Searching for biomarkers of Non-alcoholic fatty liver disease and Metabolic syndrome

Thesis submitted for the degree of Doctor, presented by:

Laura de la Cruz Villar

Supervised by:

Dr. José M. Mato de la Paz

Dr. Óscar Millet Aguilar-Galindo

Faculty of Science and Technology

Department of Biochemistry and Molecular Biology

Doctoral Program in Molecular Biology and Biomedicine

Leioa, 2020

AGRADECIEMIENTOS

En primer lugar quisiera mostrar mi agradecimiento a José M. Mato por darme la oportunidad de realizar este trabajo en CIC bioGUNE durante estos cuatro años y de haber podido formarme en diferentes laboratorios, siempre acompañada de compañeros que me han ayudado en todo momento y me han hecho sentir parte del grupo desde el principio.

En concreto, quiero agradecer a José María y a Óscar haber tenido la oportunidad de aprender de vosotros, desde metabolismo hasta RMN, y por guiarme y ofrecerme vuestra ayuda siempre que la he necesitado.

A mis compañeros de OWL, por recibirme durante mis primeros meses en País Vasco, por introducirme en LC/MS y por enseñarme cómo es el trabajo en una pequeña empresa.

A todos mis compañeros del laboratorio de RMN-1 por los buenos momentos que hemos pasado trabajando y también tomando el té. En especial, quiero agradecer a Nieves todo lo que me ha enseñado desde el momento que llegué al laboratorio, tanto profesional como personalmente. Eres una gran persona, una buenísima compañera y me alegra haber podido compartir este tiempo contigo.

A Itxaso, por los momentos de té en el atrio y los ratos (algo más largos) de té en la denda. Por todo lo que compartimos tanto dentro como fuera del laboratorio y por haber encontrado una amiga como tú.

A todo el laboratorio de metabolómica y en especial a Malu, por abrirme desde el primer momento las puertas de tu despacho y ofrecerme su ayuda siempre que la he necesitado.

A Fer y a David por ser los mejores postdocs y amigos que alguien puede tener. Ha sido una verdadera suerte conoceros, trabajar con vosotros y disfrutar de la luz que siempre lleváis allí donde estéis.

A Vir, por todo el trabajo que haces, por lo gran compañera que eres y por el cuadro que me hiciste con el cariño con el que haces siempre las cosas.

A María y a Rubén por ser los mejores amigos y compañeros de piso que uno puede tener. Ojalá nos hubiésemos encontrado antes para disfrutar más tiempo de esta peculiar familia de 5 que tenemos montada en Indautxu. Sin duda estaréis siempre en mi corazón.

Por último, agradecer a mi familia y a mis amigos el apoyo que siempre me han ofrecido y las fuerzas que me han dado en los momentos de debilidad. En especial gracias a Joni por ofrecerme un hogar en el que poder escribir y molestarte a partes iguales. Gracias a todos por creer siempre en mí y por haberme acompañado incansablemente durante este camino que ahora llega a su fin.

TABLE OF CONTENTS

OBJECTIVES

OBJECTIVES

The main objective of this work is the study of non-alcoholic fatty liver disease and metabolic syndrome to better understand and characterized the patients suffering these diseases. Besides, we try to find non-invasive biomarkers and potential diagnostic methods to carry out metabolomic studies of big populations without compromising people well-being.

1. As the etiology of non-alcoholic fatty liver disease is multifactorial, we want to emphasize that NAFLD is a syndrome more than a disease and:
    1.1. Demonstrate the existence of different subtypes of NAFLD patients.
    1.2. Categorized these subtypes depending on the mechanism causing the disease.
    1.3. Show how personalized medicine could address the specific treatments for each subtype.
    1.4. Investigate further Aramchol® mechanism of action.

2. Demonstrate that nuclear magnetic resonance can be used as a non-invasive diagnostic tool for metabolic syndrome.
    2.1. Study of OBENUTIC and PreMedEus populations as pilot project.
    2.2. Set up of human urine samples analysis by NMR and data analysis.
    2.3. Find a metabolic fingerprint that allows the discrimination between healthy people, metabolic syndrome patients and intermediate phenotypes using NMR.

CHAPTER 1

INTRODUCTION

1 INTRODUCTION

1.1 Non-alcoholic fatty liver disease (NAFLD)

The liver, which can be from 2% to 4% of the total body weight, is a key metabolic organ that governs whole body energy homeostasis being hepatic dysfunction usually associated with systemic metabolic unbalances and *vice versa*. The liver metabolically connects several tissues, such as adipose tissue and skeletal muscle. Moreover, it also regulates the blood levels for many metabolites, since most of the substances we ingest by food are absorbed by the intestine with a first pass through the liver. In turn, liver energy metabolism is controlled by multiple neuronal, hormonal and nutritional signals that regulate glucose, lipid and amino acid metabolism.

An impairment in liver metabolism regulation, can lead to non-alcoholic fatty liver disease (NAFLD)[1] among other diseases.

NAFLD is a term used to describe a range of related disorders that comprise simple steatosis, non-alcoholic steatohepatitis (NASH) with varying degrees of fibrosis, cirrhosis and hepatocellular carcinoma (HCC)[2], a liver pathology with poor prognosis. Liver damage caused by drug and/or alcohol consumption and virus-induced damage are not considered in this work. NAFLD is a growing health problem with a global prevalence of 25% being the most common liver disorder in western countries and has an annual cost of more than $100 billion only in United States. The prevalence increases in people with type 2 diabetes (T2D) (70%) and in morbid obesity (90%)[3].

The earliest stage of NAFLD is steatosis, which is characterized by the deposition of triglycerides (TGs) as lipid droplets (LD) in the cytoplasm of the hepatocytes. It is considered steatosis when more than 5% of the liver contains LD[4]. NASH is distinguished from steatosis by the presence of inflammation and hepatocyte injury, as it will be explained in more detail in **section 1.3**. Approximately 25% of patients with simple steatosis progress to NASH, of whom 25% progress to cirrhosis and of those that develop cirrhosis at least 1-2% per year develop HCC[5,6,7], which is the 4-27% of these patients[8] (**Figure 1**).

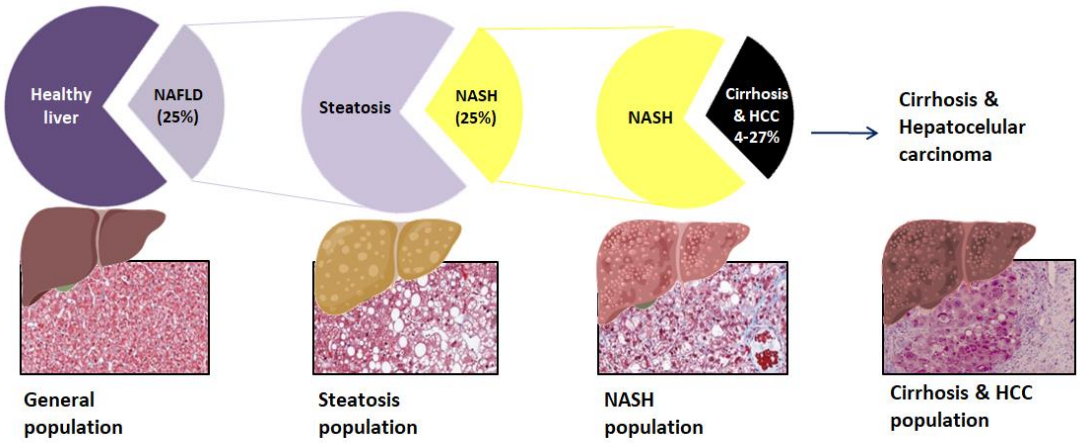


Figure 1. Progression from Healthy liver to HCC in humans

NAFLD may progress to liver cancer with no previous fibrosis or histologic NASH, or develop fibrosis without going to the NASH stage[6,9].

During evolution of human beings, we have been changing our diet incorporating more energetic aliments. This fact was favorable and advantageous in primitive societies where the daily energetic demand was very high for most of individuals and caloric diets helped them to survive. Taking into account differences in body size, it has been estimated that our energy expenditure per kg of body weight is on average, less than 40% as compared to our prehistoric ancestors[10]. Sedentary lifestyles acquired in today's modern societies and the consumption of high-caloric food, have paved the way for the development of metabolic diseases (among others) lowering the quality of life in terms of health. In this context, the last century's changes in lifestyle and eating habits have made NAFLD the most common liver disease, tightly related with T2D, obesity and dyslipidemia[6,11].

In fact, NAFLD may be considered as a hepatic manifestation of the metabolic syndrome (MetS). MetS is a set of metabolic abnormalities comprising T2D, increased fasting plasma glucose, hypertriglyceridemia, low high-density lipoprotein levels, hypertension and increased waist circumference. MetS components are highly prevalent in NAFLD patients and the risk of mortality is higher when more components of MetS are present[12]. MetS will be explained further in **section 1.4**.

1.1.1 *De novo* lipogenesis (DNL)

The liver is the main organ which converts carbohydrates into fatty acids (FAs). FAs are biomolecules that provide energy and act as constitutive building blocks in biological membranes. FAs are very energetic molecules (9 kcal/g) and provide much more energy than

other biological molecules such as proteins or carbohydrates (4 kcal/g). FAs are stored into LD as TGs but they can also have other fates, as described in this section.

Accumulation of FAs is an adaptive evolutionary mechanism in many animals that let them obtain energy in periods of food starvation[13]. For example, migratory birds store large quantities of TGs in the liver as an energy source in preparation for prolonged seasonal flights. This is in contrast to humans, where fatty liver is maladaptive and has severe clinical consequences.

DNL is transcriptionally regulated mainly by two transcription factors: the sterol regulatory element binding protein 1 (SREBP1) activated by insulin, and the glucose sensor carbohydrate response element-binding protein (ChREBP), activated by glucose[14]. The peroxisome proliferator-activated receptor gamma (PPARγ), is another transcription factor involved in lipogenesis; hepatic PPARγ stimulates the expression of genes involved in FA uptake, trafficking and TG biosynthesis increasing liver lipid levels[1].

Through glycolysis, glucose is metabolized to pyruvate which is then imported to the mitochondria and further decarboxylated by the pyruvate dehydrogenase complex (PDC) that converts it into acetyl-coA. Citrate synthase (CS) combines acetyl-coA and oxaloacetate to form citrate that is exported to the cytoplasm where the ATP-citrate lyase (ACL) splits it into acetyl-CoA and oxaloacetate. This oxaloacetate is metabolized by malic enzyme generating pyruvate and NADPH which is needed to synthesize palmitic acid (16-carbon FA) by the FA synthase (FAS). Once in the cytoplasm, acetyl-CoA carboxylase (ACC) carboxylates the acetyl-coA generating malonyl-CoA, which together with NADPH are precursors for the synthesis of palmitic acid. Palmitic acid is elongated by fatty acyl-CoA elongase family members (Elovl) to generate long chain FA (LCFA) (more than 16 carbon units) and these are desaturated by stearoyl-CoA desaturases 1 (SCD1). SCD1 products are monounsaturated fatty acids (MUFAs) being oleate and palmitoleate important regulators of glucose and lipid metabolism in the liver[15]. In addition to DNL, the adipose tissue and the diet are the other two major sources of hepatic FA (**Figure 2**).

Once synthetized, the four major fates of hepatic FAs are: mitochondrial β-oxidation, biosynthesis of other lipids, esterification and storage of TGs into LD and the assembly of TGs into very-low density lipoproteins (VLDL) exported into blood (**Figure 2**).

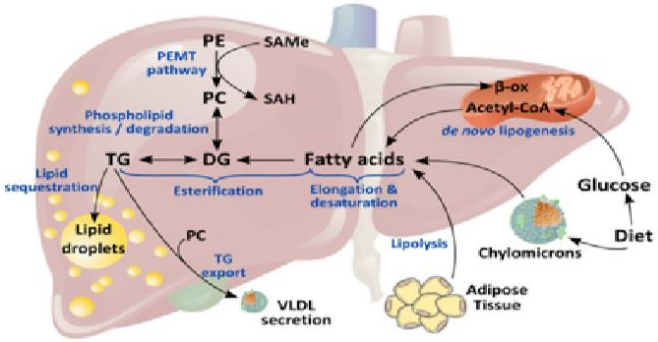


Figure 2. Synthesis of TG in the liver. TG can be synthetized by 4 different routes: by desaturation, elongation and esterification of FA, or by the phosphatidylethanolamine N-methyltransferase (PEMT) pathway that generates phosphatidylcholine (PC) from phosphatidylethanolamine (PE)

1.1.2 Triglycerides and Very-Low-Density Lipoproteins (VLDL)

The main mechanism by which liver stores and exports FAs is by the assembly of TGs into LD and VLDL. As TGs are insoluble in water, they can be stored to high levels without causing adverse osmotic or colloidal effects on cells[2].

Under non-pathological conditions, the liver stores TGs in small amounts and exports them in the form of VLDL to deliver FAs to peripheral organs such as muscle or adipose tissue (depending on the nutritional status). The assembly of FAs as TGs is protective against FA-mediated hepatotoxicity[16].

The glycerol-3 phosphate (G3P) pathway is the principal route of TGs synthesis in most of mammalian cell types (**Figure 3**). The first step of this pathway is the esterification of long-chain acyl-CoA to G3P catalyzed by mitochondrial and microsomal G3P acyltransferase (GPAT) enzymes, and it is a rate-limiting step that generates lysophosphatidic acid (LPA) molecules. Acylglycerol-3-phosphate acyltransferases (AGPAT) present in endoplasmic reticulum (ER) membranes, acylate these LPA molecules to form phosphatidic acid (PA) that can have two fates: to be converted into cytidine diphosphate diacylglycerol (CDP-DG), that is a substrate for the synthesis of cardiolipins and certain glycerolphopholpids[17,18], or be dephosphorylated by phosphatidate phosphohydrolase (PAP or lipin) to form diacylglycerol (DG). DG is a precursor for the synthesis of TGs, phosphatidylcholine (PC) and phosphatidylethanolamine (PE). The final step of TGs synthesis is the acylation of DG by the action of DG acyltransferase (DGAT)[16].

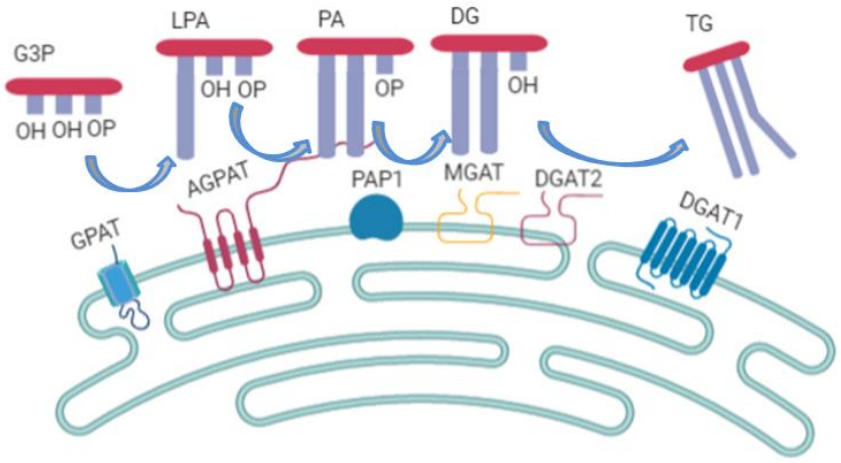


Figure 3. Schematic representation of TG synthesis from G3P in the endoplasmic reticulum membrane.

As shown in **Figure 2**, TGs also can be formed by the phosphatidylethanolamine N-methyltransferase (PEMT) pathway that uses S-adenosylmethionine (SAMe) to methylate PE and generates PC[19].

Although TGs accumulation is the main cause of NAFLD, there are more toxic lipid species contributing to NAFLD and NASH progression that are currently under study[20].

VLDLs are particles enriched in TGs also constituted by apolipoprotein (Apo) B-100 and C (ApoC), cholesterol esters and phospholipids. They are built in the liver to be delivered from it and reach the adipose tissue for storage and the muscle for oxidation and energy. VLDL particles are assembled in two steps: the first step is carried out in the ER lumen, where the microsomal triglyceride transfer protein (MTP) incorporates TGs into apoB100 (apoB48 in rodents) while they are being translated by ribosomes and translocated to the ER membrane. The second step includes further TG packaging into the nascent apoB100-containing particles forming the mature VLDL particles[16]. Within hepatocytes, MTP promotes ApoB-100 fusion with TGs and VLDLs formation and export[14]. Besides, *de novo* biosynthesis of PC is tightly related with VLDL assembly and secretion; and alterations in the PC-generating pathways also alter these processes. The gene methionine adenosyltransferase 1A (*MAT1A)* and its protein products (MATI and MATIII) are necessary for proper VLDLs assembly and secretion because as we will see later, are the enzymes responsible of SAMe synthesis; and this molecule acts as the methyl donor for PC formation from PE, which is needed for VLDLs synthesis and assembly[21].

1.1.3 Fatty acid oxidation

FAs from de DNL, circulating lipids or derived from hydrolysis of hepatic TGs stores can be oxidized by multiple pathways. FA β-oxidation takes place in the mitochondrial matrix and in normal conditions is the main pathway for the oxidation of the majority of FAs found in hepatocytes including: short (<C4), medium (C4-C12) and long-chain (C12-C20) FAs. Through peroxisomal β-oxidation, long-change dicarboxylic acids and very-long-change FAs are metabolized to complete the oxidation process in mitochondria and genetic ablation of enzymes involved in peroxisomal β-oxidation leads to hepatic steatosis[22]. Under normal conditions, cytochrome P450 4A (CYP4A) Ω-oxidation of FAs, which takes place in the microsomes, is a minor pathway for FA oxidation.

In normal conditions, FA β-oxidation occurs preferentially in the mitochondria rather than in the peroxisomes. Through β-oxidation of FA, acetyl-CoA is generated serving as a substrate for oxidative phosphorylation (OXPHOS), ketone body formation and other biochemical reactions such as DNL or gluconeogenesis[14]. Carnitine-palmityl-transferase 1 (CPT-1) is the main check point that controls the uptake of FAs into the mitochondrial matrix for oxidation and it is inhibited by malonyl-CoA[23,24]. PPARα is expressed at high levels in the liver and is involved in the regulation of mitochondrial and extra-mitochondrial FA oxidation[16].

1.1.4 Lipid homeostasis alterations

In a non-pathological state, i.e. after a meal, dietary fat is delivered to the liver in form of chylomicrons (**Figure 2**). Besides, the products of glucose metabolism are used to synthesize FAs in the liver through DNL and hepatocytes incorporate long-chain fatty acids in form of TGs, phospholipids and cholesterol esters (as LD) or secrete them into the circulation as VLDLs. On the other hand, mitochondrial FA β-oxidation has a very important role because provides energy to the hepatocytes, generates ketone bodies that are delivered into the circulation to provide metabolic fuel to extrahepatic tissues during fasting, and burn the excess of FAs in the liver.

NAFLD arises as a consequence of defects in diverse metabolic pathways that lead to an imbalance between the DNL and the uptake of FAs from circulation and the capacity of the liver to oxidize FAs and their elimination as TGs in form of VLDLs[25].

1.1.5 Lipid species in liver disease

Lipids are composed by a few structural “building blocks” but due to the high combinatorial side chain possibilities, there are more than 100,000 lipid molecules[26] and it is known that different lipid species belonging to the same class are differently associated with health and

disease states[27,28,29]. Regarding this, and unlike to what happens with discrete metabolites such as glucose, lipids can be built up from dozens of different molecules combined in different ways. For example, TGs result from the combination of various acyl chains esterified to glycerol so the molecular diversity is high. In this context, many studies are trying to uncover the lipid signatures of disease with promising results in insulin resistance[29] and NAFLD[27]. Determining the lipid profile of NAFLD may help to stablish groups of NAFLD patients with specific lipid metabolic profiles that better respond to a specific treatment.

1.1.6 Glucose and fructose metabolism in lipogenesis and NAFLD

Dietary carbohydrates drive DNL and, as mentioned before, glycolysis and lipogenesis are connected since the main glycolysis product (pyruvate) provides a carbon source for lipogenesis. GCK (or hexokinase, HK) is the enzyme that catalyzes the first reaction of glycolysis and a variant of the glucokinase regulatory protein (GCKR) that negatively regulates it, has been found to be associated with hepatic steatosis and hyperglycemia in obese patients[1,30].

Glucose catabolism through the pentose phosphate pathway provides additional NADPH to carry out DNL. The main enzymes catalyzing the generation of NADPH (glucose-6-phosphate dehydrogenase and 6-phosphogluconate dehydrogenase) are likely to be involved in the regulation of lipogenesis (**Figure 4**).

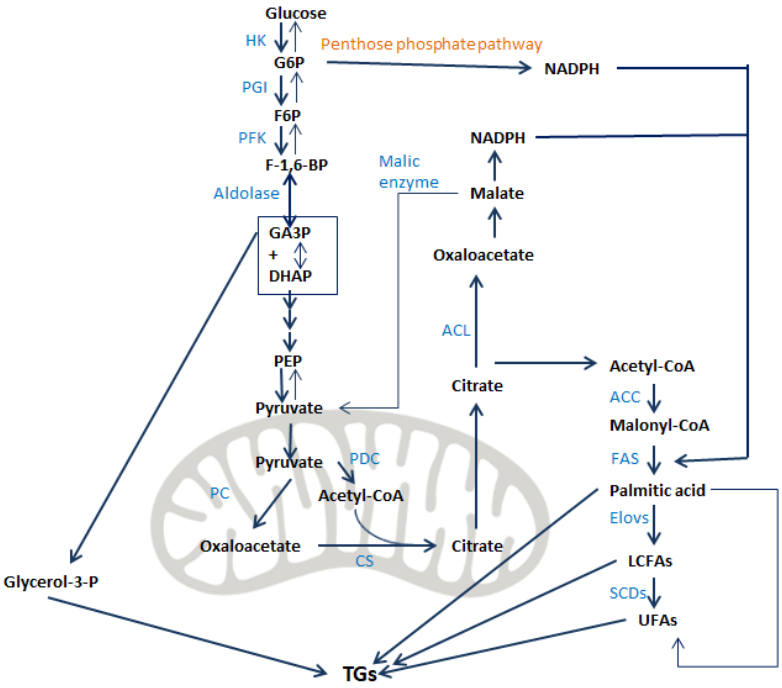


Figure 4. Lipogenic pathways. In blue appear lipogenic enzymes: ACC: acetyl-CoA carboxylase, ACL: ATP-citrate lyase, Elovs: fatty acyl-CoA elongases, FAS: fatty acid synthase, SCDs: stearoyl-CoA desaturases. Many steps are reversible by gluconeogenic enzymes.

Metabolized fructose provides fuel for central carbon metabolism[31]. Diets rich in sugar, particularly high fructose corn syrup (which is present in most daily products people consume: cookies, breakfast cereal, soft drinks, etc.), may lead to development of NASH by impairing insulin sensitivity[32]. Many studies demonstrated that fructose is lipogenic and stimulates TG synthesis. For instance, long-term administration of fructose to rats results in hepatic steatosis with an increase of 198% in hepatic TGs and 89% increase in hepatic cholesterol concentration[33]; ducks fed high fructose diets develop fatty liver[34]; and splanchnic perfusion studies demonstrate that fructose produces higher rates of TGs secretion from the liver than equimolar amounts of glucose[35].

The metabolism of fructose and glucose are different because before converging in the glycolytic pathway, fructose metabolism involves its phosphorylation by the action of fructokinase (ketohexokinase, KHK), generating fructose-1-phosphate (F1P) using ATP. In contrast to GCK, the phosphorylation of fructose to F1P by KHK is specific for fructose and the high activity of this enzyme in the liver could result in hepatic ATP depletion. This could be one of the mechanisms by which excess fructose consumption cause NAFLD[32].

1.1.7 NAFLD diagnosis and current therapies

NAFLD diagnosis is heterogeneous and relies on a variety of assessment tools that include radiological tests such as ultrasonography, blood test of liver enzymes and liver biopsy[36]. Nowadays, liver biopsy is the gold standard to diagnose NAFLD and identify the presence of NASH and fibrosis in patients. Due to the invasiveness of the technique, and taking into account the complications that may cause in the patient, numerous ongoing studies[37] use multiomics approaches[38] for the identification of noninvasive NAFLD biomarkers in blood to provide a diagnose of the disease and to check on its progression and response to therapy. Metabolomics, genomics, transcriptomics and proteomics are promising fields that steadily progress towards this goal[13].

Despite representing the most common cause of chronic liver disease in western countries and the huge investment carried out by the pharmaceutical industry, so far, there is no treatment for NAFLD approved by the Food and Drug Administration agency (FDA) or the European Medicine Agency (EMA). As mentioned before, NAFLD is tightly related to diabetes and obesity so, the current recommendations of clinicians are based in weight loss and control of diabetes as it is shown to slow down the progression of the disease[3].

To control diabetes, insulin sensitizers or glitazones are used. Glitazones are antidiabetic drugs[39] that promote differentiation of insulin-resistant pre-adipocytes into proliferative insulin-sensitive adipocytes through the direct activation of PPARγ avoiding the delivery of FAs to the

liver and enhancing FA uptake in these adipocytes. Their importance came since insulin resistance (IR) is a feature present in almost every NASH patient. That said, their beneficial effects are short-living after treatment and undesired side effects (bone loss and weight gain) have been reported[40]. Other insulin sensitizers acting as agonist of different metabolic pathways are being used as metabolic pathways activators as for example: agonist of FXR transcription factor (implicated in bile acid and cholesterol metabolism and homeostasis) or similarly, agonists of PPARα and PPARδ have shown to reduce or inhibit hepatic DNL and increase FA oxidation, reducing liver inflammation and fibrosis[41,42]. PPARα/δ agonist in NASH trials have shown improvement in dyslipidemia, inflammation, liver function and IR[43,44]. There is also a study showing the hepatoprotective effect of vitamin E and the possible positive effects of this vitamin in hepatic disease[45] but while interesting, these results are not conclusive.

Several possible therapies with different mechanisms of action are currently under clinical development[46]. Actually, there are more than 40 molecules[47] in clinical trials for NASH treatment, some of them in phase III like obeticholic acid and arachidyl-amido cholanoic acid (Aramchol) (**Figure 5**). Obeticholic acid or the bile acid derivative 6-ethylchenodeoxycholic acid is a potent activator of the FXR that reduces fibrosis and liver fat content in animal models of fatty liver disease[41,48]. On the other hand, a phase IIa study, showed that 3 months administration of Aramchol to patients with NAFLD significantly reduced liver fat content[49]. The phase IIb has been successful too and the molecule is currently in phase III of the study.

In the interim, until effective treatments are found, patients are told to endeavor to change their bad lifestyle habits, incorporate healthy food in the diet and practice exercise. Currently, this is the only therapeutic strategy that can stop (and sometimes improve) the progression of NAFLD[6,7].

1.1.7.1 Aramchol mechanism of action

3β-arachidyl-amido, 7α-12α-dihydroxy, 5β-cholan-24-oic acid (Aramchol; Trima Israel Pharmaceutical Products Ltd, Maabarot, Israel) is a synthetic lipid resulting from the conjugation of a bile acid and a saturated FA, cholic acid and arachidic acid respectively through an amide bond[49]. As we mentioned before, SCD1 is a key enzyme involved in FA metabolism in the liver and in *in vitro* models, Aramchol achieves a 70-83% inhibition of this enzyme activity. SCD1 inhibition promotes FA β-oxidation and decreases FA synthesis resulting in decreased hepatic storage of FA esters and TGs[49,50,51]. Moreover, Aramchol administration to animal models fed with a high-fat diet, significantly reduced hepatic fat content[52].

GALMED Pharmaceuticals (Tel Aviv, Israel) is a clinical-stage biopharmaceutical company focused in the research and development of therapies based in Aramchol for the treatment of NASH. In the phase IIb study in patients with NASH, ARamchol for the REsolution of STeatohepatitis (ARREST; study NCT 02279524), GALMED evaluated the efficacy and safety of two doses of Aramchol (placebo, 400mg/day and 600mg/day) in 247 NASH patients successfully. Currently, Aramchol is in phase III clinical trial (**Figure 5**) and ARMOR study is being carried out in a total of 2000 subjects.

Although there are multiple therapeutic targets being evaluated and also other molecules in phase III of study (**Figure 5**), in this work we have investigated the effect of Aramchol in primary hepatocytes and mice model of NASH. So, promising previous results together with the fact that the phase IIb in humans was successful make Aramchol a possible and reliable drug for NASH treatment.

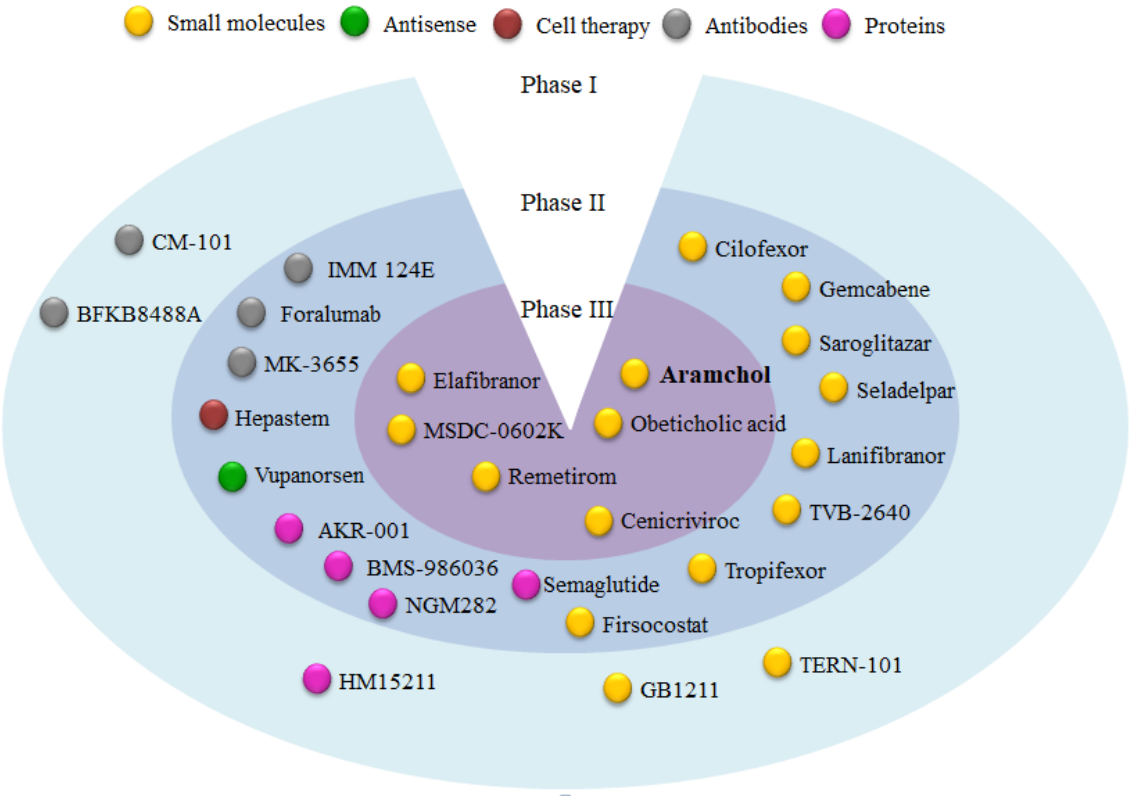


Figure 5. Clinical trials of significant interest for NAFLD/NASH therapies. Modified from IQVIA institute study 2019[47].

1.2 Methionine and SAMe metabolism in the liver

In the past and present centuries there have been many studies looking for the relationship between impaired methionine metabolism and the progression of liver disease. The first to show this link were Banting and Best, the discoverers of insulin, which in early 1930s observed that choline could prevent the development of steatosis in pancreatectomized diabetic dogs treated with insulin. Later work carried out in mice and rats revealed that diets deficient in methionine, choline as well as other methyl group donors (betaine, folates) produced liver steatosis; and that the prolonged administration of these diets, caused steatohepatitis, fibrosis and HCC. This work linked for the first time diabetes with steatosis and also manifested the importance of methyl groups in normal liver function and metabolism[53,54.] Later, Kinsel demonstrated that plasma methionine (intravenously injected) clearance was delayed in patients with cirrhosis linking liver disease with hypermethionemia[55]. All these early observations have been confirmed by studies in patients and mice lacking key enzymes involved in methionine and folate metabolism. Besides, the use of methionine metabolites like SAMe in animal models has been shown to be hepatoprotective[56].

Methionine is a sulfur-containing amino acid that plays a key role in the liver. It is mostly metabolized by this organ to be converted in SAMe (also named SAM or AdoMet) by the action of the enzyme methionine adenosyl transferase (MAT I and MAT III) using ATP as co-substrate[57]. In liver disease, elevated hepatic levels of methionine are due to a decrease in MAT activity, which consequently causes a reduction in hepatic SAMe[58]. SAMe is the most important biological methyl donor and, although it is produced in all cell types, is mainly synthesized and catabolized in the liver. Approximately 50% of dietary methionine is metabolized into SAMe by the liver where about 85% of all transmethylation reactions occur[59,60]. As SAMe is the major biological methyl donor, it is involved in transmethylation reactions, polyamine synthesis and in the transsulfuration pathway that generates glutathione (GSH), an important antioxidant agent of the cell. SAMe is also the key methyl donor in PC synthesis, required for VLDL formation and export of TGs from the liver[61]. In the transmethylation pathway SAMe donates its methyl group to sugars, phospholipids, proteins, RNA and DNA by means of specific methyltransferases[56,53,62]. A reduction in SAMe and GSH levels has been detected in many different types of liver diseases; and treatment with SAMe in patients with alcoholic liver cirrhosis has been found to rise GSH levels and increase survival[63,59].

Methionine serves as the precursor of other sulfur-containing amino acids such as homocysteine, cysteine and taurine, via the methionine cycle (**Figure 6**). The sulfur atom of methionine has a methyl group covalently bound which may be transferred to a large variety of acceptor molecules upon its activation to form SAMe. This reaction is catalyzed by the enzyme

MAT, which uses ATP as a co-substrate. SAMe can be demethylated by numerous methyltransferases (MTs), being the most abundant in the liver glycine N-methyltransferase (GNMT), to be converted in S-adenosylhomocysteine (SAH) which is a competitive inhibitor of many methyltransferases, although not GNMT. GNMT is an important enzyme since maintains the ratio SAMe/SAH which is an indicator of the methylation capacity of the cell[64]. SAH hydrolase catalyzes a reversible reaction which converts SAH in adenine and homocysteine, the latter being able to enter into two pathways: the remethylation pathway and the transsulfuration pathway.

The regeneration of methionine from homocysteine through the remethylation pathway can be catalyzed by two different enzymes: betaine homocysteine methyltransferase (BHMT) and methionine synthase (MS). BHMT is an enzyme exclusive of liver and renal tissues and needs betaine as co-substrate. MS-mediated remethylation of homocysteine is coupled to the folate cycle and requires vitamin B_{12}. 5-methyltetrahydrofolate (5-MTHF) is the methyl donor used by MS to convert homocysteine into methionine and tetrahydrofolate (THF). THF is then converted to 5,10-MTHF by the enzyme 5,10-MTHF synthase (MTHFS), and the enzyme MTFH reductase finally regenerates 5-MTHF completing the folate cycle.

As mentioned above, homocysteine can also undergo the transsulfuration pathway where it is used as a substrate of cystathionine β-synthase (CBS), which requires vitamin B6 as cofactor, to generate cysteine and GSH[65,66]. The transsulfuration pathway is very important in the liver due to its high activity in this organ. The entrance of homocysteine in the transsulfuration pathway or in the remethylation pathway is regulated by SAMe. Thus, SAMe inhibits the remethylation enzymes MTHF and MS, whereas activates CBS, the first step of the transsulfuration pathway[67,68].

SAMe may also be used for polyamine synthesis, where it is first decarboxylated by the enzyme SAMe decarboxylase. The aminopropyl group of SAMe is then transferred to putrescine to form spermidine and then a second aminopropyl group is transferred to spermidine to form spermine. These reactions generate 2 molecules of 5'-methylthioadenosine (MTA). MTA is and inhibitor of methylation reactions, S-adenosylhomocysteine hydrolase (SAHH) activity and polyamine synthesis; and through the methionine salvage pathway MTA may be metabolized to regenerate methionine[69].

When there is an alteration in the expression of any of the key enzymes involved in the methylation cycle (MAT, GNMT, SAHH, MS, BHMT, CBS, MTHFR), the cycle gets dysregulated and the alteration in the homeostatic concentration of its products trigger liver disease. Both, increased levels of SAMe (caused by the downregulation of GNMT) and decreased SAMe levels (caused by the downregulation of MAT1A) cause liver steatosis that

progresses to steatohepatitis, fibrosis and finally liver cancer[70]. These results indicate that SAMe concentration in the liver must be tightly regulated to prevent the development of liver disease.

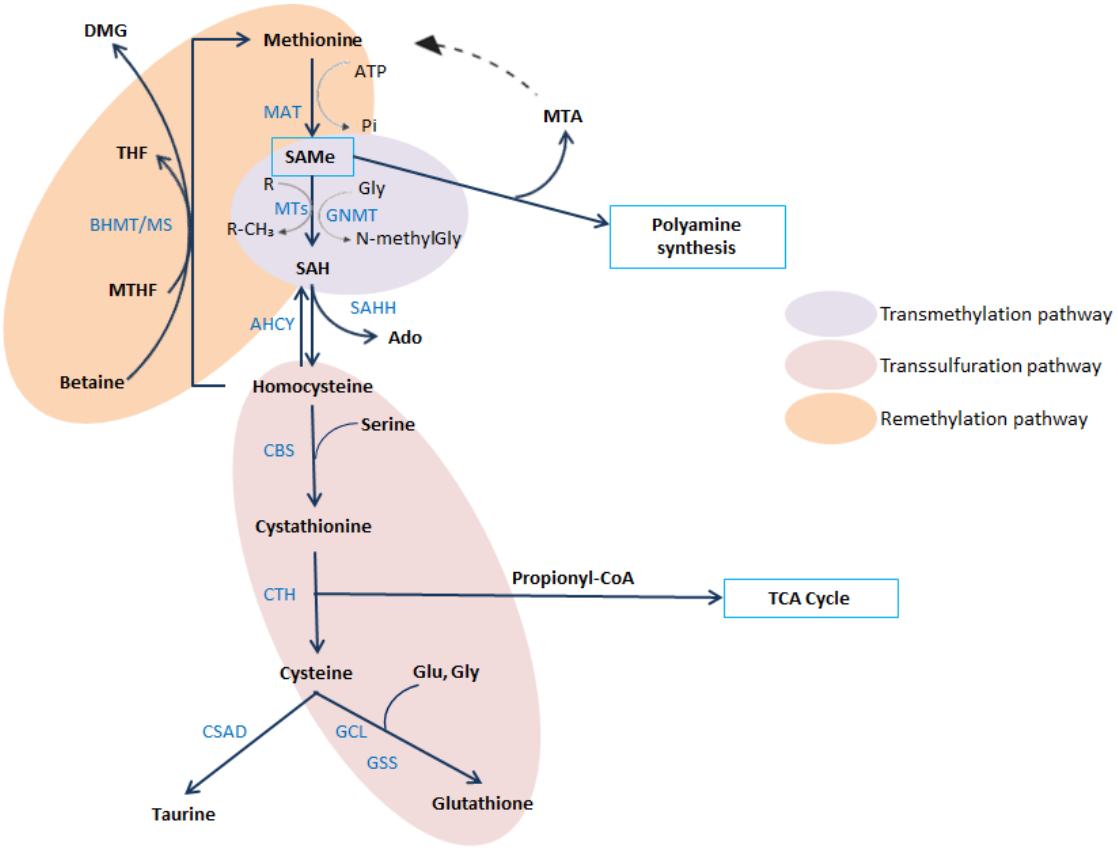


Figure 6. Methionine cycle and SAMe metabolism in the liver

1.2.1 Animal models of NAFLD

There are several animal models (genetic and dietary) that allow the study of NAFLD. Unfortunately, the majority of the models available do not reflect the sequence of histopathological features that characterize NAFLD progression in humans: steatosis, inflammation, ballooning and fibrosis.

1.2.1.1 Genetic models of NAFLD

Genetic murine models of NAFLD are generally based in knocking down enzymes involved in lipid and carbohydrate metabolism, as for example Ob/Ob mice and db/db mice, which are two important classical genetic models based on absence or deficiency of leptin signaling (a hormone that regulates lipogenesis and appetite). Ob/ob mice have a mutation in the leptin gene and develop IR, obesity, steatosis, hyperglycemia and hyperinsulinemia but do not progress to

steatohepatitis nor develop fibrosis and a second stimulus is needed[71]. Db/db mice have a leptin receptor that is not functional so these mice are insulin resistant, obese and have hepatic steatosis but do not progress spontaneously to steatohepatitis requiring also a second stimulus[71].

Lldlr-KO mice, in turn, are knock outs in low density lipoprotein receptor (LDLr). These mice are a proper model for atherosclerosis[72] but in combination with a high fat diet (HFD), they also develop NASH[73]. Actually, they develop higher plasma levels of cholesterol and phenotypically shows metabolic features similar to that of high-risk NASH patients, including insulin resistance[48,74].

Other animal model are the sterol regulatory element-binding protein (SREBP)-1c transgenic mice which develop NASH and insulin resistance but also followed by a decrease in adipose tissue mass[27,75]. Finally, there are other models that develop steatosis through the disruption of FA oxidation (PPARα-KO mice) or peroxisomal *β*-oxidation (AOX null mice).

In our laboratory we work with two genetic murine models of NAFLD knocked-out in two genes that, *a priori*, do not seem to be directly related with lipid metabolism: *MAT1A*-KO[76] and *GNMT*-KO[77] mice. These murine models have disrupted key enzymes in methionine and SAMe metabolism[76,78] and they are characterized by spontaneous development of steatosis and NASH. Alterations in these enzymes have been described in NAFLD patients. As a deficiency in MAT and GNMT enzymes leads to a deficiency and excess of SAMe respectively, and both facts cause liver disease, it seems clear that SAMe levels must be stable to avoid NAFLD development.

The mechanisms by which *MAT1A* is involved and related with SAMe metabolism are explained in next sections.

1.2.1.2 Methionine adenosyltransferase

SAMe is the principal biological methyl donor and also a key metabolite in polyamine biosynthesis[41]. Methionine adenosyltransferases (MATs) are enzymes that catalyze the formation of SAMe transferring the methyl group of methionine to the adenosine moiety of ATP liberating its tripolyphosphate moiety. *MAT1A* and *MAT2A* are the two genes (located in different chromosomes) that in mammals encode for α_1 and α_2, the two homologous MAT catalytic subunits. *MAT1A* is expressed in adults and in differentiated liver whereas *MAT2A* is expressed in fetal liver and extrahepatic tissues and is associated with rapid growth or dedifferentiation. α_1 subunits can form dimers (MAT III) or tetramers (MAT I), whereas α_2 forms tetramers (MAT II). The MAT gene is very well conserved throughout evolution as is one of the genes required for survival of organisms[39,79,80].

In several liver diseases as HCC, *MAT1A* is downregulated while *MAT2A* is overexpressed. In murine liver, loss of *MAT1A* and increased *MAT2A* expression have been found in situations of rapid cell growth such as liver regeneration after partial hepatectomy[81] and hepatocyte dedifferentiation to fibroblast in culture[82].

DNA methylation has an important implication in the regulation of MAT genes and there is a clear relation between liver cancer and the switch in gene expression from *MAT1A* to *MAT2A*[83]. The promoter of *MAT1A* has been found to be hypermethylated in two CpG sites in extrahepatic tissues, fetal liver, cirrhosis, HCC and hepatoma cell lines[70,59]. On the contrary, the promoter of *MAT2A* has been found to be hypomethylated in HCC[84].

When dealing with MAT enzymatic activity and regulation, its various isoforms differ in their kinetic and regulatory properties. MAT I and MAT II have low K_m values for methionine (23 µM-1 mM and 4-10 µM respectively) whereas the K_m of MAT III is high (215 µM-7 mM)[53,85]. MAT II is strongly inhibited by physiological concentrations of SAMe (IC_{50}= 60 µM) whereas MAT I is slightly inhibited (IC_{50}= 400 µM) and MAT III, on the contrary is strongly activated by SAMe[86,87]. On the other hand, MAT I and MAT III are reversibly inactivated by nitric oxide (NO) and hydroxyl radicals and mechanisms as NO and ROS production are implicated in the regulation of its enzymatic activity by switching MAT I and MAT III to inactive conformational forms[59].

1.2.1.2.1 *MAT1A* knock-out mouse model

MAT1A-KO mice lack *MAT1A* gene and consequently show a deficiency in MAT I/ MAT III enzymes. As partially explained in **section 1.2.1.1**, these mice are characterized by having reduced SAMe levels compared to WT, increased blood levels of methionine, decreased GSH levels and high levels of hepatic TGs[83]. All these facts, together with deficient VLDLs formation[21], contribute to steatosis development in this animal model. The lack of this gene results in the spontaneous development of NASH at 8 months and HCC at 18 months.

