UNIVERSITY OF LATVIA

Identification of molecular pathophysiological mechanisms of non-alcoholic fatty liver disease (NAFLD)

Doctoral thesis

Department of Internal Medicine

Faculty of Medicine

Author: Ēriks Šmagris

Supervisor: Dr. med. Valdis Pīrāgs

Riga 2018

Abstract

Non-Alcoholic Fatty Liver Disease (NAFLD) affects approximately 30% of the adult population in Western countries. The disease encompasses a spectrum ranging from an excess of fat in the liver (steatosis) to liver inflammation (steatohepatitis), fibrosis, cirrhosis, and ultimately hepatocellular carcinoma [1].

To identify molecular mechanisms causing hepatic fat accumulation, inflammation and fibrosis, we performed the first NAFLD Genome-wide association study (GWAS) in 2008. This GWAS, and a follow-up GWAS in 2014, identified two common single-nucleotide polymorphisms (SNPs) associated with NAFLD. The polymorphisms in the genes PNPLA3 (Ile148Met; I148M; rs738409) and TM6SF2 (Glu167Lys; E167K; rs58542926) were associated with the higher liver fat content and elevated blood liver enzymes (ALT). Additionally, the TM6SF2 E167K polymorphism was associated with lower plasma levels of LDL cholesterol, triglycerides and alkaline phosphatase (ALKP)[2].

In order to understand the mechanisms by which genetic variation in PNPLA3 and TM6SF2 cause NAFLD, we studied these protein variants in cell and animal models. The results of these studies are described in the present thesis, and in three original research articles published between 2014 and 2016.

We found that the PNPLA3 protein resides on the surface of liver lipid droplets and is highly regulated by fasting/refeeding. In mice genetically engineered to express the steatogenic human PNPLA3 I148M variant, the Pnpla3 I148M protein accumulates on the lipid droplet surface to a much greater degree than the wildtype protein.

The TM6SF2 E167K variant causes protein instability and a 60% faster degradation compared to the wild type [2]. Inactivating Tm6sf2 in mice caused accumulation of hepatic fat and reduced plasma levels of triglycerides, similar to the observations in human carriers of TM6SF2 E167K [3]. TM6SF2 resides in the endoplasmic reticulum (ER) and Golgi apparatus (Golgi) and is involved in very low-density lipoprotein (VLDL) generation and secretion [2-4].

In conclusion, the studies presented in this thesis describe the generation and initial phenotyping of animal models for the two most important common genetic causes of human NAFLD. The studies provide initial insights into the mechanisms by which PNPLA3 and TM6SF2 influence NAFLD. Future mechanistic studies in these animal models are likely to

further uncover these mechanisms. A deeper understanding of the genetic causes of NAFLD could aid in the development of new methods to treat or prevent this common disease.

The research underlying the present thesis was conducted: Faculty of Medicine, University of Latvia, Riga, Latvia in collaboration with the Department of Molecular Genetics, University of Texas Southwestern Medical Center at Dallas, Texas, USA, from 2010 to 2016.

Thesis submitted to the Faculty of Medicine, University of Latvia, Riga, Latvia.

Supervisor: Dr. med. Prof. Valdis Pīrāgs, Department of Internal Diseases, University of Latvia

Opponents:

- Dr. med. Prof. Juris Pokrotnieks, Department of Internal Diseases, Riga Stradins University, Latvia
- Dr. med. Prof. Mārcis Leja, Faculty of Medicine, University of Latvia, Latvia
- Dr. biol. Prof. Aldons J. Lūsis, Department of Human Genetics, David Geffen School of Medicine, University of California Los Angeles, California, USA

The doctoral thesis accepted for the commencement of the degree of Dr. Medicine.

Commencement date: 13.03.2018

Table of contents

Abbreviations

AKT2 - AKT Serine/Threonine Kinase 2

ALT - Alanine transaminase

ALKP – Alkaline phosphatase

ApoB - Apoprotein B

ATGL – Adipose triglyceride lipase (PNPL2)

AU - Arbitrary Units

bp – base pairs (of DNA)

ChREBP – Carbohydrate-responsive element-binding protein

CNX – Calnexin; CANX

COPII – Specific coat protein complex II

CRISPR/CAS9 – Clustered regularly interspaced short palindromic repeats

C_t value – quantitative RT-PCR cycle threshold value

DHS – Dallas Heart Study

EBP - Emopamil Binding Protein

ECL - Enhanced Chemiluminescence

ER – Endoplasmic Reticulum

FAS – Fatty acid synthase

FPLC – Fast protein liquid chromatography

GGT - Gamma-glutamyl transpeptidase

HCC – Hepatocellular carcinoma

^{1}H-MRS - Proton Magnetic Resonance Spectroscopy

HTGC - Hepatic triglyceride content

IRB - Institutional review board

KLF6 – Kruppel-like factor 6

LDL-C – Low-density lipoprotein cholesterol

LPL - Lipoprotein lipase

LXR – Liver X receptor

MTOR –Mechanistic Target of Rapamycin

MTTP – Microsomal triglyceride transfer protein

NAFLD – Non-alcoholic fatty liver disease

NCAN – Neurocan

PC – Phosphatidylcholine (in study generalized as phospholipids)

PDI – Protein disulfide-isomerase

PCR - Polymerase chain reaction

PNPLA3 – Patatin-like phospholipase domain-containing protein 3

SAR1b – Secretion Associated Ras-Related GTPase 1B

SOD2 - Superoxide dismutase 2, mitochondrial

shRNA – Short/small hairpin RNA

SNP – Single-nucleotide polymorphism

SREBP1c – Sterol regulatory element-binding protein 1c

RNAseq – RNA sequencing

RPKM - Reads Per Kilobase of transcript per Million Mapped reads

RPLP0 (36B4) – Ribosomal Protein Lateral Stalk Subunit P0

TG – Triglycerides

TM6SF2 – Transmembrane 6 superfamily member 2

VLDL – Very low density lipoprotein

LDLR - Low density lipoprotein receptor

WB – Western Blot

WT –Wild type, most common allele

Novelty of chosen area of research

Non-alcoholic fatty liver disease (NAFLD) is the most common liver disease afflicting Western populations. In the US, approximately 20-30% of adults and 10-20% of children have elevated fat content in the liver [5-7]. A similar tendency is also observed in Asia, where the incidence of NAFLD is rapidly increasing [8]. Complications of NAFLD are now among the leading causes of liver transplantation in the Western world; it is predicted that NAFLD will become the most common cause of liver transplantation within the next ten years [7]. Obesity is the main causal risk factor for NAFLD. However, there is significant inter-individual variation in susceptibility to developing NAFLD. There are patients with the severe metabolic syndrome, obesity and type 2 diabetes, but only slightly elevated liver fat. On the other hand, some lean individuals develop NAFLD. Part of the variability in susceptibility to NAFLD is due to genetic factors. The fact that NAFLD clusters in families further points to the presence of a genetic component to the disease. An understanding of the genetic component of NAFLD could help to predict individual predisposition to the disease and may help to guide preventive measures or treatment.

This thesis is a compilation of the author`s and colleague's discoveries regarding the mechanisms by which genetic variation in PNPLA3 and TM6SF2 influence the risk of NAFLD.

The studies underlying this doctoral thesis for the first time showed that the two most common human genetic forms of NAFLD are caused by two not related gene polymorphisms, in two extremely different proteins, by two different pathways.

Aims of the dissertation

- To identify the association of the TM6SF2 E167K variant with fat accumulation and inflammation in the liver and with changes in plasma lipids and biochemical markers of liver damage in humans.

- To explore the functional effects of TM6SF2 E167K in cell models.

- To develop animal models lacking TM6SF2, in order to study the function of TM6SF2.

- To generate and characterize Pnpla3 I148M *knock-in* mice as a new NAFLD model for studies of human NAFLD caused by PNPLA3 I148M.

Research Limitations

This study has been performed outside the Republic of Latvia and the European Union. This is making difficult to generalize the study data to all the world, because there may be diet, genetics and other factors that either enhance or decrease NAFLD expression in certain populations. The study does not give a complete answer to the exact mechanism by which TM6SF2 and PNPLA3 proteins regulate the fat content in the liver.

Hypothesis of study

Genetic polymorphisms could be the important factor in ethology of human non-alcoholic fatty liver disease (NAFLD) and genome-wide association studies (GWAS) could help to discover those genetic variants. Using cell and animal models could be helpful for future identification of molecular pathophysiological mechanisms of human non-alcoholic fatty liver disease (NAFLD).

Introduction

During the early evolution, eukaryotic cells chose triglycerides (TG) as the main substance of energy storage in cells. That is due to the chemical inertness of TGs and their ability to contain much more kilocalories per unit of weight than carbohydrates or proteins. TG accumulates in specific tissues (white adipose tissue), but there are exceptions. For example, migratory birds before long migrations can accumulate enormous amounts of TG in the liver, which helps them fly long distances with minimal interruptions [6].

In humans, TGs predominantly accumulate in white adipose tissue. The accumulation and release of TGs are regulated hormonally; the process can be very fast and work in both directions. Other human organs have little capacity for TG accumulation. However, the human liver and other organs do store some lipid droplets intracellularly. These lipids in tissues are for local use and are less hormonally regulated.

Thus, white adipose tissue can be considered as a reservoir for TG accumulation. However, if the storage capacity of the reservoir is exceeded, triglycerides start to accumulate in other tissues, and this leads to local tissue inflammation, insulin resistance and/or dyslipidemia [9].

In the last few decades, the prevalence of obesity has reached epidemic proportions in Western countries. Increased body mass index (BMI) is increasingly common not only in adults but also in children [10, 11].

Metabolic syndrome, which is often accompanied by obesity, also includes the accumulation of fat in the liver. Mild TG accumulation in the liver is often regarded as benign and reversible. However, increasing amounts of fat in the liver increases the risk of progression to later stages of NAFLD (steatohepatitis and end-stage liver disease).

Although the main cause of NAFLD for many years was deemed to be the metabolic syndrome, in recent years more and more data show that NAFLD at least partially is genetically determined. The idea of a genetic component emerged from clinical observations that NAFLD (with or without complications) is affecting certain families [12]. Moreover, the prevalence of NAFLD varies by ethnicity (NAFLD is more common in Hispanics than in African American and European Americans)[13]. These observations formed the basis for several genome-wide association studies (GWAS) aimed at discovering genetic variants influencing NAFLD.

Why should NAFLD be investigated? In Western countries, increased liver fat content is found in 20-30% of adults [8, 9] and 10-20% of children [10]. As fat accumulation in the liver increases the risk of developing inflammation of the liver, cirrhosis and hepatocellular carcinoma, NAFLD will more and more increase the burden on health care budgets. In recent years, NAFLD has become one of the leading causes of liver transplantation in Western countries [10].

1. Literature overview

1.1. Basics of NAFLD

1.1.1. Definition of NAFLD

NAFLD as a distinct disease was first described in 1980 [14]. Patients with obesity and type 2 diabetes had unusual liver damage with hepatomegaly and elevated plasma ALT levels. Liver histology showed a massive accumulation of fat in hepatocytes with lobular hepatitis and fibrosis, similar to what is observed in alcohol-related steatohepatitis [18].

Histologically, human NAFLD is defined by the presence of macrovesicular steatosis in more than 5% of hepatocytes in individuals with a history of less than 20 g/day ethanol intake [15]. The differentiation of macrovesicular from microvesicular steatosis is done by histological assessment, and the two may coexist in more advanced disease [16].

Non-alcoholic fatty liver disease (NAFLD) is considered to be part of the metabolic syndrome. The NAFLD diagnosis is a diagnosis of exclusion; it is made after excluding other causes of steatosis and/or liver damage such as chronic viral hepatitis, drug hepatitis (amiodarone, tamoxifen and others), autoimmune hepatitis, hemochromatosis, Wilson's disease and excessive alcohol consumption. NAFLD includes 1) simple fat accumulation in the liver without inflammation, 2) fatty liver with inflammation - steatohepatitis (NASH), 3) fatty liver or steatohepatitis with fibrosis/cirrhosis. In the past, liver cirrhosis caused by NAFLD was often termed ‘cryptogenic cirrhosis’. It is considered that about 50% of all cases of cryptogenic cirrhosis are in fact caused by NAFLD [17]. The main complication of a long lasting steatohepatitis with/without fibrosis/cirrhosis is hepatocellular carcinoma.

1.1.2. Pathogenesis of NAFLD

NAFLD starts with accumulation of neutral lipids (triglycerides) in the liver. Such accumulation is often an innocent process that may resolve with time. However, in some cases, the process can also progress - causing inflammation (NASH), premature death of hepatocytes, activation of liver stellate (Ito) cells with increased collagen I synthesis and development of fibrosis in the liver.

The causes of NAFLD can be divided into pre- and intra-hepatic. Pre-hepatic causes of NAFLD include hyperphagia, obesity, lipodystrophies, insulin resistance and microbial

dysbiosis, which causes an influx of substrate (free fatty acids, glucose, acetate and possible others) into the liver [18].

Triglycerides enter the liver from 3 sources - from food intake (chylomicrons), white adipose tissue lipid stores (entering the liver as free fatty acids, glycerol) and from hepatic TG synthesis (*de novo* lipogenesis). The synthesis and modulation of triglycerides partially take place in the endoplasmic reticulum (ER), and partially on lipid droplets (LD). Postprandial (after food intake) insulin stimulates TG synthesis in the liver, via the cascade: insulin → insulin receptor → AKT2 →mTOR→ LXR→ SREBP1c. Without SREBP1c, ChREBP also has potent lipogenic activity [6].

Incoming or newly synthesized triglycerides in the liver are further modified and stored in lipid droplets or secreted from the liver into the circulation as VLDL particles. Several essential proteins are involved in this process: Apoprotein B (ApoB), MTTP, PDI and SAR1b. In the fasting condition, lipids in the liver are broken down for the energy production. This process involves such proteins as ATGL (PNLPA2) and CGI-56 (ABHD5). In theory, any interference with the function of the proteins mentioned above may result in increased triglyceride accumulation in the liver.

Also, phospholipid metabolism interacts with the metabolism of triglycerides and fatty acids. Phospholipids compose all cell membranes and metabolic defects in this pathway may cause fat accumulation in the liver [19].

Oxidation of free fatty acids takes place in the mitochondria and some other organelles. Defects in these organelles can also cause NAFLD.

Changes in the intestinal microflora have been hypothesized to also play a role in the development of NAFLD. Bacterial metabolic products (such as lipopolysaccharides, ethanol and possible others) are directly absorbed from the intestines through the portal vein into the liver, where they might cause inflammation and metabolic disorders [20].

Little is known about processes that cause inflammation (NASH) in the context of hepatic steatosis. Some data indicate that stellate cells play a major role in the regulation of inflammatory processes in the liver. These cells can produce collagen and are involved in retinol (vitamin A) metabolism in the liver [21].

Similarly, little is known about how fat accumulation in the liver causes liver fibrosis. A few proteins have been suggested to play a direct role in the development of fibrosis of the

human liver. For example, mutations in the genes encoding SOD2 and KLF6 are associated with increased fibrosis in NAFLD [22, 23].

Finally, epigenetic effects have been suggested to play a role in NAFLD development [24].

1.1.3. Epidemiology

Non-alcoholic fatty liver disease is widespread in the world. Because of the use of inaccurate methods for detection of hepatic lipids and the use of small study groups, the precise prevalence of NAFLD is not known. The prevalence of NAFLD ranges from 2.8% to 46% in different populations [1]. By using the most accurate non-invasive method - proton magnetic resonance spectroscopy (^{1}H-MRS), it was shown that the TG content in the liver can vary widely [13, 25]. When the amount of TG is measured in the livers of a specific group of patients with low risk of excessive liver fat accumulation (patients without type 2 diabetes, severe obesity, hypertriglyceridemia, without chronic hepatitis C, use of specific medications, etc.), and using 95 percentile as the highest normal amount of TG in the liver, hepatic TG level up to 5.56% (55.6 mg TG/1 g of liver tissue) was considered as normal [25]. By using this criterion, it was found that the incidence of NAFLD in the population varies depending on race and nationality - it was most common in Hispanics (45%), followed by European Americans (33%) whereas only 24% of Afro-Americans had NAFLD. The prevalence of NAFLD is on average twice as high in men than in women (42% vs. 24%) NAFLD was present in at least 50% of persons with severe obesity [26].

1.1.4. Diagnostics of NAFLD

The diagnosis of NAFLD can be divided into clinical (hepatomegaly), biochemical (ALT and AST), histological (biopsy) and radiological (ultrasonography, computed tomography, magnetic resonance, proton magnetic resonance spectroscopy (^{1}H-MRS)). Although liver biopsy is considered to be the golden standard for the diagnosis of NAFLD, it is an unpleasant procedure for the patient, and an invasive method with possible complications. Also, liver biopsy gives information on only about 1/50000 of liver [27]. However, liver biopsy remains a necessary procedure for two situations: initial diagnosis of liver disease and rapid progressive cases of NAFLD.

Ultrasonography (US), a cheap and widely used method in the clinic, is only partially capable of determining the exact amount of lipid accumulation in the liver. It is also a performer dependent method, and not well suited for detection of liver inflammation and fibrosis. Computed

tomography (CT) and magnetic resonance imaging (MRI) are more accurate than ultrasonography for NAFLD diagnostics. MRI, although more expensive, recently gained a more significant role in the determination of liver damage. MRI is sometimes called "virtual liver biopsy," as at the same time it is possible to visualize the entire liver, to quantify the accumulation of lipids, iron and the amount of fibrosis in the liver [27].

However, the best method for evaluation of the accumulation of lipid in the liver is proton magnetic resonance spectroscopy (^{1}H-MRS). Although the method is expensive, it is very accurate. This method has been useded in some genome-wide association studies [26, 28, 29].

1.1.5. Complications of NAFLD

Put simply, the risk of NAFLD transition to every next stage can describe by a "law of ten." For example, the transition from simple hepatic steatosis to steatohepatitis occurs in 10% of patients within ten years. The same proportion and time ratio remains for the transition from a steatohepatitis to liver cirrhosis, and from cirrhosis to hepatocellular carcinoma (HCC). More precisely, the shift to the following stage in 10 years occurs in 10-25% of patients at a given stage [30]. However, liver cirrhosis develops in only 1-3% of patients with simple hepatic steatosis [31]. Of course, the most serious complication of NAFLD is a hepatocellular carcinoma, which usually occurs in the background of long-term inflammation, fibrosis and/or cirrhosis in the liver [1]. In recent years, NAFLD complications are becoming one of the leading causes of liver transplantation in the Western world. It is expected that in the year 2020, complications of NAFLD will have become the primary cause of liver transplantation in the West [10][32].

1.1.6. Treatment of NAFLD

There is currently no specific treatment for NAFLD. The most common recommendations are the same as for the treatment of the metabolic syndrome - weight loss, diabetes control and increased physical activity. Some medical treatments for NAFLD have been recommended, for example fish oil, betaine, ursodeoxycholic acid, vitamin E and other antioxidants [33]. However, the studies on these treatments have been small; the results are controversial. The lack of an effective medical treatment can in part be explained by the relatively weak understanding of the pathophysiology of NAFLD.

1.2. NAFLD study groups in the world

In the United States and other countries numerous groups of researchers specifically study the pathogenesis of NAFLD. Even though the research methods differ; the overall goal is to identify genes involved in the pathogenesis of NAFLD. One important driving factor in NAFLD research is genome (or exome) wide association studies (GWAS/EWAS). These studies look for correlations between liver lipids and specific single nucleotide polymorphisms (SNP). For the moment, the largest NAFLD study group is the Genetics of Obesity-Related Liver Disease Consortium (GOLD), which includes several major independent research groups such as the Framingham Heart Study (FRAM), the old order Amish study, the Family Heart Study (FamHS) and The Age, Gene/Environment Susceptibility- Reykjavik Study (AGES) [34]. This consortium has a large sample size, which makes it possible to identify associations for relatively rare genetic variants in the population. Several key discoveries of genes involved in NAFLD have been made in the Dallas Heart Study (DHS). The DHS is a prospective, multi-ethnic study of the general population in Dallas, Texas. In the year 2008 liver TG levels were measured for approximately 2700 participants from the Dallas area by using the most accurate non-invasive method, magnetic resonance spectroscopy. Two GWAS/exome association studies have been performed in this cohort [28]. There are also several studies that specifically collected patients with liver fibrosis/cirrhosis patients.

1.3. Genetic aspects of NAFLD

As mentioned above, heritable factors play an important role in NAFLD [12, 35]. Methods to study the genetics of NAFLD vary. A cheaper and therefore more frequently used method is the study of families affected by NAFLD. The cornerstone of this approach is to identify families in which NAFLD is present in several generations, especially NAFLD with complications such as cirrhosis or HCC. After exclusion of other causes of liver damage and collection of genetic material, the next step is to choose the best available method to detect a genetic determinant of the disease. Further investigations can be divided into:

1) Hypothesis-driven research – the comparison of one or more specific mutation(s) in one or more candidate genes between individuals with and without the disease of interest. The most common method is to select genes that have been hypothesized to play a role in the disease, for example, SOD2 for NAFLD. The exons of the gene can be sequenced and presence or absence of specific mutations can be compared between patients and healthy persons. In the simplest cases,

the sequencing is done only for particular, already known genes or mutations. The main problem in such studies is the restriction to work on already known genes or SNPs; it is very difficult to identify new causal genes.

2) Hypothesis-free (or -generating) studies, which rely on genotyping of SNPs across the genome, or on sequencing of the exome or genome.

Common limitations of family studies include a limited sample size, too few family members and/or generations accessible, and the complicating issue of reduced genetic penetrance. Last but not at least there are problems directly related to the interpretation of genetic variants/mutations, such as difficulties to interpret the clinical significance of a particular mutation, especially in genes of unknown function. Also, mutations can be very rare (family-specific, or even personal), making it difficult to interpret their effect specific to the family. In that case, additional laboratory studies with cell cultures and animals can sometimes be helpful. Another approach is to look for the same mutations in the same gene in other families with a similar disease phenotype [36].

Problems with family studies of NAFLD genetics stimulated the search for alternatives. For example, studies in large groups of patients with certain diseases (e.g., patients with NAFLD), and a control group (patients without NAFLD). Similarly, as in the family studies, those studies can be hypothesis-driven (studying only possible candidate genes or single SNPs), or hypothesis-free (or -generating) studies such as genome-wide association studies (GWAS). GWAS have the potential to identify numerous variants associated with the disease or trait of interest. Unfortunately, GWAS-identified variants rarely cause a change in protein function (for example, a protein truncation, a codon frameshift or a damage of protein`s catalytic site or any other important part of the protein). Again, as in the case of family studies of NAFLD, understanding the mechanism underlying GWAS associated SNPs often requires additional research in cells/animals. In addition, GWAS-associated variants usually have to be replicated in a different study cohort; this helps to confirm or refute the significance of the identified associations.

1.3.1. PNPLA3 I148M polymorphism and NAFLD

The first genetic variant discovered to be associated with NAFLD in the general population is the I148M polymorphism in patatin like phospholipase domain-containing protein 3

(PNPLA3) [28]. Later studies repeatedly confirmed the association of this gene variant not only with hepatic steatosis and NASH [37, 38], but also with liver cirrhosis [39] and hepatocellular carcinoma [40]. In addition to the association with NAFLD, the PNPLA3 I148M variant also shows a significant correlation with alcoholic fatty liver disease (AFLD) [41-43]. PNPLA3 is a lipase, an enzyme which digests lipids. It is upregulated in the refed state, which is opposite to most other lipases that are typically expressed in the fasting condition [44]. In humans, PNPLA3 is highly expressed in the liver, less in white adipose tissue, and localizes on the surface of cytosolic lipid droplets. The expression of the PNPLA3 protein is regulated by the nuclear receptor SREBP1c [44, 45]. Despite the use of several animal and cell models [46-48], it is still unknown why the PNPLA3 I148M variant causes NAFLD. At this, time there are several hypotheses: changes in retinol (vitamin A) metabolism in hepatic stellate cells [21, 49, 50], disturbances in the metabolism of phospholipids [51], or the accumulation of the Pnpla3 I148M variant on the surface of lipid droplets [47].

1.3.2. TM6SF2 E167K polymorphism and NAFLD

Several years before the finding that the TM6SF2 variant E167K associates with NAFLD, it had been shown in several GWAS that genetic variation in the NCAN gene region is associated with fatty liver [34, 52, 53]. However, it was not until the year 2014 that the causal mutation at the NCAN gene region was discovered to be TM6SF2 E167K [2].

This thesis presents mechanistic data from several animal and cell models on TM6SF2 E167K and PNPLA3 I148M polymorphism and their association with NAFLD.

1.3.3. Other known polymorphisms associated with NAFLD

Several polymorphisms (SNPs) have been associated with NAFLD in human populations [54]. Polymorphisms in the genes GCKR (rs780094), PPP1R3B (rs4240624), LYPLAL1 (rs12137855) and TRIB1 (rs2954021) have been associated with fatty liver. The latest discovery is a SNP adjacent to the gene MBOAT7 (rs641738) that was found to be associated with both alcoholic fatty liver disease (ATAS) [55] and NAFLD [56-58]. Many of these polymorphisms are in genes that regulate the metabolism of carbohydrates or phospholipids, but may also act on lipid droplets or secretion of VLDL.

Also, it should be noted that polymorphisms that influence NAFLD have been identified in other animals, e.g. mice [59] and fish [60].

1.4.1. Animal models of NAFLD

During the last 2 decades animal models have become increasingly used to characterize the molecular mechanisms of candidate genes associated with human disease. While GWAS and similar methods can be used to identify associations between SNPs and NAFLD in the general population, further studies are required to determine the underlying gene/protein/SNP function. Studies on humans are expensive, often very complex and limited by ethical concerns. Studies on monkeys, dogs or similar large animals are expensive, and frequently cause conflicts with animal rights activists. In recent years the main animal model for NAFLD studies has been rodents (mice, or less commonly, rats) or rarely fish (zebrafish). Reviews of mice NAFLD models have been published recently[18, 61].

1.4. NAFLD animal models

1.4.1. Diet models

Numerous studies have used specific diets to cause NAFLD in animals. The aim of these studies is to mimic human fatty liver histologically or biochemically in an animal model. The largest part of these diets is designed for rodents and consists of high amounts of triglycerides, cholesterol and/or carbohydrates. The most commonly used diets are high-fat diet, high sucrose diet, Western style diets (with increased amount of triglycerides and cholesterol) and combined diets, such as high sucrose/fat diets [62].

The use of these rodent diets causes obesity, hyperinsulinemia, insulin resistance and severe fatty liver [47]. Even though the mice fed with these diets develop massive lipid accumulation in the liver, the next stages of NAFLD –inflammation, fibrosis, cirrhosis or liver cancer develop only after long-term use (up to half of the lifetime of rodent). This could be due not only to metabolic characteristics of rodents, but also to the intestinal flora or non-natural diet composition. A high-sucrose diet is particularly useful for promoting NAFLD in animal models. In a high-sucrose diet, more than 60% of calories come from sucrose, and the diet contains almost no lipids. The major advantage of this diet is its ability to cause accumulation of fat in the rodent liver (without inflammation or fibrosis) in 4 or even 2 weeks, without causing obesity or insulin resistance [47]. Since the diet has a minimal amount of fat, it is not recommended to be used for more than eight weeks, as animals suffer from severe skin lesions (author's personal observation).

Initial studies of the so-called MCD diet (methionine/choline deficiency diet) created a lot of enthusiasm [63]. Two weeks after starting the diet, rodents develop a severe accumulation of lipids in the liver followed by inflammation and fibrosis [64]. Unfortunately, by using this diet, mice markedly lose weight and after 12 weeks of eating this diet, mice become prominently dystrophic. In recent years, attempts have been made to return to using MCD diets by using a combination of high-fat/choline deficiency diets [65]. However, since it is unknown how the methionine/choline deficiency causes NAFLD in rodents and how well this model recapitulates human NAFLD, most of the NAFLD research laboratories try to avoid the use of this diet.

1.4.2. Genetic models of animals

One of the most widely used animal genetic models of NAFLD is leptin deficient mice (ob/ob mice) [66]. Deficiency of leptin causes hyperphagia, with marked obesity, hyperinsulinemia and fat accumulation in the liver. Unfortunately, as with any hormonal deficiency, leptin deficiency does not recapitulate the pathophysiology of human NAFLD.

A potential model for NAFLD studies could be specific strains of mice (BXD60/RwwJ; KK/HLJ and possibly others), which have spontaneously occurring fat accumulation in the liver. Although these models do not fully reflect the human NAFLD changes in the liver, they can perhaps be useful for studying the mechanisms underlying the development of NASH and liver fibrosis [59].

2. Methods and materials

2.1. Dallas Heart Study (DHS-II) and replication studies

2.1.1. Study characteristics

Dallas Heart Study 1 (DHS-1) began in the year 2000 as a group of researchers at the University of Texas Southwestern Medical Center at Dallas decided to set up a multi-ancestry, population-based probability sample study of Dallas County residents. The main research direction was coronary heart disease, but the collected data were also available for other researchers [67]. In 2007 the study was converted from a cross-sectional to a longitudinal study. All participants were invited for a repeat evaluation (DHS-2). The sampling process was designed to obtain a probability sample of residents that was representative of the target population and to identify subjects to be recruited later for mechanistic substudies of hypertension and dyslipidemia. By power calculations for several such studies, the recruitment target was set at 3,000 subjects

completing all phases of the initial data collection. The DHS-2 cohort was augmented by voluntary participation of friends and spouses of the original participants. DHS was approved by the Institutional Review Board (IRB) of the University of Texas Southwestern Medical Center at Dallas, and all subjects provided written informed consent. Each participant completed a detailed staff-administered survey, including questions about socioeconomic status, medical history and medication use. Ancestry was self-reported. A total of 4,708 DHS participants (52% non-Hispanic African Americans, 29% non-Hispanic whites, 16% Hispanics and 3% other ancestries) who underwent clinical examination during DHS-1 or DHS-2 were included in the present analysis. DHS-2 inclusion criteria were: age 35-70 years, the subject was a participant of the DHS-1 and consent to participate in the study. The study involved an equal number of males and females and 50% Afro-Americans. At 2008 the DHS-2 NAFLD study group was completed (52% of the patients themselves described as African-American, 29% European Americans, 17% Hispanics, and 2% other) and enrolled 2.971 patients were involved, aged from 30 to 65 years [28].

Findings in DHS were replicated in two independent cohorts: the Dallas Biobank and the Copenhagen Study. The Dallas Biobank is a repository of DNA and plasma samples from individuals ascertained at various locations in north-central Texas. All participants were over 18 years of age and gave written informed consent for inclusion in the database. Individuals for this study were European Americans who completed a preventive medicine examination at the Cooper Clinic in Dallas, Texas between 2008 and 2012 after signing an informed consent that was approved by the Cooper Research Institute IRB.

The Copenhagen Study combines participants from two studies of the Danish population: the Copenhagen General Population Study (CGPS) and the Copenhagen City Heart Study (CCHS) [68]. CGPS and CCHS are prospective studies of the Danish general population initiated, respectively, in 2003 and 1976–1978, with ongoing enrollment. Individuals were selected by the National Danish Civil Registration System to reflect the adult Danish population aged 20 to 100+ years old. Data were obtained from a self-administered questionnaire reviewed together with an investigator on the day of attendance, a physical examination and blood samples including DNA extraction. Blood samples for DNA extraction and biochemical analyses were drawn on the day of enrollment in CGPS (2003–2011) and on the day of enrollment at the CCHS examinations in 1991–1994 and 2001–2003. Studies were approved by IRBs and Danish Ethical Committees and were conducted according to the Declaration of Helsinki. Written informed

consent was obtained from participants. All participants were white and of Danish descent, as determined by the National Danish Person Registration System. There was no overlap of individuals between the studies. We included 64,000 consecutive participants from CGPS and 9,532 participants from CCHS, yielding a total of 73,532 participants.

2.1.2. Clinical measurements

BMI was calculated as weight measured in kilograms divided by squared height in meters. HOMA-IR was calculated from fasting concentrations of glucose and insulin. Plasma levels of HDL-C and triglycerides were determined by β quantification, and LDL-C concentrations were estimated using the Friedewald equation [69]. Liver enzyme levels (ALT, AST and ALP) were determined by enzymatic assays.

2.1.3. Detection of liver triglycerides

Hepatic triglyceride content was measured with proton magnetic resonance spectroscopy (^{1}H-MRS) as previously described [13, 25]. Hepatic triglyceride measurements were available for a total of 2,815 participants (including 2,287 of the DHS-1 and 528 of the DHS-2 participants). Measurements of liver TG were obtained with subjects on a 1.5 Tesla Gyroscan INTERA whole body system (Philips Medical Systems). Details of liver ^{1}H-MRS measurements were described elsewhere [25]. Previously in the special patient study group with low risk of fatty liver, the "upper limit of normal" liver TG level was determined by ^{1}H-MRS. The 95th percentile of liver TG in these subjects was 5.56%, which corresponds to a hepatic triglyceride level of 55.6 mg/g. With this value as a cutoff, the prevalence of hepatic steatosis in Dallas County was estimated to be 33.6% [25].

2.2. Genotyping and quality control

Genomic DNA was extracted from circulating leukocytes. A total of 4,625 DHS-1 and DHS-2 participants were genotyped using the Illumina Infinium HumanExome BeadChip, which captured a total of 247,870 markers, including functional exonic variants (>90%), disease-associated tag markers from recently published GWAS, ancestry-informative markers and other markers. Genotypes were called using Illumina GenomeStudio software. Number quality control filters were applied to samples and variants before analysis. Individuals were excluded for the following reasons: a call rate of <99% (n = 25) or duplicate discordance (n = 1). Variants were excluded on the basis of a call rate of <99% (n = 1,795) or a deviation from Hardy-Weinberg

equilibrium in African Americans with P < 0.0001 (n = 221). Also, variants that were monomorphic in our study population (n = 61,947) or had a single heterozygote (n = 27,705) were removed from the analysis. The further filter of variants was done by minor allele frequency. After exclusions, a total of 156,202 variants and 4,591 individuals were available for analysis (including 138,374 variants in 2,735 individuals with measures of HTGC). Genotyping of the Dallas Biobank Sample and the Copenhagen Study was performed using a TaqMan assay (dbSNP rs58542926; Applied Biosystems, C_89463510_10).

2.3. Genetic association analysis

To account for possible population stratification, principal components of ancestry were computed by markers with a minor allele frequency of >0.1% in the combined sample using EIGENSTRAT20 software version 4.2. Each sequence variant was tested for association with liver TG using linear regression, assuming an additive genetic model (with genotypes coded as 0, 1 or 2) and adjusted for age, sex, BMI and the four leading principal components of ancestry. For very rare variants (<8 carriers), the analysis was performed using the Wilcoxon rank-sum test, without covariate adjustment. A power transformation (1/4) was applied to liver fat measurements to achieve approximate normality of the residuals. By the number of tests performed, a significance threshold was set at 1×10^{-7} for each variant to maintain a family-wise type I error rate of 5%. A quantile-quantile plot of $-\log^{10}$ P-values did not show systematic deviation from the expected distribution. In the follow-up and replication analyses, the association between the TM6SF2 variant encoding p.Glu167Lys and clinical traits were tested using linear regression with adjustment for age, sex, BMI and ancestry (where appropriate). For hepatic fat content, an additional model was tested adjusted for HOMA-IR and alcohol consumption. For variables with skewed distributions, a logarithm (BMI, ALT, AST, ALP, triglycerides) or square-root (LDL) transformation was applied before analysis to ensure that the residuals were approximately normal and had constant variance. All reported P values are two-sided.

2.4. Detection of a conservation of amino acid sequence

For approval or exclusion of each potentially causal protein-coding SNP, evolutionary conservation of the amino acid sequence of protein was analyzed. Protein sequences were compared using Clustal Omega's free online service (http://www.ebi.ac.uk/Tools/msa/clustalo/). Protein sequences were obtained from the NCBI Gene resource

(http://www.ncbi.nlm.nih.gov/gene/), trying to collect as complete as possible a representative collection of proteins - from humans to the birds, amphibians, fish, or even plants.

2.5. Detection of protein expression in tissues

To determine whether the protein of interest is expressed (assuming that protein levels correlate with RNA levels) in the liver and potentially involved in a liver metabolism, several databases were used. The first database was created from 8 human liver biopsies(obtained at the Department of Invasive Radiology, University of Texas Southwestern Medical Center at Dallas, Texas, USA), where mRNA (non-ribosomal RNA) expression levels were determined using RNA sequencing (RNAseq). RNA expression levels were normalized and compared using RPKM (Reads Per Kilobase of transcript per Million Mapped reads) units. Relative expression was determined using reference gene expression, such as RPLP0.

The second database was created from RNAseq data of mice liver samples. Mice were fasted/refed with a high sucrose diet as described in the Mice studies section. The murine liver was removed, frozen in liquid nitrogen and then stored at -80ºC until further use. RNA was isolated from 30 mg of a liver sample of each animal using an RNA isolation kit (RNeasy Plus Universal Kit, QIAGEN, Germany). Levels of RNA transcripts were determined by using RNA sequencing (RNAseq) [70]. RNA expression levels were normalized and compared using RPKM units. The relative expression was determined by comparing values with those of the reference gene (RPLP0). The last database was from another published study that shows the relative RNA expression levels in various mouse tissues [71].

By comparing these three databases, the potential NAFLD causing protein (containing a particular SNP) was selected.

2.6. Mice studies

Animals were maintained on a 12-h-light/12-h-dark cycle and fed ad libitum Teklad Rodent Diet 2016 (Harlan Teklad), a high-fat diet (D12451; 45% kcal from lard, Research Diets), or a high-sucrose diet (HSD) (no.901683; 74% kCal from sucrose; MP Biomedicals) with free access to water. All experimental protocols were approved by the University of Texas Southwestern Medical Center Institutional Animal Care and Research Committee. For the fasting-refeeding experiments, chow-fed male C57BL/6N mice (n=4/group, age 8 weeks) were

fasted for 18h (2:00 PM to 8:00 AM) then refed with a high-sucrose diet (HSD) (no. 901683, MP Biomedicals) for 6 h (8:00 AM to 2:00 PM) for 3 days. On the morning of the fourth day, mice in the fasted group were deprived of food for an additional 6 h (total: 24 h fast) and the re-fed group was given food (HSD) for 6 h. The two groups of mice were sacrificed simultaneously and the livers collected. For all other measurements, mice were fed ad libitum overnight and the food was withdrawn at 7.00 am. Blood and tissue samples were taken at 11.00 am, 4 h after food withdrawal.

Total body fat and lean body mass were determined by nuclear magnetic resonance (MRI) using a Minispec analyzer MQ10 (Bruker).

2.7. Tm6sf2 knockout mice generation and validation

Mouse embryonic stem (ES) cells in which one Tm6sf2 allele was inactivated were purchased from The Knockout Mouse Project (KOMP), Davis, CA. The cells were injected into C57BL/6N (B6N-Tyrc- Brd/BrdCrCrl, albino, Charles River Laboratories) early-stage mouse embryos and the embryos were implanted into the uteri of C57BL/6N albino female mice. Chimeric males generated from these embryos were selected, and those with the greatest proportion of mutant cells (assessed by coat color) were bred with C57BL/6N albino females (Charles River Laboratories). Black offspring were genotyped and backcrossed with C57BL/6N females. Offspring of heterozygous (Tm6sf2+/-) matings were used for all of the experiments described in this study. DNA extracted from the tails of mice was PCR amplified using two forward oligonucleotides: 5`- taattatgtctttctcccctatgccc-3` (WT) and 5`-ttatacgaagttatggtctgagctcg-3` (KO) and a reverse primer: 5`-ggcagaggcaggcaagtttc-3`. PCR products were size-fractionated on 2% agarose gels and visualized with ethidium bromide.

