# INDUCED PLURIPOTENT STEM CELL DERIVED LIVER MODEL FOR THE STUDY OF PNPLA3-ASSOCIATED NON-ALCOHOLIC FATTY LIVER DISEASE

Samantha Grace Tilson
Murray Edwards College

Cellular Genetics
Wellcome Sanger Institute
University of Cambridge

This thesis is submitted for the degree of Doctor of Philosophy
August 2020

DECLARATION

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the Preface and specified in the text. It is not substantially the same as any that I have submitted, or, is being concurrently submitted for a degree or diploma or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text. I further state that no substantial part of my thesis has already been submitted, or, is being concurrently submitted for any such degree, diploma, or other qualification at the University of Cambridge or any other University or similar institution except as declared in the Preface and specified in the text.

In accordance with the Degree Committee for the Faculty of Biology guidelines, this thesis is does not exceed 60,000 words in length (excluding figures, photographs, tables, appendices, and bibliography).

Signed:_____

Date:_____

Samantha G. Tilson

August 31, 2020

Cambridge

SUMMARY

Non-alcoholic fatty liver disease (NAFLD) is now the leading cause of chronic liver disease in the developed world, afflicting approximately one in four adults globally. NAFLD is defined by the accumulation of fat within the liver which ranges in severity from simple steatosis to non-alcoholic steatohepatitis (NASH), cirrhosis, and hepatocellular carcinoma. Until recently, NAFLD has been considered to be a consequence of metabolic syndrome; however, recent studies suggest that genetic factors may also influence disease onset and progression. Accordingly, genome wide association studies have linked the I148M variant in the gene coding for Patatin-like phospholipase domain-containing protein 3 (PNPLA3) with NAFLD aggravation without underlying metabolic disease. However, despite over a decade of research, the function of PNPLA3 and its role in the pathogenesis of NAFLD remains largely obscure. The lack of clarity regarding the function and disease associations of PNPLA3 is due in large part to the lack of a comprehensive human model of PNPLA3-associated NAFLD. In order to address this need, we have developed an *in vitro* model that takes advantage of the unique properties of human induced pluripotent stem cells (hiPSC) and the CRISPR/CAS9 gene editing technology.

We first used CRISPR/CAS9 to generate hiPSC lines with either a knock-out (PNPLA3^{KO}) of the *PNPLA3* gene or with the I148M variant knocked in (PNPLA3^{I148M}). The resulting cells were then differentiated into hepatocytes and grown in 3D culture conditions to improve their maturity and functionality. The differentiated cells were treated with either monounsaturated or saturated free fatty acids to induce NAFLD-like phenotypes (lipid accumulation and lipid toxicity, respectively) and characterized by various functional, genomic, and lipidomic assays. The genetically edited sublines showed similar differentiation efficiency toward hepatocytes as untargeted cells indicating that changes in PNPLA3 activity does not affect hepatic development. Following treatment with free fatty acids, PNPLA3^{KO} cells showed higher lipid accumulation than untargeted cells as well as an altered pattern of response to lipid-induced stress. Indeed, PNPLA3^{KO} cells were resistant to saturated fatty acid-induced lipotoxicity. Furthermore, lipidomic analyses suggest that the PNPLA3 edited cells may be avoiding cell death by shuttling the saturated fatty acids into triglycerides rather than other metabolic pathways. These findings were initially incongruent with the human

disease; however, despite their resistance to lipid induced stress, the PNPLA3^{KO} cells downregulated all phases of drug metabolizing enzymes which made them more susceptible to other forms of hepatotoxicity such as ethanol. The PNPLA3^{I148M} cells exhibited an intermediate phenotype between untargeted and PNPLA3^{KO} cells.

These results demonstrate for the first time in a fully human model that the I148M variant in PNPLA3 is a loss of function variant. This loss of PNPLA3 activity leads to the global downregulation of metabolic pathways likely due to the sequestration of fatty acids in triglycerides caused by reduced lipid droplet remodelling capacity in these hepatocytes. Our results indicate that patients carrying the I148M variant have lower hepatic metabolic activity which causes steatosis, reduced susceptibility to lipotoxicity, and increased susceptibility to other forms of hepatotoxicity which may contribute to NAFLD progression. This novel system provides the first opportunity to study the role of PNPLA3 in the development and progression of human NAFLD *in vitro*.

Acknowledgements

Winston Churchill once said, "Success is stumbling from failure to failure with no loss of enthusiasm." I feel this quote perfectly summarizes the experience of obtaining a PhD. Most days, science is choosing to move forward in the face of immanent failure. I don't think that I can claim that I never lost any enthusiasm throughout this process; however, this thesis is a testament that even though my PhD journey was, by no means, an unadulterated success, I did manage to fail in the right direction. There are so many people who have contributed to my journey in large and small ways. This would never have been possible without each and every one of you.

First, I would like to thank my supervisors Dr. T. Jake Liang and Professor Ludovic Vallier. Both of you took a chance on me, a naïve engineer straight out of undergrad with little experience in the field of biology. You provided me with the space to grow and learn. You listened to my ideas, even when they were a bit far-fetched, you allowed me to make mistakes, but you were always there to offer constructive feedback and help guide me in the right direction. I thank you both for your flexibility as I followed the science down a rabbit hole and my project morphed from our original intentions.

Carola, thank you for teaching me everything I know about growing stem cells, microscopy, and liver disease. You were a constant fountain of knowledge and a great sounding board for trouble shooting and experimental design. Even though we had some struggles and disagreements, I thank you for taking time to train me and push me to my fullest potential. But most of all, thank you for your support both in and out of the lab. Anna, thank you for taking me under your wing and being the first to make me feel welcomed in a foreign land. Thank you for trying to teach me Italian and laughing with me about the state of American politics. To the Kneating Club and the rest of the LRM, thank you for the support, and the laughs, and everything each of you did to make my time on the island a bit more bearable.

To Amarpreet, my first friend in Cambridge, I would not have survived without your love and support. I didn't know when I walked into Clover House on my first day in the UK that the fast-talking Mancunian who I could barely understand would turn into a life-long friend. You were always there for me when I needed to rant or cry or laugh. You bought

me a chocolate cake when Murray Edwards gave me moose and it was these little acts of kindness that got me through some of my darkest times. I have always admired your strength, boldness, and conviction. Thank you for always being a light in what was often a very dark tunnel.

Maren, I can't thank you enough for all of the support you have given me throughout my time at NIH. When I grow up, I can only hope to be half of the scientist and mentor that you are. Zongyi, thank you for always bringing laughter and levity to the lab. You have the admirable ability to brighten any room or life you walk into. To Seung Bum, thank you for commiserating with me about stem cell work as well as collaborating with me and supporting me. To the rest of the LDB, thank you all for welcoming me into this supportive and collaborative branch. The creativity and intelligence of this group is inspiring, and I have grown exponentially as a scientist for having been a part of it.

To Lindsey and Katelyn, thank you for being the two best PhD partners in crime and friends I could ask for. We have been through so much together over the past couple of years and I couldn't have done it without you both. Thank you for your unending support whether it be talking through experimental designs over lunch, group drafting emails, or just commiserating about science together. I admire you both as scientists and as humans and I can't think of anyone I would rather be going through the soul-crushing adventure of a PhD with.

Last but certainly not least, I would like to thank my family, without whom, I would never have reached the finish line. To my mom and dad, words cannot express my gratitude. You have pushed me to fulfil my greatest potential and supported me no matter how far away from you my academic journey has taken me. This accomplishment is as much yours as it is mine because you always made sure I kept "my eye on the prize". Thank you. To Nan, losing you was one of the hardest things I had to endure in my PhD. But I know that you have been watching over me throughout this journey and I hope I've made you proud. To Mike, thank you for being my rock throughout the entire thesis writing process for being my quarantine partner. Your unwavering support has been a lifeline in these uncertain and crazy times. To all of you and more, thank you. Please share in this success, because without you, none of this would have been possible. Here's to the next chapter of failing in the right direction.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND ACRONYMS