1.2.1.3 Dietary models of NAFLD

There are several dietary models to study the onset and progression of NAFLD. Ideally, animal models should completely reflect the features of human disease such as steatosis, ballooning, inflammation and fibrosis. However, no existing model can reproduce the entire human NAFLD phenotype. These dietary murine models include the traditional high-fat and high-fructose diets that causes fatty liver by excess of fat and fructose[88], and the methionine-choline deficient (MCD) diet model that causes fatty liver due to the deficiency in methionine and choline[71]. As mentioned above, methionine and choline are precursors of SAMe and PC, respectively, and

their deficiency compromises methylation reactions, antioxidant response and VLDLs formation[71].

Rodents fed with the MCD diet develop rapidly (within 2 weeks) steatosis, inflammation, release of transaminases, fibrosis and cell death. However, unlike humans they experiment a severe weight loss of up to 35% over 4 weeks and are not insulin resistant.

High fat diets (HFD) have been used to induced NAFLD and metabolic alterations in rodents. Compared to MCD diet, HFD requires longer times to induce liver steatosis and produces less severe liver injury than the MCD diet model. In both models, the damage depends on the mouse strain used[89,90]. HFD is composed mostly by dietary fat (71%) and depending on the duration of the feeding time, induces steatosis, oxidative stress, expression of proinflammatory cytokines[91,92], dyslipidemia, increased expression of lipogenesis regulators (SREBP1c, LXR), obesity and impaired glucose tolerance, which links this model diet tightly with MetS too.

As mentioned before, fructose consumption in humans causes features of MetS. In animal models, high fructose diet also induces features similar to MetS such as weight gain, hypertriglyceridemia, hypertension and IR. These effects are not observed when using other simple sugars such as glucose[88].

1.3 Non-alcoholic steatohepatitis (NASH)

The progression of NAFLD to NASH in humans comprises a series of behavioral, genetic and metabolic factors that are often tightly related with the MetS[93,94] while not completely understood. The progression of NAFLD has been proposed to be the result of the "multiple hit theory". According to this theory, there are 3 hits: during the first hit the accumulation of TG into LD takes place in more than 5% of hepatocytes. This first step of liver steatosis is usually associated with benign prognosis and may be reversible but it makes the liver prone to suffer a second hit which includes oxidative stress (ROS formation), mitochondrial dysfunction, release of pro-inflammatory cytokines[95,96] and gut-derived bacterial endotoxemia, which switches the liver to a necroinflammatory stage or NASH. The third hit includes impaired hepatocyte regeneration and the involvement of palatine-like phospholipase 3 (*PNPLA3*) gene[3].

On the other hand, NASH may or not be associated with fibrosis[97,98,99] which, at an early stage could be reversible. However, if maintained over time, fibrosis may progress to cirrhosis and HCC. As the prevalence of NAFLD is increasing worldwide, NASH-related cirrhosis and HCC will become the leading cause of liver transplantation in a near future being already the first cause of liver transplantation in women[100].

It is important to note that NASH is not a simple disease and it has a multifactorial etiology. Genetic, environmental, nutritional and behavioral factors can cause NASH (**Figure 7**).

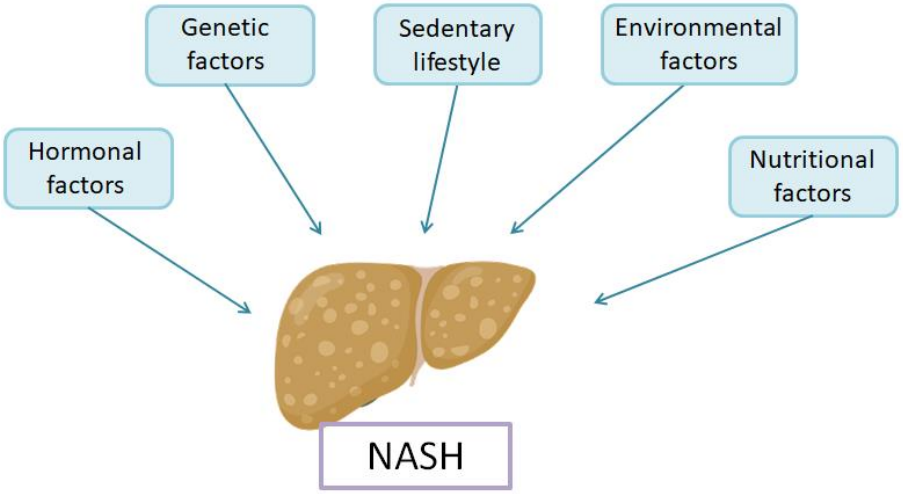


Figure 7. Several factors that may cause NASH

Due to the multifactorial etiology of NASH, different subtypes of patients (developing NASH through different mechanisms) may exist. If this were the case, the development of treatments adapted to every subtype might be required for controlling the disease in an effective manner.

Regarding this matter, new ways of diagnosis are needed. Personalized medicine and metabolomics are promising fields on this topic as we will see in **sections 1.4.3** and **1.5**.

1.4 Metabolic syndrome

The high prevalence of NAFLD is associated with the rise in MetS and obesity. MetS has elevated socioeconomic cost and is considered a worldwide epidemic[101]. The components of MetS are: IR, hyperglycemia or high fasting glucose levels (> 110 mg/dL), obesity (body mass index BMI > 30 kg/m²), dyslipidemia (TGs >> 150 mg/dL or low cHDL < 34,79 mg/dL in men and < 38,66 mg/dL in women), high arterial blood pressure (≥ 140/90 mmHg or medical treatment) and microalbuminuria (**Figure 8**). As shown in **Table 1**, there are different definitions for this disorder such as that of the World Health Organization (WHO), the National Cholesterol Education Program: Adult Treatment Panel III (NCEP:ATPIII) and the International Diabetes Federation (IDF). In this thesis we used the definition provided by the WHO.

WHO	ATPIII	IDF
Fasting glycemia ≥ 110mg/dL plus 2 or more of the following factors:	3 or more of the following factors:	
Obesity: BMI › 30 kg/m²	Obesity: waist circumference > 102 cm in men > 88 cm in women	Obesity: waist circumference depending on ethnic group.
TGs >> 150 mg/dL or cHDL < 34,79 mg/dL in men cHDL <38,66 mg/dL women	TGs ≥ 150 mg/dL cHDL < 40 mg/dL in men cHDL < 50 mg/dL women	TGs ≥ 150 mg/dL cHDL < 40 mg/dL in men cHDL < 50 mg/dL women
Hypertension: blood pressure ≥140/90 mmHg or medical treatment	Hypertension: blood pressure ≥130/85 mmHg	Hypertension: blood pressure ≥130/85 mmHg or medical treatment
Microalbuminuria: albumin excretion of 20µg/min	Fasting glycemia ≥ 110mg/dL or T2D	Fasting glycemia ≥ 100mg/dL or T2D

Table 1. Parameters to define metabolic syndrome following World Health Organization (WHO), Adult treatment panel III (ATPIII) and International Diabetes Federation (IDF)

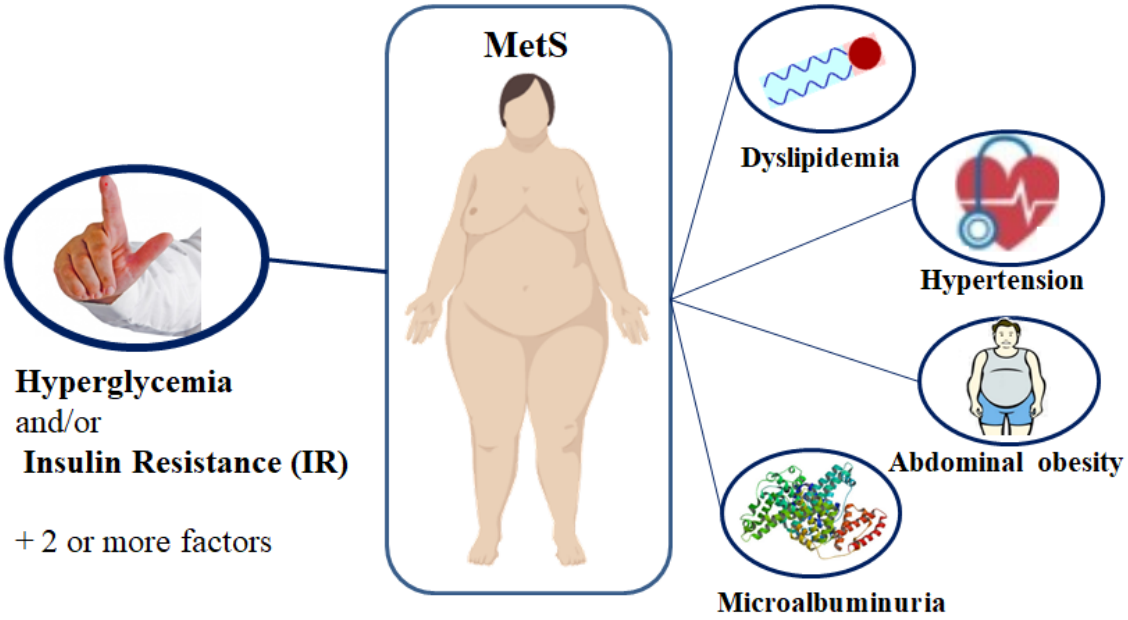


Figure 8. Schematic representation of MetS according to WHO

1.4.1 Obesity

Obesity is a heterogeneous condition defined by an excess of body fat. It can be estimated by the body mass index (BMI), a ratio of weight over height being expressed in kilograms per square meter[102].

High sugar consumption in western countries is the main cause of obesity[3]. Obesity and overweight are serious public health problems affecting a high number of people around the world and it is the main etiologic factor that predisposes to MetS and IR[103]. It also plays a key role in the development of complex diseases as T2D and some types of cancer. In Caucasian adults, overweigh and obesity are tightly associated with rise in mortality[104]. Most NAFLD patients are obese or overweight and have associated MetS (although lean individuals may also develop features of MetS and NAFLD)[105]. Diet and moderate physical activity (30 minutes of aerobic exercise 4 times a week) are a good method to lose weight and it is recommended to these patients, as there are many studies showing the benefit of weight loss also in NAFLD[106].

1.4.2 Insulin resistance

IR is a pathologic condition where organs fail to respond to insulin causing hyperglycemia setting the stage to organ damage and failure if maintained over time[107]. The multiple metabolic effects of insulin are mediated by different tissue-specific actions that include changes in protein phosphorylation and function as well as changes in gene expression. In a non-pathological situation, when blood glucose concentration is rising, pancreatic-β-cells release insulin which lowers blood glucose levels by increasing its uptake by muscle and adipose tissue cells and suppress hepatic glucose production. Adipocytes are one of the most insulin-responsive cells and insulin mechanism of action regulates numerous aspects of its biology: stimulates lipoprotein lipase activity incorporating FAs derived from circulating lipoproteins and also promotes TGs accumulation by inhibiting lipolysis and stimulating lipogenesis in mature adipocytes[103].

IR in obesity and T2D is manifested by decreased insulin-stimulated glucose transport and metabolism in adipocytes and skeletal muscle and by impaired suppression of hepatic glucose output[108]. Obesity and IR are tightly related since the more fat content (measured as BMI), the higher the risk for T2D and IR: this implies that the amount of body fat has an impact on insulin sensitivity[109,110]. It must be taken into account that not all fat depots in the body contribute the same to IR. Intra-abdominal fat depots (or central obesity) is much more strongly linked to IR, cardiovascular disease and T2D than peripheral fat depots[111].

1.4.3 Personalized medicine

Personalized or precision medicine refers to the individualized treatments that use new diagnostics and therapeutics that better fit the needs of a patient (or a group of patients) depending on their genetic, metabolic features and biomarker characteristics[112]. Personalized medicine has raised interest during the last decades and it is increasingly applied for diagnosis, clinical stratification and treatment of MetS and its associated risk, such as NAFLD, T2D and their complications[113].

As abovementioned, there is not a specific treatment approved for MetS or NAFLD, although an impressive effort is being currently applied in that direction. Due to the multifactorial etiology of these diseases and the possible existence of different subtypes, it is adequate and timely to characterize a large population of individuals to obtain a large data base that could help to define each subtype of patients and the most appropriate treatment. In this context, we aim to find subgroups of patients with similar genetic and/or metabolic features that could respond in a similar and more effective way to a specific drug or treatment.

To that end, we require non-invasive techniques of diagnosis and metabolomics is a promising one.

1.5 Metabolomics

Disease is a biological process that operates through interconnected and complex interactions between genes, RNA, proteins, metabolites and coupled to environmental factors[114]. Thus, an integrative view at the subcellular level is required to understand the molecular basis of the disease. Several “omic” sciences have emerged during the last years in order to have a holistic vision of this particular problem. The term “omic” refers to any type of study that gives collective information of a biological system. Nowadays the most representative omics fields include genomics, transcriptomics, proteomics and metabolomics, as well as other new omics sciences that are constantly emerging such as lipidomics, fluxomics, epigenomics, glycomics, foodomics, etc. Each omic technique is complementary for the understanding of biological systems and also expand the information generated in other fields or provided by other approaches[115].

Metabolites are biomolecules of low molecular weight (<1500 Da) that play central roles in biology serving as metabolic intermediates, energy sources, signaling molecules and building blocks. Their relative levels give information of biological functions and define the phenotype of biological systems in response to environmental or genetic changes[115].

Metabolomics can be defined as the study of all the metabolites present in a biological system and their variations in reaction to time and/or genetic or external stimuli. It aims to provide comparative and semi-quantitative information about all the metabolites present and its abundance. Metabolomics has applications in a wide variety of research fields including drug toxicity studies[116], functional genomics[117] and disease diagnosis[118].

The complete set of metabolites found in a cell, tissue, organ or organism can also be defined as the metabolome. The metabolome is dynamic and changes rapidly affected by genetic and environmental modifications so, it reflects the state of the system studied at a given time point[119]. The metabolomic field appeared in 1998[120] and it is becoming more and more popular in the last times since the metabolome can be considered the molecular phenotype of an organism (**Figure 9**) or biological system. This does not happen with the genome, as it is not always transcribed, with the transcriptome or with the proteome. The transcriptome does not necessarily correlate with the proteome and this, although translated, may or not be functionally active. So, metabolomics by reflecting the final substrates and products of the cellular metabolism of living organisms, allow a better understanding of the phenotype of a specific biological condition and disease state[115]. Definitely, as metabolomics allows us to obtain a metabolomic fingerprint[121] of a particular condition it is possible to identify potential biomarkers of disease. These biomarkers can be used clinically as diagnostic, screening or prognosis tools and studies to characterize the "normal" metabolome are required to determine the intra- and inter-individual variability and then compare them with "disease" metabolome[114].

Metabolomic experiments have many advantages but the most remarkable is, that compared with other techniques, a higher sample throughput is possible and, if combined with automation of sample preparation and loading, a measurement instrument can be measuring almost 24h a day. A problem is, however, how to analyze all the data generated and we will talk about data analysis tools in **section 1.5.5.**

All this is very useful in the case of systemic diseases as for example, MetS, obesity or T2D, which are disorders affecting the whole body and in order to understand the biochemical changes that take place in one organ, it is necessary to understand how it interacts with systemic metabolism[122].

Currently, for the study of metabolomics there are two main tools: Nuclear Magnetic Resonance (NMR) and Mass Spectrometry (MS).

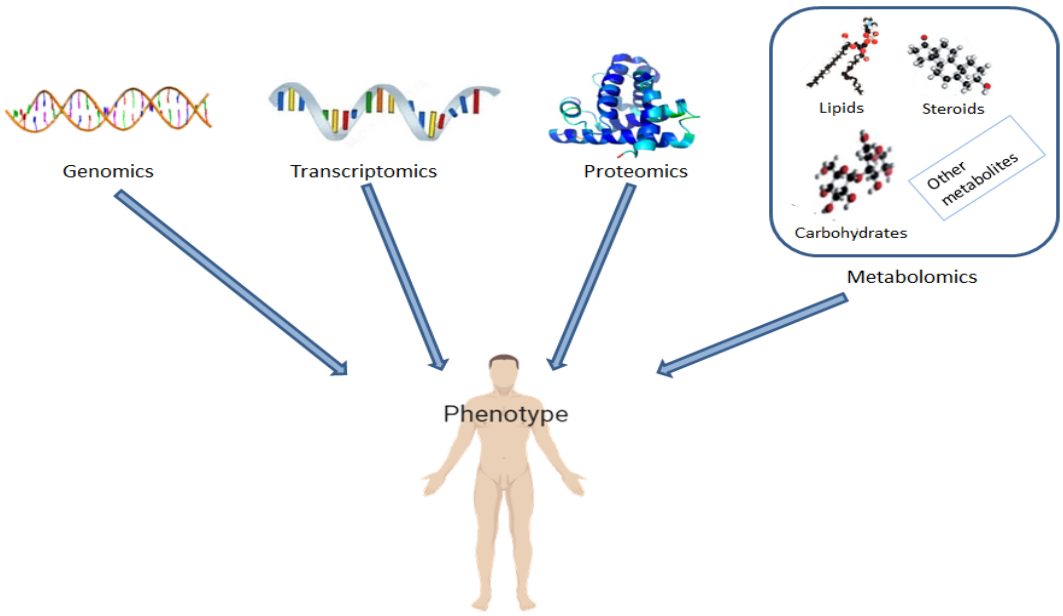


Figure 9. Main "omics" studied in biology

1.5.1 Mass spectrometry

MS is an analytical technique that measures the mass of a molecule by measuring the mass-to-charge ratio (m/z)[114]. It relies on identifying metabolites according to their mass after they have been ionized and accelerated across a distance. In GC/MS metabolites are first vaporized into the gas phase and introduced into a column placed in an oven where the temperature raises across a time cycle, making molecules vaporize at different times and then, obtaining a chromatographic separation. In LC/MS metabolites are also separated chromatographically according to their solubility in the liquid phase against their affinity for the solid phase contained within the column with the difference that the mobile phase (o carrier) is liquid and not gas[122]. The metabolites are finally ionized and analyzed by the mass spectrometer. Several kinds of MS instruments can be used and combined such as quadrupoles, ion traps, triple quads and time-of-flight mass analyzers. Nowadays, LC/MS is the most important MS-based approach for metabolomics because the sensitivity and resolution are high, requires few amount of sample for the analysis, and gives a lot of information[115]. LC/MS is able to detect a very large number of metabolites in a given sample (several hundreds).

1.5.2 NMR-metabolomics

A complementary and key technique for metabolomic studies is NMR spectroscopy. High resolution ^{1}H NMR spectroscopy is a rapid and nondestructive method also with minimal sample preparation that can detect any molecule containing hydrogen nuclei (protons). Other advantages of this technique are that it is robust, reproducible, fast and operates in a high

dynamic range of concentrations. On the other hand, NMR-based metabolomics requires higher amounts of sample and its intrinsic sensitivity is lower than that of MS.

NMR spectroscopy is based in the principle that some nuclei possess magnetic moment that can be measured when they are exposed to a magnetic field. The experiment contains information about the chemical and physical properties of molecules or atoms in which they are contained. In other words, nuclei placed in a magnetic field absorb and re-emit electromagnetic radiation (resonance) at a specific frequency depending on the magnetic moment of the nucleus, the chemical environment and the magnetic field[114,123]. Isotopes present in biological samples sensitive to the NMR phenomenon are ^{1}H and ^{31}P at natural abundance, while ^{13}C and ^{15}N often need to be isotopically enriched. In this context, some human samples that can be analyzed include cerebrospinal fluid (CSF), plasma, serum and urine. Also it is possible to analyze intact tissue samples by using solid state NMR; specifically, for the metabolic profiling of tissue specimens ^{1}H high-resolution magic angle spinning (HR MAS ^{1}H NMR) can be used[124]. Regarding liquid samples, urine and plasma are abundant in metabolites and 1D proton spectra can be recorded fast in a few acquisitions allowing the measurement of high number of samples[125] which makes possible the study of a large number of samples per day.

MS and NMR are often complementary techniques that, when combined, can help to have a better understanding and vision of changes in whole system metabolism of a specific organism[126]. These two techniques can be used for targeted and untargeted metabolomics analysis.

1.5.3 Non-targeted metabolomics

Non-targeted metabolomics is useful for hypothesis generation as it measures as many metabolites as possible present in different samples to later compare them. The aim when using non-targeted metabolomics approaches is to have a profile of the whole metabolome present in cell, tissues or biofluids. Careful preparation of samples and proper analytical instrumentation are important to detect the maximum number of metabolites in a reproducible way. The main complication when analyzing these datasets is the identification of unknown features[127] because as there are metabolites with various physicochemical properties found from low picomolar to millimolar concentrations, it is difficult to observe all the metabolites present in a biological sample at once[115].

1.5.4 Targeted metabolomics

Targeted profiling aims to identify and quantify all detectable metabolites in a given spectra prior to subsequent data analysis[128]. This metabolic analysis is useful when we want to answer specific biological or biochemical hypotheses. This approach gives (semi)quantitative

information about a predefined list of metabolites of interest. Metabolites are identified and quantified by comparing the LC/MS or NMR spectrum of the sample of interest to a set of standards or to a spectral reference library obtained from standards. This approach requires the compounds of interest to be known *a priori*[128]. The information obtained here is more accurate than in non-targeted analysis and low-level metabolites which are often bioactive and play important roles in biological systems can also be detected[115].

It is common practice to carry out both non-targeted and targeted analysis to generate and test a hypothesis respectively.

1.5.5 Data analysis

One of the most challenging aspects in metabolomics is how to deal with the complexity and huge volume of biochemical information obtained from both, MS and NMR. Extraction of relevant information from these data is possible due to the advances in statistical and modeling methods so, biostatistics and bioinformatics are essential in this field.

Raw data is first converted into computer-readable formats compatible with software packages[114]. Dataset is then pre-processed in a specific way, depending it comes from MS or NMR, but it always includes normalization, scaling, outlier removal and imputation of missing values between others[114,123]. Subsequently, multivariate and univariate data analysis can be performed and in general, a combination of both techniques is advantageous as different and complementary information can be extracted. The usual goal of metabolomics is to identify differences between experimental groups finding differences in the metabolites present in each class. With hundreds of metabolites in each class, it is not practical to visualize changes between experimental groups by analyzing metabolites individually so, researchers use tools that allow the visualization of changes in the metabolome like: Principal Component Analysis (PCA), MetaboAnalyst web server[128], Human metabolome database, the Madison Metabolomics Consortium Database[129] and Workflow4Metabolomics[130] among others.

For example, MetaboAnalyst is an easy-to-use web server for metabolomic data analysis created to allow the users implement an accessible metabolomic data analysis. It accepts a variety of input data such as NMR peak list, binned spectra, MS peak list, compound/concentration data in several formats. The steps for a complete analysis in this source include: data upload, data processing and data integrity checking, data normalization, data analysis, data annotation (peak search and pathway mapping) and summary report download. Through MetaboAnalyst, techniques as t-test, fold change analysis, hierarchical clustering analysis, PCA, PLS-DA and more statistical and machine learning methods can be carried out[128] (**Figure 10**). It also includes

a large library of reference spectra for compound identification originally developed for the Human Metabolome Data Base (HMDB) and MetaboMiner[131].

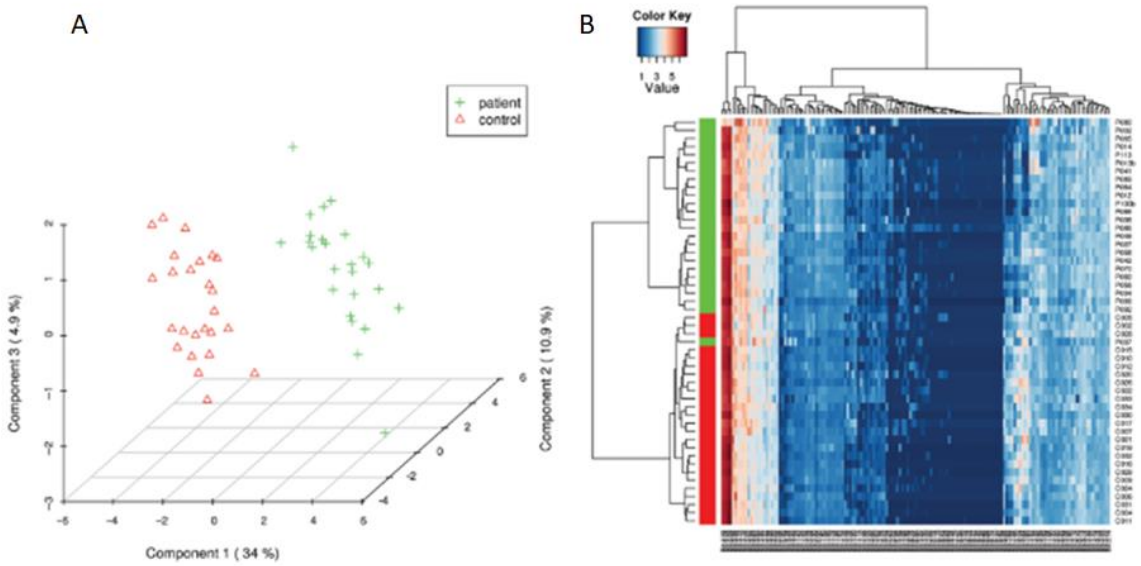


Figure 10. Examples of some graphics generated from MetaboAnalyst. (**A**) PLS-DA class separation. (**B**) heat map generated from hierarchical clustering. Modified from Xia, J *et al*. 2009[128]

Metabolome databases are growing and expanding every year, there are still a lot of metabolites unidentified in biological systems. There are some initiatives focus towards the creation of a central reporting database to share methodologies and results between laboratories[129,132].

1.5.6 Fluxomics

Fluxomics is a subset of targeted metabolomics that tries to determine metabolic fluxes. It allows the quantification of small molecule fluxes through metabolic networks providing access to the *in vivo* activity of pathways and reactions in intact living cells. It integrates *in vivo* measurements of metabolic fluxes with stoichiometric network models for the determination of absolute fluxes through large networks of the central carbon metabolism[133]. Fluxomic (as well as metabolomic) studies, need the use of NMR, liquid chromatography, mass spectrometry and the use of tracer compounds[134].

Fluxomic studies can be implemented using labelled precursors of carbon and nitrogen for example, labelled compounds that have the same chemical properties as the unlabeled ones but an increased mass. This stable labelled isotopes usually are ^{13}C, ^{15}N ^{2}H, ^{18}O and ^{34}S resulting in peaks of grater mass when analyzing them in LC/MS[135].

CHAPTER 2

EXPERIMENTAL PROCEDURES

2 EXPERIMENTAL PROCEDURES

2.1 Protein extraction and analysis

2.1.1 Total protein extraction

Plated cells were washed twice with PBS buffer and homogenized in RIPA lysis buffer (1.6 mM NaH_2PO_4, 8.4 mM Na_2HPO_4, 0.5% Azide, 0.1 M NaCl, 0.1% SDS, 0.1% Triton X-100, 5 mg/ml sodium deoxycholate) supplemented with protease and phosphatase inhibitor cocktail (Roche, Switzerland). When extracting proteins from frozen liver tissue, about 50mg of tissue was homogenized in 1 mL lysis buffer by using Precellys tissue homogenizer (Precellys, France). In all cases the lysates were centrifuged at 13000 rpm during 20 min at 4 ºC. Supernatants (protein extract) were collected and quantified for total protein content by the BioRad protein assay, or by BCA protein assay (Pierce, USA) when measuring samples with high fat content.

2.1.2 Western blotting

Protein extracts were boiled at 95 °C for 5 min in SDS-PAGE sample buffer (250 mM Tris-HCl pH 6.8, 500 mM β-mercaptoethanol, 50% glycerol, 10% SDS and bromophenol blue). An appropriate amount of protein (between 5 and 30 μg, depending on protein abundance and antibody sensitivity), were separated by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) in 8% to 12% acrylamide gels (as a function of the molecular weight of the protein of interest), using a Mini-PROTEAN Electrophoresis System (Bio-Rad). Gels were transferred onto nitrocellulose membranes by electroblotting using a Mini Trans-Blot cell (Bio-Rad). Membranes were blocked with 5% nonfat milk in TBS pH 8 containing 0.1% Tween-20 (Sigma Aldrich) (TBST-0.1%), for 1 hour at room temperature, washed three times with TBST-0.1% and incubated overnight at 4°C with commercial primary antibodies. Primary antibodies used for western-blot are detailed in **Table 2**. Membranes were then washed three times with TBST-0.1% and incubated for 1 hour at room temperature in blocking solution containing secondary antibody conjugated to horseradish-peroxidase (HRP). Immunoreactive proteins were detected by using Western Lightning Enhanced Chemiluminescence reagent (ECL, PerkinElmer, USA) and exposed to Super Rx-N X-ray films (Fuji, Japan) in a Curix 60 Developer (AGFA, Belgium). Bands were quantified by densitometry using the free image processing software ImageJ (http://rsbweb.nih.gov/ij).

Protein	Antibody reference	Supplier
ACCα/β	3676S	Cell Signaling Technology, Inc.
Phospho-ACCα/β (S79)	3661S	Cell Signaling Technology, Inc.
β-actin	A5441	Sigma-Aldrich
AMPKα	07-350	Millipore
Phospho-AMPKα (T172)	2531S	Cell Signaling Technology, Inc.
CPT1A/B	SC-393070	Santa Cruz Biotechnology
S6	2317S	Cell Signaling Technology, Inc.
Phospho-S6 (S235/236)	4857S	Cell Signaling Technology, Inc.
SCD1	2794S	Cell Signaling Technology, Inc.
p70S6K	2708S	Cell Signaling Technology, Inc.
Phospho-p70S6K (T389)	9234S	Cell Signaling Technology, Inc.

Table 2. Primary antibodies used in the western blots.

2.2 RNA isolation and real-time polymerase chain reaction

Total RNA was isolated with Trizol (Invitrogen, Carlsbad, CA). An aliquot of 1 μg of total RNA was treated with DNAse (Invitrogen) and reverse transcribed into complementary DNA using M-MLV Reverse Transcriptase (Invitrogen). Quantitative real-time polymerase chain reaction (RT-PCR) was performed using SYBR Select Master Mix (Applied Biosystems, Foster City, CA) and the Viia 7 Real-Time PCR System (Applied Biosystems). PCR was executed with the following primers: Collagen, type I, alpha 1F (COL1A1F), 5'- TTGACCAACCGAACATGACC -3'; COL1A1R, 5'- GCAGAAAGGGACTTACCCCC -3'; peroxisome proliferator-activated receptor gamma F (PPARγF), 5'- GCTGTGCAGGAGATCACAGA -3'; PPARcR, 5'- GGGCTCCATAAAGTCACCAA -3'; GAPDHF, 5'- AATGAAGGGGTCATTGATGG -3'; GAPDHR, 5'- AAGGTGAAGGTCGGAGTCA -3'. Expression levels were normalized to the level of GAPDH messenger RNA (mRNA) in each sample[136].

2.3 Cell isolation, culture and treatments

2.3.1 Primary mouse hepatocytes isolation and culture

Primary hepatocytes from wild type (C57BL/6J strain), *MAT1A*-KO and *GNMT*-KO mice were isolated by perfusion with collagenase Type I (Worthington, USA). Animals were anesthetized with isoflurane (1.5% isoflurane in O_2); then, the abdomen was opened and a catheter was inserted into the inferior vena cava. Liver was perfused with buffer A (Phosphate buffered saline (PBS), 5mM EGTA. 37 ºC and oxygenated) and the portal vein was cut. After that, liver was perfused with buffer B (PBS, 37 ºC and oxygenated) to remove EGTA, and finally perfused

with buffer C (PBS 1mM $CaCl_2$, collagenase type I, 37 ºC and oxygenated). After buffer C perfusion, liver was extracted from the body and placed into a petri dish with Minimum Essential Medium (MEM) (Gibco, USA). Gall bladder was carefully removed and then, the liver was mechanically disaggregated with forceps. The digested liver diluted in MEM was then filtered through a sterile gauze, and the filtered liver cells were collected and washed three times in MEM (400 rpm 4', 1x; and 500 rpm, 5', 2x). After the final wash, pellet containing hepatocytes was resuspended in 10% fetal bovine serum (FBS) (Gibco) MEM supplemented with 1% penicillin, streptomycin and glutamine (PSG) (Gibco), for subsequently culturing. Isolated hepatocytes were seeded over collagen type I-coated culture dishes (5 x 10^5 cells/dish) in 10% FBS MEM supplemented with 1% PSG and maintained in a 5% CO_2-95% air incubator at 37 °C. After 2-3 hours of attachment, medium was replaced for fresh 0% or 10% FBS MEM depending on the treatments that are going to be carried out.

2.3.2 Test substances and Vehicle

The test substance Aramchol was provided by Galmed Pharmaceuticals. It was originally produced by WIL Research Europe B.V. as test substance number 204147/Y. The batch number was CS11-153Am-1402. All analytes were within the limit of quantifications. In addition, heavy metals and ions were also below the upper limits (irons < 10 ppm; arsenites < 1 ppm; heavy metals < 10 ppm; sulfate < 1% and chloride < 0.1%).

Compound was dissolved in DMSO, at a concentration of 100 mM, and stored at -20 °C.

2.3.3 Aramchol: preparation of stock solutions and treated primary hepatocytes

A 20 mM stock preparation of Aramchol was prepared by diluting the 100 mM stock for in vitro treatments. The Aramchol preparations were stored at -20 °C until each use.

Primary hepatocytes were isolated as described above and allowed to attach during 3 hours. After this, culture medium was removed and replaced by serum-free MEM or serum-free methionine and choline deficient medium (MCDM), with or without DMSO (vehicle) and Aramchol (20 µM), and cultured for an additional 48 hours.

2.3.3.1 Proteomic analysis

Samples were incubated in a buffer containing 7M urea 2M Thiourea 4% CHAPS and 5mM DTT for 30 min at RT under agitation and digested following the filter-aided FASP protocol described by Wisniewski *et al*[137] with minor modifications. Trypsin was added to a trypsin:protein ratio of 1:50, and the mixture was incubated overnight at 37 °C, dried out in a RVC2 25 Speedvac concentrator (Christ), and resuspended in 0.1% FA. Peptides were desalted and resuspended in 0.1% FA using C18 stage tips (Millipore).

Samples were analyzed in a novel hybrid trapped ion mobility spectrometry – quadrupole time of flight mass spectrometer (timsTOF Pro with PASEF, Bruker Daltonics) coupled online to a nanoElute liquid chromatograph (Bruker). This mass spectrometer takes advantage of a novel scan mode termed parallel accumulation – serial fragmentation (PASEF), which multiplies the sequencing speed without any loss in sensitivity[138] and has been proven to provide outstanding analytical speed and sensibility for proteomics analyses[139]. Sample (200 ng) was directly loaded in a 15 cm Bruker nanoelute FIFTEEN C18 analytical column (Bruker) and resolved at 400 nL/min with a 100 min gradient. Column was heated to 50 °C using an oven.

Protein identification and quantification was carried out using PEAKS software (Bioinformatics solutions). Searches were carried out against a database consisting of human entries (Uniprot/Swissprot), with precursor and fragment tolerances of 20 ppm and 0.05 Da. Only proteins identified with at least two peptides at FDR < 1% were considered for further analysis. Data was loaded onto Perseus platform[140] and further processed (log_2 transformation, imputation). A t-test was applied in order to determine the statistical significance of the differences detected and heatmaps and volcano plot were generated. Protein-Protein Interaction Networks were analyzed by the tool STRING software[141].

2.3.4 Human hepatic stellate cell line LX-2

The human hepatic stellate cell line LX-2 (Millipore Corporation, Darmstadt, Germany)[142] was cultured in Dulbecco's modified Eagle's medium supplemented with 2% FBS for 30 hours, after which the medium was replaced by serum-free Dulbecco's modified Eagle's medium. After 12 hours in culture, LX-2 cells were treated with Aramchol (10 μM) for 24 hours. For cell cultures, 100 mM Aramchol solution was prepared by dissolving 70 mg in 1 mL of DMSO. This stock solution was then further diluted 10 times in DMSO before treatment of cells at a final concentration of 10 μM. The final concentration of DMSO in culture media was 0.1%.

2.3.5 BODIPY assay

Primary hepatocytes in rat collagen type I-coated coverslips cultured in methionine/choline deficient medium (MDMC) were fixed in 4% paraformaldehyde for 10 minutes and incubated with BODIPY 493/503 (Molecular Probes, Invitrogen) at 1 mg/mL (1h, RT). BODIPY immunocytofluorescence images were taken using an Axioimager D1 (Zeiss) microscope. Quantification of lipid bodies was performed using Frida Software and represented as mean area per total number of cells.

2.3.6 Cellular reactive oxygen species (ROS) production

Cellular reactive oxygen species (ROS) production in primary hepatocytes was assessed using CellROX Deep Green Reagent (ThermoFisher, Waltham, MA). The hepatocytes were loaded with 1.5 µM CellROX in 10% FBS-MEM for 30 minutes at 378Cina CO_2 incubator. The hepatocytes were then carefully washed 3 times with PBS, collected, and analyzed by flow cytometry.

2.4 Determination of hepatocyte mitochondrial membrane potential

Hepatocytes from WT and *MAT1A*-KO mice were isolate and incubated with 4 mmol/L SAMe for 4 hours. To evaluate mitochondrial membrane potential the JC-1 dye (Mitochondrial Membrane Potential Probe; Life Technologies) was used according to the manufacturer's instructions. In brief, cells were incubated in 1.5 mL culture medium containing 10 mg/mL JC-1 for 10 minutes at 37º C in a 5% CO_2 atmosphere. Then, cells were washed twice with cold PBS, trypsinized, suspended in 1 mL PBS/1% FBS and immediately analyzed by flow cytometry using an LSRII flow cytometer (BD Biosciences, San Agustin de Guadalix, Madrid, Spain) and Flowjo software (FlowJo, Ashlan, OR) . The green fluorescence emission (JC-1 monomers) was monitored at 520 nm and the red-orange fluorescence emission (J-aggregates) was monitored at 590 nm. The aggregate/monomer ratio was calculated as a percentage of the maximal potential of wild-type hepatocytes (100%).

2.5 Fluxomic analysis

2.5.1 Uniformly ^{13}C-labelled glucose

Following hepatic perfusion, primary mouse hepatocytes were plated in 10% FBS containing MEM culture medium at a density of 500 thousand hepatocytes/well in 6-well plates. Once the hepatocytes were attached, 3 hours after the seeding, culture medium was replaced with 0% FBS containing MEM supplemented with 20 µM Aramchol or DMSO (vehicle). Primary hepatocytes were incubated during 48h and after this time, the medium was replaced with 0% FBS containing DMEM without glucose, supplemented with 11mM fully ^{13}C-labeled glucose. Plates were washed with PBS three times and snap frozen with liquid nitrogen after 10, 20, 30, 45, 60, 120 and 240 minutes. Samples were stored at -80 °C until metabolite extraction. All conditions and time points were performed in sextuplicate.

For the metabolite extraction, cell pellets were collected and lysed in 500 µL of a mixture of ice-cold water/methanol (50/50 v/v%) with a tissue homogenizer (Precellys) in a single 30 seconds cycle at 6000 rpm. Subsequently 400 µL of the homogenate was transferred to a new aliquot to which 400 µL of chloroform was added. The resulting solution was shaken at 1400

rpm for 1 hour at 4 °C. Next, the aliquots were centrifuged for 30 minutes at 14000 rpm at 4°C in order to separate the phases. From the organic phase (chloroform) 200 µL per sample was transferred to fresh aliquots and placed at -80 °C for 20 minutes and evaporated with a Speedvac. The resulting pellets were resuspended in 150 µL methanol.

Samples were measured with a UPLC system (Acquity, Waters, Manchester) coupled to a Time of Flight mass spectrometer (ToF MS, SYNAPT G2, Waters). A 2.1 x 100 mm, 1.7 µm BEH amide column (Waters), thermostated at 40 °C, was used to separate the analytes before entering the MS. Solvent A (aqueous phase) consisted of 99.5% water, 0.5% formic acid and 20 mM ammonium formate while solvent B (organic phase) consisted of 29.5% water, 70% MeCN, 0.5% formic acid and 1 mM ammonium formate.

In order to obtain separation of the analytes, the following gradient was used: from 5% A to 50% A in 2.4 min in curved gradient (#8, as defined by Waters), from 50% A to 99.9% A in 0.2 min constant at 99.9% A for 1.2 min, back to 5% A in 0.2 min. The flow rate was 0.250 mL/min and the injection volume was 2 µL. All samples were injected randomly. After every 8 injections a QC low and QC high sample were injected. All the samples were injected per duplicate. The MS was operated in negative electrospray ionization mode in full scan (50 Da to 1200 Da). The sampling cone voltage was 25 V and capillary voltage was 100 V. Source temperature was set to 120 °C and capillary temperature to 450 °C. The flow of the cone and desolvation gas (both nitrogen) were set to 5 L/h and 600 L/h, respectively. A 2 ng/mL leucine-enkephalin solution in water/acetonitrile/formic acid (49.9/50/0.1 % v/v/v) was infused at 10 µL/min and used for a lock mass which was measured each 30 seconds for 0.5 seconds. Spectral peaks were automatically corrected for deviations in the lock mass.

Extracted ion traces of labeled and unlabeled species were obtained for glucose, citric acid and malic acid in a 20 mDa window. These traces were subsequently smoothed and integrated with QuanLynx software (Waters, Manchester, UK). The signals for the labeled species of a given metabolite were corrected for natural signals by the following calculation

$$S^n_{adj} = S^n_{raw} - f^n * S^0$$

where S^n_{adj} is the adjusted signal of isotope n, S^n_{raw} the raw signal for isotope, f^n the correction factor and S^0 the signal of the first ($^{13}C_0$) isotope of the metabolite.

2.6 Animal experiments

All the animal experimentation was conducted in accordance with the Spanish Guide for Care and use of Laboratory animals, and with the International Care and Use Committee Standards.

All procedures were approved by the CIC bioGUNE's Animal Care and Use Committee and the competent authority (Diputación de Bizkaia). Mice were housed in a temperature-controlled animal facility (AAALAC-accredited) with 12-hour light/dark cycles and with water *ad libitum*. All experiments with animals were approved by CIC bioGUNE's Biosafety and Bioethics Committee and the Country Council of Bizkaia.

2.6.1 SAMe treatment

Eight-month-old *MAT1A*-KO male mice (in a C57Bl/6 background) with increases in liver enzyme levels (alanine aminotransferase [ALT] and aspartate aminotransferase [AST]) and hepatic lipid accumulation (determined by ultrasound) were given orally by gavage with vehicle (water) (n = 12) or SAMe (30 mg/kg/day, n = 12; Abbott, Chicago, IL) for 8 weeks before death. Age-matched WT male sibling littermates showing normal liver serum enzyme levels and ultrasound also were treated with vehicle for the same duration (n = 11). Animals were bred and housed in the CIC bioGUNE animal unit, accredited by the Association for Assessment and Accreditation of Laboratory Animal Care International. The mice were housed in groups using high-quality wood pellet hygienic litter bedding (Lignocel HBK 1500-3000; Rettenmaier & Sönne, Rosenberg, Germany) and in the presence of enrichment materials. Animals were fed with standard commercial chow animal diet (ref. 2914; Envigo, Barcelona, Spain). Submandibular and retroorbital blood samples were collected at the beginning and at the end of the experiment. Blood samples were deposited in serum-separator gel tubes (Microtainer; Becton-Dickinson, Franklin Park, NJ) and centrifuged (6000 rpm, 15 min, 4 °C) for serum separation. Livers were removed and snap frozen in liquid nitrogen, optimal cutting temperature cryo-compound embedded, or formalin-fixed. All procedures were performed during the light cycle and were approved by the Diputación de Bizkaia upon a favorable assessment by the Institutional Animal Care and Use Committee at CIC bioGUNE[143].

2.6.2 Methionine and Choline Deficient (MCD) Diet and Aramchol treatment

We got C57BL/6 male mice (Charles River, St Germain sur l'Arbresle, France) at 8 weeks of age and allowed them to acclimate for a period of 1 week. The mice were placed in groups of 10 on a diet lacking choline and with 0.1% methionine (0.1MCD diet) for a period of 4 weeks. A control group was maintained on a regular diet with 1,030 mg/kg choline and 0.3% methionine (Teklad Global 14% Protein Rodent Maintenance diet; Envigo RMS Spain, Sant Feliu de Codines, Spain). After 2 weeks of feeding on the 0.1MCD diet, mice were divided into groups of 10 and treated by intragastric gavage with a formulation of 5 mg/kg/day of Aramchol in vehicle (1.7% carboxymethyl cellulose; BUFA, IJsselstein, the Netherlands) and 0.2% sodium lauryl sulfate (Sigma Aldrich, Steinheim, Germany) or vehicle alone. Animals kept on the normal diet were also provided the vehicle preparation. The mice were analyzed for

transaminases in blood at 0, 2, and 4 weeks of exposure to the 0.1MCD diet. All mice were sacrificed at 4 weeks. Livers were then removed and snap frozen in liquid nitrogen, optical coherence tomography cryocompound embedded, or formalin fixed.