2.8. Pnpla3 I148M and S47A Knock-in Mice Generation and validation.

Pnpla3 I148M and S47A knockin mice were generated using the Velocigene method [72]. To introduce the 148M variant, targeting constructs were generated containing exon 3, in which the codon for isoleucine 148 (ATT) was replaced with a methionine codon (ATG). In another strain of mice, we substituted serine (TCG) for alanine (GCC) at codon 47 in exon 1. Each construct contained a Neomycin (Neo) cassette flanked by loxP sites in an intron. After selection, the Neo cassettes were removed by Cre-mediated recombination, leaving a single LoxP site. The 148M

KI mice were genotyped using PCR primers that flank the loxP site (5'-CCTAATCCCTCCTTCCTTCCG-3' and 5'-CAGTGTATCCAACAGACAGCAGGC-3'), producing a 232-basepair (bp) product from the wild-type allele and a 230-bp product for the KI allele. The loxP site contains a NheI restriction site, so digestion of the PCR fragment from the KI mice generated two fragments upon digestion with NheI: 143 bp and 87 bp fragments. The S47A KI mice were genotyped using the primers 5'-TCACCTGGCGTGGGATTTC-3' and 5'-GGACAAACCTTCTGCGGCTC-3', which generated a 168 bp product for the wild-type allele and an 182 bp fragment for the mutant allele. The PCR products were separated by electrophoresis on 2% agarose gels. The mice were on a mixed background (129/Sv and C57Bl/6J). Experiments were performed in F1 mice crossed with C57Bl/6J mice and littermates were used as controls. The eight exons of Pnpla3 were PCR-amplified individually from genomic DNA from two lines each of the I148M and S47A KI mice and the PCR products were sequenced by Sanger sequencing. Hepatic RNA from the same lines of mice was converted to cDNA by SuperScript III First- Strand Synthesis System (Invitrogen) and the coding region of Pnpla3 was amplified by PCR and subjected to Sanger sequencing. The expression levels of Pnpla3 in each line were determined by Real-time PCR.

2.9. Other experiments and measurements

Liver and Plasma Chemistries

Lipids were extracted from pieces of liver (95–125 mg) using the method of Folch and Lees [73]. Tissue levels of TG, cholesterol, cholesteryl ester, free cholesterol, and phosphatidylcholine were measured using enzymatic assays (Infinity, Thermo Electron Corp., and Wako Inc.) and normalized to sample weight. Serum levels of alanine transaminase (ALT), aspartate aminotransferase (AST), triglyceride (TG), cholesterol, albumin and glucose were measured using the Vitros 250 system (GMI Inc.). Serum levels of nonesterified fatty acids (NEFA) were measured by enzymatic assay using a commercial reagent (Wako Inc.). ELISA assays were used to quantify serum insulin (Crystal Chem Inc.) and plasma PCSK9 levels. For the PCSK9 assays, LumiNunc Maxisorp assay plates (NalgeNunc) were coated with polyclonal rabbit anti-mouse PCSK9 antibody (clone IgG-551C, Jay Horton, UT Southwestern) diluted to 5 μg/ml in 100 μl of Buffer A (100 mM NaCl, 20 mM Na_3PO_4, pH 7.5). After an overnight (ON) incubation at 4°C, plates were washed 3 X with 350 μl of PBST [PBS with Tween 20 (0.05%), pH 7.4] and then

incubated in 150 µl of 0.5% BSA in Buffer A for 1 h at RT. All subsequent plate washes were conducted on a BioTek ELv405 Plate Washer. Standards and antibodies were diluted in Buffer A plus 0.5% BSA. Duplicate plasma samples (1 µl) were applied to each well. A standard curve was constructed using purified mouse PCSK9 (Jay Horton lab, UT Southwestern) serially diluted to a final concentration of 0.39 to 50 ng/ml. Plates were shaken at 37°C for 2.5 h and then the samples were washed before adding 100 µl of rabbit anti-mouse PCSK9 polyclonal antibody (IgG-552C, biotinylated with EZ-Link Sulfo-NHS-Biotin kit, Pierce). After 2 h incubation at RT, the plate was washed again. A total of 100 µl of Avidin-HRP (1:40,000; Pierce) was added and incubated for 1 h at RT. After a final wash, 100 µl of SuperSignal ELISA Pico Substrate (Pierce) was added for 1 min. Luminescence was quantified using a ThermoFisher Luminex Luminometer. Linear regression analysis of the standard curve was used to determine concentrations of PCSK9.

To analyze the fatty acid composition of different lipid classes, pieces of liver (50–60 mg) were homogenized in 5 ml of chloroform/methanol (2:1, v/v) and 1 ml of NaCl. Homogenized samples were centrifuged at 4,000Xg for 10 min at room temperature, and the lower phases were collected. One-third of the extract was dried under N_2, and the residue was resuspended in 100 µl of chloroform. Triglycerides, phospholipid (phosphatidylcholine), cholesterol esters, and nonesterified fatty acids were separated by thin layer chromatography on silica plates using hexane/diethyl ether/acetic acid (120, 30 and 3 ml, respectively) as the mobile phase. Bands corresponding to each lipid class were visualized using iodine vapor, excised, and placed in 2 ml of methanol/benzene (4:1). Acetyl chloride (200 µl) was added as the sample was vortexed. The sample was incubated at 100 °C for 1 h and then allowed to cool to room temperature. Once at room temperature, 5 ml of 6% aqueous K_2CO_3 was added slowly, followed by the addition of 1.5 ml of benzene. Samples were mixed and centrifuged at 4,000 x g for 10 min. The upper phase was removed, dried under N_2, and reconstituted in 150µl of hexane. Fatty acids were measured by gas chromatography (GC) using a Hewlett Packard 6890 series GC system. The identities of the fatty acids were determined by comparing the retention times with fatty acid standards (GLC-744, NU-Chek Prep) using pentadecanoic acid (C15:0) as an internal standard. Fatty acid concentrations were calculated based on the area of the C15:0 peaks. Sterols were extracted from bile collected from the gall bladders of mice after a 4-h fast and measured using GC-MS. Eluted sterols were identified and quantified using electron ionization-MS in single-ion-monitoring mode, as described previously [74]. Individual sterols were identified using a pattern of three

ions. The sterol levels were quantified using ions with the following m/z: cholesterol, 458; campesterol, 382; sitosterol, 396; stigmasterol, 484.

RNA Expression

RNA was isolated from indicated tissues using the RNA isolation kit (RNeasy Plus Universal Kit, QIAGEN, Germany). mRNA levels were determined by Real-Time (RT) PCR as described [75]. Mouse 36B4 was used as internal control. RNA-Seq (RNA sequencing) was done on ribosomal RNA depleted total RNA as described [76]. Relative expression levels of TM6SF2 in human tissues were assayed in cDNA samples prepared from the RNA of 22 human tissues (Human Total RNA Master Panel II, Clontech) by RT-PCR using two oppositely oriented primers. Human RPLP0 mRNA was used as an internal control.

Western Blots

Liver tissue (50-80 mg) was homogenized in RIPA buffer [150 mM NaCl, 1.0% IGEPAL CA-630, 0.5% sodium deoxycholate, 0.1% SDS and 50 mM Tris, pH 8.0 supplemented with protease inhibitors (Protease Inhibitor Cocktail, Roche)]. Lysates were cleared by centrifugation (12,000g) for 15 min at 4^{o}C, mixed with 2x Western Blot Laemmli sample buffer (65.8 mM Tris, pH 6.8, 2.1% SDS, 26.3% (w/v) glycerol, 0.01% bromophenol blue, 5% (v/v) β-mercaptoethanol), and incubated at 37ºC for 1 h. Samples were size-fractionated by 10% SDS-PAGE and proteins transferred to a PVDF membrane (Immobilin-FL, Millipore). Membranes were incubated ON at 4ºC with antibodies against the following proteins (source, catalog # and dilution of the antibodies are given in parentheses): calnexin (Enzo; ADI-SPA-860-F; 1:5000), PLIN2 (Fitzgerald Industries Int.; 20RAP002; 1:5000), LDH (Abcam; ab134187; 1:1000), PNPLA3 (R&D Systems; AF5179; 1:1000), FASN (320D, rabbit polyclonal anti-mouse antibody, gift from Jay Horton, UT Southwestern; 1:5000), ATGL (Cell Signaling; 2439S; 1:1000), SAR1b (Abcam; ab155278; 1:5000), MTTP (BD Biosciences; 612022; 1:5000), Calnexin-N terminus (Novus; NB100- 1965; 1:1000), BiP (Cell Signaling; 3183S, 1:1000), GOS28 (Abcam; ab82961; 1:1000), PLTP (Novus; NB400-106, 1:500), LSD1 (Cell Signaling; #2184; 1:1000), ApoB (Abcam; ab20737, 1:1000), CGI-58 (Abnova; H00051099-M01; 1:10,000), guinea pig anti-mouse PLIN2 (Fitzgerald Industries International; 20R-AP002; 1:10,000), V5 (Invitrogen; R960-25; 1:5,000), β-actin (Cell Signaling Technology; #12262; 1:1,000), rabbit anti-mouse HSL

(Novus Biologicals; H00051099-M01; 1:1000) polyclonal Abs. Rabbit antisera directed against PLIN3; PLIN4, and PLIN5 were kindly provided by Dr. Perry Bickel, UT Southwestern.

A polyclonal rabbit anti-mouse TM6SF2 antibody (792D) and mouse monoclonal antibodies (mAb) (clones 8B3-B7 and 8B3-E1) were raised against a KLH-conjugated peptide corresponding to the C-terminal 15 amino acids of mouse TM6SF2 (CALPSSPQDQDKKQQ obtained from GenScript). We used a goat anti-mouse IgG, Fcγ fragment specific antibody (Jackson ImmunoResearch Inc, 115-035-071, 1:3000) to detect the primary anti-mouse TM6SF2 antibody. To detect ApoB, liver lysates were prepared by homogenizing tissue in modified RIPA buffer (150 mM NaCl, 1.0% IGEPAL CA- 630, 0.5% sodium deoxycholate, 1% SDS, 1 M urea and 50 mM Tris, pH 8.0) supplemented with protease inhibitors. Lysates were size-fractionated by SDS (5%)-PAGE and then transferred to a Hybond-C membrane (GE Healthcare) at 0.1 A in 4ºC overnight (14h). Membranes were incubated with an ApoB antibody ON at 4ºC. ApoB was detected in mouse plasma by immunoblotting. Plasma (10 µl) was added to 50 µl 2X Laemmli sample buffer and 40 µl RIPA buffer, and 2 µl of the mixture was size fractionated on 4-12% Tris-Bis gels (BioRad).

To assess the processing of SRE binding protein-1c (SREBP-1c), mice were metabolically synchronized for three days (fasting 8:00 AM till 8:00 PM; refeeding 8:00 PM till 8:00 AM) and sacrificed at the end of the last refeeding cycle. Livers were collected and stored at -80ºC. A total of 100 mg of liver was homogenized on ice in 1 ml of Homogenization Buffer [20 mM Tris-Cl, pH 7.4, 2 mM $MgCl_2$, 0.25 M sucrose, 10 mM EDTA , 10 mM EGTA, 5 mM DTT and 1 mM PMSF] supplemented with 5 X protease inhibitors using a T10 mechanical homogenizer (IKA) set at maximum speed. Lysates were centrifuged for 5 min at 3000g at 4ºC, and 800 µl of supernatant was collected (the post-nuclear supernatant). The remaining pellet and supernatant were maintained at 4^{o}C (the nuclear fraction). The post-nuclear supernatant was centrifuged at 100,000g for 30 min (47,100 rpm, Sorvall S120AT2 rotor), and the pellet collected (membrane fraction). The membrane fraction was dissolved in 400 µl of RIPA buffer and the protein concentration determined using a BCA kit (Pierce). The remaining supernatant was aspirated from the nuclear fraction, and the pellet was resuspended in 1 ml Homogenization Buffer and centrifuged for 5 min at 1000g at 4ºC. The supernatant was removed and the pellet resuspended in 300 µl of Buffer B [20 mM HEPES, pH 7.6, 2.5% glycerol (v/v), 0.42 M NaCl, 1.5 mM $MgCl_2$, 1 mM EDTA; 1 mM EGTA supplemented with 5 X protease inhibitor]. The

resuspended nuclei were rocked for 45 min at 4ºC, then centrifuged at 100,000g for 30 min (47,100 rpm, Sorvall S120AT2 rotor). The supernatant (nuclear extract) was collected and the protein concentration was determined. A total 20 μg of protein was mixed with 6x SDS sample buffer (0.375 M Tris pH 6.8, 12% SDS, 60% glycerol (v/v), 0.6 M DTT, 0.06% bromophenol blue) and heated at 95ºC for 8 min. Equal volumes of nuclear and membrane fractions from each group (n=4) were pooled and 40 μg of protein loaded on SDS (10%)-PAGE gel. The blot was incubated ON at 4ºC with rabbit anti-mouse SREBP-1c antibody that was developed in the Department of Molecular Genetics, UT Southwestern. After signal detection, membranes were stripped (Restore Western Blot Stripping Buffer, Thermo Scientific) and re-blotted with antibodies against calnexin and LSD1 to determine the purity and loading consistency of each fraction. The ECL or was detected and quantified using a Licor Odyssey Fc Imager.

Some of the protein blots were visualized using Odyssey Quantitative Fluorescent Imaging System (LI-COR Biosciences, Lincoln, NE). IR Dye 680 conjugated anti-guinea pig, anti-mouse, anti-rabbit and IR Dye 800 conjugated anti-goat, anti-rabbit and anti-mouse antibodies were used to visualize the bands on the immunoblots.

Fast protein liquid chromatography (FPLC) of blood plasma

To measure plasma lipid and lipoprotein levels, blood was collected from mice fasted for 4 h, plasma from 4-5 mice was pooled (total volume 400 μl) and size-fractionated on a fast protein liquid chromatography (FPLC) Superose 6 column (GE Healthcare). A total of 42 fractions (300 μl each) were collected, and cholesterol and TG content of each fraction were measured in each fraction using enzymatic assays (Infinity, Thermo Scientific).

Secretion of VLDL and ApoB

To measure rates of VLDL-TG secretion, 5 female mice (10 weeks old) of each genotype fasted for 4 h at the beginning of the light cycle. A bolus of Triton WR-1339 (Tyloxapol; Sigma-Aldrich) (500 mg/kg) was given via the tail vein [77]. Blood samples (~40 μl) were collected from the tail vein in EDTA-coated tubes (Microvette 500 KE, Sartedt) before and 30, 60, 90 and 120 min after the injection of detergent.

ApoB secretion rates were measured as described previously [78]. Briefly, each mouse was injected with 200 μCi of [35S]-methionine (1,175 Ci/mmol) (PerkinElmer Life Sciences) and Triton WR1339 (500 mg/kg). Blood was collected from the tail vein into aprotinin-containing EDTA-coated tubes, before and 45 and 90 min after the injection. Plasma from the 90 min time point was isolated, diluted 1:20 in PBS containing protease inhibitors, and incubated with deoxycholate (100 μl of 0.15% solution) for 10 min at RT. Trichloroacetic acid (TCA) (50 μl) was added and the mixture was incubated for 30 min on ice and then centrifuged at 10,000g for 15 min at 4ºC [79]. The supernatant was removed, and the pellet resuspended, washed twice with ice-cold acetone, and centrifuged at 10,000g for 15 min at 4ºC. Pellets were air dried, dissolved in buffer containing 2% SDS and 1 M urea, and heated to 37ºC for 1 h. Samples were mixed with 6X SDS sample buffer, heated for 5 min at 95°C, and size-fractionated by 5% SDS-PAGE. Gels were dried and exposed to X-Ray film (BIOMAX XAR) for 4 days at −80°C. Films were scanned using an HP Scanjet 5590 and quantified using Licor Image Studio Software.

Fat Tolerance and Lipid Absorption Test

Chow-fed mice were fasted 16 h and then gavaged with 10 μL corn oil/g mouse body weight. Blood samples were collected from the tail vein at the indicated times, and plasma TG levels were determined using a colorimetric assay (Infinity; Thermo Scientific).

To measure TG absorption, mice were fasted for 16 h before injection of Triton WR1339 (500 mg/kg) in the tail vein. After 30 min, the mice were gavaged with 50 μl corn oil (Sigma, C8267) supplemented with and 5 μCi 9,10-3H(N) ^{3}H-labeled oleic acid (54.5 Ci/mmol; Perkin Elmer) and 1 μCi ^{14}C-labeled cholesterol (50.8 mCi/mmol; Perkin Elmer). Blood was collected from the tail vein before and 1, 2, 3, 4 and 7 h after Triton WR1339 injection and plasma levels of TG was measured by enzymatic assay. Plasma samples (10 μl) were subjected to scintillation counting to quantitate the radioactivity.

Subcellular Fractionation

Female C57BL/6N mice (age 12 weeks) were fed an HSD for 2 weeks. Mice fasted for 18 h and refed with HSD for 6 h before being sacrificed. Lipid droplets and lipid droplet proteins were isolated from the liver as previously described [47]. Livers were maintained at 4^{o}C on ice in 10 ml of Buffer C [250 mM sucrose, 20 mM Tricine, pH 7.8] plus protease inhibitors and then

homogenized using a Dounce-type glass homogenizer. Lysates were centrifuged at 17,738g and the lipid droplet fractions were collected from the top of the tube. Pellets and supernatants were resuspended in 10 ml Buffer C and centrifuged at 1000g for 10 min at 4ºC. The supernatant (post-nuclear fraction) was collected and 1 ml aliquot was diluted in 9 ml of buffer C and then subjected to ultracentrifugation in a TH-641 swinging bucket rotor at 100,000g for 45 min at 4ºC. The supernatants (cytosol fraction) were collected, and the pellets washed with PBS and then solubilized in 0.5 ml of RIPA buffer. The cytosol was concentrated to a volume of 200 µl using Amicon Ultra-15 centrifuge filters (Millipore; UFC901096) and then diluted with RIPA buffer. Laemmli Sample buffer 2X (Biorad) was added and the samples were heated at 37ºC for 1 h. Proteins from the membrane and cytosolic fractions (50 µg each) and one-tenth of the lipid droplet proteins were size fractionated by SDS(10%)-PAGE and immunoblotted overnight at 4ºC with antibodies to calnexin, lactate dehydrogenase, perilipin 2, PNPLA3, and a mouse anti-TM6SF2 mAb (clone 8B3-B7, 5 µg/ml).

Subcellular fractionation of cultured cells

Hepa1c1c7 cells (obtained from the University of California, Los Angeles) were transfected with pCMV6-XL5 (OriGene Technologies) expressing V5-tagged wild-type human TM6SF2 using Lipofectamine 2000 (Life Technologies). Cells, which tested negative for mycoplasma, were grown in DMEM supplemented with 10% FCS and 400 µM oleate conjugated to albumin. After 48 h, cells were washed three times with ice-cold PBS and disrupted in Buffer J (20 mm Tris, 1 mm EDTA, pH 7.4) plus protease inhibitors (Protease Inhibitor Cocktail, Roche), homogenates were passaged through a 25-gauge needle 20 times, and lysates were centrifuged at 4,000g for 10 min at 4 °C to pellet nuclei and unbroken cells. Postnuclear supernatants were adjusted to a concentration of Buffer J plus 20% sucrose and added to the bottom of an ultracentrifuge tube. A step gradient was constructed by overlaying the postnuclear supernatant with 5 ml of Buffer J plus 5% sucrose followed by Buffer J alone. The tube was clamped and centrifuged at 28,000 xG for 40 min at 4 °C. After centrifugation, the lipid layer was isolated by slicing the tube, combined with 20 volumes of −80 °C acetone and stored overnight at −20 °C. Precipitated proteins were pelleted by centrifugation at 4,000g for 1 h at 4 °C and resuspended in Buffer J plus 1% SDS. Lipid droplet, postnuclear supernatant (cytosol) and membrane fractions were subjected to immunoblotting using an antibody to V5 tag and ADRP. Antibodies against calnexin and β-actin were used as internal controls.

Preparation of mice liver microsomes and affinity chromatography of ER and Golgi

ER and Golgi fractions were isolated by affinity chromatography as described [80]. Magnetic beads (2 ml Dynabeads M-450 Epoxy, Invitrogen) were washed five times with ice-cold buffer D (0.1 M phosphate buffer, pH 7.9) and pelleted with a Dynal Magnetic bead separator (Invitrogen). Goat anti-rabbit IgG (500 μg, Jackson ImmunoResearch Inc.) in 1.5 ml of Buffer D was added to the washed beads and slowly rotated for 72 h at 4ºC. Beads were washed five times again with ice-cold Buffer D and divided into two equal portions. The samples were then resuspended in 1.5 ml of Buffer D. To capture the ER, one-half of the beads were incubated with rabbit anti-calnexin antibody (90 μg) (Enzo Life Sciences; ADI-SPA-860-F). To capture the Golgi, the remaining portion of washed beads was added to 20 μg of rabbit anti-GM130 antibody (GeneTex Inc; GTX61445) and rotated for 24 h at 4º C. Beads were then washed 5 times with Buffer D, and again divided into two parts (for WT and KO samples) in preparation for incubation with mouse liver microsomes.

To prepare liver microsomes, mice were entrained to a synchronized regimen for three days by fasting from 8:00 AM to 8:00 PM and refeeding with chow overnight. At the end of the final 12 h refeeding period, the mice were sacrificed and their livers were removed and kept on ice throughout the procedure. Pieces of liver (600 mg) were disrupted in 3.0 ml of Buffer E (10 mM HEPES, pH 7.4, 10 mM KCl, 1.5 mM $MgCl_2$; 5 mM sodium EDTA, 5 mM sodium EGTA, 250 mM sucrose, supplemented with protease inhibitors) using a T10 mechanical homogenizer (IKA) at maximal speed, then passed through a syringe 20X with a 20 G needle. The homogenate was centrifuged at 2,000g for 10 min at 4°C. The postnuclear supernatant was collected and centrifuged (21,000g) at 4ºC for 30 min. One-third of the pellet was resuspended in RIPA buffer. The remainder was resuspended in buffer containing 1.4 ml of Buffer D and 0.7 ml Buffer E plus protease inhibitors. A total of 1.0 ml of the mixture was slowly rotated in the presence of antibody-coated beads for 24 h at 4ºC. The supernatant was then removed, and the beads were washed 10X with BSA (5%) in PBS with protease inhibitors, then another 10X in 0.1% BSA in PBS and finally 10X with PBS alone. Washed beads were incubated in 200 μl of buffer containing 2% SDS, 1 M urea plus 50 μl of 6X SDS sample buffer. Samples were heated for 1 h at 37ºC and then the supernatants were collected and subjected to SDS (10%)-PAGE.

Expression of TM6SF2 in cultured hepatocytes

The cDNA for human TM6SF2 was cloned downstream of the cytomegalovirus promoter/enhancer elements in the pCMV6-XL5 plasmid (OriGene Technologies). Single-nucleotide changes were introduced using the QuikChange site-directed mutagenesis kit (Stratagene) and confirmed by Sanger sequencing. A sequence encoding a C-terminal V5 epitope tag (GKPIPNPLLGLDST) was placed in each plasmid construct. Human hepatoma (HuH-7) cells, which tested negative for mycoplasma, were obtained from Sigma and were transfected with plasmids expressing wild-type or mutant human TM6SF2-V5 using Lipofectamine 2000 (Life Technologies). HuH-7 cells were grown to 90% confluence in DMEM supplemented with 10% FCS plus 100 IU/ml penicillin and 100 μg/ml streptomycin for 48 h. Cells were then washed three times with ice-cold PBS, collected and divided into two aliquots. One aliquot was used to determine TM6SF2 mRNA levels by RT-PCR. The other aliquot was suspended in RIPA buffer (150 mM NaCl, 1.0% IGEPAL CA-630, 0.5% sodium deoxycholate, 0.1% SDS and 50 mM Tris, pH 8.0) supplemented with protease inhibitors (Protease Inhibitor Cocktail, Roche) and disrupted by 20 passages through a 25-gauge needle. Homogenates were centrifuged at 12,000g for 10 min at 4 °C to pellet nuclei and unbroken cells. Postnuclear supernatant protein concentrations were determined by BCA Protein Assay (Thermo Scientific). Identical amounts (25 μg) of protein were size fractionated on 10% SDS-PAGE gels and immunoblotted using an antibody to V5 tag (R960-25, Invitrogen). Proteins were visualized using Super-Signal ECL (Pierce Biotechnology).

Immunofluorescence Microscopy

Mice were anesthetized with isoflurane, and a catheter was inserted through the inferior vena cava. The liver was perfused (3 ml/min) with 30 ml of Liver Perfusion Medium (Invitrogen Inc., 17701-38) and then with 30 ml Liver Digestion Medium (Invitrogen Inc. 17701-34). Livers were excised and placed in 100 mm plastic dishes containing 20 ml of Liver Digestion Medium preheated to 37°C. The surface capsule of the liver was removed, and the hepatocytes gently dispersed. Cells were filtered through a 100 μM Nylon Mesh Cell Strainer (Falcon, Corning) and then transferred to a conical tube containing 20 ml Washing Medium [DMEM plus glucose (1000 mg/L), FCS (5%) plus 50 U/ml penicillin and 50 mg/ml streptomycin] and centrifuged at 40g at 4°C for 5 min. Supernatants were discarded and cells were resuspended in Washing Medium. Cells were seeded at a density of $0.5x10^6$ cells per well on 6- well plates containing coverslips

(Fisher) precoated with rat tail collagen (Enzo Lifesciences). After 3 h the cells were washed with PBS, fixed with 4% paraformaldehyde in PBS, and quenched in 50 mM ammonium chloride for 15 min. Cells were then permeabilized for 2 min in 0.1% (vol/vol) Triton X -100 in blocking buffer [0.4% fish skin gelatin (Sigma, G7765) in PBS]. Cells were incubated with primary antibody in blocking buffer for 1 h and with secondary antibody for 30 min at RT. Coverslips were mounted on slides using Vectashield mounting medium with DAPI (Vector Laboratories). Slides were viewed under Leica TCS-SP5 using a 63X/1.30 numerical aperture oil immersion objective. All images were exported as tiff files and compiled using Image J software. TM6SF2 was detected using a mouse mAb (clone 8B3-B7) generated as described above. The following commercial antibodies were used: rabbit polyclonal antibodies against giantin and calnexin from Abcam (Cambridge, MA), rabbit polyclonal antibody against PLIN2 from Thermo Scientific (Grand Island, NY), rabbit mAb against RCAS1 from Cell Signaling Technology (Beverly, MA). BODIPY 493/503 dye was purchased from Life Technologies (Grand Island, NY). All other reagents were purchased from Sigma-Aldrich (St. Louis, MO).

Tissue Morphology

For Oil Red O staining, livers were fixed in 4% paraformaldehyde for 24 h and then equilibrated in 10% and then 18% sucrose (all in PBS) for 24 h each at 4°C. Tissue was cryosectioned, brought to RT and air-dried for 2 h and then fixed in methanol-free paraformaldehyde (4%). Slides were washed three times with H_2O and then incubated for 10 min in 0.18% Oil Red O (Sigma) prepared in 60% isopropyl alcohol. Slides were washed five times in H_2O. Nuclei were counterstained with hematoxylin and the coverslips were affixed with aqueous mounting medium (Vector Laboratories). Sections were visualized using a Leica microscope (DM2000) at 20x, 40x and 63x magnification. The numbers and sizes of lipid droplets (original magnification X63) were determined in five randomly selected images from each slide (n = 4 mice/ group) using ImageJ (National Institutes of Health).

Electron Microscopy

To image very low density lipoprotein (VLDL) using electron microscopy, 400 µl of plasma was obtained from TMm6sf2 T and KO mice after a 4-h fast. Plasma was placed at the bottom of an ultracentrifuge tube (1 ml) and overlaid with 600 µl of KBr (1.006 g/ml) solution. Samples were centrifuged at 434,000 xG for 2 h at 16 °C in an S120-AT2 rotor. A total of 200 µl from the top

of the tube was collected (VLDL fraction). A 5-µl aliquot of the VLDL fraction was applied to a carbon-coated copper grid and stained with 2.0% uranyl acetate for 15 min [19]. Grids were visualized with a Tecnai G2 Spirit transmission electron microscope (FEI). Images of VLDL were recorded on a Gatan 2,000 x 2,000 CCD camera, and the size distribution of the particles was determined using ImageJ (10 random images/group).

Isolation of lipid droplets

Livers were collected from isofluorane-anaesthetized mice and washed in ice-cold PBS. The livers were maintained at 4^{o}C on ice in 10 ml of Buffer F (250 mM sucrose, 20 mM Tricine, pH 7.8) plus protease inhibitors (Complete Mini EDTA-free Protease Inhibitor Cocktail, Roche) and then homogenized using a Dounce-type glass-glass homogenizer. After the samples were centrifuged at 100 g for 30 min at 4°C to remove tissue debris, the supernatant (~7 ml) was collected and transferred to an ultracentrifuge tube overlayed with Buffer G (100 mM KCl, 2 mM $MgCl_2$, 20 mM HEPES, pH 7.4) plus protease inhibitors (Roche), and centrifuged at 17738g for 60 min at 4°C in a swinging rotor (TH-641 Sorvall) as described [81]. The lipids at the top of the tube were transferred to a microfuge tube and centrifuged at 13,000g for 5 min at 4°C. The aqueous phase was removed, and the lipid layer was washed twice with Buffer G, resuspended in Buffer H (150 mM NaCl, 20% (wt/vol) glycerol, 50 mM Tris-HCl, pH=7.4) and centrifuged at 13,000g for 5 min at 4°C. The buffer was removed, and the lipid droplets were suspended in cold (-20°C) acetone and kept at -20°C overnight. Samples were centrifuged at 15,000g for 30 min at 4°C, and the pellet was resuspended in a 1:1 mixture of acetone and diethyl ether (vol/vol). The samples were washed once in acetone: diethyl ether and once in diethyl ether [82]. After the second wash, the supernatant was discarded and the pellet dried under a stream of N_2 for 30 min. The dried pellet was dissolved in Buffer I (2% SDS and 1M urea in PBS) at 37°C for 1h. Protein concentrations were determined using BCA Protein Assay kit (Pierce, Rockford, IL).

Adeno-associated viral knockdown of Tm6sf2 in mice

Ten shRNA sequences (shRNA1–shRNA10) to specifically target mouse Tm6sf2 were designed using a proprietary algorithm. Synthesized 19-nt sequences starting from positions 604, 508, 470, 616, 855, 899, 247, 210, 977 and 323 of Tm6sf2 as complementary antiparallel oligonucleotides with a loop sequence (TTCAAGAGA) and BbsI overhang. Forward and reverse oligonucleotides were annealed and ligated into the BbsI site of AAV cis plasmid downstream of

the human H1 promoter (pAAV-shRNA). To evaluate the potency of these shRNAs, the ten pAAV-shRNAs were transfected into a CHO cell line that stably expresses mouse Tm6sf2 using Nucleofector (Lonza) per the manufacturer's instructions. AAV cis plasmid containing the H1 promoter and termination sequences was used as a negative control. After 48 h, cells were collected, and RNA was isolated and prepared for RT-PCR analysis. Relative mRNA expression levels were normalized to the expression of Actb (encoding β-actin). Primers and probe sets for mouse Tm6sf2 and Actb were purchased from Applied Biosystems. The two most potent shRNAs (shRNA5 and shRNA8) were selected for AAV vector production. AAV8 mTm6sf2-shRNAs and the negative control AAV vector containing the H1 promoter and termination sequences were produced by the helper-free triple plasmid transfection method as described [83]. The pAAV cis plasmid, an adenoviral helper plasmid and a chimeric packaging plasmid containing the AAV2 Rep gene and AAV8 Cap gene were cotransfected into HEK293 cells. AAV vectors were subsequently purified by two rounds of cesium chloride density gradient ultracentrifugation, and titers were determined via RT-PCR analysis. Eight-week-old male C57BL/6J mice (Jackson Laboratory) were injected via the tail vein with 1×10^{11} genome copies (GC) of control AAV8 or AAV8-shRNA5. Mice were maintained on a 12-h light/12-h dark cycle and fed the Teklad Mouse/Rat Diet 7001 chow diet (Harlan Teklad) ad libitum. After 14 d, mice were fasted for 4 h at the end of the dark cycle and then sacrificed. Tissues and plasma were collected and frozen at −80 °C.

Glucose tolerance testing

Chow, high sucrose or high-fat fed mice were fasted for 4 hours before intraperitoneal injection of glucose (1.5 g/kg). Plasma levels of glucose were monitored at the indicated times from tail using glucose strips (Contour, Bayer, Pittsburgh, PA).

3. Results

The Results consists of a compilation of several of the author`s original publications. The author's contribution to the enclosed publications is presented.

Original Paper I (presented in methods, results and discussions)

Kozlitina J.*, Smagris E.*, et al., Exome-wide association study identifies a TM6SF2 variant that confers susceptibility to nonalcoholic fatty liver disease. Nat Genet, 2014. 46(4): p. 352-6.A.

Contribution: Performed most of the experiments, helped to detect causal variant, made figures and contributed to writing and revising the manuscript.

*These authors contributed equally to this work.

Original Paper II (presented in methods, results and discussions)

Smagris, E., et al., Inactivation of Tm6sf2, a Gene Defective in Fatty Liver Disease, Impairs Lipidation but Not Secretion of Very Low Density Lipoproteins. J Biol Chem, 2016. 291(20): p. 10659-76.

Contribution: Performed most of the experimental work and data analysis, made figures and contributed to writing and revising the manuscript.

Original Paper III (presented in methods, results and discussions)

Smagris, E., et al., Pnpla3 I148M knockin mice accumulate PNPLA3 on lipid droplets and develop hepatic steatosis. Hepatology, 2015. 61(1): p. 108-18.

Contribution: Performed most of the experimental work, data analysis, made figures, partly wrote the manuscript.

3.1. TM6SF2 study

Previously, our study group found a missense variant (encoding p.Ile148Met) in patatin-like phospholipase domain–containing 3 (PNPLA3) that is strongly associated with hepatic

triglyceride content (HTGC) and with serum levels of alanine transaminase (ALT) [28]. Subsequent genome-wide association studies (GWAS) found other common SNPs associated with liver fat content [34] and levels of circulating liver enzymes [37, 84]. To identify the functional variants at these loci, we used genotyping arrays (HumanExome BeadChip, Illumina) to perform an exome-wide association study in a multi-ancestry, population-based study, the Dallas Heart Study (DHS)[67]. A total of 138,374 sequence variants that were polymorphic and passed our quality control criteria were tested for association with hepatic triglycerides (HTGC) in 2,736 DHS participants (1,324 non- Hispanic African Americans, 882 non-Hispanic, European Americans, 467 Hispanics and 63 other ancestry groups), with adjustment for age, sex, ancestry and body mass index (BMI). Two sequence variants in PNPLA3 (rs738409 and rs2281135) had the lowest association P values (4.0×10^{-16} and 6.9×10^{-12}, respectively), and the next most significant variant (rs58542926) was in TM6SF2 ($P = 5.7\times10^{-8}$) (Fig.1A).

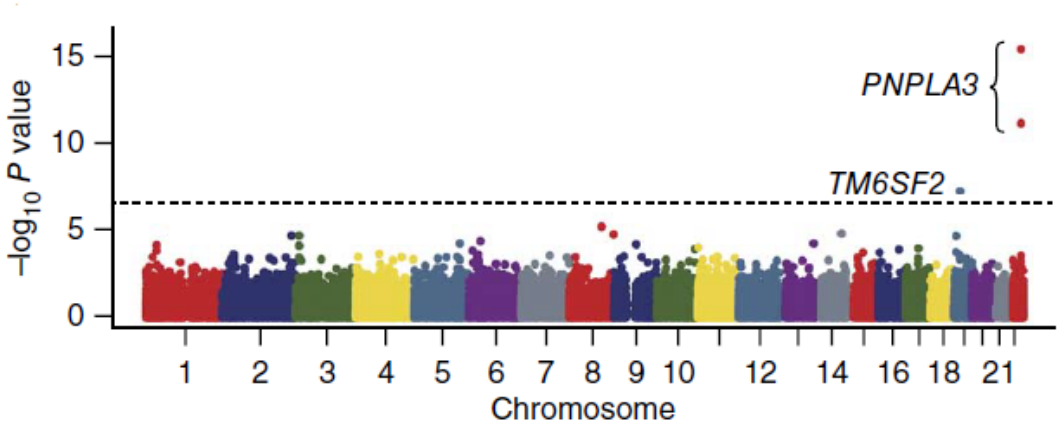


Figure 1A. Exome-wide association with HTGC (hepatic triglyceride content) in the Dallas Heart Study (DHS). *Manhattan plot showing the association of 138,374 sequence variants on the HumanExome Array (BeadChip, Illumina) with HTGC in DHS (n = 2,736). The dashed line denotes the Bonferroni-corrected significance threshold.*

No other variants exceeded the exome-wide significance threshold of 3.6×10^{-7}. After excluding these three SNPs, the quantile-quantile plot of P-values showed no systematic deviation from the expected null distribution (Fig.1B).

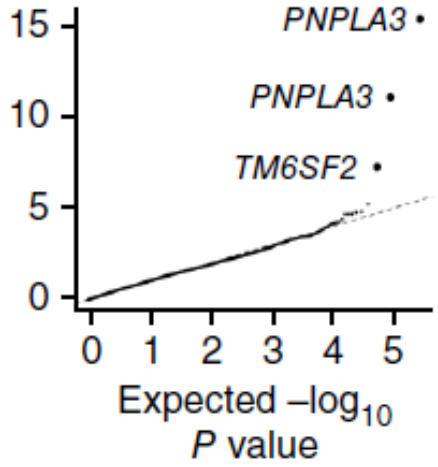


Figure 1B. The quantile-quantile plot of P values showed no systematic deviation from the expected null distribution. *Quantile-quantile plot of $-log_{10}$ P-values.*

The TM6SF2 variant was not associated with other risk factors for hepatic steatosis, including BMI, homeostatic model assessment of insulin resistance (HOMA-IR) or alcohol intake (Table 1).

Characteristic	EE[a]	EK	KK	P-value[b]
Total n	4,151	423	13	-
Female, n (%)	2,389 (57.6)	225 (53.2)	9 (69.2)	0.16
Ethnicity[c], n (%)				
Non-Hispanic African-American	2,203 (53.1)	160 (37.8)	2 (15.4)	2.1×10^{-10}
Non-Hispanic European-American	1,169 (28.2)	174 (41.1)	11 (84.6)	3.9×10^{-11}
Hispanic	676 (16.3)	69 (16.3)	0 (0)	0.35
Other	103 (2.5)	20 (4.7)	0 (0)	0.029
Age, years (mean)	45 ± 11	44 ± 11	50 ± 12	0.73
Body-mass index, kg/m^2	29.6 (26-35)	28.5 (26-34)	31.8 (28-34)	0.95
Diabetes[d], n (%)	523 (12.6)	37 (8.7)	2 (16.7)	0.14
Alcohol consumption[e] >20 g/day, n (%)	334 (8.1)	48 (11.4)	1 (7.7)	0.058
Lipid-lowering medication (statins), n (%)	359 (8.8)	33 (8.0)	0 (0)	0.17
HDL-cholesterol (mg/dL)	50.6 ± 14.9	49.9 ± 14.0	50.2 ± 18.8	0.98
HOMA-IR (U)[f]	3.0 (2-5)	2.9 (2-5)	4.6 (2-8)	0.52
Fasting glucose (mg/dL)[g]	91 (84-98)	91 (85-99)	94 (86-104)	0.15
Gamma-glutamyl transpeptidate (IU/L)	23 (17-36)	24 (16-38)	22 (12-28)	0.51
Bilirubin (mg/dL)	0.5 (0.4-0.7)	0.5 (0.4-0.7)	0.6 (0.5-0.8)	0.92
HTGC (%)[h]	3.5 (1.9-6.7)	4.4 (2.2-9.9)	15.7 (8.8-18.4)	5.6×10^{-7}

Table 1. Characteristics of Dallas Heart Study*. Participants stratified by TM6SF2 (E167K) genotype. Values are median (interquartile range), mean ± s.d., or number (%).* [a]*Missense variant (E167K) in TM6SF2.* [b]*P-values were calculated using linear regression for continuous variables (age, body-mass index, glucose, HOMA-IR) and logistic regression for categorical variables (gender, ethnicity, diabetes diagnosis, alcohol consumption, and use of lipid-lowering*

drugs). Regression models were adjusted for age, gender, ethnicity, and BMI (where appropriate). [c]*Ethnicity was self-reported as described in the Methods.* [d]*Diabetes was defined as a self-reported diagnosis of diabetes, use of prescription medication, fasting glucose ≥ 126 mg/dL or nonfasting glucose ≥ 200 mg/dL or HbA1c ≥ 6.5%.* [e]*Alcohol consumption was self-reported.* [f]*HOMA-IR, homeostatic model assessment of insulin resistance was calculated as described in the Methods.* [g]*Diabetics excluded.* [h]*P-value after adjustment for HOMA-IR, alcohol consumption and covariates listed above.*

The TM6SF2 variant associated with fatty liver is an adenine-to-guanine substitution in coding nucleotide 499, which replaces glutamate at residue 167 with lysine (c.499A>G; p.Glu167Lys). Glu167 is highly conserved among mammals and is also an acidic residue (aspartate) in birds (Fig. 1c).