A1AT: Alpha-1-anti-trypsin

ABC: ATP-binding cassette

ACC: Acetyl-CoA carboxylase

ACLY: ATP citrate lyase

ADH: Alcohol dehydrogenase

AGPAT: Acyl-sn-glycerol-3-phosphate acyltransferase

AKT: RAC-alpha serine/threonine protein kinase

ALB: Albumin

ALD: Alcoholic liver disease

ALDH: Aldehyde dehydrogenase

AP-1: Activator protein 1

APOB-100: Apolipoprotein B-100

Asp: Aspartic acid

ATF6: Activating transcription factor 6

ATP: Adenosine triphosphate

BAX: Bcl-2-associated X protein

BIP: Binding immunoglobulin protein

BMP4: Bone morphogenetic protein 4

BODIPY: Boron-dipyrromethene

c-MYC: MYC proto-oncogene

C3: Complement component 3

C9: Complement component 9

CAR: Constitutive androstane receptor

CAS9: CRISPR-associated protein 9

CAT: carnitine-acylcarnitine translocase

CCL2: Chemokine ligand 2

CCL5: Chemokine ligand 5

CD36: Cluster of differentiation 36

CDM: Chemically-defined medium

CEBPα: CCAAT/enhance-binding protein alpha

CHOP: C/EBP homologous protein

ChoRE: Carbohydrate response element

ChREBP: Carbohydrate-responsive element-binding protein

CRISPR: Clustered regularly interspaced short palindromic repeats

CRISPRi: CRISPR interference

CXCL1: Interleukin 1

CXCL8: Interleukin 8

CYP: Cytochrome P-450

DAMP: Damage associated molecular pattern

DGAT: Diglyceride acyltransferase

DME: Drug metabolizing enzyme

DNL: De novo lipogenesis

dPBS: Dulbecco's phosphate buffered saline

ECM: Extracellular matrix

ELOVL6: Elongation of very long chain fatty acids protein 6

ENPP1: Ectonucleotide phosphodiesterase 1

ER: Endoplasmic reticulum

FABP1: Fatty acid-binding protein 1

FAC: Ferric ammonium citrate

FACS: Fluorescence-activated cell sorting

FAS: Fatty acid synthase

FATP: Long-chain fatty acid transport protein

FBS: Fetal bovine serum

FFA: Free fatty acid

FGF2: Fibroblast growth factor 2

FOXO1: Forkhead box protein O1

GADD34: Growth arrest and DNA damage inducible protein

GATA4: Transcription factor GATA-4

GCKR: glucokinase regulator

GLUT1: Glucose transporter 1

Gly: Glycine

GM-CSF: Granulocyte-macrophage colony-stimulating factor

GPAT: Glycerol-3-phosphate acyltransferase

GSK3: Glycogen synthase kinase 3

GST: Glutathione S-transferase

GWAS: Genome wide association studies

H-MRS: Proton magnetic resonance spectroscopy

HBV: Hepatitis B virus

HCC: Hepatocellular carcinome

HCV: Hepatitis C virus

HDR: Homology directed repair

HGF: Hepatocyte growth factor

hiPSC: human induced pluripotent stem cell

HK: Hexokinase

HLC: Hepatocyte-Like Cell

HNF4a: Hepatocyte nuclear factor 4 alpha

HNF6: Hepatocyte nuclear factors

HSC: Hepatic Stellate Cell

HSD17B13: 17β-hydroxysteroid dehydrogenase type 13

IFNL3: Interferon lambda 3

IL1β: Interleukin 1 beta

IL6: Interleukin 6

INSIG1: Insulin induced gene 1

iPSC: Induced pluripotent stem cells

IRE1: Inositol-requiring enzyme 1

IRS-1: Insulin receptor substrate 1

JNK: c-JUN N-terminal kinases

KLF4: Kruppel-like factor 4

KO: Knock-out

LCAD: Long chain acyl dehydrogenase

LCFA: Long chain fatty acid

LDL: Low-density lipoprotein

LPAAT: Lysophosphatidic acid acyltransferase

LPC: Lysophosphatidylcholine

LPIN1: Lipin-1

LSEC: Liver Sinusoidal Endothelial Cell

LXR: Liver X receptor

MBOAT7: Membrane-bound O-acyltransferase domain containing protein 7

MCAD: Medium chain acyl dehydrogenase

MCFA: Medium chain fatty acid

MEF: Mitotically-inactivated mouse embryonic fibroblasts

MERTK: Proto-oncogene tyrosine-protein kinase MER

MGAT: Monoglyceride acyltransferase

MLX: Max-like protein X

MMP2: Matrix metalloproteinase-2

MnSOD: Superoxide dismutase

MRI: Magnetic resonance imaging

MUFA: Monounsaturated fatty acid

NAFLD: Non-alcoholic fatty liver disease

NASH: Non-alcoholic steatohepatitis

NFκB: Nuclear factor kappa-light-chain-enhancer of activated B cells

NHEJ: Nonhomologous end joining

OA: Oleic Acid

OCT4: Octamer-binding transcription factor 4

OSM: Oncostatin M

PA: Palmitic Acid

PAM: Protospacer adjacent motif

PBGD: Porphobilinogen deaminase

PCA: Principle component analysis

PERK: Protein kinase R-like endoplasmic reticulum kinase

PI3K: Phosphoinositide 3-kinase

PKC: Protein Kinase C

PNPLA3: Patatin-like phospholipase domain-containing protein 3

PPARα: Peroxisome proliferator-activated receptor α

PPARγ: Peroxisome proliferator-activated receptor γ

PPH-1: Serine/threonine-protein phosphatase

PPRE: Peroxisome proliferator response element

PROX1: Prospero homeobox protein 1

PUFA: Poly-unsaturated fatty acid

PUMA: p53 upregulated modulator of apoptosis

PVA: Polyvinyl alcohol

PXR: Pregnane X Receptor

ROS: Reactive oxygen species

RXR: Retinoid X receptor

S1P: Site-1 protease

S2P: Site-2 protease

SCAD: Short chain acyl dehydrogenase

SCAP: SREBP cleavage activating protein

SCD-1: Stearoyl-CoA desaturase-1

SCFA: Short chain fatty acid

Ser: Serine

SFA: Saturated fatty acid

sgRNA: Single guide RNA

SLC: Solute carrier family

SNP: Single nucleotide polymorphism

SOX2: SRY-box 2

SRE: Sterol response element

SREBP1c: Sterol regulatory element-binding protein 1c

ssODN: single-stranded donor oligonucleotides

SULT: Sulfotransferase

TALEN: Transcription activator-like effector nuclease

TCA: Tricarboxylic acid cycle

TG: Triglyceride

TGFβ: Transforming growth factor β

TIMP1: TIMP metallopeptidase inhibitor 1

TIMP2: TIMP metallopeptidase inhibitor 2

TM6SF2: Transmembrane 6 superfamily 2

TNFα: Tumor necrosis factor α

TRIB1: Tribbles homolog 1

TTR: Transthyretin

UBC: Ubiquitin

UC: Untargeted Control

UCP2: Mitochondrial uncoupling protein 2

UGT: Uridine-diphospho-glucuronosyltransferase

VLCFA: Very long chain fatty acid

VLDL: Very low-density lipoprotein

ZFN: Zinc Finger Nuclease

1 INTRODUCTION

1.1 Liver Structure

1.1.1 Gross and Cellular Anatomy of the Liver

The liver is an extremely complex organ with over 500 diverse functions in the human body. These functions range from secretion of blood components to xenobiotic detoxification and glucose homeostasis [1]. In order to facilitate these many diverse functions, the structure of the liver is extremely complex. The smallest functional unit of the liver is the liver lobule, a polygonal mass of tissue consisting of 3-6 portal triads surrounding a central vein [2]. Each portal triad consists of a branch of the portal vein, a branch of the hepatic artery, and a bile duct. Blood flows through hepatic sinusoids in each lobule from the portal vein to the central vein while bile flows in the opposite direction.

The hepatic sinusoids are blood vessels with a fenestrated, discontinuous endothelium composed of liver sinusoidal endothelial cells (LSECs) [2]. This discontinuous endothelium allows for bidirectional exchange between the blood and hepatocytes. The liver sinusoids as well as the space between the LSECs and hepatocytes, termed the Space of Disse, are home to several non-parenchymal hepatic cell types including Kupffer cells, the tissue resident macrophages, and hepatic stellate cells (HSCs), the liver resident fibroblasts. In addition, the Space of Disse contains several populations of antigen presenting, myeloid, and innate and adaptive lymphoid resident immune cells. These immune cells participate in maintaining organ homeostasis, innate and adaptive immunity, as well as regulating the inflammatory and fibrotic response during injury [3]. The location of these cells allows for efficient immune surveillance and clearance as well as facilitating metabolic functions.

Hepatocytes are the parenchymal cell of the liver and are responsible for the metabolic functions of the organ. As blood flows through the sinusoids, hepatocytes absorb oxygen, nutrients, xenobiotics, and toxins for metabolism, storage, and detoxification [4]. This creates an oxygen and nutrient gradient across the liver lobule which necessitates the zonation of different metabolic processes. The liver can be divided into three functional zones based upon these gradients: zone 1 (periportal), zone 2 (transition zone), zone 3 (pericentral). The periportal zone is where oxygen and nutrient rich blood enters the liver.

Thus, the oxygen and nutrient concentrations decrease along the spectrum from periportal to transition to pericentral. Metabolic processes with zonal organization include xenobiotic metabolism, lipid metabolism, and bile acid synthesis [5].