2.7 Immunostaining assays

2.7.1 Histologic staining

Sudan III: Ornithine carbamyl transferase–embedded frozen samples were sectioned, cleared with 60% isopropanol, and stained with Sudan III solution (0.5% in isopropanol Sudan III Panreac ref: 251731.1606) for 1 hour and finally cleared with 60% isopropanol. Sections were counterstained with Mayer hematoxylin (MHS32-1L; Sigma) and mounted in aqueous mounting medium for lipid quantification.

Hematoxylin & Eosin and Sirius Red: Paraffin-embedded liver samples were sectioned, dewaxed with a Xylene substitute (HS-202, Histoclear; National Diagnostics, Atlanta, GA), and hydrated. Sections were stained for 5 min with Harry's hematoxylin (HHS128-4L; Sigma) and for 15 min with aqueous eosin (HT110232-1L; Sigma) for H&E staining or with 0.01% Fast green FCF in saturated picric acid for 15 min and 0.04% Fast green For Coloring Food/0.1% Sirius red in saturated picric acid for 15 min for Sirius red staining. Samples were dehydrated and cleared with Histoclear. Finally, sections were mounted in DPX mounting media (06522, 500 mL; Sigma).

F4/80: In addition, liver sections were immunostained for detecting F4/80. F4/80 samples were unmasked with proteinase K during 15 min at room temperature. Endogenous peroxidase activity was blocked for 10 min with 3% hydrogen peroxide, then sections were blocked with 5% normal goat serum for 30 min and incubated with F4/80 (1:50, 1 hour at 37 °C, MCA497BB; Bio-Rad, Hercules, CA) followed by 30 min with anti-Rat Immpress reagent (MP-7404; Vector, Burlingame, CA). Colorimetric detection was completed with Vector Vip purple substrate (sk-4600; Vector). Slides were counterstained with Mayer Hematoxylin (MHS32-1L; Sigma), and finally samples were dehydrated, cleared, and mounted in DPX mounting media (06522-500 mL; Sigma). Smooth muscle actin samples were counterstained with 4′, 6-diamidino-2-phenylindole and mounted with Dako fluorescence mounting media (S3023; Dako, Carpinteria, CA).

For the analysis, 5 images per sample were taken with a 10× (H&E and Sirius Red), 20× (F4/80), or 40× (Sudan III) objective from an upright light microscope (Carl Zeiss AG, Oberkochen, Germany). Quantification of staining areas was performed using FRIDA software (http://bui3.win.ad.jhu.edu/frida/; Johns Hopkins University) and expressed as the percentage of stained area.

2.8 Metabolomic analysis

To determine metabolite measurements in serum, we used a LC-single-quadrupole-MS amino acid analysis system in combination with 2 separate LC-time of flight-MS-based platforms that analyzed methanol/chloroform or methanol extracts for lipid analysis. For liver samples, the previous LC/MS platforms were completed with a methanol/water extract analysis, covering polar metabolites. Metabolomics data were preprocessed using the TargetLynx application manager for MassLynx 4.1 (Waters Corp, Milford, MA).

2.8.1 Serum and liver samples

Targeted serum metabolic profiles were semiquantified as previously describe[144]. Ultra-high performance LC–single-quadrupole–MS amino acid analysis system was combined with 2 separate ultra-high performance LC–time of flight–MS – based platforms analyzing methanol and chloroform/ methanol serum extracts. Identified ion features in the methanol extract platform included FAs, oxidized FAs, acyl carnitines, N-acyl ethanolamines, bile acids, steroids, and lyso-phospholipids. The chloroform/methanol extract platform covered glycerolipids, cholesteryl esters, sphingolipids, diacyl-phospholipids, acyl-ether-phospholipids, and primary fatty acid amides. Lipid nomenclature and classification follows the LIPID MAPS convention (www.lipidmaps.org).

Liver metabolic profiles were analyzed as described[144,145]. The 3 ultra-high performance LC/MS platforms previously mentioned in serum analysis were completed with a methanol/water extract platform[145]. This platform covered polar metabolites, such as vitamins, nucleosides, nucleotides, carboxylic acids, coenzyme-A derivatives, carbohydrate precursors/derivatives, and redox-electron-carriers. Metabolomics data were preprocessed using the TargetLynx application manager for MassLynx 4.1 (Waters Corp, Milford, MA). Intrabatch normalization followed the procedure described[146].

2.8.2 Quantification of total lipids in liver

Lipids were extracted as described in[147] from homogenized livers of approximately 300mg. TGs were quantified using a kit (A. Menarini Diagnostics, Florence, Italy) and PE, PC, and DG were separated by thin-layer chromatography and quantified as described[148].

2.8.3 One central carbon metabolism measurement in liver samples

Extraction of the main metabolites belonging to the methionine pathway were performed as described[149]. To give a brief summary, livers (50 mg) were disrupted and metabolites were extracted in ice-cold methanol/water (50/50 % v/v) containing 10 mmol/L acetic acid in a tissue homogenizer (Precellys, Montigny-le-Bretonneux, France) using 2 x 20” cycles at 6000 rpm.

After centrifugation and evaporation, pellets were resuspended in water/MeCN/formic acid (40/60/0.1 %v/v/v) for injection on the ultra-high-performance LC/MS system. For the analysis of serum samples, 40 mL aliquots of serum were diluted with 40 mL water containing 0.2% formic acid (%v/v) and proteins were precipitated by addition of 120 mL of acetonitrile. Then, samples were sonicated for 10 min at 4 ºC and centrifuged at 14,000 rpm for 30 min at 4 ºC. The supernatant was injected directly onto the ultra-high-performance LC/MS system. Samples were analyzed with an ultra-high-performance LC system (Acquity; Waters, Manchester, UK) coupled to a time-of-flight–mass spectrometer (SYNAPT G2; Waters). A 2.1 x 100 mm, 1.7-mm BEH amide column (Waters), thermostated at 40 ºC, was used to separate the analytes before entering the MS. Solvent A (aqueous phase) consisted of 99.5% water, 0.5% formic acid, and 20 mmol/L ammonium formate, whereas solvent B (organic phase) consisted of 29.5% water, 70% MeCN, 0.5% formic acid, and 1 mmol/L ammonium formate.

To obtain a good separation of the analytes the following gradient was used: from 5% A to 50% A in 2.4 min in curved gradient (#8, as defined by Waters), from 50% A to 99.9% A in 0.2 min constant at 99.9% A for 1.2 min, back to 5% A in 0.2 min. The flow rate was 0.250 mL/min and the injection volume was 2 mL. All samples were injected randomly. After every 8 injections a high and a low-quality control sample was injected. If necessary, signals were corrected for signal drift during the run and all samples were injected in duplicate.

The MS was operated in positive electrospray ionization mode in full scan (50–1200 daltons) with a capillary voltage of 250 V, a sampling cone voltage of 20 V, and an extraction cone voltage of 5 V. Source and desolvation temperatures were 120 ºC and 450 ºC, respectively. Cone and desolvation gas flows were 5 and 600 L/h, respectively. The MS was tuned to a mass resolution of 20,000 full width at half maximum. Scan time was 0.2 seconds. A lock mass was used to correct for instrument fluctuations during the run. Therefore, leucine–enkephalin (2 mg/mL) was infused at 10 mL/min and its signal was measured every 40 seconds for 0.2 seconds. Extracted ion traces were obtained for methionine (m/z, 150.0589), SAMe (m/z, 399.1451), SAMe (m/z, 385.1294), MTA (m/z, 298.097), decarboxylated SAMe (m/z, 355.1552), spermine (m/z, 203.2236), spermidine (m/z, 146.1657), GSH (m/z, 308.0916), and choline (m/z, 104.1070) in a 20 millidalton window and subsequently smoothed (2 points, 2 iterations) and integrated with QuanLynx software (Waters). Peak areas were obtained from the extracted ion chromatograms and normalized to milligram of tissue or milliliter of serum.

2.9 Global DNA methylation profiles

Reduced representation bisulfite sequencing (RRBS) was used to analyze global DNA methylation. Sequencing libraries were prepared with a modification of the method described by Varela-Rey et al,[150] digesting DNA with TaqI and MspI and using the NEXTflex Bisulfite-Seq

Kit, Illumina Compatible (Bioo Scientific, Austin, TX), and Single-Read sequenced on a HiScanSQ platform (Illumina Inc, San Diego, CA) for 50 nucleotides. Resulting multiplexed FASTQ files for each sample were merged into a single FASTQ before the quality control and filtering steps. Quality control and adapter trimming were performed via FASTQC (http:// www.bioinformatics.babraham.ac.uk/projects/fastqc/) and Trim Galore (http://www.bioinformatics.babraham.ac.uk/ projects/trim_galore/), and only those with a minimum of 20 phred quality score were retained. Bisulfite reads were mapped to mm10 (Mus musculus) reference genome. Reads were mapped using Bismark, which is developed specifically to map bisulfate-treated reads[151]. In this analysis, 2 alignment mismatches (n = 2) were allowed. Resulting alignment files, in SAM format, were sorted before being used as input for the methylKit (R-package) to call DMRs, compute methylation ratios, and perform a DMR analysis[152].

2.10 Proteomics

Liver samples were prepared and analyzed by LC/MS/MS in the following way: frozen livers (n = 5) were ground while frozen in a liquid N_2 cooled cryohomogenizer (Retsch, Haan, Germany) and prepared for digestion. 750 µg of protein were denatured in a solution of 100 mmol/L TRIS-HCL, pH 8, and 8 mol/L urea. Samples then were ultrasonicated for 10 min in 10-second repeating on/off intervals of 10 seconds on and 10 seconds off (QSonica, Newtown, CT), and centrifuged at 16,000 x g for 10 min at 4ºC to remove insoluble pellets. The soluble fraction then was reduced with dithiothreitol (15 mM) for 1 hour at room temperature and alkylated with iodoacetamide (15 mmol/L) for 30 min at room temperature in the dark. Then, 75 mg protein was diluted to a final concentration of 2 mol/L urea with 100 mmol/L TRIS-HCL, pH 8, and digested overnight on a shaker at 37 ºC in 3 mg of trypsin/Lys-C mix (Promega, Madison, WI). Samples were de-salted and cleaned using HLB plates (Oasis HLB 30 mm, 5 mg sorbent; Waters).

LC/MS/MS was performed on a Dionex Ultimate 3000 NanoLC connected to an Orbitrap Elite (Thermo Fisher) equipped with an EasySpray ion source. The mobile phase A comprised 0.1% aqueous formic acid and mobile phase B comprised 0.1% formic acid in acetonitrile. Peptides were loaded onto the analytical column (PepMap RSLC C18 2 mm, 100 Å, 50 mm i.d. x 15 cm) at a flow rate of 300 nL/min using a linear AB gradient composed of 2%–25% A for 185 minutes, 25%–90% B for 5 minutes, then an isocratic hold at 90% for 5 minutes with re-equilibrating at 2% A for 10 minutes. The temperature was set to 40 ºC for both columns. Nano-source capillary temperature was set to 275 ºC and spray voltage was set to 2 kV. MS1 scans were acquired in the Orbitrap Elite at a resolution of 60,000 full width at half maximum with an automated gain control target of 1 x 106 ions over a maximum of 500 ms. MS2 spectra were

acquired for the top 15 ions from each MS1 scan in normal scan mode in the ion trap with a target setting of 1 x 104 ions, an accumulation time of 100 ms, and an isolation width of 2 daltons. Normalized collision energy was set to 35% and 1 microscan was acquired for each spectrum.

For preparative data analysis and peptide identification search, the raw MS files were converted to mzXML using MSConvert and searched against the Swiss-Prot–reviewed mouse FASTA database (33,330 proteins and decoys) using the COMET, X! Tandem native, and X! Tandem k-score search algorithms[153,154]. Target-decoy modeling of peptide spectral matches was performed with peptide prophet[155] and peptides with a probability score of > 95% from the entire experimental data set were imported into Skyline software[156] to establish a library for quantification of precursor extracted ion intensities. Precursor extracted ion intensities from each experimental file were extracted against the Skyline library, and peptide extracted ion intensities with isotope dot product scores greater than 0.8 and a minimum of 2 peptides per protein were filtered for final statistical analysis of proteomic differences[157]. Normalization of raw peptide intensities and protein level abundance inference were calculated using the linear mixed-effects model built into the open sources MSSTATs (v3.2.2) software suite[158].

2.11 Human samples

All the studies were performed in agreement with the Declaration of Helsinki and with local national laws. The Human Ethics Committee of each hospital approved the study procedures and written informed consent was obtained from all patients before inclusion in the study.

2.11.1 Biopsy-proven NAFLD patients

Our study included a total of 535 patients who underwent liver biopsy analysis (353 diagnosed of simple steatosis and 182 with a diagnosis of NASH), seen at 11 participating hospitals. Between them, 377 patients were described previously by Barr *et al*[144] and 158 additional patients were recruited since 2013 for this study by 3 hospitals that also participated in the first study (Hospital Universitario Marqués de Valdecilla, Hospital Universitario Príncipe de Asturias, and Hospital Virgen de Valme). PCA of the metabolomics data showed that patients cluster together independently of the hospital of origin as shown in **Figure 11**. All patients were recruited using the following inclusion criteria: (1) age 18-75 years; (2) not known acute or chronic disease except for obesity or T2D based on medical history, physical examination, and standard laboratory tests; and (3) alcohol consumption was less than 20 g/day for women and 30 g/day for men. Exclusion criteria included viral, autoimmune, hemochromatosis, and drug-induced causes of liver disease. All of the subjects were of Caucasian origin. For all subjects, blood was drawn under fasting conditions on the morning the diagnostic liver biopsy was

performed. Serum was separated and stored at -80 ºC until analysis. Clinical data were collected retrospectively using patient records and laboratory values obtained at the time of biopsy and it is summarized in **Table 3**. Diagnoses were established histologically in liver biopsy specimens. The histologic diagnosis of NAFLD was established by a single liver pathologist in each participating hospital using the scoring system defined by Kleiner et al[159]. After assessment, patients were classified by the pathologists into 2 histologic groups: (1) simple steatosis (hepatic steatosis alone), and (2) NASH (presence as determined by the pathologist). There are no patients with cirrhosis.

		Total		M subtype		non-M subtype		Indeterminate	
		Simple steatosis	NASH	Simple steatosis	NASH	Simple steatosis	NASH	Simple steatosis	NASH
N (%)	535	353 (66)	182 (34)	174 (66)a	88 (34)b,c	103 (61)d	68 (39)	76 (75)	26 (25)
Female	69%	70%	68%	72%a	77%b,c	62% d	57%	77%	65%
Age	44.77 ± 11.35	44.04 ± 11.54	46.07 ± 10.94	40.31 ± 10.77a,e	43.64 ± 9.68e	48.35 ± 11.29	49.30 ± 11.44	46.77 ± 10.89	45.36 ± 11.81
BMI. Kg/m²	44.76 ± 10.78	44.50 ± 10.65	45.21 ± 11.02	48.19 ± 9.71a,e	49.72 ± 8.67b,c	40.54 ± 10.34	40.66 ± 11.18f	41.12 ± 10.30	42.01 ± 12.02
Fasting glucose level, mg/dL	113.59 ± 41.26	111.39 ± 35.86	117.31 ± 48.96	112.21 ± 35.97	119.21 ± 51.93c	106.32 ± 27.47	119.89 ± 50.82f	117.10 ± 45.01	101.05 ± 21.03
Total fasting cholesterol level, mg/dL	188.71 ± 41.08	187.84 ± 41.88	190.66 ± 39.45	176.88 ± 35.29a	175.01 ± 27.60c	196.59 ± 47.31	192.11 ± 43.89	191.69 ± 39.20	213.67 ± 23.56
Fasting HDL cholesterol level, mg/dL	47.02 ± 13.70	48.90 ± 14.26	41.83 ± 10.51	45.67 ± 10.96	42.85 ± 11.80	50.06 ± 14.77	40.66 ± 10.07	52.50 ± 17.37	48.50 ± 9.47
Fasting LDL cholesterol level, mg/dL	109.81 ± 33.14	109.14 ± 33.39	111.35 ± 32.87	109.17 ± 32.89	104.17 ± 23.98	107.78 ± 34.34	113.13 ± 37.22	112.29 ± 34.36	119.50 ± 19.33
Fasting triglyceride level, mg/dL	150.33 ± 100.19	150.57 ± 106.95	149.78 ± 83.33	124.88 ± 47.39a	134.38 ± 96.15	164.48 ± 92.57	156.71 ± 79.30	170.99 ± 177.11	145.62 ± 83.08
ALT level, U/L	39.70 ± 33.09	35.38 ± 28.39	44.62 ± 37.23	29.93 ± 20.54a	34.01 ± 19.45b	39.11 ± 32.68	55.56 ± 45.07	38.22 ± 30.97	52.76 ± 51.28

Table 3. Clinicopathologic characteristics of the NAFLD patients included in the study. All diagnoses were established histologically in liver biopsy specimens. Additional classification as M subtype, non-M subtype, or indeterminate was based on results detailed in the "Subclassification of NAFLD patients" section. Values are given as percentages or means ± SD of the mean. Results that were significantly different ($p < 0.05$) among the M subtype, non-M subtype, or indeterminate groups are indicated. BMI = body mass index; HDL = high-density lipoprotein; LDL = low-density lipoprotein. a = significant at $p < 0.05$, between M subtype and non-M subtype patients with simple steatosis. b = significant at $p < 0.05$, between M subtype and non-M subtype patients with NASH. c = significant at $p < 0.05$, between M subtype and indeterminate patients with NASH. d = significant at $p < 0.05$, between non-M subtype and indeterminate patients with simple steatosis. e = significant at $p < 0.05$, between M subtype and indeterminate patients with simple steatosis. f = significant at $p < 0.05$, between non-M subtype and indeterminate patients with NASH.

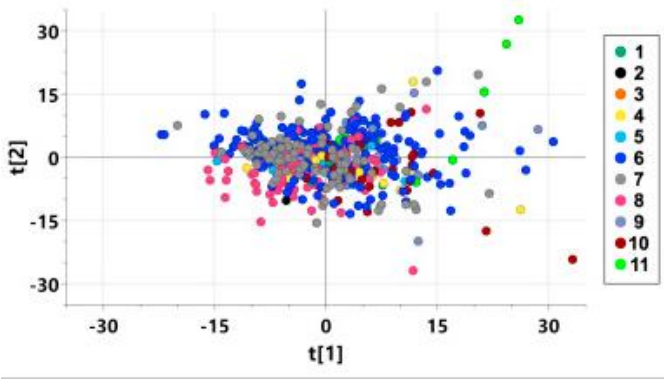


Figure 11. PCA of the 535 samples from the 11 participating hospitals showing that all patients cluster together independently of the hospital they belong. Every hospital is represented by a color and each dot represents a sample. (1) Hospital Universitario "12 de Octubre", Madrid, Spain; (2) Hospital Clínic, Barcelona, Spain; (3) Hospital del Tajo, Aranjuez, Spain; (4) Hospital General Universitario Gregorio Marañón, Madrid, Spain; (5) Pitié-Salpêtrière Hospital, Paris, France; (6) Hospital Universitario Marqués de Valdecilla, Santander, Spain; (7) Hospital Universitario Príncipe de Asturias, Madrid, Spain; (8) Hospital Universitario Reina Sofía, Córdoba, Spain; (9) Hospital Universitario Santa Cristina, Madrid, Spain; (10) Hospital Nuestra Señora de Valme and Hospital Universitario Virgen Macarena y Virgen del Rocío, Sevilla, Spain; and (11) Hospital Clínico Virgen de la Victoria, Málaga, Spain.

2.11.2 MetS and general population urine samples

Urine samples from general population of Valencia and Euskadi were obtained from University of Valencia and the company Osarten respectively for this project. Moreover, metadata from each patient including lifestyle habits, biochemical and anthropometric parameters were sent to CIC bioGUNE. These individuals belong to two projects: OBENUTIC and PreMedEus that will be explained in next section.

The metadata from people participating in OBENUTIC and PreMedEus projects were correlated with the ^{1}H spectra obtained from measuring these urine samples in a 600MHz spectrometer (600MHz AVANCE IIIHD IVDr spectrometer of BRUKER) to look for the parameters that best could define MetS. These two cohorts of patients are described further below.

2.11.2.1 OBENUTIC cohort

Obesidad, Nutrición y Tecnologías de la Información y Comunicación (**OBENUTIC**) is a study carried out in the Preventive and Public Health department of the Faculty of medicine of Valencia. OBENUTIC is a study of obesity (with controls and cases) originally designed to improve and automatize data collection in epidemiologic nutritional studies. The total sample size of this study was constituted for 1668 individuals (general population) between 18 and 80 years old from the Community of Valencia (Spain) and female and male were included. 633 participants were recruited between 2012 and 2017 by Rocío Barragan, a PhD student from this department of the University of Valencia where urine samples were collected at the Department of Preventive Medicine and Public Health. The period of inclusion of patients was from February to October 2011 following the following inclusion and exclusion criteria:

Inclusion criteria:

- Any gender
- Age between 18 - 60 years old
- Residents of the Community of Valencia
- Body mass index (BMI) between 20-35 kg/m²
- Agreed to participate in the study

Exclusion criteria:

- To suffer from any infectious-contagious disease, physical or mental incapacitating illness.
- Pregnancy or lactation.
- Diagnosis of cancer, thyroid disorders and/or type I diabetes
- Cushing's disease
- Use of medication that alter the concentration of blood chemistry or lipid profile.

The variables analyzed in OBENUTIC included sociodemographic data (age, gender, level of studies, birth place), family and personal medical history (anthropometric measures, tobacco and alcohol consumption, physical activity, food consumption, mediterranean diet adherence, etc) and biochemical (fasting glycemia, TGs, creatinine, cHDL, etc).

CIC bioGUNE received 511 urine samples from the project OBENUTIC (together with correlated biochemical data for each patient). Urine samples were measured by NMR (600MHz AVANCE IIIHD IVDr spectrometer of BRUKER) to study MetS by obtaining and analyzing the bins and metabolites present in the urine of this subgroup of OBENUTIC cohort. ^{1}H monodimensional spectrum was obtained for each sample and the information obtained by the NMR spectrometer was correlated with the biochemical and anthropometric data. Metadata of these patients was provided by Dr. Dolores Corella (Universidad de Valencia) in a database made in SPSS including personal history (pathological and no pathological) biochemical and behavioural data of this general population.

Metadata included:

- Anthropometric data (BMI, high, weight)
- Biochemical data (fasting blood glucose, TGs blood levels, cHDL)
- Blood pressure
- Medicine consumption
- Lifestyle behavior (regular exercise, smoking)

To sum up, urine samples previously collected and transported to the NMR-metabolomics platform of CIC bioGUNE were processed and analyzed as described in **section 2.11.2.3** and

analyzed by NMR in the 600MHz AVANCE IIIHD (IVDr) BRUKER spectrometer used only for metabolic screening.

2.11.2.2 PreMedEus cohort

Osarten is a company which aggregates other companies and offers them a service of prevention looking for the security and healthcare of the workers belonging to the associated companies. Osarten and CIC bioGUNE collaborate in a project called "**Metabotipo de la población laboral activa de la CAPV a través del análisis metabolómico de suero y orina por RMN"** which, has as main objective, improve the diagnosis of diseases like MetS and check its progression along the time. The project includes 10.000 urine samples and other 10.000 serum samples for the study of metabolomic features in general population from Basque Country. People participating in this study are between 18 and 70 years old, with equal proportion of males and females. There is only one exclusion criteria: people participating should not had suffered a serious illness like cancer or ictus in the 3 months preceding the sample collection.

From this huge cohort, 75 urine samples were selected (according to the criteria we have been working with) to feed OBENUTIC cohort, as there was not enough representation of MetS condition in the 511 OBENUTIC urine samples. In this work, these 75 samples were renamed as "Precision Medicine Euskadi" (**PreMEdEus**), as it is coming from the project aforementioned, tightly related with precision medicine research in Basque Country.

As we show in the results section (**Figure 42**), PCA were carried out to confirm that samples coming from the two selected cohorts, OBENUTIC and PreMedEus, were a heterogeneous group with no substantial differences that could difficult the analysis of this samples when considering them as a unique group.

2.11.2.3 Urine sample preparation

Urine samples were prepared as follows: urine samples (which have been stored at -80 ºC) were carefully thawed at room temperature (30 min for ≤ 1mL volume sample). Samples were centrifuged for 5 min at room temperature and 5000 rcf and 540 µL of urine were taken into Eppendorf tubes of 1.5 mL volume). 60 µL of BRUKER's urine buffer (pH 7.4) containing 1.5M KH_2PO_4 phosphate buffer and 2 mM NaN_3 in D_2O + 0.1% TSP was added, and samples were shaken strongly for mixing. All the volume was introduced into the NMR tubes (5 mm). These samples were analyzed in a 600 MHz AVANCE IIIHD (IVDr) provided with a BBI probe-head, automatic sample charger (Sample-jet) and 3 channels: ^{1}H, ^{2}H and X. ^{1}H mono-dimensional spectra and J-resolved 2D were obtained for sample and binning at 0.03 ppm were generated and normalized by the total area. Taking together bins and metadata, treatment of data and statistical analysis was carried out as we will see in **section 2.11.3**.

2.11.2.4 WHO metabolic syndrome definition

We work with the definition of MetS provided by the WHO: diabetes or impaired glucose tolerance; fasting blood glucose > 110 mg/dL plus 2 or more of the following factors:

- Obesity: BMI > 30 kg / m²
- Dyslipidemia: TGs >>> 150 mg/dL or cHDL < 34.79 mg/dL in men and < 38.66 mg/dL in women.
- Hypertension ≥ 140/90 mmHg or drug treatment
- Albuminuria: albumin excretion 20 µg/min

2.11.3 Treatment of data and statistical analysis

From OBENUTIC project 500 samples were analyzed at the end as some samples were excluded due to a wrong water suppression in NMR spectra (TSP > 1.3), or because metadata from Valencia was missing for some patients. 75 MetS samples more from PreMedEus cohort were included to balance the samples size of each class.

2.11.3.1 Outlier elimination

Besides, outliers were detected using the multivariant detection algorithm DBSCAN (https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.121.9220): bins are scaled by Pareto method and samples which (according to the algorithm) have no accessible other 4 samples at a minimal distance of 10 are selected as outliers.

2.11.3.2 Metadata unification

Regarding to metadata, it was unified renaming the variables for both datasets to match. Units and format were also matched.

2.11.3.3 Adopted criteria for metabolic syndrome

As said before, the criteria adopted for selecting samples as MetS was carried out following the WHO criteria (**section 2.11.2.4**).

In our case a little modification of this definition was made because the data for “albuminuria” was not available in the data sheet that we received from Valencia. On the other hand, it is considered that there is a problem of diabetes/glucose when fasting glucose ≥ 110 mg/dL or there is evidence of diabetes medication consumption. So, a sample is considered “MetS” when is “diabetic” and have at least other 2 factors (obesity, dyslipidemia and hypertension). If the sample does not meet the diabetic criteria and either have the other factors is it classified as “asymptomatic”.

2.11.3.4 Untargeted analysis data (bins)

Each spectrum was read from 0.5 ppm to 9.5 ppm. Water region comprising between 4.7-5.0 ppm was eliminated. The result was divided in 290 bins of 0.03 ppm. The result was normalized by the total intensity of the spectrum after water region elimination.

2.11.3.5 Targeted analysis data (metabolites)

Identification and quantification of metabolites were carried out automatically using the software B.I.QUANT-UR-b de BRUKER. 50 different metabolites could be quantified. In the cases when metabolites could not be identified, because they were below the limit detection, they adopted a value of 0. These cases were modified to have half the detection limit as value. The results were presented normalized by creatinine in mmol/mmol creatinine (with the exception of creatinine which was expressed in its absolute unit in mmol).

2.11.3.6 Final metadata

The final metadata is summarized in **Table 4**. This analysis was generated with the R library *compareGroups*, version 1.4.0.

	Female n = 319	**Male n = 219**
age	46.87 ± 12.77	47.72 ± 13.71
MetS	30 (9.40%)	72 (32.88%)
Other	124 (38.87%)	97 (44.29%)
Asymptomatic	165 (51.72%)	50 (22.83%)
whoms_diabetes	43 (13.48%)	83 (37.90%)
whoms_obesity	99 (31.03%)	107 (48.86%)
whoms_dyslipidemia	57 (17.87%)	84 (38.36%)
whoms_hypertension	87 (27.27%)	122 (55.71%)
BMI	27.57 ± 5.62	29.86 ± 4.84
glucose_mg_dL	94.44 ± 18.09	108.70 ± 30.40
triglycerides_mg_dL	106.21 ± 55.56	146.54 ± 87.86
cholesterol_HDL_mg_dL	63.52 ± 13.76	49.54 ± 11.49
syst_bp_mmHg	120.84 ± 16.61	132.64 ± 15.85
diast_bp_mmHg	76. 96 ± 9.70	82.02 ± 11.84
diabetes	13 (4.53%)	8 (5.48%)
med_for_diabetes	8 (2.62%)	29 (14.95%)
hypertension	86 (27.48%)	122 (55.96%)
med_for_hypertension	37 (12.05%)	43 (22.05%)

Table 4. Metadata for final MetS cohort.

CHAPTER 3

RESULTS

3 RESULTS

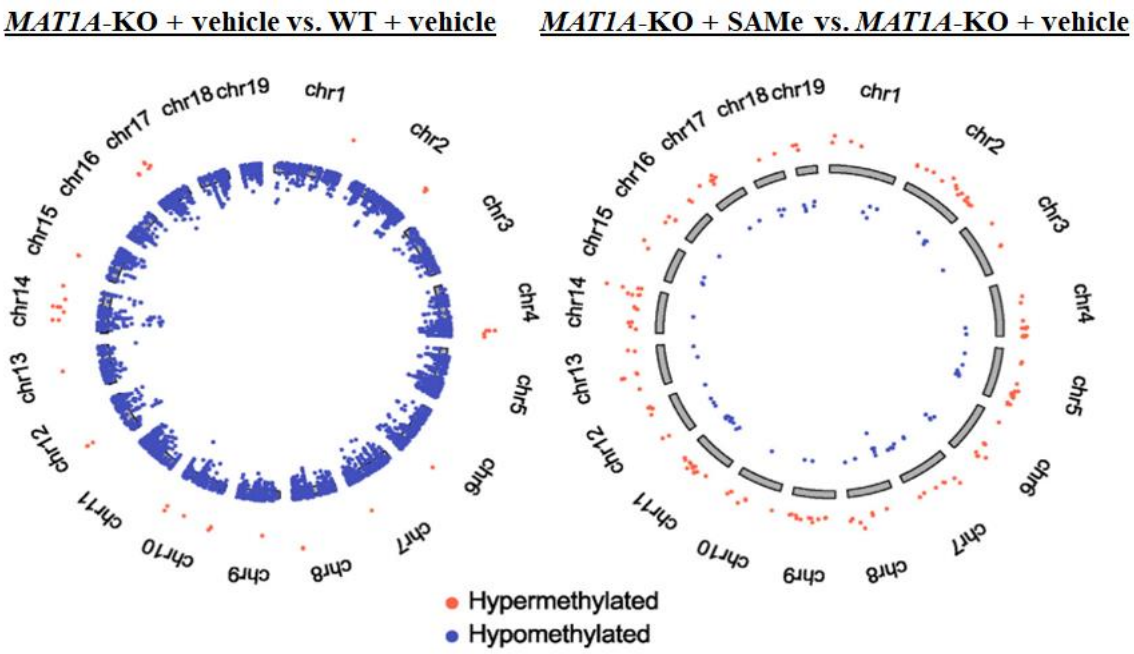
The main aim of this work has been the search for metabolomic biomarkers of NASH and MetS. As NASH is considered by many the hepatic manifestation of MetS, its study may be also useful to understand NASH development and progression.

3.1 Overall effect of SAMe administration in *MAT1A*-KO mice

3.1.1 DNA methylation is altered as a function of SAMe concentration

As we explained above in **section 1.2**, SAMe is the main biological methyl donor of the cell and alterations in its homeostatic levels lead to hepatic disorders. Yet, SAMe half-life is approximately 5 minutes[160] and orally administered SAMe is barely bioavailable due to the rapid hepatic metabolism[161]. Previous work in the laboratory showed that liver SAMe content increased rapidly after SAMe intraperitoneal injection, reaching a peak at 15 minutes and decaying to basal levels 4 hours after injection[162]. Consistently, in our study we found that *MAT1A*-KO mice receiving SAMe or vehicle (water) during two months before sacrifice had no significant differences in the concentration of SAMe in serum and liver, when sacrificed 24 hours after the last administration of SAMe.

As reported in previous studies[76], in *MAT1A*-KO mice, deletion of *MAT1A* lead to a reduction of hepatic SAMe and, as a consequence of that, lower levels of DNA methylation were found. DNA hypomethylation in the liver is related with human NASH and NAFLD progression[163]. Thus, we analyzed differentially methylated DNA regions (DMRs) in *MAT1A*-KO mice given SAMe compared to *MAT1A*-KO mice given vehicle (**Figure 12**). SAMe administration increases DNA methylation: regions that are hypomethylated in *MAT1A*-KO mice partially restore methylation levels when given SAMe while additional regions appear to be also hypermethylated in SAMe-treated mice.



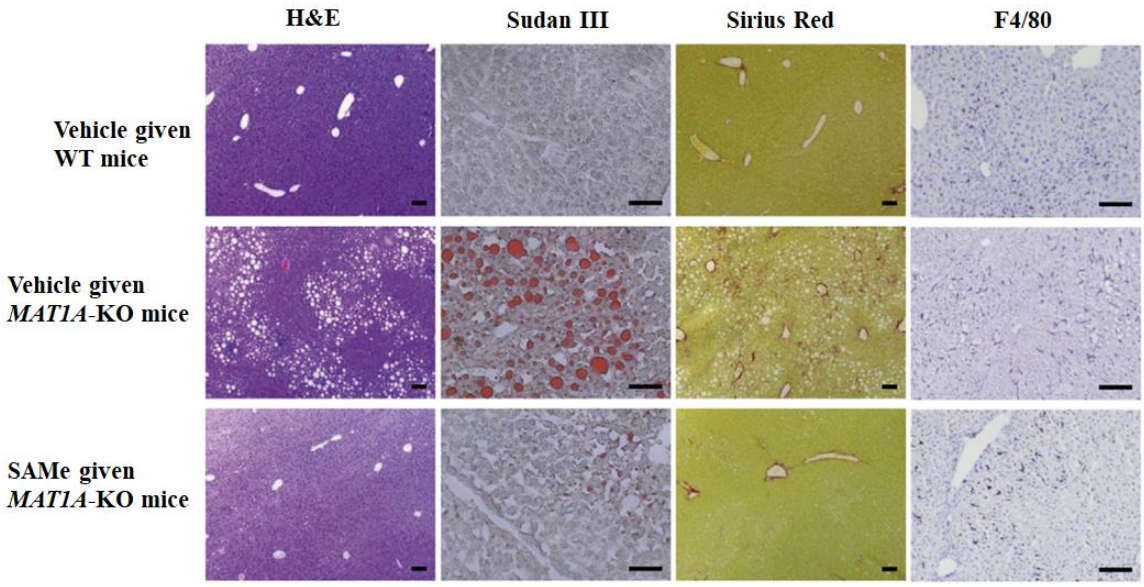
DMR*	*MAT1A*-KO + vehicle vs. WT + vehicle	*MAT1A*-KO + SAMe vs. WT + vehicle	*MAT1A*-KO + SAMe vs. MAT1A-KO + vehicle
Hypermethylated	43	53	161
Hypomethylated	6678	5330	69
Total	**6721**	**5383**	**230**

Figure 12. Effect of SAMe treatment on global DNA methylation in *MAT1A*-KO mice. Circos plots of the DMRs between *MAT1A*-KO + vehicle and WT + vehicle (left) and *MAT1A*-KO + SAMe and *MAT1A*-KO + vehicle (right). Red spots and blue spots represent hypermethylated and hypomethylated regions, respectively. The table shows the number of hypermethylated, hypomethylated and total DMRs in each comparative performed (*MAT1A*-KO + vehicle vs WT + vehicle; *MAT1A*-KO + SAMe vs WT + vehicle, and *MAT1A*-KO + SAMe vs *MAT1A*-KO + vehicle). *Filtering parameters: Q-value < 0.05, methylation difference > 20%.

3.1.2 SAMe administration improves liver histology and liver function in *MAT1A*-KO mice

SAMe homeostasis is central to a normal liver function, as illustrated by *MAT1A*-KO mice, which show increased susceptibility to develop fatty liver, steatohepatitis and HCC[164,165]. Previous work have shown that methionine derived metabolites have hepatoprotective properties in rodents[166] so here we checked if *MAT1A*-KO mice given SAMe showed an improvement in liver function and histology. At the end of eight weeks of administration of SAMe and vehicle to *MAT1A*-KO mice, animals were sacrificed and liver histology performed to determine if SAMe produced histological differences between these two groups of animals.

Our results indicated that *MAT1A*-KO mice receiving SAMe showed a marked reduction (11-fold) in liver lipid accumulation as compared to *MAT1A*-KO mice that received vehicle (p = 2e-04) as shown by Sudan III red-stained area quantification (**Figure 13**).



	Sudan III (a.u ± SD)	Sirius Red (a.u ± SD)	F4/80 (a.u ± SD)
WT + vehicle	1.1 ± 1*	1.4 ± 0.8*	0.7 ± 0.9
***MAT1A*-KO + vehicle**	21.5 ± 14.7	3.4 ± 2.8	1.3 ± 0.9
***MAT1A*-KO + SAMe**	2.0 ± 2.9*	1.3 ± 0.7*	0.4 ± 0.3*

Figure 13. Effect of SAMe administration on histology of *MAT1A*-KO mice liver. Here we show images of H&E, Sudan III red, Sirius red and F4/80 immunofluorescence staining of liver tissues after eight weeks of SAMe (30 mg/kg/day) or vehicle administration. Sizing bars correspond to 100 mm for H&E and Sirius Red, and 50 mm for Sudan III and F4/80. Quantitative analyses are shown in the table. Results that were significantly different (* $p < 0.05$) from vehicle-given *MAT1A*-KO mice are indicated. Data shown represent the mean of 12 vehicle-given *MAT1A*-KO, 12 SAMe-given *MAT1A*-KO, and 11 vehicle-given WT animals.

Mice given SAMe also showed a reduction in liver fibrosis and inflammation as showed by the quantification of Sirius red-stained areas (3-fold compared with *MAT1A*-KO given vehicle; p = 2.6e-02) and F4/80 (3-fold compared with *MAT1A*-KO given vehicle; p = 3e-03) respectively (**Figure 13**).

To determine if liver function was improved by SAMe administration, we measured serum transaminases levels before and after treatment. SAMe treatment in *MAT1A*-KO mice reduces the liver enzyme activity of alanine aminotransferase (ALT) and aspartate aminotransferase

(AST) as compared with those receiving only the vehicle (**Table 5**). Cholesterol and TGs levels also decreased in SAMe-treated mice (**Table 5**).

	ALT (UL ± SD)	AST (UL ± SD)	Cholesterol (mg/dL ± SD)	TGs (mg/dL ± SD)
WT + vehicle (before)	31 ± 5	88 ± 21	147 ± 17	136 ± 27
WT + vehicle (after)	20 ± 11*	43 ± 16*	110 ± 25*	105 ± 25
***MAT1A*-KO + vehicle (before)**	103 ± 49	72 ± 20	100 ± 22	150 ± 41
***MAT1A*-KO + vehicle (after)**	137 ± 105	107 ± 46*	107 ± 27	121 ± 38*
***MAT1A*-KO + SAMe (before)**	220 ± 255	128 ± 94	106 ± 29	170 ± 55
***MAT1A*-KO + SAMe (after)**	41 ± 23*	64 ± 35*	94 ± 37	114 ± 42*

Table 5. Effect of SAMe administration on serum parameters: ALT, AST, cholesterol and TGs. For each group of animals (WT + vehicle, *MAT1A*-KO + vehicle and *MAT1A*-KO + SAMe), results that were significantly different before and after administration are indicated (* $p < 0.05$). 12 SAMe-treated *MAT1A*-KO, 12 vehicle-given *MAT1A*-KO and 11 vehicle-given WT mice were analyzed.

3.2 SAMe depletion alters 1-carbon metabolism

In **section 1.2** we introduced the relationship between methionine and SAMe metabolism with 1-carbon metabolism (1CM). 1CM circulates one carbon unit from different nutrients like choline or folate and amino acids such as glycine, methionine, threonine and serine to generate a large variety of outputs including, DNA methylation, synthesis of polyamines, NADPH, GSH and nucleotides and methylation of PE rich in polyunsaturated FA (PUFA) to generate PC rich in PUFA (**Figure 14**)[167].

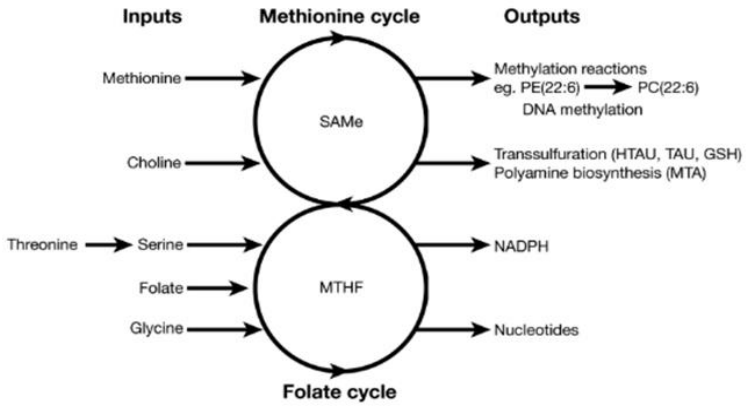


Figure 14. SAMe depletion alters 1CM**.** Schematic representation of 1CM which circulates one carbon unit from different inputs (methionine, choline, serine, threonine, and glycine), via SAMe and MTHF, into a large variety of outputs, such as DNA and phospholipid methylation, GSH, polyamines, NADPH, and nucleotide synthesis.

MAT1A deletion leads to the hepatic accumulation of methionine upstream metabolites, like serine (Ser) and threonine (Thr), PE molecules containing docosahexaenoic acid [(PE (22:6)], and MTHF, also showing abnormal content of enzymes related to 1CM such as AHCY, aldehyde dehydrogenase 1a1 (ALDH1A1), BHMT, CBS, cysteine sulfinic acid decarboxylase (CSAD), cystathionine g-lyase (CTH), dimethylglycine dehydrogenase (DMGDH), MAT2A and serine dehydratase (SDS) (**Figure 15 A, B**). Besides, *MAT1A* deletion causes a reduction in downstream metabolites such as MTA, a biomarker of polyamine synthesis, hypotaurine (HTAU), TAU, and GSH, key biomarkers of the transsulfuration pathway; PC containing docosahexaenoic acid [PC(22:6)] and NADPH levels (**Figure 15 A, B**).

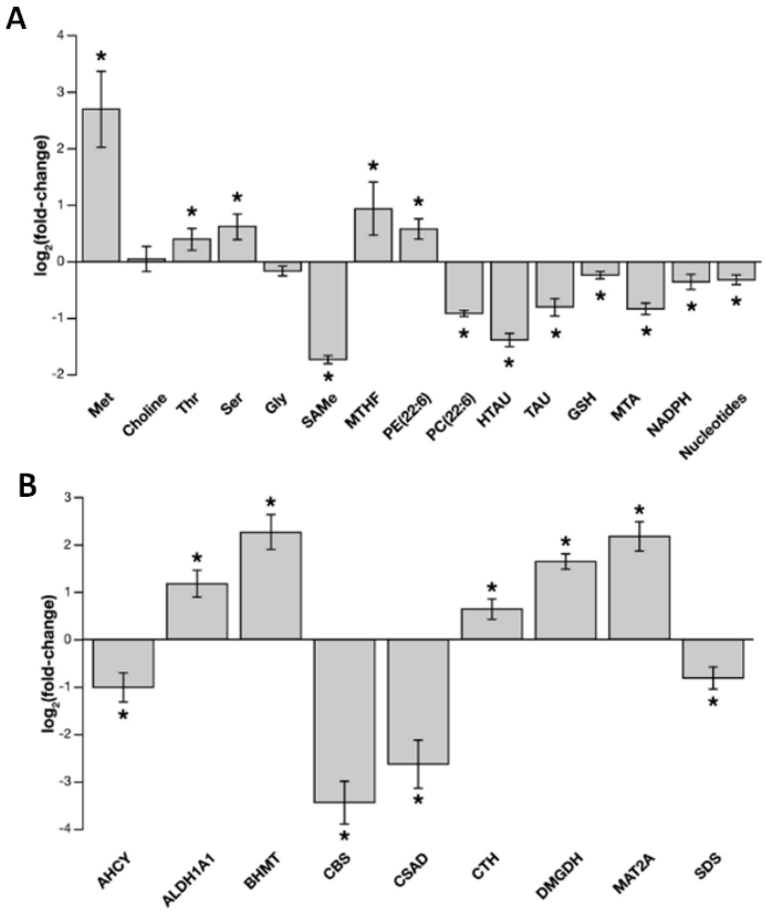


Figure 15. (A) Relative fold-change (log_2) in the hepatic content of the main metabolites involved in 1CM in *MAT1A*-KO as compared with WT mice. *MAT1A* deletion induced a reduction in hepatic SAMe content and downstream metabolites, such as PC(22:6), MTA, GSH, HTAU and TAU, NADPH, and nucleotides. *MAT1A* ablation also resulted in the accumulation of methionine and upstream metabolites, such as Ser, Thr, MTHF, and PE (22:6). **(B)** Relative fold-change (log_2) in the protein content of enzymes involved in hepatic 1CM in *MAT1A*-KO as compared with WT mice. *MAT1A* deletion led to abnormal protein content of numerous enzymes involved in 1CM. AHCY, ALDH1A1, BHMT, CBS, CSAD, CTH, DMGDH, MAT2A, SDS, * $p < 0.05$.