Human	150-	**M**SI**LVFLT**G**NILGKYSS**E**IRPAFF**L**TIPYL**LVP**C**WAG**MKV**F**S**Q**PR	-194
Chimpanzee	150-	**M**SI**LVFLT**G**NILGKYSS**E**IRPAFF**L**TIPYL**LVP**C**WAG**MKV**F**S**Q**PR	-194
Monkey	150-	**M**SI**LVFLT**G**NILGKYSS**E**IRPAFF**L**TIPYL**LVP**C**WAG**VRV**F**S**Q**PR	-194
Mouse	150-	**M**SI**LVFLP**G**NILGKYSS**E**MRPTFF**L**AILYM**LVP**C**WAG**VRI**F**N**Q**SR	-194
Rat	150-	**M**SL**LVFLP**G**NILGKYSS**E**IRPTFF**L**AILYM**LVP**C**WAG**MRI**F**N**Q**SP	-194
Cow	150-	**M**SI**LVFLP**G**NILGKYSS**E**IRPTFF**L**TIPFL**LVP**C**WAG**LRV**F**N**Q**TQ	-194
Wolf	150-	**M**TI**LVFLP**G**NILGKYSS**E**VRPTLF**L**AIPYV**LVP**C**WAG**MRI**F**N**Q**SR	-194
Chicken	148-	**M**SI**LVFLL**G**NLIGKYSS**D**ISAAFL**L**NLPYV**LIP**I**WAG**TRL**F**Q**Q**PR	-192

Figure 1C. Evolutionary conservation of TM6SF2. *Bolded letters denote residues that are conserved among all species shown. Residue 167 is shown in red.*

The frequency of the TM6SF2 variant encoding p.Glu167Lys was higher in individuals of European ancestry (7.2%) than in African Americans (3.4%) or Hispanics (4.7%). Carriers of the TM6SF2 variant encoding p.Glu167Lys had elevated mean and median HTGC (hepatic triglyceride content) in all three ancestry groups, although the difference in triglyceride content relative to non-carriers did not reach statistical significance in Hispanics, likely owing to the lower number of Hispanic participants and the lower frequency of the variant in this group (Fig. 1D).

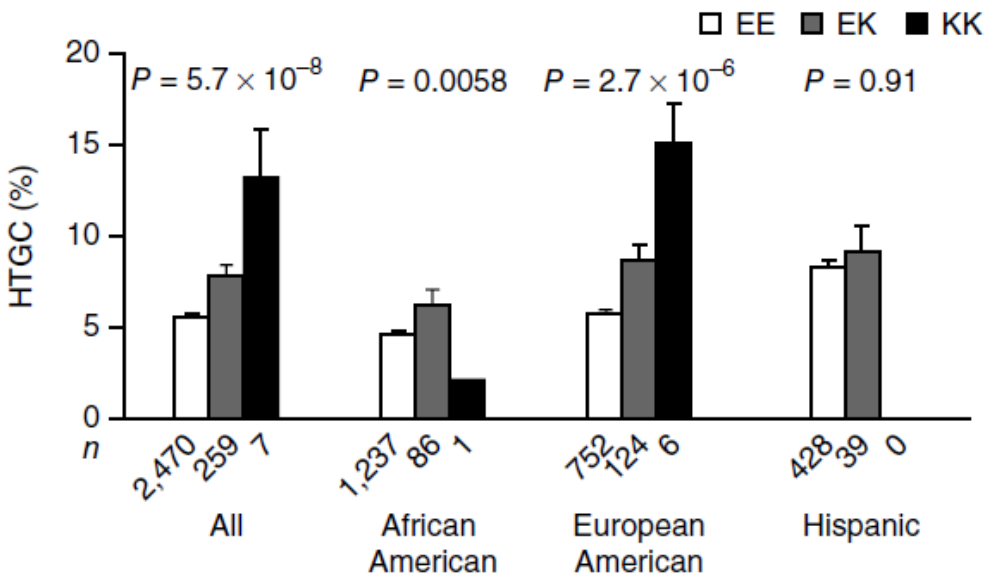


Figure 1D. Mean HTGC (± s.e.m.) by TM6SF2 genotype (rs58542926) in the Dallas Heart Study. *Association was tested using linear regression with adjustment for age, sex, ancestry and BMI. The number of individuals with each genotype is indicated below the corresponding bar: E, allele encoding Glu167; K, allele encoding Lys167.*

The association remained significant after adjusting for ethanol intake and HOMA-IR (P = 5.6 $\times 10^{-7}$). The effect of the TM6SF2 variant encoding p.Glu167Lys on HTGC was independent of the effect mediated by the PNPLA3 rs738409 polymorphism; we found no evidence for statistical interaction between the two risk alleles (P = 0.62). Previously, SNPs near TM6SF2 were found to be associated with NAFLD [34, 52, 85]. The variant at the locus that was most strongly associate with HTGC in the largest GWAS (total of 2.4 million imputed or assayed SNPs) was in NCAN (rs2228603) [34]. The TM6SF2 variant encoding p.Glu167Lys remained robustly associated with HTGC (P = 1.3 × 10^{-5}) after conditioning on rs2228603, as well as on other SNPs from the region on the array (Table 2).

Locus	SNP	MAF (%)	Median Hepatic TG % by genotype			Beta	*P* Uncondi-tional	P_{SNP} condi-tional on rs58542926	$P_{rs58542926}$ conditional on SNP
			0	1	2				
NCAN	rs2228603 (P92S)	3.7	3.5	3.9	12.0	0.076	0.002	0.79	1.3×10^{-5}
NCAN	rs150701551 (V548L)	0.64	3.6	3.3	-	0.033	0.57	0.52	5.5×10^{-8}
NCAN	rs10426537 (S838N)	3.9	3.6	2.8	3.2	-0.060	0.013	0.015	6.8×10^{-8}
NCAN	rs1064395	29.7	3.7	3.5	3.1	0.007	0.51	0.20	3.1×10^{-8}
TM6SF2	rs58542926 (E167K)	5.0	3.5	4.5	15.7	0.115	5.7×10^{-8}	NA	NA
SUGP1	rs10401969	12.4	3.6	3.6	3.3	0.035	0.014	0.26	7.2×10^{-7}
SUGP1	rs17751061 (R290H)	7.9	3.5	3.9	2.4	0.021	0.24	0.12	3.7×10^{-8}
SUGP1	rs138023808 (D104D)	0.7	3.6	3.3	-	0.027	0.62	0.56	5.5×10^{-8}
CILP2, PBX4	rs16996148	11.8	3.6	3.3	5.0	0.025	0.084	0.22	1.2×10^{-7}
CILP2, PBX4	rs17216525	5.3	3.5	4.2	18.1	0.073	0.0005	0.24	1.6×10^{-5}

Table 2. Conditional association analysis for variants in the chromosome 19 region including NCAN, TM6SF2, SUGP1, CILP2, and PBX4. *Variants within 10 kb of the coding sequences of NCAN, TM6SF2, and SUGP1 genotyped using the ExomeChip array, with a minor allele frequency of >0.5%, and other variants at the 19p13 locus reported in previous GWAS, are listed in the table. Genotypes were coded as 0, 1, and 2 for wild-type homozygotes, heterozygotes, and variant allele homozygotes, respectively. Each SNP was tested for association with hepatic TG content using linear regression including ancestry, age, gender, and BMI as covariates. The conditional analysis was also adjusted for TM6SF2 rs58542926 genotype. NA, not applicable.*

Conversely, conditioning on the TM6SF2 variant encoding p.Glu167Lys abolished the association between the NCAN variant (rs2228603) and HTGC (conditional P = 0.8). RT-PCR analysis of cDNA prepared from a panel of human tissues indicated that TM6SF2 is most highly expressed in the small intestine, liver and kidney and is present at lower levels in other tissues (Fig. 1E).

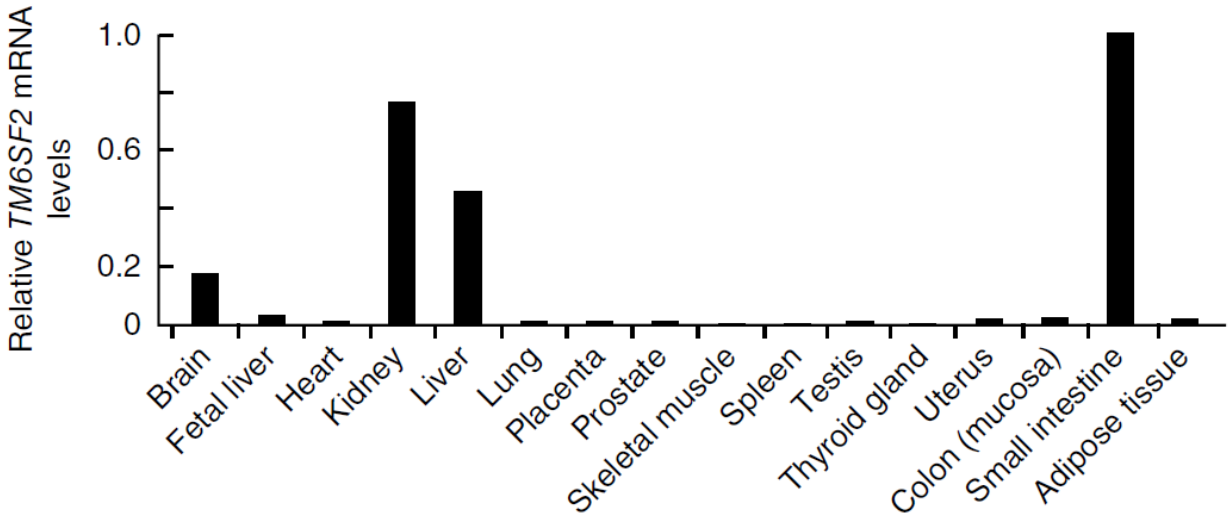


Figure 1E. Levels of TM6SF2 mRNA in human tissues. *Quantitative RT-PCR was performed on mRNA extracted from human tissues (Clontech). Each bar represents the average of a triplicate measurement expressed as a fraction of the Ct value obtained from the tissue expressing the highest level (small intestine). Values were normalized to the levels of the RPLP0 transcript.*

The TM6SF2 variant encoding p.Glu167Lys was also associated with a significant increase in serum ALT activity, consistent with increased hepatic injury (Table 3 and 4).

Study	Trait	EE[a]	EK[a]	KK[a]	*P* value
DHS					
	n	4,151	423	13	–
	ALT (U)	23.5 ± 19.9	25.8 ± 18.0	29.6 ± 26.8	0.014
	AST (U)	24.3 ± 20.6	25.1 ± 19.6	25.6 ± 16.5	0.17
	ALP (U)	71.6 ± 26.4	68.0 ± 20.9	63.6 ± 16.4	0.031
	LDL-C (mg/dl)	109 ± 36	105 ± 34	94 ± 39	0.005
	Triglycerides (mg/dl)	123 ± 102	118 ± 91	130 ± 66	0.037
Dallas Biobank (European Americans)					
	n	7,416	1,112	57	–
	ALT (U)	35.9 ± 16.3	36.5 ± 15	44.8 ± 23.8	0.003
	AST (U)	27 ± 13.8	26.9 ± 10.7	30.2 ± 13.5	0.22
	ALP (U)	69 ± 18.9	66.9 ± 17.3	65.9 ± 19.1	1.2×10^{-4}
	LDL-C (mg/dl)	107 ± 34	105 ± 32	97 ± 32	0.029
	Triglyceride (mg/dl)	112 ± 79	105.3 ± 67	97.5 ± 54	6.7×10^{-5}
Copenhagen Study					
	n	61,279	11,700	553	–
	ALT (U)	22.7 ± 17.2	23.8 ± 16.0	26.7 ± 20.9	7.6×10^{-14}
	AST (U) (*n* = 8,487)	22.8 ± 13.2	23.9 ± 18.6	30.0 ± 36.6	1.8×10^{-4}
	ALP (U)	91.1 ± 37.5	89.7 ± 39.1	85.0 ± 34.2	4.3×10^{-7}
	LDL-C (mg/dl)	128 ± 39	127 ± 38	112 ± 36	4.7×10^{-14}
	Triglyceride (mg/dl)	152 ± 106	148 ± 96	133 ± 90	3.7×10^{-9}

Table 3. Association between the nonsynonymous TM6SF2 variant (p.Glu167Lys), liver enzymes and plasma lipid levels in DHS, the Dallas Biobank and the Copenhagen Study. *Values are means ± s.d. Association was tested using linear regression with adjustment for age, sex, BMI and ancestry (where appropriate). A logarithm transformation was applied to traits with non-normal distributions. AST was only measured in a subset of the Copenhagen Study (n = 8,487). Ancestry breakdown for DHS is provided in Table 4. ALT, alanine aminotransferase; AST, aspartate transaminase; ALP, alkaline phosphatase; LDL-C, low-density lipoprotein–cholesterol.* [a]*EE, homozygotes for the allele encoding Glu167; EK, heterozygotes; KK, homozygotes for the allele encoding Lys167.*

Trait	Ethnicity	EE[a]	EK	KK	P-value[b]
HTGC (%) mean	DHS	5.79 (0.14)	8.10 (0.60)	13.20 (2.63)	5.7×10^{-8}
	Non-Hispanic African-American	4.78 (0.17)	6.52 (0.86)	2.15 (-)	0.0058
	Non-Hispanic European-American	5.86 (0.25)	8.63 (0.98)	15.04 (2.23)	2.7×10^{-6}
	Hispanic	8.35 (0.41)	9.14 (1.38)	-	0.91
HTGC (%) median	DHS	3.46 (2-7)	4.49 (2-10)	15.70 (9-18)	5.2×10^{-5}
	Non-Hispanic African-American	3.15 (2-5)	4.17 (2-7)	2.15 (-)	0.020
	Non-Hispanic European-American	3.50 (2-7)	4.29 (2-9)	16.87 (13-19)	0.0044[6]
	Hispanic	4.74 (3-11)	6.22 (3-13)	-	0.77
ALT (U)	Non-Hispanic African-American	22 ± 20.1	23.3 ± 16.8	13.0 ± NA	0.78
	Non-Hispanic European-American	23.7 ± 16.4	26.2 ± 18.5	32.4 ± 28.4	0.084
	Hispanic	27.7 ± 25	30.7 ± 19.9	-	0.015
AST (U)	Non-Hispanic African-American	24.1 ± 17.3	25.5 ± 27.7	18.0 ± NA	0.19
	Non-Hispanic European-American	23.9 ± 17.4	24.4 ± 12.1	27.3 ± 17.5	0.95
	Hispanic	25.7 ± 33	26.2 ± 12.6	-	0.23
ALP (U)	Non-Hispanic African-American	72.6 ± 29	71.4 ± 21.3	61.0 ± NA	0.063
	Non-Hispanic European-American	66.6 ± 20.5	64.3 ± 20.4	63.4 ± 17.8	0.87
	Hispanic	77.2 ± 26.3	71.3 ± 19.2	-	0.12
LDL (mg/dL)	Non-Hispanic African-American	108.1 ± 37	101.7 ± 34.8	61.0 ± NA	0.033
	Non-Hispanic European-American	110.8 ± 34.9	108.7 ± 33.9	96.3 ± 41.2	0.22
	Hispanic	108.7 ± 33.1	104.8 ± 33.6	-	0.21
Triglyceride (mg/dL)	Non-Hispanic African-American	107.2 ± 95.1	96.9 ± 55.2	67.0 ± NA	0.041
	Non-Hispanic European-American	136.2 ± 98.4	126.3 ± 76.6	132.2 ± 68	0.26
	Hispanic	149.6 ± 118.7	144.6 ± 160.8	-	0.46
HDL (mg/dL)	Non-Hispanic African-American	52.6 ± 15	52 ± 13.2	65.0 ± NA	0.85
	Non-Hispanic European-American	49.7 ± 15.4	49.5 ± 15.3	50.6 ± 19.2	0.76
	Hispanic	46.4 ± 11.5	46.9 ± 11.5	-	0.73

Table 4. Liver enzymes and plasma lipid levels in Dallas Heart Study. *Participants stratified by TM6SF2 (E167K) genotype and (self-reported) ethnicity.*

This finding is similar to what was previously observed with the PNPLA3 variant encoding p.Ile148Met, which is strongly associated with both higher HTGC and elevated ALT activity [28, 37]. To confirm the association with NAFLD, we performed association studies in two additional cohorts: the Dallas Biobank (n = 8,585 European Americans) and a cohort from Copenhagen (referred to as the Copenhagen Study in this manuscript) that included both the Copenhagen City Heart Study (CCHS) and the Copenhagen General Population Study (CGPS) (n = 73,532), (Tables 5 and 6).

Characteristic	EE[a]	EK	KK	P-value[b]
Total n	7,416	1,112	57	-
Female, n (%)	2,265 (30.5)	345 (31.0)	18 (31.6)	0.72
Age, years	53±11	54±11	54±10	0.069
Body-mass index, kg/m^2	26.2 (24-29)	26.1 (24-29)	26.1 (24-29)	0.71
Diabetes[c], n (%)	173 (2.3)	27 (2.4)	2 (3.5)	0.62
Lipid-lowering medication (statins), n (%)	443 (6.0)	70 (6.3)	6 (10.5)	0.37
Fasting glucose (mg/dL)[d]	92 (87-98)	93 (88-99)	91 (86-96)	0.041
Bilirubin (mg/dL)	0.56 (0.40-0.76)	0.56 (0.42-0.74)	0.54 (0.39-0.76)	0.71
HDL-cholesterol (mg/dL)	57.1±17.2	57.8±16.7	56.3±19.3	0.14

Table 5. Characteristics of the Dallas BioBank participants stratified by TM6SF2 (E167K) genotype. *Values are median (interquartile range), mean ± s.d., or number (%).* [a] *Missense variant (E167K) in TM6SF2.* [b] *P-values were calculated using linear regression for continuous*

variables (age, body-mass index, glucose, HOMA-IR) and logistic regression for categorical variables (gender, ethnicity, diabetes diagnosis, alcohol consumption, and use of lipid-lowering drugs). Regression models were adjusted for age, gender, and BMI, where appropriate. [c]*Diabetes was defined as a self-reported diagnosis of diabetes by a physician, use of prescription medication, fasting glucose ≥ 126 mg/dL or non-fasting glucose ≥ 200 mg/dL or HbA1c ≥ 6.5%.* [e]*Alcohol consumption was self-reported.* [d]*Diabetics excluded.*

Characteristics	EE	EK[a]	KK	P-value[b]
Total n	61,279	11,700	553	-
Age, years	58 (47-67)	58 (47-67)	56 (45-66)	0.02
Female, n (%)	34,075 (56)	6,377 (55)	315 (57)	0.07
Body mass index, kg/m^2	26 (23-28)	26 (23-28)	25 (23-28)	0.14
Diabetes[c]	3,025 (5)	588 (5)	34 (6)	0.40
Physical activity[c]	28,858 (48)	5,459 (47)	267 (49)	0.60
Alcohol consumption[c]	10,874 (18)	2,141 (18)	103 (19)	0.31
Gamma-glutamyltransferase, U/L	28 (21-42)	28 (21-42)	28 (21-42)	0.16
Bilirubin, µmol/L	10 (8-13)	10 (8-13)	11 (8-14)	0.87
Lipid-lowering therapy[c](statins), n (%)	5,693 (9)	1,023 (9)	45 (8)	0.13
HDL (mg/dL)	62.2 ± 19.7	62.1 ± 19.6	62.6 ± 19.4	0.67

Table 6. Baseline characteristics of participants in the Copenhagen General Population Study and the Copenhagen City Heart Study combined (The Copenhagen Study), stratified by TM6SF2 (E167K) genotype. *Values are median (interquartile range), mean ± s.d., or number (%).* [a]*Missense variant (E167K) in TM6SF2.* [b]*P-values were calculated using Kruskal-Wallis analysis of variance or Pearson's χ2-test.* [c]*Diabetes, physical activity in leisure time, alcohol consumption, and lipid-lowering therapy were self-reported, dichotomized, and defined as diabetes (self-reported disease, use of insulin, use of anti-diabetic medication, and/or non-fasting plasma glucose > 11.0 mmol/L) versus no diabetes, physical activity (four hours or more per week of light physical activity in leisure time versus less than four hours), alcohol consumption (>14/21 versus ≤14/21 units per week in women/men; 1 unit=12g alcohol), and use of lipid-lowering drugs (yes/no), mainly statins (>97%).*

As in DHS, the TM6SF2 variant encoding p.Glu167Lys was associated with significantly higher serum activity of ALT in both cohorts (Table 1). Mean serum aspartate transaminase (AST) activity was also higher in individuals homozygous for the TM6SF2 variant encoding p.Glu167Lys in both of these larger cohorts, but the increase only reached significance in the Copenhagen Study. These findings further support the hypothesis that the TM6SF2 variant

encoding p.Glu167Lys is associated with NAFLD and are consistent with the notion that the variant compromises hepatic integrity. No genotype-specific differences were observed in plasma levels of bilirubin or γ-glutamyltransferase (Tables 1 and 6), but, in all three populations, the TM6SF2 variant encoding p.Glu167Lys was associated with a significant reduction in the serum activity of alkaline phosphatase (ALP). The TM6SF2 variant encoding p.Glu167Lys was associated with highly significant reductions in plasma levels of triglycerides and LDL-C in the DHS, Dallas Biobank and Copenhagen Study cohorts (Table 1). Other SNPs in the region (rs10401969, rs16996148 and rs17216525)12–14 were associated with differences in plasma levels of triglycerides and LDL-C, but TM6SF2 remained associated with HTGC when we controlled for the genotypes at these loci. Conversely, controlling for the TM6SF2 variant encoding p.Glu167Lys abolished the signals for these SNPs (Table 2). Plasma levels of high-density lipoprotein–cholesterol (HDL-C) were not associated with the TM6SF2 variant encoding p.Glu167Lys (Table 1). Thus, the p.Glu167Lys substitution in TM6SF2 results in an increase in HTGC and a decrease in plasma levels of liver-derived triglyceride-rich lipoproteins. The specific biological role of TM6SF2 is not known. The protein is predicted to have seven transmembrane domains (TMHMM 2.0) [86], but it does not contain any known functional domains [87]. To assess the effect of the p.Glu167Lys substitution on the expression and localization of TM6SF2, we expressed human wild-type and mutant proteins in a human hepatoma cell line, HuH-7. Levels of TM6SF2 mRNA were comparable in cells expressing wild-type and mutant TM6SF2 (Fig. 2A, left), but levels of the mutant protein were reduced by 46% (Fig. 2A, right). Similar results were obtained in cells solubilized in 3% SDS and 8 M urea (data not shown).

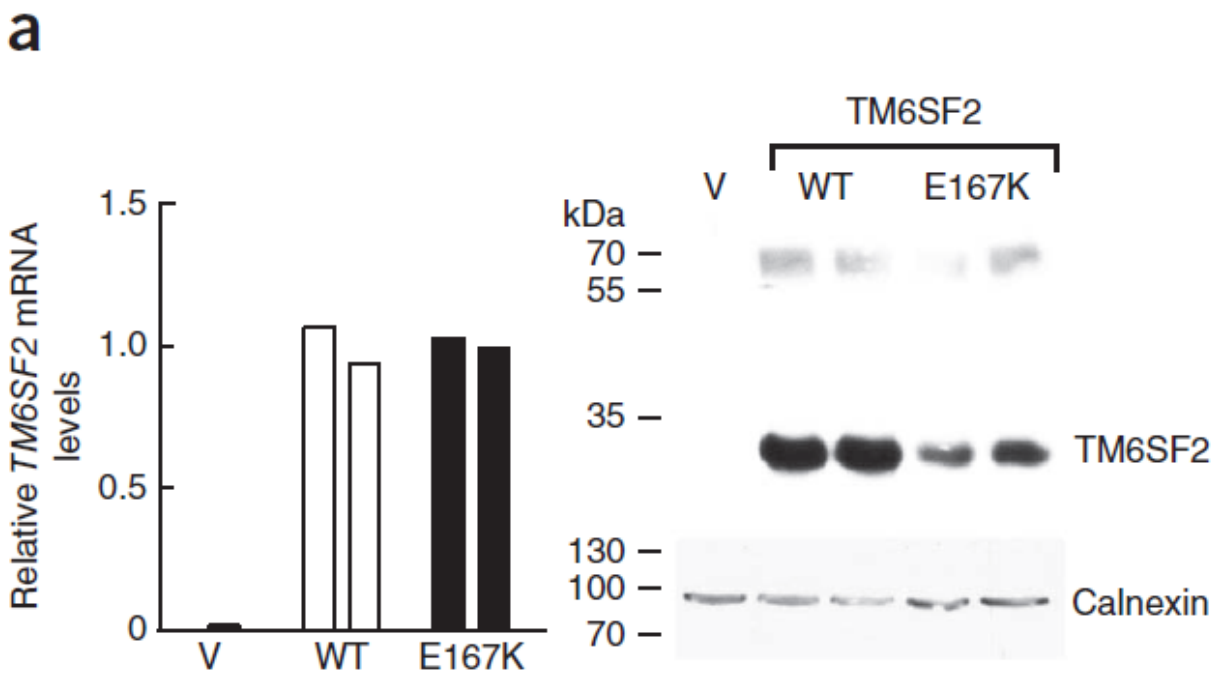


Figure 2A. Expression of TM6SF2 in cultured hepatocytes. *Plasmids encoding wild-type (WT) and Glu167Lys human TM6SF2 were expressed in HuH7 cells. Two days after transfection, TM6SF2 mRNA levels were measured using RT-PCR (left). Cells were collected and solubilized in RIPA buffer. Immunoblotting was performed and quantified using a LI-COR Odyssey infrared imaging system (as described in Methods) (right). The experiment was performed twice with similar results. Blots shown are representative of two independent experiments. V, vector.*

When lysates from Hepa1c1c7 cells expressing human TM6SF2-V5 were fractionated on sucrose density gradients, TM6SF2 was recovered exclusively in the membrane fraction (Fig. 2B). These findings suggest that TM6SF2 is a polytopic membrane protein and that the Glu167Lys variant form is misfolded and undergoes accelerated intracellular degradation.

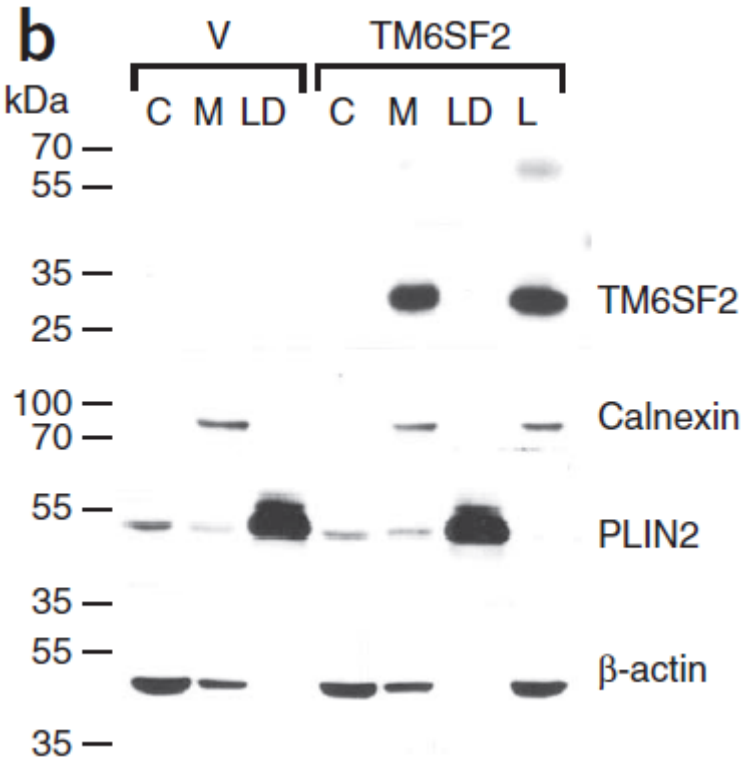


Figure 2B. Expression of TM6SF2 in cultured hepatocytes. *Recombinant wild-type TM6SF2 was expressed in Hepa1c1c7 cells. After two days, cells were fractionated and subjected to immunoblotting as described in Methods. C, cytosol; M, membranes; LD, lipid droplet; L, whole-cell lysate. The experiment was performed twice with similar results. Blots shown are representative of two independent experiments.*

To directly assess the effect of loss of TM6SF2 function on HTGC, we used recombinant adeno-associated viral (AAV) vectors expressing short hairpin RNAs (shRNAs) to selectively reduce Tm6sf2 transcript levels in the livers of mice. Expression of two different shRNAs targeting mouse Tm6sf2 resulted in >90% knockdown of Tm6sf2 mRNA (Fig. 3A, left), without any changes in Tm6sf2 mRNA levels in adipose tissue or small intestine (Fig. 3B).

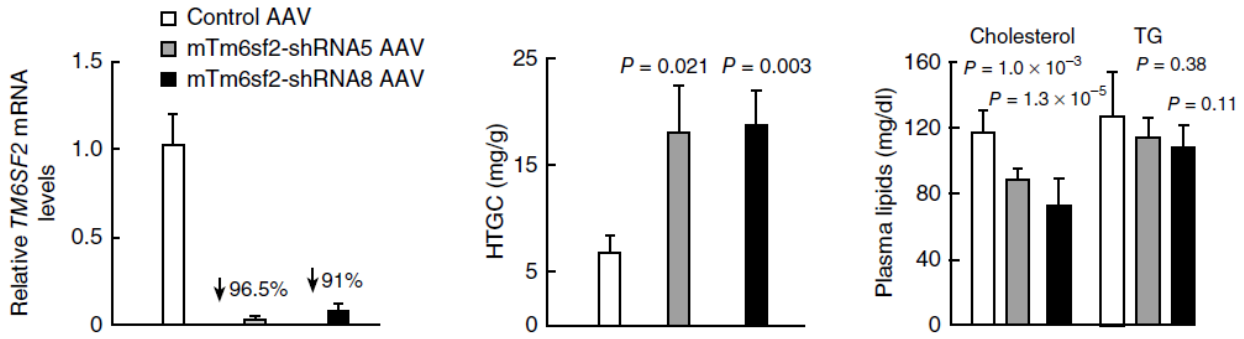


Figure 3A. Knockdown of Tm6sf2 expression in mouse liver is associated with increased HTGC. *AAVs expressing two different shRNAs (shRNA5 and shRNA8) to Tm6sf2 or vector alone were administered intravenously into the tail veins of 8-week-old chow-fed C57BL/6J male mice (n = 8 per group). After 2 weeks, livers were collected, and the levels of Tm6sf2 mRNA were measured using RT-PCR (left). Lipids were extracted from the livers of the AAV-treated mice, and triglyceride levels were quantified using enzymatic assays (middle). Plasma levels of cholesterol and triglycerides (TG) were measured in ad libitum–fed AAV-treated mice at the end of the dark cycle (right). The experiment was performed three times with similar results. Error bars, s.d.*

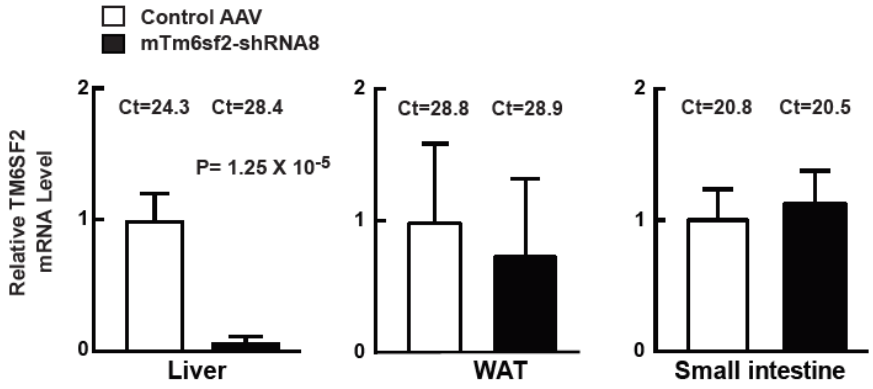


Figure 3B. Short hairpin RNA-mediated knockdown of Tm6sf2 in mice. *AAV vector alone or AAV expressing a shRNA directed against mTm6sf2 mRNA were injected into the tail veins of 8 week old chow-fed C57Bl/6J male mice (n=5/group). After two weeks, the mice were sacrificed and the liver, white adipose tissue (WAT) and small intestine were harvested. Levels of Tm6sf2 mRNA were measured using Real-Time PCR and normalized to m36B4 levels. The differences in mean expression levels were compared using a two-sample t-test.*

Empty AAV and AAV expressing nonfunctional shRNA were used as controls (Fig. 3C, upper). Hepatic inhibition of Tm6sf2 increased HTGC by threefold (Fig. 3C, middle left) and significantly decreased plasma levels of cholesterol in mice fed a chow diet ad libitum (Fig. 3C, middle right). Plasma triglyceride levels tended to be lower in knockdown mice fed ad libitum (Fig. 3C, bottom).

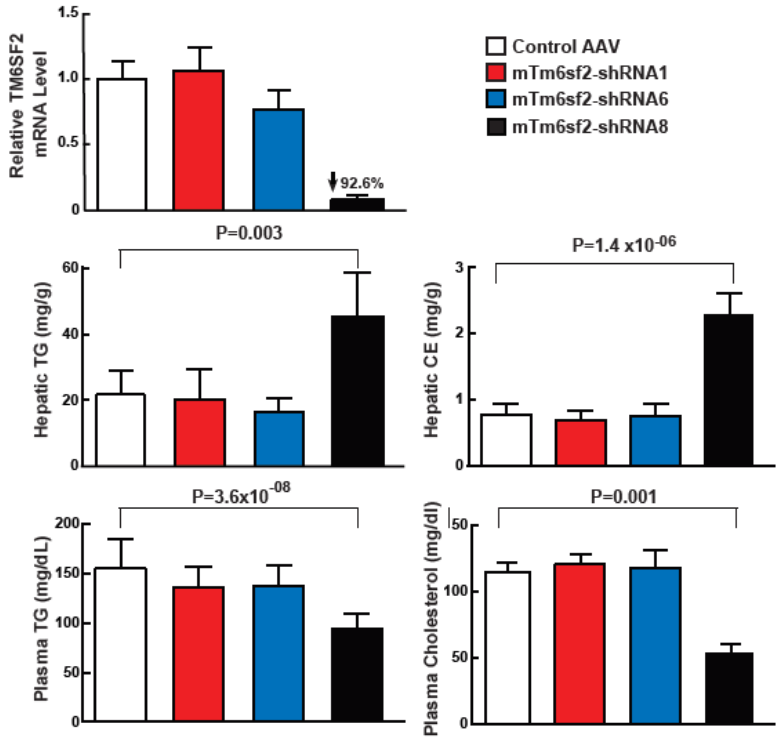


Figure 3C. Short hairpin RNA-mediated knockdown of Tm6sf2 in mice. *AAV expressing 2 nonfunctional shRNAs (shRNA1 and shRNA6), one functional shRNA (shRNA8) or AAV vector alone were injected intravenously into 8-week old chow-fed C57BL/6J male mice (n=6/group). All three shRNAs were directed against mouse Tm6sf2, but the only shRNA8 knocked down the levels of Tm6sf2 mRNA with high efficiency in cells. Two weeks after the injection, mice were fasted for 4 h, the livers were harvested and the levels of TM6SF2 mRNA were measured using Real-Time PCR. Values were normalized to m36B4 and expressed relative to the mean value of the AAV-only treated mice. Plasma was collected in the same experiment and TG and cholesterol levels measured as described in Methods. Differences in mean expression levels were compared using a two-sample T-tests. Ct, cycle threshold value.*

Fast protein liquid chromatography (FPLC) fractionation of plasma lipoproteins isolated after a 4-h fast showed that Tm6sf2 knockdown reduced the cholesterol content of both the LDL and HDL fractions and the triglyceride content of the VLDL fraction (Fig. 3D).

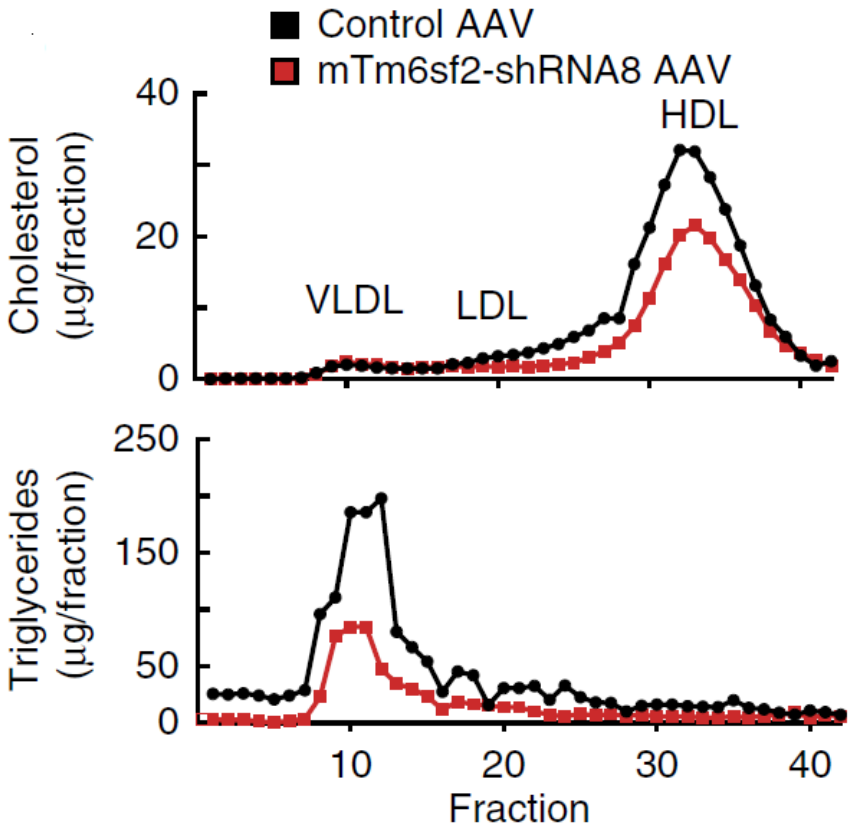


Figure 3D. Knockdown of Tm6sf2 expression in mouse liver is associated with increased HTGC. *Plasma lipoproteins from chow-fed C57BL/6J male mice (n = 6 per group) after a 4-h fast were fractionated using FPLC as described in Methods. Cholesterol and triglyceride levels were measured enzymatically in each column fraction. The experiment was performed twice with similar results.*

Mean serum ALT levels were unchanged in knockdown mice compared to controls (mean ± s.d.: 136 ± 8 U (shRNA6) and 131 ± 4 U (shRNA8) versus 132 ± 8 U (control); P = 0.33 and P = 0.68, respectively). Higher HTGC together with lower plasma cholesterol and triglyceride levels is consistent with a defect in VLDL secretion. To determine the effect of Tm6sf2 knockdown on VLDL secretion, we inhibited intravascular lipoprotein lipase, which hydrolyzes the triglyceride content of VLDL, and measured the rate of accumulation of triglycerides in plasma. Triglyceride accumulation was markedly reduced in the Tm6sf2 knockdown mice (mean ± s.d.: 4.5 ± 1.6 mg/dl/min (shRNA8) versus 11.1 ± 0.6 mg/dl/min (control); Fig. 3E).