As bile flows in the opposite direction of blood in the liver, bile acid synthesis mainly occurs in the pericentral area. Bile is produced by hepatocytes before being secreted into bile canaliculi which are lined by cholangiocytes. Cholangiocytes line the entirety of the biliary system and are responsible for transporting and modifying bile. The unique structure of the liver maintains a barrier between the blood and bile using tight junctions between hepatocytes [4]. A schematic of the liver lobule can be found in Figure 1.1.
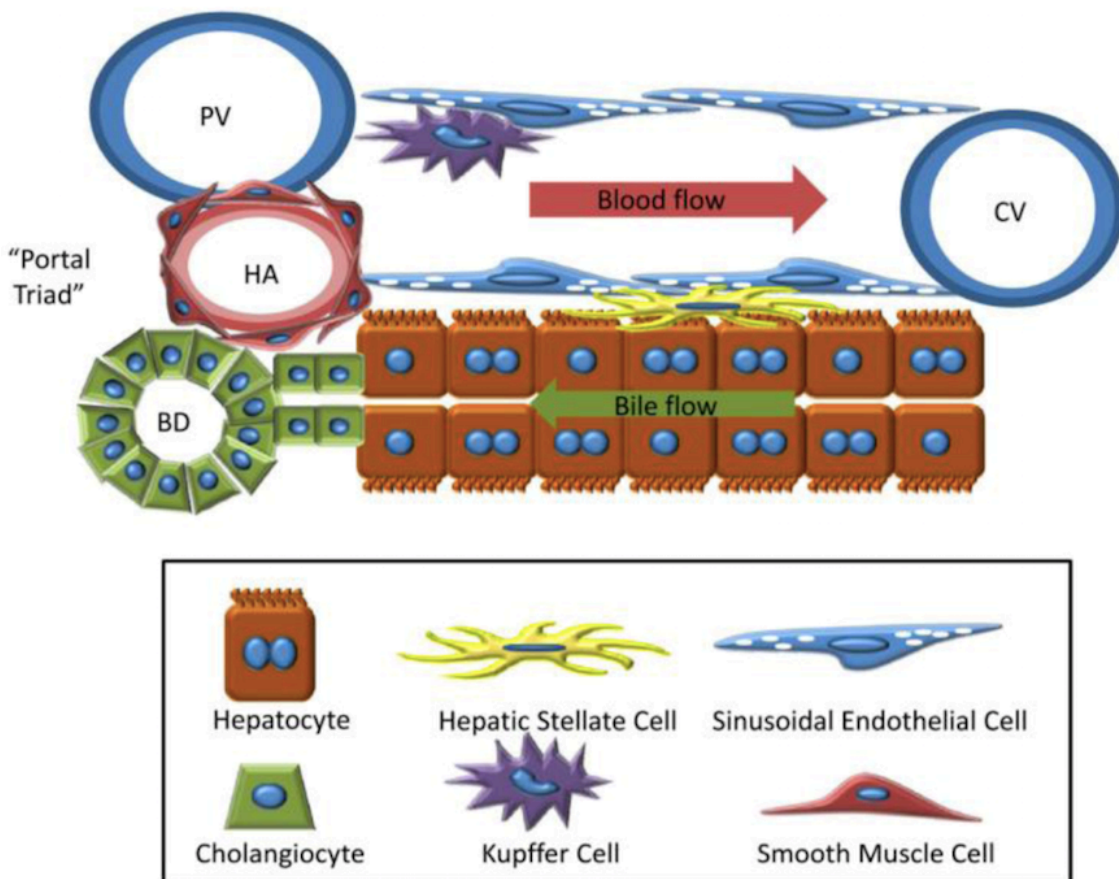
Figure 1.1 Organization of the liver lobule.

Cords of hepatocytes extend between the portal triad and the central vein. Blood flows through liver sinusoids lined by liver sinusoidal endothelial cells from the portal triad to the central vein. Kupffer cells reside within the sinusoid while hepatic stellate cells are located in the Space of Disse, the space between the LSECs and hepatocytes. Bile is secreted by hepatocytes and flows through bile canaliculi in the opposite direction as blood to a bile duct lined by cholangiocytes. (Figure adapted from Kang et al. Cells 2012) [1]

1.1.2 Hepatocytes

Hepatocytes are the most abundant cell type in the liver comprising approximately 60% of the total cell population and 80% of the total volume of the organ [6]. Hepatocytes are responsible for liver functions, including but not limited to, synthesis of serum proteins, storage of vitamins, intermediate metabolism of amino acids, lipids, and carbohydrates, xenobiotic detoxification, and bile acid synthesis.

In order to maintain their functionality, hepatocytes are polarized with three distinct membrane domains: sinusoidal (basal), lateral, and canalicular (apical). The sinusoidal membrane domain is in contact with the blood and non-parenchymal cells to facilitate export of serum proteins and import of oxygen, nutrients, and toxins from the blood. This surface has microvilli that extend into the Space of Disse, interact with HSCs, and occasionally extend through the fenestrae of the LSECs directly into the blood. These microvilli increase the surface area of the hepatocyte and enhance the capacity for receptor and transport functions. The lateral surface forms the blood-bile barrier by fusing hepatocytes together with tight junctions. The canalicular domain forms a bile canalicular network between adjacent hepatocytes to transport bile produced by hepatocytes to the bile ducts [4, 7]. Bile is a complex fluid that contains organic and inorganic substrates in an alkaline solution. Bile aids in digestion and absorption of lipids and lipid soluble vitamins while simultaneously eliminating waste products such as bilirubin and cholesterol.

The unique vascular organization of the liver exposes hepatocytes to a lobular gradient of oxygen, nutrients, toxins, and hormones. As a result, hepatocytes are heterogeneous in both their structure and function along this gradient. Hepatocytes in the periportal region are exposed to a high concentration of nutrients and oxygen, as a result, they are small, participate in gluconeogenesis and fatty acid degradation, and have a low capacity for xenobiotic detoxification. Alternately, pericentral hepatocytes are larger and are responsible for bile acid synthesis and xenobiotic metabolism. Transitional hepatocytes have structure and function intermediate between periportal and pericentral hepatocytes [5]. A schematic of metabolic zonation can be found in Figure 1.2. Hepatocytes are unique

among terminally differentiated epithelial cells as they maintain the ability to divide throughout adult life. This ability to re-enter the cell cycle, in part, grants the liver its unique ability to regenerate following injury [6].
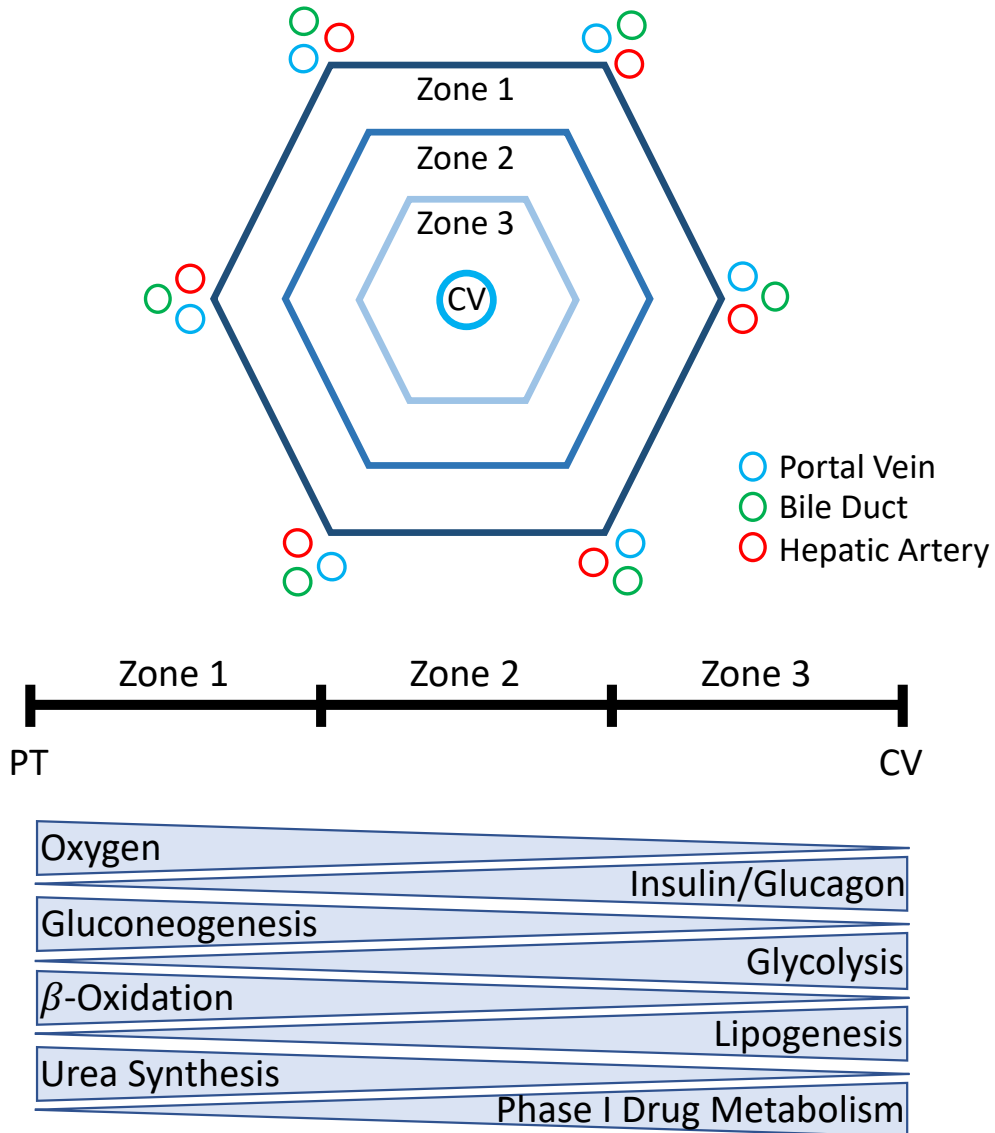
Figure 1.2 Metabolic Zonation.