3.3 SAMe depletion activates processes as FA uptake, FA desaturation and esterification, impairing FA oxidation and VLDL secretion

Our results showed that SAMe depletion is associated with hepatic accumulation of FAs, DG and TGs (**Figure 16 and 17A**) so we examined for proteins involved in lipid metabolism that could be differentially expressed in *MAT1A*-KO and WT mice. Through proteomic data analysis, we found that some of these proteins are differentially expressed in *MAT1A*-KO. For example, the content of SCD1 (the rate limiting enzyme in the synthesis of MUFAs) significantly augments in *MAT1A*-KO mice. This is a relevant result since MUFAs make the majority of TGs and membrane phospholipids. Other enzymes involved in TGs biosynthesis such as 1-acylglycerol-3-phosphate O-acyltransferase 2 (AGPAT2) and diacylglycerol acyltransferase 2 (DGAT2) also augment in the liver of *MAT1A*-KO mice (**Figure 17B**). Similarly, the content in the FA transporter CD36 which overexpression correlates with TGs accumulation in human NAFLD[168], again significantly increase (**Figure 17B**).

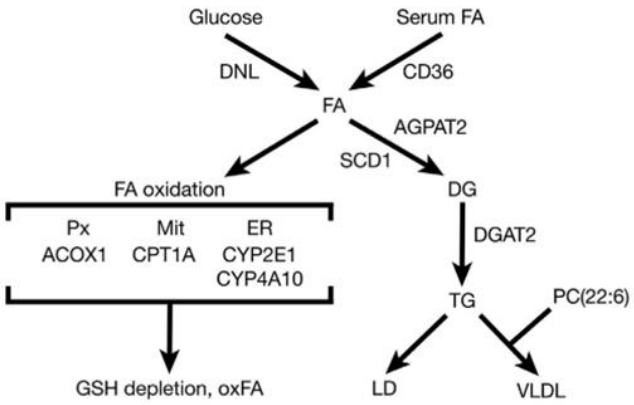


Figure 16. Schematic representation of hepatic lipid metabolism**.** Hepatic FAs originate from serum and through de novo lipogenesis (DNL). FA can be either oxidized in the mitochondria (Mit) or esterified to form TGs, which can follow different pathways. Px: peroxisomes; ER: endoplasmic reticulum; Mit: mitochondria; ACOX: acyl-CoA oxidase 1; cytochrome (CYP); oxFA: oxidized fatty acids, carnitine palmitoyltransferase 1a (CPT1A); palmitoylcarnitine (AC16:0).

As already mentioned **in section 1.1.2**, TGs are the mechanism by which the liver stores FAs and can have different fates: they can be stored in LD, used to form other lipids (ceramides, phospholipids and cholesteryl esters) or be exported into blood as VLDLs. PEMT has been demonstrated to be a key enzyme necessary for proper VLDL assembly and export as its deficiency rises the probability of develop fatty liver in mice[169]. Other studies showed that *MAT1A*-KO mice have impaired VLDL assembly and reduced VLDL secretion[170]. Here we found that SAMe depletion causes a reduction of PC (22:6)/PC and PC (20:4)/ PE (20:4) ratios, which are indicators of reduced PEMT activity (**Figure 17A**).

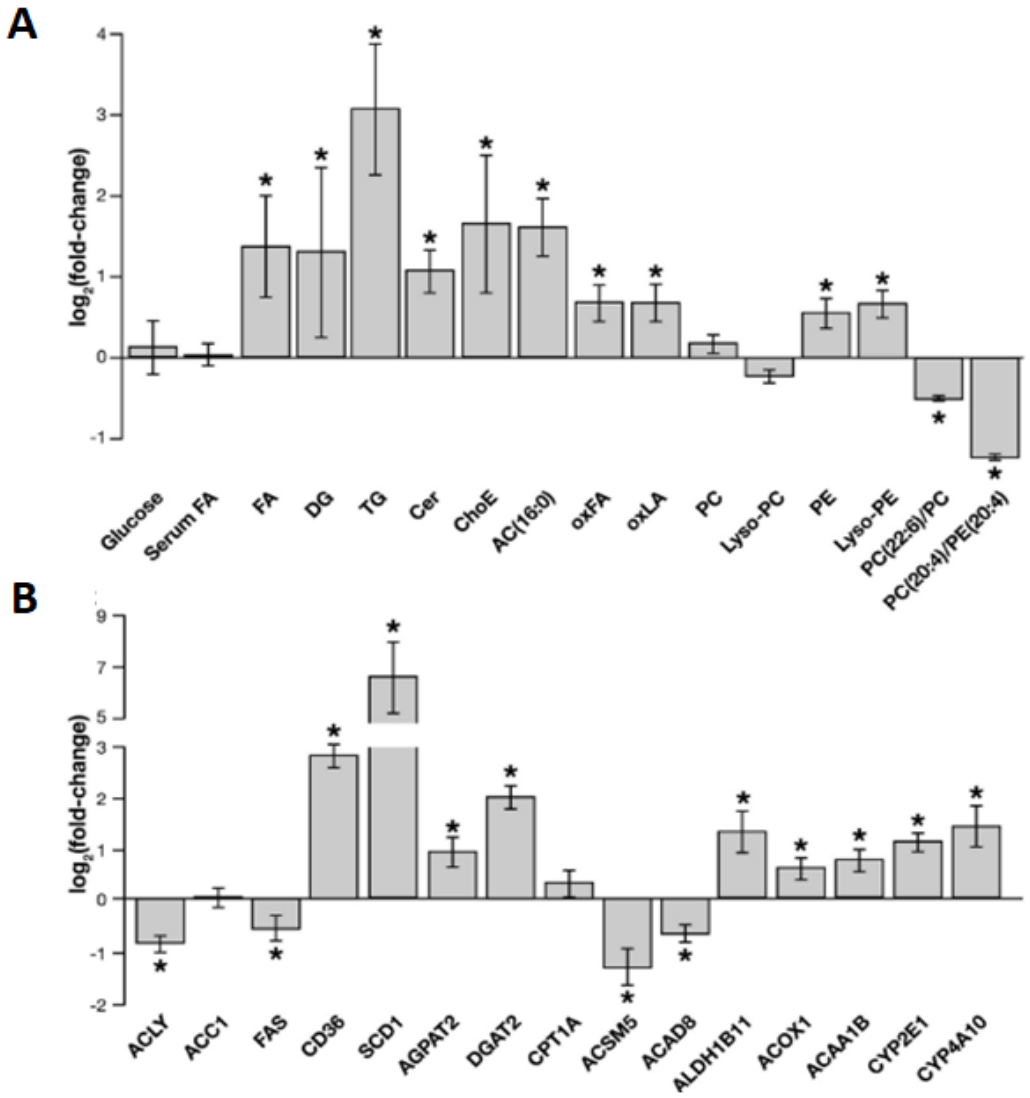


Figure 17. SAMe depletion activates FA uptake and esterification, whereas FA oxidation and VLDL secretion are impaired. **(A)** Relative fold-change (log2) in the hepatic content of the main metabolites involved in lipid metabolism in *MAT1A*-KO as compared with WT mice. AC (16:0): palmitoylcarnitine; oxFA: oxidized FA; oxLA: linoleic acid (18:2) derived oxidized FA; lyso-PC: lyso-phosphatidylcholine; lyso-PE: lyso-phosphatidylethanolamine; PC(22:6)/PC: ratio PC with docosahexaenoic acid/total PC; PC(20:4)/PE(20:4) and ratio PC/PE with arachidonic acid. **(B)** Relative fold-change (log_2) in the content of proteins involved in liver lipid metabolism in *MAT1A*-KO as compared with WT mice. DNL enzymes: ACLY: citrate lyase; ACC1: acetyl-CoA carboxylase 1; FAS: fatty acid synthase. FA transport: CD36, fatty acid translocase. FA esterification: SCD1, stearoyl-CoA desaturase 1; AGPAT2: 1-acylglycerol-3-phosphate O-acyltransferase 2; and DGAT2: diacylglycerol acyltransferase 2. Mitochondrial FA b-oxidation: CPT1A, carnitine palmitoyltransferase 1a; ACSM5: acyl-CoA synthetase medium chain family member 5; ACAD8: acyl-CoA dehydrogenase family member 8; ALDH1B1: aldehyde dehydrogenase 1 family member B1. Peroxisomal FA β-oxidation: ACOX1, acyl-CoA oxidase 1; ACAA1B: acyl-CoA acetyltransferase. Endoplasmic reticulum FA ω-oxidation: CYP2E1 and CYP4A10. * $p < 0.05$.

Our results also show that while having a normal content of carnitine palmitoyltransferase (CPT1A), its product palmitoylcarnitine (AC 16:0, the rate limiting substrate in mitochondrial FA oxidation), is accumulated (**Figure 17 A, B**) suggesting an impairment of FA oxidation in *MAT1A*-KO mice. In line, several enzymes involved in catalyzing the oxidation of FAs in the mitochondria such as acyl-CoA synthetase medium chain family member 5 (ACSM5) and acyl-CoA dehydrogenase family member 8 (ACAD8) decrease content in *MAT1A*-KO mice livers. In contrast, levels of acyl-CoA oxidase 1 (ACOX1) and acyl-CoA acetyltransferase (ACAA1B) are higher. These enzymes are key for peroxisomal FA oxidation. We also found higher levels on enzymes involved in endoplasmic reticulum FA oxidation such as cytochrome P450 enzyme 2E1 (CYP2E1) and cytochrome P450 enzyme A10 (CYP4A10) (**Figure 17 B**). Moreover, the rise in oxidized FAs correlates with the increased protein content of aldehyde dehydrogenase 1 family member B1 (ALDH1B1), a key mitochondrial enzyme related with lipid peroxidation (**Figure 17 A, B**).

Finally, when considering the central enzymes involved in DNL, we found that the content of acetyl-CoA carboxylase 1 (ACC1) does not change while citrate lyase (ACLY) and FA synthase (FAS) levels decrease (**Figure 17 B**).

3.3.1 SAMe depletion alters mitochondrial polarization in *MAT1A*-KO mice

As we observed many alterations in FA oxidation, we wanted to know if mitochondrial integrity was affected by SAMe depletion in *MAT1A*-KO. We used flow cytometry to look for differences in murine hepatocytes (*MAT1A*-KO *vs* WT), after incubated for 4 hours with SAMe (4 mmol/L) or vehicle. As shown in **Figure 18**, membrane polarization is impaired in *MAT1A*-KO hepatocytes and SAMe treatment restores it.

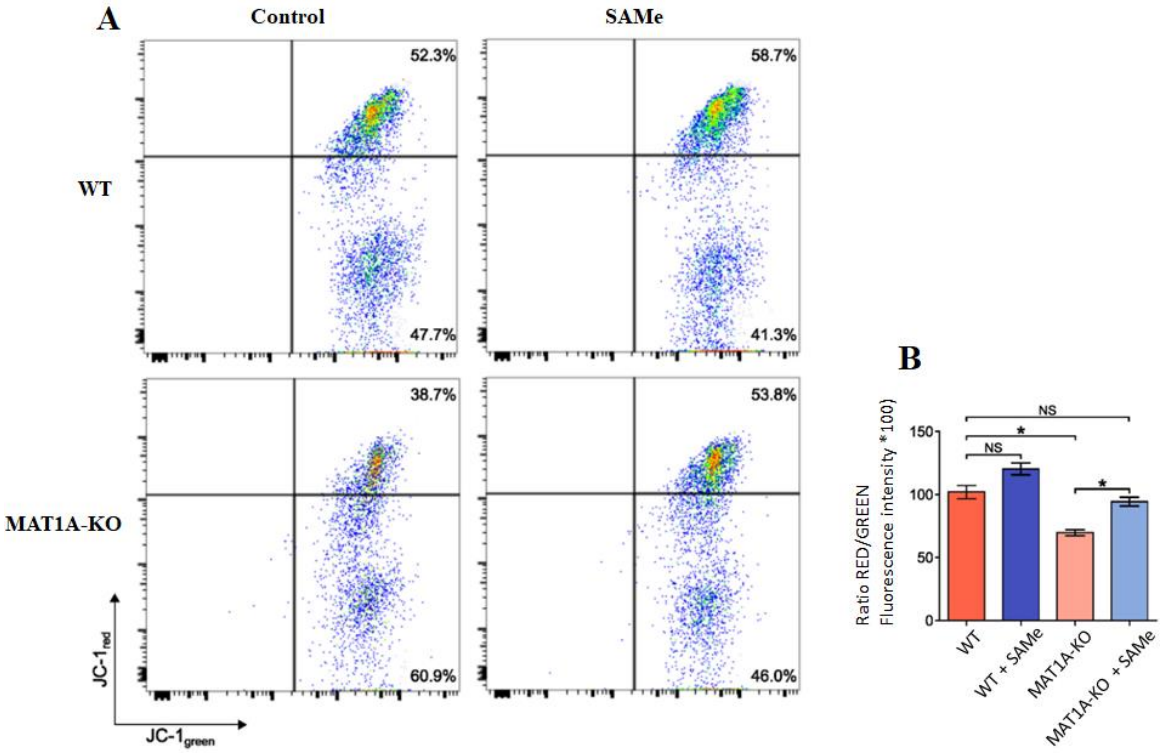


Figure 18. Mitochondrial membrane polarization was restored in *MAT1A*-KO mice upon incubation with SAMe. (A) Representative flow cytometric analysis of WT and *MAT1A*-KO hepatocytes treated with SAMe or vehicle. The percentage of JC-1red high/JC-1green high and JC-1red low/JC-1green high populations is presented. The data are representative of 2 independent experiments performed in triplicate. (B) WT and *MAT1A*-KO hepatocytes were incubated for 4 hours with 4 mmol/L SAMe or vehicle. The cells then were stained with 10 mg/mL of the dye JC-1, and assessed for red and green fluorescence by flow cytometry. The data represent the variation of the ratio between red and green fluorescence intensity relative to WT control hepatocytes (100%) performed in triplicate and are representative of 2 independent experiments. *Student t-test, * $p < 0.05$.

3.4 *MAT1A*-KO mice metabolomic analysis reveals a fingerprint also present in 50% of NAFLD patients.

3.4.1 Serum metabolomic profiling compares well with differences observed in the hepatic metabolism of *MAT1A*-KO mice

As mentioned, *MAT1A*-KO mice spontaneously develop NASH due to their reduced levels of SAMe. It is extremely educational to understand whether NAFLD patients will present metabolic hepatic alterations that are similar to the *MAT1A*-KO mice ones. To answer this question, we first investigated if the metabolomic serum and liver profiles in *MAT1A*-KO mice correlate or not.

Using LC/MS we measured the metabolites present in the liver and in the serum of *MAT1A*-KO and WT mice and compared them. We first created independent lists of significant metabolites, with the variation (fold-change) and p-value for each one of them (**Supplementary Table 1**). Subsequent comparison of both datasets revealed the common metabolites, and we conclude that serum metabolomic profile largely reflects the metabolism in the liver in *MAT1A*-KO mice (**Figure 19**). This correlation is statistically significant (p = 1e-04).

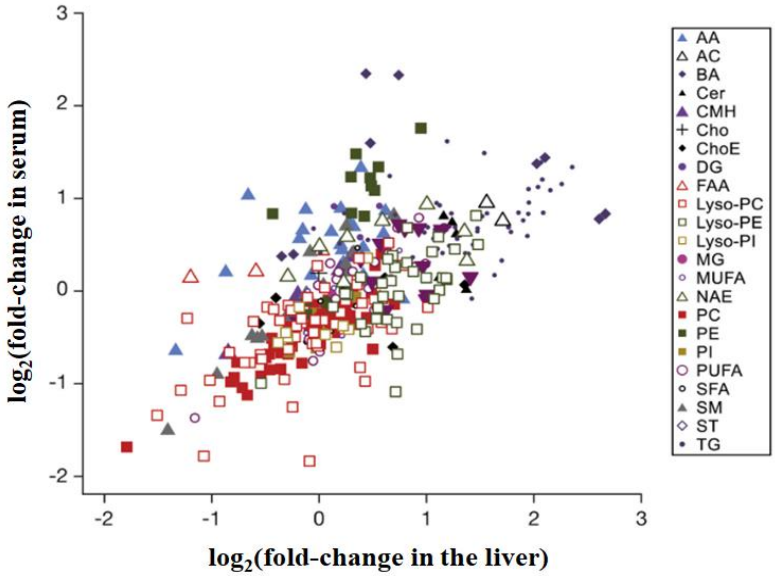


Figure 19. Serum metabolomic profile reflects hepatic metabolism**.** Comparison of liver and serum metabolomic profiles of *MAT1A*-KO mice. Each point represents the $\log_2$(fold-change) of individual metabolites of *MAT1A*-KO compared with WT mice in serum and liver. A list with the $\log_2$ (fold-change) and p-value for each metabolite in serum and liver is given in **Supplementary Table 1**. $R^2 = 0.45$; p = 1e-04. AA; amino acid; AC: acyl carnitine; BA: bile acid; Cer: ceramide; CMH: monohexosylceramide; Cho: cholesterol; ChoE: cholesteryl ester; FAA: fatty acyl amide; lyso-PC: lyso-phosphatidylcholine; lyso-PE: lyso-phosphatidylethanolamine; lyso-PI: lyso-phosphatidylinositol; MG: monoglyceride; MUFA: monounsaturated fatty acid; NAE: N-acylethanolamines; PI: phosphatidylinositol; PUFA: polyunsaturated fatty acid; SFA: saturated fatty acid; SM: sphingomyelin; ST: steroid.

3.4.2 Sub-classification of NAFLD patients

From the pool of significant metabolites in **Figure 19**, we selected the top 50 serum metabolites differing significantly between *MAT1A*-KO and WT mice (**Figure 20 A**). In this set, some metabolites increase/decrease in *MAT1A*-KO as compared to WT mice, as shown in the figure. We used this metabolic profile to analyze the serum metabolome of a cohort of 535 biopsy-proven NAFLD patients (**section 2.11.1**) in order to determine whether there are similitudes between *MAT1A*-KO mice and these patients. By Silhouette cluster analysis[171], we found that NAFLD patients were subclassified in 2 main clusters: the first one shows a serum metabolomic profile similar to that found in *MAT1A*-KO mice (we denominate it M-subtype) while the second one shows an alternative serum metabolomic profile (called non-M subtype) (**Figure 20 B**).

3.4.3 Validation of the two NAFLD patient's profiles associated with *MAT1A*-KO

3.4.3.1 Validation for randomize partition and comparison of estimation and validation cohorts

To validate these results, we observed the following procedure (**Figure 21**): first we randomly partitioned the human sample dataset into 2 cohorts (50/50), named estimation and validation cohorts. Since 353/182 patients were diagnosed with steatosis/NASH, the estimation and validation cohorts were generated abiding the same proportional representation also using the same protocol for gender. Next, we performed a clustering analysis (based on the selected top 50 metabolites differing between *MAT1A*-KO and WT mice, **Figure 21**) and patients were classified into two subtypes: M-subtype and non-M subtype.

Based on the entire metabolic profile of the human serum samples (328 metabolites) and using univariate analysis we identified and validated the metabolites that were significantly different between NASH and simple steatosis, named biomarkers. Estimation and validation cohorts assisted in the validation of the biomarkers that were further stressed to a 1000-fold repetition of the random partition, each time with a proportional representation of steatosis/NASH and males/females. This iterative process allowed determining the frequency distribution of metabolites that significantly discriminate between steatosis and NASH in both subtypes: M and non-M, (**Figure 21**). Metabolites showing a reproducibility of ≥ 70% (at least 700 times in 1000 repetitions) in the test were selected and sorted out according to the NASH subtype reproducibility and p-value (**Supplementary Table 2**).

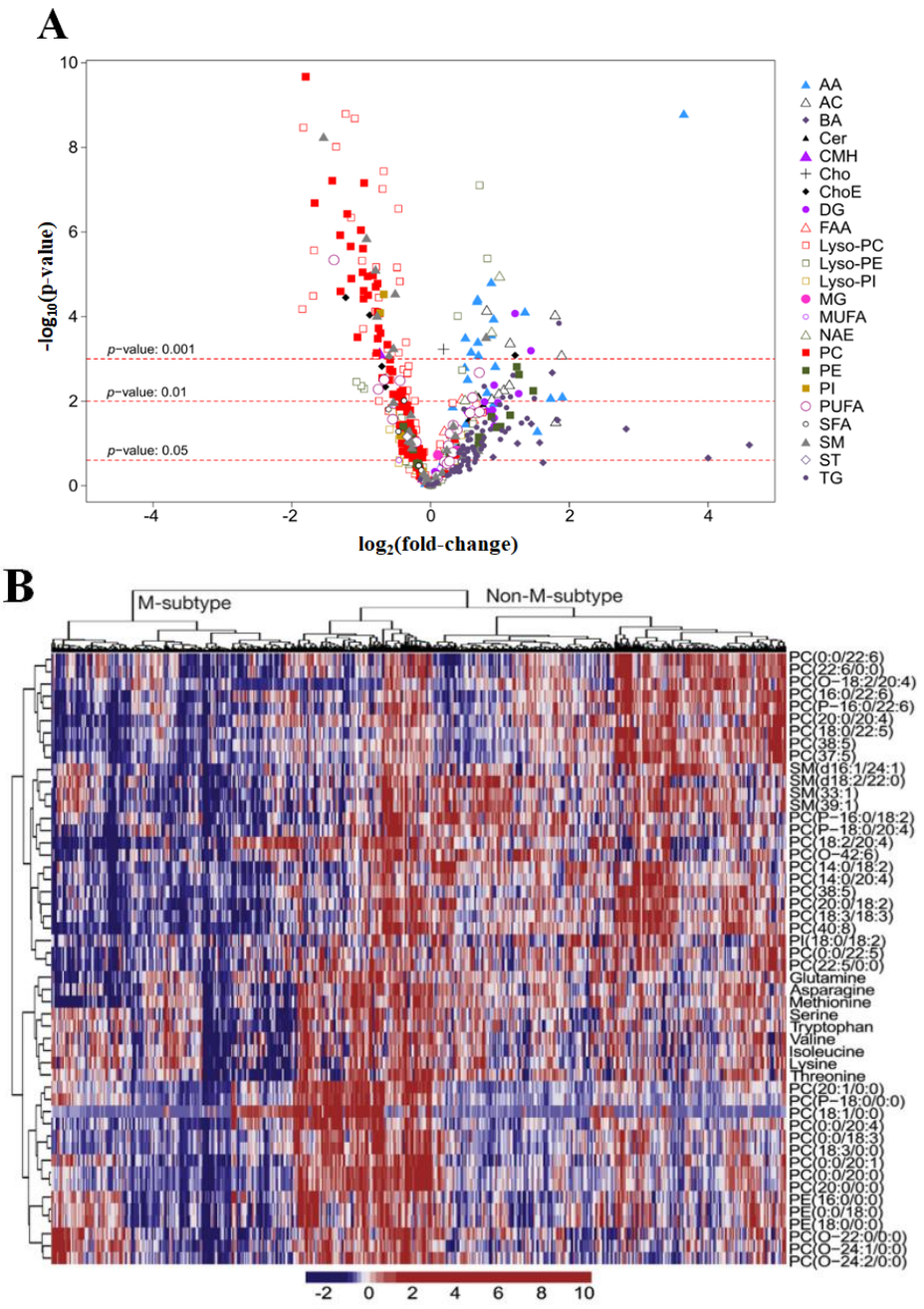


Figure 20. Identification of a subset of NAFLD patients showing a *MAT1A*-KO serum metabolomic profile. **(A)** Volcano plot representation indicating the -$\log_{10}$(p-value) and $\log_2$(fold-change) of individual serum metabolic ion features of *MAT1A*-KO compared with WT mice. AA, amino acid; AC, acyl carnitine; BA, bile acid; Cer, ceramide; CMH, monohexosylceramide; Cho, cholesterol; ChoE, cholesteryl ester; FAA, fatty acyl amide; lyso-PC, lyso-phosphatidylcholine; lyso-PE, lyso-phosphatidylethanolamine; lyso-PI, lyso-phosphatidylinositol; MG, monoglyceride; MUFA, monounsaturated fatty acid; NAE, N-acylethanolamine; PI, phosphatidylinositol; PUFA, polyunsaturated fatty acid; SFA, saturated fatty acid; SM, sphingomyelin; ST, steroid. **(B)** Heatmap representation of the serum metabolomic profile from 535 patients with biopsy-confirmed NAFLD. Each data point corresponds to the relative ion abundance of a given metabolite (vertical axis) in an individual patient's serum. Metabolite selection is based on the top 50 serum metabolites that differentiated more significantly between *MAT1A*-KO and WT mice. The hierarchical clustering is based on the optimum average silhouette width, dividing the classification of the samples into 2 groups: the first cluster resembles the serum metabolomic profile observed in the *MAT1A*-KO mice (M subtype), and the second cluster shows a different metabolomic profile (non-M subtype).

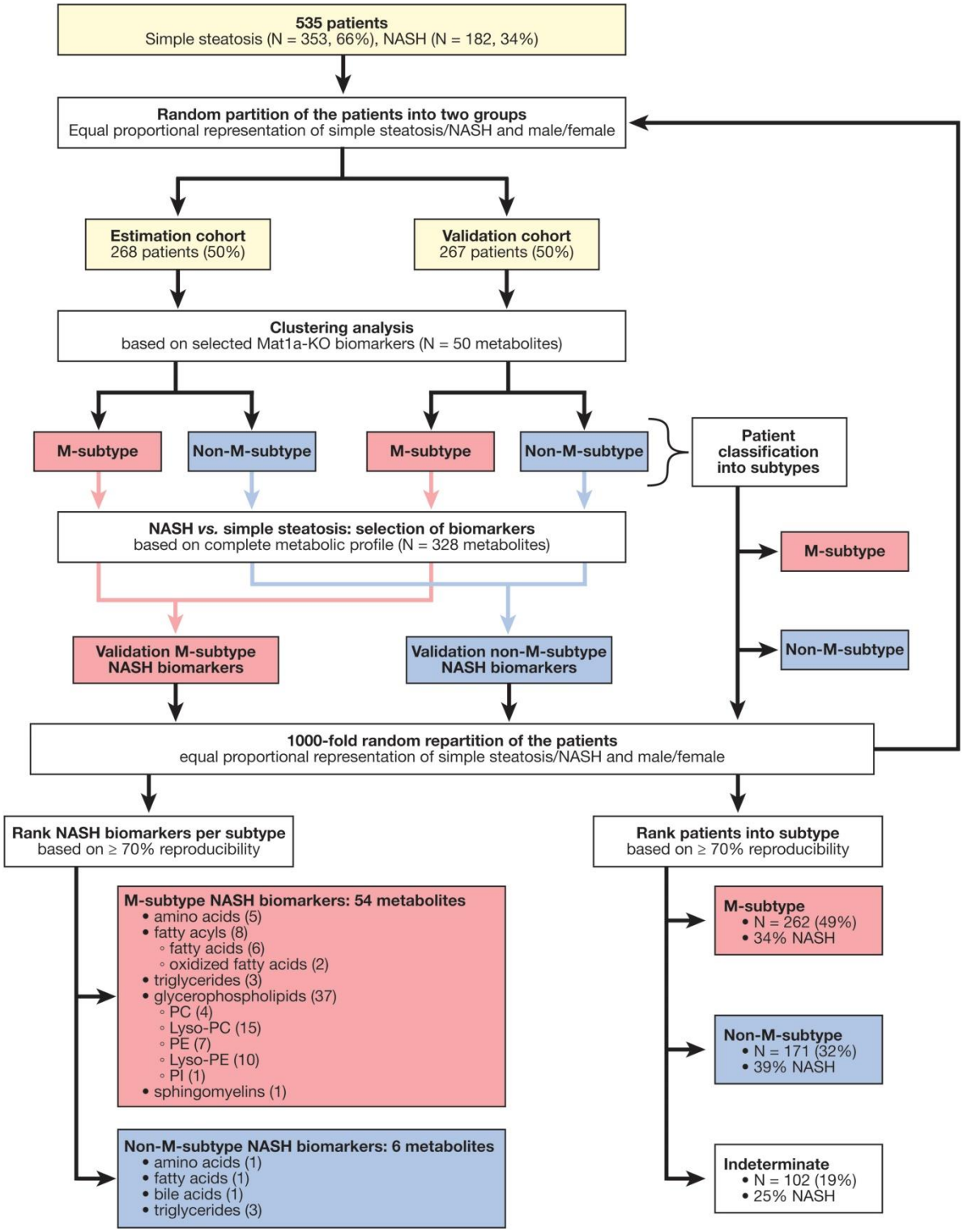


Figure 21. Scheme for the identification and validation of NAFLD subtypes and NASH biomarkers. Lyso-PC, lyso-phosphatidylcholine; lyso-PE, lyso-phosphatidylethanolamine; PI, phosphatidylinositol.

NASH biomarkers for the M-subtype include 54 metabolites: 8 fatty acids (6 FA and 2 oxFA), 37 phospholipids (4 PC, 7 PE, 15 lyso-PC, 10 lyso-PE and 1 phosphatidylinositol), 3 TGs and 5 aminoacids (**Figure 21** and **Supplementary table 2a**). A subset of 29 of these metabolites have a reproducibility of ≥ 90%. Interestingly, 25 out of the 54 biomarkers are lyso-phospholipids. On the other hand, for non-M subtype we found 6 NASH biomarkers including 1 amino acid, 1 FA, 1 bile acid, and 3 TGs (**Figure 21** and **Supplementary table 2b**).

Following the criteria based on ≥ 70% reproducibility we calculated the frequency distribution of the NAFLD patients into both subtypes M and non-M. 262 patients out of 535 were classified as M subtype (49%), 171 as non-subtype (32%) and the remaining 102 (19%) could not be classified in neither of subtypes because their reproducibility was less than 70%, so we called them "indeterminate group". **Supplementary table 3** summarizes the serum metabolites that are associated with the *MAT1A*-KO serum metabolomic profile in the human subtype. The percentage of NASH patients in the M-subtype group and in the total cohort is the same (34%), lower that the percentage of NASH patients with non-M subtype (39%) and higher that the indeterminate group (25%) (**Table 3**).

There were more women than men in the steatosis and NASH groups of M-subtype when compared with the non-M subtype but this introduced no bias in the information that can be obtained from the serum metabolome, as deduced from the PCA analysis by gender (**Figure 22**).

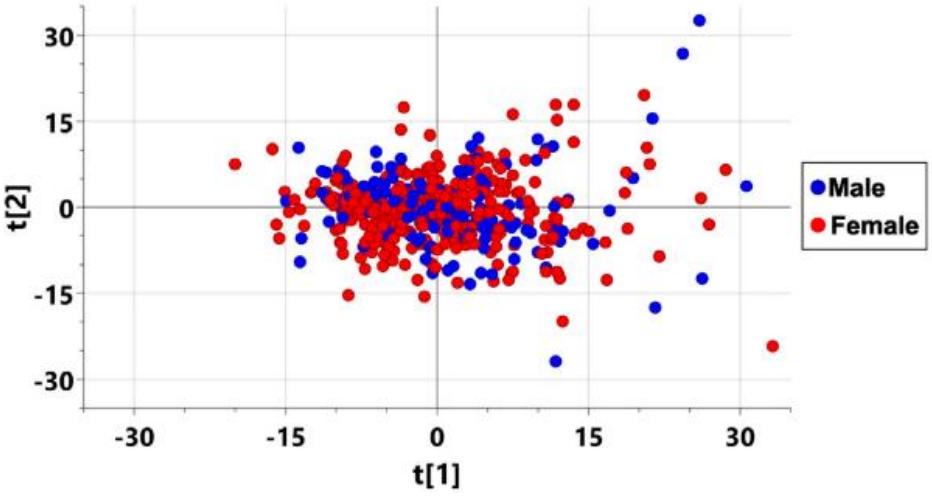


Figure 22. PCA plot of human serum samples classified according to gender. PCA analysis showed that males (30%) and females (70%) clustered together. Each individual is represented by 1 dot and the color corresponds to females (red) and males (blue).

3.4.3.2 Validation with methionine choline deficient (MCD) fed-mice NASH model

To further validate the classification of patients between the two groups (M and non-M type) we repeated the analysis with the patient's cohort but, this time, comparing their metabolomes with the metabolome of the MCD deficient mice (a dietary model of NASH), which also shows a deficiency in SAMe.

To avoid body weight loss caused by the canonical MCD diet[172], we supplemented the diet with 0.1% methionine. The resulting 0.1MCD diet also induces steatosis, inflammation and fibrosis (**Figure 23 A,B**). Body weight is lower as compared to control group but it gets stable after the first 2-3 days (**Figure 23 C**). The increase of ALT and AST enzymes in serum is similar to the levels found in the canonical MCD diet-fed mice (**Figure 23 D**). 0.1MCD diet fed-mice showed a reduction in hepatic SAMe and an increase in SAH (**Figure 24 A, B**) which means a reduction in SAMe/SAH ratio, known to cause inhibition of PEMT. As mentioned in **section 1.1.2**, PEMT is the enzyme that generates PC by the N-methylation of PE and converts PE rich PUFAs to PUFA-rich PCs, ultimately required to synthetize VLDLs. Consistently, mice fed with 0.1MCD show an increase in hepatic PE (22:6) and decrease PC (22:6) levels (**Figure 24 A, B**), the main phospholipid in the outer coat of VLDLs. VLDL impaired formation and export results in a TGs accumulation in the liver. Finally, 0.1MCD diet also causes reduction in serum TGs levels, PC (22:6) serum levels (as compared to those fed a normal diet) and reduce levels of hepatic GSH and cystathionine (**Figure 24 A, B**). GSSG levels are increased in 0.1MCD mice; being the reduced GSH/GSSG ratio a biomarker of oxidative stress[173].

At the protein expression level, several GSH-consuming enzymes such as GSH peroxidase 1 (GPX1), GSH S-transferases 1-3 (GSTM 1-3) and thioredoxin reductase 1 (TXNRD1) increase content in 0.1MCD diet-fed animals as compared to chow diet and the content of GSH synthase (GS) also increases (**Figure 24 A, C**). The first enzyme connecting SAMe metabolism with the transsulfuration pathway is CBS, which is reduced in 0.1MCD mice. In turn, in 0.1MCD mice, glutamate-cysteine ligase (GCL) catalytic subunit levels increase but not the amount of the corresponding modifier subunit (the regulatory subunit of GCL, **Figure 24 A, C**). These results suggest that the fall in GSH and the reduction in GSH/GSSG ratio are due to increased oxidative stress and impaired oxidative response.

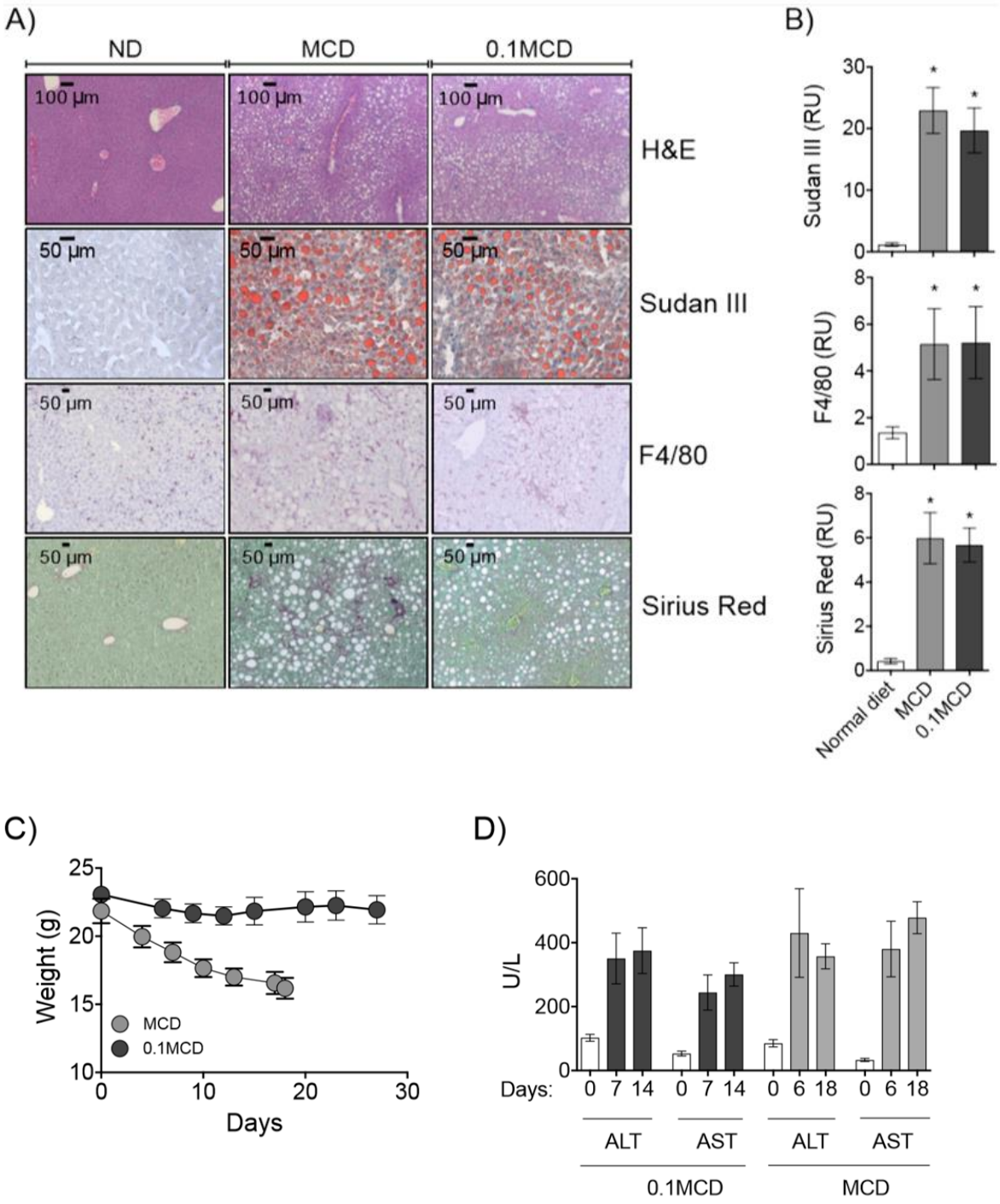


Figure 23. **(A)** Histology of livers from WT mice fed with normal diet (ND), canonical MCD diet (MCD) and MCD diet with 0.1% of methionine (0.1MCD). H&E, Sudan III red, Sirius red and F4/80 immunofluorescence staining of livers are shown. **(B)** Sudan III red, Sirius red and F4/80 quantification. **(C)** Body weight loss in MCD and 0.1MCD-fed mice along 30 days. **(D)** Serum levels of ALT and AST enzymes in mice fed the 0.1MCD diet and the canonical MCD diet.

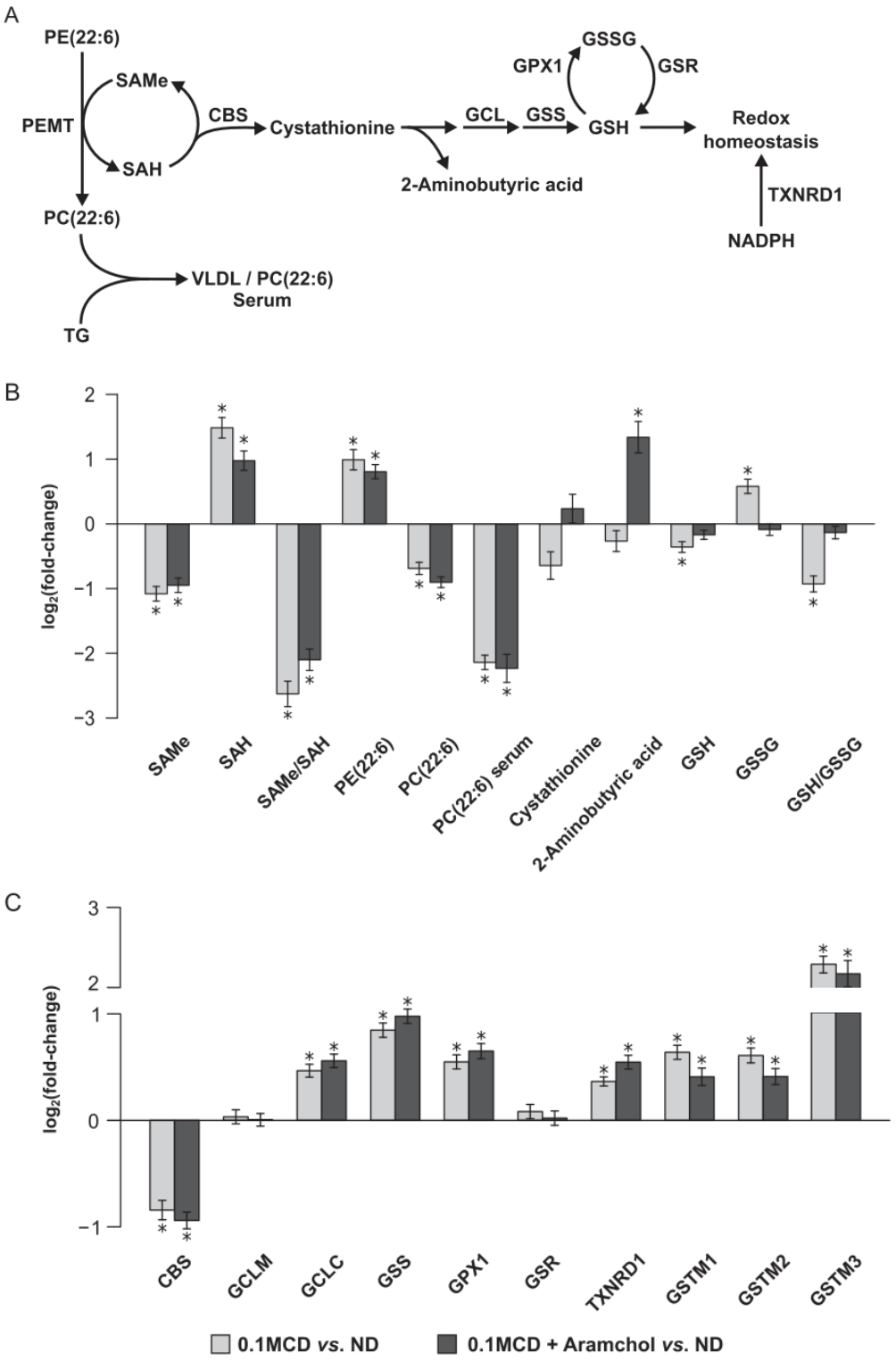


Figure 24. Hepatic SAMe metabolism in mice fed a 0.1MCD diet and treated with vehicle or with Aramchol. **(A)** Schematic representation of hepatic SAMe metabolism. **(B)** Relative fold change (log_2) in the hepatic content of the main metabolites involved in hepatic SAMe metabolism in mice fed a 0.1MCD diet compared to mice fed a normal diet. **(C)** Relative fold change (log_2) in the protein content of enzymes involved in hepatic SAMe metabolism in 0.1MCD-fed mice treated with vehicle or with Aramchol compared to mice fed a normal diet. Data were represented as mean ± SEM. * $p < 0.05$.

We next investigated if the cohort of 535 NAFLD-biopsy-proven patients showed alterations in liver metabolism similar to that found in 0.1MCD mice. To that end, we compared the liver and serum metabolome of 0.1MCD mice and those fed a control diet and generated two lists of metabolites (for the liver and serum metabolomes) showing the fold changes and p-values for each metabolite (**Supplementary Table 4**). Then, we analyzed if the serum metabolic profile of 0.1MCD mice reflected their liver metabolism by comparing both sets of data and we found that this system is defined by a bi-linear regression plots, one made by the pool of TGs (R = 0.865, p-value = 2.5e-12) and the second one made by all lipids but TGs (R = 0.576, p-value = 2.2e-16) (**Figure 25**).