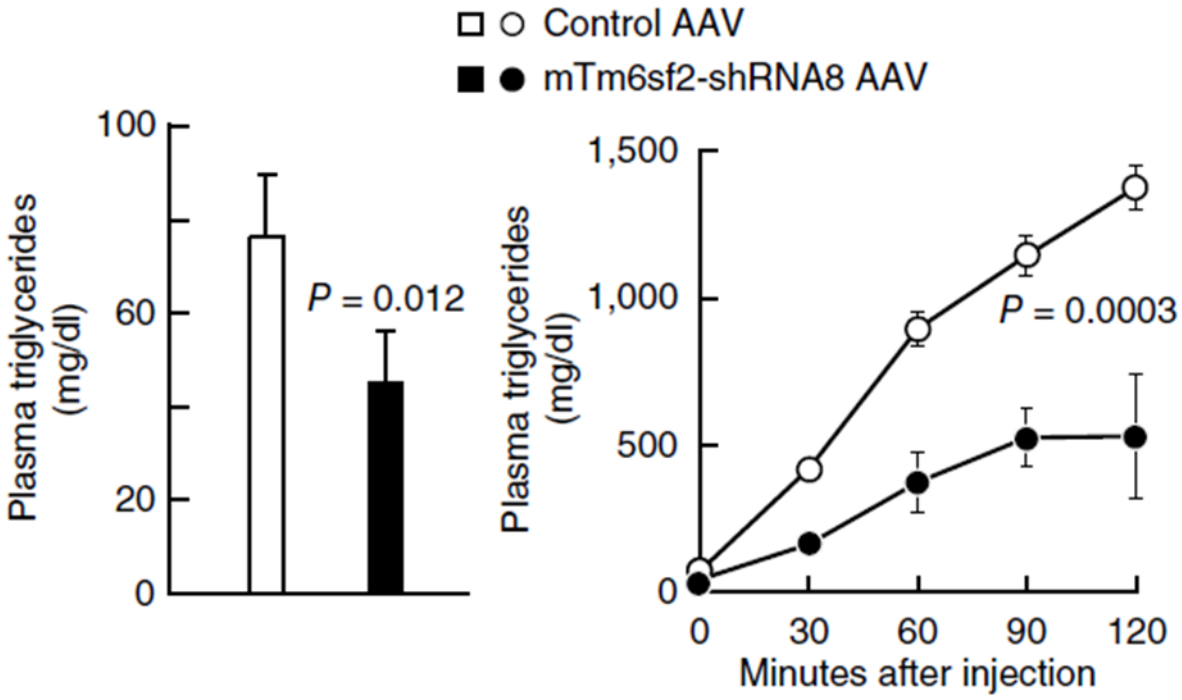


Figure 3E. Knockdown of Tm6sf2 expression in mouse liver is associated with increased HTGC. *Hepatic VLDL secretion in Tm6sf2 knockdown mice. Plasma triglyceride levels after a 4-h fast (left) and triglyceride accumulation after injection with Triton WR-1339 (right) were measured in chow-fed male mice (8 weeks old; n = 4 per group) treated with mTm6sf2-shRNA8 AAV or control AAV as described in Methods. Two weeks after infection, mice were fasted for 4 h and injected via the tail vein with Triton WR-1339 (500 μg/g body weight). Blood was sampled at the indicated times, and plasma triglyceride levels were measured. Mean triglyceride levels at each time point are shown. Slopes of the lines were calculated by least-squares regression and compared using a t-test. The experiment was performed twice with similar results. Error bars, s.d..*

These data indicate that TM6SF2 normally acts to promote VLDL secretion and suggest that the increased HTGC associated with the Glu167Lys TM6SF2 variant in humans results from a reduction in TM6SF2 function. High-sucrose diets, which increase hepatic triglyceride synthesis, exacerbated the effects of Tm6sf2 knockdown on HTGC (Fig. 4). In sucrose-fed mice, levels of triglycerides and cholesterol esters were higher in the liver (Fig. 4A) and lower in plasma (Fig. 4B). Oil Red O staining, which stains neutral lipids, showed an increase in the number and size of lipid droplets in the knockdown mice (Fig. 4C).

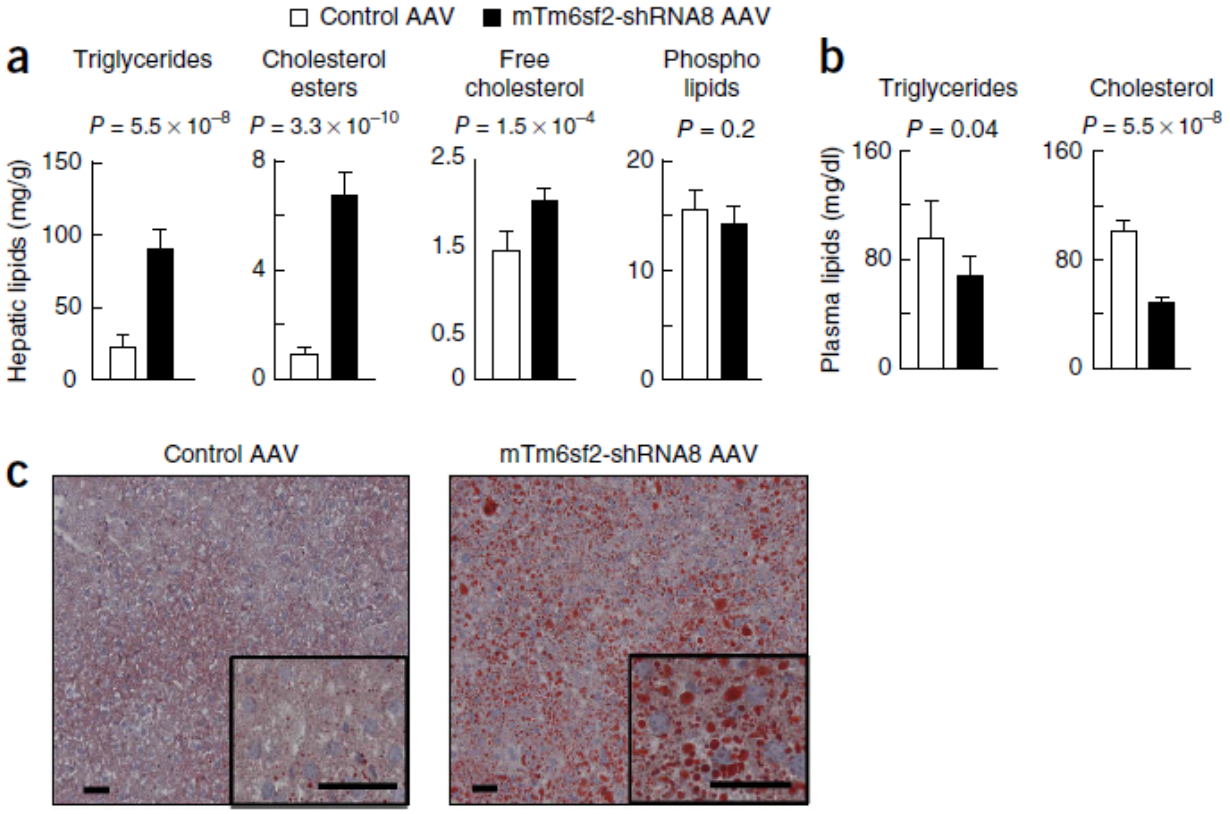


Figure 4. Sucrose feeding in Tm6sf2 knockdown mice. *AAV expressing shRNA to mTm6sf2 or vector alone were administered intravenously into the tail veins of 8-week-old C57BL/6J male mice (n = 6 per group), and mice were fed a high-sucrose diet for 4 weeks. (a) Lipids were extracted from the livers of the AAV-treated mice, and hepatic triglyceride levels were quantified using enzymatic assays. Error bars, s.d. (b) Plasma levels of cholesterol and triglycerides in Tm6sf2 knockdown mice were measured enzymatically. Differences in mean hepatic triglyceride concentration were compared using two-sample t-tests. Error bars, s.d. (c) Sections from the liver tissue of the same mice stained with Oil Red O as described in Methods. Magnification: main image ×20, inset ×64. Scale bars, 40 μm. The experiment was performed twice with similar results.*

Therefore, knockdown of Tm6sf2 selectively in mouse liver recapitulated the effects on HTGC and plasma lipids of the Glu167Lys TM6SF2 variant observed in humans. TM6SF2 is expressed at the highest levels in the human intestine (Fig. 1E), which is the source of dietary-derived triglyceride-rich lipoproteins (chylomicrons).

To elucidate the biologic function of TM6SF2 and the pathogenic mechanisms of TM6SF2-associated liver disease we generated mice in which expression of the gene was ablated by insertional mutagenesis using a gene trap vector as described in Methods.

To determine the tissue distribution of TM6SF2 expression in mice, we measured levels of TM6SF2 mRNA in tissues collected from female wildtype (WT) C57Bl/6N mice after a 4 h fast (Fig. 5A). TM6SF2 expression was highest in the liver and small intestine.

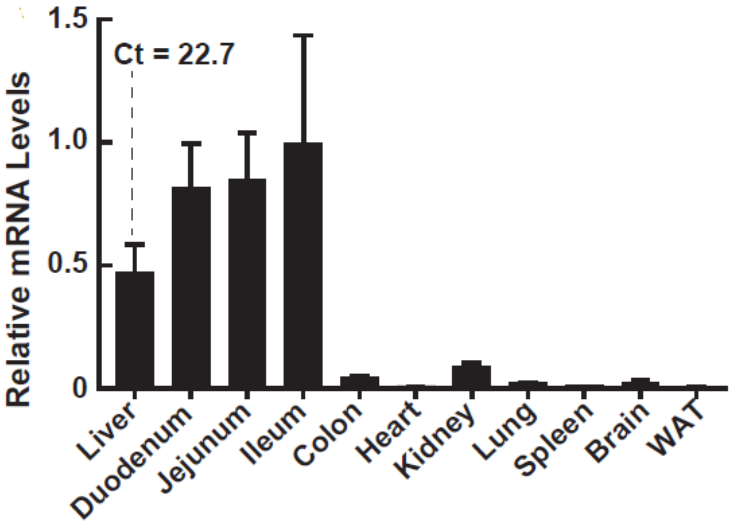


Figure 5A. Expression of Tm6sf2. *Total RNA was extracted from the indicated tissues of WT female mice (n=3, age=14 week) after a 4 h fast and subjected to quantitative RT-PCR as described in Methods. The mean (±SEM) levels of TM6SF2 transcript in each tissue are expressed relative to the expression level in the ileum, which was arbitrarily set to 1.*

To determine if food intake alters TM6SF2 expression, we collected the jejunum and liver from mice fasted for 24 h, and from mice fasted for 18 h and then refed for 6 h. TM6SF2 protein levels were ~10-fold higher in jejunum than in liver (Fig. 5B) and were not significantly affected by fasting or refeeding in either tissue (Fig. 5B and 5C). Levels of TM6SF2 mRNA were slightly increased in response to refeeding in one experiment, but this change was not observed consistently (data not shown) and was not reflected in the levels of TM6SF2 protein (Fig. 5C). Expression levels of fatty acid synthase (FAS) and adipocyte TG lipase (ATGL) were assayed as positive controls for the feeding protocol and showed the expected changes (Fig. 5C).

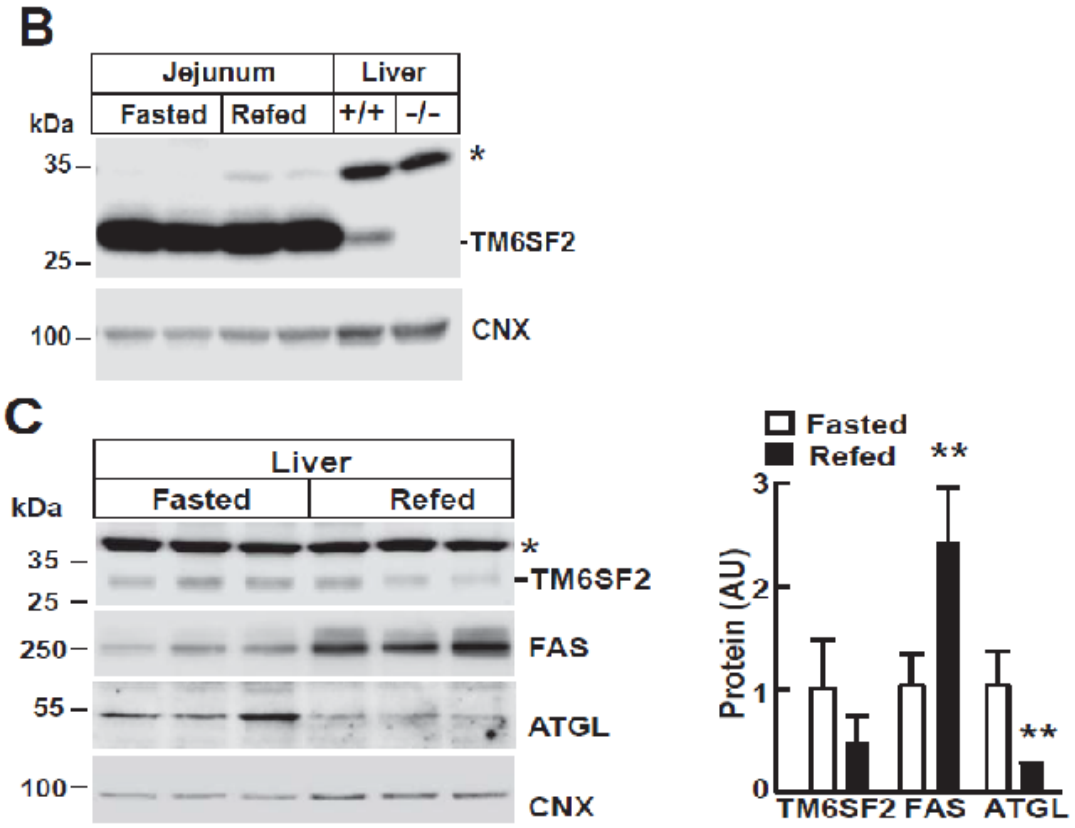


Figure 5B and C. Expression and regulation of Tm6sf2. *(B) Regulation of TM6SF2 expression in response to fasting and refeeding. Mice were entrained to a synchronized feeding regimen for 3 days and then sacrified after a 24 hour fast (fasted) or after 18 h of fasting and 6 h of refeeding with high sucrose diet (refed) (4 male mice, age=8 weeks). Jejunal proteins (60 µg/well) were size fractionated by 10% SDS-PAGE, and immunoblot analysis was performed using antibodies against TM6SF2 and calnexin. (C, left) Liver proteins (60 µg/well) from the experiment described in Fig. 5B were size-fractionated on a 10% SDS-PAGE. Fatty acid synthase (FAS) and adipose TG lipase (ATGL) and were used as positive controls for fasting and refeeding and calnexin (CNX) was used as loading control in this experiment. (C, right) Immunoblot signals were quantified using a Licor Odyssey Fc Imager. **P<0.01*

To determine the effect of stably inactivating Tm6sf2 in all tissues, we established Tm6sf2 -/- mice from two independent embryonal stem (ES) cell lines. The map of the genetically modified locus is shown in Fig. 5D. All Tm6sf2 KO mice used in the experiments described in this study were offspring from matings between heterozygotes (Tm6sf2 +/-) derived from one of these lines (EPD0097_3_F02), and the major findings were validated in mice derived from the second line (EPD0097_3_E02) (EPD0097_3_E02 mice data not shown).

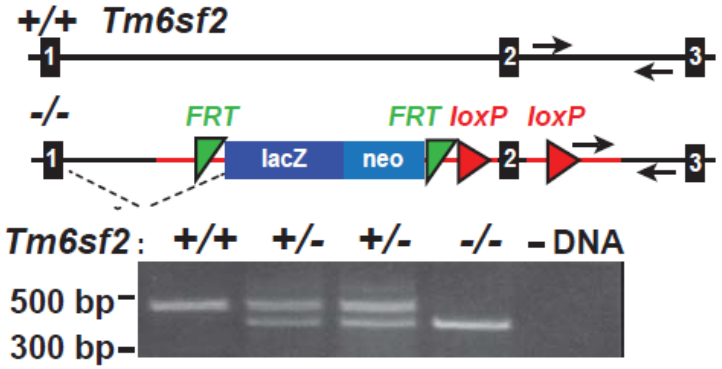


Figure 5D. Generation of Tm6sf2 -/- mice. *Tm6sf2-/- mice were generated as described in Methods. Genotyping was performed by PCR using oligonucleotides (arrows) to amplify a 470 bp (WT) or 400 bp (KO) fragment from genomic DNA.*

The liver of Tm6sf2-/- mice expressed almost no TM6SF2 mRNA (Fig. 5E, left) and had no detectable TM6SF2 protein (Fig. 5E, right). Both male and female KO mice were fertile. Litters from matings between heterozygous mice (Tm6sf2+/-) were comparable to those obtained from WT mice of the same strain (5.3 ±2.3 vs. 5.0 ± 2.5 pups/litter). The sex ratios were similar among WT, Tm6sf2+/- and Tm6sf2-/- offspring. Although fewer KO offspring were obtained from heterozygous parents than expected by chance, the difference did not reach statistical significance (P = 0.28) (data not shown).

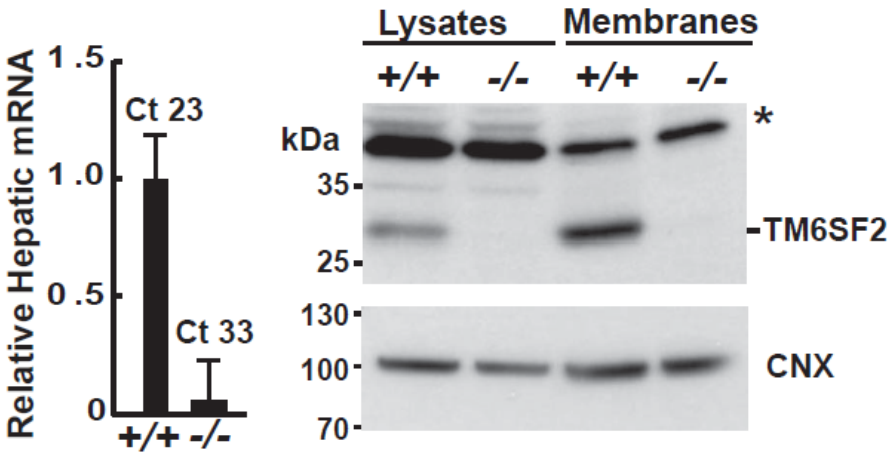


Figure 5E. Expression of Tm6sf2 in Tm6sf2 KO mice. *Left: RNA was isolated from livers of 4h fasted male WT and KO mice (n=3 male mice/group, 14 weeks old) and TM6SF2 expression was determined by quantitative RT-PCR as described in Methods. The level of TM6SF2 transcript in WT mice was arbitrarily set to 1. Right: Immunoblot analysis of hepatic TM6SF2 in 7 week-old*

*female WT and KO mice. Liver lysates and membranes were prepared as described in the Methods. Aliquots of each fraction (50 μg) were size-fractionated by SDS-PAGE and immunoblotting was performed using a rabbit anti-mouse TM6SF2 polyclonal antibody (1:1000) as described in Methods. Calnexin (CNX) served as a loading control for the experiment. * - nonspecific band. All experiments were repeated at least once, and the results were similar. Ct, cycle threshold. Values are means +/- SEM.*

Inactivation of Tm6sf2 did not adversely affect postnatal development: body weights, fat mass, lean mass and liver weights were similar in 13-week old chow-fed male KO and WT mice (Fig. 6A). Similar findings were seen in female mice and mice derived from the second ES line (Clone EPD0097_3_E02, data not shown). Serum glucose, insulin and non-esterified fatty acid (NEFA) levels after a 4-h fast (Fig. 6B), and serum glucose levels after a 24 h fast (92±10 vs. 90±5 mg/dl, P=0.86; n=4 male mice) (data not shown), did not differ between the strains. Biliary sterol levels were similar in KO and WT animals (Fig. 6C). As expected, body weights and fasting serum insulin levels were increased after the mice consumed a high-fat diet for 12 weeks, but the increases were similar in WT and KO animals (Fig. 6D and E).

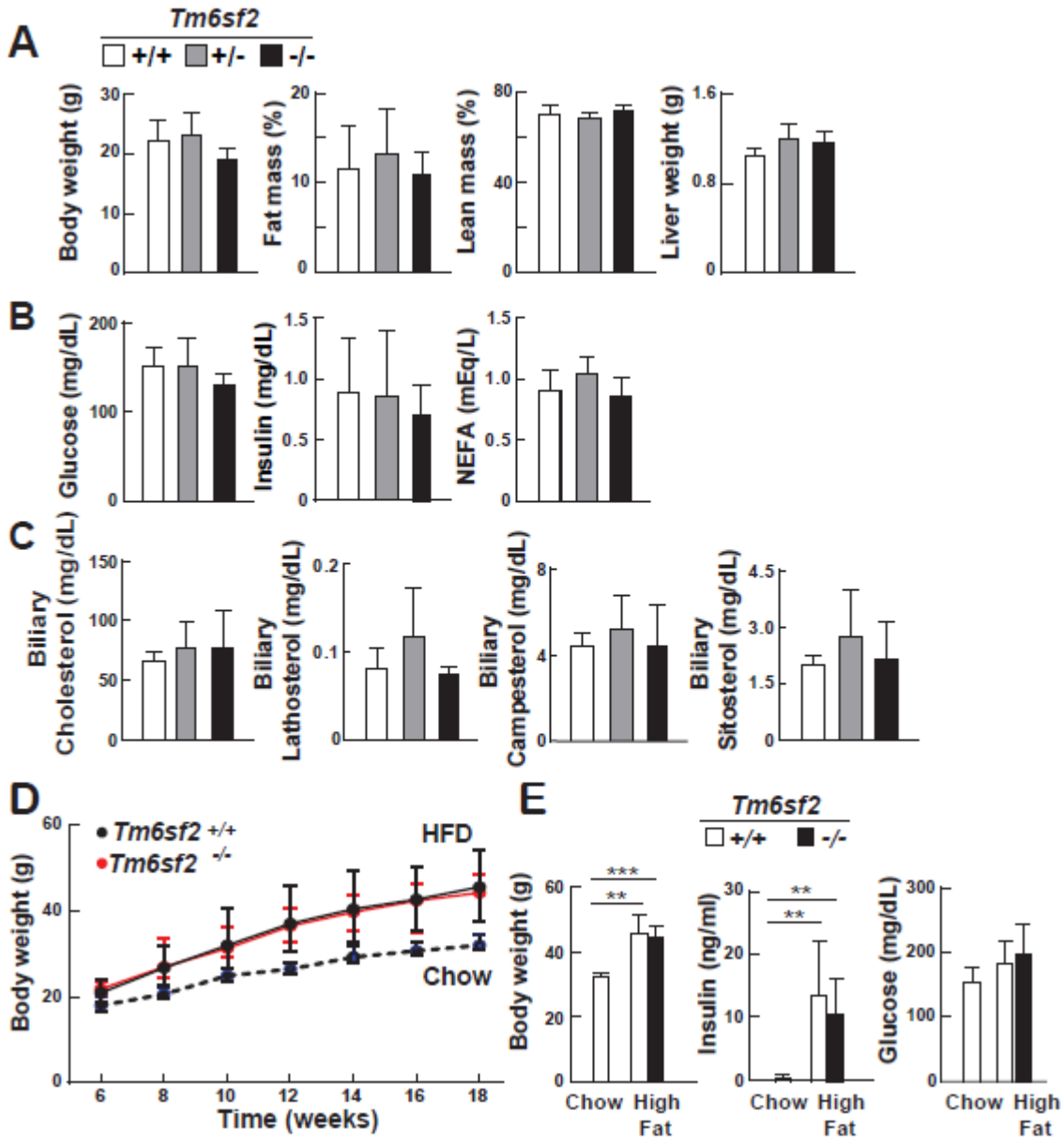
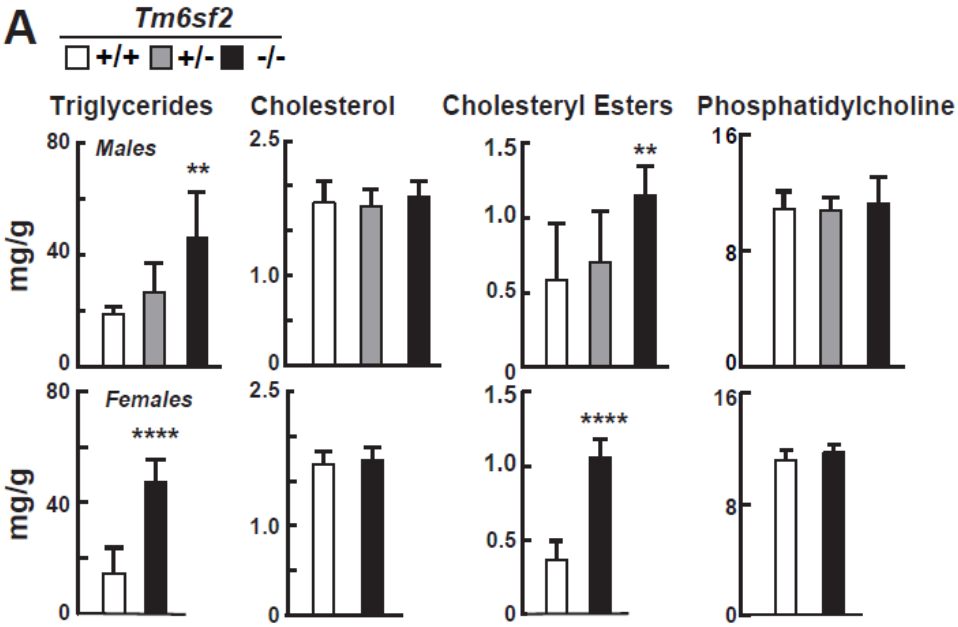


Figure 6. Body composition (A), plasma chemistries (B), bile (C) and responses to a high-fat diet (D and E) in WT and Tm6sf2-/- mice. *(A) Fat mass and lean body mass of chow-fed male mice (9 weeks old, n=5) were measured by nuclear magnetic resonance (NMR) using a Minispec analyzer (Bruker). Body and liver weights were measured after the mice were sacrificed at 13 weeks of age. (B) Plasma glucose, insulin, and non-esterified fatty acids (NEFA) were measured after 4 h fast in the same mice at 11 weeks, as described in the Methods. (C) The composition of 13 weeks old male mice bile, 4h fasting. (D) Male WT and KO mice (age = 6 weeks, n=5 /group) were switched from a chow to a high-fat diet at 6 weeks of age and fed with HFD for 12 weeks. A similar number of WT mice were simultaneously fed a chow diet. Body weights were measured at*

*2 weeks intervals, and final body weights plus plasma glucose and insulin levels (E) were measured at 18 weeks after a 4 h fast. Values are means +/- SEM. **P<0.01, ***P<0.001*

Levels of TG and cholesteryl esters (CE) were increased in livers of chow-fed Tm6sf2-/- male mice compared to their WT littermates (Fig. 7A, top). Heterozygous mice had levels of TG and CE that were intermediate between those of WT and homozygous KO animals, which is consistent with the inactivating allele having a co-dominant effect. More noticeable increases in neutral lipids were seen in chow-fed female Tm6sf2-/- mice (Fig. 7A, bottom). Hepatic levels of free cholesterol and phosphatidylcholine were similar in WT and KO animals of both sexes. Oil Red O staining of liver sections (Fig. 7B) revealed an increase in the number (417 vs. 47 droplets per 8404 μm^2 slide field) and the median size (2.2 versus 1.6 μm^2, P=0.023) of lipid droplets (LDs) in the KO animals (Fig. 7C). To limit artifacts due to nonspecific staining, a droplet area of 1 μm^2 was set as a minimum threshold. The size distribution of droplets was skewed to the left in both strains, with most droplets being in the smallest detectable size range (1-2 μm^2).



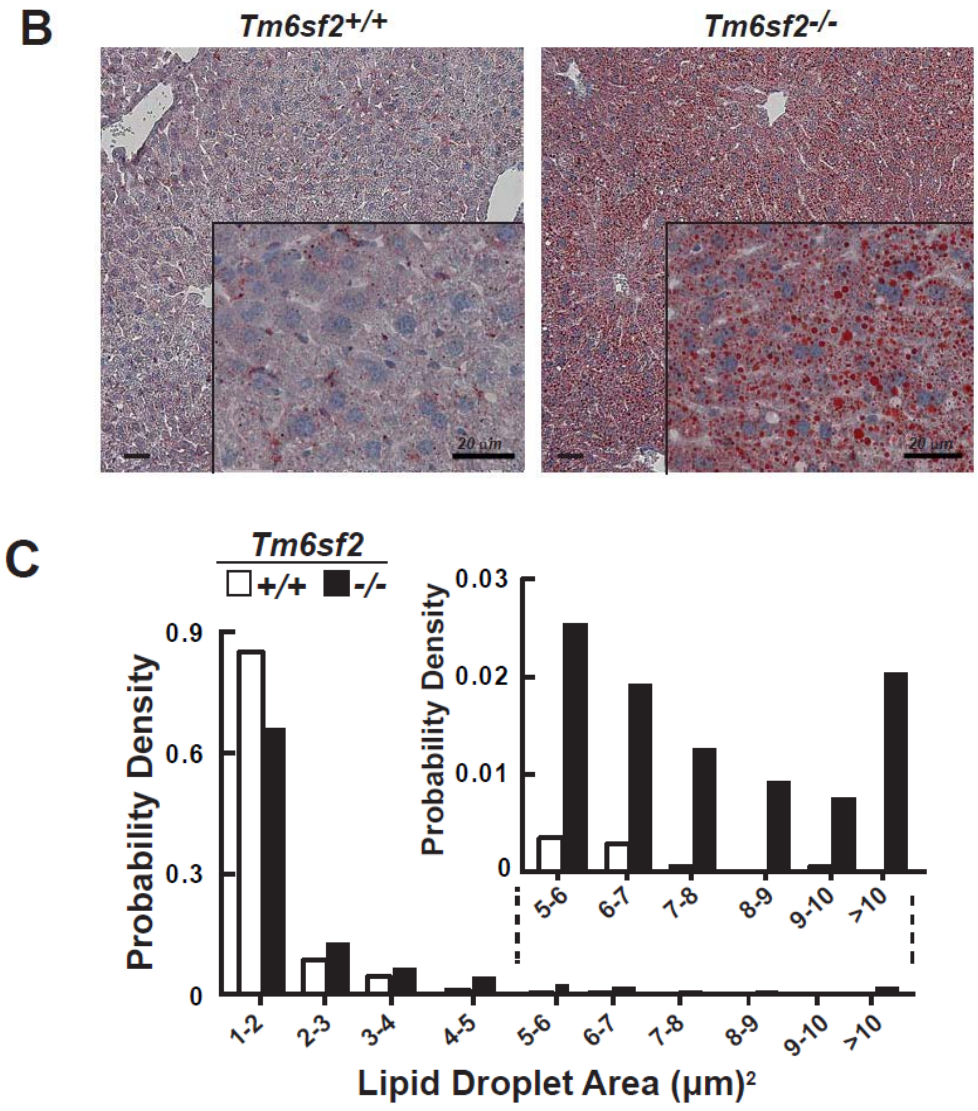


Figure 7. Hepatic lipid content and LD size distribution in Tm6sf2 -/- mice. *(A) Hepatic lipid levels were measured in 13 week-old chow-fed male mice (5/group)(upper panel) and 11 week-old female mice (6/group) (lower panel) using enzymatic assays. Bars are means ± SEM. (B) Liver sections from WT and Tm6sf2-/- mice were stained with Oil Red O and hematoxylin and visualized using a Leica microscope (DM2000) at 10X and 40X (inset) magnification. (C) Size distribution of male hepatic LDs in WT and KO mice. LD sizes were determined from images of Oil Red O stained liver sections using ImageJ software as described in the Methods. The experiment was repeated twice, and the results were similar. **P<0.01 and ****P<0.0001*

In contrast to the increase in neutral lipid content and LD size, we found no differences in immunodetectable ApoB-48 or ApoB-100, the major structural proteins in very low density lipoproteins (VLDL), in liver lysates of KO and WT mice (Fig. 8A). Thus, inactivation of

Tm6sf2 -/- led to the accumulation of neutral lipids in the liver without changing the steady state levels of ApoB-100 or ApoB-48. Genetic variation in TM6SF2 is associated with steatohepatitis in humans [2]. Therefore, we screened the KO mice for early stigmata of hepatic injury. Serum levels of ALT, a liver enzyme released into the circulation in response to hepatocyte injury, were higher in male KO mice than in WT animals (Fig. 8B), although levels of hepatic transcripts encoding genes associated with inflammation and fibrosis and the unfolded protein response were similar in KO and WT animals (Fig. 8C). Hematoxylin and eosin staining of liver sections did not reveal pathological stigmata of liver injury; no mononuclear infiltrates, ballooning of hepatocytes or fibrin deposition was seen in the livers of the KO mice (data not shown).

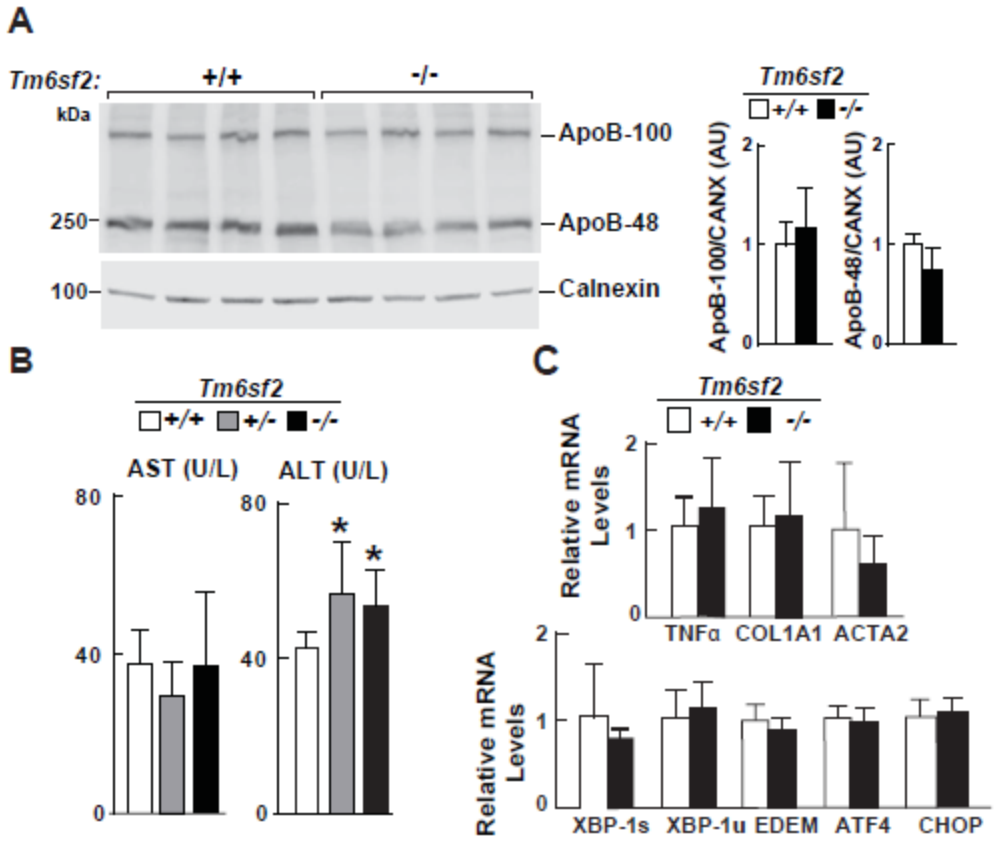


Figure 8. Hepatic ApoB levels (A), liver function tests (B) and relative hepatic levels of selected transcripts encoding proteins involved in activation of fibrosis and ER stress (C) in male WT and Tm6sf2 -/- mice described in Fig.7A. *(A) Immunoblot analysis was performed on liver protein (50 μg) using a rabbit anti-mouse ApoB polyclonal antibody (1:1000, Abcam) and ECL (SuperSignal West Pico Kit, Thermo Scientific). The ECL signal was visualized using a Licor imager (Odyssey Fc Imager) and analyzed using Licor Image Studio Software. (B) Plasma levels*

*of aspartate aminotransferase (AST) and alanine aminotransferase (ALT) in chow-fed male mice. (C) RNA levels were detected using quantitative Real-Time PCR (qRT-PCR), normalized to levels of 36B4 and expressed relative to the levels in the WT animals (n=5). TNFα, tumor necrosis factor alpha; COL1A1, collagen alpha-1(I) chain; ACTA2, actin, aortic smooth muscle; XBP1s/u, X-box-binding protein 1 spliced/unspliced; ATF4, Activating Transcription Factor 4; EDEM, ER Degradation Enhancer, Mannosidase Alpha-Like 1; CHOP(DDIT3), C/EBP-homologous protein. Values are means +/- SEM. *P<0.05*

Hepatic expression of genes involved in the transcriptional regulation, synthesis, storage, lipolysis and secretion of lipids and lipoproteins was quantitated by RT-PCR after a 4 h fast in chow-fed mice (Fig. 9A). Expression of PNPLA3 was markedly reduced in the KO mice. PNPLA3 is a direct target of SREBP-1c, which is an insulin responsive transcription factor that upregulates fatty acid synthesis [44]. Levels of mRNA encoding other SREBP-1c target genes were unchanged in males, and modestly reduced in female KO mice (Fig. 9A).

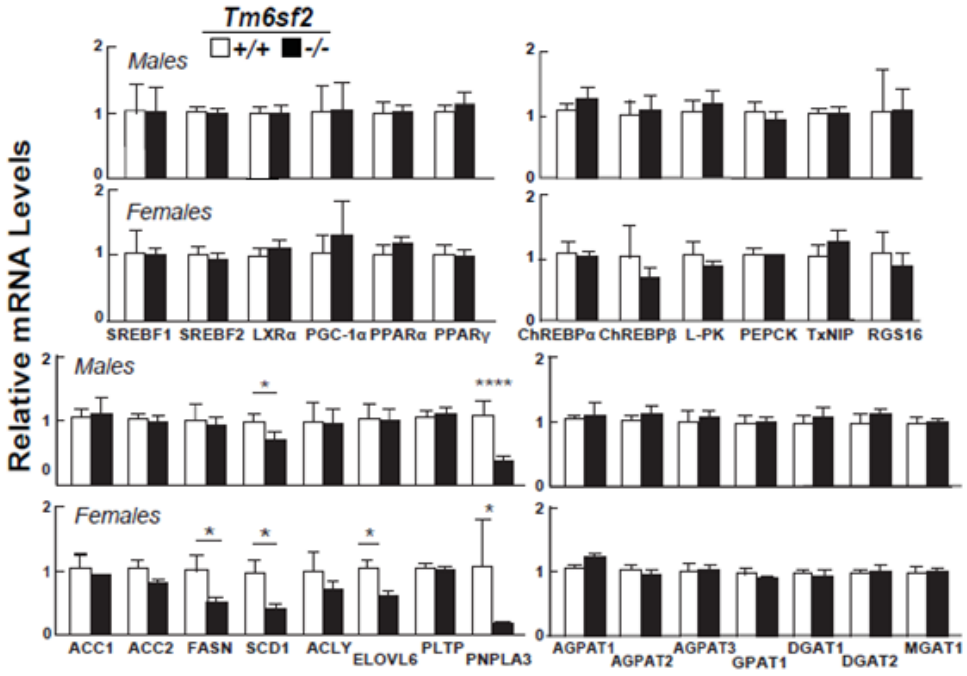


Figure 9A. Relative mRNA levels of selected genes involved in cholesterol and triglyceride metabolism in the livers of WT and Tm6sf2-/- mice. *Quantitative RT-PCR assays were performed to assess the relative levels of selected mRNAs in the liver of the 13 week-old chow-fed male mice (5/group) and 11 week-old chow-fed female mice (6/group) described in the legend to*

*Fig. 7A. Expression levels were normalized to levels of 36B4 and expressed relative to levels of WT transcript. Values are means +/- SEM. The official gene symbols were used for all the genes with the following exceptions: ATGL, PNPLA2; LXRα, NR1H3; PGC1α, PPARGC1A; L-PK, PKLR; PEPCK, PCK1; ChREBP, MLXIPL. *P<0.05 and ****P<0.0001.*

No differences in the hepatic levels of either the precursor or mature form of SREBP-1c were found in livers from male (Fig.5B) or female (data not shown) WT and KO mice.