The liver lobule can be divided into three functionally different metabolic zones. Zone 1 also known as the periportal region has high levels of oxygen and hepatocytes in this zone are responsible for gluconeogenesis, β-oxidation, and urea synthesis. Zone 2 is labelled the transition zone and hepatocytes in this zone represent a functional intermediate between Zone 1 and Zone 3 hepatocytes. Zone 3 is known as the pericentral region due to its proximity to the central vein. Hepatocytes in this region are exposed to low levels of oxygen and high levels of hormones and are responsible for glycolysis, lipogenesis, and phase one drug metabolism.

1.1.3 Cholangiocytes

Cholangiocytes, also known as biliary epithelial cells, represent only 3-5% of the cells in the liver. These cells line the biliary tree and are responsible for transporting and modifying bile. Cholangiocyte secretions account for approximately 40% of the total volume of bile. As the bile is transported from canaliculi, through progressively larger bile ducts, to the gallbladder, the bile is alkalinized and diluted by cholangiocytes in a series of secretory and absorptive processes. Cholangiocytes possess cilia on their apical plasma membrane that extends into the bile duct lumen to detect changes in bile flow, composition, and osmolality.

Cholangiocytes are heterogeneous throughout the biliary tree. Large cholangiocytes line the larger ducts and are responsible for mucin secretion and hormone regulated bile secretion while small cholangiocytes line the smaller ducts and have high proliferative capacity and considerable plasticity. Upon liver injury, cholangiocytes can participate in the regenerative process either by repairing bile ducts or transdifferentiating into hepatocytes in a process termed ductular reaction. Though this process can contribute to liver regeneration, in many chronic liver diseases, prolonged ductular reactions may contribute to disease progression and fibrosis. The study of these reactions have been hindered because the type and mechanism of ductular reaction is dependent on the injury, disease, and model system used; therefore, additional studies are needed to fully elucidate the role of these reactions in liver regeneration and disease [8-10]. Additionally, cholangiocytes stimulate angiogenesis in response to nutritional and functional demands of proliferating cells. It is believed that small cholangiocytes are responsible for this function. Cholangiocytes also participate in the immune system by secreting chemokines and cytokines to localize and coordinate immune responses [11, 12].

1.1.4 Hepatic Stellate Cells

Hepatic stellate cells (HSCs) represent approximately 5-8% of liver cells. These cells reside in the Space of Disse in close proximity to both hepatocytes and LSECs. These cells have long, branching cytoplasmic processes that allow them to cover the entire hepatic sinusoidal microcirculatory network [13]. In their quiescent state, HSCs are responsible for vitamin A storage in the form of retinyl esters. 50-80% of retinol within

the body is stored in the liver and 80-90% of those retinoids are stored in lipid droplets in HSCs [14].

When the liver faces a chronic injury that overwhelms the normal regenerative processes, HSCs become activated. These activated HSCs lose their retinoid stores, acquire a highly proliferative myofibroblast-like phenotype, and become the primary producer of extracellular matrix components. Activated HSCs produce mostly collagen 1 and are the cell type thought to be responsible for liver fibrosis. HSCs can be activated by mediators secreted by immune cells and necrotic hepatocytes. Additionally, activated HSCs can recruit immune cell transmigration into the parenchymal space thereby further enhancing hepatic injury [15]. A schematic of HSC activation can be found in Figure 1.3.
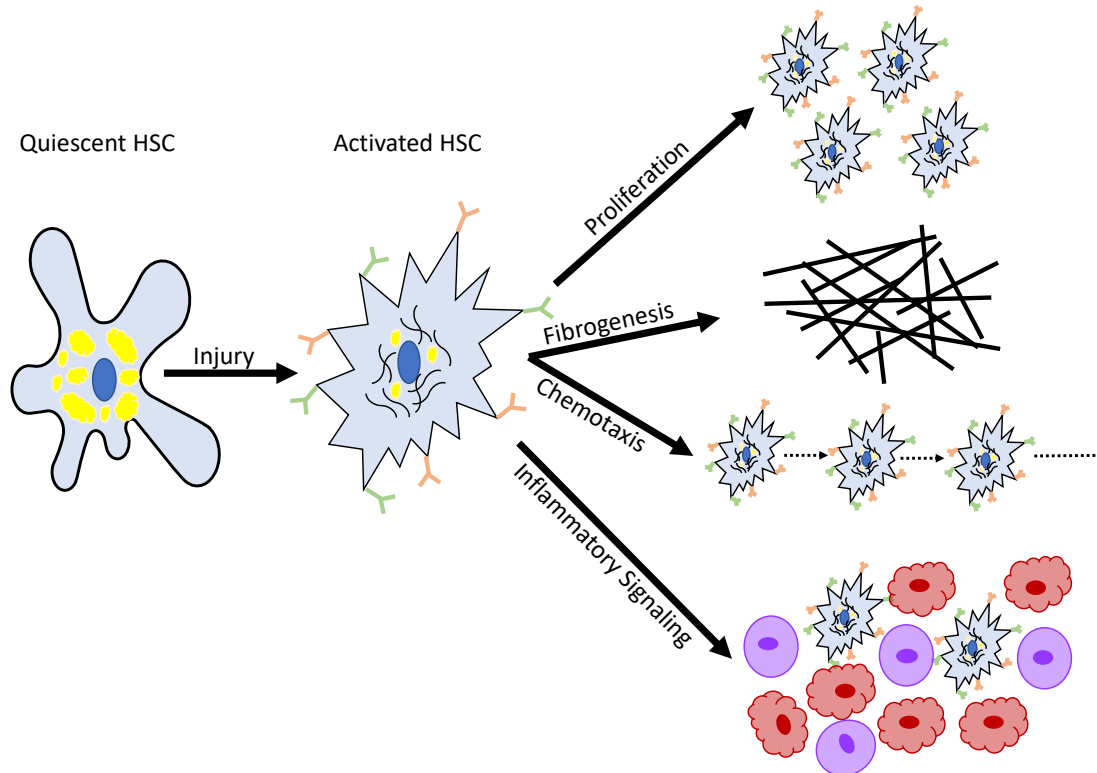
Figure 1.3 Activation of HSCs.

Quiescent HSCs are responsible for storing the majority of the body's retinol in lipid droplets. However, upon liver injury, these HSCs can become activated which causes several phenotypic changes. Activation of HSCs causes the depletion of intracellular retinol stores, increases proliferation and chemotaxis, stimulates fibrogenesis through production of irregular ECM components, and produces pro-inflammatory cytokines that modulate the inflammatory response to the injury.

1.1.5 Kupffer Cells

Kupffer cells are the resident liver macrophages. These cells are the largest tissue resident macrophage population in the body and make up approximately 12% of the cells within the liver. These cells reside in the sinusoidal lumen attached to LSECs. They function in the removal of gut- and environment-derived toxins, microorganisms, senescent and damaged red blood cells, and circulating neoplastic cells. The blood flow through sinusoids is slow which prolongs the interaction between Kupffer cells and portal blood which enhances the removal capacity of these cells. Kupffer cells can be activated by LPS and the complement system. Upon activation, Kupffer cells produce large quantities of inflammatory mediators which contribute to liver injury. The distribution of Kupffer cells is variable throughout the liver lobule with the highest number in the periportal region (~40%) and the lowest in the pericentral region (~25%). The Kupffer cells in the periportal region are the largest, perform the most phagocytosis, and produce the most cytokines of all the regions [16, 17].

