</sup> using enzyme commission numbers. The resulting HMR database contains 3,765 genes, 6,007 metabolites (3,160 unique metabolites) and 8,181 reactions, and 74% of the reactions associated to one or more genes. The generated HMR 2.0 database is the most comprehensive resource for the human-related biochemical reactions and includes all of the genes, metabolites and reactions in the recently published models (Supplementary Table 8). It also includes all of the genes and associated reactions in recently published generic human model Recon2 (ref. 16). In the HMR 2.0 database, proteins encoded by genes are classified into eight different compartments, including cytosol, nucleus, endoplasmic reticulum, Golgi apparatus, peroxisome, lysosome, mitochondria and extracellular space. The HMR 2.0 database is available at <http://www.metabolicatlas.com> in SBML (Systems Biology Mark-up Language) format.

To construct a simulation-ready HMR 2.0 database, first it was tested so that all individual reactions, except pool reactions, were mass balanced. Second, it was guaranteed that high-energy compounds cannot be generated from low-energy compounds (such as ATP from ADP). This allowed us to test the thermodynamic constraints and the reversibility of the reactions. Third, the gap identification and gap-filling capabilities of the RAVEN Toolbox<sup>23</sup> were used to guide targeted literature studies to keep the number of dead-end reactions to a minimum. The production of all metabolites in the model was tested using artificial reactions (Supplementary Data 10). Artificial reactions were used to ensure the connectivity and were not included during the simulations and network-dependent analysis.

**Consensus GEMs for hepatocytes.** GEMs provide biologically meaningful mechanistic basis for the genotype-phenotype relationships, yet it is necessary to have functional cell-type GEMs to identify the metabolic differences between different states. We reconstructed *iHepatocytes2322* by merging recently generated GEM for hepatocytes *iHepatocyte1154* and previously published liver models<sup>1,15,21,22</sup>. *iHepatocyte1154* was generated from the HMR database using the INIT algorithm, which allows for automated reconstruction of GEMs based on the cell-type-specific proteome in the HPA (<http://www.proteinatlas.org>)<sup>20</sup>.

We incorporated differentially expressed genes (*t*-test, *P*-value < 0.001) in NAFLD patients in our reconstruction process, as *iHepatocytes2322* is used for the analysis of NAFLD patient data. Moreover, we integrated biochemical knowledge about hepatocyte metabolism and a large number of additional clinical data into the model. The resulting *iHepatocytes2322* is the largest cell/tissue-type-specific GEM and contains 2,322 genes, 5,686 metabolites (2,895 unique metabolites) in 8 different compartments and 7,930 reactions. In the model, 74% of the reactions are associated to one or more genes. *iHepatocytes2322* was validated with 256 known biological functions of hepatocytes (Supplementary Data 5), based on the definitions by Gille *et al.*<sup>1</sup>, by using the checkTasks function in the RAVEN Toolbox<sup>23</sup>.

**Transcriptomics data for NAFLD.** To study the appearance of human NAFLD through the changes in global gene expression, microarray data for liver samples were retrieved from ArrayExpress public repository (accession code E-MEXP-3291). The data include samples from 45 different subjects and the samples were diagnosed as healthy (*n* = 19), steatotic (*n* = 10), NASH with FL (*n* = 9) and NASH without FL (*n* = 7)<sup>34</sup>. The age, gender and disease state of the patients are presented (Supplementary Data 8).

The steatotic samples did not demonstrate significant gene expression changes compared with normal samples, as similarly reported in the plasma metabolic profiling of subjects with NAFLD, steatosis and NASH<sup>50</sup> (Supplementary Fig. 1). Hence, we performed the pair-wise comparison analysis of the gene expression to compare NASH with and without FL samples versus healthy (Fig. 3) and steatotic (Supplementary Fig. 2) samples using Piano R package<sup>66</sup>. Identified metabolic differences between the NAFLD patients through *iHepatocytes2322* provided detailed information comparing the enrichment of differentially expressed genes in the KEGG pathways (Supplementary Fig. 6). Differentially expressed genes in NASH with and without FL samples versus healthy and steatotic samples enriched in metabolism-related KEGG pathways, including steroid biosynthesis, oxidative phosphorylation, valine, leucine and isoleucine degradation, peroxisome, pyrimidine metabolism, pentose phosphate pathway and FA biosynthesis.

**Data availability.** HMR 2.0 database (Supplementary Data 11) and GEM for hepatocytes *iHepatocytes2322* (Supplementary Data 12) is publicly available in the SBML format at Human Metabolic Atlas (<http://www.metabolicatlas.org>).