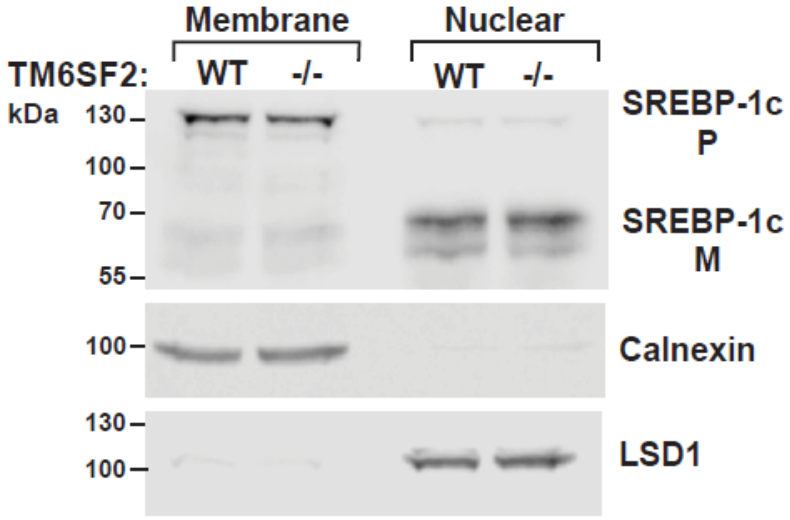


Figure 9. The hepatic SREBP-1c cleavage in Tm6sf2 KO mice. *Nuclear and membrane fractions were isolated from livers of 18 weeks old refed male mice (n=4) by ultracentrifugation as described in Methods. Lysates from each mouse were pooled and 40 μg of pooled protein was size separated on SDS (10%) PAGE gels. Proteins were transferred to nitrocellulose membranes and blotted with rabbit anti-mouse mSREBP-1c antibody. The bands were visualized by ECL and quantified using a Licor Odyssey Fc Imager. The membranes were then stripped and reblotted with antibodies against Calnexin and LSD1. The experiment was repeated with 13-week old females, and the results were similar.*

No changes were seen in the major SREBP- 2 target genes in male KO mice, whereas females had reductions in both HMG-CoA synthase (HMGCS1) and HMG-CoA reductase (HMGCR) mRNA levels (Fig. 10A), but not in the mRNAs encoding other enzymes in the cholesterol biosynthetic pathway (Fig. 10B). These findings indicate that absence of TM6SF2 in chow-fed animals increases hepatic TG and CE content with only modest effects on the transcriptional control of fatty acid and cholesterol metabolism.

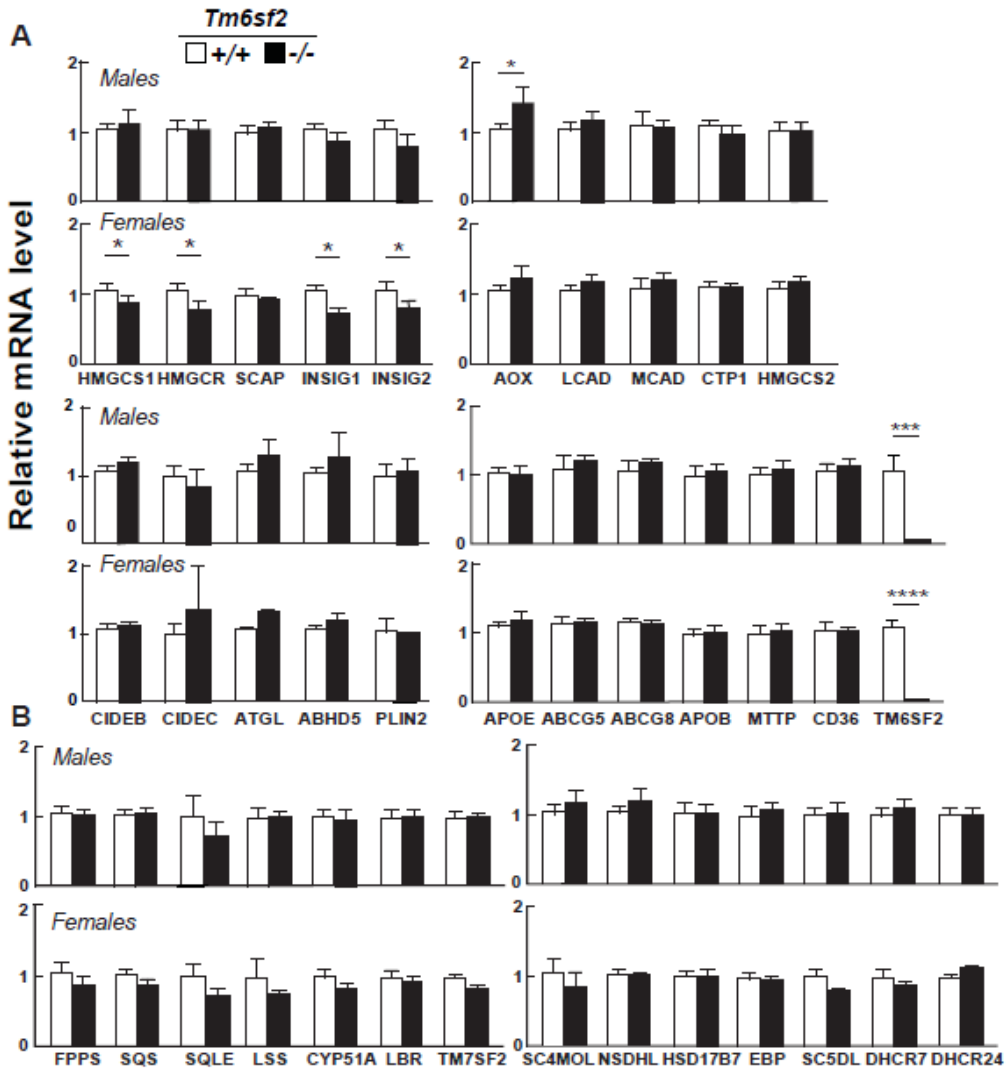


Figure 10. Relative mRNA levels of selected genes involved in cholesterol metabolism in the livers of WT and Tm6sf2-/- mice. *(A and B) Quantitative RT-PCR assay was performed to assay the relative levels of selected mRNA levels in livers of the mice described in the legend to Fig. 7A and Fig. 9A. Expression levels were normalized to levels of 36B4 and expressed relative to levels of WT transcript. Values are means +/- SEM. The official gene symbols were used for all the genes with the following exceptions: AOX, ACOX1; MCAD, ACADM; FPPS, FDPS; SQS, FDFT1; SC4MOL, MSMO1. *P<0.05, ***P<0.001 and ****P<0.0001.*

Like the liver, the intestine is a major site of lipoprotein synthesis. To determine if TM6SF2 deficiency leads to accumulation of neutral lipids in the intestine, we challenged the KO mice and their WT littermates with a high-fat diet (45% lard, 5.5% soybean oil, 17% sucrose, and 19% casein) for 3 months. Oil red O staining of the jejunum was performed after a 4 h fast. LDs were seen along the basolateral surfaces of enterocytes covering surfaces of the villi, but not in

the crypts of the KO mice (Fig. 11A). Staining was also apparent in the *lamina propria* of the villi. No lipid accumulation was apparent in enterocytes of WT mice. A similar staining pattern was seen in the duodenum and ileum (data not shown). To determine whether TM6SF2 plays a role in TG absorption, we challenged chow-fed WT and KO mice with an acute bolus of corn oil (10 μL corn oil/g body weight) delivered by gavage. Plasma TG levels were measured at the indicated times. The excursion of plasma TG following the gavage was similar in magnitude in WT and KO mice, but peak levels occurred later in the KO mice (1.5±0.5 vs 2.7±0.5 hours, P=0.0003) (Fig. 11B). This finding suggests that in Tm6sf2 KO mice TG absorption is preserved, but somewhat delayed. To more precisely compare rates of absorption of neutral lipid between the mouse strains, we inhibited lipoprotein lipase (LPL) by injecting Triton WR-1339 into the tail vein of the mice. We then administered by gavage a solution of corn oil containing ^{3}H-labeled oleic acid and ^{14}C-labeled cholesterol and measured the appearance of the labeled lipids in the blood. TG and cholesterol absorption were modestly but reproducibly decreased in the Tm6sf2-/- mice, especially in the early time points (Fig. 11C).

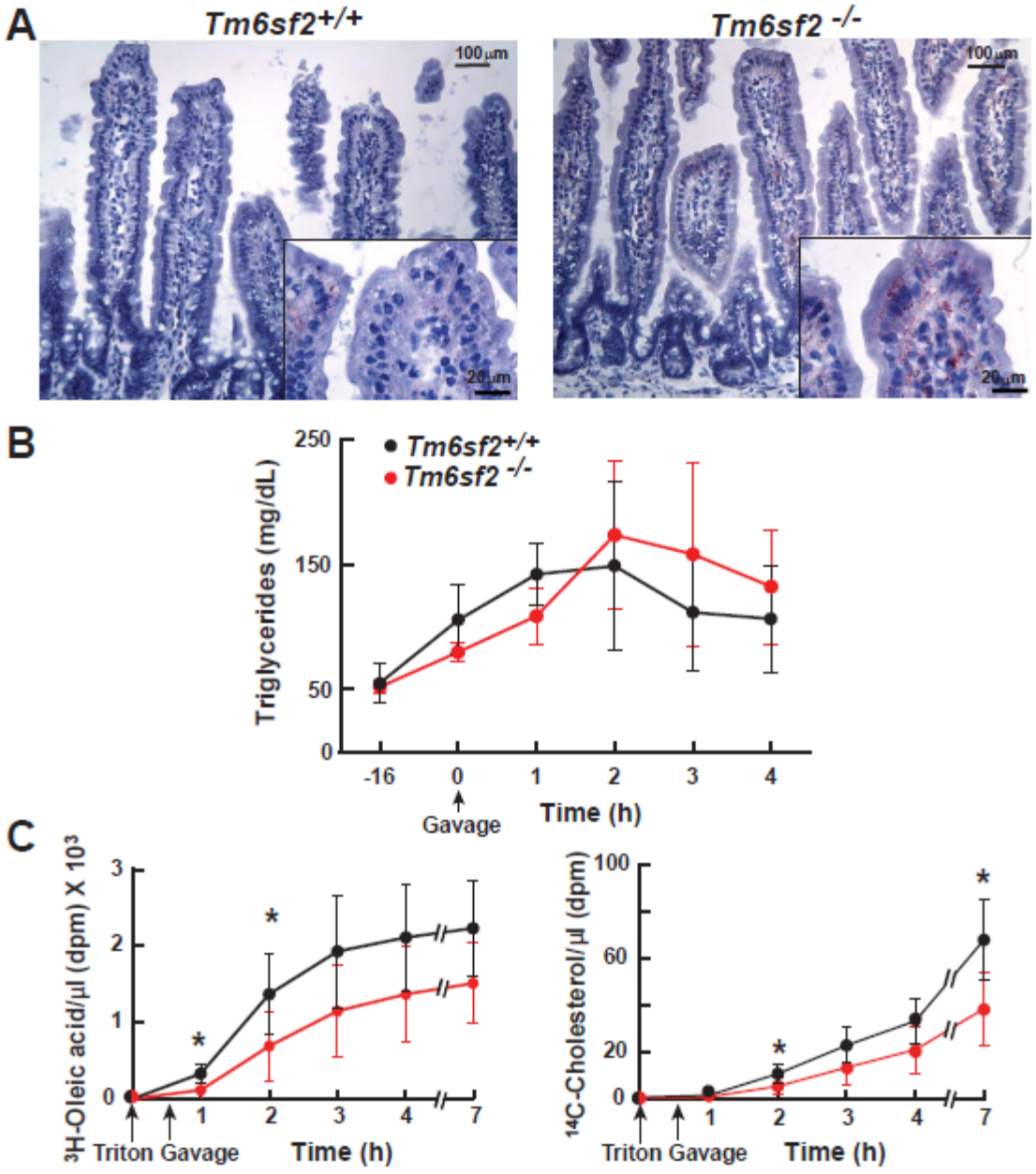


Figure 11. Intestinal accumulation and absorption of lipid in Tm6sf2 -/- mice. *(A) Jejunal sections prepared from male mice (18 weeks old) fed a high-fat diet for 12 weeks were stained with Oil Red O and visualized using a Leica microscope (DM2000) at 10X and 63X (inset) magnification as described in Methods. (B) Chow-fed male mice (n=4/group, 16 week-old) were fasted for 16 h and then gavaged with corn oil (10 μL/g). Blood was collected at the indicated times after gavage and plasma levels of TG was measured by enzymatic assay. The experiment was repeated, and the results were similar. (C) Chow-fed female mice (n=6/group, 12 week-old) were fasted for 16 h before the injection of Triton WR1339 (500 mg/kg) into the tail vein. A total of 30 min after the injection, mice were gavaged with corn oil (50 μl) supplemented with 1 μCi*

*[^{14}C]-cholesterol (50.8 mCi/mmol) and 5 μCi [^{3}H]-labeled oleic acid (54.5 Ci/mmol). Blood was collected at the indicated times before and after injection of Triton WR1339, and the radioactivity in the plasma was measured by scintillation counting at the shown time points. *P<0.05*

To define the subcellular location of TM6SF2, we co-stained primary hepatocytes of WT mice with antibodies against TM6SF2 and proteins that localize predominantly to ER (calnexin), the cis/medial Golgi (RCAS1), and Golgi matrix (giantin) (Fig. 12A). Anti-TM6SF2 antibody showed some degree of co-staining with each marker, indicating that the polytopic protein is present in both ER and Golgi complex. To confirm these results, we separated mouse liver microsomes by immunoaffinity chromatography on magnetic beads coated with anti-calnexin (ER) and anti-GM130 (Golgi complex) antibodies (Fig. 12B). As expected, the immuno-isolated ER fraction was enriched in BiP (GRP78, HSPA5) and calnexin, and depleted of Golgi SNAP receptor complex member 1 (Gos28, GOSR1), a protein located predominantly in the Golgi complex [88]. Reciprocal patterns of expression of these proteins were seen in the Golgi fraction. Immunodetectable TM6SF2 was detected in both the ER and Golgi fractions.

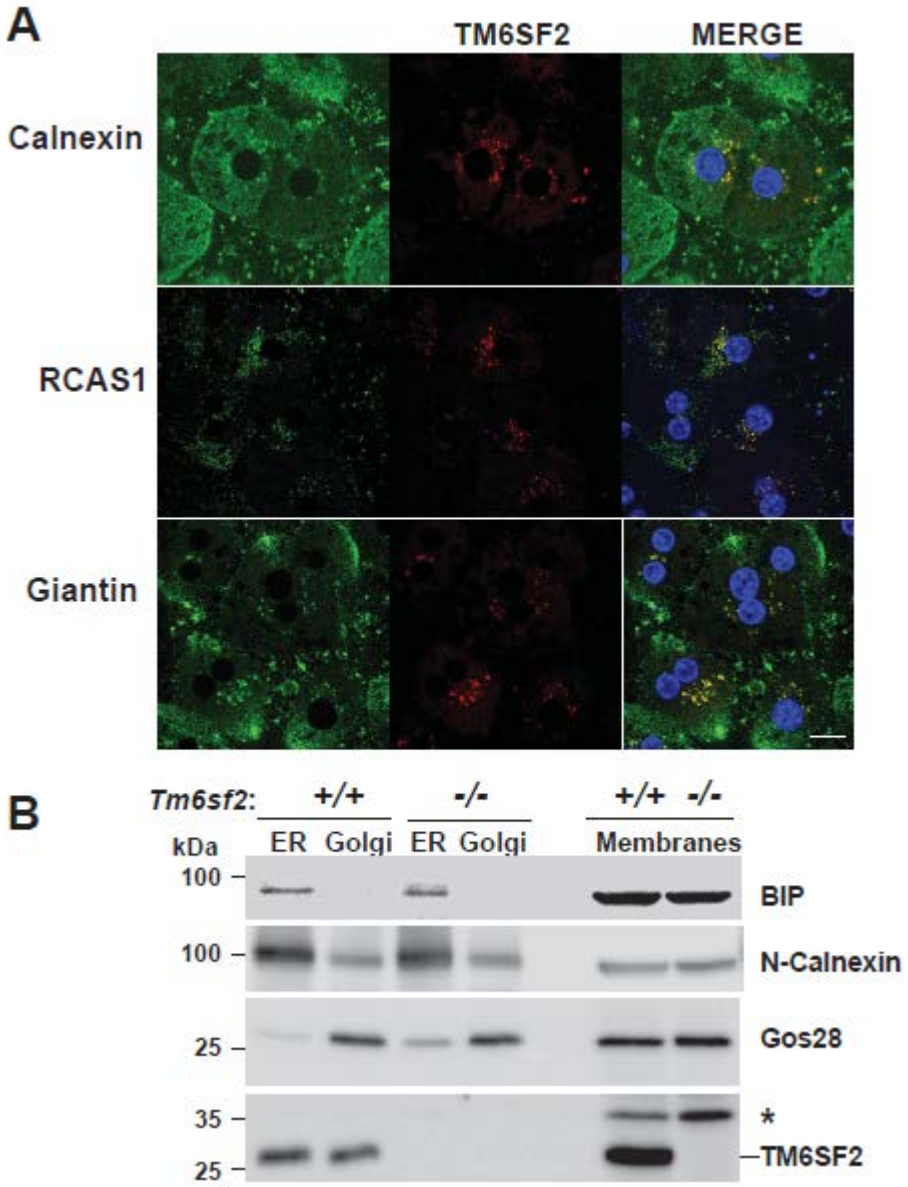


Figure 12. Subcellular localization of TM6SF2. *(A) Primary hepatocytes from 8-week old female mice on a chow diet were plated on collagen-coated coverslips for 4h, fixed and stained with antibodies against markers for the ER (calnexin, CANX), cis-Golgi [receptor binding cancer antigen expressed on SiSo cells (RCAS1)] and Golgi, (Giantin, GOLGBI), (green, left column) and TM6SF2 (red, middle column). The merged signal from both channels (yellow, right column) shows subcellular co-localization. All images were taken using a 63X oil immersion objective. Scale bar, 10 μm. (B) Immunoaffinity isolation of ER and Golgi complex from mouse liver. ER and Golgi fractions were prepared from mouse liver microsomes by immuno-affinity chromatography as described in Methods. Microsome membranes were dissolved in RIPA buffer, and equal volumes were separated on SDS (10%)-PAGE and immunoblotting as described in*

*Methods. BiP, Binding immunoglobulin protein; Gos28, Golgi SNAP receptor complex member 1; *, nonspecific band.*

To determine if the neutral lipid in the liver co-localizes with TM6SF2, we stained primary hepatocytes from WT and KO mice with BODIPY, a neutral lipid dye and an antibody to TM6SF2. The two signals did not colocalize in WT animals (Fig.13A). We confirmed that the antibody for TM6SF2 was specific since no signal was seen in hepatocytes from the KO animals. To determinate if the lipid that accumulates in the liver is in lipid droplets, we stained sections with a protein marker of LDs, perilipin 2 (PLIN2) (Fig. 13B). The neutral lipid co-localized with the PLIN2 staining. No co-localization of BODIPY staining was seen with markers from the ER (calnexin, CANX) or Golgi complex (giantin, GOLGBI) (data not shown). These findings are consistent with the bulk of neutral lipid in KO mice being located in cytoplasmic LDs.

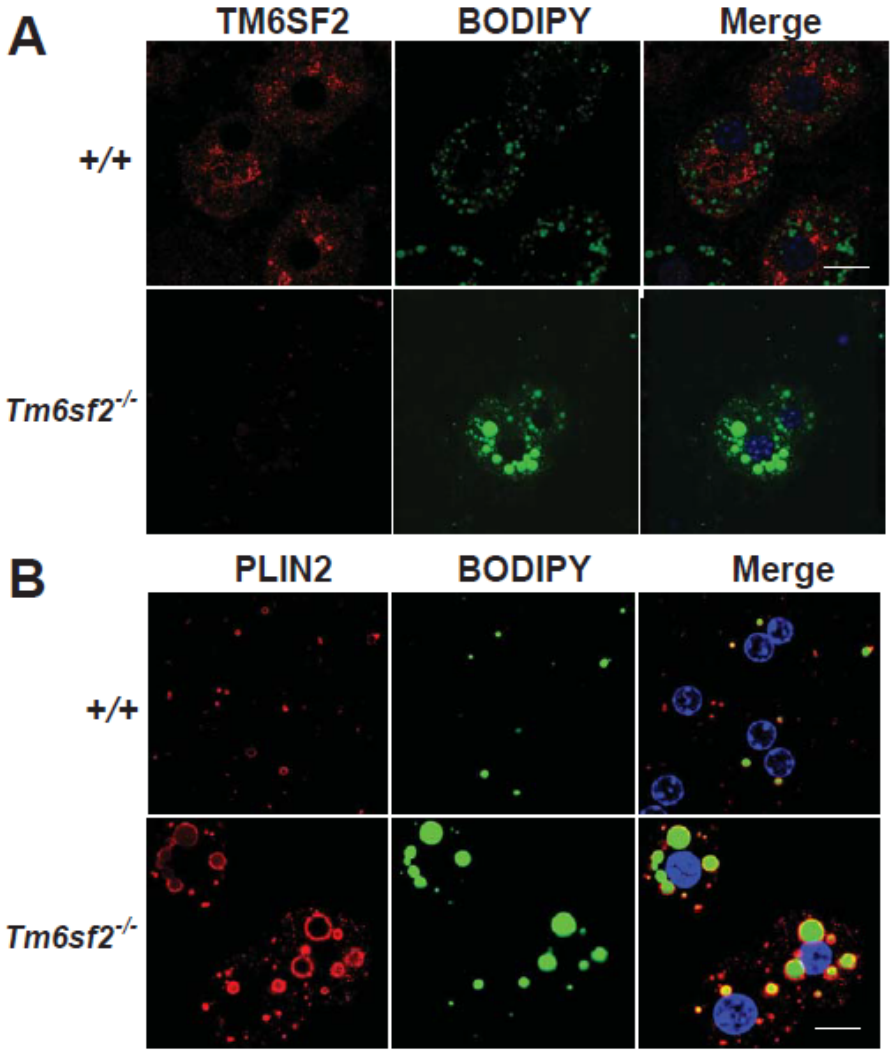


Figure 13A and B. Localization of TM6SF2 and neutral lipids in mouse primary hepatocytes. *(A) Primary hepatocytes from 8-week old female WT mice that were fed chow ad lib were plated on coverslips for 4 h, fixed, and stained with BODIPY and with an antibody against TM6SF2 and (B) the LD marker perilipin 2 (PLIN2). All images were taken using a 63X oil immersion objective. Scale Bar, 10 μm.*

The exclusion of TM6SF2 from LDs was confirmed by cell fractionation studies. In livers of WT mice fed sucrose to stimulate LD formation, PNPLA3 was detected exclusively in the LD fraction, whereas all immunodetectable TM6SF2 was located in the membrane fraction (Fig. 13C). Therefore, unlike PNPLA3, TM6SF2 is not a LD protein.

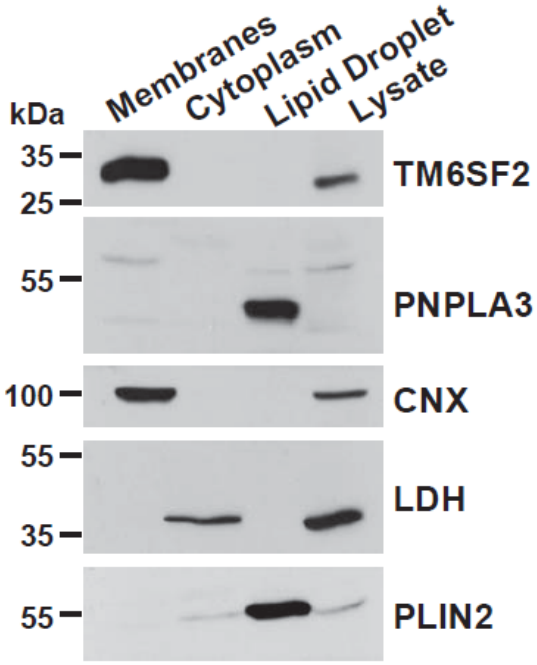


Figure 13C. Subcellular localization of TM6SF2. *Female C57Bl/N mice (14-weeks old) were fed a high-sucrose diet for 2 weeks. Feeding was synchronized for 3 days, and then the mice were sacrificed at the end of the feeding cycle. Livers were homogenized and the LDs, membranes, and cytosol were separated by ultracentrifugation as described in Methods. Aliquots of proteins from the membrane and cytosolic fractions (50 μg each) and one-tenth of the LD protein was subjected to SDS (10%)-PAGE and immunoblotting as described in the Methods. Calnexin (CNX), lactate dehydrogenase (LDH) and perilipin 2 (PLIN2) were used as controls for the ER, cytosolic, and LD fractions, respectively. The experiments were repeated, and the results were similar. *, nonspecific band.*

To determine if the accumulation of lipid in the liver of Tm6sf2-/- mice are due to impaired TG secretion, as suggested by our previous experiments [2], we examined and compared plasma lipid and lipoprotein levels in KO and WT mice. Previously, we showed that inactivation of TM6SF2 in the liver using shRNAs was associated with a significant reduction in plasma levels of cholesterol and TG [2]. In the Tm6sf2-/- mice, plasma levels of cholesterol were reduced, reflecting a decrease in cholesterol in both the LDL and HDL fractions (Fig. 14A).

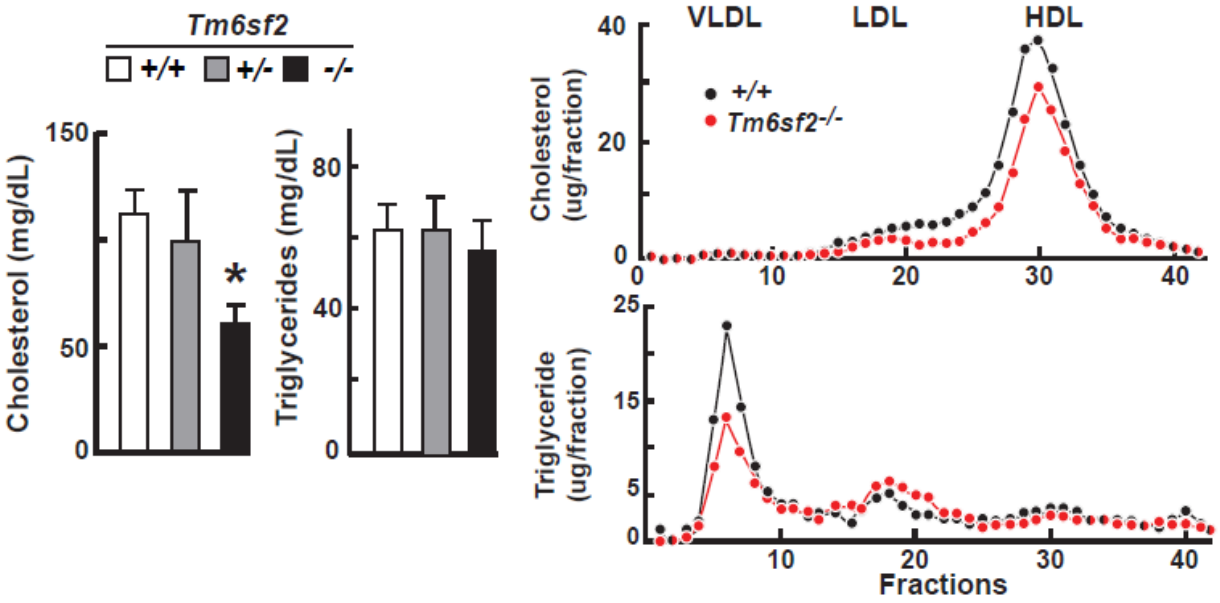


Figure 14A. Levels of lipids in plasma and fractions of lipoproteins in Tm6sf2 -/- mice. *(Left) Plasma TG and cholesterol levels were measured in chow-fed male WT, Tm6sf2+/- and Tm6sf2 -/- mice (n=5/group, 13-week-old) using enzymatic assays. (Right) Fast protein liquid chromatography (FPLC) profiles of plasma samples pooled from WT and Tm6sf2-/- mice (4 male mice/group, 10-week-old). Cholesterol (top) and TG (bottom) was measured in each fraction. Experiments were repeated twice with similar results.*

Despite the reduction in cholesterol, circulating levels of ApoE, a ligand for both the LDLR and LDLR related protein [89], and of ApoA1 and ABCA1, proteins required for HDL production [90], were similar in the two strains (data not shown). No differences were found in LDLR and SCARB1 (SR-B1), the major receptors that mediate clearance of LDL or HDL, respectively (data not shown). Also, circulating levels of PCSK9, a protein that regulates expression of hepatic LDLR [91], the major clearance pathway for ApoB-containing lipoproteins, were similar in the two groups of mice (Fig. 14B). Tm6sf2-/- mice had less TG in the VLDL

fraction, but higher TG levels in the LDL fraction compared to WT mice (Fig. 14A, right). Despite the reduction in VLDL-TG, the levels of ApoB- 100 were similar in the KO and WT animals (Fig. 14C). Surprisingly, the levels of ApoB-48 were significantly and reproducibly higher in plasma of Tm6sf2-/- mice (Fig. 14C). The reason for the increase in ApoB-48 levels in the KO mice is not known.

The decrease in VLDL-TG relative to plasma ApoB-100 levels implies a reduction in VLDL particle size in the Tm6sf2-/- mice. To directly assess VLDL particle size, VLDL were isolated by ultracentrifugation and visualized by electron microscopy using negative staining (Fig. 14D, left). The sizes of the VLDL particles varied over a wide range with strong leftward skewing in WT and in KO mice (Fig. 14D, right). The median particle diameter was smaller in KO than in WT animals (31 versus 38 nm, $P=1.3 \times 10^{-86}$). KO animals had more small particles (20-40 nm) and very few large particles (>50 nm) relative to WT mice.

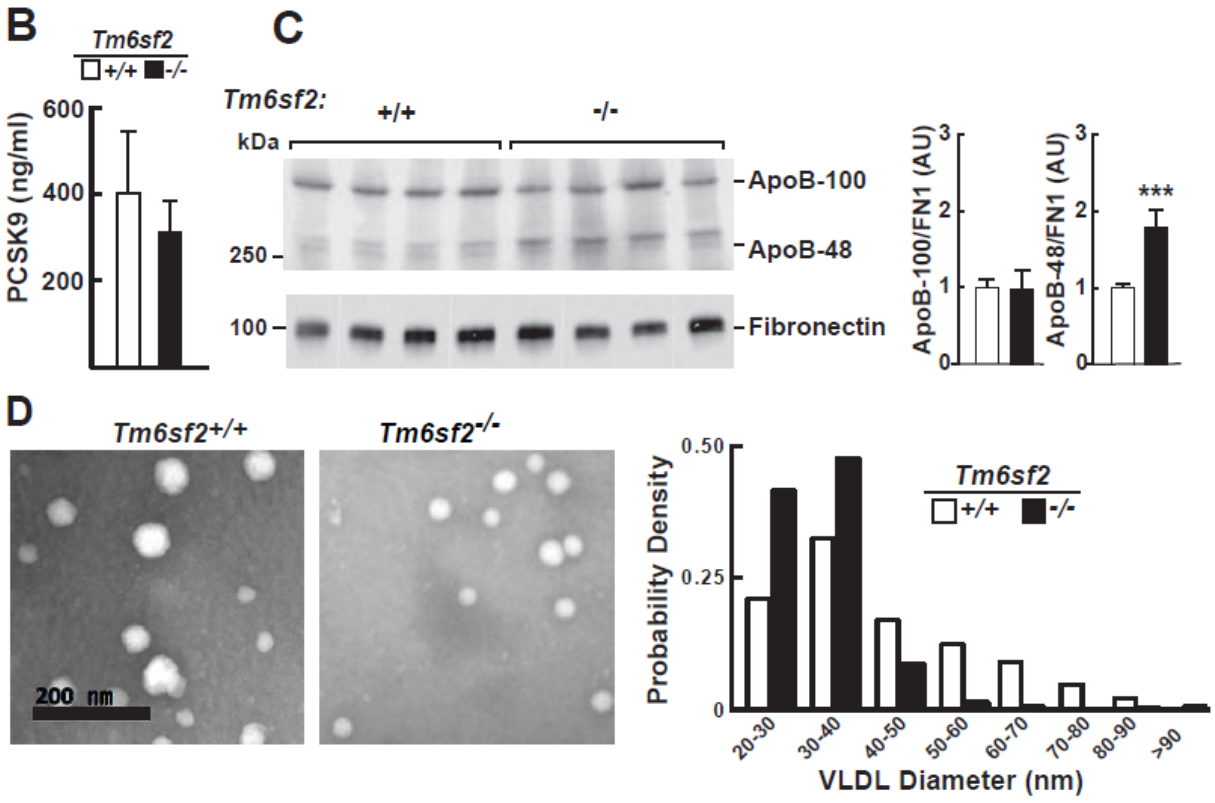


Figure 14B, C and D. Plasma levels PCSK9, ApoB in Tm6sf2 -/- mice and plasma VLDL size. *(B) Levels of plasma PCSK9 in 11-week-old male chow-fed male mice. Mice (n=5) were metabolically synchronized for 3 days by fasting from 8:00 AM - 8:00 PM and refed overnight. Blood was collected after the last refeeding period (at 8:00 AM) and the plasma levels of PCSK9*

*were detected using an ELISA as described in Methods (C) Plasma (0.2 µl) was size-fractionated on a 4-12% gradient SDS-PAGE and levels of ApoB-48 and ApoB-100 were determined by immunoblot analysis using a rabbit polyclonal antibody (Abcam, ab20737, 1:1000). The signal was detected and quantified using a Licor Odyssey Fc Imager. Fibronectin was used as a loading control. Values are means (± SEM). (D) VLDL particles from the plasma of WT and KO female mice (n=3, 20 week-old) were visualized by electron microscopy as described in Methods. The size distribution of VLDL particles in 10 randomly selected images was analyzed and compared using ImageJ software. *P<0.05, ***P<0.001*

The observation that KO mice have smaller VLDL with reduced TG content but normal levels of ApoB-100 and ApoB-48 is consistent with a model in which the liver (and possibly also the intestine) secretes ApoB containing lipoproteins that are not fully lipidated. Alternatively, the clearance of large TG-rich lipoproteins may be selectively enhanced in the KO mice.

To establish if the low VLDL-TG in the KO mice was due to reduced VLDL-TG synthesis or to accelerated VLDL-TG clearance, we inhibited LPL using Triton WR-1339 (500 mg/kg) and monitored the appearance of TG in the circulation (Fig. 15A). The rate of TG secretion (determined from the slope of plasma TG levels plotted against time) was markedly reduced in KO animals when compared to WT littermates (225 ± 102 vs. 600 ± 95 mg/dL/h, P <0.0001).

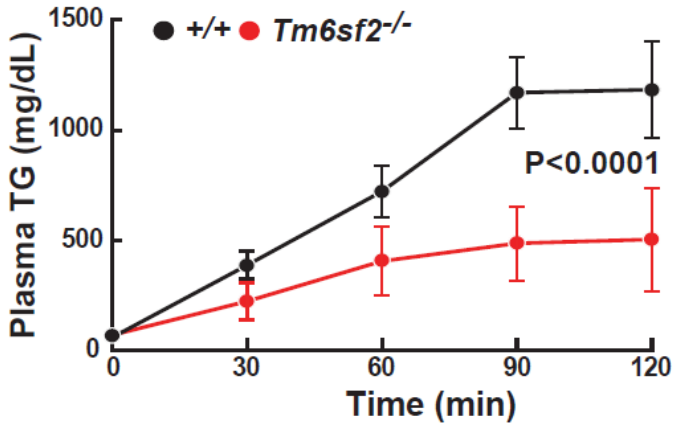


Figure 15A. VLDL secretion in WT and Tm6sf2 -/- mice. *The VLDL-TG secretion experiment. Triton WR-1339 (500 µg/g) was injected into the tail veins of chow-fed female mice (n= 4/group, 9-week-old) after 4h fasting. Blood was collected at the indicated time points, and plasma TG levels were measured. Mean TG levels at each point are shown. Rates of VLDL-TG secretion*

were determined by least-squares regression of plasma TG levels plotted against time. Slope estimates were determined from the linear portion of each graph, and compared using a Student's t-test.

To determine if the reduction in VLDL-TG accumulation was due to a defect in VLDL particle synthesis or secretion, we examined the levels of key proteins involved in these processes.

Chylomicron and VLDL synthesis are both severely compromised in two recessive disorders causing severe hypocholesterolemia: chylomicron retention disease, which is caused by Sar1b deficiency [92], and abetalipoproteinemia, which is due to deficiency of microsomal triglyceride transfer protein (MTTP) [93]. Levels of both SAR1B and MTTP did not differ between WT and KO mice (Fig. 15B). Synthesis of ApoB-100 containing lipoproteins is also compromised in mice lacking phospholipid transfer protein (PLTP), the enzyme that transfers phospholipids onto the nascent VLDL particle [94]. No differences in levels of PLTP proteins were found in the KO and WT mice (Fig. 15B). To determine if ApoB secretion was decreased in a similar manner to VLDL-TG, which would be consistent with a defect in whole particle secretion, we injected mice with ^{35}S-methionine and Triton WR-1339 and measured the incorporation of radiolabel into circulating ApoB- 48 and ApoB-100 after 90 min. The accumulation of label in ApoB-100 was similar in WT and KO mice (Fig. 15C). This finding, together with the similar levels of ApoB-100 in both the liver (Fig. 8A) and the plasma (Fig. 14C) is consistent with similar rates of ApoB-100 secretion in WT and KO mice. Thus, TM6SF2 absence does not compromise the secretion of VLDL particles; rather it reduces the lipidation of each particle. Interestingly, accumulation of label in ApoB-48 was greater in KO than in WT animals and comparable to the increase in plasma levels of ApoB-48 (Fig. 14C).

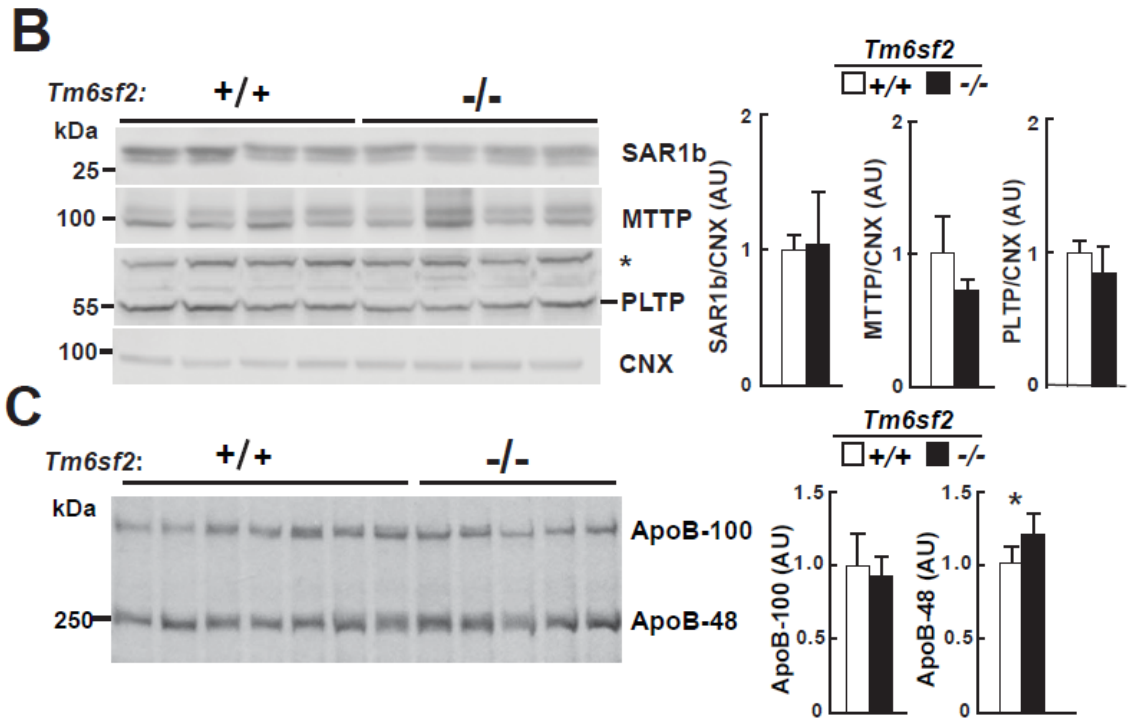


Figure 15B and C. Liver VLDL related protein levels and VLDL secretion in WT and Tm6sf2 -/- mice. *(B) Immunoblot analysis of Secretion Associated, Ras-Related GTPase 1B (SAR1b), microsomal TG transfer protein (MTTP), and phospholipid transfer protein (PLTP). Liver lysates (40 μg) were size-fractionated by 10% SDS–PAGE and then incubated with primary and secondary antibodies, as described in Methods. * - nonspecific band. (C)* ^{35}S *methionine incorporation in ApoB. Male mice (n=5, age=9 weeks) were synchronized on a 3-day regimen of fasting (8:00 AM - 8:00 PM) and refeeding (8:00 PM - 8:00 AM). At 8:00 AM on day 4 food was withdrawn and mice were fasted for 4 h and then injected via the tail vein with 200 μCi of* ^{35}S*-methionine (1,175 Ci/mmol) and Triton WR1339 (500 mg/kg). Blood was collected from the tail vein before and 45 and 90 min after the injection. Plasma samples from the 90 min point were treated with deoxycholate/trichloroacetic acid to precipitate proteins. The pellets were washed with acetone and dissolved in 200 μl of 2% SDS/1M urea buffer. Loading buffer (50 μl of 5x WB loading buffer) was added, and 20 μl aliquots were size-fractionated by 5% SDS-PAGE. Gels were dried and exposed to X-ray film (BIOMAX XAR) for 4 days in -80ºC. Band intensity was quantified using the Licor Image Studio Software. The experiment was repeated, and the results were similar. * P < 0.05*

In liver-specific Mttp-/- mice, no ApoB-100 is detected in the plasma, but levels of ApoB-48 are similar to WT animals [95], perhaps because lipoproteins containing ApoB-48 require less

lipidation for secretion. These data support the hypothesis that livers of Tm6sf2-/- mice secrete normal numbers of poorly lipidated lipoproteins. This finding suggests that TM6SF2 plays a role in the bulk lipidation step that leads to maturation of VLDL particles in the liver.