1.1.6 Liver Sinusoidal Endothelial Cells

Liver sinusoidal endothelial cells (LSECs) are the second most abundant cell type of the liver comprising approximately 20% of all cells in the organ. LSECs are positioned between the blood and hepatocytes and comprise the liver sinusoids. The LSECs create a discontinuous endothelium with fenestrae that act as a sieve. The fenestrae measure approximately 150-175 nm and occupy 6-8% of the total endothelial surface. Porosity and fenestration diameter increase across the liver lobule from the periportal region to the pericentral region. LSEC protect parenchymal cells from direct contact with the blood by blocking passage of particles larger than 200 nm. Only smaller remnants can enter the Space of Disse and be metabolized by hepatocytes. The diameter of fenestrae is influenced by many factors and certain drugs can lead to permanent defenestration of LSECs which causes atherosclerosis and hyper-lipidemia by preventing the removal of lipoproteins from the blood [18].

These cells also possess endocytic capabilities and act as an important blood clearance system. They express at least 5 types of scavenger receptors to remove biomacromolecules, proteins, polysaccharides, lipids, and nucleic acids. LSECs regulate the migration of immune cells into the parenchymal space thereby mediating the immune

reaction to liver injury. As antigen presenting cells, LSECs are also a major contributor to the innate immunity in the liver [19].

1.2 Metabolic Functions of the Liver

1.2.1 Glucose Homeostasis

The liver is the metabolic hub of the body controlling the storage and availability of glucose to other tissues. The liver is the major source of post-prandial glucose uptake with 30-40% of portal vein glucose being absorbed on the first pass. This glucose can then be utilized by several different pathways. The glucose can be metabolized through glycolysis to meet the immediate energy needs of the organ. Alternately, the glucose may enter one of two storage pathways: de novo lipogenesis or glycogen synthesis. In addition to its role in post-prandial glucose storage, the liver is also capable of de novo synthesis of glucose from precursors such as lactate, amino acids, and glycerol in a process called gluconeogenesis [20].
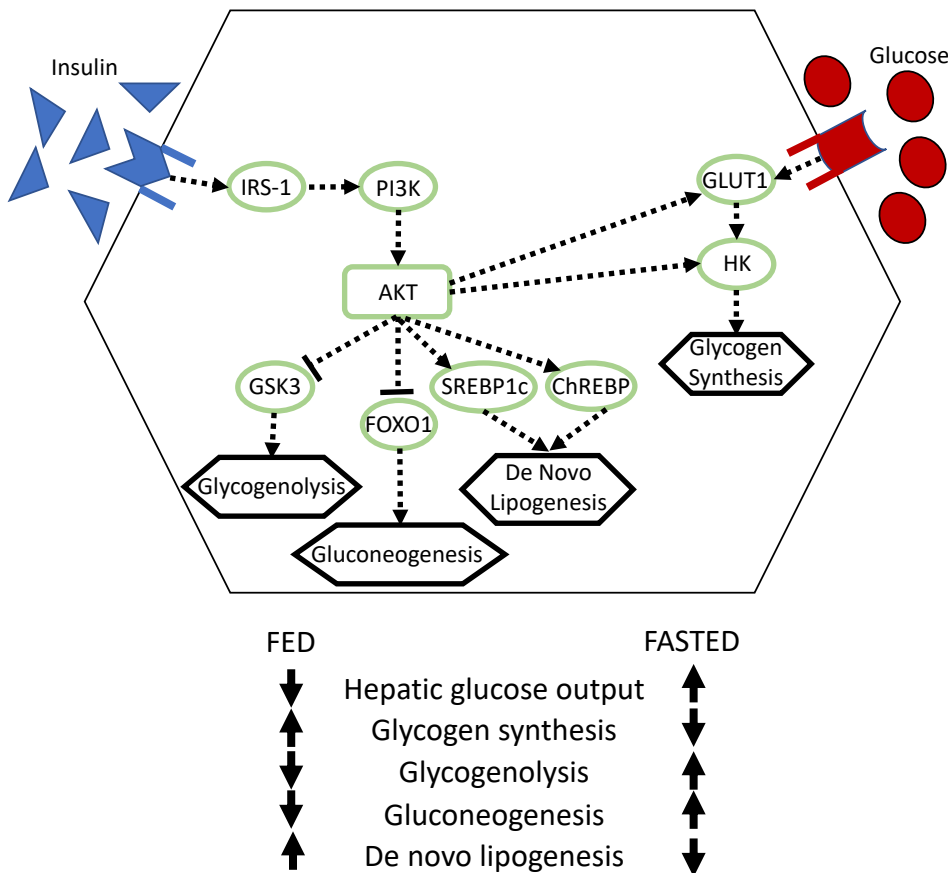
Figure 1.4 Glucose Homeostasis.

Schematic of AKT signalling in response to insulin and glucose. The binding of insulin to its receptor causes the phosphorylation of IRS-1 which in turn activates PI3K. PI3K activation then leads to the phosphorylation of AKT which has several downstream effects. AKT inhibits GSK3 and FOXO1 which leads to the downregulation of glycogenolysis and gluconeogenesis, respectively. Alternately, AKT activates SREBP1c and ChREBP which stimulates de novo lipogenesis. AKT also activates glycolysis and glycogen synthesis by upregulating the expression of GLUT1 and HK. In the fed state, glucose storage pathways such as glycogen synthesis and de novo lipogenesis are upregulated. In the fasted state, glucose production pathways such as glycogenolysis and gluconeogenesis are upregulated to meet the metabolic needs of the body.

Glucose homeostasis in the liver is under tight nutritional and hormonal control. In the post-prandial state, glucose and insulin levels are high which leads to the activation of the AKT signalling cascade. AKT activation leads to the upregulation of SREBP1c and ChREBP which are the master regulators of the de novo lipogenesis pathway. In addition, AKT inhibits glycogenolysis and gluconeogenesis through phosphorylation of GSK3 and FOXO1, respectively [21]. In the fasted state, insulin levels are low which releases the inhibitory effect on FOXO1. This transcription factor activates de novo glucose production by upregulating the gluconeogenic gene program and increasing the flux of gluconeogenic precursors to the liver. FOXO1 also inhibits glycolysis and de novo lipogenesis to prevent glucose utilization by the liver in this state [22, 23]. In the fasted state, glucagon levels are also very high. Glucagon activates glycogen phosphorylase which leads to glycogenolysis [24, 25]. In all, during times of high glucose availability, the liver uptakes and stores excess glucose while in times of low glucose availability, the liver generates and releases glucose for use in other tissues. A summary of the signalling pathways involved in glucose homeostasis can be found in Figure 1.4.

1.2.2 Lipid Metabolism

The liver is also a major site of lipid metabolism within the body. Lipids in the liver can be derived from three major sources: de novo lipogenesis, non-esterified fatty acids from the lipolysis of adipose tissues, and dietary fats transported to the liver in chylomicrons. Fatty acids within the liver have three major fates: storage as triglycerides in lipid droplets, secretion in very low-density lipoproteins (VLDL), or oxidation for energy utilization [26].
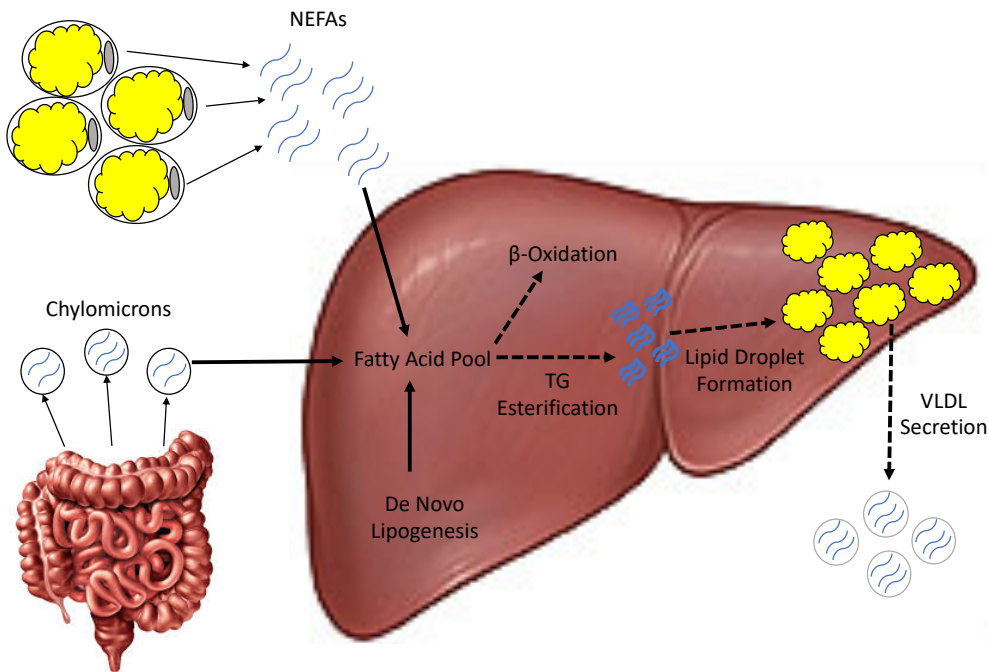
Figure 1.5 Fatty acids in the liver.