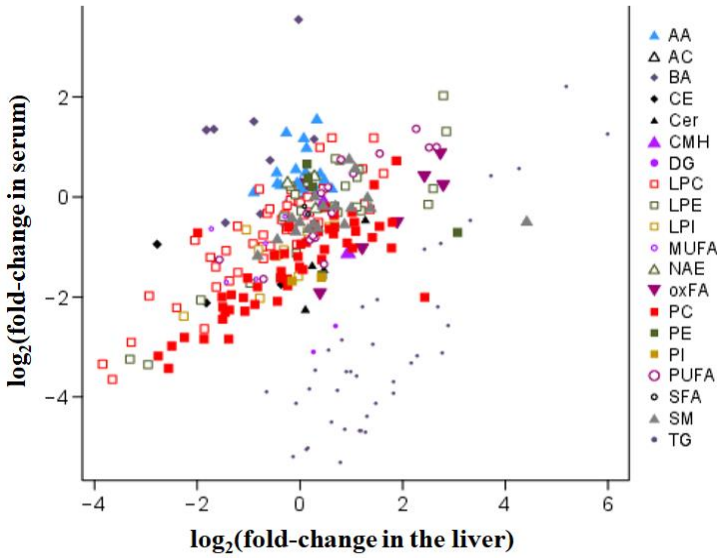


Figure 25. Serum metabolomic profile reflects hepatic metabolism in 0.1MCD diet fed-mice. Comparison of liver and serum metabolomic profiles of 0.1MCD diet fed-mice. Each point represents the log_2(fold-change) of individual metabolites of 0.1MCD diet fed-mice compared with normal diet-fed mice in serum and liver. AA; amino acid; AC: acyl carnitine; BA: bile acid; Cer: ceramide; CMH: monohexosylceramide; DG: diglyceride; LPC: lyso-phosphatidylcholine; LPE: lyso-phosphatidylethanolamine; LPI: lyso-phosphatidylinositol; MUFA: monounsaturated fatty acid; NAE: N-acylethanolamines; oxFA: oxidized fatty acid; PC: phosphatidylcholine; PE: phosphatidylethanolamine; PI: phosphatidylinositol; PUFA: polyunsaturated fatty acid; SFA: saturated fatty acid; SM: sphingomyelin; TG: triglycerides.

Afterwards, we selected the top 50 serum metabolites in discriminating between 0.1MCD and normal diet-fed mice (**Figure 26 A**) and used this metabolic signature to classify the 535 biopsy-proven NAFLD patients. Again by Silhouette cluster analysis, patients were subclassified into two main clusters: one cluster with a metabolic profile similar to the one found in 0.1MCD mice and the other showing a different profile (**Figure 26 B**).

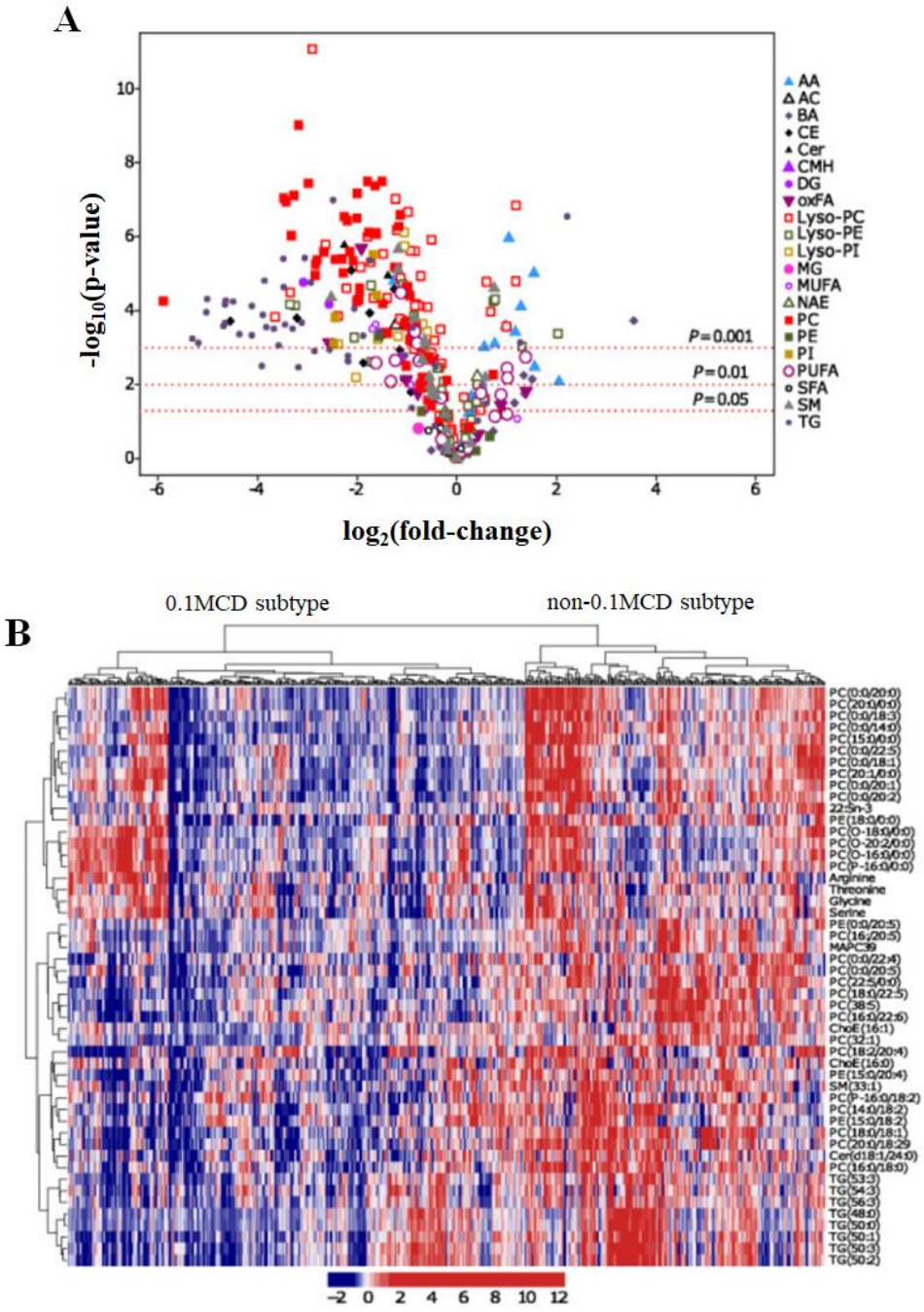


Figure 26. Identification of a subset of patients with NAFLD showing a serum metabolomic profile similar to that of the 0.1MCD mouse model. **(A)** Volcano plot representation indicating the $-\log_{10}$(p-value) and $\log_2$(fold change) of individual serum metabolic ion features of mice fed a 0.1MCD diet compared to mice fed a normal diet. **(B)** Heatmap representation of the serum metabolomic profile from 535 patients with biopsy-confirmed NAFLD. The scale indicates (vertical axis) the relative ion abundance of each metabolite in the serum extract of a given subject with respect to that found in the rest of the study population, 0 being the mean value. The hierarchical clustering is based on optimum average silhouette width obtaining two main clusters; the first (left-hand cluster) resembles the serum metabolomic profile observed in mice fed a 0.1MCD diet (0.1MCD subtype), while the second cluster (right-hand cluster) shows a different metabolomic profile (non-0.1MCD subtype).

For validation of this classification we followed, once more, the procedure developed and explained in **section 3.4.3.1,** that consists in carry out random partition of samples, 50/50, in two cohorts (estimation and validation) and perform the Silhouette cluster analysis aforementioned to see which patients have a serum metabolic profile similar to the 0.1MCD diet fed-mice profile. Then, we repeated this random partition 1000 times and patients showing a metabolic profile resembling the 0.1MCD mice model ≥ 700 times were selected.

Out of this analysis, 250 patients (out of 535, 47%) had a metabolic phenotype similar to the one found in 0.1MCD diet fed-mice and we name it 0.1MCD subtype. On the other hand, 147 patients (27%) had a different profile so we call them non-0.1MCD subtype and the remaining 138 patients (26%) were indeterminate as they were classified < 700 times as 0.1MCD-subtype and non-0.1MCD subtype indistinctly (**Figure 27 A**). Remarkably, 90% (n = 225) of the patients classified here as 0.1MCD subtype were also classified as M-subtype when compared with *MAT1A*-KO mice metabolic profile and we call this group M^+ subtype (**Figure 27 A**). On the other hand, 73% (n = 107) patients classified as non-0.1MCD subtype were also classified as non-M subtype when comparing them with *MAT1A*-KO mice model, so we call this group non-M^+ subtype (**Figure 27 A**). A third miscellanea group, called I^+ subtype, encloses the remaining 203 patients that do not belong to any of the two subtypes (**Figure 27 A**).

These results support the hypothesis that low levels and impaired synthesis of SAMe in the liver may be common feature in NAFLD patients.

Finally, we compared the serum metabolomic profiles of the M^+ and non-M^+ subtypes (already validated) and found that 93.6% of all the analyzed amino acids and lipids were significantly reduced in the M^+ subtype (**Figure 27 B**). The analyzed lipids included saturated FAs, MUFAs, PUFAs, DGs, TGs, ethanolamine and choline glycerophospholipids (diacyl, plasmalogens, and lyso phospholipids), bile acids, cholesteryl esters and sphingolipids (ceramides and sphingomyelins). The serum levels of glycine, methionine, serine and threonine (the amino acids that feed 1CM to form SAMe) and the levels of taurine (the main output of SAMe catabolism) are reduced in serum of M^+ subtype patients as compared to non-M^+ subtype patients. The I^+ subtype showed higher levels of serum amino acids and lipids than M^+ subtype but lower than the non-M^+ subtype (**Figure 27 B**).

These results suggest that M^+ subtype NAFLD patients (as well as the *MAT1A*-KO and 0.1MCD mice models) may have impaired VLDL formation and export. The secretion of other lipid vesicles (as exosomes) may be affected too in M^+ subtype patients compared to the non-M^+ subtype.

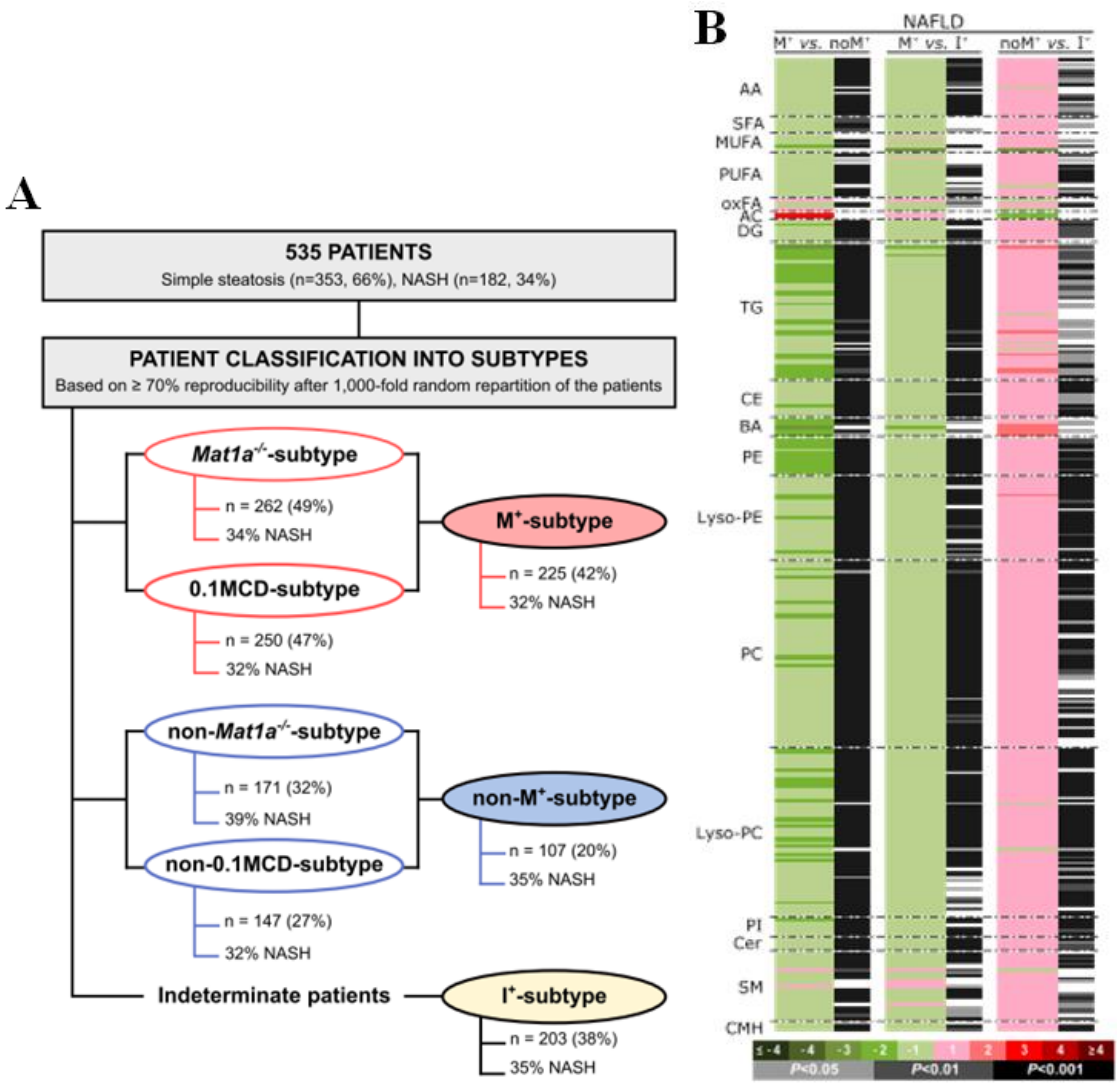


Figure 27. Scheme for the identification and validation of NAFLD subtypes M^+ and non-M^+ subtype. **(A)** Schematic representation summarizing the results obtained from the metabocentric analysis**. (B)** Heatmap representing the comparisons of the serum metabolic profiles of validated M^+, non-M^+, and I^+ groups, displaying the $\log_2$(fold change) and unpaired student t-test. Abbreviations: AA, amino acids; AC, acyl carnitines; BA, bile acids; CE, cholesteryl esters; Cer, ceramides; CMH, monohexosylceramides; oxFA, oxidized fatty acids; MG, monoglycerides; NAE, N-acylethanolamines; SFA, saturated fatty acids; SM, sphingomyelins.

3.4.4 Hepatic lipid and protein content alterations in 0.1MCD diet fed-mice

We analyzed the lipid composition that was altered in mice fed 0.1MCD diet to find a hepatic lipid accumulation of TGs, DGs, cholesteryl esters and FAs as well as oxidized FAs. Noteworthy, one of these metabolites are oxidized derivatives of linoleic acid (oxLA), that have been proven to be a human NASH biomarker[174] (**Figure 28 A, B**).

On the other hand, we also wanted to investigate which proteins related to hepatic lipid metabolism may be altered in 0.1MCD mice. Proteomics data analysis showed that 0.1MCD mice had increased levels for a set of proteins as compared to controls. Such set included the CD36 FA transporter, AGPAT3, CPT1A, HADH A and B, UCP2 (mitochondrial uncoupling protein 2), and CYP4A14 (microsomal cytochrome P450 family 4, subfamily a, polypeptide 14) (**Figure 28 C**), which agrees with previous findings[175,176]. On the other hand, other proteins like SCD1, ACC1, ACC2, ACLY and FAS showed decreased concentration.

HADHA and HADHB are subunits of mitochondrial trifunctional protein which catalyzes the last three steps of mitochondrial β-oxidation of long chain FAs. In turn, AGPAT3 is a key enzyme involved in TG synthesis since it is involved in phospholipid metabolism converting lysophosphatidic acid (LPA) into phosphatidic acid (PA). Otherwise, the protein content of other enzymes involved in DNL such as ACLY, ACC1, ACC2 and FAS were slightly diminished (**Figure 28 C**). Finally, ACC2 is the enzyme that synthetizes mitochondrial malonyl-CoA (an inhibitor of CPT1A) and as its protein content was reduced, it indicates that MCD-mediated reduction of SCD1 stimulates FA oxidation through the β-oxidation pathway, which agrees with previous studies[175].

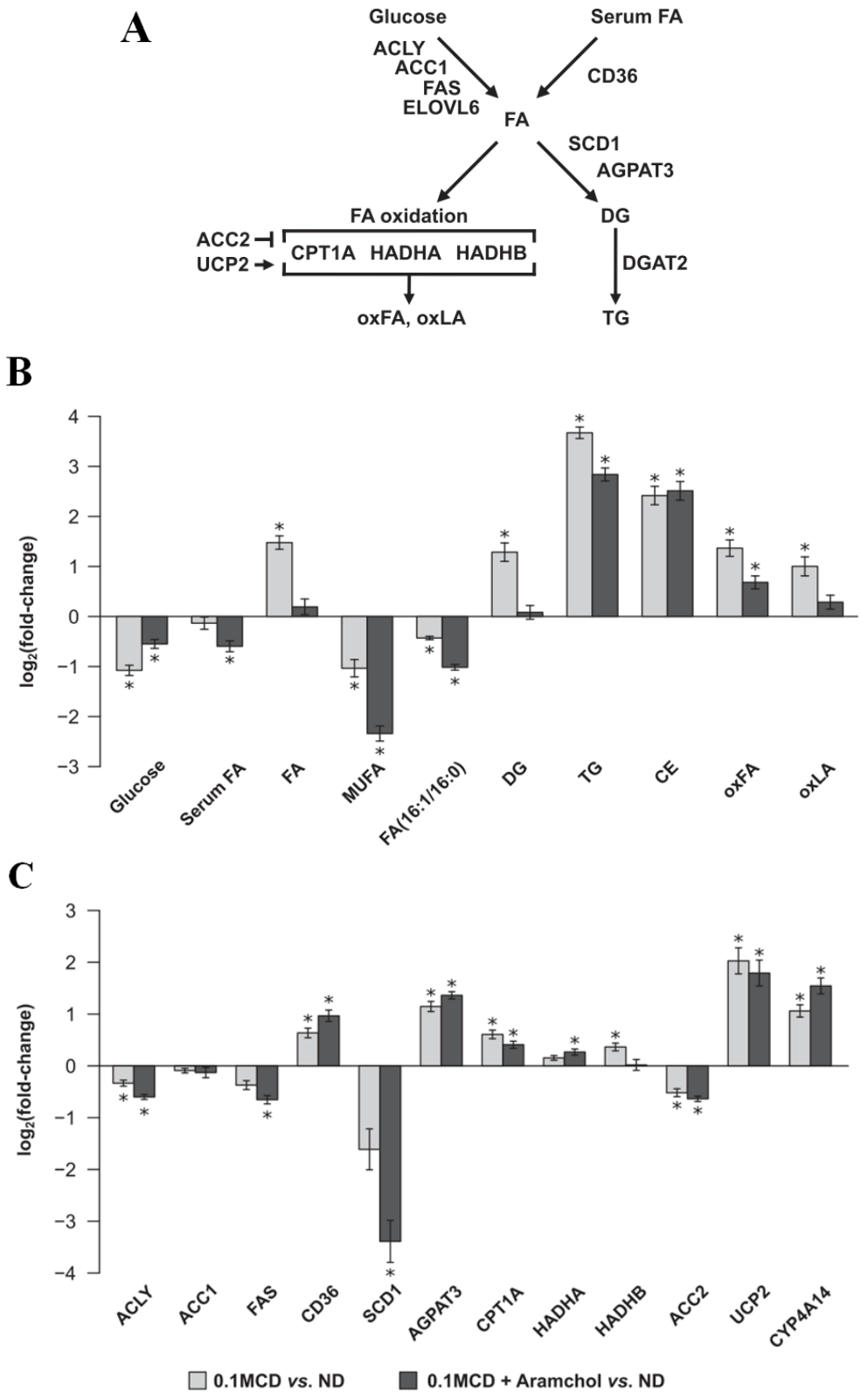


Figure 28. Hepatic lipid metabolism in 0.1MCD mice treated with vehicle or with Aramchol. (A) Schematic representation of hepatic lipid metabolism. **(B)** Relative fold change (log_2) in the hepatic content of the main metabolites involved in lipid metabolism in mice fed a 0.1MCD diet and treated with vehicle or with Aramchol compared to mice fed a normal diet. **(C)** Relative fold change (log_2) in the content of proteins involved in liver lipid metabolism in mice fed a 0.1MCD diet and treated with vehicle or with Aramchol compared to mice fed a normal diet.

3.5 Aramchol administration reduces NASH features in 0.1MCD diet fed-mice.

As mentioned previously (**section 1.1.7.1**), Aramchol is a lipidic molecule resulting from the synthetic conjugation of a bile acid and a saturated FA. It was proven that Aramchol inhibits SCD1 activity promoting FA β-oxidation and decreasing FA synthesis[49]. To see if Aramchol treatment reduces steatohepatitis, we tested the compound (5mg/kg/day vs vehicle) during 2 weeks in a murine model of NASH (0.1MCD mice, already fed during 2 weeks).

As depicted in **Figure 29**, mice treated with Aramchol showed improved liver histology when compared to mice given vehicle, as indicated by hematoxylin and eosin histology and confirmed by Sudan III, Sirius Red (showing less lipid accumulation in liver), F4/80 and CD64 (reporting less liver inflammation). COL1A1 western-blotting staining and quantification (**Figure 30 A**) showed that 0.1MCD mice treated with Aramchol had reduced levels of COL1A1 and that these levels were comparable to those in mice fed with a normal diet (**Figure 30 A**).

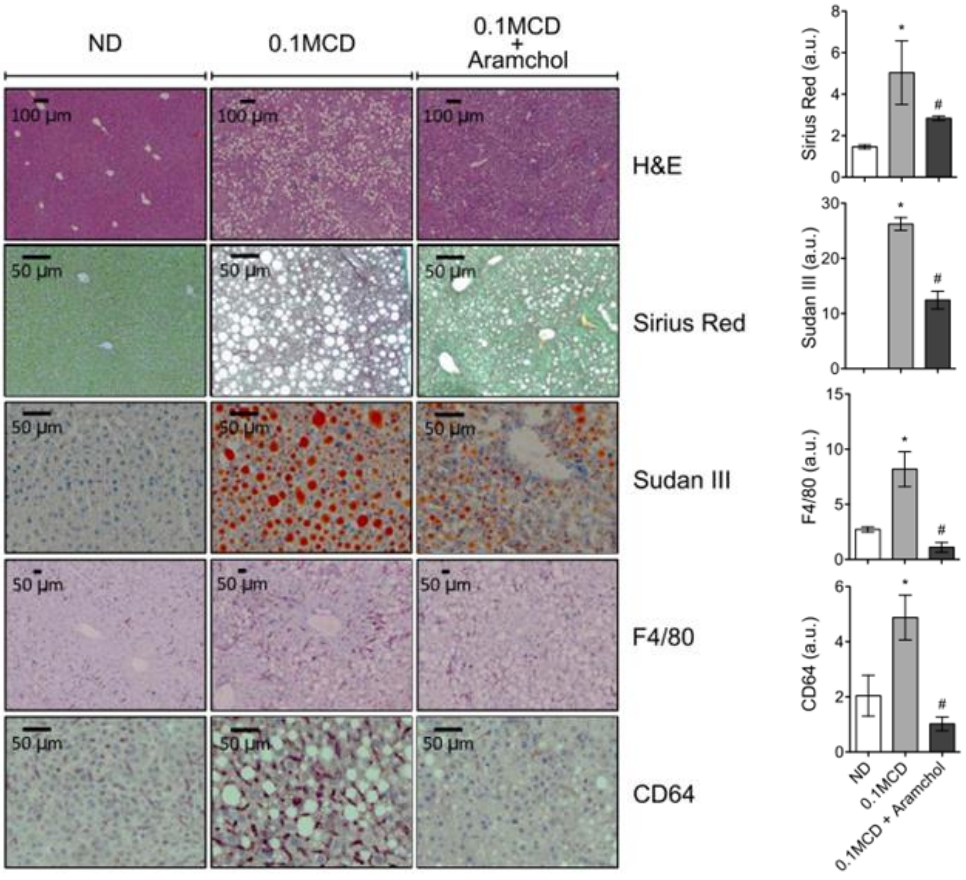


Figure 29. Aramchol ameliorates liver histology in 0.1MCD-fed mice. Hematoxylin and eosin staining, Sirius Red, and Sudan III staining and quantification of positive areas; F4/80 and CD64 immunostaining and quantification of positive areas.

Measurements of the serum liver enzymes AST and ALT showed that feeding 0.1MCD diet to mice for 4 weeks caused a significant increase of these enzymes when compared with normal-diet mice. Such increase is also found in mice treated with Aramchol, but it is not statistically significant (**Figure 30 B**). 0.1MCD mice also showed a reduction in serum TGs levels, unrecovered upon Aramchol treatment (**Figure 30 B**).

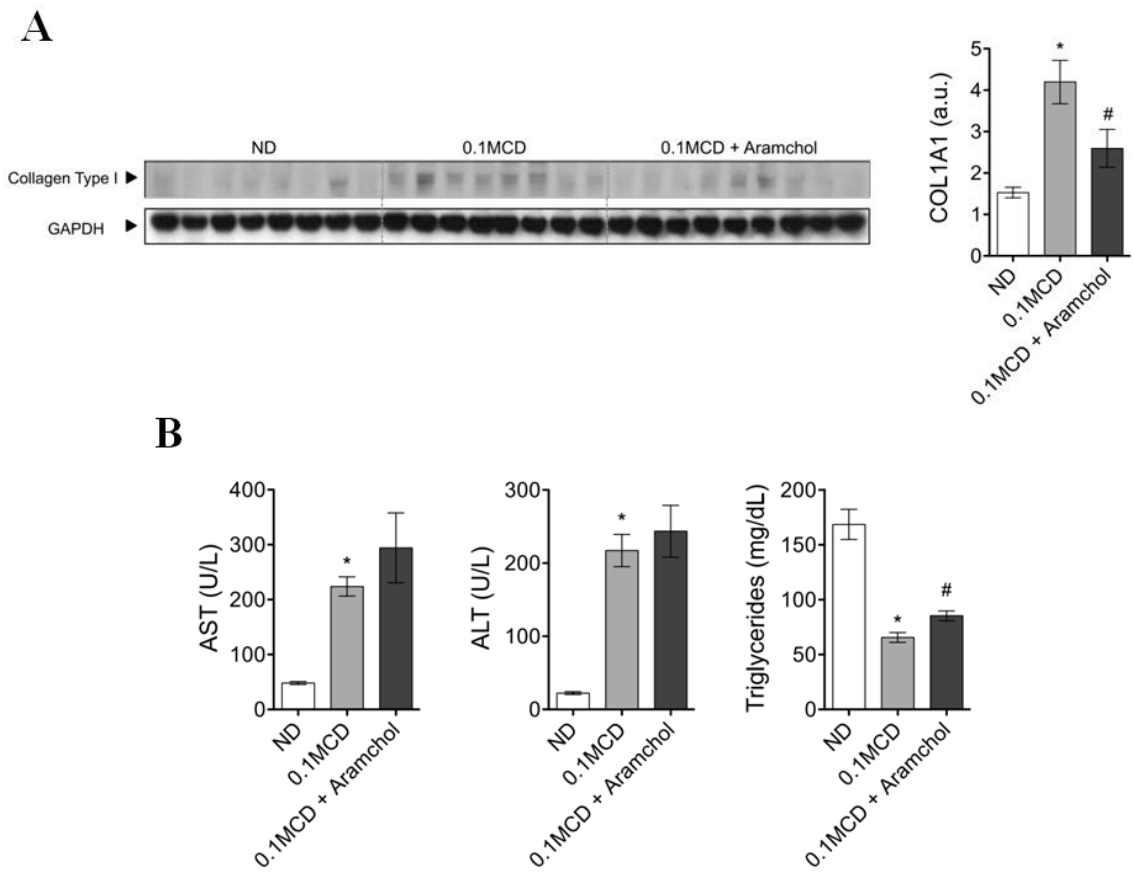


Figure 30. Aramchol treatment reduces COL1A1 protein content but did not affect serum enzymes or TGs in 0.1MCD diet fed-mice. **(A)** Collagen type I protein content in the liver of mice was assessed by immunoblotting using GAPDH as the loading control. **(B)** Serum levels of ALT, AST, and TGs. * $p < 0.05$ versus control; # $p < 0.05$ versus MCD. Data were represented as mean 6 SEM. Abbreviations: ND, normal diet; a.u., arbitrary units.

In agreement with the decrease in fibrosis and inflammation we found that Aramchol treated 0.1MCD mice, showed a recovery to normal levels in some metabolites such as FAs, oxFAs (**Figure 28 B**), cystathionine and GSH as well as in the GSH/GSSG ratio (**Figure 24 B**). Administration of Aramchol to 0.1MCD mice did not caused significant effects in protein content of enzymes involved in GSH synthesis and redox regulation (CBS, GCL catalytic subunit, GCL modifier subunit, GS, GSR, GPX1, GSTM1-3, and TXNRD1) (**Figure 24 C**). Yet, aminobutyric acid (a biomarker of the flux through transsulfuration pathway)[177] was markedly increased in animals treated with Aramchol (**Figure 24 B**).

The levels of enzymes involved in DNL, FA uptake, TG synthesis and FA β-oxidation were similar in 0.1MCD mice treated with Aramchol or vehicle with the exception of SCD1, whose reduced levels were further validated by measuring the FA (16:1)/(16:0) ratio and the content of MUFAs (**Figure 28 B, C**).

3.6 Identification of serum biomarkers associated with Aramchol treatment

As we mentioned in previous sections, there is a need to identify reliable biomarkers that differ between several kind of NAFLD patients and healthy people but also check their progression under treatment.

To that end, a volcano plot analysis that compared serum metabolome of mice treated with Aramchol versus vehicle (**Supplementary Table 5**) uncovered up to 91 metabolites able to discriminate between both mice (**Figure 31 A**, **Supplementary Table 5**). Alternatively, when comparing the liver metabolome of these two groups of mice, 30 metabolites differ significantly between the 2 groups (**Figure 31 B**). These metabolites include: 2-aminobutiric acid (that increases after treatment with Aramchol), 3 MUFAs, 10 PUFAs, 4 lyso-PE, 9 lyso-PC and 3 lyso-PI, whose levels decrease in mice treated with Aramchol (**Figure 31 B**).

To evaluate the effect of dose-response on the potential biomarkers, we tested two doses of Aramchol (1 mg/kg/day and 5 mg/kg/day) in serum and liver. The 1 mg/kg/day dose did not ameliorate liver histology (data not shown) and, as shown in **Figure 31 B**, it only results in a modification in 2 lyso-PCs in serum and 3 lyso-PCs and 2 lyso-PI in the liver.

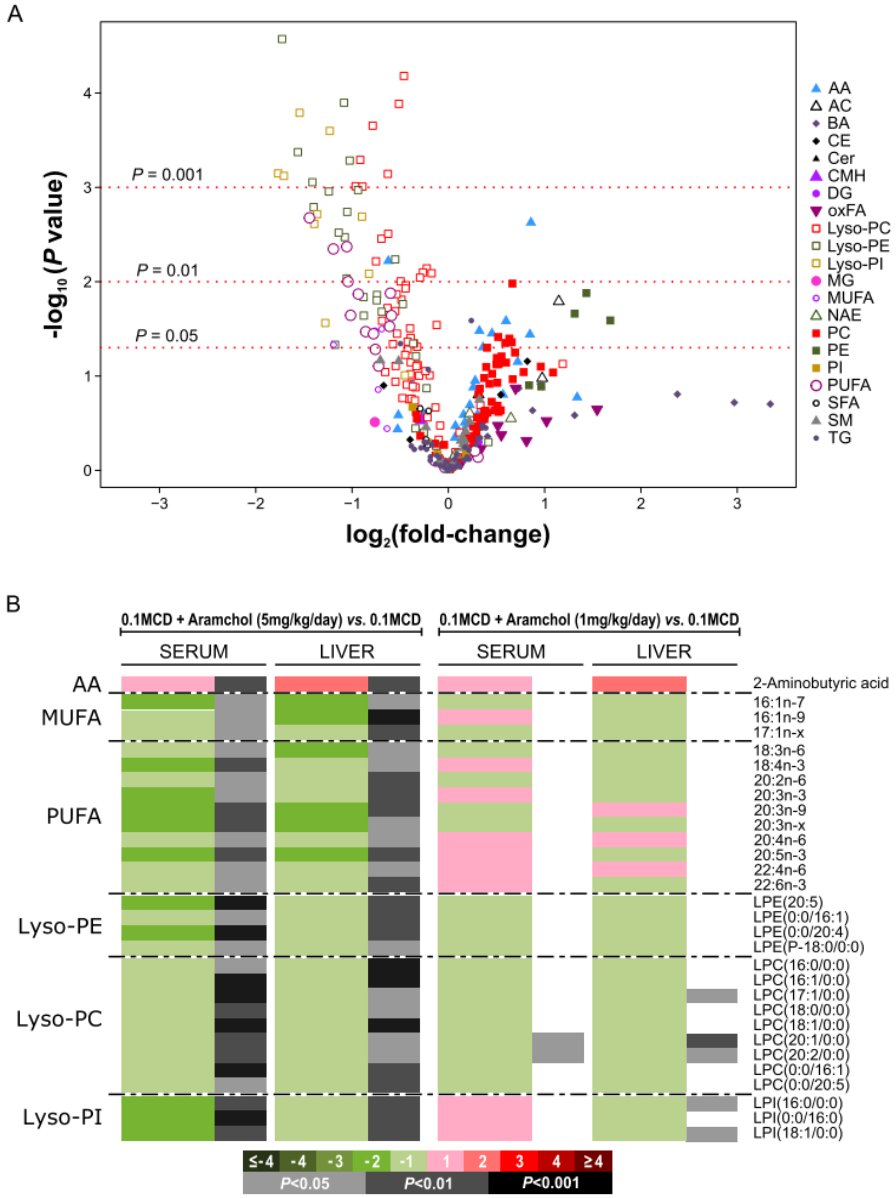


Figure 31. Identification of serum biomarkers associated with Aramchol treatment. **(A)** Volcano plot representation indicating the $-\log_{10}$(p value) and $\log_2$(fold change) of individual serum metabolic features of mice fed a 0.1MCD diet and treated with Aramchol (5 mg/kg/day) compared to untreated mice fed a 0.1MCD diet. **(B)** Heatmap of the 30 serum metabolites found to reflect the effect of Aramchol in the liver. This graph shows the $\log_2$(fold change) together with the unpaired student t-test of each metabolite. For the comparisons, log-transformed ion abundance ratios are depicted as represented by the scale. Darker green and red colors indicate higher drops and elevations of the metabolite levels, respectively, in every comparison. Gray lines correspond to significant fold changes of individual metabolites; darker gray colors have been used to highlight higher significances ($p < 0.05$, $p < 0.01$, $p < 0.001$). Metabolites have been ordered in the heatmap according to their carbon number and unsaturation degree of their acyl chains. Abbreviations: AA, amino acids; AC, acyl carnitines; BA, bile acids; CE, cholesteryl esters; Cer, ceramides; CMH, monohexosylceramides; LPE, lyso-PE; LPC, lyso-PC; LPI, lyso-PI; oxFA, oxidized fatty acids; MG, monoglycerides; NAE, N-acylethanolamines; SFA, saturated fatty acids; SM, sphingomyelins.

3.7 Aramchol improves steatosis and oxidative stress in MCD medium-exposed-hepatocytes and reduced *COL1A1* mRNA in human stellate cells

Regarding to lipid accumulation and cellular oxidative stress, we analyzed the effect of Aramchol in mouse hepatocytes cultured in MCD medium. Hepatocytes were cultured during 48 hours in serum-free MEM or MCD medium, with or without Aramchol (10 μM extracellular concentration). We found that hepatocytes incubated with MCD medium increase neutral lipid accumulation (as determined by BODIPY staining) and that this accumulation is prevented by Aramchol treatment (**Figure 32 A**). Total cell ROS in hepatocytes exposed to MCD medium increases 4.5-fold and this effect is significantly reduced under treatment with Aramchol (**Figure 32 B**).

On the other hand, we incubated LX-2 human stellate cells in absence or presence of Aramchol (10 μM extracellular concentration) for 24 hours to find that *COL1A1* mRNA content decreases while *PPARγ* transcript increases (**Figure 32 C**), which is interesting since PPARγ is a negative regulator of type I collagen expression[178]. Besides, *PPARγ* expression in hepatocytes positively correlates with fat accumulation induced by diabetes or obesity[179].

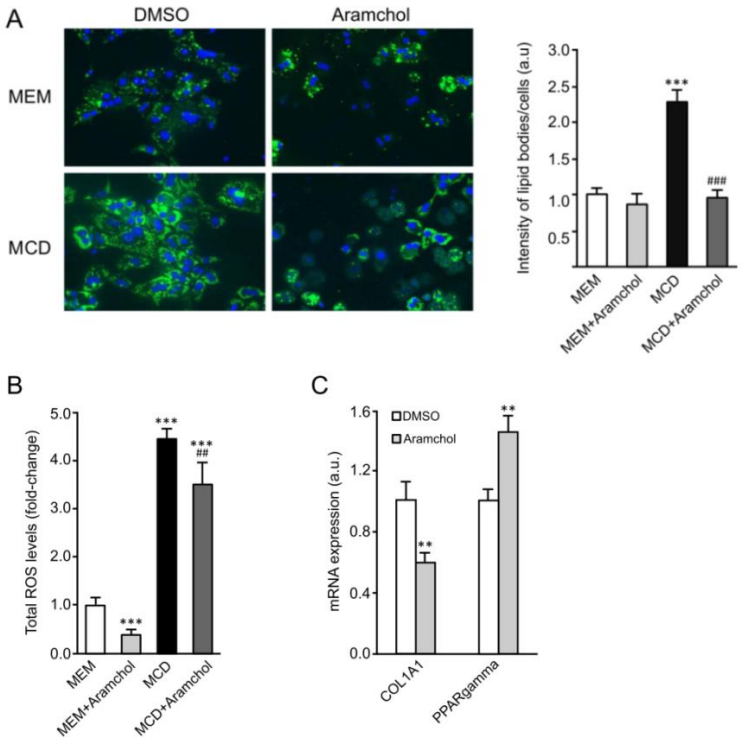


Figure 32. Aramchol attenuated lipid accumulation and ROS production in hepatocytes and collagen production in LX-2 cells. **(A)** BODIPY staining and quantification of the intensity of lipid bodies per cell. **(B)** CellROX Deep Green Reagent loading and total ROS levels quantification by FACS. (C) LX-2 mRNA expression of *COL1A1* and *PPARγ* was assessed by quantitative PCR. ** p < 0.01 versus control; ***p < 0.001 versus control; ##p < 0.01 versus MCD; ###p < 0.001 versus MCD. Data were represented as mean 6 SEM. Abbreviation: FACS, fluorescence-activated cell sorting.

3.8 Effect of Aramchol in the regulation of liver glucose metabolism

As mentioned in previous sections, our experiments show that Aramchol treatment causes down-regulation of SCD1 in 0.1 MCD diet-fed mice, leading to a decrease in FAs and TGs in the liver, reverting fibrosis, ameliorating inflammation, and improving FA β-oxidation by increasing the flux through transsulfuration pathway.

On the other hand, in a one-year study involving 247 NASH patients enrolled in the ARamchol for the REsolution of STeatohepatitis (ARREST) clinical trial (NCT 02279524) (**Figure 33**), Aramchol significantly reduced liver fat, improved histology and hepatic biochemistry. At week 52, placebo-treated patients exhibited an increase in glycated hemoglobin (HbA1c), while those treated with Aramchol (400 and 600 mg/day) showed a reduction in HbA1c (**Figure 34**). The differences from placebo were statistically significant, suggesting that Aramchol also targets glucose metabolism.

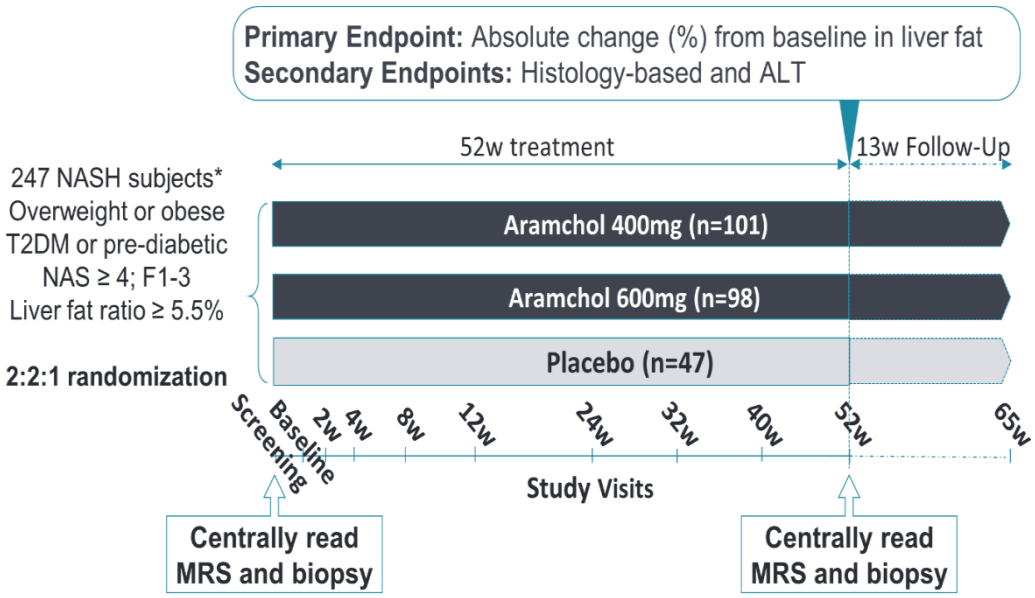


Figure 33. ARREST design. A phase IIb, double-blind, multinational, multicenter randomized, controlled clinical trial to evaluate the efficacy and safety of two Aramchol doses versus placebo in patients with NASH.

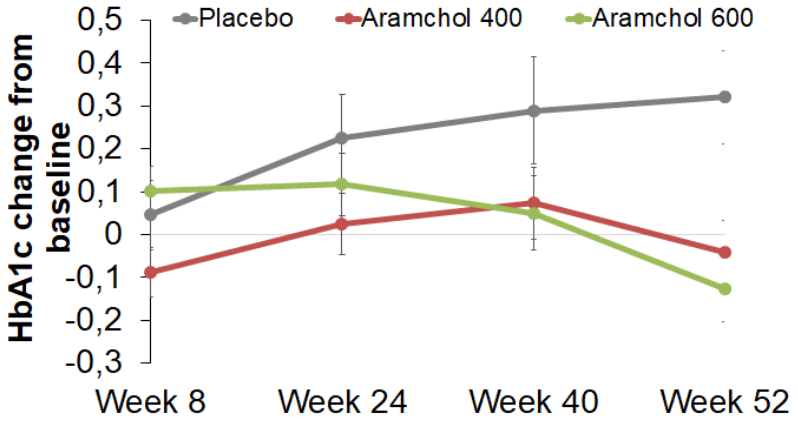


Figure 34. Analysis of HbA1c at week 52 of ARREST clinical trial revealed an increase in the placebo (pbo) group and a significant reduction in patients treated with Aramchol (400 or 600 mg/day). Aramchol 400mg *vs* pbo: p = 0.006. 600mg *vs* pbo: p = 0.001.

The promising results obtained with Aramchol in 0.1MCD mice model and LX-2 human stellate cells[136] combined with putative mechanism of Aramchol targeting SCD1 and glucose metabolism, made us focused in testing the effect of Aramchol in glucose metabolism *in vitro*. In addition, the metabolomic dataset obtained in the previous study conducted with 0.1MCD NASH preclinical model, was re-evaluated in terms of the glycolysis and gluconeogenesis-related metabolites.

3.8.1 Effect of Aramchol in catabolic and anabolic pathways

We treated primary WT hepatocytes cultured in MCD or control medium with Aramchol or DMSO (20 µM extracellular concentration) for 48h. Western blot analysis showed the pathways targeted by Aramchol (**Figure 35**).

As abovementioned, Aramchol targets SCD1[180], the enzyme that accomplish the first rate-limiting step in TGs synthesis, also improving the metabolic alterations that characterize NASH (accumulation of lipids, lipotoxicity and oxidative stress) and reverting fibrosis (collagen production). As expected, SCD1 level lower in control medium with Aramchol and the same behavior is observed in MCD medium with Aramchol (**Figure 35**). Besides, we found that Aramchol increases the levels of the phosphorylated (activated) form of 5' adenosine monophosphate-activated protein kinase (AMPK) in both culture media (**Figure 35**). AMPK is a key enzyme associated to energy homeostasis in the cell. Carnitine palmitoyltransferase I (CPT1A/B) is a mitochondrial enzyme that mediates the transport of long chain FAs across the mitochondrial membrane. Our results show that CPT1A/B protein levels increase in control

medium with Aramchol, while the increase in MCD medium with Aramchol is less pronounced (**Figure 35**).

We also analyzed the levels of acetyl-CoA carboxylase (ACC), an enzyme involved in the biosynthesis of FAs, in presence or absence of Aramchol and found that its inactive form (p-ACC) slightly but significantly increases in both culture mediums with Aramchol (**Figure 35**).

On the other hand, we also analyzed the levels of p70S6k and S6. p70S6k is a serine/threonine kinase and S6 ribosomal protein is its substrate target[181]. Phosphorylation of S6 ribosomal protein induces protein synthesis at the ribosome and importantly it is inhibitory for autophagy. Total S6 protein level decreases in control medium with Aramchol and the same behavior is observed in MCD medium with Aramchol (**Figure 35**). In addition, P-S6/S6K ratio level significantly decreases in hepatocytes in control medium with Aramchol and in hepatocytes in MCD medium with Aramchol (**Figure 36**). Also, P-p70S6K/p70S6K ratio level significantly decreases in hepatocytes cultured in control medium with Aramchol, otherwise in MCD medium with Aramchol the P-p70S6K/p70S6K ratio seems not to be altered (**Figure 35**), (**Figure 36**).