The annotation of the presence or absence of protein targets in hepatocytes together with the high-resolution images is publicly available through the HPA (<http://www.proteinatlas.org>).

## References

- Gille, C. *et al.* HepatoNet1: a comprehensive metabolic reconstruction of the human hepatocyte for the analysis of liver physiology. *Mol. Syst. Biol.* **6**, 411 (2010).
- Baffy, G., Brunt, E. M. & Caldwell, S. H. Hepatocellular carcinoma in non-alcoholic fatty liver disease: an emerging menace. *J. Hepatol.* **56**, 1384–1391 (2012).
- Neuschwander-Tetri, B. A. & Caldwell, S. H. Nonalcoholic steatohepatitis: summary of an AASLD single topic conference. *Hepatology* **37**, 1202–1219 (2003).
- Bedogni, G., Kahn, H. S., Bellentani, S. & Tiribelli, C. A simple index of lipid overaccumulation is a good marker of liver steatosis. *BMC Gastroenterol.* **10**, 98 (2010).
- Lanpher, B., Brunetti-Pierri, N. & Lee, B. Inborn errors of metabolism: the flux from Mendelian to complex diseases. *Nat. Rev. Genet.* **7**, 449–460 (2006).
- Chan, S. Y. & Loscalzo, J. The emerging paradigm of network medicine in the study of human disease. *Circ. Res.* **111**, 359–374 (2012).
- Mardinoglu, A. & Nielsen, J. Systems medicine and metabolic modelling. *J. Intern. Med.* **271**, 142–154 (2012).
- Shoae, S. *et al.* Understanding the interactions between bacteria in the human gut through metabolic modeling. *Sci. Rep.* **3**, 2532 (2013).
- Varemo, L., Nookaew, I. & Nielsen, J. Novel insights into obesity and diabetes through genome-scale metabolic modeling. *Front. Physiol.* **4**, 92 (2013).
- Mardinoglu, A., Gatto, F. & Nielsen, J. Genome-scale modeling of human metabolism—a systems biology approach. *Biotechnol. J.* **8**, 985–996 (2013).
- Nielsen, J. Systems biology of lipid metabolism: from yeast to human. *FEBS Lett.* **583**, 3905–3913 (2009).
- Duarte, N. C. *et al.* Global reconstruction of the human metabolic network based on genomic and bibliomic data. *Proc. Natl Acad. Sci. USA* **104**, 1777–1782 (2007).
- Ma, H. *et al.* The Edinburgh human metabolic network reconstruction and its functional analysis. *Mol. Syst. Biol.* **3**, 135 (2007).
- Hao, T., Ma, H. W., Zhao, X. M. & Goryanin, I. Compartmentalization of the Edinburgh Human Metabolic Network. *BMC Bioinformatics* **11**, 393 (2010).
- Agren, R. *et al.* Reconstruction of genome-scale active metabolic networks for 69 human cell types and 16 cancer types using INIT. *PLoS Comput. Biol.* **8**, e1002518 (2012).
- Thiele, I. *et al.* A community-driven global reconstruction of human metabolism. *Nat. Biotechnol.* **31**, 419–425 (2013).
- Mardinoglu, A. *et al.* Integration of clinical data with a genome-scale metabolic model of the human adipocyte. *Mol. Syst. Biol.* **9**, 649 (2013).
- Machado, M. V. & Cortez-Pinto, H. Non-invasive diagnosis of non-alcoholic fatty liver disease - a critical appraisal. *J. Hepatol.* **58**, 1007–1019 (2012).
- Newgard, C. B. Interplay between lipids and branched-chain amino acids in development of insulin resistance. *Cell Metab.* **15**, 606–614 (2012).
- Uhlen, M. *et al.* Towards a knowledge-based Human Protein Atlas. *Nat. Biotechnol.* **28**, 1248–1250 (2010).
- Jerby, L., Shlomi, T. & Ruppin, E. Computational reconstruction of tissue-specific metabolic models: application to human liver metabolism. *Mol. Syst. Biol.* **6**, 401 (2010).
- Bordbar, A. *et al.* A multi-tissue type genome-scale metabolic network for analysis of whole-body systems physiology. *BMC Syst. Biol.* **5**, 180 (2011).
- Agren, R. *et al.* The RAVEN toolbox and its use for generating a genome-scale metabolic model for penicillium chrysogenum. *PLoS Comput. Biol.* **9**, e1002980 (2013).
- Chan, C., Berthiaume, F., Lee, K. & Yarmush, M. L. Metabolic flux analysis of cultured hepatocytes exposed to plasma. *Biotechnol. Bioeng.* **81**, 33–49 (2003).