There are three primary sources of fatty acids within the liver: diet, adipose tissue, and hepatic de novo lipogenesis. Fatty acids from the diet are transported to the liver in chylomicrons while fatty acids from adipose tissue are transported in the form of non-esterified fatty acids. Upon entering a hepatocyte, the fatty acids can be metabolized through β-oxidation or esterified into triglycerides for storage in lipid droplets or secretion as VLDL.

1.2.2.1 De Novo Lipogenesis

De novo lipogenesis (DNL) converts excess carbohydrates to fatty acids for energy storage. Glucose is first converted to pyruvate before entering the Krebs Cycle to yield citrate. Citrate is then broken down into acetyl-CoA and oxaloacetate. Acetyl-CoA is then converted to malonyl-CoA by acetyl-CoA carboxylase (ACC) in the rate limiting step. Fatty acid synthase (FAS) then adds successive molecules of malonyl-CoA to an acetyl-CoA primer to generate palmitic acid. Palmitic acid is a saturated fatty acid with a 16-carbon chain. Palmitic acid can then be desaturated by stearoyl-CoA desaturase (SCD-1) to form a monounsaturated fatty acid. Alternately, it can be elongated by elongation of very long chain fatty acids protein 6 (ELOVL6) which catalyses the addition of 2 carbon units to fatty acid chains [26]. These fatty acids can then be esterified into triglycerides for storage or oxidized in the mitochondria for energy production.

Samantha G. Tilson - January 2021

De novo lipogenesis is tightly regulated by nutritional, hormonal, and genetic factors. The master regulators of de novo lipogenesis are the transcription factors SREBP1c and ChREBP [21, 27]. SREBP1c is regulated by intracellular sterol levels, insulin, and glucagon. SREBP1c is initially produced as a large precursor molecule that is bound to the ER membrane. SCAP is a sterol sensing protein that acts as an escort protein for SREBP1c from the ER to the Golgi Apparatus. When sterols are abundant, SCAP undergoes a conformational change and is unable to bind to SREBP; however, when sterols are absent, SCAP binds to SREBP1c and facilitates its translocation to the Golgi. Following this translocation, SREBP1c is proteolytically cleaved from the membrane by S1P and S2P before translocating to the nucleus. Nuclear SREBP1c activates lipogenic genes by binding the sterol response elements in their promoter/enhancer regions [27, 28]. ChREBP is a glucose responsive transcription factor that regulates the expression of lipogenic genes specifically induced by carbohydrates. Glucose availability controls both the nuclear translocation and functionality of ChREBP. Under fasting conditions, ChREBP is phosphorylated which sequesters it to the cytosol. ChREBP also has a glucose sensing module that controls its ability to bind to DNA. In low glucose concentrations, the low glucose inhibitory domain causes a conformational change that prevents ChREBP from binding to DNA. High glucose concentrations result in both dephosphorylation and restoration of DNA binding capacity. ChREBP and its functional partner MLX then form a heterodimer that binds to the ChoRE motifs in target lipogenic genes [29, 30]. Both SREBP1c and ChREBP activation leads to the upregulation of de novo lipogenesis genes including ACC, FAS, SCD1, and ELOVL6. An overview of this transcription regulation can be found in Figure 1.6.
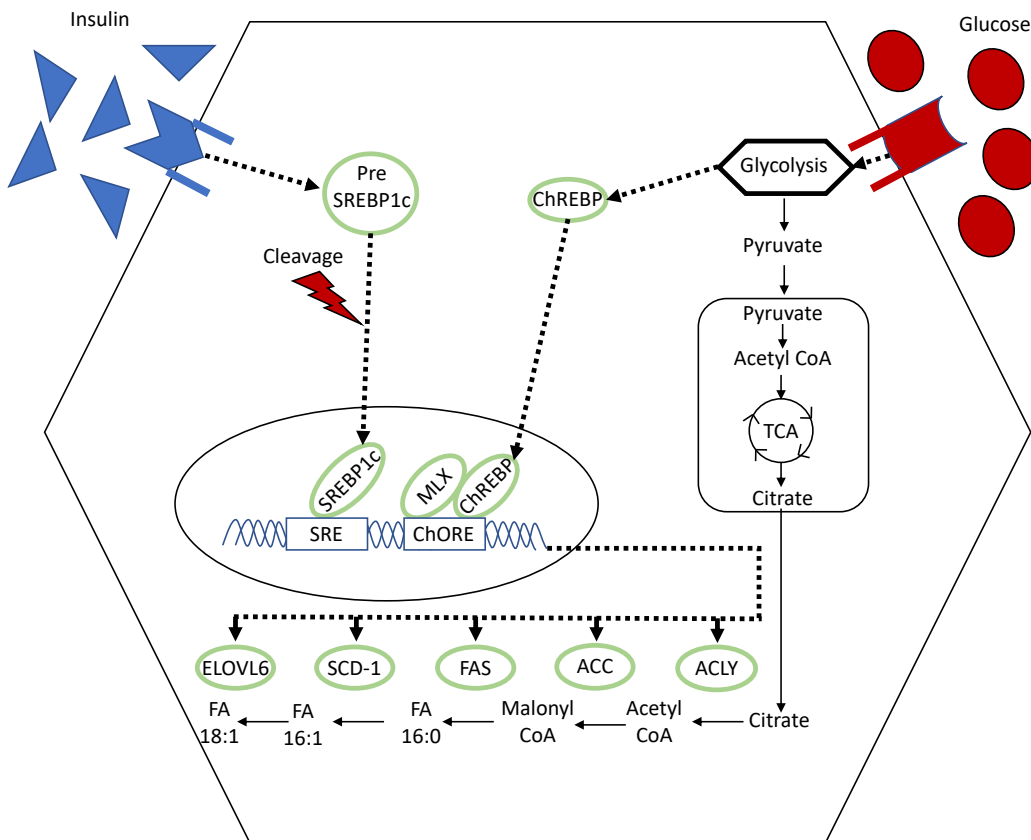
Figure 1.6 De Novo Lipogenesis.

Transcriptional regulation of de novo lipogenesis in hepatocytes. High levels of insulin cause increased transcription and nuclear translocation of SREBP1c where it binds to sterol response elements and promotes the transcription of downstream lipogenesis genes such as ACLY, ACC, FAS, SCD-1, and EVOVL6. High glucose concentrations have the dual effect of activating glycolysis to increase the availability of de novo lipogenesis precursors such as citrate as well as activating ChREBP which translocates to the nucleus and binds to carbohydrate response elements along with its functional partner MLX in order to stimulate the transcription of de novo lipogenesis genes.

1.2.2.2 Lipid Storage in Lipid Droplets

Lipids within hepatocytes derived from exogenous or endogenous pathways must often be stored for later use. This storage is accomplished using a specialized organelle called a lipid droplet. Lipid droplets contain a pool of neutral lipids surrounded by a phospholipid monolayer membrane decorated with various lipid droplet proteins. These droplets serve as an immediate energy source for the hepatocyte. However, lipid droplets are not inert storage receptacles for fat. These organelles are extremely dynamic and play a vital role in lipid homeostasis and intracellular signalling. The protein coating of the lipid droplet is tightly controlled and can be efficiently altered in order to regulate lipid

droplet biogenesis and balance the metabolic needs of the cell at any given time. Additionally, lipid droplets may function as a type of chaperone to transport specific lipid species to different cellular locations for specific functions [31-34]. Free fatty acids, especially saturated fatty acids, can be harmful to hepatocytes by inducing oxidative stress and mitochondrial dysfunction. Conversion of these metabolically harmful species to metabolically inert triglycerides in lipid droplets represents an adaptive mechanism to protect against lipotoxicity [35]. However, if acquisition of lipids outpaces their utilization and/or secretion, excessive numbers of lipid droplets can accumulate within the cell [26]. This accumulation of lipid droplets causes hepatic steatosis which can result in insulin resistance and metabolic dysfunction.