These results indicate stimulation of catabolic pathways and reduction of anabolism.

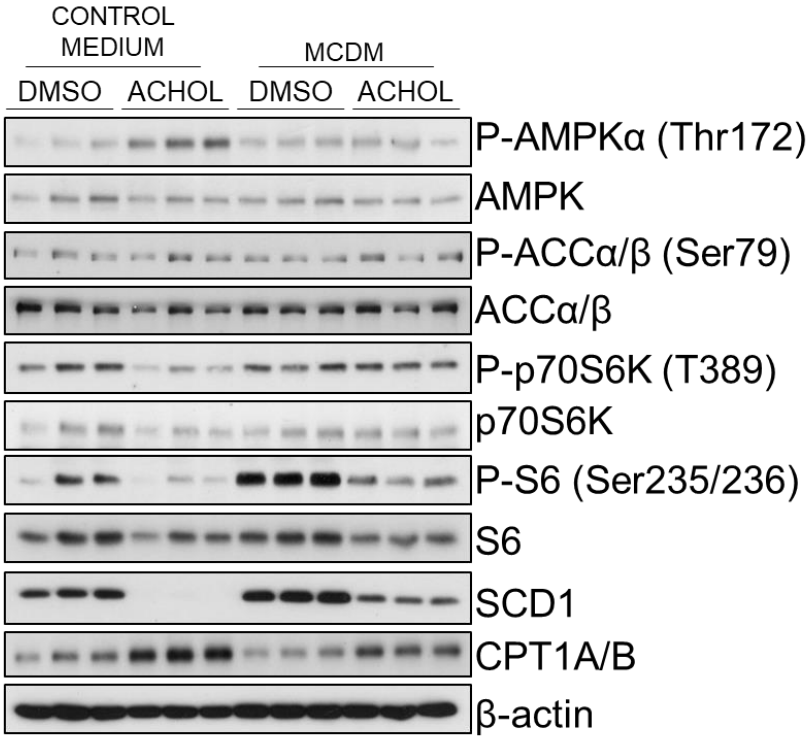


Figure 35. Aramchol targets metabolic pathways. Western blot analysis performed in hepatocytes treated 48h with Aramchol (20µM) cultured with control or MCD medium.

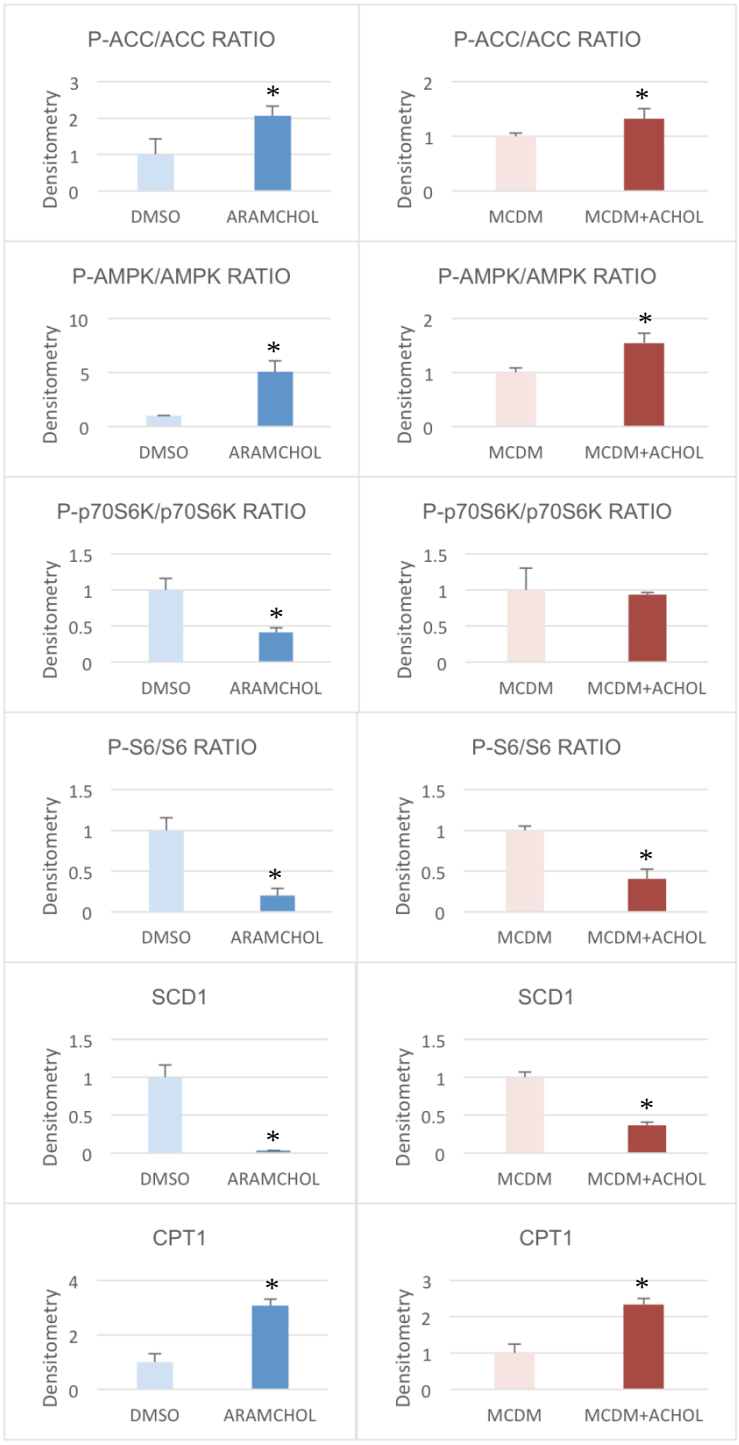


Figure 36. Densitometry analysis of the western blots performed in hepatocytes treated 48h with Aramchol (20µM) cultured with control or MCD medium. * $p < 0.05$.

3.8.2 Effect of Aramchol in key metabolic and signaling pathways

Protein extracts from murine hepatocytes isolated from three-month old male WT mice and cultured 48h in control medium (DMSO) or Aramchol (20 µM extracellular concentration), were analyzed by LC/MS. More than 3000 proteins were identified, from which 220 were differentially expressed between the two experimental groups. In the volcano shown in **Figure 37**, differentially expressed proteins according to the up or down regulation are classified by colors depending on their function. Differentially expressed selected proteins were analyzed with STRING software and classified in key biological functions showing a reduction of translation and fibrosis; and activation of lipid droplet clearance, fatty acid oxidation, oxidative phosphorylation, antioxidant response and the tricarboxylic acid cycle (TCA cycle).

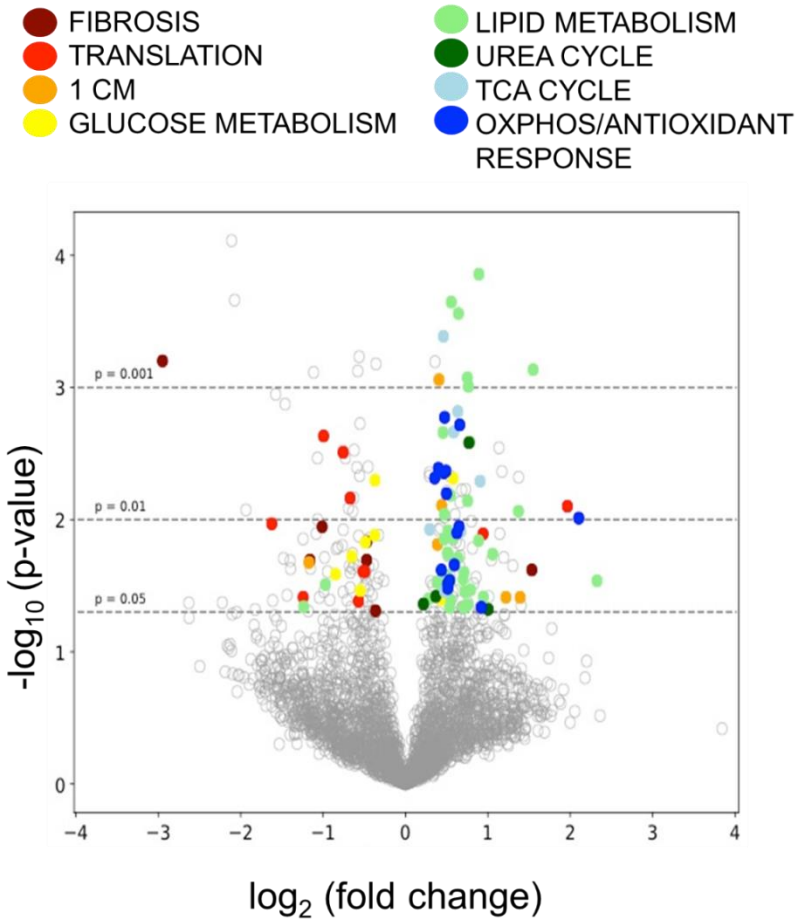


Figure 37. Volcano plot elaborated with the proteins identified by proteomics analysis and differentially expressed between hepatocytes treated 48h with Aramchol (20µM) or DMSO in control medium. Proteins are classified in different colors corresponding with their functions.

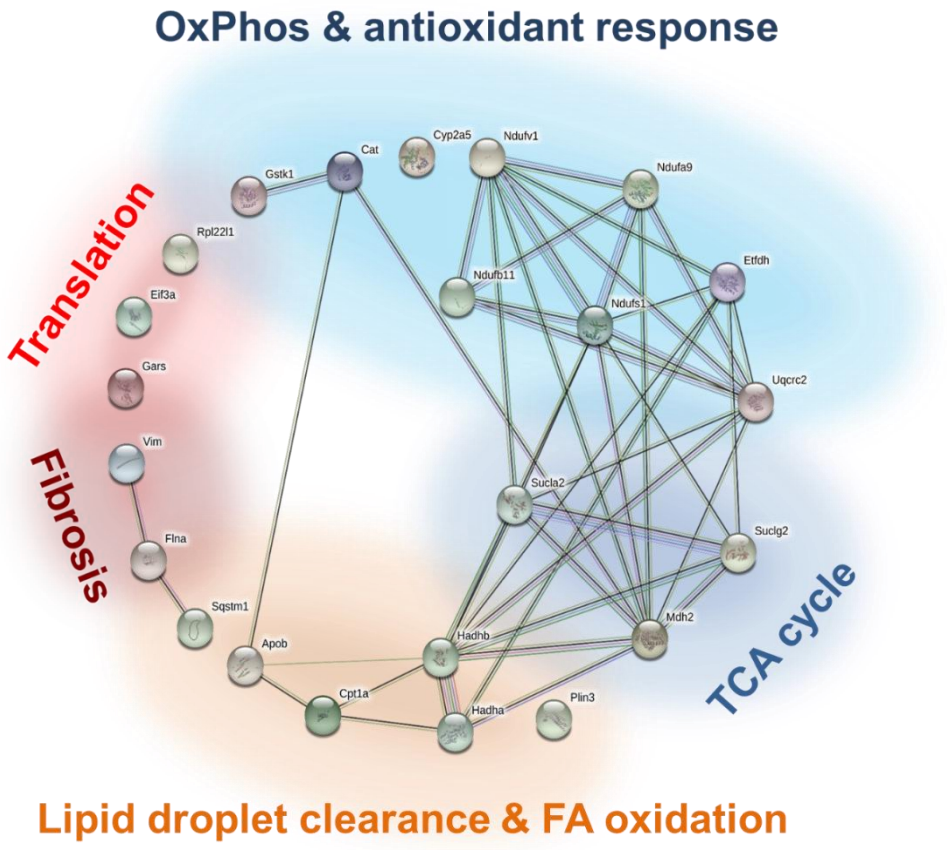


Figure 38. STRING analysis elaborated with selected proteins identified by proteomics analysis and differentially expressed between hepatocytes treated 48h with Aramchol (20µM) or DMSO in control medium. Proteins were classified corresponding with key biological processes.

3.8.3 Effect of Aramchol in TCA cycle activity

Based on the results of the proteomic analysis, we evaluated the effect of Aramchol in the TCA cycle by culturing primary mouse hepatocytes in control medium (DMSO) or medium with Aramchol (20 µM extracellular concentration) for 48h. Subsequently, we incubated these cells with ^{13}C uniformly labeled glucose for 2h or 4h. All samples were analyzed by LC/MS and extracted ion traces of labelled and unlabeled glucose, citric acid (not shown) and malic acid (**Figure 39**) were obtained for both experimental groups.

The results reflect an increase in the number of rounds that malate remains in the TCA cycle, indicating a reduction in *cataplerosis*[182], which may be linked to reduced gluconeogenesis (**Figure 39**).

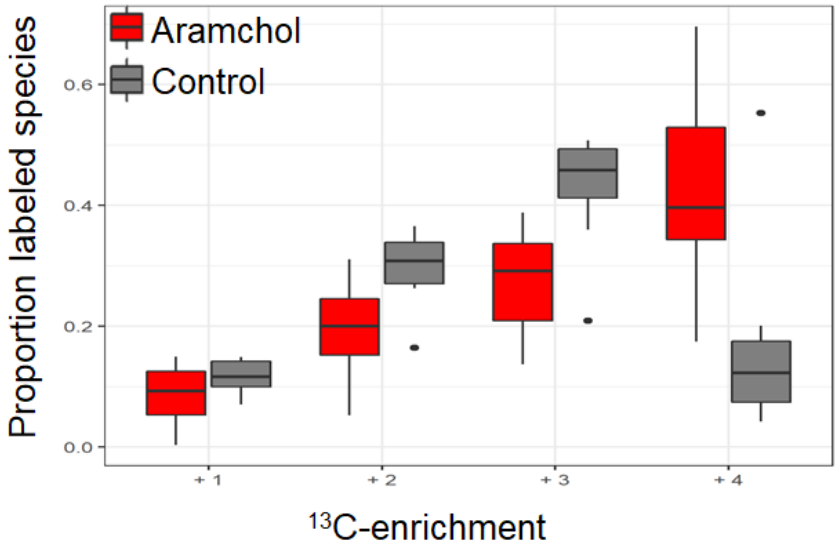


Figure 39. Aramchol reduces cataplerosis. Proportion of ^{13}C-labeled species of malate in hepatocytes treated or not with Aramchol (20µM) for 48h and incubated with fully labeled glucose for 2h.

3.8.4 Aramchol treatment improves glucose homeostasis in 0.1MCD diet fed-mice

As previously shown (**section 3.6**), 0.1 MCD mice liver metabolomics analysis was performed to investigate the potential benefit of Aramchol on fibrosis. The doses of Aramchol given to mice were 1mg/kg/day and 5mg/kg/day.

Therapeutic use of Aramchol in 0.1MCD mice produced a reduction in liver steatosis, fibrosis and inflammation.

The metabolic analysis of the 0.1MCD mice liver samples also revealed a reduction of liver glucose, glucose 6-phosphate (G6P), fructose 6-phosphate (F6P), UDP-glucose (UDPG), ribulose 5-phosphate (Ru5P) fructose 1, 6-bisphosphate (FBP) and pyruvate (PYR) as compared to mice fed with a normal diet. Under Aramchol treatment the levels of these metabolites increase (with a stronger effect when using 5mg/kg/day) although without getting the levels found in mice fed with normal diet. These results indicate that Aramchol administration improves glycolysis/gluconeogenesis in a NASH mice model, in a dose-dependent manner (**Figure 40**).

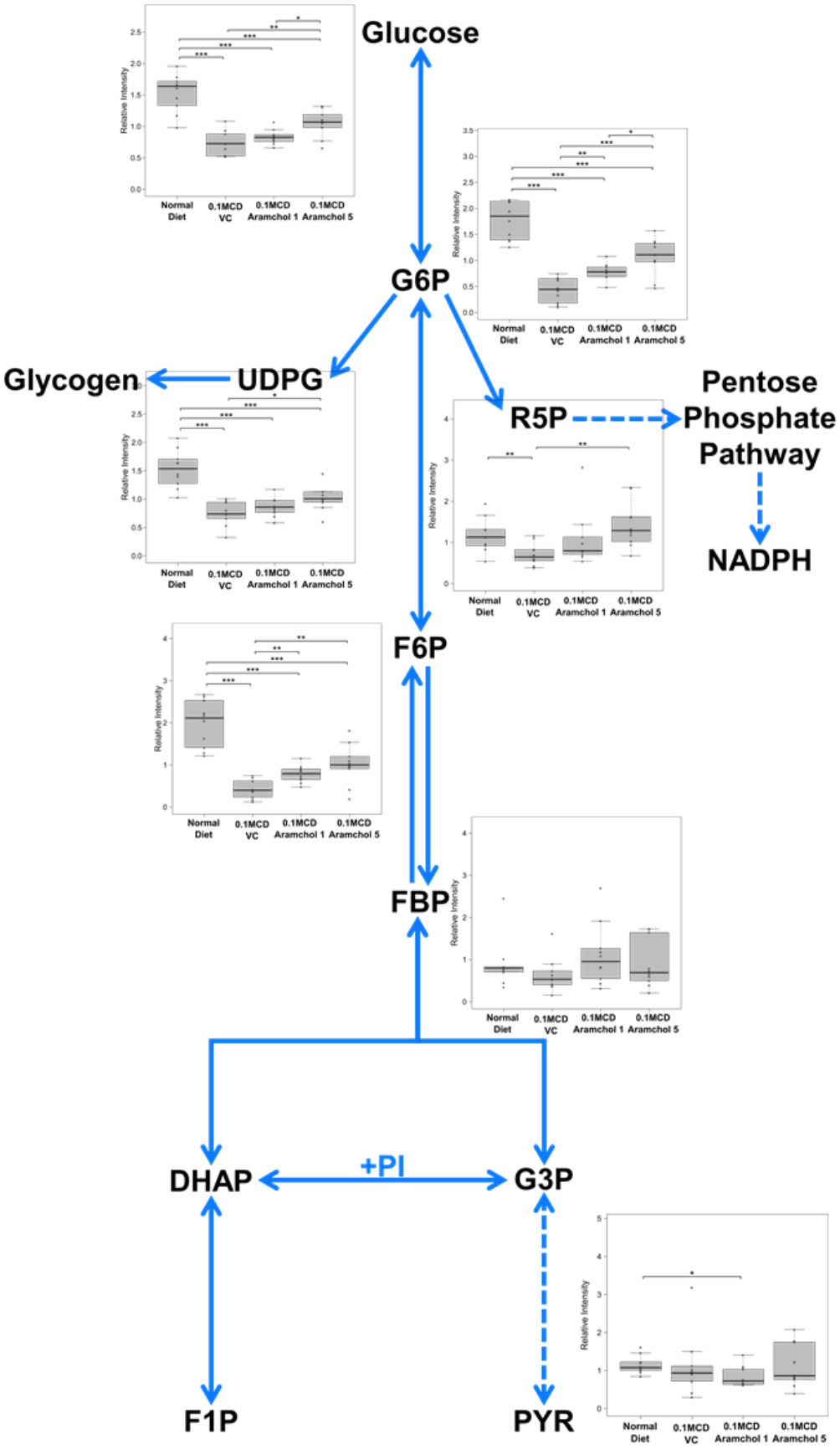


Figure 40. Aramchol improves liver glucose pathway metabolites in mice fed 0.1MCD diet. * $p < 0.05$, * $p < 0.01$. * $p < 0.001$.

3.9 Aramchol inhibits glycerophosphocholine (GPC) and lysophosphatidylcholine (LPC) formation in hepatocytes

We also wanted to check if Aramchol had an effect in VLDL synthesis and secretion so we treated WT hepatocytes cultured in MCD medium, with Aramchol (20 µM extracellular concentration) or vehicle during 48h. After this time, we replaced the media by ^{13}C-labeled-glucose at 0', 60', 120' and 240'. Aramchol blocks the formation of glycerophosphocholine (GPC) and lysophosphatidylcholine (LPC) in both experimental groups (**Figure 41**).

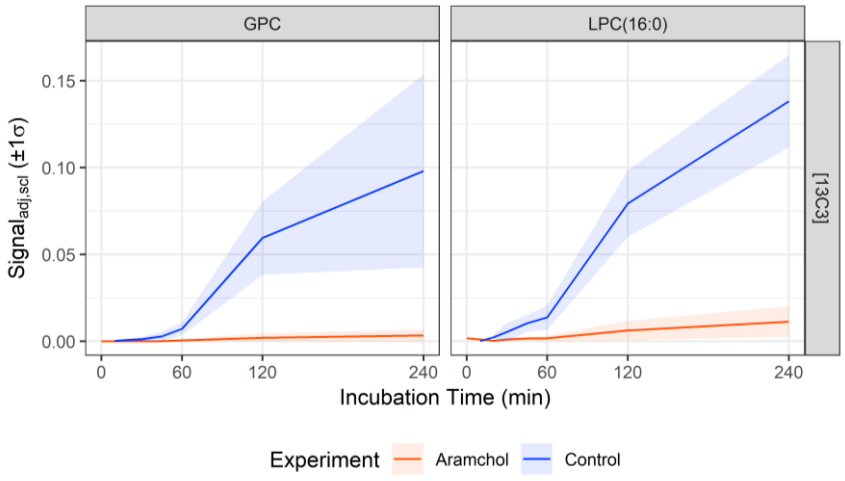


Figure 41. Aramchol 20 µM inhibits the formation of GPC and LPC in hepatocytes cultured in MCD medium. ^{13}C-labeled glucose was added at 0', 60', 120' and 240'.

3.10 Searching for biomarkers of MetS in human urine samples

MetS is a complex disorder caused by several altered factors, tightly related with NAFLD (which in some cases is considered its hepatic manifestation). There are several similar (yet not identical) definitions for MetS. These definitions are mostly based on symptomatology and not based in a metabolic fingerprint so we aimed to identify a biochemical fingerprint for this disease analyzing which are the metabolites present in urine samples from healthy people (here called "asymptomatic") and MetS patients. The final goal is to provide a rigorous and universal definition of this syndrome.

Using NMR spectroscopy, we measured urine samples from 511 people participating in the OBENUTIC study in the University of Valencia (Spain) and correlated the ^{1}H monodimensional NMR spectra with the metadata containing biochemical and anthropometric parameters of each patient. We classified patients, using the definition given by the WHO, in MetS, asymptomatic (no parameters altered) and intermediate (people with some altered parameters but not suffering MetS).

In a preliminary treatment of data, we found that from the OBENUTIC cohort only 43 patients had MetS so the cohort required to be supplemented in the number of MetS patients to balance the presence of each condition in the analysis. To do so, we selected 75 MetS patients from the cohort PreMedEus provided by the company Osarten and included the information from their urine samples measurements for this study. The first step was to check whether it is feasible to analyze simultaneously the OBENUTIC and PreMedEus cohorts. PCA of spectral bins shows no differences between samples coming from either project (**Figure 42**), validating this procedure.

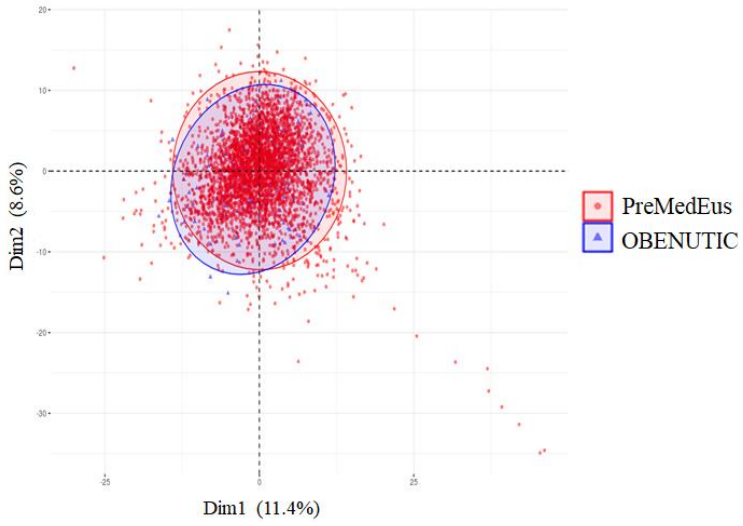


Figure 42. PCA of samples taken from OBENUTIC and PreMedEus participating patients

3.10.1 New classification of profiles based in 4 bit and WHO criteria

We performed a new classification of samples from people participating in those projects based in the definition given by the WHO and also based in 4 bits codified as binary (0 = no, 1 = yes) (**Figure 43, and Table 6**). We made this new classification renaming the variables to make both data sets match. The parameters for this new definition are high fasting glucose (> 110 mg/dL) as the mandatory requirement for MetS, obesity (BMI > 30 kg/m^2), dyslipidemia (if TGs > 150 mg/dL or cHDL < 34.75 mg/dL in men or < 38.61 mg/dL in women) and hypertension (systolic pressure ≥ 140 mmHg or diastolic pressure ≥ 90 mmHg or medication). We do not have data for microalbuminuria.

In bit 1 appears "1" if the patient has fasting glucose > 110 mg/dL and "0" if not (**Figure 43**). This is similar for all bits (bit 2 for obesity, bit 3 for dyslipidemia, bit 4 for hypertension) (**Figure 43**) and finally, all the combinations gave us 16 profiles, 4 of them classified as MetS according to the WHO (**Table 6**).

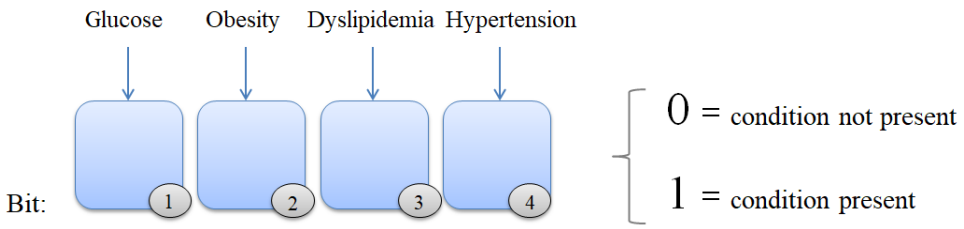


Figure 43. Bit position system schematic explanation. The new classification is based in 4 bits position and the MetS definition given by the WHO.

0000	Asymptomatic (n=215)	1001	diabetes + dyslipidemia (n=8)
0001	Hypertension (n=43)	1010	diabetes + hypertension (n=5)
0010	Dyslipidemia (n=23)	1100	diabetes + obesity (n=5)
0100	Obesity (n=45)	0111	obesity + dyslipidemia + hypertension (n=26)
1000	Diabetes (n=6)	**1011**	diabetes + dyslipidemia + hypertension (n=25)
0011	dyslipidemia + hypertension (n=7)	**1101**	diabetes + obesity + hypertension (n=31)
0101	obesity + hypertension (n=44)	**1110**	diabetes + obesity + dyslipidemia (n=21)
0110	obesity + dyslipidemia (n=9)	**1111**	diabetes + obesity + dyslipidemia + hypertension (n=25)

Table 6. Classification of samples based in 4 bit and WHO definition. From all combinations 16 different profiles were obtained including asymptomatic people (0000), 4 MetS profiles (in red) and 11 intermediate phenotypes. The sample size (n) of each condition is shown.

3.10.2 Identification of urine metabolomic biomarkers of the MetS by NMR

Starting from the spectra dataset of human urine samples, we first checked whether bins and metabolites provided different solutions to discriminate between MetS and the asymptomatic group so, using the R library "EnhancedVolcano" (version 1.2.0), we generated two volcano plots: one using the spectral bins (**Figure 44 A**) (untargeted analysis) and the other one using the metabolites identified (targeted analysis) (**Figure 44 B**).

In the untargeted analysis, we found that the bins differing most in MetS patients compared to asymptomatic people were: 3.41, 3.44, 3.47, 3.50, 3.53, 3.71, 3.83, 3.89, 4.64, 4.67, 5.24, 7.13, 9.17, 9.50 and 0.74 (**Figure 44 A**).

For the targeted analysis we only took the 23 metabolites that were quantified at least in 10% of the samples. These metabolites were: acetic acid, acetoacetic acid, acetone, alanine, allantoin, betaine, citric acid, creatine, creatinine, D-glucose, dimethylamine, formic acid, glycine, guanidinoacetic acid, hippuric acid, N,N-dimethylglycine, oxaloacetic acid, proline betaine, succinic acid, tartaric acid, taurine, trigonelline, valine. The metabolites differing significantly between both conditions were: D-glucose, betaine, alanine, dimethylglycine, tartaric acid (increased in MetS) and glycine (decreased in MetS) (**Figure 44 B**).

The p-values were calculated using the Wilcoxon or Mann-Whitney test[183], which are non-parametric test used as alternative to the parametric t-test. The latter assumes that the data has a normal distribution and it could not be the case. The p-values obtained were adjusted for multiple comparisons using the FDR method[184]. It was considered statistically significant the cases where the adjusted p-value was < 0.05. Regarding the fold change, a limit of abs($\log_2$fold change) ≥ 0.4 was established to generate the vertical discontinued lines (**Figure 44**).

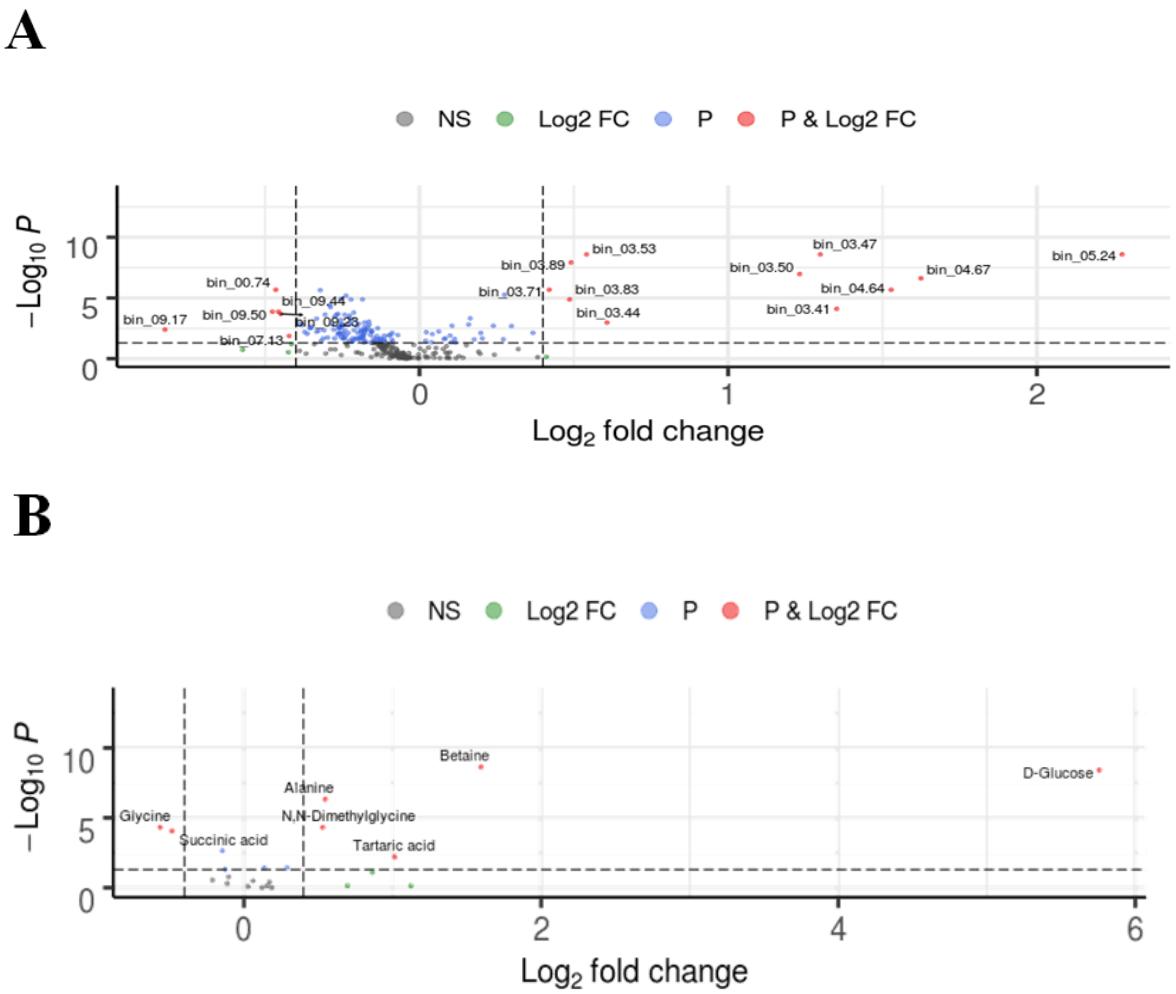


Figure 44. Volcano plot (univariate analysis) for MetS *vs* asymptomatic group**.** (A) Untargeted analysis showing the Log_2 fold change and p-values of the bins differing between both groups. (B) Targeted analysis showing the Log_2 fold change and p-values of the metabolites differing between two groups. p-values are calculated from Wilcoxon test and adjusted with FDR method. In red appear the metabolites that change significantly in MetS patients. Significance level = 0.05.

With libraries of R *ade4* (version 1.7-13) we generated a PCA analysis using the spectral bins (standard centering and scaling) obtained from measuring the human urine samples by NMR. Here we show the PCA by group (MetS *vs* asymptomatic) with the first and second component (**Figure 45**). It seems that the MetS group is positioned along the horizontal component while the asymptomatic group is positioned along the vertical component. In any case, this plot does not provide much information as we could not see a clear separation between both groups.

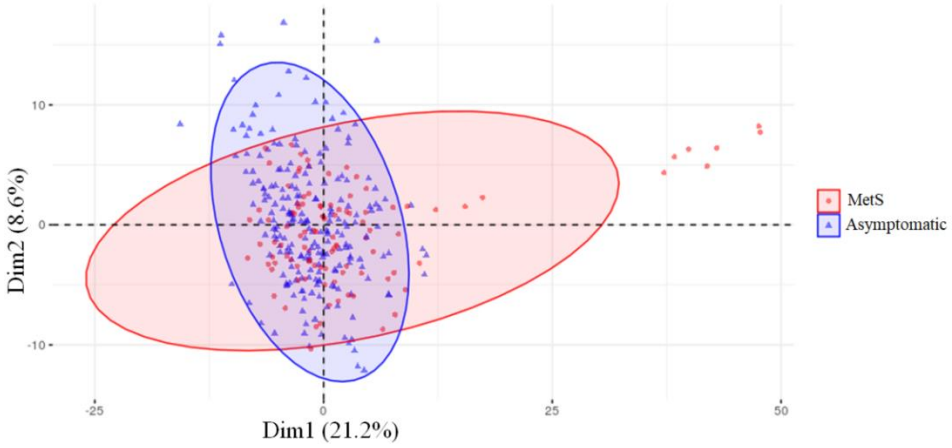


Figure 45. PCA by group (Mets *vs* asymptomatic, in red and blue respectively) showing the first and second component. No differences were found between these two groups with this unsupervised analysis.

As we did not obtain a separation between MetS patients and asymptomatic individuals by PCA (**Figure 45**), we performed an Orthogonal Partial Least-Squares Discriminative Analysis (OPLS-DA) (**Figure 46**) which, unlike PCA, is a supervised analysis that forces separation between experimental groups based in their intrinsic features. To generate this OPLS-DA we used the R *ropls* (version 1.16.0) package.

In this case we can see a clear separation between MetS patients and asymptomatic individuals (**Figure 46**). We obtained a R^2 = 0.651 and a Q^2 = 0.398. According to Chin (1998)[185] and Henseler et al. (2009)[186], the range of R^2 is in between 0 and 1, and the higher level, the higher predictive accuracy of the model. A value of 0.651 indicates a medium-high predictive accuracy.

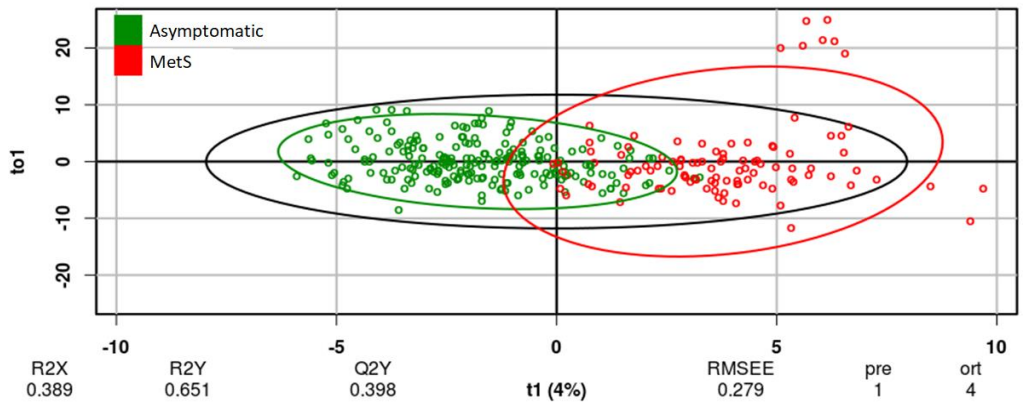


Figure 46. OPLS-DA of MetS patients and asymptomatic individuals. There is a clear separation between MetS and asymptomatic groups. The R^2 = 0.651 indicates a high predictive accuracy of the model and the Q^2 = 0.398 indicates if there is overfitting.

Q^2 is another important factor to consider, that helps to detect model overfitting[187]. A Q^2 value much lower than R^2 means that the model works properly with the training data but it will not work with new data, because the model is overfitted and it does not generalize well. In our case, Q^2 and R^2 are largely equivalent. These results indicated that this model explains the 65.1% of the variability and it is not too much overfitted.

To test whether the obtained results were not due to random, we made permutation of classes (**Figure 47**). The Q^2 values obtained in these permutations (black dots in bottom left) were always lower than the real Q^2 value obtained in the model when the correct classes were assigned properly (black dot which marks the value in Similarity = 1) (**Figure 47**). This was similar for the R^2 values generated in the permutations (grey dots in the left): they appear below the real R^2 value (grey point which marks the value in Similarity = 1, **Figure 47**). In summary, the R^2 and Q^2 values obtained in the OPLS-DA are not due to random but to the intrinsic metabolomic differences between MetS patients and asymptomatic individuals.

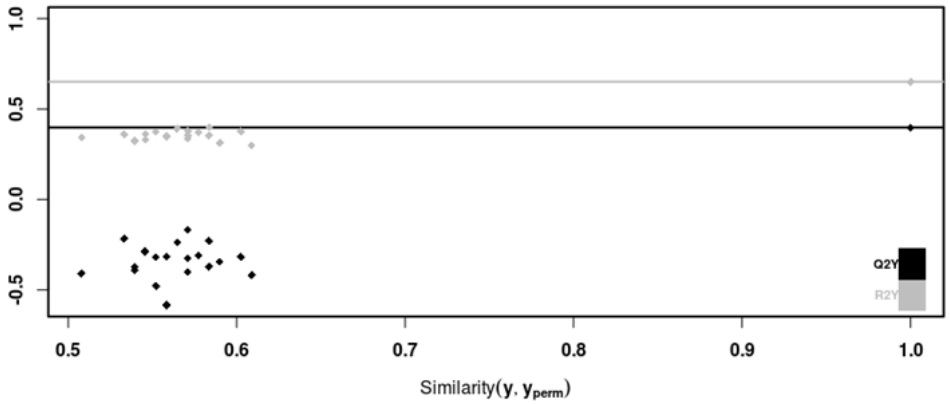


Figure 47. Permutation test for OPLS-DA. The simulations showed that the Q^2 values obtained are always below the value obtained in the model for the correct classes (black line).

In **Figure 48**, we can see the loading plot showing the bins that most contribute to the separation of MetS patients and asymptomatic people in the OPLS-DA (**Figure 46**).

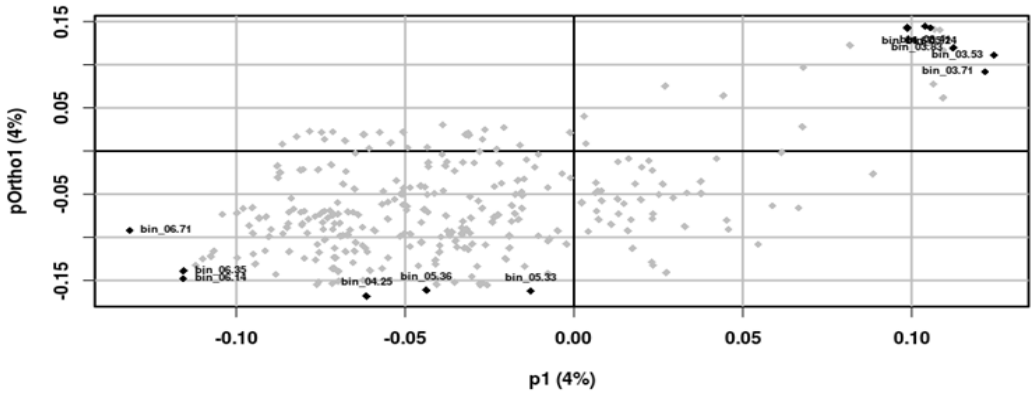


Figure 48. Loading plot representing the bins contributing to differences between MetS and asymptomatic spectra. The most contributing bins are in black: 3.41, 3.53, 3.71, 3.83, 4.67, 5.24, 6.35, 6.14, 4.25, 5.36 and 5.33.

Finally, we tested if our model was optimum using a ROC curve. We used the R package pROC (version 1.15.0) that allowed us to generate a ROC curve for this training model (**Figure 49**). The best point for the balance between sensitivity and specificity is maximum is 0.404.

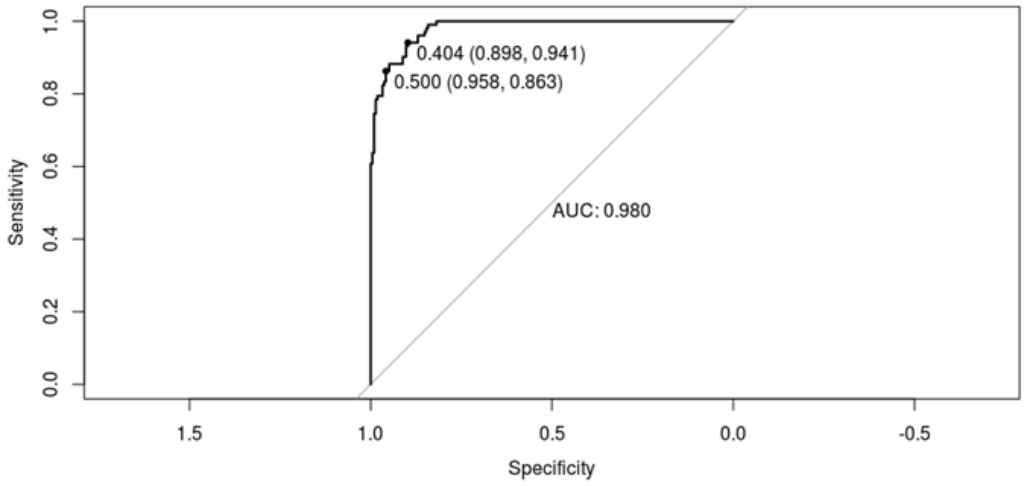


Figure 49. ROC curve for our training model. The graph shows the ROC curve and also the best sensitivity-specificity point (4.04).

The Area Under the Curve value of this ROC curve (AUC = 0.980) indicates that our training model is explicative and can distinguish between the different parts of the NMR spectrum that differs between the MetS and asymptomatic cohorts, bypassing the intermediate conditions. It also tells us that our model fits well with the data from the cohort of patients we have been working with.

Taken together, these results suggest that increasing the number of samples we may be able to create an algorithm to discriminate each condition of MetS when measuring urine samples by NMR.

CHAPTER 4

DISCUSSION

4 DISCUSSION

The current doctoral thesis is framed in the context of the development of precision medicine tools to improve diagnosis, treatment and prognosis of patients suffering NASH and MetS.

NAFLD is the most common liver disorder in western countries and it is a global health problem with a current prevalence of about 25%, rising during last decades. Tightly associated with NAFLD and NASH is the rise in obesity and MetS, which are also to be considered current worldwide epidemics[101]. NAFLD is associated with MetS risk factors (IR, obesity, hypertension and dyslipidemia) and comprise a spectrum of alterations ranging from steatosis to NASH, accompanied in some cases with fibrosis[2]. However, the precise molecular bases of these associations are poorly understood.

In this work we focused our attention in NASH and MetS, both, complex diseases with a multifactorial etiology that develop and progress due to the interaction of behavioral, cellular and metabolic factors. This is why NASH is a syndrome rather than a disease. Currently there are no treatments available for these disorders, with the only management for NASH consisting in healthy lifestyle habits acquisition. However, regarding NASH treatment, there are more than 40 molecules[47] in clinical trials, some of them in phase III stage like Aramchol®. The results obtained in the present work shed light about Aramchol effect and mechanism of action.