Mice lacking the lysophosphatidylcholine acyltransferase LPCAT3 have several features in common with Tm6sf2-/- mice, including hepatosteatosis, reduced VLDL TG secretion, smaller VLDL particles [19, 96] and accumulation of lipid in enterocytes [97]. Genetic disruption of Lpcat3 sharply reduces levels of arachidonic acid-containing phosphatidylcholines in the liver and circulating VLDL. Thus, we measured the distribution of fatty acids in lipids extracted from livers and plasma of KO and WT mice. Although some small differences in the fatty acid composition of phospholipids were observed in WT and KO mice, the arachidonic acid content of the phospholipid fraction was remarkably similar in the two groups (Fig. 16).

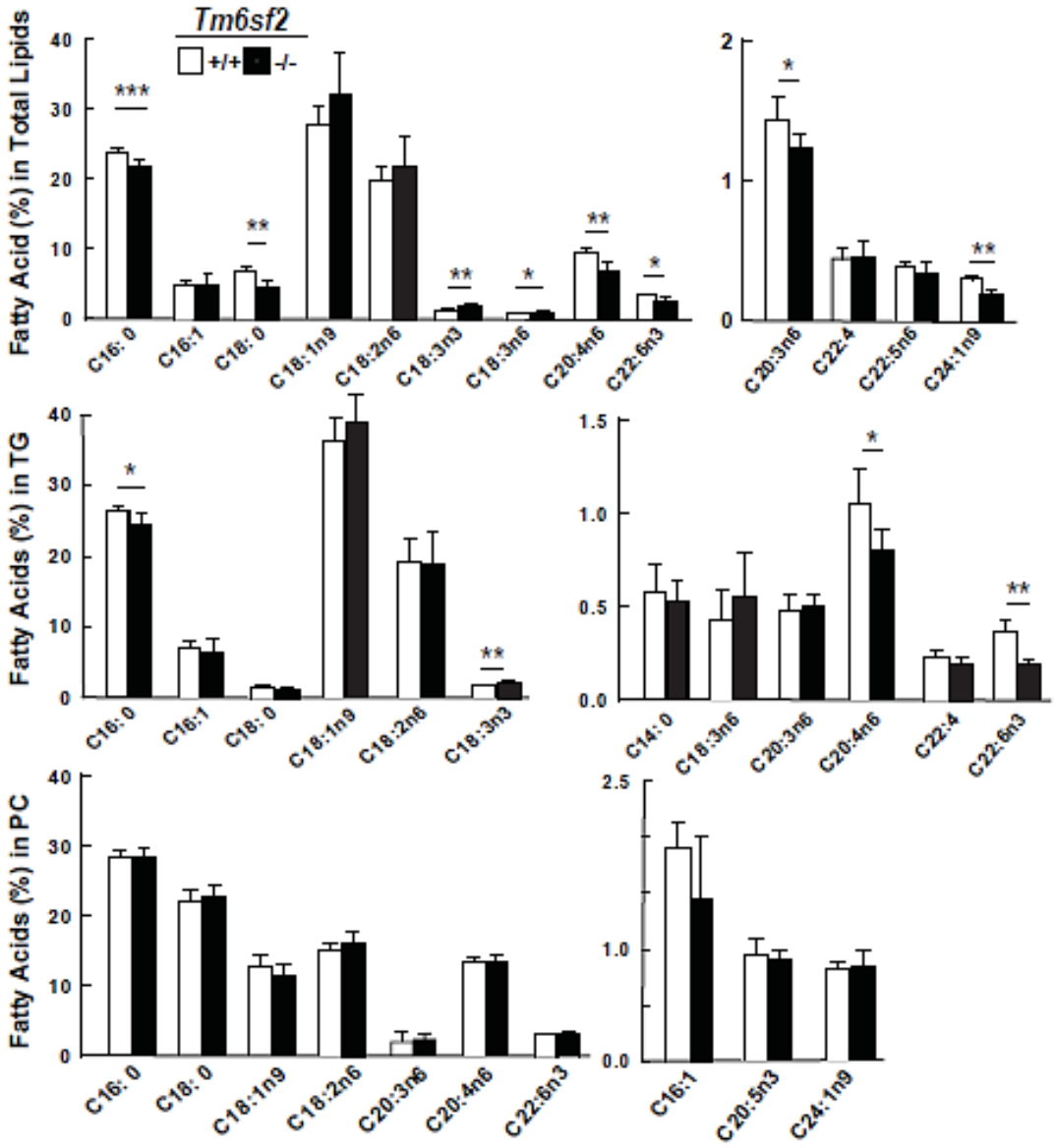


Figure 16. Fatty acid composition of hepatic lipids in WT and Tm6sf2-/- mice. *Livers were collected from chow-fed 13-week- old male mice (n=5) after a 4 h fast. The TG and PC fractions were separated by thin layer chromatography (TLC), hydrolyzed and derivatized with trimethylsilane. The fatty acid methylesters were quantified by gas chromatography (GC). Each value represents the mean ± SEM. *P < 0.05, **P<0.01 and ***P<0.001*

3.2. PNPLA3 study

Previously, the metabolic studies on the PNPLA3^{148MTg} (apoE promoter driven human PNPLA3 I148M transgenic) mice in our laboratory provided evidence consistent with PNPLA3-148M conferring both a loss and a gain, of function. Glycerol release from TG was reduced in primary hepatocytes from PNPLA3^{148MTg} mice when compared to the PNPLA3^{WTTg} mice. The endogenous fatty acid synthesis was significantly increased in both transgenic lines of mice when compared to nontransgenic animals, with the levels of TG synthesis being higher in PNPLA3^{148MTg} animals [46]. Although the PNPLA3^{148MTg} mice developed hepatic steatosis, this model has two significant limitations: 1) human PNPLA3 is only 68% identical to the mouse protein and may have non-physiological effects in mice, and 2) the PNPLA3 transgene is expressed constitutively at very high levels (ApoE promoter) and does not exhibit the ~80-fold changes in response to food intake that characterize the endogenous gene [44].

To develop an animal model that more accurately recapitulates the human phenotype associated with PNPLA3-148M, we altered codon 148 of Pnpla3 in the mouse genome from ATT to ATG, substituting methionine for isoleucine (Pnpla3^{148M}) (Fig. 16A). We also generated a second line of mice in which the catalytic serine at residue 47 was replaced with alanine (Pnpla3^{47A} mice). Sanger sequencing confirmed that the mutations were introduced without altering other sequences in the gene (Fig. 16B) and RT-PCR indicated that expression of the mutant and wild-type alleles was similar in mice fed a chow diet and that transcript levels increased by similar amounts when the two lines were fed a high-sucrose diet (74% of kCal from sucrose) (Fig. 16C).

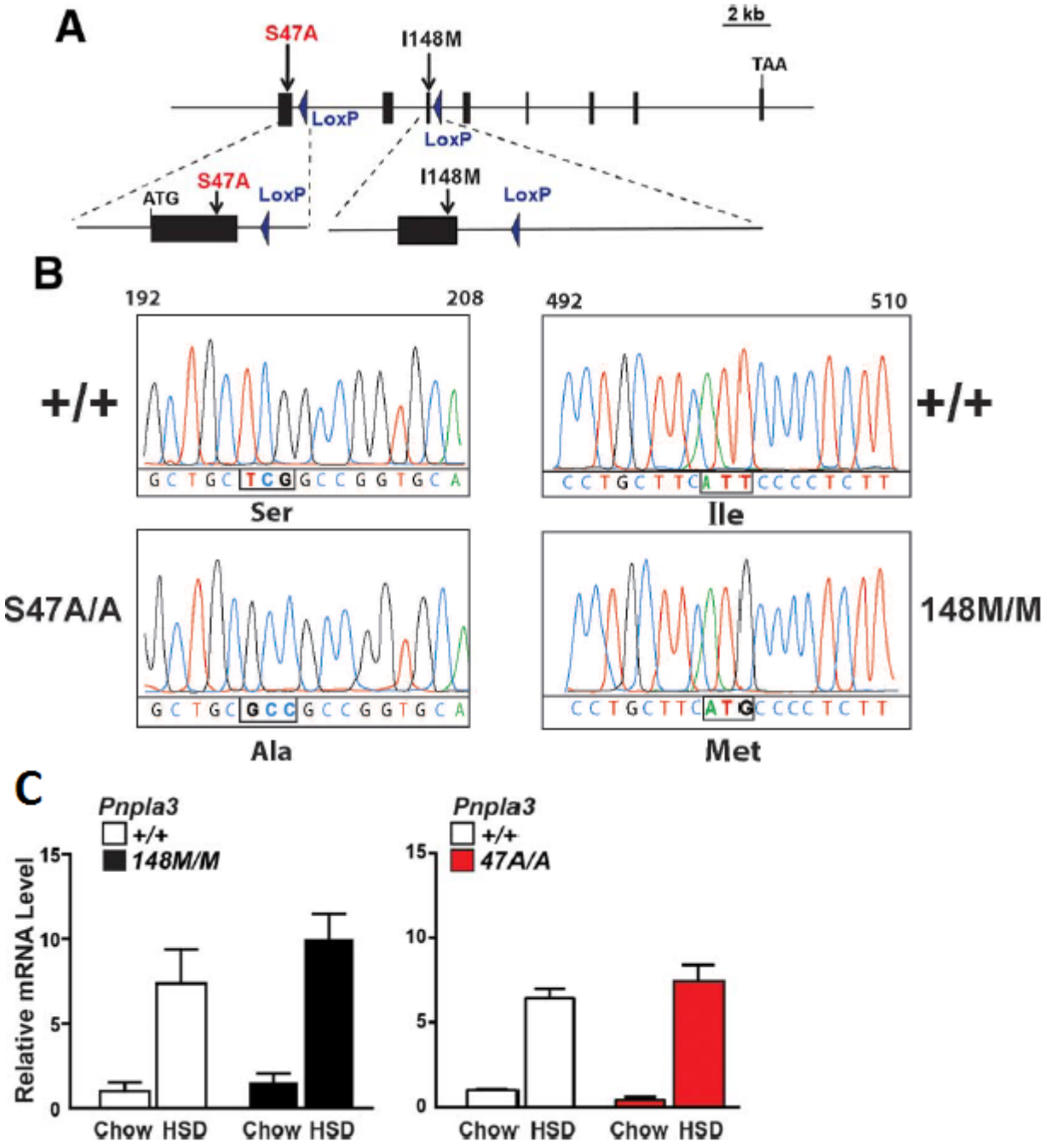


Figure 16. Generation of Pnpla3^{148M} and Pnpla3^{47A} KI mice. *(A) Diagram of Pnpla3 targeting strategy. Homologous recombination was used to replace either exon 1 (S47A) or exon 3 (I48M) with the corresponding mutant sequences. The selectable marker (Neo) located between two LoxP sites was then excised with Cre recombinase, leaving one LoxP site in the flanking intron. (B) Sanger sequencing of I148M and S47A KI mutations introduced in mouse Pnpla3. Genomic DNA was extracted from the tails of the KI mice and used for PCR-mediated Sanger sequencing*

of the regions of mouse Pnpla3 containing the bp substitution in codons 148 and 47, respectively. (C) Pnpla3 RNA levels in livers from female mice (n=4/group) consuming a chow diet or a high sucrose diet (HSD) for 4 weeks. Livers were collected at the end of the feeding cycle. Levels of Pnpla3 mRNA were determined using real-time PCR and normalized to values of chow-fed wild-type mice.

The KI mice were born in the expected Mendelian ratios and both litter sizes and gender distributions were unaltered by the introduction of the mutations. No differences in body weight, liver weight, fat mass, or lean mass were detected between the two strains of KI mice and their wild-type littermates (Fig. 17).

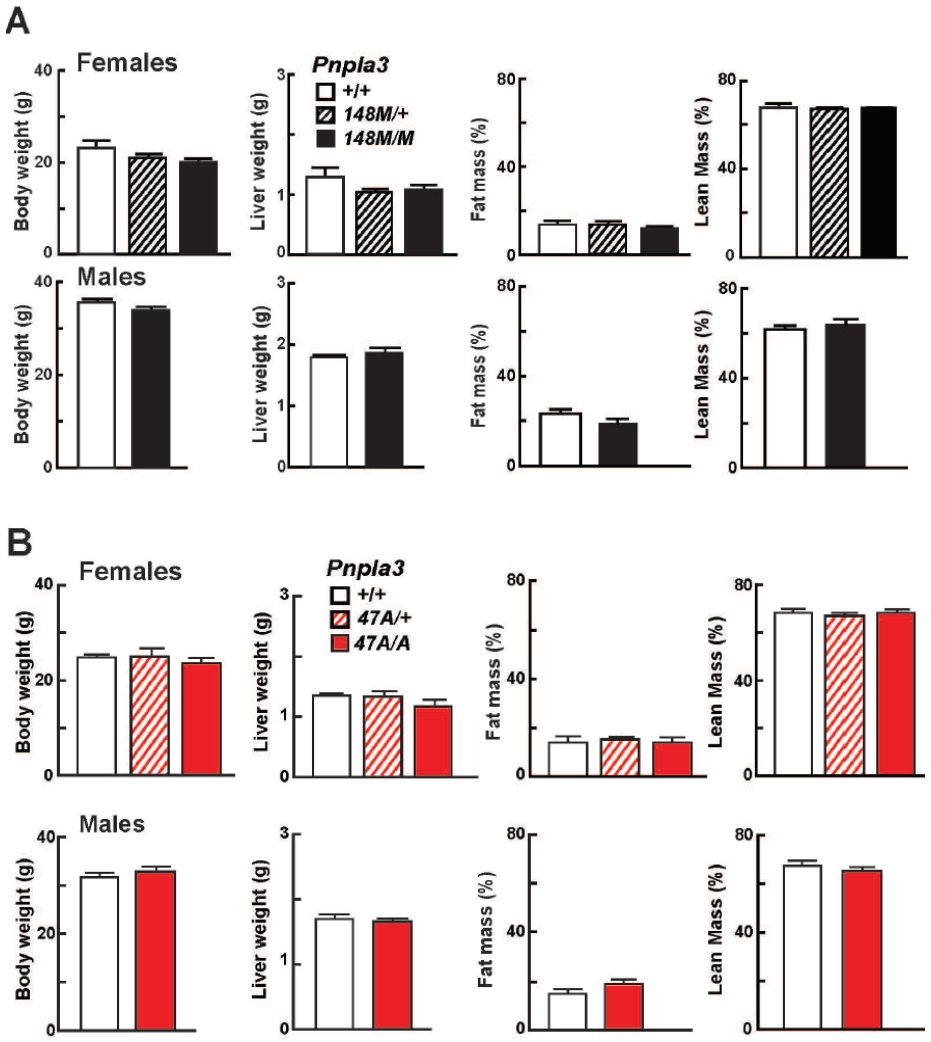


Figure 17. Body composition of wild-type (+/+), Pnpla3$^{148M/M}$ and Pnpla3$^{S47A/A}$ mice. *Female and male mice of the indicated genotypes were fed a high-sucrose diet for 4 weeks (n=5-*

6/group). Total lean mass and fat mass were measured by NMR and normalized to body weight as described in the Methods.

Morphological analysis of the adipose tissue revealed no differences in adipocyte size or conformation (Fig. 18).

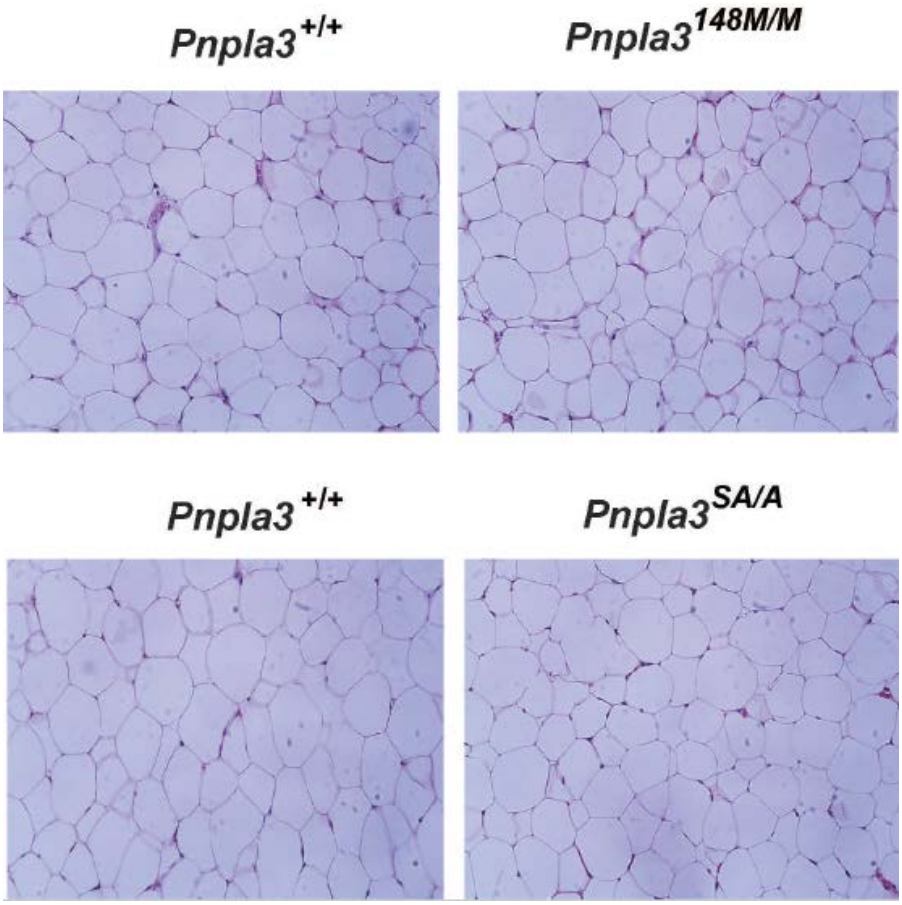
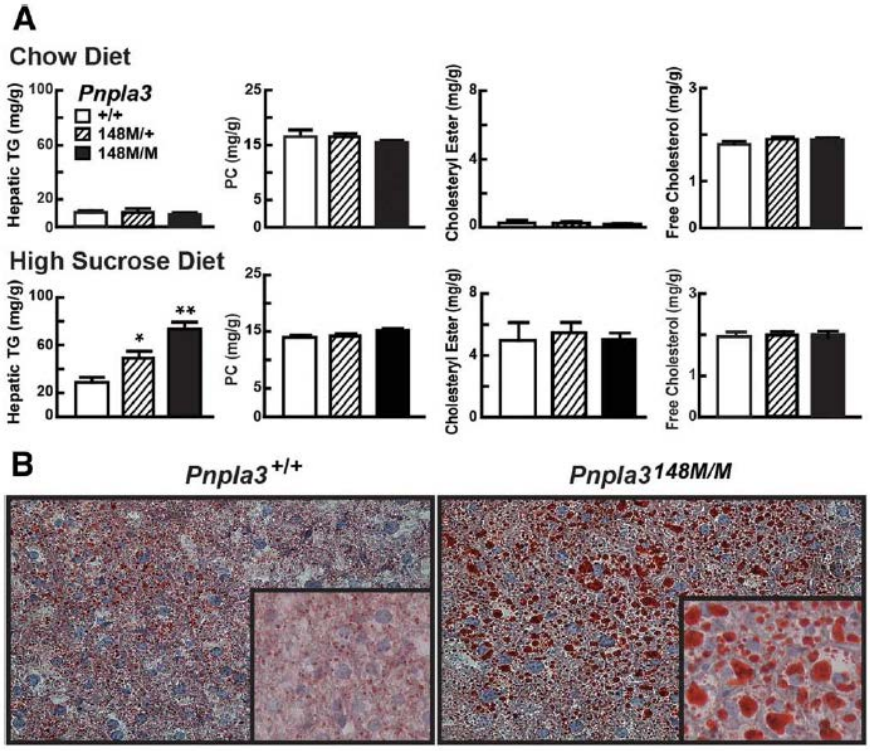


Figure 18. Sections from white adipose tissue (WAT) of wild-type (+/+), Pnpla3$^{148M/M}$ and PNPLA3S$^{47A/A}$ mice. *11-week-old female mice (n=4-5/group) of the indicated genotypes were fed a high-sucrose diet for 4 weeks. Adipose tissue was collected and stained with hematoxylin and eosin as described in the Methods. Magnification 20X.*

On a chow diet, the levels of TG, phosphatidylcholine (PC), cholesteryl esters, and free cholesterol in the livers of female mice that were heterozygous (Pnpla3$^{148M/+}$) or homozygous (Pnpla3$^{148M/M}$) for the Pnpla3^{148M} allele did not differ significantly from those of wild-type littermates (Pnpla3$^{+/+}$) (Fig. 19A, top). After 4 weeks on a high-sucrose diet, hepatic TG content was higher in female Pnpla3148$^{M/+}$ and Pnpla3148$^{M/M}$ mice than in wild-type animals (Fig. 19B, bottom). Levels of hepatic TG in heterozygous mice were intermediate between those of wild-

type and homozygous KI animals. Similar responses were observed in male mice (data not shown). No differences in circulating lipid levels were seen between the two lines (Fig. 20). Hematoxylin and eosin staining of liver sections from the Pnpla3$^{148M/M}$ mice revealed no differences in cellular morphology (data not shown). Oil Red O staining of neutral lipids was markedly increased in Pnpla3$^{148M/M}$ mice (Fig. 18B), and the size distribution of their lipid droplets was shifted to the right (Fig. 19C), reflecting a striking increase in the number of large droplets (>20 μm^2). In wild-type mice, the median area of the droplets was 4.3 μm^2 and fewer than 5% were greater than 22 μm^2. The median area of lipid droplets in the Pnpla3$^{148M/M}$ mice (8.7 μm^2) was double that of their wild-type littermates, and 28% were greater than 22 μm^2. Despite the increased lipid content of the livers in the Pnpla3$^{148M/1}$ and Pnpla3$^{148M/M}$ mice, no changes in circulating liver enzymes were observed in these animals (Fig. 19D).



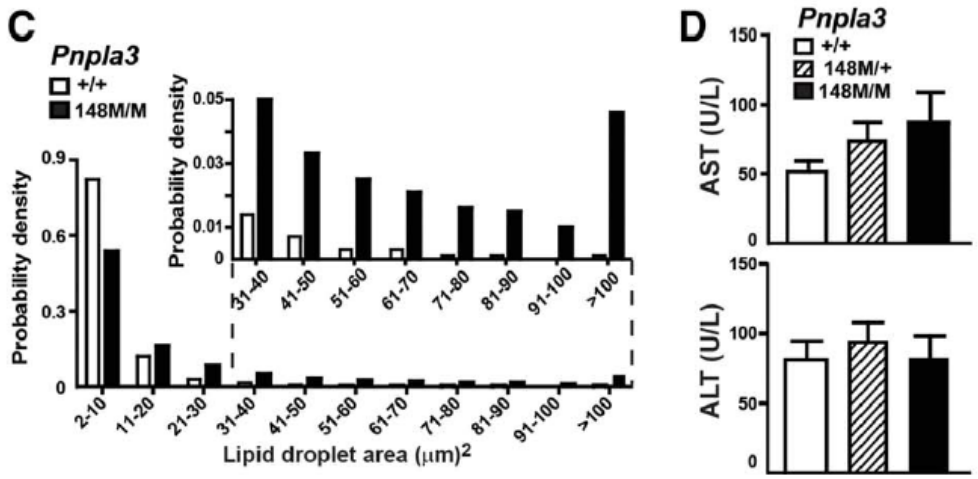


Figure 19. Hepatic TG content in Pnpla3^{I148M} KI mice. *Data in panels A-D are from the same experiment (n=5 mice/group). Tissue and plasma were collected at the end of the feeding cycle. (A) Hepatic lipid levels were measured in 12-week-old female wild-type (+/+), heterozygous (148M/+), and homozygous (148M/M) KI mice fed a chow diet or high-sucrose diet (HSD) for 4 weeks. (B) Liver sections from wild-type (I/I), and 148M/M KI mice on high sucrose diets were stained with Oil Red O and viewed using a Leica microscope (DM2000) (magnification: 20x and inset 63x). (C) Size distributions of hepatic lipid droplets in wild-type and Pnpla3148$^{M/M}$ mice. Oil Red O-stained slides were analyzed using ImageJ as described in Methods. (D) Plasma levels of aspartate aminotransferase (AST) and alanine aminotransaminase (ALT) in wild-type, 148M/I, and 148M/M KI mice. Values are means ± SEM. The experiment was performed twice with similar results *P<0.05, **P<0.001.*

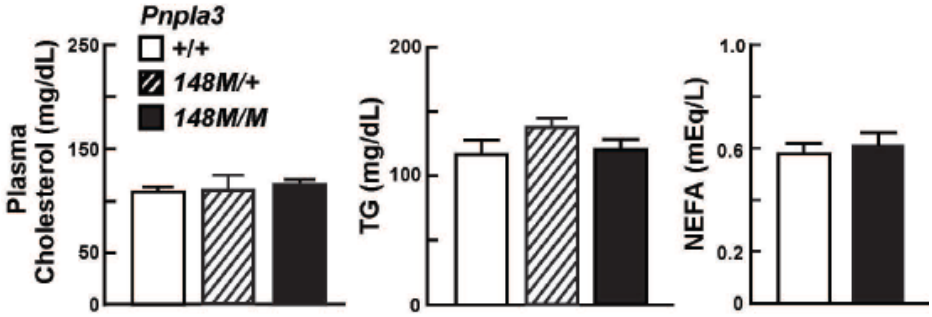
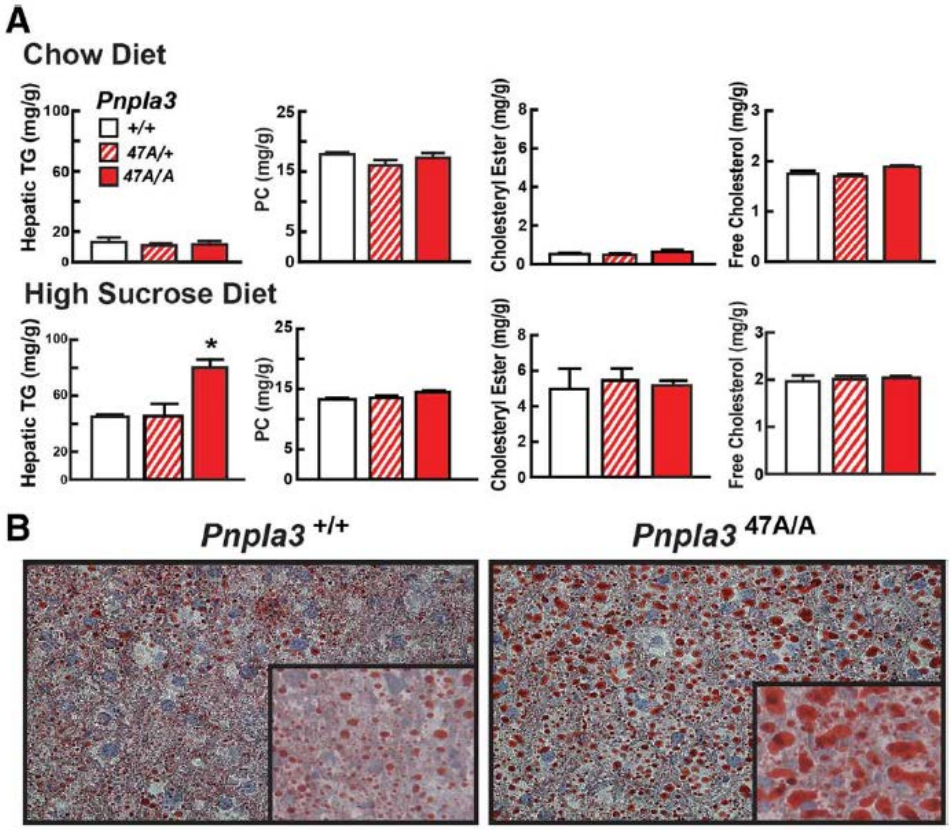


Figure 20. Plasma lipid and liver enzyme levels in Pnpla3$^{148M/M}$ and wild-type littermates. *Plasma was collected at the end of feeding cycle from 12-week-old female mice fed a high-sucrose diet for 4 weeks (5/group). Values are mean ± SEM.*

Similar results were obtained in Pnpla3^{S47A} KI mice. On chow diets the lipid levels of Pnpla3$^{47A/+}$ and Pnpla3$^{47A/A}$ KI mice were similar to those of their wild-type littermates in liver (Fig. 21A, top) and plasma (Fig. 22). On the high-sucrose diet, hepatic TG content was significantly higher in both the female (Fig. 21A, bottom) and male (data not shown) Pnpla347$^{A/A}$ mice than in their wild-type littermates. The hepatic TG content of mice heterozygous for the S47A substitution did not differ significantly from that of wild-type animals. Hepatic levels of PC, cholesteryl esters and free cholesterol did not differ between the lines on either diet (Fig. 21A) and no differences in circulating liver enzymes (Fig. 21D), or plasma lipids (Fig. 22) were apparent on the high-sucrose diet. The increase in hepatic TG content was associated with an increase in the number of large, Oil Red O-positive lipid droplets in the livers of the Pnpla3$^{S47A/A}$ animals (Fig. 21B, C).



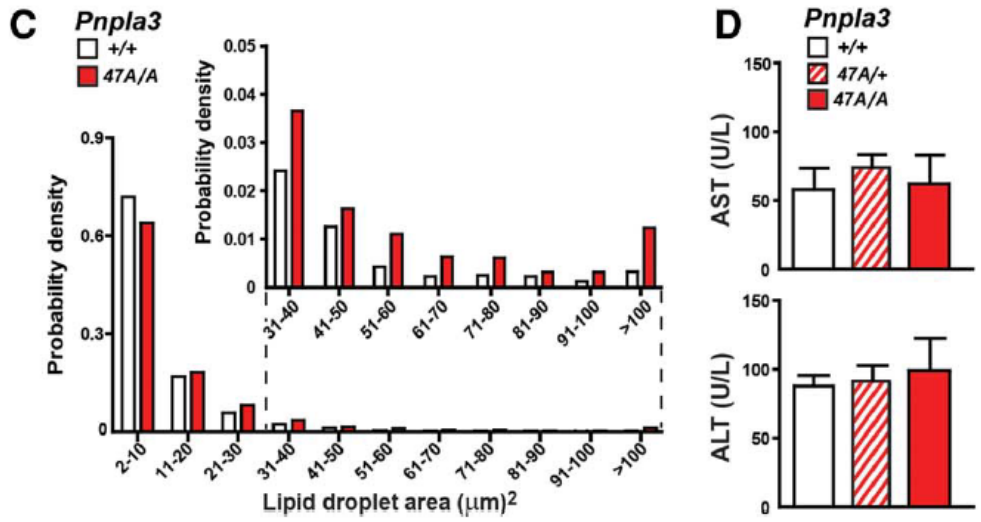


Figure 21. Hepatic TG content in Pnpla3^{S47A} mice. *Data in panels A-D are from the same experiment (n=7 mice/group). Tissue and plasma samples were collected at the end of the feeding cycle. The experiment was repeated with similar results. (A) Hepatic lipid levels were measured in 13- week-old female wild-type (+/+), S47A heterozygous (S47A/+) and homozygous (S47A/A) KI mice fed a chow diet or high-sucrose diet (HSD) for 4 weeks. (B) Liver sections from sucrose-fed wild-type and S47A/A KI mice were stained with Oil Red O and viewed using a Leica microscope (DM2000) (magnification: 20x, inset 63x). (C) Size distributions of hepatic lipid droplets in wild-type and homozygous (S47A/A) KI mice. Oil Red O-stained slides were analyzed using ImageJ as described in the Methods. (D) Plasma levels of liver enzymes (AST and ALT). Values are means ±SEM. *P=0.001.*

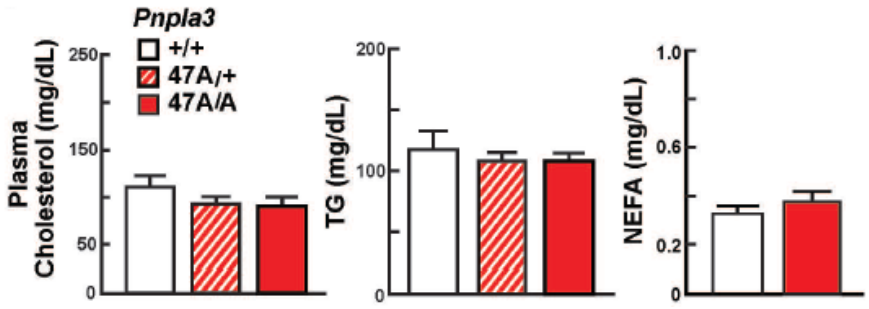


Figure 22. Plasma lipid and liver enzyme levels in PNPLA3$^{S47A/A}$ and wild-type littermates. *Plasma was collected at the end of feeding cycle from 12-week-old female mice fed a high-sucrose diet for 4 weeks (5/group). Values are mean ± SEM.*

High-level expression of human PNPLA3-148M in livers of transgenic mice was associated with increased mRNA levels of sterol regulatory element binding protein-1c (SREBP-

1c), which orchestrates the up-regulation of the fatty acid synthesis pathway, [98] and several of its target genes,[46] suggesting that the elevated TG content associated with the PNPLA3-148M variant is due to accelerated TG synthesis. In mice consuming high-sucrose diets, hepatic mRNA levels of SREBP-1c and its target genes were similar in Pnpla3$^{148M/M}$, Pnpla3$^{47A/A}$ and wild-type animals (Fig. 23). We also found no strain-specific changes in genes involved in TG lipolysis (hormone sensitive lipase [HSL], adipocyte TG lipase [ATGL]), fatty acid oxidation, hepatic inflammation, or fibrosis.

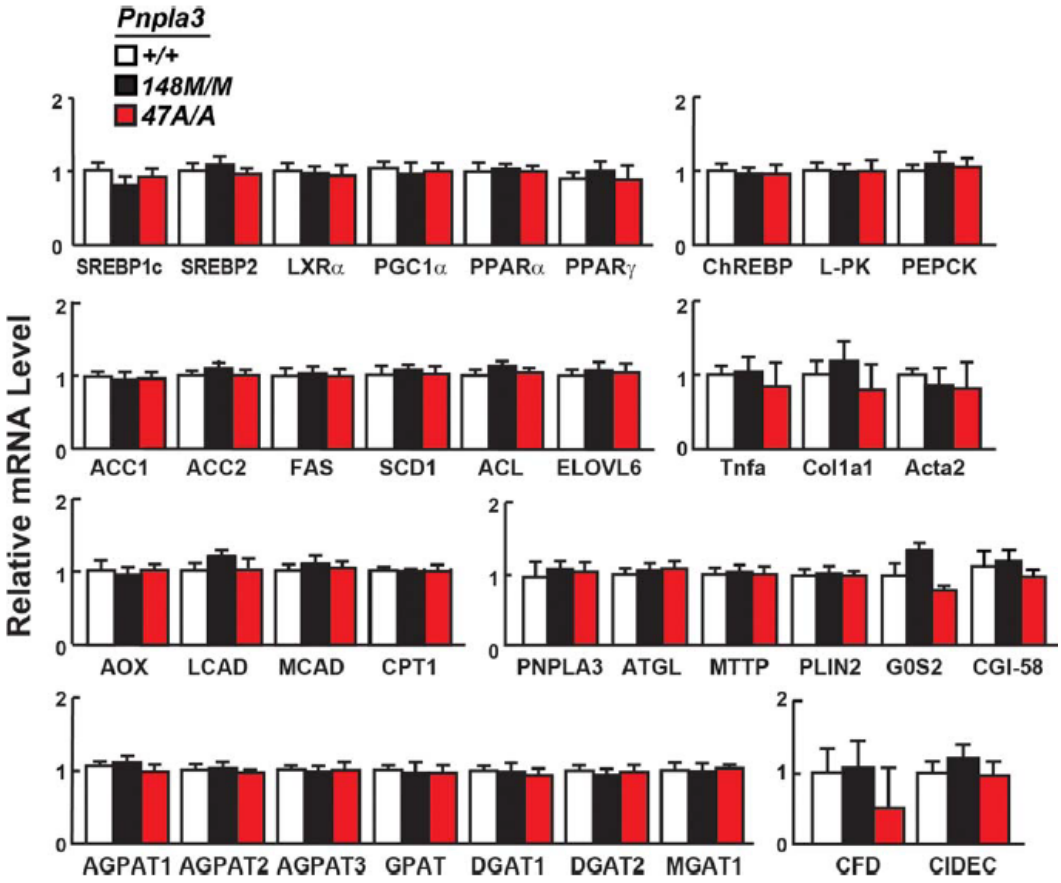


Figure 23. Relative mRNA levels of selected genes in livers of Pnpla3$^{148M/M}$ and Pnpla3$^{47A/A}$ mice with wild-type littermates (n=4/group). *The mice used in this experiment are described in the legends to Figs. 19. and 21.. mRNA levels were quantified by real-time PCR, normalized to levels of m36B4, and expressed relative to levels in wild-type mice. Values are means±SEM. SREBP1c, sterol regulatory element binding protein 1 isoform C; SREBP2, sterol regulatory element binding protein 2; LXRα, liver X receptor alpha; PGC-1α, PPARc coactivator 1α; PPARα, peroxisome proliferator-activated receptor alpha; PPARγ, peroxisome proliferator-activated receptor gamma; ChREBP, carbohydrate responsive element-binding protein; L-PK,*

liver pyruvate kinase; PEPCK, phosphoenolpyruvate carboxykinase; ACC1, acetyl-CoA carboxylase-1; ACC2, acetyl-CoA carboxylase 2; FAS, fatty acid synthase; SCD1, stearoyl-CoA desaturase-1; ACL, ATP citrate lyase; ELOVL6, ELOVL family member 6; TNFα, tumor necrosis factor alpha; COL1a1, collagen type-1a1; ACTA2, α-smooth muscle actin; AOX, acyl-CoA oxidase-1; LCAD, long-chain acyl-CoA dehydrogenase; MCAD, medium-chain acyl-CoA dehydrogenase; CPT1, carnitine palmitoyltransferase 1; ATGL, adipose triglyceride lipase; MTTP, microsomal TG transfer protein; PLIN2, perilipin 2; G0S2, G0/G1switch 2; CGI-58, comparative gene identification 58; AGPAT1-3, 1-acylglycerol-3-phosphate O-acyltransferase 1-3; GPAT, glycerol-3-phosphate acyltransferase; DGAT1, diglyceride acyltransferase-1; DGAT2, diglyceride acyltransferase-2; MGAT1, monoacylglycerol O-acyltransferase 1; CFD, complement factor D (adipsin); CIDEC, cell death-inducing DFFA-like effector C. The experiment was repeated twice and the results were similar.