1.2.2.3 VLDL Secretion

VLDLs are lipoprotein particles that are assembled and secreted by the liver when plasma lipid levels are too low or the lipid content within the liver is too high. These lipoproteins contain triglycerides, cholesterol, and phospholipids. VLDLs are assembled in a two-step process that is dependent on the availability of the protein apolipoprotein B-100 (APOB-100). VLDL can be separated from the other major lipoprotein, chylomicrons, by the APOB protein that is used as a scaffolding. VLDL are built upon APOB-100 while chylomicrons are assembled on APOB-48 [36]. Following synthesis, ApoB-100 is translocated through the rough ER. During this translocation, ApoB-100 acquires a small number of triglycerides, phospholipids, and cholesterol esters and forms a small, dense VLDL-precursor. This precursor then goes through a maturation process in the smooth ER where it is fused with lipid droplets [37, 38]. The fully lipidated VLDL is then transported to the plasma membrane and secreted into circulation. These VLDL particles can then be hydrolysed in the peripheral vascular bed releasing the fatty acids for use by other peripheral tissues.

1.2.2.4 Fatty Acid β-Oxidation

Given that fatty acids produce the most energy per gram of any macronutrient, in times of low energy availability, they can be broken down and utilized for ATP production. β-oxidation is the catabolic process by which fatty acids are degraded into acetyl-CoA subunits that can then enter the citric acid cycle to produce energy. There are two major β-oxidation pathways. The mitochondrial β-oxidation pathway is the primary pathway

and is responsible for the degradation of short, medium, and long-chain fatty acids [39]. Whereas degradation of long and very long-chain fatty acids occurs via the peroxisomal β-oxidation pathway [40].
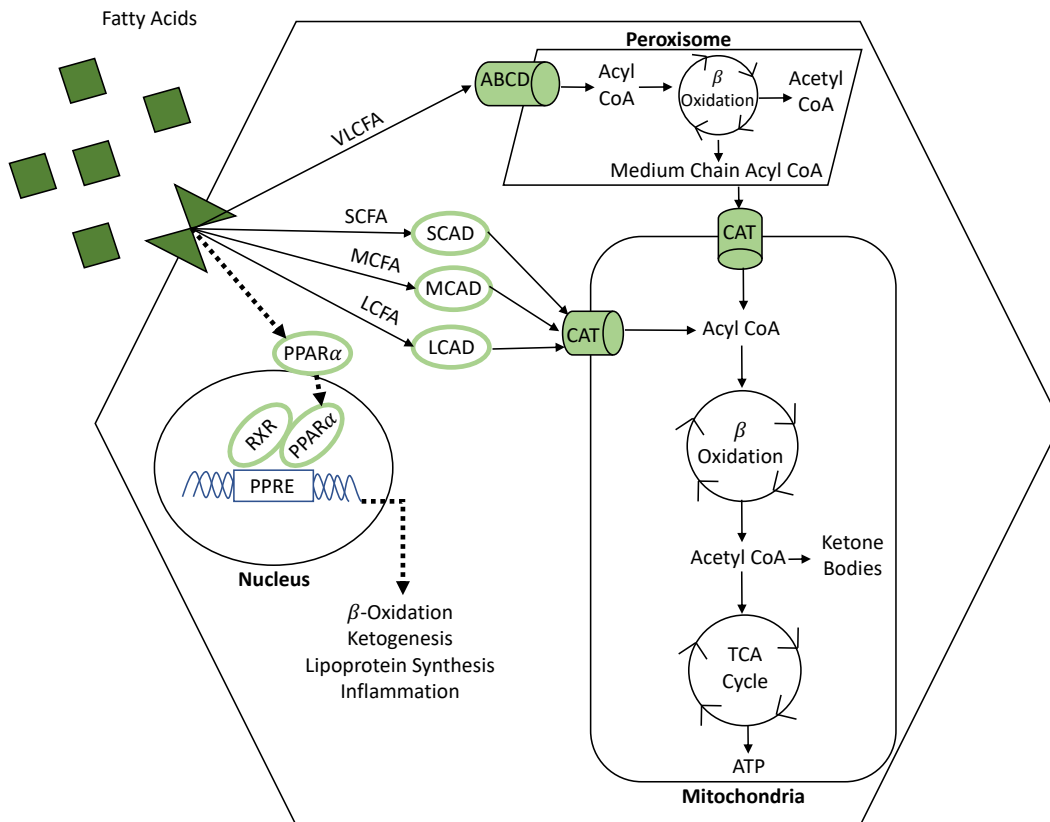
Figure 1.7 β-Oxidation.

Schematic of mitochondria and peroxisomal β-oxidation. Upon entering the cell, fatty acids activate PPARα which binds in conjunction with RXR to peroxisome proliferator response elements in target genes to upregulate the expression of genes involved in β-oxidation, ketogenesis, VLDL synthesis, and inflammation. The length of the fatty acid determines if it will first enter the peroxisome or the mitochondria for oxidation. Short, medium, and long chain fatty acids are activated into acyl-CoA esters by their respective acyl-CoA dehydrogenase enzymes before entering the mitochondria through the carnitine shuttle. In the mitochondria, the acyl CoA is slowly shortened by releasing two-carbon acetyl CoA subunits at a time. These acetyl CoA subunits are then released as ketone bodies or enter the TCA cycle to produce ATP. Alternately, very long chain fatty acids cannot be transported into the mitochondria. Therefore, these fatty acids are partially oxidized in the peroxisome until they are shortened enough to enter the mitochondria for full oxidation.

The first step in mitochondrial β-oxidation is the activation of fatty acids into acyl-CoA esters [41]. These acyl-CoA esters can then be transported across the inner mitochondrial

membrane by the carnitine shuttle. Following entry into the mitochondria, the acyl-CoA is progressively shortened by releasing two carboxy-terminal carbon atoms as acetyl-CoA subunits. These subunits can then be condensed into ketone bodies for export to other organs or enter the citric acid cycle to be further oxidized into water and carbon dioxide to produce ATP. The shortened acyl-CoA then re-enters the fatty oxidation cycle until it has been fully oxidized [42]. Peroxisomal β-oxidation is used to degrade very long chain fatty acids because the mitochondria lack the ability to convert these fatty acid species into acyl-CoA esters which prevents them from being imported into the mitochondria. The energy generated in peroxisomal β-oxidation dissipates as heat and therefore is less energy efficient than mitochondrial β-oxidation [43, 44]. For this reason, peroxisomal β-oxidation does not run to completion. Rather, acyl-CoA esters are shortened to the point they can be imported into the mitochondria for completion. An overview of the mitochondrial and peroxisomal β-oxidation pathway can be found in Figure 1.7.

Fatty acid β-oxidation is nutritionally and transcriptionally controlled. Peroxisome proliferator-activated receptor α (PPARα) is the master regulator of all the β-oxidation pathways [41, 45, 46]. PPARα is a nuclear receptor protein that, upon activation by peroxisome proliferators, functions as a transcription factor. When activated, PPARα heterodimerizes with the retinoid X receptor (RXR) and binds the peroxisome proliferator hormone response element in target genes to activate the β-oxidation gene program. PPARα is activated by the presence of excess saturated and polyunsaturated fatty acids. Under fasting conditions, there is an increase in flux of non-esterified fatty acids from the adipose tissue into the liver which results in the activation of PPARα signalling cascade [47-49]. In addition to activating β-oxidation, PPARα also plays a role in lipoprotein synthesis, inflammatory responses, and liver cancer.

1.2.3 Drug Metabolism

The liver plays an extremely crucial role in metabolizing xenobiotics and toxins. The liver metabolizes these drugs in order to either modify, bioactivate, detoxify, and/or excrete them. Hepatic drug metabolism is divided into three phases: modification, conjugation, and excretion [50]. Phase I reactions are generally facilitated by the cytochrome P450 (CYP) superfamily that function as monooxygenases. There are 57 CYP isoforms that