When investigating the molecular basis of NASH, we focused on the effect of SAMe depletion in *MAT1A*-KO mice. SAMe has a central role in many biological reactions acting as the main biological methyl donor in the cell. The liver is the principal responsible of its homeostasis and when there is an excess or defect in hepatic SAMe levels, liver disease triggers[70]. Our results validated previous studies showing that *MAT1A* deletion leads to a reduction in hepatic SAMe levels[76], as we found that hepatic DNA methylation levels were diminished by SAMe depletion in *MAT1A*-KO mice. This is remarkable, since liver DNA hypomethylation has been correlated with human NASH[163]. SAMe deficiency upon *MAT1A* deletion alters hepatic metabolism downstream to SAMe, 1CM, FAs metabolism, VLDL secretion and mitochondrial function[143]. In here, we found that SAMe deficiency set the stage to NASH development and progression in *MAT1A*-KO mice by inducing hypomethylation of DNA, altering a set of biological processes that leads to disease. In this context, we can conclude that *MAT1A* accomplishes a key role in maintaining cellular homeostasis in the liver.

Hepatic methylation levels recover in *MAT1A*-KO mice after eight-month SAMe administration. This is accompanied by an improvement in liver histology and function, probably by ameliorating VLDLs formation, export and GSH synthesis, which coincide with

studies showing that SAMe administration in animal models have hepatoprotective effects[56] and that patients with alcoholic liver cirrhosis treated with SAMe increased GSH levels and survival[59,63].

A low amount of methionine (0.1%) is sufficient to prevent body weight lose but without affecting the development of NASH. Our results using 0.1MCD diet models show reduced levels of SAMe accompanied by an accumulation of TGs in the liver, caused by the reduced secretion of hepatic TGs as VLDLs (due to the impaired synthesis of PC-PUFAs through the PEMT pathway), and also, due to the increased levels of proteins involved in lipid synthesis (AGPAT3) and uptake (CD36). In addition, the increase in FAs oxidized species like oxLA, a proven human NASH biomarker[174], goes together with a reduction in the GSH/GSSG content likely due to the active utilization of GSH since GSH consuming enzymes content are also increased. This agrees with previous studies showing that this mice model present elevated oxidative stress and has low plasma TGs levels due to impaired VLDLs export[172]. Taken together, we conclude that these are the causes contributing to hepatic FAs accumulation, lipotoxicity and increased FAs oxidation in these mice, as suggested by the abnormal protein content of CPT1A, ACC2 and UCP2.

We could also identify the existence of at least 2 metabolic subtypes of NAFLD patients: M^+ subtype and non-M^+ subtype. The M^+ subtype shows a serum metabolomic profile similar to that found in *MAT1A*-KO and 0.1MCD mice. The metabolomic profile of *MAT1A*-KO mice reflecting their hepatic metabolism supports the idea that patients with a M^+ phenotype have lower MAT1A activity.

M^+ subtype patients are not more prone to develop NASH than non-M^+ subtype patients as both phenotypes were found in simple steatosis and NASH in approximately the same proportions. This suggests that impaired SAMe occurs early in the development of NAFLD. Regarding the M^+ subtype, our results validate the existence of this phenotype based on its similarities with the serum metabolomic profile of *MAT1A*-KO and 0.1MCD mouse models. Both mice, although have some metabolic and molecular differences (as for example, the low levels of SCD1 in 0.1MCD compared to the increased content of this enzyme found in *MAT1A*-KO), they show many similarities and are characterized by impaired export of VLDLs, low content of GSH and reduced synthesis of SAMe, all these facts, leading to NASH development and progression. These findings agree with the observation that impaired synthesis of SAMe and lower rate of transmethylation is a frequent feature in human NAFLD[188,189]. So, we can conclude that SAMe deficiency must be a common feature in NASH patients and also a key factor in NASH development. Besides, since SAMe treatment improved NASH in *MAT1A*-KO mice, it seems

reasonable that NASH patients belonging to M^+ subtype may benefit from this treatment or others that will be develop in the future.

Our analysis also revealed a second metabolic signature based on metabolites of 1CM like methionine, PC, PE and their lyso forms, that allows the separation of NAFLD patients (from M^+ subtype) into simple steatosis and NASH. As some of these metabolites are biomarkers of impaired 1CM, these results support the idea that altered SAMe levels trigger the switch from benign steatosis to NASH, which also agrees with NASH patients in advanced stage usually showing reduced expression of *MAT1A*[188].

The non-M^+ subtype, has a different metabolomic profile. According to Morrison *et al*. 2018[48], this subtype shows a serum metabolome that resembles the profile of HFD-fed Ldlr-KO mice, a model for atherosclerosis that under this diet, develops NASH.

The importance of these findings is that, if we characterize subgroups of patients according to its metabolomic profile, we can identify the mechanisms underlying the disease in each case and, thus, find specific drugs and treatments to effectively suit each of these subtypes in an independent manner.

In vivo studies in this work, showed how 0.1MCD mice treated with Aramchol during two weeks (5mg/kg/day), improved features of NASH and recovered normal levels of some metabolites such as GSH, FAs, oxFAs and aminobutyric acid, which is a biomarker of transsulfuration pathway[190]. SCD1 expression and activity is lower after Aramchol treatment, in agreement with previous findings where Aramchol inhibits SCD1 activity promoting FA β-oxidation and decreases FA synthesis[49]. In agreement with a reduction in fibrosis, COL1A1 western blot analysis showed the reduced levels of this protein in the livers of 0.1MCD mice treated with Aramchol, supporting its role in improving NASH.

In connection with steatosis and oxidative stress induced by 0.1MCD diet, *in cellulo* experiments recapitulate the *in vivo* findings, since, Aramchol reduces cellular lipid accumulation and cellular ROS in hepatocytes incubated in MCD medium. Moreover, we found that Aramchol negatively regulates collagen production in LX-2 human stellate cells. At this point and taking these results together, we found that Aramchol performs its antisteatotic, anti-inflammatory and antifibrotic effects by targeting SCD1, COL1A and increasing the flux through transsulfuration pathway.

On the other hand, we identified a set of serum and lipid biomarkers associated with Aramchol treatment in 0.1MCD mice that are dose dependent (5 mg/kg/day). These biomarkers include lipids such as several MUFAs, PUFAs and lyso forms of PE, PC and PI. Lysophospholipid species are associated with lipotoxicity and NASH progression[191,192], and a reduction induced by

Aramchol enhances the idea of its therapeutic effect in NASH treatment. Importantly, the reduction of MUFAs agrees with the mechanism of action for Aramchol (targeting SCD1).

On the other hand, as ARREST study results suggest that Aramchol also targets the glucose metabolism, we performed a set of *in vitro* experiments to go further in the study of the mechanism of action of Aramchol. We found that this drug is a regulator of liver glucose metabolism and homeostasis as well.

First of all, western blot analysis revealed that Aramchol treatment stimulates catabolic pathways and reduces anabolism in hepatocytes. LC/MS results obtained from Aramchol treated hepatocytes, support and complement these results as they underline the reduction in the concentration of proteins associated to processes such as fibrosis, protein translation and glucose metabolism, while proteins involved in lipid metabolism, antioxidant response, oxidative phosphorylation and TCA cycle were increased. TCA cycle is the central pathway for the metabolism of amino acids, fatty acids and carbohydrates in order to generate cellular energy and, studies in humans and mice, indicate that alterations in its activity may play a central role in NAFLD pathogenesis[193]. We showed how Aramchol treatment reduced *cataplerosis* since malate remained more rounds in the cycle (as compared to untreated hepatocytes), which is linked to reduced gluconeogenesis, being part of Aramchol therapeutic effect. Accordingly, *in vivo* experiments showed that Aramchol treatment improves glucose homeostasis (in a dose dependent manner, being greater at 5mg/kg/day), because glucose pathway metabolites are reduced as reflected by the liver metabolic analysis of 0.1MCD mice. Finally, it is known that, one of the mechanisms causing NASH in 0.1MCD mice is the impaired VLDL-TG synthesis and secretion due to the reduction synthesis of PC-PUFAs through PEMT pathway[13], so we performed *in cellulo* experiments in hepatocytes incubated in MCD medium, where Aramchol treatment blocks the formation of GPC and LPC, which would improve VLDL synthesis and secretion due to the higher content of PC available for the synthesis of these molecules. This also, could be part of Aramchol therapeutic effect.

Thus, all these results point out that Aramchol administration, down-regulates SCD1 enzyme leading to a decrease in FAs and TGs content in the liver, reverting fibrosis, improving inflammation and enhancing FA β-oxidation by increasing flux through transsulfuration pathway. Moreover, Aramchol improves glycolysis/gluconeogenesis homeostasis in the liver by reducing *cataplerosis* in TCA cycle and consequently reduces gluconeogenesis. All these findings together with the fact that this drug also improves VLDL synthesis and export summarize the mechanism of action of Aramchol. The results obtained in this work together with previous promising results obtained in mouse models and in clinical trials, convert Aramchol into a potential and reliable drug for NASH treatment. However, it remains to be

determined whether patients that improve under Aramchol treatment belong to M^+ subtype, non-M^+ subtype or both.

Lastly, in this work we also aimed to identify urine biomarkers of MetS using NMR spectroscopy that would allow the diagnoses of this syndrome in a quantitative and precise manner. Urine is a very handy biofluid to work with since the preparation of urine samples for analysis is easy as compared to other biofluids and it can be collected non-invasively in large amounts and by repeated sampling[194]. Urine reflects many biochemical changes occurring in the body as is the primary way in which the body eliminates water-soluble waste compounds, so, it is a good tool for looking for new biomarkers using metabolomic approaches[195,196]. Urine metabolomics is a widely used tool for biomarker research in medicine and currently, there are studies characterizing the variations of urine metabolome under basal conditions to avoid confounding effects in cohort studies[194]. This is important to stablish which variations in the metabolome are due to normal physiological reasons, which are due to disease and then, differentiate rigorously cohorts of patients according to it.

Here, we carried out a preliminary study (that is part of a much more ambitious project including 10.000 urine and 10.000 serum samples) to set up an analytical workflow that allowed us to discriminate asymptomatic people and MetS patients according to their urine metabolomic profile, although intermediate phenotypes remained undetectable.

The univariate analysis of spectral bins, showed a signal present around 3 ppm that, according to information available in HMDB, is very likely D-glucose (although it will be confirmed in a near future by spiking analysis, a common strategy used to identify metabolites in NMR-based metabolomics[197]). This suggests that D-glucose is higher in MetS patients, and this fact, highlights the role of glucose and diabetes in MetS. This result is supported by the volcano plot of metabolites, where we can see clearly that glucose is higher in MetS patients, and validates the WHO definition for this disorder. Moreover, the volcano plot for metabolites revealed augmented levels of betaine and N,N-dimethylglycine and lower levels of glycine in MetS patients. These are metabolites of 1CM and, although it should be further analyzed, it suggest that deranged 1CM could be one of the causes linking MetS and NAFLD, as impaired 1CM is also found in this disease[198]. This finding would support the hypothesis of that NAFLD is part of the spectrum of MetS, an idea sustained by many researchers[199,200].

OPLS-DA clearly discriminates between MetS patients and asymptomatic people bypassing intermediate conditions. The R^2 value, tells us the percentage of variability that our model can explain with the training data and Q^2 explains how our model generalize with data different to training dataset. If Q^2 were much lower than R^2, it would mean that the model is overfitted. In our case, $Q^2 = 0.398$ is not much lower than $R^2 = 0.651$, which means that our model is not

overfitted but however, it could generalize better. On the other hand, the loading plot showing which bins contribute to the separation of these two conditions, supports the results obtained in the volcano plot, as bins around 3 ppm are probably D-glucose, which is a key factor in the development of MetS.

A ROC curve is a fundamental evaluation tool in clinical medicine and research that gives information about the accuracy of the classification of patients in correct groups[201] when using diagnostic test with binary outcome (positive or negative test result)[202]. Sensitivity and specificity of the model is plotted and this curve discriminate the true state of subjects finding also an optimal cut-point value. The AUC value is a measure of accuracy that gives the probability of correctly classify information[203].We obtained an AUC = 0.980 which indicates that our training model classify patients accurately into MetS and asymptomatic group.

The aim of a model is to use it for: explication and/or prediction. In this first stage of our study, the importance of our model is its explicative capacity. In other words, in this point, our model can see which metabolic features are important to distinguish between MetS and asymptomatic people. Once the metabolic bases of MetS are better understood, we could go one step forward and focus on the predictive capacity of the model, to be able of blindly detect MetS from urine samples. So, the future work would be: 1) to validate the model in a new cohort of patients to see if AUC value changes and thus, ultimately to obtain a realistic estimation of the model performance, and 2) to increase the number of samples analyzed to see if intermediate phenotypes can be discriminated.

Taking these results together, we implemented an analytical method that allowed the discrimination between MetS patients and asymptomatic group according to their urine metabolomic fingerprint. Although we were looking for a method that allowed us to differentiate also the intermediate phenotypes, with this sample size it was not possible. These results suggest that, analyzing a bigger cohort, we could develop an algorithm to diagnose, in a continuous way, all the intermediate conditions between asymptomatic people and MetS patients.

In summary, our results unveiled the existence of 2 subtypes of NAFLD patients and highlight the importance of *MAT1A* and SAMe in the pathogenesis of this syndrome. In the context of NASH, we uncovered the mechanism of action of Aramchol and showed that, this drug, improved liver histology and function in animal models and improved glucose homeostasis in 0.1MCD mice and human patients. Its efficacy in murine models coupled to its efficacy in previous phases of clinical trial makes Aramchol a great candidate for NASH treatment in humans. On the other hand, we implemented an analytical workflow that, strongly supported by bioinformatics, allowed the discrimination of MetS patients and asymptomatic people through

the analysis of human urine samples by NMR spectroscopy. In this regard, future research will be directed toward the implementation of an algorithm that may allow the diagnosis of MetS using this procedure. Moreover, we will intend to provide a new definition of MetS based in the urine metabolic fingerprint. Finally, our results clearly demonstrated the capability that LC/MS and NMR spectroscopy have as metabolomic tools, to characterize biofluids and categorize patients into different subgroups, which is useful in terms of diagnosis and treatment.

CONCLUSIONS

5 CONCLUSIONS

- MAT1A protein is an integrator of cellular metabolic homeostasis and its deletion affects hepatic metabolism upstream and downstream of SAMe altering processes like biosynthesis of lipids, proteins and mitochondrial function.
- In *MAT1A*-KO mice, SAMe depletion causes DNA hypomethylation which agrees with that found in advanced stage NAFLD patients.
- Our results reveal the existence of 2 human NAFLD metabolic phenotypes: M^+ and non-M^+. M^+ subtype shows a serum metabolic profile similar to that found in *MATA1A*-KO and 0.1MCD mice.
- SAMe deficiency may be a common feature in NASH patients and an important factor in NASH development.
- SAMe administration to *MAT1A*-KO mice improves NASH but if NAFLD patients with an M^+ subtype could benefit from SAMe treatment should be clinically proven.
- 0.1MCD mice treated with Aramchol reduced steatohepatitis and fibrosis which supports the potential use of Aramchol in NASH treatment for the M^+ subtype of patients. Moreover, a metabolic serum profile for Aramchol treatment was identified in 0.1MCD mice receiving this drug.
- Aramchol decreases SCD1 content, increases the flux through transulfuration pathway improving the antioxidant response, activates TCA cycle and stimulates catabolic pathways while reduces anabolism.
- Aramchol also activates LD clearance and improves hepatic glucose metabolism in a dose dependent manner and reduces cataplerosis and consequently gluconeogenesis.
- Our results unveiled the potential capacity of Aramchol to be a NASH treatment in humans for suppressing steatohepatitis development and progression.
- We implemented a procedure to analyze human urine samples by NMR spectroscopy that allows discriminating MetS patients from an asymptomatic cohort, although intermediate phenotypes for this condition remain undetectable, probably because a higher number of samples are needed.
- Our results support the concept that metabolomics is a useful tool to detect biomarkers associated to diseases, detect metabolites associated to a specific treatment and definitely has a huge potential as method to develop non-invasive diagnostic and monitoring tools.

6 SUMMARY

Non-alcoholic fatty liver disease (NAFLD) is the leading cause of chronic liver disease in western countries, and currently is one of the most common causes of liver disease. NAFLD global prevalence is about 25% and it increases in patients with T2D and obesity, which are other two common epidemics spreading in our societies tightly related with NAFLD. NAFLD is the term used to refer to several related disorders that include steatosis, non-alcoholic steatohepatitis (NASH), fibrosis, cirrhosis and HCC. Steatosis is the earliest stage of the disease and it happens when more of the 5% of the liver is fat. Hepatic steatosis starts when processes like DNL and the uptake of FAs from circulation saturate the capacity of the liver to eliminate them by forming and exporting VLDL and to oxidize them in the mitochondria. NASH is the advanced form of NAFLD and it is characterized by the presence of steatosis together with inflammation, ballooning with or without liver fibrosis. The mechanism for the progression of NASH involves multiple parallel hits and processes including inflammation, mitochondrial dysfunction and oxidative stress, which make the liver prone to suffer further damage. Thus, having a better understanding of the consequences of these processes and handling the functioning of these pathways may be and effective approach to prevent the development and progression of NASH.

So, NAFLD is a consequence of defects in diverse metabolic pathways that lead to hepatic accumulation of TGs and the features of these aberrations may determine if NAFLD progresses or not to NASH. Moreover, as different causes can lead towards NASH development, we hypothesized that different NAFLD subtypes should exists and that each subtype would be characterized by a specific liver and serum metabolome.

Regarding NAFLD diagnosis, the diagnosis is heterogeneous and relies mostly in blood test of liver enzymes and liver biopsy been this one, the most reliable way to diagnose it. However liver biopsy presents some disadvantages because it is a very invasive technique that may compromise the welfare of patients and also the fact that, the final result depends on the liver portion taken in the biopsy, which could no reflect the status of the entire organ. That is why new diagnostic tools should be developed.

With regard to NASH current therapies, there is no treatment approved by the Food and Drug Administration or the European Medicine Agency. The current recommendations are based in weight loss and a healthy lifestyle because they are the only known factors that can stop the progression of the disease. However, although there is no approved therapy for NASH, currently there are over 40 molecules in clinical development, some of them in phase III of study like Aramchol.

SUMMARY

Aramchol is a synthetic lipid molecule resulting from conjugate a bile acid and a saturated fatty acid and in this work we aimed to investigate its mechanism of action as well as its effect in NASH.

Here we worked with two different NASH animal models, a genetic NASH model (*MAT1A*-KO mice) and a dietary NASH model (0.1MCD diet-fed mice). In the case of *MAT1A*-KO mice, the lack of *Mat1a* caused a deficiency in SAMe and metabolites upstream and downstream this point were dysregulated, affecting processes such as DNA and phospholipid methylation, biosynthesis of lipids, proteins, mitochondrial polarization and function, as well as polyamine and GSH synthesis, which is the main cellular antioxidant. On the other hand, we worked with 0.1MCD diet-fed mice. These mice were fed a diet lacking choline and containing 0.1% of methionine avoiding the weight loss that occurs with the canonical MCD diet without affecting NASH development. As well as *MAT1A*-KO mice, this model is characterized by reduced synthesis of SAMe, low levels of GSH content and impaired formation and export of VLDL. Both mice models spontaneously develop NASH, although 0.1MCD mice develop the disease much sooner than *MAT1A*-KO mice.

In these work we examined the liver and serum metabolic and proteomic features of these two murine models of NASH. First of all, we compared the serum and liver metabolic profiles of *MAT1A*-KO mice and found that serum metabolome reflected their hepatic metabolism. Afterwards, we selected the top 50 metabolites changing more significantly between *MAT1A*-KO and WT mice and made a list. Then, we compared the serum metabolome of a cohort of 535 NAFLD patients (353 with simple steatosis and 182 with NASH) with the serum metabolome of *MAT1A*-KO mice and identified a serum metabolic signature associated with these mice that also was present in 49% of NAFLD patients. We validate these results by random partition and comparison of estimation and validation cohorts obtained from the 535 samples we worked with, but also, we validated these results by carrying out the same procedure explained above but this time by comparing the serum metabolome of 0.1MCD diet-fed mice with the serum metabolome of the 535 NAFLD patients. Our results revealed the existence of 2 NAFLD human subtypes: M^+ subtype and non-M^+ subtype.

Patients with an M^+ subtype have a serum metabolic profile similar to that found in *MAT1A*-KO mice and 0.1MCD mice, which suggest that these patients may benefit from the treatments that ameliorate NASH features in these mice models, such as SAMe and Aramchol (or other treatments that may develop in the future). In relation to this, we found that *MAT1A*-KO mice treated with SAMe, ameliorated liver function and histology and recovered the methylation of specific DNA regions that were hypomethylated in mice given vehicle alone. If humans with an M^+ subtype could benefit from SAMe treatment should be determine.

On the other hand, we tested the effect of Aramchol in 0.1MCD mice, observing an improvement in liver function and histology in mice treated with this molecule. Serum and liver biomarkers associated with Aramchol treatment in 0.1MCD mice were also identified by LC/MS.

Aramchol is in phase III of clinical trial and is a very promising molecule with a big potential use for NASH treatment. In phase IIb of study Aramchol showed to improve fibrosis and hepatic biochemistry with no signal of toxicity in humans. It also reduced the glycated hemoglobin (HbA1C) content, observed in patients receiving placebo, in a dose dependent manner, suggesting that glucose metabolism in targeted by Aramchol. Analyzing metabolites such as glucose, glucose-6-P and fructose-6-P, present in the livers of mice fed a control diet or 0.1MCD diet treated with Aramchol or vehicle, we confirmed that Aramchol treatment improved hepatic glucose homeostasis in 0.1MCD mice.

Aramchol was known to target and reduce the activity of SCD1, which is a key enzyme that catalyzes the first reaction that leads to TG synthesis. In this work, we confirmed this in many ways, and also by western blot analysis of protein extracts obtained from murine primary hepatocytes incubated in control and MCD medium treated with Aramchol or vehicle (DMSO). Thus, we found that this molecule reduces the protein content of SCD1, but also, affects other proteins including: 5' adenosine monophosphate-activated protein kinase (AMPK), carnitine palmitoyltransferase I (CPT1A/B), acetyl-CoA carboxylase (ACC), P70-S6 kinase and S6 (a ribosomal protein component of the 40S ribosomal subunit). Aramchol treatment in hepatocytes increased the levels of CPT1A/B, the activated form of AMPK (p-AMPK) and the inactivated form of ACC (p-ACC). Conversely, Aramchol decreased total protein content of S6, and p-P70-S6K/ P70-S6K ratio levels. These suggest that Aramchol would stimulate catabolic pathways reducing anabolic processes.

Again, in primary hepatocytes cultured with control or MCD media and treated with Aramchol or vehicle, more than 3000 proteins were quantified and we selected those that were differentially expressed between both conditions. These results showed that Aramchol affects key biological functions reducing translation and fibrosis while activating FA β-oxidation, LD clearance, oxidative phosphorylation, antioxidant response and the TCA cycle.

We also evaluated specifically the effect of Aramchol in the TCA cycle by fluxomics analysis using uniformly ^{13}C-labeled glucose. We found that under Aramchol treatment, the ^{13}C-labeled species of malate present in hepatocytes treated with Aramchol were higher compared to controls. The increase in the number of rounds malate remains in TCA cycle indicates a reduction in cataplerosis and, consequently, a reduction in gluconeogenesis. Moreover, we found that Aramchol blocked the formation of LPC and GPC which suggest an increase in the

content of PC favoring hepatic VLDL-TG synthesis and secretion. All the aforementioned results provide a better understanding of Aramchol`s therapeutic effect.

On the other hand, in this work we also searched for biomarkers of the MetS which is a complex disorder defined by a cluster of related risk factors that increase the probability of suffering other related diseases like T2D and cardiovascular disease. MetS is tightly associated with NAFLD and has elevated socioeconomic cost, being considered a worldwide epidemic together with obesity. There is not a consensus to define MetS and many organizations give similar but different definitions for this condition. We validated the definition given by the WHO that considers that a patient has MetS when the levels of fasting glucose are above 110 mg/dl as mandatory requirement, together with the alteration of 2 or more of the following parameters: obesity (body mass index BMI > 30 kg/m²), dyslipidemia (TGs >> 150 mg/dL or low cHDL < 34,79 mg/dL in men and < 38,66 mg/dL in women), high arterial blood pressure (≥ 140/90 mmHg or medical treatment) and microalbuminuria (albumin excretion of 20µg/min).

All the current definitions for MetS are based in a conjunction of biochemical and anthropometric data. In this work, we aimed to find a metabolic fingerprint for MetS to provide a universal definition for this disorder based in quantitative metabolic patterns.

To do so, we worked with 575 human urine samples coming from two different studies and belonging to the general population from Valencia and Euskadi. Together with urine samples, we had correlated metadata containing biochemical information for each person in the cohorts. So, we focused in the development of a method that, strongly supported by bioinformatics tools, allowed us to differentiate metabolically between asymptomatic individuals, MetS patients and the intermediate phenotypes between both conditions using the metabolites present in the urine of these people.

We created a new classification based in 4 bits codified as binary (0 = condition not present, 1 = condition present) and considering the MetS factors given by the definition of the WHO with the exception of microalbuminuria. We obtained 16 profiles in total, 4 of them for MetS patients (1011, 1101, 1110 and 1111), 1 for asymptomatic individuals (0000) and the rest for the intermediate phenotypes.

On the other hand, urine samples were measured in a 600 MHz AVANCE IIIHD (IVDr) Bruker Spectrometer and ^{1}H monodimensional spectra were obtained for each sample. Every sprectrum was fragmented in bins of 0.03 ppm and normalized by the total area.

With these data, we performed univariate analysis (targeted and untargeted) and found which metabolites and bins differed most between MetS and asymptomatic individuals, highlighting the role of the glucose in this syndrome.

Besides, multivariate analysis was carried out and OPLS-DA revealed that there is a clear separation between asymptomatic people and MetS patients although with this number of samples we could not differentiate intermediate phenotypes.

On the other hand, we tested the predictive accuracy of our model and checked if there was overfitting. Our present results show that our training model is explicative, can distinguish between different parts of the NMR spectra and can separate the two extreme conditions based in the metabolic fingerprint obtained by measuring human urine samples by NMR.

Although preliminary, these results are promising and suggest that increasing the number of samples, it would be possible to differentiate the intermediate phenotypes and create a diagnosis algorithm that would allow diagnosing MetS in a continuous way.

To develop a diagnosis strategy using metabolomics, we have been working with LC/MS and NMR to perform metabolomic analysis of big cohorts of patients to first, characterize their metabolic profile in different biological samples with a specific metabolic fingerprint and secondly, carry out the statistical analysis of the data generated to classify people into different profiles.

Due to the multifactorial etiology of NAFLD and MetS, metabolomics is a very useful tool to study and characterized big populations of patients suffering these disorders. Metabolomics allow us to obtain the all the metabolites present in a biological sample, so, analyzing samples coming from people suffering metabolic disease, it is possible to identify specific biomarkers for these diseases. Besides, it is possible to check the effect of specific treatments and also the progression of disease during lifetime.

Metabolomics and precision medicine are two synergistic approaches because through metabolomics we can identify different subtypes of patients and by the understanding the mechanisms inherent to their disease, we could find and provide personalized and effective treatments.

7 RESUMEN

La enfermedad del hígado graso no alcohólica (EHGNA) es una de las mayores causas de enfermedad hepática crónica en los países desarrollados y actualmente, una de las principales causas de mortalidad. La prevalencia a nivel mundial de esta enfermedad está alrededor de un 25% aunque aumenta en pacientes con diabetes tipo 2 y obesidad, dos enfermedades estrechamente relacionadas con EHGNA que son consideradas a su vez epidemias de nuestra sociedad moderna. La enfermedad del hígado graso no alcohólica es un término que engloba diferentes desordenes hepáticos, entre los que se encuentran la esteatosis hepática, la esteatohepatitis no alcohólica (EHNA), la fibrosis, la cirrosis y el carcinoma hepatocelular (CHC). Se considera que un hígado tiene esteatosis cuando más del 5% del órgano es grasa. La esteatosis hepática es la fase más temprana de EHGNA y comienza cuando los procesos de síntesis y captación de ácidos grasos por parte del hígado superan la capacidad de éste para oxidarlos en la mitocondria o enviarlos a la circulación en forma de lipoproteínas de muy baja densidad (VLDLs). La EHNA, es la fase avanzada de la enfermedad del hígado graso no alcohólica y se caracteriza, además de por la presencia de esteatosis, por la presencia de inflamación con o sin fibrosis. Los mecanismos por los cuales la EHNA progresa incluyen múltiples procesos entre los que se encuentran la inflamación, la disfunción mitocondrial y el estrés oxidativo, procesos que hacen que el hígado sea más propenso a sufrir un daño hepático mayor. Por lo tanto, entender mejor estos mecanismos y las consecuencias de estos procesos así como manipular su funcionamiento, puede ser una forma efectiva de prevenir el desarrollo y la progresión de esta enfermedad.

EHGNA surge como consecuencia de fallos en diferentes rutas metabólicas que conllevan a la acumulación de triglicéridos en el hígado. Las características de estos fallos podrían determinar si EHGNA progresa a EHNA o no. Además, como diferentes causas puede conducir al desarrollo de EHNA, nuestra hipótesis es que existen diferentes subtipos de pacientes de EHGNA y que cada subtipo debe estar caracterizado por un perfil metabólico concreto.

Por otra parte, el diagnóstico de EHGNA es heterogéneo y en la actualidad se lleva a cabo midiendo las enzimas hepáticas en sangre y mediante biopsia, que actualmente es el método más fidedigno. La biopsia presenta inconvenientes ya que es un método invasivo que compromete el bienestar del paciente y además, el resultado final depende de la porción de hígado tomada, que puede no reflejar el estado del hígado en su totalidad. Por ello, el desarrollo de nuevas técnicas de diagnóstico no invasivas es necesario.

En cuanto a su tratamiento, actualmente no existe ningún medicamento aprobado por la agencia estadounidense de administración de medicamentos y alimentos (FDA) ni por la agencia

europea de medicamentos (EMA). A los pacientes se les recomienda que bajen de peso y lleven un estilo de vida saludable, ya son los únicos factores que detienen la progresión de EHNA en la actualidad. A pesar de que aún no hay un tratamiento aprobado hay más de 40 moléculas que están en ensayo clínico en la actualidad, algunas de ellas en fase III como el Aramchol. Esta molécula, es un lípido formado al conjugar un ácido biliar y un ácido graso saturado y en este trabajo nos hemos centrado en investigar su mecanismo de acción y su efecto en EHNA.

También hemos trabajado con dos modelos animales de EHNA, un modelo genético (ratones *MAT1A*-KO) y un modelo nutricional (ratones 0.1MCD). Los ratones *MAT1A*-KO tienen una mutación en el gen *Mat1a* que les causa deficiencia crónica en S-adenosilmetionina (SAMe) y esta deficiencia, hace que se desregulen procesos como la metilación del ADN, la biosíntesis de lípidos y proteínas y la respuesta antioxidante de la célula. Por otra parte, los ratones 0.1MCD son alimentados con una dieta deficiente en colina que contiene un 0.1% de metionina para evitar la pérdida de peso que ocurre con la dieta deficiente en colina y metionina canónica. Al igual que los ratones *MAT1A*-KO, los ratones 0.1MCD tienen bajos niveles de SAMe, de GSH y deficiencias en la formación y transporte de moléculas VLDL. Ambos modelos desarrollan EHNA de forma espontánea, aunque los ratones 0.1MCD lo hacen mucho más rápido.

En este trabajo, hemos analizado el perfil metabolómico de estos dos modelos de EHNA y los hemos comparado con el perfil metabolómico de pacientes con EHGNA. Primero vimos que en los ratones *MAT1A*-KO, el perfil metabólico sérico refleja el metabolismo hepático. Entonces, seleccionamos los 50 metabolitos de suero que diferían más entre los ratones *MAT1A*-KO y los controles y generamos una lista. Después comparamos el perfil metabólico de estos ratones con el perfil metabólico de pacientes con EHNA y vimos que un 49% de los pacientes presentaba un perfil muy similar a estos ratones. Validamos estos resultados mediante un método de aleatorización de muestras, pero además lo validamos realizando el mismo análisis que acabamos de describir, en ratones 0.1MCD donde obtuvimos un resultado muy similar. Estos resultados revelaron la existencia de 2 subtipos diferentes de EHGNA: subtipo M^+ y subtipo no M^+.

Los pacientes pertenecientes al subtipo M^+ tienen un perfil metabólico sérico similar al de los ratones *MAT1A*-KO y 0.1MCD por lo que podrían beneficiarse de los tratamientos que mejoren la enfermedad en estos modelos animales. En relación a esto, tratamos ratones *MAT1A*-KO con SAMe y vimos una mejora en la histología de hígado, así como una mejora de la función hepática y una recuperación en los niveles de metilación de determinadas regiones del ADN. Si un tratamiento con SAMe podría ser efectivo en los pacientes del subtipo M^+, es algo que aún está por determinar.

Por otra parte, analizamos el efecto del Aramchol en ratones 0.1MCD y vimos que también mejoraba la histología y la función del hígado en los ratones tratados con esta molécula. Además, también identificamos en estos ratones, biomarcadores asociados al tratamiento con Aramchol.

El Aramchol está actualmente en fase III de estudio clínico y es una molécula con muchas posibilidades para el tratamiento de EHNA en humanos. En la fase IIb el Aramchol mejoró la fibrosis y la bioquímica demostrando no ser tóxico en humanos. Además, reduce la cantidad de hemoglobina glicosilada (HbA1C) en sangre, molécula que se observa en valores altos en los pacientes de EHGNA que reciben el placebo. Este resultado sugiere que el Aramchol afecta al metabolismo de la glucosa. Cuando analizamos los hígados de ratones 0.1MCD en comparación a los controles y observamos metabolitos como la glucosa, glucosa-6-P y fructosa-6-P (entre otros), vemos que el Aramchol mejora el metabolismo hepático de la glucosa.

Gracias a estudios anteriores, sabíamos que el Aramchol reducía la actividad de la enzima Stearoyl-CoA desaturasa I (SCD1), la primera enzima de la ruta de síntesis de triglicéridos (TGs). En este trabajo hemos confirmado esto de diferentes formas, y también por el análisis, mediante western blot, de extractos proteicos obtenidos de hepatocitos primarios de ratón incubados con medio control o medio MCD tratados con Aramchol o vehículo. De esta forma vimos que Aramchol reduce la cantidad de SCD1 pero también afecta a otras proteínas como la AMPK, CPT1A/B, ACC, quinasa P70-60 y la proteína ribosoma S6. En hepatocitos, el tratamiento con Aramchol aumenta los niveles de CPT1A/B, de la forma activa de AMPK (p-AMPK) y de la forma inactiva de ACC (p-ACC). Al contrario, el tratamiento con Aramchol reduce los niveles los ratios pS6/S6, y pP70-S6K/ pP70-S6K. Esto indica que el Aramchol estimula las vías catabólicas y reduce las anabólicas.

De nuevo, en experimentos con hepatocitos primarios de ratón incubados con medio control o medio MCD tratados con Aramchol o vehículo, identificamos más de 3000 proteínas diferentes y seleccionamos las que se expresaban de forma más diferencial entre ambas condiciones. Nuestros resultados muestran que el Aramchol afecta a procesos biológicos clave de forma diferente ya que reduce las proteínas relacionadas con procesos de traducción y fibrosis, y activa procesos como la oxidación de ácidos grasos, eliminación de gotas lipídicas, fosforilación oxidativa y el ciclo de los ácidos tricarboxílicos (TCA por sus siglas en inglés).

Como vimos que el Aramchol aumentaba la actividad del ciclo de los ácidos tricarboxílicos, evaluamos más en detalle el efecto de esta molécula en este ciclo. Mediante experimentos de fluxómica y utilizando glucosa uniformemente marcada con el isotopo ^{13}C vimos que cuando los hepatocitos son tratados con Aramchol, las especies de malato marcadas con ^{13}C se quedan durante mas rondas dentro del ciclo tricarboxílico en comparación con los controles no tratados.

El hecho de que el malato se mantenga durante mas rondas en el ciclo indica que no sale de él para formar parte de otros procesos biológicos, o lo que es lo mismo, esto indica que hay una reducción de la cataplerosis y en consecuencia, una reducción en la gluconeogénesis.

Siguiendo el mismo diseño experimental, vimos que el Aramchol bloquea la formación de lyso-phosphatidylcholina y glycerophosphocholina. Esto podría suponer que los niveles de phosphatidylcholina aumentan y por lo tanto hubiese una mejora en la formación de moléculas de VLDL así como en su transporte. Todos estos resultados han aportado luz sobre el mecanismo de acción del Aramchol.

Por otra parte, en este trabajo también hemos investigado sobre biomarcadores de Síndrome Metabólico (SM), que es un desorden complejo definido por un conjunto de factores relacionados que aumentan el riesgo de sufrir otras enfermedades relacionadas como diabetes tipo 2 o enfermedad cardiovascular. El SM está muy asociado con la EHGNA y también tiene un coste socioeconómico muy elevado, además de ser considerado una epidemia a nivel mundial. Actualmente no existe una única definición para el SM y diferentes organizaciones aportan definiciones similares (aunque diferentes). Nosotros trabajamos con la definición que da la Organización Mundial de la Salud (OMS) que considera que un paciente sufre SM cuando los niveles de glucosa en ayunas están por encima de 110 mg/dL, como requisito obligatorio, y además tiene alterados dos o más de los siguientes parámetros: obesidad (IMC > 30 kg/m^2), dislipidemia (TGs >> 150 mg/dL o cHDL < 34,79 mg/dL en hombres y < 38,66 mg/dL en mujeres), presión arterial alta (≥ 140/90 mmHg o tratamiento) y microalbuminuria (excreción de albumina de 20µg/min).

Todas las definiciones actuales de SM están basadas en una mezcla de valores bioquímicos y antropométricos. En este trabajo, nuestro objetivo es encontrar una huella metabólica asociada a este síndrome para aportar una definición universal basada en patrones metabólicos.

Para realizar esta tarea, trabajamos con 575 muestras de orina humana procedentes de población general de Valencia y Euskadi. Junto con las muestras de orina, nos facilitaron datos médicos y bioquímicos de cada paciente. Así, contando con esta información, nuestro objetivo fue desarrollar un método que, apoyado fuertemente en la bioinformática, nos permitiese diferenciar entre personas asintomáticas (para SM), pacientes de SM y los casos intermedios entre estas dos condiciones usando las diferencias metabólicas presentes en las muestras de orina. Generamos un método para clasificar a los pacientes en clases. Este método binario se basa en definir en 4 bits (o posiciones) los parámetros del SM tomando como referencia la definición de la OMS. 0 significa condición no presente y 1, condición presente. Así obtuvimos 16 perfiles en total, 4 para SM (1011, 1101, 1110 and 1111), 1 para las personas asintomáticas (0000) y el resto para los fenotipos intermedios.

Por otra parte, las muestras de orina de cada paciente fueron medias en un espectrómetro de 600 MHz de Bruker se obtuvieron espectros monodimensionales de protón para cada muestra. Cada espectro se fragmentó en bins de 0.03 ppm y se normalizó por el área total.

Con esta información realizamos análisis univariantes (dirigidos y no dirigidos) donde vimos qué metabolitos y bins diferían más entre los pacientes con SM y las personas asintomáticas. Estos resultados, resaltaron la importancia de la glucosa en este desorden metabólico.

Además, realizamos análisis multivariantes como el *Orthogonal Partial Least Squares Discriminant Analysis*, donde vimos que hay una separación clara entre personas asintomáticas y pacientes con SM. Sin embargo, no pudimos diferenciar a las personas con fenotipos intermedios.

Por otra parte, comprobamos cual era la capacidad predictiva de nuestro modelo. También comprobamos que no hubiese overfiting. Los resultados que tenemos en este momento nos indican que nuestro modelo de entrenamiento tiene una buena capacidad explicativa y que además, puede distinguir entre diferentes partes del espectro de orina obtenido por RMN, siendo capaz de diferenciar a los pacientes de síndrome metabólico y a las personas asintomáticas para esta condición.

Aunque preliminares, estos resultados son muy prometedores ya que sugieren que si aumentamos el número de muestras analizadas, podremos distinguir de forma continua los fenotipos intermedios y crear un algoritmo que en el futuro, ayude a determinar el estatus metabólico de los pacientes.

Para desarrollar una estrategia de diagnóstico mediante el uso de la metabolómica, hemos trabajado con cromatografía líquida-espectrometría de masas y con resonancia magnética nuclear. Hemos realizado un análisis metabolómico de grandes cohortes de pacientes para primero, caracterizar su perfil metabólico en diferentes tipos de muestras biológicas y después analizar estadísticamente esas diferencias para clasificar a estos pacientes en diferentes grupos según su perfil. Debido a la etiología multifactorial de la EHGNA y del SM, la metabolómica es una herramienta muy útil para estudiar y caracterizar grandes poblaciones de pacientes que padecen estos trastornos. Debido a que estas herramientas nos permiten obtener huellas metabólicas de condiciones biológicas concretas, es posible identificar biomarcadores específicos de una enfermedad y además, nos permiten realizar su seguimiento a lo largo del tiempo o bajo tratamiento. La metabolómica y la medicina de precisión son muy compatibles ya que mediante la metabolómica podemos detectar subtipos de pacientes y mediante la comprensión de los mecanismos que subyacen a su enfermedad, podremos generar tratamientos personalizados y efectivos.

8 BIBLIOGRAPHY

1. Rui, L. Energy metabolism in the liver. *Compr. Physiol.* **4**, 177–97 (2014).

2. Cohen, J. C., Horton, J. D. & Hobbs, H. H. Human fatty liver disease: Old questions and new insights. *Science (80-.).* **332**, 1519–1523 (2011).

3. Ahmed, M. Non-alcoholic fatty liver disease in 2015. *World J. Hepatol.* **7**, 1450 (2015).

4. Szczepaniak, L. S. *et al.* Magnetic resonance spectroscopy to measure hepatic triglyceride content: prevalence of hepatic steatosis in the general population. *Am J Physiol Endocrinol Metab* **288**, 462–468 (2005).

5. Diehl, A. M. & Day, C. Cause, Pathogenesis, and Treatment of Nonalcoholic Steatohepatitis. *N. Engl. J. Med.* **377**, 2063–2072 (2017).

6. Suzuki, A. & Diehl, A. M. Nonalcoholic Steatohepatitis. *Annu. Rev. Med.* **68**, 85–98 (2017).

7. Younossi, Z. M. Non-alcoholic fatty liver disease – A global public health perspective. *J. Hepatol.* **70**, 531–544 (2019).

8. Takuma, Y. & Lai, M. Nonalcoholic steatohepatitis-associated hepatocellular carcinoma: Our case series and literature review Yoshitaka Takuma, Kazuhiro Nouso. *World J Gastroenterol* **16**, 1436–1441 (2010).

9. Cholankeril, G., Patel, R., Khurana, S. & Satapathy, S. K. Hepatocellular carcinoma in non-alcoholic steatohepatitis: Current knowledge and implications for management. *World J. Hepatol.* **9**, 533 (2017).

10. Rainer Hambrecht, S. G. Essay Hunter-gatherer to sedentary lifestyle. *Lancet* **366**, S60–S61 (2005).

11. Trovato, F. M. *et al.* Fatty liver disease and lifestyle in youngsters: diet, food intake frequency, exercise, sleep shortage and fashion. *Liver Int.* **36**, 427–33 (2016).

12. Golabi, P. *et al.* Components of metabolic syndrome increase the risk of mortality in nonalcoholic fatty liver disease (NAFLD). *Medicine (Baltimore).* **97**, e0214 (2018).

13. Mato, J. M., Alonso, C., Noureddin, M. & Lu, S. C. Biomarkers and subtypes of deranged lipid metabolism in non-alcoholic fatty liver disease. *World J. Gastroenterol.* **25**, 3009–3020 (2019).

14. Vacca, M., Allison, M., Griffin, J. & Vidal-Puig, A. Fatty Acid and Glucose Sensors in Hepatic Lipid Metabolism: Implications in NAFLD. *Semin. Liver Dis.* **35**, 250–261 (2015).

15. Miyazaki, M. *et al.* Hepatic Stearoyl-CoA Desaturase-1 Deficiency Protects Mice from Carbohydrate-Induced Adiposity and Hepatic Steatosis. *Cell Metab.* **6**, 484–496 (2007).

16. Alves-Bezerra, M. & Cohen, D. E. Triglyceride metabolism in the liver. *Compr Physiol* **8**, 1–8 (2017).

17. Shindou, H., Hishikawa, D., Harayama, T., Yuki, K. & Shimizu, T. Recent progress on acyl CoA: lysophospholipid acyltransferase research. *J. Lipid Res.* **50**, S46–S51 (2009).

18. Heacock, A. M. & Agranoff, B. W. CDP-diacylglycerol synthase from mammalian tissues. *Biochim. Biophys. Acta - Lipids Lipid Metab.* **1348**, 166–172 (1997).

19. Martínez-Uña, M. *et al.* Excess S-adenosylmethionine reroutes phosphatidylethanolamine towards phosphatidylcholine and triglyceride synthesis. *Hepatology* **58**, 1296–1305 (2013).

20. Kawano, Y. *et al.* Identification of Lipid Species Linked to the Progression of Non-Alcoholic Fatty Liver Disease. *Curr. Drug Targets* **16**, 1293–300 (2015).