- Rector, R. S., Thyfault, J. P., Wei, Y. Z. & Ibdah, J. A. Non-alcoholic fatty liver disease and the metabolic syndrome: An update. *World J. Gastroenterol.* **14**, 185–192 (2008).
- Ascha, M. S. *et al.* The incidence and risk factors of hepatocellular carcinoma in patients with nonalcoholic steatohepatitis. *Hepatology* **51**, 1972–1978 (2010).
- Miller, M. H., Ferguson, M. A. & Dillon, J. F. Systematic review of performance of non-invasive biomarkers in the evaluation of non-alcoholic fatty liver disease. *Liver Int.* **31**, 461–473 (2011).
- McHutchison, J. G. *et al.* Measurement of serum hyaluronic acid in patients with chronic hepatitis C and its relationship to liver histology. Consensus Interferon Study Group. *J. Gastroenterol. Hepatol.* **15**, 945–951 (2000).
- Pearce, S. G., Thosani, N. C. & Pan, J.-J. Noninvasive biomarkers for the diagnosis of steatohepatitis and advanced fibrosis in NAFLD. *Biomark. Res.* **1**, 7 (2013).
- Anderson, N. & Borlak, J. Molecular mechanisms and therapeutic targets in steatosis and steatohepatitis. *Pharmacol. Rev.* **60**, 311–357 (2008).
- Wieckowska, A. *et al.* In vivo assessment of liver cell apoptosis as a novel biomarker of disease severity in nonalcoholic fatty liver disease. *Hepatology* **44**, 27–33 (2006).
- Shen, J. *et al.* Non-invasive diagnosis of non-alcoholic steatohepatitis by combined serum biomarkers. *J. Hepatol.* **56**, 1363–1370 (2012).
- Griffiths, W. J. *et al.* Targeted metabolomics for biomarker discovery. *Angew. Chem. Int. Ed. Engl.* **49**, 5426–5445 (2010).
- Lake, A. D. *et al.* Analysis of global and absorption, distribution, metabolism, and elimination gene expression in the progressive stages of human nonalcoholic fatty liver disease. *Drug Metab. Dispos.* **39**, 1954–1960 (2011).
- Fisher, C. D. *et al.* Hepatic cytochrome P450 enzyme alterations in humans with progressive stages of nonalcoholic fatty liver disease. *Drug Metab. Dispos.* **37**, 2087–2094 (2009).
- Kleiner, D. E. *et al.* Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology* **41**, 1313–1321 (2005).
- Patil, K. R. & Nielsen, J. Uncovering transcriptional regulation of metabolism by using metabolic network topology. *Proc. Natl Acad. Sci. USA* **102**, 2685–2689 (2005).
- Fromenty, B. & Pessayre, D. Inhibition of mitochondrial beta-oxidation as a mechanism of hepatotoxicity. *Pharmacol. Ther.* **67**, 101–154 (1995).
- Mensenkamp, A. R., Havekes, L. M., Romijn, J. A. & Kuipers, F. Hepatic steatosis and very low density lipoprotein secretion: the involvement of apolipoprotein E. *J. Hepatol.* **35**, 816–822 (2001).
- Greco, D. *et al.* Gene expression in human NAFLD. *Am. J. Physiol. Gastrointest. Liver Physiol.* **294**, G1281–G1287 (2008).
- Anstee, Q. M. & Day, C. P. S-adenosylmethionine (SAME) therapy in liver disease: a review of current evidence and clinical utility. *J. Hepatol.* **57**, 1097–1109 (2012).
- Nobili, V. *et al.* Hyaluronic acid predicts hepatic fibrosis in children with nonalcoholic fatty liver disease. *Transl. Res.* **156**, 229–234 (2010).
- Kalsch, J. *et al.* Evaluation of biomarkers of NAFLD in a cohort of morbidly obese patients. *J. Nutr. Metab.* **2011**, 369168 (2011).
- Younossi, Z. M. *et al.* A genomic and proteomic study of the spectrum of nonalcoholic fatty liver disease. *Hepatology* **42**, 665–674 (2005).
- Afratis, N. *et al.* Glycosaminoglycans: key players in cancer cell biology and treatment. *FEBS J.* **279**, 1177–1197 (2012).
- Farese, R. V. & Walther, T. C. Lipid droplets finally get a little R-E-S-P-E-C-T. *Cell* **139**, 855–860 (2009).
- Gorden, D. L. *et al.* Increased diacylglycerols characterize hepatic lipid changes in progression of human nonalcoholic fatty liver disease; comparison to a murine model. *PLoS One* **6**, e22775 (2011).
- Maddocks, O. D. *et al.* Serine starvation induces stress and p53-dependent metabolic remodelling in cancer cells. *Nature* **493**, 542–546 (2013).