Previously, we found a significant reduction in the proportion of long-chain polyunsaturated fatty acids in liver TG from mice expressing high levels of human PNPLA3-148M [46]. The fatty acid profiles of hepatic TG, cholesteryl esters, and phospholipids were remarkably similar in the KI mice and their wild-type littermates (Fig. 24). Stearate (C18:0) was modestly decreased in the liver TG fraction of Pnpla3$^{148M/M}$ mice, while palmitoleate and linoleate were slightly higher in liver TGs of Pnpla3$^{S47A/A}$ animals, but no consistent differences were observed in the two KI strains. Thus, the increased hepatic TG content in the two KI strains was not associated with any obvious alteration in the spectrum of fatty acids of hepatic glycerolipids.

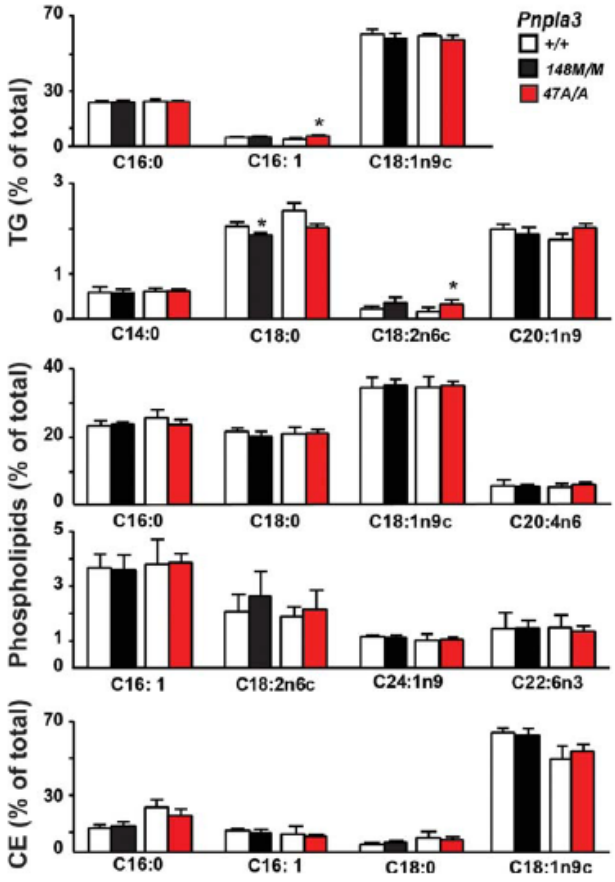


Figure 24. Fatty acid composition of hepatic lipids in wild-type, Pnpla3I$^{148M/M}$, and Pnpla3$^{S47A/A}$ KI mice. *Lipids were extracted from livers of 12-week-old female mice (n=6/group) fed a high-sucrose diet for 4 weeks as described in the Methods. The TG, phospholipid, and cholesterol ester fractions were separated by TLC, hydrolyzed, and derivatized with trimethylsilane. The fatty acid methylesters were quantified by gas chromatography. Each value represents the mean ±SEM. *P<0.05. The experiment was repeated, and the results were similar.*

On a high-sucrose diet, levels of PNPLA3 mRNA were similar in the livers of wild-type and KI mice (Fig. 25A). Liver TG content was 2-fold higher in KI mice (Fig. 25B). Thus, the differences in hepatic TG associated with sucrose feeding were not due to differences arising from altered PNPLA3 mRNA levels. Quantitative immunoblotting revealed a dramatic (40-fold) increase in PNPLA3 levels in lipid droplets from Pnpla3$^{148M/M}$ and Pnpla3$^{47A/A}$ mice compared to their wild-type littermates (Fig. 25C, D). No differences in levels of ATGL or HSL, the two major TG lipases in lipid droplets, were seen among the four groups. In contrast to these findings, the levels of CGI-58, a cofactor of ATGL that was barely detectable in the wild-type animals,

were increased ~5-fold in lipid droplets from both Pnpla3$^{148M/M}$ and Pnpla3$^{47A/A}$ mice. The major differences in levels of both PNPLA3 and CGI-58 were not accompanied by changes in levels of the three perilipins that are most abundant in hepatic lipid droplets: PLIN2, PLIN3, and PLIN5. The selective accumulation of CGI-58 in the lipid droplets suggested that PNPLA3, which resembles ATGL, may sequester CGI-58 on the lipid droplet, resulting in a reduction in TG hydrolase activity by ATGL.

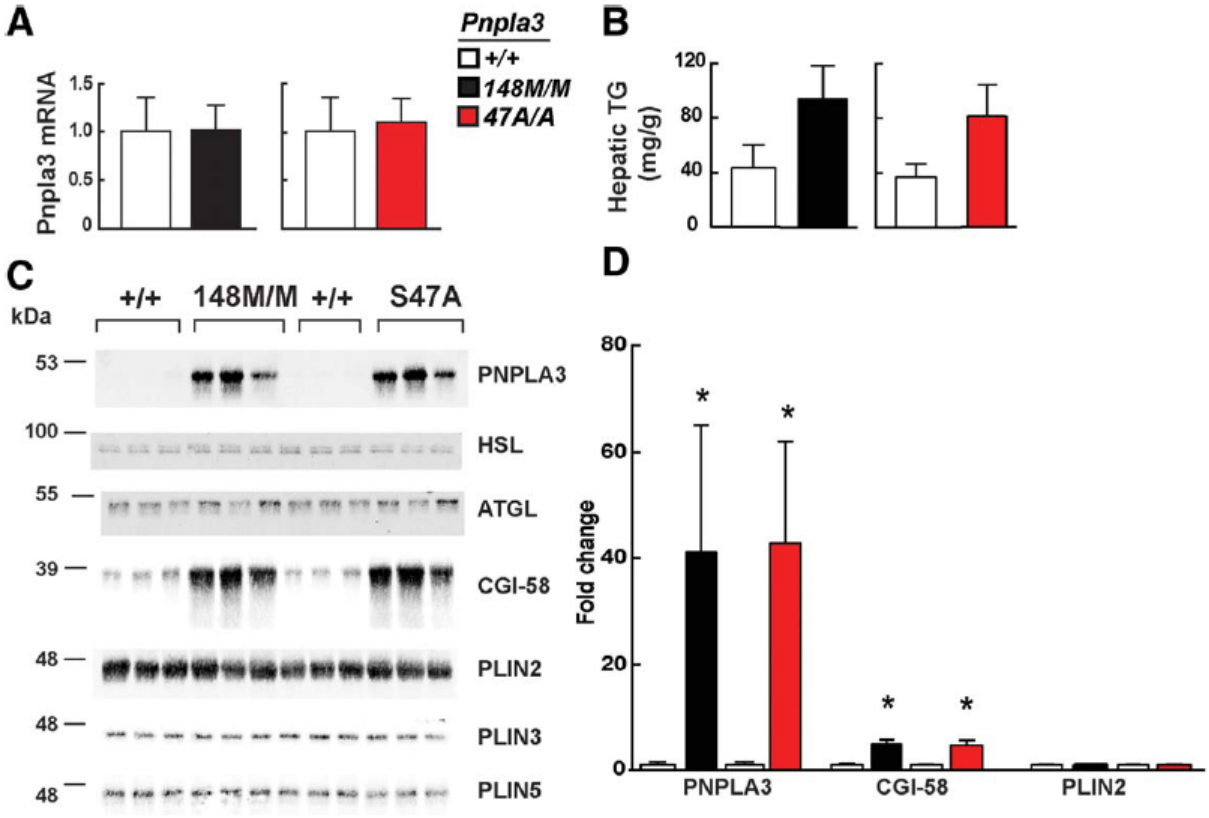


Fig. 25. Accumulation of PNPLA3 and CGI-58 protein on hepatic lipid droplets of Pnpla3$^{148M/M}$ and Pnpla3$^{47A/A}$ KI mice. *Livers were harvested from 11-week-old female wild-type (+/+), Pnpla3$^{148M/M}$, and Pnpla3$^{47A/A}$ mice (n=5/group) fed a high-sucrose diet for 4 weeks. (A) Hepatic PNPLA3 mRNA levels were determined using real-time PCR, as described in the Methods. (B) Hepatic TG levels were measured by an enzymatic assay as described in the Methods. (C) Quantitative immunoblot analysis of lipid droplets (LD) isolated from 3 mice per group. A total of 3 µg of LD protein was size-fractionated on a 10% SDS-PAGE gel and immunoblot analysis was performed as described in the Methods. (D) Immunoblots were quantitated using Licor infrared fluorescent imaging as outlined in the Methods. Values are mean ± SEM. *P<0.05. The experiment was repeated twice, and the results were similar.*

To determine if PNPLA3^{148M} accumulates secondary to the development of fatty liver, we measured PNPLA3 mRNA and protein in livers from mice fed a high-fat diet for 12 weeks. Hepatic levels of PNPLA3 mRNA were similar in wild-type and KI animals (Fig. 26A), but in both lines, the levels were much lower on the high-fat diet (C_t=24.5) than on the high-sucrose diet (C_t=22). Despite the 3 to 4-fold increase in hepatic TG levels in fat-fed mice (Fig. 26B), PNPLA3 protein was barely detectable in the lipid droplets of wild-type or the KI animals (Fig. 26C). Quantitative immunoblotting of lipid droplets from the fat-fed mice indicated that PNPLA3 levels were 6-fold higher in the Pnpla3$^{148M/M}$ mice than in their wild-type littermates (Fig. 26C, D). Nonetheless, PNPLA3 levels in both strains were far lower in fat-fed than in sucrose-fed animals. Levels of CGI-58 were increased 2.5-fold in the KI mice.

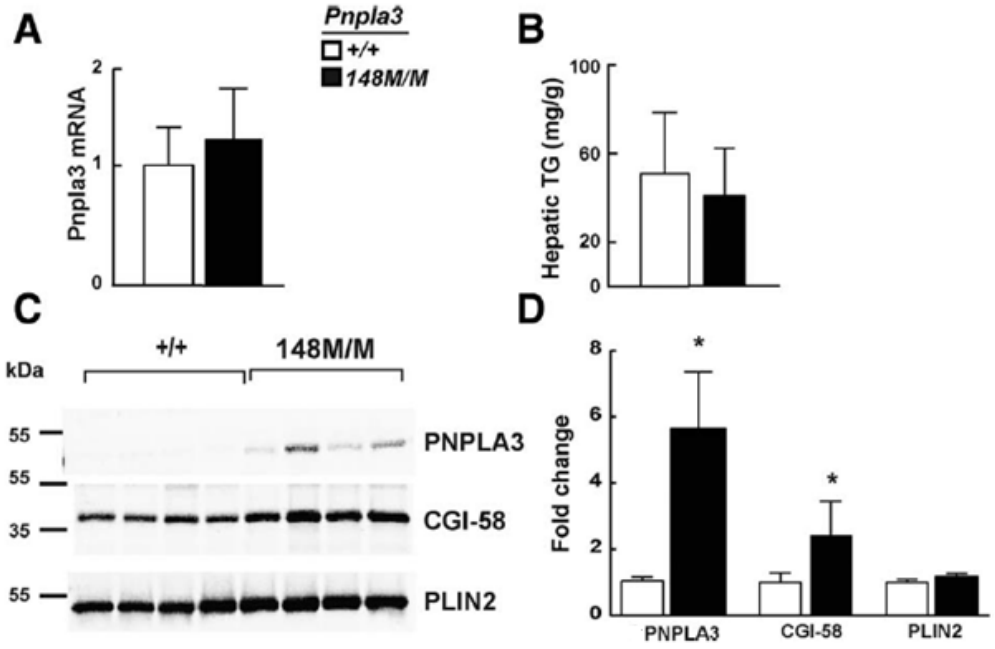


Figure 26. PNPLA3 on lipid droplets from fat-fed mice. *Livers were harvested from 17-week-old female wild-type (+/+) and Pnpla3$^{148M/M}$ mice (n=5/group) fed a high fat diet for 12 weeks. (A) Hepatic PNPLA3 mRNA levels were determined using real-time PCR, as described in the Methods. (B) Hepatic TG levels were measured by an enzymatic assay as described in the Methods. (C) Immunoblot analysis of lipid droplets isolated from 4 mice per group was performed as described in the legend to Fig. 25, and (D) quantified as described in the Methods. Values are means ±SEM. *P<0.05.*

Despite the feeding with High Fat Diet of PNPLA3^{148M} and (+/+) mice for 12 weeks, we did not find any changes in body mass or insulin resistance in those animals (Fig. 26). Also, none of the used diets did change glucose tolerance tests in PNPLA3^{148M} and (+/+) mice (Fig. 27).

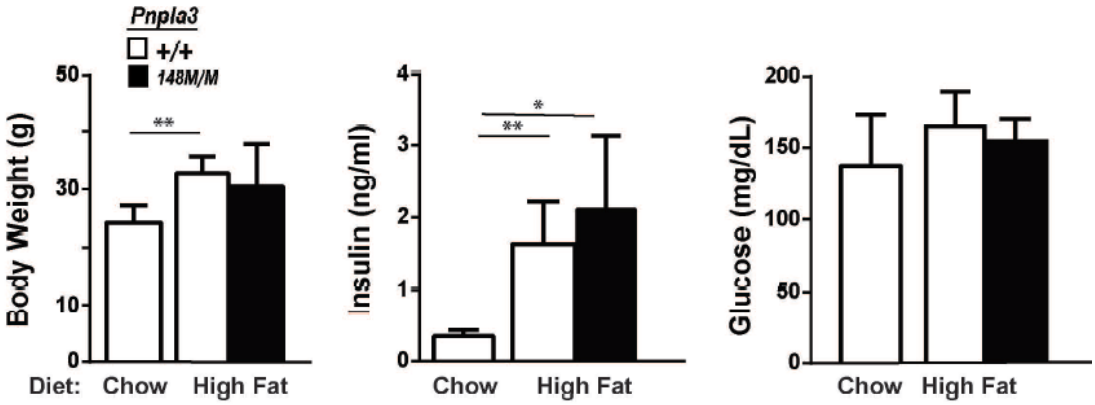


Figure 26. Pnpla3$^{148M/M}$ KI mice do not have increased insulin resistance. *Wild-type and Pnpla3$^{148M/M}$ KI female mice (n=5/group, 17 weeks old) were fed a high-fat diet for 12 weeks. Venous blood was obtained from the tail vein of mice after a 4 hour fast, and the levels of insulin and glucose were measured as described in the Methods. *P < 0.05, **P < 0.01.*

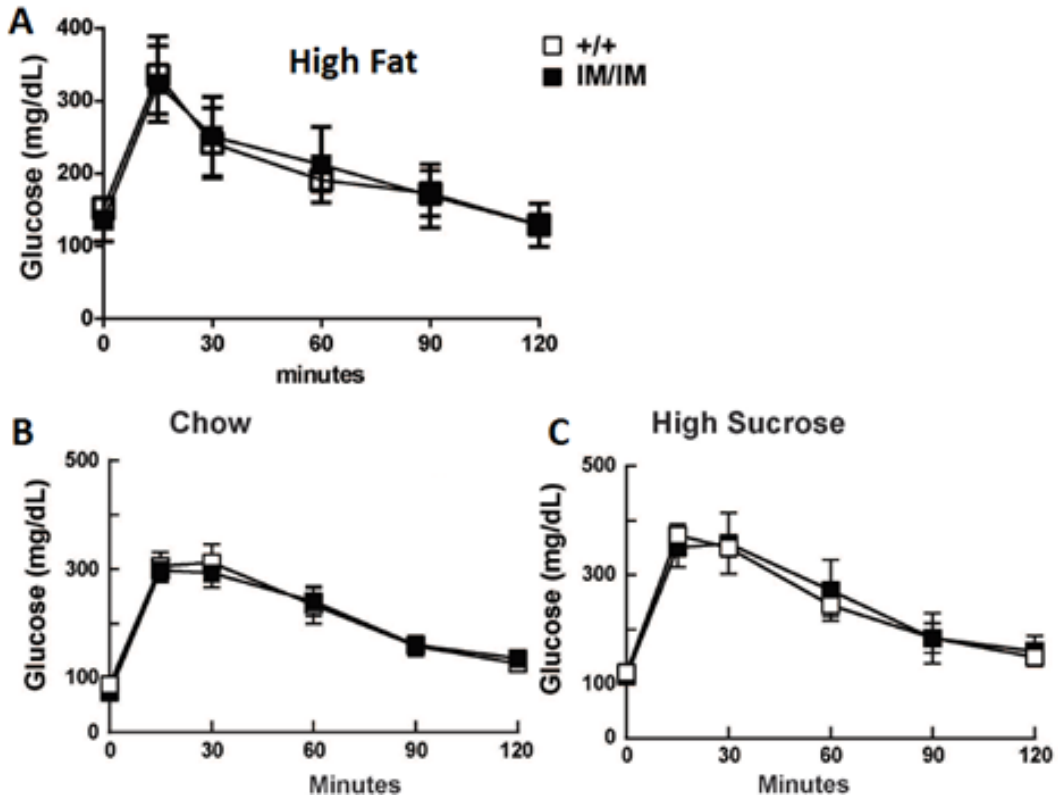


Figure 27. Pnpla3$^{148M/M}$ KI mice and per oral glucose tolerance tests (GGT). *Glucose tolerance tests were performed on Wild-type and Pnpla3$^{148M/M}$ KI female mice fed (A) a high fat diet for 10 weeks mice (n=5/group), (B) a chow diet or (C) a high-sucrose diet for 4 weeks, as described in the Methods. Intraperitoneal glucose (1.5g/kg) was given, and the blood was sampled at the indicated times. The experiment was repeated, and the results were similar.*

To determine if the accumulation of CGI-58 on the lipid droplets of the KI was causally related to the accumulation of TG, we examined the levels of CGI-58 on the lipid droplets of mice expressing mutant human PNPLA3. If lipid accumulation in the mutant KI mice is due to sequestration of CGI-58 by PNPLA3, a similar increase in CGI-58 should be seen in the PNPLA3^{148MTg} mice. In mice fed a high-sucrose diet for 4 weeks, levels of mouse PNPLA3 and CGI-58 mRNA were similar in livers from the transgenic and wild-type animals (Fig. 28A). The level of human PNPLA3 mRNA was ~30% higher in the PNPLA3^{148MTg} mice than in the mice expressing the wild-type human PNPLA3 transgene, although this difference was not statistically significant. The levels of human PNPLA3 were increased dramatically in liver lysates (Fig. 29) and in hepatic lipid droplets isolated from mice expressing the PNPLA3^{I148M} transgene when compared to the mice expressing wild-type PNPLA3 (Fig. 28B). The levels of CGI-58 on hepatic lipid droplets did not differ significantly between the PNPLA3^{-148M} and PNPLA3^{-WT} transgenic mice.

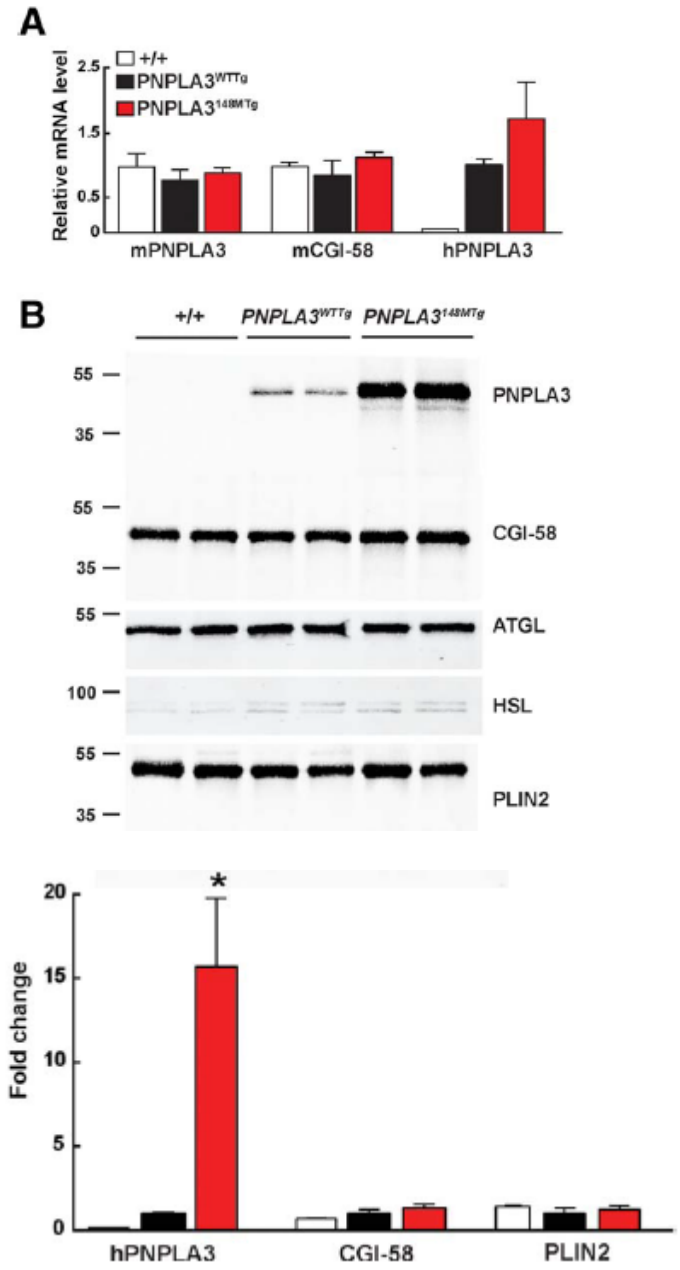


Figure 28. Accumulation of PNPLA3 and CGI-58 protein on hepatic lipid droplets of PNPLA3^{WTTg} and PNPLA3^{148MTg} mice. *Lipid droplets and mRNA were prepared from livers of 11-week-old male mice after 4 weeks on a high-sucrose diet. (A) Relative levels of endogenous and human PNPLA3 and CGI-58 mRNA in wild-type (+/+), PNPLA3^{WTTg}, and PNPLA3^{148MTg} transgenic mice. (B) Immunoblot analysis of lipid droplet proteins from wild-type (+/+), PNPLA3^{WTTg}, and PNPLA3^{148MTg} transgenic mice. *P<0.01. The experiment was repeated twice, and the results were similar.*

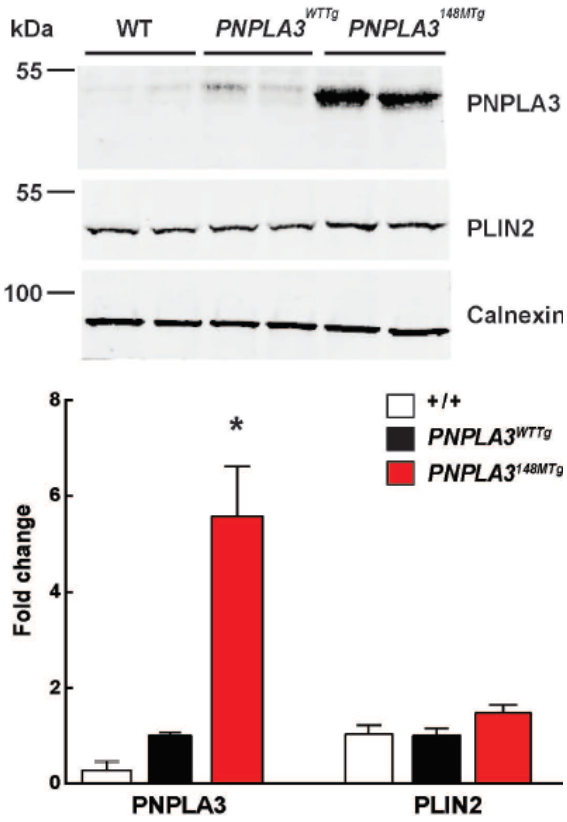


Figure 29. PNPLA3 levels are increased in liver lysates from PNPLA3^{148MTg} mice. *Livers were homogenized in RIPA buffer. A total of 40 µg of protein was size-fractionated on a 10% SDS-polyacrylamide gel and immunoblotted with antibodies to human PNPLA3 and both mouse PLIN2 and calnexin. The same liver samples were used to isolate the lipid droplets utilized in the experiments shown in Fig. 28.*

Thus, the accumulation of CGI-58 is not a prerequisite for the development of hepatic steatosis in association with expression of mutant PNPLA3.

4. Discussion

The studies of our and other laboratories show that genome- and exome-wide association studies can uncover new genes associated with human NAFLD, potentially providing new insights into the pathogenesis of NAFLD.

A major finding in an exome-wide association NAFLD study performed by our laboratory in DHS is the discovery of the association of the TM6SF2 variant Glu167Lys with NAFLD. The TM6SF2 variant encoding p. Glu167Lys was also associated with higher circulating levels of alanine transaminase, a marker of liver injury, and with lower levels of low-density lipoprotein–cholesterol (LDL-C), triglycerides and alkaline phosphatase. This is different from NAFLD associated with the PNPLA3 I148M variant, where almost no extrahepatic phenotypes were observed [28]. Subcellular separation of Hepa1c1c7 cells showed that PNPLA3 and TM6SF2 are localized in two distinct fractions: TM6SF2 in the membrane fraction and PNPLA3 on lipid droplets. It seems unlikely that these proteins interact; rather they cause NAFLD by two independent pathways.

The overexpression of a recombinant TM6SF2 protein in cultured hepatocytes showed that TM6SF2 Glu167Lys variant protein was expressed less than 50% of wild-type TM6SF2. In the same experiment, RNA levels were similar between both forms, indicating that the TM6SF2 E167K protein is less stable *in vivo*. The reason for this instability is unknown. These observations indicate that TM6SF2 Glu167Lys causes a deficiency of the TM6SF2 protein and that this deficiency promotes NAFLD.

Experiments with adeno-associated virus–mediated short hairpin RNA (shRNA AAV) knockdown of Tm6sf2 in mice increased liver triglyceride/cholesterol content and decreased very-low-density lipoprotein (VLDL) secretion with lower plasma TG and cholesterol levels. This observation mimics the findings in human TM6SF2 E167K carriers.

The tissue distribution of human TM6SF2 showed the highest RNA expression in the small intestines, liver and kidneys. Taken together, these human and mice data indicate that TM6SF2 activity is required for normal VLDL secretion and that impaired TM6SF2 function causally contributes to NAFLD.

To elucidate the association of PNPLA3 I148M and TM6SF2 E167K polymorphisms with NAFLD, we used genetically modified mice.

A major finding of our Tm6sf2 knockout (KO) mice studies is that the absence of TM6SF2 causes hepatic steatosis and elevated ALT levels in the absence of a dietary challenge. We provide evidence that TM6SF2 is required for normal lipidation of TG-rich lipoproteins in the liver, and possibly in the intestine, but is not needed for secretion of ApoB-containing lipoproteins. It was shown that TM6SF2, which is predicted to have 8-10 transmembrane helices, is located in both the ER and the Golgi complex, whereas the lipid that accumulates in the livers of mice expressing no TM6SF2 is located in LDs. The large decrease in TG secretion in KO mice once more indicates that TM6SF2 plays a role in VLDL assembly.

There is no evidence for impaired secretion of ApoB-100 and ApoB-48 into the circulation in TM6SF2 KO mice. Thus, TM6SF2 is not required for the initial lipidation of ApoB, for the packaging of VLDL into COPII vesicles to exit the ER, or for secretion of VLDL particles from cells. Rather, the decreased size and TG:APOB ratio in circulating TG-rich lipoproteins of KO mice is consistent with the hypothesis that these mice secrete normal numbers of VLDL particles that are relatively lipid-poor. Taken together, our results indicate that TM6SF2 promotes the addition of neutral lipid to nascent ApoB-containing lipoproteins in hepatocytes, and also possibly in enterocytes. The possible functional roles of TM6SF2 are summarized in the legend to Fig. 30.

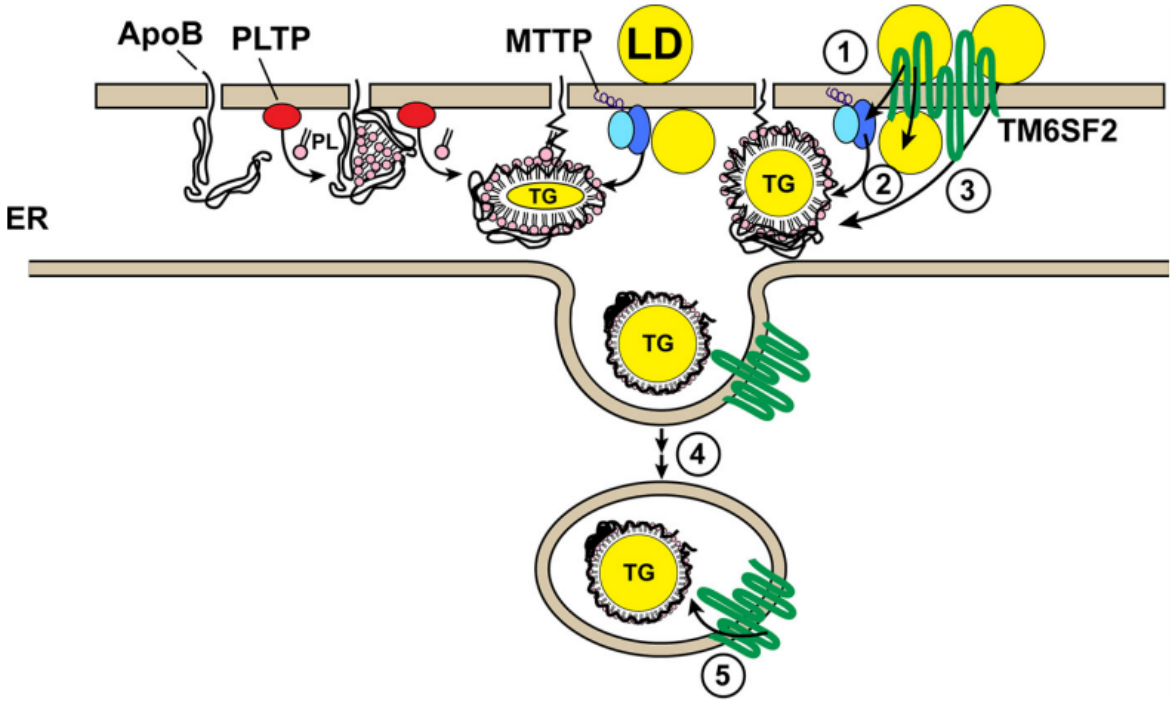


Figure 30. Schematic illustration of the role of the TM6SF2 in VLDL lipidation. *VLDL synthesis is initiated in the ER with the co-translational addition of phospholipids to ApoB. The addition of TG to the particle also begins in the ER, a process that requires microsomal triglyceride transfer protein (MTTP). The partially lipidated VLDL particle is packaged into COPII vesicles and exported to the Golgi where they appear to undergo further "bulk phase" lipidation. Our present data are most consistent with a model in which TM6SF2 promotes this bulk phase lipidation, either by transporting neutral lipids from lipid droplets to the particle by transferring lipids to MTTP ①, to neutral LD in the ER lumen ②, or directly to the nascent VLDL particle ③. Alternatively, TM6SF2 could participate in the transfer of lipid to the particle en route to or within the Golgi complex ④ and ⑤.*

Despite a noticeable reduction in VLDL-TG secretion, mice lacking TM6SF2 had no apparent defect in energy metabolism, or in the metabolism of free fatty acids in the liver. This finding is consistent with our prior observation that body weight and insulin sensitivity are not altered in humans with loss-of-function mutations in TM6SF2 [2, 99]. The observation that hepatic expression of TM6SF2 was not affected by nutritional status, together with the finding that hepatic expression of genes involve in fatty acid metabolism are modestly altered in the KO mice is consistent with the hypothesis that TM6SF2 plays a constitutive role in efflux of hepatic fatty acids in the form of VLDL-TG, rather than in the biosynthesis or oxidation of fatty acids in hepatocytes.

The only gene whose expression was markedly affected by disruption of TM6SF2 was PNPLA3, an LD protein that we previously found to be associated with liver fat content [28]. Deletion of PNPLA3 in mice does not lead to accumulation of liver fat [48, 100]; therefore, the low PNPLA3 transcript levels in TM6SF2 KO mice are likely to be a consequence, rather than a cause, of the increase in hepatic TG in these animals. PNPLA3 is the transcription factor SREBP-1c target gene, and it is possible that the PNPLA3 gene is sensitive to small changes in SREBP-1c activity. Alternatively, PNPLA3 may be subject to regulation by other, as yet unidentified factors that are altered in the Tm6sf2 KO mice. The specific biochemical action of TM6SF2 in VLDL remains to be defined. Mutations in proteins that mediate the initial lipidation [93, 101] and COPII packaging [102] of VLDL markedly reduce secretion of ApoB. Since ApoB secretion was not decreased in Tm6sf2 KO mice, it is unlikely that TM6SF2 is required for any of the early steps in VLDL assembly. VLDL particles are further enriched with TG as they

traverse the ER/Golgi secretory pathway (45). The mechanism by which additional TG molecules are added to particles has not been fully defined. One hypothesis is that the lipid-poor VLDL fuses with LDs in the microsomal lumen [103]. Ultrastructural studies of hepatocytes from liver-specific Mttp KO mice reveal a fundamental role for MTTP in the formation of large LDs in the microsomal lumen [95], but the incorporation of TG from lumenal droplets into nascent VLDL does not require MTTP [104]. It is tempting to speculate that TM6SF2 acts downstream of MTTP to facilitate bulk lipidation of the nascent VLDL particle, by transferring neutral lipids either from cytoplasmic or intraluminal LDs to nascent VLDL particles (Fig. 30). Alternatively, TM6SF2 may have an enzymatic activity. Despite the markedly reduced production of VLDL-TG in the Tm6sf2 KO mice, plasma levels of TG in these animals were similar to those of their WT littermates. This result differs from our previous finding that inactivation of Tm6sf2 in mouse liver using shRNA decreases plasma levels of TG [2]. The reasons for the different effects of TM6SF2 inactivation on plasma TG levels in these two mouse models is not clear at this time. It is possible that the shRNAs that we used to knockdown TM6SF2 have off-target effects in the liver. Alternatively, permanent inactivation of TM6SF2 may result in compensatory changes in TG metabolism that mitigate changes in plasma TG levels. Since VLDL-TG secretion is reduced in the Tm6sf2-/- mice, the normal TG levels in these mice imply a decreased rate of clearance of TG. Fisher et al. [105] showed that large VLDL are much better substrates for LPL, the enzyme that mediates hydrolysis of VLDL-TG. The VLDL particles in the Tm6sf2 KO mice are smaller than those of WT mice and may be hydrolyzed more slowly, and thus remain longer in the circulation, than do VLDL from WT mice. It is also possible that rapid initial hydrolysis of VLDL-TG in the KO mice leads to redistribution of TG to LDL which are then cleared more slowly from the circulation.

We also observed accumulation of neutral lipids in the intestine, the other major site of expression of TM6SF2, and of synthesis/secretion of ApoB-containing lipoproteins. For reasons that are not clear, neutral lipids accumulate in the stroma of villus as well as in the absorptive cells of the small intestine (Fig. 11A). It is interesting that TM6SF2 expression is high in all three regions of the small intestine, in contrast to other proteins involved in lipid absorption and chylomicron synthesis, such as NPC1L1, MTTP and Sar1b [106-108]. TM6SF2 may have a function independent of its role in the formation of ApoB-containing lipoproteins. Additional studies will be required to fully characterize the role of TM6SF2 in the intestine. The most striking difference in the plasma lipid levels of KO and WT mice was in the level of cholesterol

(Fig. 14A). This pattern mirrors the effect of the TM6SF2-E167K variant observed in humans [2]. The changes in plasma cholesterol levels in Tm6sf2-/- mice were not accompanied by any major changes in the hepatic expression of genes encoding proteins involved in cholesterol synthesis (Fig. 10) or of the major receptor that removes LDL particles from the plasma, LDLR (data not shown). The reduction in HDL-C without a concomitant reduction in ApoA1 (data not shown) would be consistent with an increase in activity of SR-BI, which selectively removes cholesterol esters from circulating lipoproteins [109]. However, no differences were detected in SR-BI levels in KO animals when compared to WT littermates (data not shown). Mice lacking MTTP in the liver also have low plasma HDL-C levels, presumably due to reduced availability of VLDL lipids for transfer to HDL. The low HDL-C in Tm6sf2-/- mice may be due to decreased secretion of VLDL lipids in these animals. The reciprocal effects of TM6SF2 inactivation on the accumulation of TG in cytoplasmic LDs and the secretion of VLDL-TG by the liver suggests that the two processes are tightly coupled. This finding has significant implications for current models of LD function. LDs are usually viewed as organelles for localized energy storage. However, under physiological conditions, hepatocytes of WT mice secrete significant amounts of TG (6 mg/h in the present study assuming a plasma volume of 1 mL). Given a hepatic TG content of 10-20 mg, the turnover of TG in cytoplasmic LDs is rapid. Hepatocytes only accumulate TG when the influx of fatty acids exceeds the capacity of cells to package and secrete TG in VLDL. When the influx of fatty acids is reduced, liver TG content decreases rapidly [110]. These findings suggest that in contrast to the large unilocular LDs in adipocytes, which are adapted for long-term, high volume TG storage, the small multilocular LDs in hepatocytes function to rapidly but transiently buffer free fatty acids *en route* to oxidation in mitochondria or secretion in VLDL-TG. Continued loading of hepatocytes with fatty acids, as occurs in obesity, elicits a maladaptive inflammatory response that can progress to fibrosis and ultimately, cirrhosis. The increase in serum ALT levels in Tm6sf2-/- mice suggests that hepatic steatosis in these animals is associated with injury. This finding contrasts with models of PNPLA3-I148M associated steatosis, which has increased levels of hepatic fat but not serum enzymes [47]. The reason for this difference is not clear. The steatosis and transaminitis in Tm6sf2-/- mice were observed on standard chow diets. In contrast, the PNPLA3- I148M knock-in mice did not develop steatosis unless they were placed on a high-sucrose diet, which causes an increase in hepatic fatty acid synthesis and PNPLA3 levels [47].

Distinct extracellular lipid transport systems that utilize lipoproteins evolved more than 900 million years ago. They include the apolipoproteins circulating through the hemolymph of insects, vitellogenins of oviparous animals, and the apoB-lipoproteins secreted by vertebrates [111]. It was hypothesized that the phospholipid transfer activity was the most ancient activity associated with MTP and that the neutral lipid transfer activity was acquired during evolution [111, 112]. Apoprotein B (ApoB) protein evolved only in vertebrates, and it is believed that ApoB and MTTP evolutionary evolve in the same time to transfer neutral lipids [113]. Although a lot in known about the synthesis and secretion of lipoproteins, there are still gaps in the knowledge of this process.

By using software for the prediction of protein domains (such as the NCBI Conserved Domain https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi), there are no known protein domains in the TM6SF2 protein. However, a new protein domain termed 'EXPERA' was described recently. Besides TM6SF2, this domain is also found in its closest homolog TM6SF1 and two other proteins involved in cholesterol metabolism - EBP and MAC30 (TMEM97) [114]. However, the biological function of these proteins that are predicted to have an EXPERA domain may differ greatly. For example, EBP (3-beta-hydroxysteroid-delta-8, delta-7 isomerase) is involved in the biosynthesis of cholesterol. Its deficiency is associated with two human genetic traits: X-linked dominant chondrodysplasia punctata and MEND syndrome; the former is characterized by multiple neurologic and dysmorphic defects [115]. The functions of MAC30 and TM6SF1 are not completely clear, but there is data supporting that MAC30 is involved in the metabolism of cholesterol in lysosomes [116]. TM6SF1 also is localized in lysosomes, but its function has not previously been studied [117]. In contrast, TM6SF2 is mainly localized in the ER and Golgi apparatus, and its function is associated not only with cholesterol but also triglycerides. Thus TM6SF2 appears to be quite different from the rest of EXPERA domain-containing proteins.