21. Cano, A. *et al.* Methionine adenosyltransferase 1A gene deletion disrupts hepatic VLDL assembly in mice. *Hepatology* **54**, 1975–1986 (2011).

22. Fan, C. Y. *et al.* Steatohepatitis, spontaneous peroxisome proliferation and liver tumors in mice lacking peroxisomal fatty acyl-CoA oxidase. Implications for peroxisome proliferator-activated receptor alpha natural ligand metabolism. *J. Biol. Chem.* **273**, 15639–45 (1998).

23. Rolo, A. P., Teodoro, J. S. & Palmeira, C. M. Role of oxidative stress in the pathogenesis of nonalcoholic steatohepatitis. *Free Radic. Biol. Med.* **52**, 59–69 (2012).

24. Pessayre, D. & Fromenty, B. NASH: a mitochondrial disease. *J. Hepatol.* **42**, 928–940 (2005).

25. Barr, J. *et al.* Obesity dependent metabolic signatures associated with nonalcoholic fatty liver disease progression NIH Public Access. *J Proteome Res* **11**, 2521–2532 (2012).

26. Jha, P. *et al.* Genetic Regulation of Plasma Lipid Species and Their Association with Metabolic Phenotypes. *Cell Syst.* **6**, 709-721.e6 (2018).

27. Puri, P. *et al.* The plasma lipidomic signature of nonalcoholic steatohepatitis. *Hepatology* **50**, 1827–38 (2009).

28. Quehenberger, O. & Dennis, E. A. The human plasma lipidome. *N. Engl. J. Med.* **365**, 1812–23 (2011).

29. Rhee, E. P. *et al.* Lipid profiling identifies a triacylglycerol signature of insulin resistance and improves diabetes prediction in humans. *J. Clin. Invest.* **121**, 1402–11 (2011).

30. Santoro, N. *et al.* Variant in the glucokinase regulatory protein (GCKR) gene is associated with fatty liver in obese children and adolescents. *Hepatology* **55**, 781–789 (2012).

31. Bu, P. *et al.* Aldolase B-Mediated Fructose Metabolism Drives Metabolic Reprogramming of Colon Cancer Liver Metastasis. *Cell Metab.* **27**, 1249-1262.e4 (2018).

32. Ouyang, X. *et al.* Fructose Consumption as a Risk Factor for Non-alcoholic Fatty Liver Disease. **48**, 993–999 (2009).

33. Ackerman, Z. *et al.* Fructose-induced fatty liver disease: hepatic effects of blood pressure and plasma triglyceride reduction. *Hypertens. (Dallas, Tex. 1979)* **45**, 1012–8 (2005).

34. Davail, S. *et al.* Effects of Dietary Fructose on Liver Steatosis in Overfed Mule Ducks. *Horm. Metab. Res.* **37**, 32–35 (2005).

35. Wolfe, B. M., Ahuja, S. P. & Marliss, E. B. Effects of intravenously administered fructose and glucose on splanchnic amino acid and carbohydrate metabolism in hypertriglyceridemic men. *J. Clin. Invest.* **56**, 970–977 (1975).

36. Vernon, G., Baranova, A. & Younossi, Z. M. Systematic review: the epidemiology and natural history of non-alcoholic fatty liver disease and non-alcoholic steatohepatitis in adults. *Aliment. Pharmacol. Ther.* **34**, 274–285 (2011).

37. Iruarrizaga-Lejarreta, M. *et al.* Emerging Circulating Biomarkers for The Diagnosis and Assessment of Treatment Responses in Patients with Hepatic Fat Accumulation, Nash and Liver Fibrosis. in *Translational Research Methods in Diabetes, Obesity, and Nonalcoholic Fatty Liver Disease* 423–448 (Springer International Publishing, 2019). doi:10.1007/978-3-030-11748-1_16

38. Pirola, C. J. & Sookoian, S. Multiomics biomarkers for the prediction of nonalcoholic fatty liver disease severity. *World J Gastroenterol* **24**, 1601–1615 (2018).

39. Saraf, N., Sharma, P. K., Mondal, S. C., Garg, V. K. & Singh, A. K. Role of PPARg2 transcription factor in thiazolidinedione-induced insulin sensitization. *Journal of Pharmacy and Pharmacology* **64**, 161–171 (2012).

40. Ratziu, V., Goodman, Z. & Sanyal, A. Current efforts and trends in the treatment of NASH. *J. Hepatol.* **62**, S65–S75 (2015).

41. Neuschwander-Tetri, B. A. *et al.* Farnesoid X nuclear receptor ligand obeticholic acid for non-cirrhotic, non-alcoholic steatohepatitis (FLINT): a multicentre, randomised, placebo-controlled trial. *Lancet (London, England)* **385**, 956–65 (2015).

42. Wagner, M., Zollner, G. & Trauner, M. Nuclear bile acid receptor farnesoid X receptor meets nuclear factor-κB: New insights into hepatic inflammation. *Hepatology* **48**, 1383–1386 (2008).

43. Cariou, B. *et al.* Dual Peroxisome Proliferator-Activated Receptor / Agonist GFT505 Improves Hepatic and Peripheral Insulin Sensitivity in Abdominally Obese Subjects. *Diabetes Care* **36**, 2923–2930 (2013).

44. Staels, B. *et al.* Hepatoprotective effects of the dual peroxisome proliferator-activated receptor alpha/delta agonist, GFT505, in rodent models of nonalcoholic fatty liver disease/nonalcoholic steatohepatitis. *Hepatology* **58**, 1941–1952 (2013).

45. Federico, A. *et al.* A new silybin-vitamin E-phospholipid complex improves insulin resistance and liver damage in patients with non-alcoholic fatty liver disease: preliminary observations. *Gut* **55**, 901–2 (2006).

46. Cassidy, S. & Syed, B. A. Nonalcoholic steatohepatitis (NASH) drugs market. *Nat. Rev. Drug Discov. 2016 1511* **15**, 745–746 (2016).

47. IQVIA Institute for Human Data Science. The Changing Landscape of Research and Development. *Iqvia* 31 (2019).

48. Morrison, M. C. *et al.* Obeticholic Acid Modulates Serum Metabolites and Gene Signatures Characteristic of Human NASH and Attenuates Inflammation and Fibrosis Progression in Ldlr-/-.Leiden Mice. *Hepatol. Commun.* **2**, 1513–1532 (2018).

49. Safadi, R. *et al.* The Fatty Acid-Bile Acid Conjugate Aramchol Reduces Liver Fat Content in Patients With Nonalcoholic Fatty Liver Disease. *Clin. Gastroenterol.*

Hepatol. **12**, 2085–2091 (2014).

50. Dobrzyn, A. & Ntambi, J. M. Stearoyl-CoA desaturase as a new drug target for obesity treatment. *Obes. Rev.* **6**, 169–174 (2005).

51. Dobrzyn, P. *et al.* Stearoyl-CoA desaturase 1 deficiency increases fatty acid oxidation by activating AMP-activated protein kinase in liver. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 6409–14 (2004).

52. Konikoff, F. M. & Gilat, T. Effects of Fatty Acid Bile Acid Conjugates (FABACs) on Biliary Lithogenesis: Potential Consequences for Non-Surgical Treatment of Gallstones. *Curr. Drug Targets - Immune, Endocr. Metab. Disord.* **5**, 171–175 (2005).

53. Mato, J. M., Martínez-Chantar, M. L. & Lu, S. C. S-adenosylmethionine metabolism and liver disease. *Ann. Hepatol.* **12**, 183 (2013).

54. Best, C. H., Hershey, J. M. & Huntsman, M. E. The effect of lecithine on fat deposition in the liver of the normal rat. *J. Physiol.* **75**, 56–66 (1932).

55. Kinsell, L. W., Harper, H. A., Barton, H. C., Hutchin, M. E. & Hess, J. R. STUDIES IN METHIONINE AND SULFUR METABOLISM. I. THE FATE OF INTRAVENOUSLY ADMINISTERED METHIONINE, IN NORMAL INDIVIDUALS AND IN PATIENTS WITH LIVER DAMAGE 1, 2. *J. Clin. Invest* **27**, 677–688 (1948).

56. Mato, J. M., Martínez-Chantar, M. L. & Lu, S. C. Methionine Metabolism and Liver Disease. *Annu. Rev. Nutr.* **28**, 273–293 (2008).

57. CATONI, G. L. S-Adenosylmethionine; a new intermediate formed enzymatically from L-methionine and adenosinetriphosphate. *J. Biol. Chem.* **204**, 403–16 (1953).

58. Duce, A. M., Ortíz, P., Cabrero, C. & Mato, J. M. S-adenosyl-L-methionine synthetase and phospholipid methyltransferase are inhibited in human cirrhosis. *Hepatology* **8**, 65–8

59. Mato, J. M., Corrales, F. J., LU, S. C. & Avila, M. A. S-Adenosylmethionine: a control switch that regulates liver function. *FASEB J.* **16**, 15–26 (2002).

60. Mudd, S. H. & Poole, J. R. Labile methyl balances for normal humans on various dietary regimens. *Metabolism* **24**, 721–735 (1975).

61. Manley, G. *et al.* Methionine and Protein Metabolism in Non Alcoholic Steatohepatitis: Evidence for Lower Rate of Transmethylation of Methionine. *PLoS One* **32**, 736–740 (2017).

62. Petrossian, T. C. & Clarke, S. G. Uncovering the Human Methyltransferasome. *Mol. Cell. Proteomics* **10**, M110.000976 (2011).

63. Mato, J. M. *et al.* S-Adenosylmethionine in alcoholic liver cirrhosis: a randomized, placebo-controlled, double-blind, multicenter clinical trial. *J. Hepatol.* **30**, 1081–1089 (1999).

64. Finkelstein, J. D. Metabolic regulatory properties of S-adenosylmethionine and S-adenosylhomocysteine. *Clin. Chem. Lab. Med.* **45**, 1694–1699 (2007).

65. Lu, S. C. Regulation of hepatic glutathione synthesis: current concepts and controversies. *FASEB J.* **13**, 1169–83 (1999).

66. Lu, S. C. Regulation of glutathione synthesis. *Mol. Aspects Med.* **30**, 42–59 (2009).

67. Mato, J. M., Alvarez, L., Ortiz, P. & Pajares, M. A. S-adenosylmethionine synthesis: molecular mechanisms and clinical implications. *Pharmacol. Ther.* **73**, 265–80 (1997).

68. Prudova, A. *et al.* S-adenosylmethionine stabilizes cystathionine beta-synthase and modulates redox capacity. *Proc. Natl. Acad. Sci.* **103**, 6489–6494 (2006).

69. Avila, M. A., García-Trevijano, E. R., Lu, S. C., Corrales, F. J. & Mato, J. M. Methylthioadenosine. *Int. J. Biochem. Cell Biol.* **36**, 2125–2130 (2004).

70. Avila, M. A. *et al.* Reduced mRNA abundance of the main enzymes involved in methionine metabolism in human liver cirrhosis and hepatocellular carcinoma. *J. Hepatol.* **33**, 907–914 (2000).

71. Van Herck, M. A., Vonghia, L. & Francque, S. M. Animal models of nonalcoholic fatty liver disease—a starter's guide. *Nutrients* **9**, 1–13 (2017).

72. Veseli, B. E. *et al.* Animal models of atherosclerosis. *Eur. J. Pharmacol.* **816**, 3–13 (2017).

73. Jacobs, A., Warda, A.-S., Verbeek, J., Cassiman, D. & Spincemaille, P. An Overview of Mouse Models of Nonalcoholic Steatohepatitis: From Past to Present. in *Current Protocols in Mouse Biology* **6**, 185–200 (John Wiley & Sons, Inc., 2016).

74. Morrison, M. C. *et al.* Intervention with a caspase-1 inhibitor reduces obesity-associated hyperinsulinemia, non-alcoholic steatohepatitis and hepatic fibrosis in LDLR−/−.Leiden mice. *Int. J. Obes.* **40**, 1416–1423 (2016).

75. Ibrahim, S. H., Ch, M. B. B., Hirsova, P., Malhi, H. & Gores, G. J. Animal Models of

Nonalcoholic Steatohepatitis: Eat, Delete, and Inflame. *Dig. Dis. Sci.* **61(5)**, 1325–1336 (2016).

76. Lu, S. C. *et al. Methionine adenosyltransferase 1A knockout mice are predisposed to liver injury and exhibit increased expression of genes involved in proliferation*. *Pnas* **98**, (2001).

77. Luka, Z., Capdevila, A., Mato, J. M. & Wagner, C. A Glycine N-methyltransferase Knockout Mouse Model for Humans with Deficiency of this Enzyme. *Transgenic Res.* **15**, 393–397 (2006).

78. Luz Martínez-Chantar, M. *et al. Loss of the Glycine N-Methyltransferase Gene Leads to Steatosis and Hepatocellular Carcinoma in Mice NIH Public Access*. *Hepatology* **47**, (2008).

79. Ramani, K. & Lu, S. C. Methionine adenosyltransferases in liver health and diseases. *Liver Res* **1(2)**, 103–111 (2017).

80. Gil, B. *et al.* Differential expression pattern ofS-adenosylmethionine synthetase isoenzymes during rat liver development. *Hepatology* **24**, 876–881 (1996).

81. Huang, Z.-Z., Mao, Z., Cai, J. & Lu, S. C. Changes in methionine adenosyltransferase during liver regeneration in the rat. *Am. J. Physiol. Liver Physiol.* **275**, G14-21 (1998).

82. García-Trevijano E. R, Latasa M. U, Carretero M. V, Berasain C, Mato J. M, A. M. A. S-Adenosylmethionine regulates MAT1A and MAT2A gene expression in cultured rat hepatocytes: a new role for S-adenosylmethionine in the maintenance of the differentiated status of the liver. *FASEB J.* **14**, 2511–2518 (2000).

83. Cai, J., Sun, W., Hwang, J., Stain, S. C. & Lu, S. C. Changes in S-adenosylmethionine synthetase in human liver cancer: Molecular characterization and significance. *Hepatology* **24**, 1090–1097 (1996).

84. Yang, H. *et al.* Role of promoter methylation in increased methionine adenosyltransferase 2A expression in human liver cancer. *Am. J. Physiol. Liver Physiol.* **280**, G184–G190 (2001).

85. Lu, S. C. S-Adenosylmethionine. *Int. J. Biochem. Cell Biol.* **32**, 391–395 (2000).

86. Sullivan, D. M. & Hoffman, J. L. Fractionation and kinetic properties of rat liver and kidney methionine adenosyltransferase isozymes. *Biochemistry* **22**, 1636–1641 (1983).

87. Halim, A. B., LeGros, L., Geller, A. & Kotb, M. Expression and functional interaction of the catalytic and regulatory subunits of human methionine adenosyltransferase in mammalian cells. *J. Biol. Chem.* **274**, 29720–5 (1999).

88. Kanuri, G. & Bergheim, I. In vitro and in vivo models of non-alcoholic fatty liver disease (NAFLD). *Int. J. Mol. Sci.* **14**, 11963–80 (2013).

89. Omagari, K. *et al.* Effects of a Long-Term High-Fat Diet and Switching from a High-Fat to Low-Fat, Standard Diet on Hepatic Fat Accumulation in Sprague-Dawley Rats. *Dig. Dis. Sci.* **53**, 3206–3212 (2008).

90. Varela-Rey, M. *et al.* Non-alcoholic steatohepatitis and animal models: Understanding the human disease. *Int. J. Biochem. Cell Biol.* **41**, 969–976 (2009).

91. Takahashi, Y., Soejima, Y. & Fukusato, T. Animal models of nonalcoholic fatty liver disease/nonalcoholic steatohepatitis. *World J. Gastroenterol.* **18**, 2300–8 (2012).

92. Buettner, R., Schölmerich, J. & Bollheimer, L. C. High-fat Diets: Modeling the Metabolic Disorders of Human Obesity in Rodents*. *Obesity* **15**, 798–808 (2007).

93. Marra, F., Gastaldelli, A., Svegliati Baroni, G., Tell, G. & Tiribelli, C. Molecular basis and mechanisms of progression of non-alcoholic steatohepatitis. *Trends Mol. Med.* **14**, 72–81 (2008).

94. Siegel, A. B. & Zhu, A. X. Metabolic syndrome and hepatocellular carcinoma: Two growing epidemics with a potential link. *Cancer* **115**, 5651–5661 (2009).

95. Day, C. P. From fat to inflammation. *Gastroenterology* **130**, 207–10 (2006).

96. Kershaw, E. E. & Flier, J. S. Adipose Tissue as an Endocrine Organ. *J. Clin. Endocrinol. Metab.* **89**, 2548–2556 (2004).

97. Teli, M. R., James, O. F. W., Burt, A. D., Bennett, M. K. & Day, C. P. The natural history of nonalcoholic fatty liver: A follow-up study. *Hepatology* **22**, 1714–1719 (1995).

98. Takahashi, Y. & Fukusato, T. Histopathology of nonalcoholic fatty liver disease/nonalcoholic steatohepatitis. *World J. Gastroenterol.* **20**, 15539–48 (2014).

99. BRUNT, E. Pathology of nonalcoholic steatohepatitis. *Hepatol. Res.* **33**, 68–71 (2005).

100. Noureddin, M. *et al.* NASH Leading Cause of Liver Transplant in Women: Updated Analysis of Indications For Liver Transplant and Ethnic and Gender Variances. *Am. J.*

Gastroenterol. **113**, 1649–1659 (2018).

101. Kassi, E., Pervanidou, P., Kaltsas, G. & Chrousos, G. Metabolic syndrome: definitions and controversies. *BMC Med.* **9**, 1741–7015 (2011).

102. Keys, A., Fidanza, F., Karvonen, M. J., Kimura, N. & Taylor, H. L. Indices of relative weight and obesity. *Int. J. Epidemiol.* **43**, 655–665 (2014).

103. Kahn, B. B. & Flier, J. S. Obesity and insulin resistance. *J. Clin. Invest.* **106**, 473–81 (2000).

104. Kitahara, C. M. *et al.* Association between Class III Obesity (BMI of 40–59 kg/m2) and Mortality: A Pooled Analysis of 20 Prospective Studies. *PLoS Med.* **11**, e1001673 (2014).

105. Younes, R. & Bugianesi, E. NASH in Lean Individuals. *Semin. Liver Dis.* **39**, 86–95 (2019).

106. Tilg, H. & Moschen, A. Weight loss: cornerstone in the treatment of non-alcoholic fatty liver disease. *Minerva Gastroenterol. Dietol.* **56**, 159–67 (2010).

107. American Diabetes Association, A. D. Diagnosis and classification of diabetes mellitus. *Diabetes Care* **37**, S81-90 (2014).

108. Reaven, G. M. Pathophysiology of insulin resistance in human disease. *Am. Soc. Clin. Investig.* **75**, 473–486 (1995).

109. COLDITZ, G. A. *et al.* WEIGHT AS A RISK FACTOR FOR CLINICAL DIABETES IN WOMEN. *Am. J. Epidemiol.* **132**, 501–513 (1990).

110. Després, J. P. Body fat distribution and risk of cardiovascular disease: An update. *Circulation* **126**, 1301–1313 (2012).

111. Kissebah, A. H. & Krakower, G. R. Regional adiposity and morbidity. *Physiol. Rev.* **74**, 761–811 (1994).

112. Larry Jameson, J. & Longo, D. L. Precision Medicine—Personalized, Problematic, and Promising. *Obstet. Gynecol. Surv.* **70**, 612–614 (2015).

113. Shapiro, H., Suez, J. & Elinav, E. Personalized microbiome-based approaches to metabolic syndrome management and prevention. *J. Diabetes* **9**, 226–236 (2017).

114. Tan, S. Z., Begley, P., Mullard, G., Hollywood, K. A. & Bishop, P. N. Introduction to

metabolomics and its applications in ophthalmology. *Eye* **30**, 773–783 (2016).

115. Kim, S. J., Kim, S. H., Kim, J. H., Hwang, S. & Yoo, H. J. Understanding Metabolomics in Biomedical Research. *Endocrinol Metab* **31**, 7–16 (2016).

116. Kaddurah-Daouk, R., Kristal, B. S. & Weinshilboum, R. M. Metabolomics: A Global Biochemical Approach to Drug Response and Disease. *Annu. Rev. Pharmacol. Toxicol.* **48**, 653–683 (2008).

117. Bino, R. J. *et al.* Potential of metabolomics as a functional genomics tool. *Trends Plant Sci.* **9**, 418–25 (2004).

118. Moolenaar, S. H., Engelke, U. F. H. & Wevers, R. A. Proton nuclear magnetic resonance spectroscopy of body fluids in the field of inborn errors of metabolism. *Ann. Clin. Biochem.* **40**, 16–24 (2003).

119. Goodacre, R., Vaidyanathan, S., Dunn, W. B., Harrigan, G. G. & Kell, D. B. Metabolomics by numbers: acquiring and understanding global metabolite data. *Trends Biotechnol.* **22**, 245–252 (2004).

120. Oliver, S. G., Winson, M. K., Kell, D. B. & Baganz, F. Systematic functional analysis of the yeast genome. *Trends Biotechnol.* **16**, 373–8 (1998).

121. Raamsdonk, L. M. *et al.* A functional genomics strategy that uses metabolome data to reveal the phenotype of silent mutations. *Nat. Biotechnol.* **19**, 45–50 (2001).

122. Griffin, J. L. & Vidal-Puig, A. Current challenges in metabolomics for diabetes research: a vital functional genomic tool or just a ploy for gaining funding? *Physiol. Genomics* **34**, 1–5 (2008).

123. Dunn, W. B., Broadhurst, D. I., Atherton, H. J., Goodacre, R. & Griffin, J. L. Systems level studies of mammalian metabolomes: the roles of mass spectrometry and nuclear magnetic resonance spectroscopy. *Chem. Soc. Rev.* **40**, 387–426 (2011).

124. Moestue, S., Sitter, B., Bathen, T. F., Tessem, M.-B. & Gribbestad, I. S. HR MAS MR spectroscopy in metabolic characterization of cancer. *Curr. Top. Med. Chem.* **11**, 2–26 (2011).

125. Lane, A. N. *The Handbook of Metabolomics*. **17**, (2012).

126. Atherton, H. J. *et al.* A combined ^{1}H-NMR spectroscopy- and mass spectrometry-based metabolomic study of the PPAR-α null mutant mouse defines profound systemic

changes in metabolism linked to the metabolic syndrome. *Physiol. Genomics* **27**, 178–186 (2006).

127. Fiehn, O. *et al.* Plasma metabolomic profiles reflective of glucose homeostasis in non-diabetic and type 2 diabetic obese African-American women. *PLoS One* **5**, e15234 (2010).

128. Xia, J., Psychogios, N., Young, N. & Wishart, D. S. MetaboAnalyst: a web server for metabolomic data analysis and interpretation. *Nucleic Acids Res.* **37**, W652–W660 (2009).

129. Go, E. P. Database Resources in Metabolomics: An Overview. *J. Neuroimmune Pharmacol.* **5**, 18–30 (2010).

130. Giacomoni, F. *et al.* Gene expression. *Gene Expr.* **31**, 1493–1495 (2015).

131. Xia, J., Bjorndahl, T. C., Tang, P. & Wishart, D. S. MetaboMiner – semi-automated identification of metabolites from 2D NMR spectra of complex biofluids. *BMC Bioinformatics* **9**, 507 (2008).

132. Bino, R. J. *et al.* Potential of metabolomics as a functional genomics tool. *Trends Plant Sci.* **9**, 418–25 (2004).

133. Winter, G. & Krömer, J. O. Fluxomics - connecting 'omics analysis and phenotypes. *Environ. Microbiol.* **15**, 1901–1916 (2013).

134. Cubero-Leon, E., Minier, C., Rotchell, J. M. & Hill, E. M. Metabolomic analysis of sex specific metabolites in gonads of the mussel, Mytilus edulis. *Comp. Biochem. Physiol. Part D Genomics Proteomics* **7**, 212–219 (2012).

135. Chokkathukalam, A., Kim, D.-H., Barrett, M. P., Breitling, R. & Creek, D. J. Stable isotope-labeling studies in metabolomics: new insights into structure and dynamics of metabolic networks. *Bioanalysis* **6**, 511–24 (2014).

136. Iruarrizaga-Lejarreta, M. *et al.* Role of Aramchol in steatohepatitis and fibrosis in mice. *Hepatol. Commun.* **1**, 911–927 (2017).

137. Wiśniewski, J. R., Zougman, A., Nagaraj, N. & Mann, M. Universal sample preparation method for proteome analysis. *Nat. Methods* **6**, 359–362 (2009).

138. Meier, F. *et al.* Parallel accumulation-serial fragmentation (PASEF): Multiplying sequencing speed and sensitivity by synchronized scans in a trapped ion mobility device.

J. Proteome Res. **14**, 5378–5387 (2015).

139. Meier, F. *et al.* Online parallel accumulation–serial fragmentation (PASEF) with a novel trapped ion mobility mass spectrometer. *Mol. Cell. Proteomics* **17**, 2534–2545 (2018).

140. Tyanova, S., Temu, T. & Cox, J. The MaxQuant computational platform for mass spectrometry-based shotgun proteomics. *Nat. Protoc.* **11**, 2301–2319 (2016).

141. Szklarczyk, D. *et al.* STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **47**, D607–D613 (2019).

142. Xu, L. *et al.* Human hepatic stellate cell lines, LX-1 and LX-2: New tools for analysis of hepatic fibrosis. *Gut* **54**, 142–151 (2005).

143. Alonso, C. *et al.* Metabolomic Identification of Subtypes of Nonalcoholic Steatohepatitis. *Gastroenterology* **152**, 1449-1461.e7 (2017).

144. Barr, J. *et al.* Obesity-Dependent Metabolic Signatures Associated with Nonalcoholic Fatty Liver Disease Progression. *J. Proteome Res.* **11**, 2521–2532 (2012).

145. Martínez-Uña, M. *et al.* Excess S-adenosylmethionine reroutes phosphatidylethanolamine towards phosphatidylcholine and triglyceride synthesis. *Hepatology* **58**, 1296–1305 (2013).

146. Martínez-Arranz, I. *et al.* Enhancing metabolomics research through data mining. *J. Proteomics* **127**, 275–288 (2015).

147. Bligh, E. G. & Dyer, W. J. A RAPID METHOD OF TOTAL LIPID EXTRACTION AND PURIFICATION. *Can. J. Biochem. Physiol.* **37**, 911–917 (1959).

148. Ruiz, J. I. & Ochoa, B. Quantification in the subnanomolar range of phospholipids and neutral lipids by monodimensional thin-layer chromatography and image analysis. *J. Lipid Res.* **38**, 1482–9 (1997).

149. van Liempd, S., Cabrera, D., Mato, J. M. & Falcon-Perez, J. M. A fast method for the quantitation of key metabolites of the methionine pathway in liver tissue by high-resolution mass spectrometry and hydrophilic interaction ultra-performance liquid chromatography. *Anal. Bioanal. Chem.* **405**, 5301–5310 (2013).

150. Varela-Rey, M. *et al.* S-adenosylmethionine Levels Regulate the Schwann Cell DNA Methylome. *Neuron* **81**, 1024–1039 (2014).

151. Krueger, F. & Andrews, S. R. Bismark: a flexible aligner and methylation caller for Bisulfite-Seq applications. *Bioinformatics* **27**, 1571–1572 (2011).

152. Akalin, A. *et al.* methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biol.* **13**, R87 (2012).

153. Eng, J. K., Jahan, T. A. & Hoopmann, M. R. Comet: An open-source MS/MS sequence database search tool. *Proteomics* **13**, 22–24 (2013).

154. Craig, R. & Beavis, R. C. TANDEM: matching proteins with tandem mass spectra. *Bioinformatics* **20**, 1466–1467 (2004).

155. Keller, A., Nesvizhskii, A. I., Eugene Kolker, and & Aebersold, R. Empirical Statistical Model To Estimate the Accuracy of Peptide Identifications Made by MS/MS and Database Search. *Anal. Chem* **74**, 5383–5392 (2002).

156. MacLean, B. *et al.* Skyline: an open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* **26**, 966–968 (2010).

157. Schilling, B. *et al.* Platform-independent and label-free quantitation of proteomic data using MS1 extracted ion chromatograms in skyline: application to protein acetylation and phosphorylation. *Mol. Cell. Proteomics* **11**, 202–14 (2012).

158. Choi, M. *et al.* MSstats: an R package for statistical analysis of quantitative mass spectrometry-based proteomic experiments. *Bioinformatics* **30**, 2524–2526 (2014).

159. Kleiner, D. E. *et al.* Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology* **41**, 1313–21 (2005).

160. Mudd, S. H. & Poole, J. R. Labile methyl balances for normal humans on various dietary regimens. *Metabolism* **24**, 721–735 (1975).

161. Loehrer, F. M., Schwab, R., Angst, C. P., Haefeli, W. E. & Fowler, B. Influence of oral S-adenosylmethionine on plasma 5-methyltetrahydrofolate, S-adenosylhomocysteine, homocysteine and methionine in healthy humans. *J. Pharmacol. Exp. Ther.* **282**, 845–50 (1997).

162. Lu, S. C. *et al.* S-adenosylmethionine in the chemoprevention and treatment of hepatocellular carcinoma in a rat model. *Hepatology* **50**, 462–471 (2009).

163. Murphy, S. K. *et al.* Relationship Between Methylome and Transcriptome in Patients With Nonalcoholic Fatty Liver Disease. *Gastroenterology* **145**, 1076–1087 (2013).

164. Mato, J. M., Martínez-Chantar, M. L. & Lu, S. C. S-adenosylmethionine metabolism and liver disease. *Ann. Hepatol.* **12**, 183–189 (2013).

165. Lu, S. C. & Mato, J. M. S-Adenosylmethionine in cell growth, apoptosis and liver cancer. *J. Gastroenterol. Hepatol.* **23**, S73–S77 (2008).

166. Mato, J. M., Martínez-Chantar, M. L. & Lu, S. C. Methionine Metabolism and Liver Disease. *Annu. Rev. Nutr.* **28**, 273–293 (2008).

167. Lu, S. C. & Mato, J. M. S-adenosylmethionine in liver health, injury, and cancer. *Physiol. Rev.* **92**, 1515–42 (2012).

168. Greco, D. *et al.* Gene expression in human NAFLD. *Am. J. Physiol. Liver Physiol.* **294**, G1281–G1287 (2008).

169. Vance, D. E. Physiological roles of phosphatidylethanolamine N-methyltransferase. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* **1831**, 626–632 (2013).

170. Cano, A. *et al.* Methionine adenosyltransferase 1A gene deletion disrupts hepatic VLDL assembly in mice. *Hepatology* **54**, 1975 (2011).

171. Rousseeuw, P. J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* **20**, 53–65 (1987).

172. Anstee, Q. M. & Goldin, R. D. Mouse models in non-alcoholic fatty liver disease and steatohepatitis research. *Int. J. Exp. Pathol.* **87**, 1–16 (2006).

173. Lu, S. C. Regulation of glutathione synthesis. *Mol. Aspects Med.* **30**, 42–59 (2009).

174. Feldstein, A. E. *et al.* Mass spectrometric profiling of oxidized lipid products in human nonalcoholic fatty liver disease and nonalcoholic steatohepatitis. *J. Lipid Res.* **51**, 3046–3054 (2010).

175. Rizki, G. *et al.* Mice fed a lipogenic methionine-choline-deficient diet develop hypermetabolism coincident with hepatic suppression of SCD-1. *J. Lipid Res.* **47**, 2280–2290 (2006).

176. Machado, M. V. *et al.* Mouse models of diet-induced nonalcoholic steatohepatitis reproduce the heterogeneity of the human disease. *PLoS One* **10**, 1–16 (2015).

177. Medici1, V., Janet M.Peerson, Sally P. Stabler3, Samuel W. French4, J. F., Gregory III5, Maria Catrina Virata1, Antony Albanese6, Christopher L. Bowlus1, S., Devaraj7, Edward A. Panacek8, Nazir Rahim1, John R. Richards8, Lorenzo Rossaro1, A. &

Halsted1, C. H. Impared homocysteine transsulfuration is an indicator of ALD. *J Hepatol.* **53**, 551–557 (2010).

178. Zhao, C. *et al.* PPARγ agonists prevent TGFβ1/Smad3-signaling in human hepatic stellate cells. *Biochem. Biophys. Res. Commun.* **350**, 385–391 (2006).

179. Lee, Y. K., Park, J. E., Lee, M. & Hardwick, J. P. Hepatic lipid homeostasis by peroxisome proliferator-activated receptor gamma 2. *Liver Res.* **2**, 209–215 (2018).

180. Leikin-Frenkel, A. *et al.* Fatty Acid Bile Acid Conjugate Inhibits Hepatic Stearoyl Coenzyme A Desaturase and Is Non-atherogenic. *Arch. Med. Res.* **41**, 397–404 (2010).

181. Ferrari, S., Bandi, H. R., Hofsteenge, J., Bussian, B. M. & Thomas, G. Mitogen-activated 70K S6 kinase. Identification of in vitro 40 S ribosomal S6 phosphorylation sites. *J. Biol. Chem.* **266**, 22770–22775 (1991).

182. Owen, O. E., Kalhan, S. C. & Hanson, R. W. The key role of anaplerosis and cataplerosis for citric acid cycle function. *J. Biol. Chem.* **277**, 30409–30412 (2002).

183. Fay, M. P. & Proschan, M. A. Wilcoxon-Mann-whitney or t-test? *Stat. Surv.* **4**, 1–39 (2010).

184. Benjamini, Y. & Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B* **57**, 289–300 (1995).

185. Chin, W. W. Issues and Opinion on Structural Equation Modeling. *Manag. Inf. Syst. Q.* **22**, (1998).

186. Henseler, J., Ringle, C. M. & Sinkovics, R. R. The use of partial least squares path modeling in international marketing. *Adv. Int. Mark.* **20**, 277–319 (2009).

187. Hair, J. F., Sarstedt, M., Ringle, C. M. & Mena, J. A. An assessment of the use of partial least squares structural equation modeling in marketing research. *J. Acad. Mark. Sci.* **40**, 414–433 (2012).

188. Moylan, C. A. *et al.* Hepatic gene expression profiles differentiate presymptomatic patients with mild versus severe nonalcoholic fatty liver disease. *Hepatology* **59**, 471–482 (2014).

189. Kalhan, S. C. *et al.* Methionine and protein metabolism in non-alcoholic steatohepatitis: Evidence for lower rate of transmethylation of methionine. *Clin. Sci.* **121**, 179–189 (2011).

190. Medici, V. *et al.* Impaired homocysteine transsulfuration is an indicator of alcoholic liver disease. *J. Hepatol.* **53**, 551–557 (2010).

191. Myoung, S. H. *et al.* Lysophosphatidylcholine as a death effector in the lipoapoptosis of hepatocytes. *J. Lipid Res.* **49**, 84–97 (2008).

192. Neuschwander-Tetri, B. A. Hepatic lipotoxicity and the pathogenesis of nonalcoholic steatohepatitis: The central role of nontriglyceride fatty acid metabolites. *Hepatology* **52**, 774–788 (2010).

193. Sunny, N. E., Parks, E. J., Browning, J. D. & Burgess, S. C. Excessive hepatic mitochondrial TCA cycle and gluconeogenesis in humans with nonalcoholic fatty liver disease. *Cell Metab.* **14**, 804–810 (2011).

194. Thévenot, E. A., Roux, A., Xu, Y., Ezan, E. & Junot, C. Analysis of the Human Adult Urinary Metabolome Variations with Age, Body Mass Index, and Gender by Implementing a Comprehensive Workflow for Univariate and OPLS Statistical Analyses. *J. Proteome Res.* **14**, 3322–3335 (2015).

195. Ryan, D., Robards, K., Prenzler, P. D. & Kendall, M. Recent and potential developments in the analysis of urine: A review. *Analytica Chimica Acta* **684**, 17–29 (2011).

196. Zhang, T. & Watson, D. G. A short review of applications of liquid chromatography mass spectrometry based metabolomics techniques to the analysis of human urine. *Analyst* **140**, 2907–2915 (2015).

197. Kumar Bharti, S. & Roy, R. Metabolite Identification in NMR-based Metabolomics. *Curr. Metabolomics* **2**, 163-173(11) (2014).

198. Teng, Y. W., Mehedint, M. G., Garrow, T. A. & Zeisel, S. H. Deletion of betaine-homocysteine S-methyltransferase in mice perturbs choline and 1-carbon metabolism, resulting in fatty liver and hepatocellular carcinomas. *J. Biol. Chem.* **286**, 36258–36267 (2011).

199. Liangpunsakul, S. & Chalasani, N. Unexplained elevations in alanine aminotransferase in individuals with the metabolic syndrome: Results from the Third National Health and Nutrition Survey (NHANES III). *Am. J. Med. Sci.* **329**, 111–116 (2005).

200. CH Kim, Z. Y. Nonalcoholic fatty liver disease: A manifestation of the metabolic syndrome. *Cleve. Clin. J. Med.* **75**, (2008).

201. Zweig, M. H. & Campbell, G. Receiver-operating characteristic (ROC) plots: A

fundamental evaluation tool in clinical medicine. *Clin. Chem.* **39**, 561–577 (1993).

202. Hajian-Tilaki, K. Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. *Casp. J. Intern. Med.* **4**, 627–635 (2013).

203. Daskivich, T. *et al.* Differences in online consumer ratings of health care providers across medical, surgical, and allied health specialties: Observational study of 212,933 providers. *J. Med. Internet Res.* **20**, 29–36 (1982).

9 PUBLICATIONS

Cristina Alonso, David Fernández-Ramos, Marta Varela-Rey, Ibon Martínez-Arranz, Nicolás Navasa, Sebastiaan M. Van Liempd, José L. Lavín Trueba, Rebeca Mayo, Concetta P. Ilisso, Virginia G. de Juan, Marta Iruarrizaga-Lejarreta, Laura delaCruz-Villar, Itziar Mincholé, Aaron Robinson, Javier Crespo, Antonio Martín-Duce, Manuel Romero-Gómez, Holger Sann, Julian Platon, Jennifer Van Eyk, Patricia Aspichueta, Mazen Noureddin, Juan M. Falcón-Pérez, Juan Anguita, Ana M. Aransay, María Luz Martínez-Chantar, Shelly C. Lu, and José M. Mato. **Metabolomic Identification of Subtypes of Nonalcoholic Steatohepatitis**. *Gastroenterology* 152, 1449-1461 (2017).

Marta Iruarrizaga-Lejarreta, Marta Varela-Rey, David Fernández-Ramos, Ibon Martínez-Arranz, Teresa C. Delgado, Jorge Simón, Virginia Gutiérrez-de Juan, Laura delaCruz-Villar, Mikel Azkargorta, José L. Lavin, Rebeca Mayo, Sebastiaan M. Van Liempd, Igor Aurrekoetxea, Xabier Buque, Donatella Delle Cave, Arantza Peña, Juan Rodríguez-Cuesta, Ana M. Aransay, Felix Elortza, Juan M. Falcón-Pérez, Patricia Aspichueta, Liat Hayardeny, Mazen Noureddin, Arun J. Sanyal, Cristina Alonso, Juan Anguita, María Luz Martínez-Chantar, Shelly C. Lu, and José M. Mato. **Role of Aramchol in Steatohepatitis and Fibrosis in Mice**. *Hepatology Communications*. 1, 911-927 (2017).

Communication poster

Laura delaCruz-Villar, David Fernández-Ramos, Fernando Lopitz-Otsoa, Mikel Azkargorta, Marta Iruarrizaga-Lejarreta, Jon Bilbao, Diana Cabrera, Sebastiaan M. van Liempd, Cristina Alonso, Félix Elortza, Shelly C. Lu, Liat Hayardeny, Tali Gorfine and José M. Mato. (November 2019). **Aramchol, a SCD1 Regulator, Improves Liver Glucose Homeostasis in NASH**. American Association for the study of liver diseases. Boston, United States.

Cristina Alonso, Marta Iruarrizaga-Lejarreta, Laura delaCruz-Villar, Itziar Mincholé, Ibon Martinez-Arranz and José M. Mato. (April 2018). **Translational approach to NASH diversity: learning from mouse models**. European Association for the Study of the Liver. Paris, France.

Laura dela Cruz-Villar, Nieves Embade, Tammo Diercks, Oscar Millet. (June 2018). **A NMR-Based metabolomics approach to the research of biomarkers in metabolic syndrome patients**. VIII Ibero-American NMR Meeting. Lisboa, Portugal.

10 ABBREVIATIONS

AAALAC: Association for Assessment and Accreditation of Laboratory Animal Care

ACC: acetyl-CoA carboxylase

ACL: ATP - citrate lyase

AdoMet: S-adenosylmethionine

AGPAT: acylglycerol-3-phosphate acyltransferases

ALT: alanine aminotransferase

AMPK: AMP-activated protein kinase

Apo: apolipoprotein

Aramchol: arachidyl-amido cholanoic acid

ARP: actin related protein

AST: aspartate aminotransferase

ATP: adenosine triphosphate

BHMT: betaine homocysteine methyltransferase

BMI: body mass index

CBS: cystathionine beta synthase

cDNA: complementary DNA

CDP-DG: cytidine diphosphate diacylglycerol

cHDL: cholesterol high density lipoprotein

ChREBP: carbohydrate response element binding protein

CPT 1: Carnitine-palmityl-transferase 1

CS: citrate synthase

CYP4A: cytochrome P450 4A

DG: diacylglycerol

DGAT: diacylglycerol acyltransferase

DMEM: Dulbecco's Modified Eagle Medium

DMSO: dimethyl sulfoxide

DNA: deoxyribonucleic acid

DNL: de novo lipogenesis

Elovs: fatty acyl-CoA elongases

ENTPD 5: Ectonucleoside Triphosphate Diphosphohydrolase

ER: endoplasmic reticulum

F1P: fructose-1-phosphate

F6P: fructose-6-phosphate

FA: fatty acid

FAS: fatty acid synthase

FBP: fructose- 1, 6- bisphosphate

FBS: fetal bovine serum

FD: fructose diet

G3P: glycerol-3-phosphate

G6P: glucose-6-phosphate

GAPDH: Glyceraldehyde-3-Phosphate Dehydrogenase

GC: gas chromatography

GCK: glucokinase

GCKR: glucokinase regulatory protein

GNMT: glycine N-methyltransferase

GPAT: glycerol-3-phosphate acyltransferase

GSH: glutathione

H&E: Hematoxylin & Eosin

HCC: hepatocellular carcinoma

HCD: high cholesterol diet

HFD: high-fat diet

HK: hexokinase

HMDB: Human Metabolome Data Base

HR-MAS: high-resolution magic angle spinning

HRP: horseradish-peroxidase

IDF: International Diabetes Federation

Ig: immunoglobulin

IP: immunoprecipitation

IR: insulin resistance

J-res: J-resolved

kg: kilogram

KHK: ketohexokinase

KO: knock out

LC: liquid chromatography

LPA: lysophosphatidic acid

LXR: liver X receptor

M: molar

MAT: methionine adenosyl-transferase

MCDD: methionine-choline deficient diet

MDMC: methionine/choline deficient medium

MEM: minimum essential medium

MetS: metabolic syndrome

MS: mass spectrometry

MS': methionine synthase

MT: methyltransferase

MTA: methylthioadenosine

MTHF: methyltetrahydrofolate

MTP: microsomal triglyceride transfer protein

$NADPH/NADP^+$: nicotine amide dinucleotide phosphate

NAFLD: non-alcoholic fatty liver disease

NASH: non-alcoholic steatohepatitis

NCEP-ATPIII: National Cholesterol Education Program Adult Treatment Panel III

NMR: Nuclear Magnetic Resonance

NO: nitric oxide

NPC2: Niemann-Pick type C2

OBENUTIC: Obesidad, Nutrición y Tecnologías de la Información y Comunicación

OPLS-DA: Orthogonal Partial Least Squares Discriminant Analysis

OXPHOS: oxidative phosphorylation

P70S6k: P70 S6 kinase

PA: phosphatidic acid

PAP: phosphatidate phosphohydrolase

PC: phosphatidylcholine

PCA: principal component analysis

PCR: polymerase chain reaction

PDC pyruvate dehydrogenase complex

PE: phosphatidylethanolamine

PEMT: phosphatidylethanolamine methyltransferase

PNPLA3: palatine-like phospholipase 3

PPAR: peroxisome proliferator activated receptor

PRMT: protein arginine methyltransferase

PSG: penicillin, streptomycin and glutamine

PYR: pyruvate

RNA: ribonucleic acid

ROS: reactive oxygen species

rpm: revolutions per minute

RT: room temperature

RT-qPCR: real time quantitative polymerase chain reaction

Ru5P: ribulose 5-phosphate

SAH: S-adenosylhomocysteine

SAHH: S-Adenosylhomocysteine hydrolase

SAMe: S-adenosylmethionine

SCD: stearoyl-CoA desaturases

SD: standard deviation

SDS-PAGE: sodium dodecyl sulphate-polyacrylamide gel electrophoresis

SREBP1: sterol regulatory element binding protein 1

T2D: type 2 diabetes

TCA: tricarboxylic acid

TG: triacylglyceride

THF: tetrahydrofolate

TSP: trimethylsilylpropanoic

VLDL: very low density lipoproteins

WHO: World Health Organization

WT: wild type

11 LIST OF FIGURES

12 LIST OF TABLES