- Gulsen, M. *et al.* Elevated plasma homocysteine concentrations as a predictor of steatohepatitis in patients with non-alcoholic fatty liver disease. *J. Gastroenterol. Hepatol.* **20**, 1448–1455 (2005).
- Kalhan, S. C. *et al.* Plasma metabolomic profile in nonalcoholic fatty liver disease. *Metabolism* **60**, 404–413 (2011).
- Possemato, R. *et al.* Functional genomics reveal that the serine synthesis pathway is essential in breast cancer. *Nature* **476**, 346–350 (2011).
- Chaneton, B. *et al.* Serine is a natural ligand and allosteric activator of pyruvate kinase M2. *Nature* **491**, 458–462 (2012).
- Becker, S. A. & Palsson, B. O. Context-specific metabolic networks are consistent with experiments. *PLoS Comput. Biol.* **4**, e1000082 (2008).
- Shlomi, T., Cabili, M. N., Herrgard, M. J., Palsson, B. O. & Ruppin, E. Network-based prediction of human tissue-specific metabolism. *Nat. Biotechnol.* **26**, 1003–1010 (2008).
- de Koning, T. J. *et al.* Prenatal and early postnatal treatment in 3-phosphoglycerate-dehydrogenase deficiency. *Lancet* **364**, 2221–2222 (2004).
- Girardgl, A., Forestie, M. & Robin, P. Long-term adaptation of weanling rats to high dietary levels of methionine and serine. *J. Nutr.* **102**, 209 (1972).
- Verhoef, P. *et al.* Dietary serine and cystine attenuate the homocysteine-raising effect of dietary methionine: a randomized crossover trial in humans. *Am. J. Clin. Nutr.* **80**, 674–679 (2004).
- Lund, K., Merrill, D. K. & Guynn, R. W. The reactions of the phosphorylated pathway of L-serine biosynthesis: thermodynamic relationships in rabbit liver *in vivo*. *Arch. Biochem. Biophys.* **237**, 186–196 (1985).
- Frayling, T. M. *et al.* Metabolite quantitative trait Loci (mQTL) and their role in type 2 diabetes and insulin sensitivity. in *American Diabetes Association Meeting* (Diabetes, 2011).
- Kanehisa, M., Goto, S., Furumichi, M., Tanabe, M. & Hirakawa, M. KEGG for representation and analysis of molecular networks involving diseases and drugs. *Nucleic Acids Res.* **38**, D355–D360 (2010).
- Croft, D. *et al.* Reactome: a database of reactions, pathways and biological processes. *Nucleic Acids Res.* **39**, D691–D697 (2011).
- Harkewicz, R. & Dennis, E. A. Applications of mass spectrometry to lipids and membranes. *Annu. Rev. Biochem.* **80**, 301–325 (2010).
- Romero, P. *et al.* Computational prediction of human metabolic pathways from the complete human genome. *Genome Biol.* **6**, R2 (2005).
- Apweiler, R. *et al.* Ongoing and future developments at the Universal Protein Resource. *Nucleic Acids Res.* **39**, D214–D219 (2011).

65. Cotter, D., Maer, A., Guda, C., Saunders, B. & Subramaniam, S. Lmpd: lipid maps proteome database. *Nucleic Acids Res.* **34**, D507–D510 (2006).
66. 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 Res.* **41**, 4378–4391 (2013).

### Acknowledgements

This work was financially supported by the Knut and Alice Wallenberg Foundation and Chalmers Foundation.

### Author contributions

A.M. and R.A. updated the HMR 2.0 database and reconstructed *iHepatocytes2322*. R.A. and A.M. performed the analysis of clinical data. C.K. and A.A. re-evaluated the

proteomics data for hepatocytes. J.N. and M.U. conceived the project. A.M., R.A. and J.N. wrote the paper and all authors were involved in editing the paper.

### Additional information

**Supplementary Information** accompanies this paper at <http://www.nature.com/naturecommunications>

**Competing financial interests:** The authors declare no competing financial interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**How to cite this article:** Mardinoglu, A. *et al.* Genome-scale metabolic modelling of hepatocytes reveals serine deficiency in patients with non-alcoholic fatty liver disease. *Nat. Commun.* **5**:3083 doi: 10.1038/ncomms4083 (2014).