It is also unclear why the TM6SF2 protein is expressed in the kidney. Previous studies have shown that mouse renal tubular epithelial cells synthesize MTTP and ApoB proteins. It has been hypothesized that the secretion of renal ApoB-containing lipoproteins protects the kidneys from harmful fat accumulation in tubular epithelial cells [118]. A screening of our Tm6sf2 KO and WT mice (plasma creatinine and renal histology - hematoxylin-eosin staining) showed no significant changes in the kidney of these animals (unpublished data, data not shown).

Even though we found that the TM6SF2 167K protein variant is more unstable than the common 167E variant, it is not exactly known what causes this protein instability. It is also unknown whether the instability is associated only with the rapid disappearance of the protein from the cell or if the E167K variant is also functionally deficient. Mice genetically engineered to express the 167K variant might be used to answer these questions. Similar models were used successfully by us before – for example, Pnpla3 I148M or S47A knockin mice [47]. At least one other group is currently working on TM6SF2 E167K mouse models by using the CRISPR/CAS9 genome editing system [119]. This group recently reported that TM6SF2 KO mice did not develop NAFLD, in contrast to our findings. Potential reasons for these discrepancies include possible side effects of the CRISPR/CAS9 system (*off-target effects*), or the effect of specific diets.

Although the lipid accumulation in the liver is one of the components of the metabolic syndrome in humans, we did not identify any weight gain or insulin resistance in our Tm6sf2 KO mice model. Identical observations were made in the TM6SF2 E167K NAFLD studies in DHS [2], and other, more specialized human studies [99]. The Pnpla3 I148M mutation is also associated with fat accumulation in the liver, unchanged body weight and no insulin resistance. These associations are observed in mice [46, 47], as well as in humans [99]. Thus, NAFLD could be seen as a local metabolic syndrome in the liver, which in contrast to the whole-body metabolic syndrome is not associated with severe insulin resistance.

The major finding in our Pnpla3148M/M KI mice studies is that those mice develop diet-dependent hepatic steatosis. The I148M substitution did not lead to increased hepatic TG levels in chow-fed mice, but elicited a 2 to 3-fold elevation in liver fat content in sucrose-fed animals. This finding is congruent with the observation in humans that the PNPLA3-148M variant is associated with appreciably greater liver fat accumulation in obese than in lean individuals [28]. Moreover, the increase in liver TG of KI mice was comparable in magnitude to that observed in humans homozygous for the PNPLA3-148M variant. Taken together, these data provide the first direct evidence that physiological expression of the PNPLA3-148M variant causes hepatic steatosis in mice. The phenotypic differences among PNPLA3 knockout, transgenic, and KI mice reiterate the limitations of genetically modified animals as models of human pathophysiology. Nonetheless, comparisons among these lines provide new insights into the biochemical sequelae of the I148M substitution and allow significant refinement of current models of how the 148M variant promotes steatosis. The accumulation of hepatic TG in Pnpla3^{148M} mice indicates that the

steatosis observed previously in PNPLA3^{148M} transgenic mice [46] is not an artifact of high-level expression of the human protein. Furthermore, the similar phenotypes of the Pnpla3$^{47A/A}$ and PNPLA3$^{148M/M}$ KI indicate that the I148M substitution is equivalent to a loss of catalytic function, and that the association with steatosis is not a unique property of the 148M isoform. However, catalytic inactivation of PNPLA3 is not equivalent to the loss of protein expression (as illustrated by the lack of steatosis in Pnpla3 knockout mice). Rather, PNPLA3 associated hepatic steatosis in mice requires the presence of the catalytically inactive protein, not simply the absence PNPLA3 activity. Does the 148M variant also confer a gain-of-function? In the gain-of-function model proposed by Zechner and colleagues, the I148M substitution promotes TG synthesis by increasing the LPAAT activity of the enzyme [51]. A limitation of this model is that it implies an enzymatic activity that is independent of the catalytic serine, since Pnpla3^{S47A} mice also express the fatty liver phenotype. Moreover, transgenic mice overexpressing wild-type human PNPLA3 did not have increased hepatic fat content, as might have been expected if the protein has LPAAT activity. Furthermore, rates of phosphatidic acid turnover and TG synthesis in transgenic mice expressing the 148M isoform were similar to those of mice expressing wild-type PNPLA3 [46]. An alternative gain-of-function hypothesis is that PNPLA3-148M promotes the formation of a signaling molecule that stimulates TG synthesis. In mice overexpressing human PNPLA3-148M protein in the liver, mRNA levels of several SREBP-1c target genes involved in fatty acid TG synthesis were significantly increased. However, no such changes were seen in mRNA expression profiles of the 148M KI mice. Thus, the steatosis in the KI mice appears not be due to the production of a signaling molecule that activates a transcriptional program promoting fatty acid and TG synthesis. We also found significant differences in the composition of hepatic TG fatty acids in transgenic mice that could potentially alter SREBP-1c activity [46]. These changes were not observed in KI mice and are therefore unlikely to be causally related to the steatosis associated with the 148M variant. In both humans and mice, the PNPLA3-148M isoform promotes liver TG accumulation under dietary conditions that expose the liver to high levels of insulin: obesity and insulin resistance in humans, and sucrose feeding in mice. Moreover, the effect of the 148M variant in humans is amplified by diets rich in carbohydrates [120]. Sucrose feeding may promote hepatic steatosis by increasing fatty acid and TG synthesis in the liver and also by stimulating PNPLA3 expression. Increasing the availability of fatty acids to the liver is not sufficient to explain the effects of the 148M variant on liver TG content: Pnpla3^{148M} KI mice fed a high-fat diet, which increases fatty acid flux to the liver but upregulates PNPLA3

expression only modestly, did not develop hepatic steatosis. Conversely, high-level expression of PNPLA3-148M increased liver TG content even in chow-fed animals [46]. Thus, the diet-dependence of PNPLA3-148M-associated steatosis may reflect the need for diet-induced stimulation of PNPLA3 expression. Additional studies will be required to fully define the relationship between PNPLA3 genotype, diet composition, and liver fat content in both mice and humans. Kumashiro *et al.* recently reported that antisense oligonucleotide-mediated knockdown of PNPLA3 prevents hepatic steatosis, improves glucose tolerance, and increases hepatic insulin sensitivity in fat-fed rats [121]. However, chronic inactivation (or overexpression) of PNPLA3 in mice was not associated with changes in glucose homeostasis [46]. Nor did we find any association between the PNPLA3-148M isoform and insulin resistance in this study. Our results are similar to those of multiple studies in humans, including three population-based studies, which found no association between the I148M variant and either diabetes or HOMA-IR [28, 34, 122]. Moreover, euglycemic clamp studies failed to identify defects in glucose homeostasis in individuals with the I148M variant [123-125]. A striking and unanticipated consequence of sucrose feeding in the KI mice was a marked accumulation of mutant PNPLA3 protein on lipid droplets. Despite similar levels of hepatic PNPLA3 mRNA, lipid droplets from livers of both 148M and 47A KI animals had 40-fold more PNPLA3 protein than wild-type animals. The accumulation of inactive PNPLA3 on lipid droplets may cause hepatic TG accumulation associated with the I148M substitution. The inactive protein may limit TG hydrolysis, possibly by restricting droplet access to, or by sequestering, a factor required for lipolysis. Recently the specific process on lipid droplet (LD) named “protein crowding” was described, where some of LD proteins have properties to displace other proteins from the surface of LD [126]. Immunoblot analysis indicated that the accumulation of mutant PNPLA3 was not associated with a reduction in ATGL or hormone-sensitive lipase (HSL) on lipid droplets, but CGI- 58 (ABHD5), an essential cofactor of ATGL [127], was 5- fold more abundant on droplets from livers of 148M KI mice than on those of their wild-type littermates. Thus, we speculated that PNPLA3-148M promotes hepatic steatosis by sequestering CGI-58, thereby limiting hydrolysis of lipid droplet TG by ATGL. The possible roles of PNPLA3 I148M are summarized in the legend to Fig. 31.

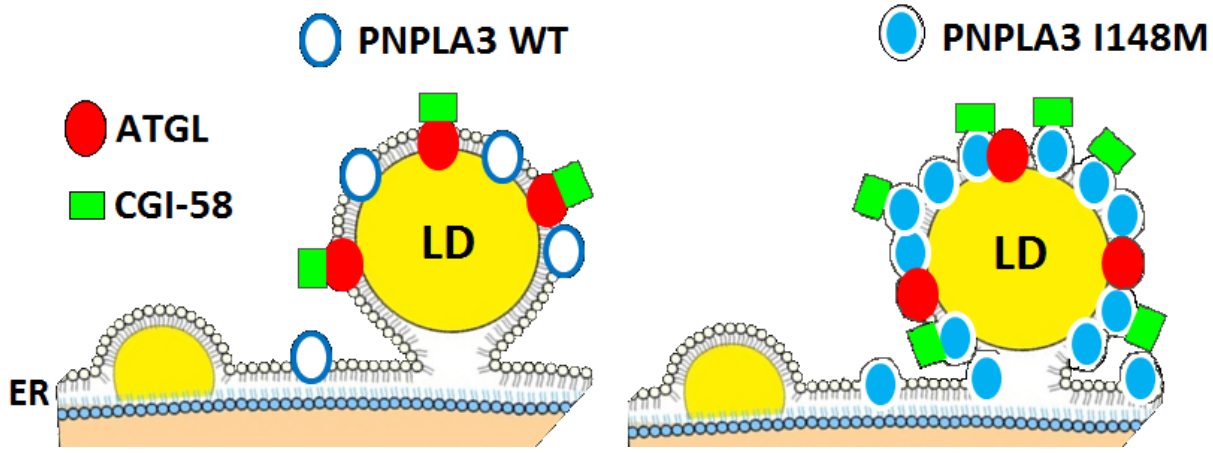


Figure 31. Schematic illustration of the possible PNPLA3 wild type and I148M protein activity on the lipid droplet (LD). *Left – Wild type PNPLA3 digests lipids of the surface of cytosolic lipid droplets next to ATLG (PNPLA2) and its cofactor CGI-58 (ABHD5). Right - PNPLA3 I148M variant accumulates on the surface of LD and displaces ATGL/CGI-58, disrupting normal LD surface protein environment.*

The marginal increase in CGI-58 accumulation on lipid droplets of PNPLA3^{148MTg} mice despite a 2-fold increase in liver fat content and massive accumulation of the mutant PNPLA3 protein argues against this hypothesis. The I148M substitution also resulted in a marked rightward shift in the size distribution of hepatic lipid droplets, such that very large droplets (>100 μm^2) comprised less than 0.1% of the total in wild-type mice, but almost 5% in the Pnpla3148M KI animals. Since the volume of these very large droplets (750 μm^2) is 40 times larger than that of the median-sized droplet (19 μm^2), the major portion of hepatic TG in Pnpla3148M mice is contained in a small fraction of the droplets. This finding raises the possibility that the 148M isoform alters the partitioning of proteins among different-sized lipid droplets, rather than between the lipid droplets and cytosol. Partitioning of proteins among diverse lipid droplets would not be detected by the methods used in this study. Experiments are under way to further characterize the protein and lipid composition of the lipid droplets in which PNPLA3 resides. We cannot exclude the possibility that the accumulation of mutant PNPLA3 on lipid droplets is a consequence, rather than a cause, of hepatic steatosis. Steatosis was not associated with high levels of the 148M isoform in fat-fed mice, but PNPLA3 mRNA was markedly reduced in livers of these animals. In cultured cells, the addition of free fatty acids to the medium induces lipid droplet formation and stabilizes PNPLA3, leading to increased expression of the protein [44]. In mice, expression of the 148M isoform is associated with an

increase in large lipid droplets, which may stabilize the mutant allele. Further studies will be required to elucidate the relationships between PNPLA3 accumulation, lipid droplet size, and hepatic steatosis.

5. Conclusions

- In recent NAFLD GWAS three variants are associated with higher liver fat levels at the exome-wide significance level of 3.6×10^{-7}: two in patatin-like phospholipid domain containing protein 3 (PNPLA3), an established locus for NAFLD, and one (encoding p.Glu167Lys) in TM6SF2 (Transmembrane 6 Superfamily Member 2), a gene of unknown function. The TM6SF2 variant encoding p.Glu167Lys is also associated with higher circulating levels of alanine transaminase, a marker of liver injury, and with lower levels of low-density lipoprotein–cholesterol (LDL-C), triglycerides and alkaline phosphatase in 3 independent populations (n > 80,000).

- The TM6SF2 Glu167Lys variant destabilizes the TM6SF2 protein, leading to lower levels of this protein in the cells. Chronic inactivation of Tm6sf2 in mice is associated with hepatic steatosis, hypocholesterolemia and transaminitis, without any associated changes in glucose homeostasis, thus recapitulating the phenotype observed in humans. TM6SF2 is required to mobilize neutral lipids for VLDL assembly, but is not needed for secretion of ApoB-containing lipoproteins. Despite TM6SF2 being located in the ER and Golgi complex, the lipids that accumulate in its absence reside in lipid droplets.

- A sequence polymorphism (rs738409, I148M) in PNPLA3 is strongly associated with nonalcoholic fatty liver disease (NAFLD. In Pnpla3 I148M knockin mice challenged with a high-sucrose diet, liver fat levels increase 2 to 3-fold compared to wild-type littermates, without any associated changes in glucose homeostasis. The growth in liver fat is accompanied by a 40-fold increase in Pnpla3 protein on hepatic lipid droplets, with no increase in hepatic PNPLA3 messenger RNA (mRNA). This protein accumulation on the surface of lipid droplets could be the cause of human PNPLA3 I148M variant associated NAFLD.

- Both animal models of human NAFLD - Tm6sf2 knockout and Pnpla3 I148M knock in mice could serve as excellent models for the studies of the mechanistic basis of human NAFLD.

6. Main thesis of defense

Thesis I

The human TM6SF2 protein polymorphism (rs58542926; Glu167Lys) is significantly associated with lipid accumulation and inflammation in the human liver as well with lower plasma levels of plasma low-density lipoprotein–cholesterol (LDL-C), triglycerides and alkaline phosphatase.

Thesis II

Human TM6SF2 protein variant 167Lys, compared to the most common TM6SF2 variant (167Glu) is less stable in tissues and cause protein deficiency.

Thesis III

The developed model of Tm6sf2 deficiently in mouse, recapitulate the phenotype of the TM6SF2 E167K variant in humans, with fatty liver, low plasma cholesterol levels and no changes in body weight and insulin resistance. Current data shows that Tm6sf2 protein participates in VLDL generation in the endoplasmic reticulum and Golgi apparatus.

Thesis IV

Established Pnpla3 I148M *knock in* mouse model partially mimics PNPLA3 I148M (rs738409) polymorphism in humans with lipid accumulation in liver and no changes in plasma lipids, body weight or insulin resistance. Pnpla3 I148M variant protein accumulates on hepatic lipid droplet surface; a phenomenon that may be causally important for its NAFLD-promoting effect.

7. Acknowledgements

I would like to thank my postdoctoral fellowship mentors and principal investigators in UTSW at Dallas, Texas - Helen Hobbs, MD and Jonathan Cohen, Ph.D. For the excellent guiding and training in basic science.

I would like to thank Dr. Stefan Stender, MD Ph.D. for critical evaluation of this manuscript.

Also, I would like to thank Christina Zhao, Fang Xu, Liangcai Nie and Stephanie Spaeth for technical assistance and John Shelton for helpful discussions.

Special thanks go to Ojara Veides fellowship program for initial support of my studies aboard.

8. References

1. Argo, C.K. and S.H. Caldwell, *Epidemiology and natural history of non-alcoholic steatohepatitis.* Clin Liver Dis, 2009. **13**(4): p. 511-31.
2. Kozlitina, J., et al., *Exome-wide association study identifies a TM6SF2 variant that confers susceptibility to nonalcoholic fatty liver disease.* Nat Genet, 2014. **46**(4): p. 352-6.
3. Smagris, E., et al., *Inactivation of Tm6sf2, a Gene Defective in Fatty Liver Disease, Impairs Lipidation but Not Secretion of Very Low Density Lipoproteins.* J Biol Chem, 2016. **291**(20): p. 10659-76.
4. Mahdessian, H., et al., *TM6SF2 is a regulator of liver fat metabolism influencing triglyceride secretion and hepatic lipid droplet content.* Proc Natl Acad Sci U S A, 2014. **111**(24): p. 8913-8.
5. Williams, R., et al., *Implementation of the Lancet Standing Commission on Liver Disease in the UK.* Lancet, 2015. **386**(10008): p. 2098-111.
6. Cohen, J.C., J.D. Horton, and H.H. Hobbs, *Human fatty liver disease: old questions and new insights.* Science, 2011. **332**(6037): p. 1519-23.
7. Temple, J.L., et al., *A Guide to Non-Alcoholic Fatty Liver Disease in Childhood and Adolescence.* Int J Mol Sci, 2016. **17**(6).
8. Li, Y., et al., *Genetic variant in PNPLA3 is associated with nonalcoholic fatty liver disease in China.* Hepatology, 2012. **55**(1): p. 327-8.
9. McCullough, A.J., *The clinical features, diagnosis and natural history of nonalcoholic fatty liver disease.* Clin Liver Dis, 2004. **8**(3): p. 521-33, viii.
10. Popkin, B.M. and C.M. Doak, *The obesity epidemic is a worldwide phenomenon.* Nutr Rev, 1998. **56**(4 Pt 1): p. 106-14.
11. Slyper, A.H., *The pediatric obesity epidemic: causes and controversies.* J Clin Endocrinol Metab, 2004. **89**(6): p. 2540-7.
12. Schwimmer, J.B., et al., *Heritability of nonalcoholic fatty liver disease.* Gastroenterology, 2009. **136**(5): p. 1585-92.
13. Browning, J.D., et al., *Prevalence of hepatic steatosis in an urban population in the United States: impact of ethnicity.* Hepatology, 2004. **40**(6): p. 1387-95.
14. Ludwig, J., et al., *Nonalcoholic steatohepatitis: Mayo Clinic experiences with a hitherto unnamed disease.* Mayo Clin Proc, 1980. **55**(7): p. 434-8.
15. Hubscher, S.G., *Histological assessment of non-alcoholic fatty liver disease.* Histopathology, 2006. **49**(5): p. 450-65.
16. Tandra, S., et al., *Presence and significance of microvesicular steatosis in nonalcoholic fatty liver disease.* J Hepatol, 2011. **55**(3): p. 654-9.
17. Ratziu, V., et al., *Survival, liver failure, and hepatocellular carcinoma in obesity-related cryptogenic cirrhosis.* Hepatology, 2002. **35**(6): p. 1485-93.
18. Mann, J.P., R.K. Semple, and M.J. Armstrong, *How Useful Are Monogenic Rodent Models for the Study of Human Non-Alcoholic Fatty Liver Disease?* Front Endocrinol (Lausanne), 2016. **7**: p. 145.
19. Rong, X., et al., *Lpcat3-dependent production of arachidonoyl phospholipids is a key determinant of triglyceride secretion.* Elife, 2015. **4**.
20. Quigley, E.M. and H.P. Monsour, *The Gut Microbiota and Nonalcoholic Fatty Liver Disease.* Semin Liver Dis, 2015. **35**(3): p. 262-9.
21. Mondul, A., et al., *PNPLA3 I148M Variant Influences Circulating Retinol in Adults with Nonalcoholic Fatty Liver Disease or Obesity.* J Nutr, 2015. **145**(8): p. 1687-91.
22. Miele, L., et al., *The Kruppel-like factor 6 genotype is associated with fibrosis in nonalcoholic fatty liver disease.* Gastroenterology, 2008. **135**(1): p. 282-291 e1.

23. Al-Serri, A., et al., *The SOD2 C47T polymorphism influences NAFLD fibrosis severity: evidence from case-control and intra-familial allele association studies.* J Hepatol, 2012. **56**(2): p. 448-54.
24. Pruis, M.G., et al., *Maternal western diet primes non-alcoholic fatty liver disease in adult mouse offspring.* Acta Physiol (Oxf), 2014. **210**(1): p. 215-27.
25. 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, 2005. **288**(2): p. E462-8.
26. Portillo-Sanchez, P., et al., *High Prevalence of Nonalcoholic Fatty Liver Disease in Patients With Type 2 Diabetes Mellitus and Normal Plasma Aminotransferase Levels.* J Clin Endocrinol Metab, 2015. **100**(6): p. 2231-8.
27. Kinner, S., S.B. Reeder, and T. Yokoo, *Quantitative Imaging Biomarkers of NAFLD.* Dig Dis Sci, 2016. **61**(5): p. 1337-47.
28. Romeo, S., et al., *Genetic variation in PNPLA3 confers susceptibility to nonalcoholic fatty liver disease.* Nat Genet, 2008. **40**(12): p. 1461-5.
29. Llaurado, G., et al., *Liver fat content and hepatic insulin sensitivity in overweight patients with type 1 diabetes.* J Clin Endocrinol Metab, 2015. **100**(2): p. 607-16.
30. McCullough, A.J., *Pathophysiology of nonalcoholic steatohepatitis.* J Clin Gastroenterol, 2006. **40 Suppl 1**: p. S17-29.
31. Day, C.P., *Natural history of NAFLD: remarkably benign in the absence of cirrhosis.* Gastroenterology, 2005. **129**(1): p. 375-8.
32. Bellentani, S. and M. Marino, *Epidemiology and natural history of non-alcoholic fatty liver disease (NAFLD).* Ann Hepatol, 2009. **8 Suppl 1**: p. S4-8.
33. Anderson, M.L., *Treating NASH.* J Gastroenterol Hepatol, 2006. **21**(1 Pt 1): p. 14.
34. Speliotes, E.K., et al., *Genome-wide association analysis identifies variants associated with nonalcoholic fatty liver disease that have distinct effects on metabolic traits.* PLoS Genet, 2011. **7**(3): p. e1001324.
35. Wagenknecht, L.E., et al., *Correlates and heritability of nonalcoholic fatty liver disease in a minority cohort.* Obesity (Silver Spring), 2009. **17**(6): p. 1240-6.
36. Stender, S., et al., *Adult-onset liver disease and hepatocellular carcinoma in S-adenosylhomocysteine hydrolase deficiency.* Mol Genet Metab, 2015. **116**(4): p. 269-74.
37. Yuan, X., et al., *Population-based genome-wide association studies reveal six loci influencing plasma levels of liver enzymes.* Am J Hum Genet, 2008. **83**(4): p. 520-8.
38. Sookoian, S., et al., *A nonsynonymous gene variant in the adiponutrin gene is associated with nonalcoholic fatty liver disease severity.* J Lipid Res, 2009. **50**(10): p. 2111-6.
39. Nischalke, H.D., et al., *The PNPLA3 rs738409 148M/M genotype is a risk factor for liver cancer in alcoholic cirrhosis but shows no or weak association in hepatitis C cirrhosis.* PLoS One, 2011. **6**(11): p. e27087.
40. Liu, Y.L., et al., *Carriage of the PNPLA3 rs738409 C >G polymorphism confers an increased risk of non-alcoholic fatty liver disease associated hepatocellular carcinoma.* J Hepatol, 2014. **61**(1): p. 75-81.
41. Tian, C., et al., *Variant in PNPLA3 is associated with alcoholic liver disease.* Nat Genet, 2010. **42**(1): p. 21-3.
42. Muller, T., et al., *Distinct, alcohol-modulated effects of PNPLA3 genotype on progression of chronic hepatitis C.* J Hepatol, 2011. **55**(3): p. 732-3.
43. Trepo, E., et al., *Common polymorphism in the PNPLA3/adiponutrin gene confers higher risk of cirrhosis and liver damage in alcoholic liver disease.* J Hepatol, 2011. **55**(4): p. 906-12.
44. Huang, Y., et al., *A feed-forward loop amplifies nutritional regulation of PNPLA3.* Proc Natl Acad Sci U S A, 2010. **107**(17): p. 7892-7.

45. Huang, Y., J.C. Cohen, and H.H. Hobbs, *Expression and characterization of a PNPLA3 protein isoform (I148M) associated with nonalcoholic fatty liver disease.* J Biol Chem, 2011. **286**(43): p. 37085-93.
46. Li, J.Z., et al., *Chronic overexpression of PNPLA3I148M in mouse liver causes hepatic steatosis.* J Clin Invest, 2012. **122**(11): p. 4130-44.
47. Smagris, E., et al., *Pnpla3I148M knockin mice accumulate PNPLA3 on lipid droplets and develop hepatic steatosis.* Hepatology, 2015. **61**(1): p. 108-18.
48. Basantani, M.K., et al., *Pnpla3/Adiponutrin deficiency in mice does not contribute to fatty liver disease or metabolic syndrome.* J Lipid Res, 2011. **52**(2): p. 318-29.
49. Pingitore, P., et al., *PNPLA3 overexpression results in reduction of proteins predisposing to fibrosis.* Hum Mol Genet, 2016. **25**(23): p.5212-5222
50. Bruschi, F.V., et al., *The PNPLA3 I148M variant modulates the fibrogenic phenotype of human hepatic stellate cells.* Hepatology, 2017. **65**(6): p.1875-1890
51. Kumari, M., et al., *Adiponutrin functions as a nutritionally regulated lysophosphatidic acid acyltransferase.* Cell Metab, 2012. **15**(5): p. 691-702.
52. Hernaez, R., et al., *Association between variants in or near PNPLA3, GCKR, and PPP1R3B with ultrasound-defined steatosis based on data from the third National Health and Nutrition Examination Survey.* Clin Gastroenterol Hepatol, 2013. **11**(9): p. 1183-1190 e2.
53. Palmer, C.N., et al., *Paradoxical lower serum triglyceride levels and higher type 2 diabetes mellitus susceptibility in obese individuals with the PNPLA3 148M variant.* PLoS One, 2012. **7**(6): p. e39362.
54. Kahali, B., B. Halligan, and E.K. Speliotes, *Insights from Genome-Wide Association Analyses of Nonalcoholic Fatty Liver Disease.* Semin Liver Dis, 2015. **35**(4): p. 375-91.
55. Buch, S., et al., *A genome-wide association study confirms PNPLA3 and identifies TM6SF2 and MBOAT7 as risk loci for alcohol-related cirrhosis.* Nat Genet, 2015. **47**(12): p. 1443-8.
56. Mancina, R.M., et al., *The MBOAT7-TMC4 Variant rs641738 Increases Risk of Nonalcoholic Fatty Liver Disease in Individuals of European Descent.* Gastroenterology, 2016. **150**(5): p. 1219-1230 e6.
57. Viitasalo, A., et al., *Association of MBOAT7 gene variant with plasma ALT levels in children: the PANIC study.* Pediatr Res, 2016. **80**(5): p.651-655
58. Luukkonen, P.K., et al., *The MBOAT7 variant rs641738 alters hepatic phosphatidylinositols and increases severity of non-alcoholic fatty liver disease in humans.* J Hepatol, 2016. **65**(6):1263-1265
59. Hui, S.T., et al., *The genetic architecture of NAFLD among inbred strains of mice.* Elife, 2015. **4**: p. e05607.
60. Kim, S.H., et al., *A post-developmental genetic screen for zebrafish models of inherited liver disease.* PLoS One, 2015. **10**(5): p. e0125980.
61. Anstee, Q.M. and R.D. Goldin, *Mouse models in non-alcoholic fatty liver disease and steatohepatitis research.* Int J Exp Pathol, 2006. **87**(1): p. 1-16.
62. Kimura, Y. and M. Sumiyoshi, *High-fat, high-sucrose, and high-cholesterol diets accelerate tumor growth and metastasis in tumor-bearing mice.* Nutr Cancer, 2007. **59**(2): p. 207-16.
63. Rinella, M.E. and R.M. Green, *The methionine-choline deficient dietary model of steatohepatitis does not exhibit insulin resistance.* J Hepatol, 2004. **40**(1): p. 47-51.
64. Chiappini, F., et al., *Hepatic and serum lipid signatures specific to nonalcoholic steatohepatitis in murine models.* Sci Rep, 2016. **6**: p. 31587.
65. Luo, X.Y., et al., *IFN-gamma deficiency attenuates hepatic inflammation and fibrosis in a steatohepatitis model induced by a methionine- and choline-deficient high-fat diet.* Am J Physiol Gastrointest Liver Physiol, 2013. **305**(12): p. G891-9.
66. Paz-Filho, G., et al., *Molecular pathways involved in the improvement of non-alcoholic fatty liver disease.* J Mol Endocrinol, 2013. **51**(1): p. 167-79.

67. Victor, R.G., et al., *The Dallas Heart Study: a population-based probability sample for the multidisciplinary study of ethnic differences in cardiovascular health.* Am J Cardiol, 2004. **93**(12): p. 1473-80.
68. Teslovich, T.M., et al., *Biological, clinical and population relevance of 95 loci for blood lipids.* Nature, 2010. **466**(7307): p. 707-13.
69. Friedewald, W.T., R.I. Levy, and D.S. Fredrickson, *Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge.* Clin Chem, 1972. **18**(6): p. 499-502.
70. Wagner, G.P., K. Kin, and V.J. Lynch, *Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples.* Theory Biosci, 2012. **131**(4): p. 281-5.
71. Lattin, J.E., et al., *Expression analysis of G Protein-Coupled Receptors in mouse macrophages.* Immunome Res, 2008. **4**: p. 5.
72. Valenzuela, D.M., et al., *High-throughput engineering of the mouse genome coupled with high-resolution expression analysis.* Nat Biotechnol, 2003. **21**(6): p. 652-9.
73. Folch, J., M. Lees, and G.H. Sloane Stanley, *A simple method for the isolation and purification of total lipides from animal tissues.* J Biol Chem, 1957. **226**(1): p. 497-509.
74. Wang, J., et al., *Relative roles of ABCG5/ABCG8 in liver and intestine.* J Lipid Res, 2015. **56**(2): p. 319-30.
75. Engelking, L.J., et al., *Overexpression of Insig-1 in the livers of transgenic mice inhibits SREBP processing and reduces insulin-stimulated lipogenesis.* J Clin Invest, 2004. **113**(8): p. 1168-75.
76. Wilhelm, B.T. and J.R. Landry, *RNA-Seq-quantitative measurement of expression through massively parallel RNA-sequencing.* Methods, 2009. **48**(3): p. 249-57.
77. Jones, C., et al., *Disruption of LDL but not VLDL clearance in autosomal recessive hypercholesterolemia.* J Clin Invest, 2007. **117**(1): p. 165-74.
78. Wang, Y., et al., *Hepatic ANGPTL3 regulates adipose tissue energy homeostasis.* Proc Natl Acad Sci U S A, 2015. **112**(37): p. 11630-5.
79. Koontz, L., *TCA precipitation.* Methods Enzymol, 2014. **541**: p. 3-10.
80. Vitale, N., et al., *Localization of ADP-ribosylation factor domain protein 1 (ARD1) in lysosomes and Golgi apparatus.* Proc Natl Acad Sci U S A, 1998. **95**(15): p. 8613-8.
81. Ding, Y., et al., *Isolating lipid droplets from multiple species.* Nat Protoc, 2013. **8**(1): p. 43-51.
82. Brasaemle, D.L. and N.E. Wolins, *Isolation of lipid droplets from cells by density gradient centrifugation.* Curr Protoc Cell Biol, 2006. **Chapter 3**: p. Unit 3 15.
83. Chen, Z., et al., *AAV8-mediated long-term expression of human LCAT significantly improves lipid profiles in hCETP;Ldlr(+/-) mice.* J Cardiovasc Transl Res, 2011. **4**(6): p. 801-10.
84. Chambers, J.C., et al., *Genome-wide association study identifies loci influencing concentrations of liver enzymes in plasma.* Nat Genet, 2011. **43**(11): p. 1131-8.
85. Gorden, A., et al., *Genetic variation at NCAN locus is associated with inflammation and fibrosis in non-alcoholic fatty liver disease in morbid obesity.* Hum Hered, 2013. **75**(1): p. 34-43.
86. Krogh, A., et al., *Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes.* J Mol Biol, 2001. **305**(3): p. 567-80.
87. Marchler-Bauer, A., et al., *CDD: a Conserved Domain Database for the functional annotation of proteins.* Nucleic Acids Res, 2011. **39**(Database issue): p. D225-9.
88. Hay, J.C., et al., *Localization, dynamics, and protein interactions reveal distinct roles for ER and Golgi SNAREs.* J Cell Biol, 1998. **141**(7): p. 1489-502.
89. Herz, J. and D.K. Strickland, *LRP: a multifunctional scavenger and signaling receptor.* J Clin Invest, 2001. **108**(6): p. 779-84.
90. Phillips, M.C., *New insights into the determination of HDL structure by apolipoproteins: Thematic review series: high density lipoprotein structure, function, and metabolism.* J Lipid Res, 2013. **54**(8): p. 2034-48.

91. Horton, J.D., J.C. Cohen, and H.H. Hobbs, *PCSK9: a convertase that coordinates LDL catabolism.* J Lipid Res, 2009. **50 Suppl**: p. S172-7.
92. Jones, B., et al., *Mutations in a Sar1 GTPase of COPII vesicles are associated with lipid absorption disorders.* Nat Genet, 2003. **34**(1): p. 29-31.
93. Wetterau, J.R., et al., *Absence of microsomal triglyceride transfer protein in individuals with abetalipoproteinemia.* Science, 1992. **258**(5084): p. 999-1001.
94. Okazaki, H., et al., *LXR-SREBP-1c-phospholipid transfer protein axis controls very low density lipoprotein (VLDL) particle size.* J Biol Chem, 2010. **285**(9): p. 6801-10.
95. Raabe, M., et al., *Analysis of the role of microsomal triglyceride transfer protein in the liver of tissue-specific knockout mice.* J Clin Invest, 1999. **103**(9): p. 1287-98.
96. Hashidate-Yoshida, T., et al., *Fatty acid remodeling by LPCAT3 enriches arachidonate in phospholipid membranes and regulates triglyceride transport.* Elife, 2015. **4**.
97. Li, Z., et al., *Deficiency in lysophosphatidylcholine acyltransferase 3 reduces plasma levels of lipids by reducing lipid absorption in mice.* Gastroenterology, 2015. **149**(6): p. 1519-29.
98. Horton, J.D., J.L. Goldstein, and M.S. Brown, *SREBPs: transcriptional mediators of lipid homeostasis.* Cold Spring Harb Symp Quant Biol, 2002. **67**: p. 491-8.
99. Lallukka, S. and H. Yki-Jarvinen, *Non-alcoholic fatty liver disease and risk of type 2 diabetes.* Best Pract Res Clin Endocrinol Metab, 2016. **30**(3): p. 385-95.
100. Chen, W., et al., *Patatin-like phospholipase domain-containing 3/adiponutrin deficiency in mice is not associated with fatty liver disease.* Hepatology, 2010. **52**(3): p. 1134-42.
101. Manchekar, M., et al., *Phospholipid transfer protein plays a major role in the initiation of apolipoprotein B-containing lipoprotein assembly in mouse primary hepatocytes.* J Biol Chem, 2015. **290**(13): p. 8196-205.
102. Ouguerram, K., et al., *Low rate of production of apolipoproteins B100 and AI in 2 patients with Anderson disease (chylomicron retention disease).* Arterioscler Thromb Vasc Biol, 2012. **32**(6): p. 1520-5.
103. Yao, Z., et al., *Microsome-associated lumenal lipid droplets in the regulation of lipoprotein secretion.* Curr Opin Lipidol, 2013. **24**(2): p. 160-70.
104. Kulinski, A., S. Rustaeus, and J.E. Vance, *Microsomal triacylglycerol transfer protein is required for lumenal accretion of triacylglycerol not associated with ApoB, as well as for ApoB lipidation.* J Biol Chem, 2002. **277**(35): p. 31516-25.
105. Fisher, R.M., et al., *Human triacylglycerol-rich lipoprotein subfractions as substrates for lipoprotein lipase.* Clin Chim Acta, 1995. **236**(1): p. 7-17.
106. Engelking, L.J., et al., *Blockade of cholesterol absorption by ezetimibe reveals a complex homeostatic network in enterocytes.* J Lipid Res, 2012. **53**(7): p. 1359-68.
107. Lin, M.C., et al., *Cloning and regulation of hamster microsomal triglyceride transfer protein. The regulation is independent from that of other hepatic and intestinal proteins which participate in the transport of fatty acids and triglycerides.* J Biol Chem, 1994. **269**(46): p. 29138-45.
108. Davis, H.R., Jr., et al., *Niemann-Pick C1 Like 1 (NPC1L1) is the intestinal phytosterol and cholesterol transporter and a key modulator of whole-body cholesterol homeostasis.* J Biol Chem, 2004. **279**(32): p. 33586-92.
109. Acton, S.L., et al., *Expression cloning of SR-BI, a CD36-related class B scavenger receptor.* J Biol Chem, 1994. **269**(33): p. 21003-9.
110. Guan, H.P., et al., *Accelerated fatty acid oxidation in muscle averts fasting-induced hepatic steatosis in SJL/J mice.* J Biol Chem, 2009. **284**(36): p. 24644-52.
111. Rava, P. and M.M. Hussain, *Acquisition of triacylglycerol transfer activity by microsomal triglyceride transfer protein during evolution.* Biochemistry, 2007. **46**(43): p. 12263-74.
112. Davis, R.A., *Evolution of processes and regulators of lipoprotein synthesis: from birds to mammals.* J Nutr, 1997. **127**(5 Suppl): p. 795S-800S.

113. Smolenaars, M.M., et al., *Molecular diversity and evolution of the large lipid transfer protein superfamily.* J Lipid Res, 2007. **48**(3): p. 489-502.
114. Sanchez-Pulido, L. and C.P. Ponting, *TM6SF2 and MAC30, new enzyme homologs in sterol metabolism and common metabolic disease.* Front Genet, 2014. **5**: p. 439.
115. Derry, J.M., et al., *Mutations in a delta 8-delta 7 sterol isomerase in the tattered mouse and X-linked dominant chondrodysplasia punctata. jderry@immunex.com.* Nat Genet, 1999. **22**(3): p. 286-90.
116. Ebrahimi-Fakhari, D., et al., *Reduction of TMEM97 increases NPC1 protein levels and restores cholesterol trafficking in Niemann-pick type C1 disease cells.* Hum Mol Genet, 2016. **25**(16): p.3588-3599
117. Tam, W.Y., L. Jiang, and K.M. Kwan, *Transmembrane 6 superfamily 1 (Tm6sf1) is a novel lysosomal transmembrane protein.* Protoplasma, 2015. **252**(4): p. 977-83.
118. Krzystanek, M., et al., *Expression of apolipoprotein B in the kidney attenuates renal lipid accumulation.* J Biol Chem, 2010. **285**(14): p. 10583-90.
119. Fan, Y., et al., *Hepatic Transmembrane 6 Superfamily Member 2 Regulates Cholesterol Metabolism in Mice.* Gastroenterology, 2016. **150**(5): p. 1208-18.
120. Davis, J.N., et al., *Increased hepatic fat in overweight Hispanic youth influenced by interaction between genetic variation in PNPLA3 and high dietary carbohydrate and sugar consumption.* Am J Clin Nutr, 2010. **92**(6): p. 1522-7.
121. Kumashiro, N., et al., *Role of patatin-like phospholipase domain-containing 3 on lipid-induced hepatic steatosis and insulin resistance in rats.* Hepatology, 2013. **57**(5): p. 1763-72.
122. Speliotes, E.K., et al., *PNPLA3 variants specifically confer increased risk for histologic nonalcoholic fatty liver disease but not metabolic disease.* Hepatology, 2010. **52**(3): p. 904-12.
123. Kantartzis, K., et al., *Dissociation between fatty liver and insulin resistance in humans carrying a variant of the patatin-like phospholipase 3 gene.* Diabetes, 2009. **58**(11): p. 2616-23.
124. Santoro, N., et al., *A common variant in the patatin-like phospholipase 3 gene (PNPLA3) is associated with fatty liver disease in obese children and adolescents.* Hepatology, 2010. **52**(4): p. 1281-90.
125. Kotronen, A., et al., *A common variant in PNPLA3, which encodes adiponutrin, is associated with liver fat content in humans.* Diabetologia, 2009. **52**(6): p. 1056-60.
126. Kory, N., et al., *Protein Crowding Is a Determinant of Lipid Droplet Protein Composition.* Dev Cell, 2015. **34**(3): p. 351-63.
127. Lass, A., et al., *Adipose triglyceride lipase-mediated lipolysis of cellular fat stores is activated by CGI-58 and defective in Chanarin-Dorfman Syndrome.* Cell Metab, 2006. **3**(5): p. 309-19.