have been identified in humans, each with a unique substrate affinity and functionality [51]. Each CYP isoform can be induced or inhibited by various endogenous and exogenous substrates. This pronounced variability makes phase I drug metabolism particularly complex. The second phase of xenobiotic metabolism is conjugation. Enzymes reduce the reactivity and increase the polarity of the drug by adding charged and bulky groups such as methyl, acetyl, sulfate, glutathione, glycine, and glucuronic acid. This increased polarity primes the substrate for the third phase, excretion. The substrates are excreted from the cell through specialized transporters in the plasma membrane. These substrates can either be excreted into the serum or into bile for disposal. A schematic of hepatic drug metabolism is shown in Figure 1.8.
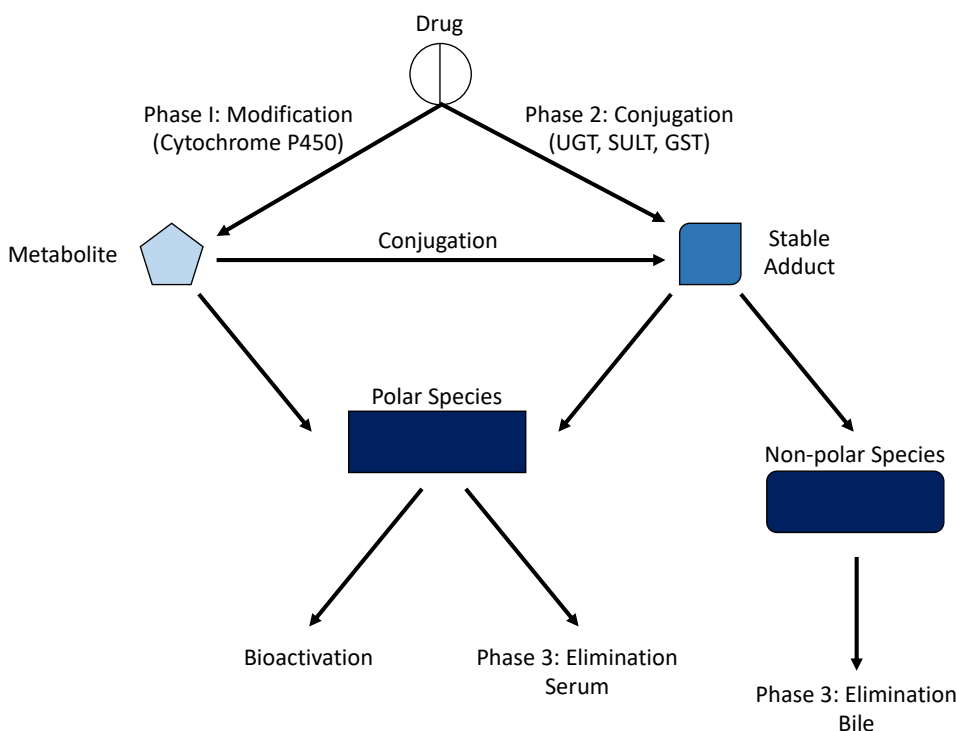
Figure 1.8 Hepatic drug metabolism.

Basic overview of hepatic drug metabolism. Upon entering a hepatocyte, a drug may be modified by Phase I enzymes such as CYP enzymes to produce a metabolite or conjugated by Phase 2 enzymes to form a stable adduct. Following the first two phases of metabolism, the resulting metabolite can be categorized as either polar or non-polar. Non-polar species are excreted into the bile for faecal elimination. Polar species can then undergo further metabolic processes that lead to its bioactivation or they can be secreted into the serum for elimination in the urine. Secretion into either the bile or the serum is facilitated by phase 3 enzymes.

1.3 Non-Alcoholic Fatty Liver Disease

Non-alcoholic fatty liver disease (NAFLD) is now the leading cause of chronic liver disease in the developed world [52-54]. In basic terms, NAFLD arises from an imbalance between triglyceride acquisition and removal. NAFLD is thought to be the hepatic manifestation of the metabolic syndrome and ranges from simple steatosis to non-alcoholic steatohepatitis (NASH). Simple steatosis is defined as the benign accumulation of fat within the liver. NASH is a progressive liver injury that is characterized by a profound inflammatory response that can develop into fibrosis, cirrhosis, and hepatocellular carcinoma. Only a small subset of patients will progress from simple steatosis to NASH and the mechanism causing this progression has yet to be fully elucidated [55]. Factors that lead to increased metabolic dysfunction and obesity such as high fat diet and sedentary lifestyle contribute to NAFLD development; however, there is a growing appreciation for the role of genetics and epigenetics in NAFLD disease progression [56].

1.3.1 Public Health Burden of NAFLD

NAFLD is defined as the presence of hepatic steatosis, defined as fat accumulation in at least 5% of hepatocytes, when no other cause for hepatic fat accumulation is present [52, 57]. The epidemiology of NAFLD is difficult to quantify given the lack of a reliable, affordable, non-invasive method to measure liver fat that can be used to precisely ascertain the number of affected individuals worldwide. Currently, several different techniques are used to quantify liver fat (ultrasound, elastography, H-MRS, MRI, biopsy, etc) with varying levels of accuracy and accessibility [57]. Given these discrepancies, the reported prevalence of NAFLD ranges from 6-35% [52]. The prevalence is highest in western industrialized countries with a median prevalence of 25-26% [52, 58]. The global prevalence is summarized in Figure 1.9.
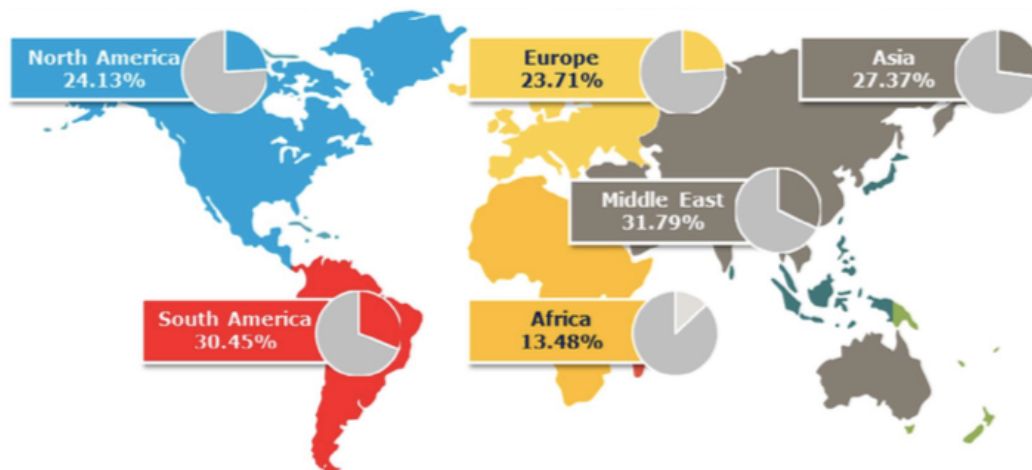
Figure 1.9 Global prevalence of NAFLD.

The prevalence of NAFLD is difficult to quantify but it is currently estimated to affect approximately one in four adults globally. The highest prevalence is in the Middle East and South America while the lowest prevalence is in Africa. (Figure adapted from Younossi et al. Hepatology 2018) [59]

Some patients with simple steatosis will go on to develop more severe forms of the disease such as NASH and fibrosis. It is difficult to estimate what percentage of NAFLD patients suffer from NASH given the invasive nature of liver biopsy which is necessary to confirm this diagnosis. However, studies estimate that up to one in three patients with NAFLD have biopsy confirmed NASH. Of the patients with NASH, 20-30% will progress to severe fibrosis within 10 years and 10-29% will develop cirrhosis. In patients with NASH-induced cirrhosis, 4-27% will go on to develop hepatocellular carcinoma [55, 56]. A schematic of NAFLD progression can be found in Figure 1.10. NASH is projected to become the leading indication for liver transplantation in the United States within the next 10 years [60]. However, 70% of patients that receive a liver transplantation for NASH go on to develop NAFLD in the transplanted liver [61].
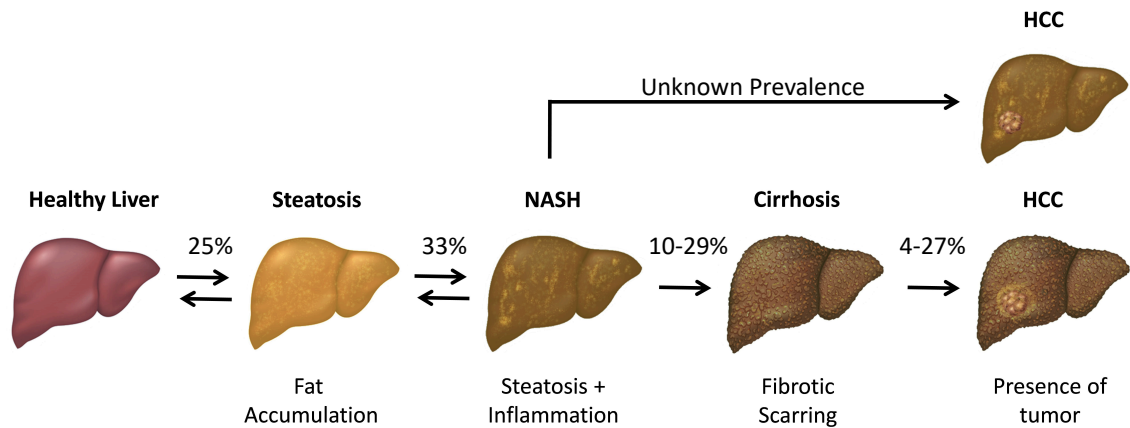
Figure 1.10 NAFLD progression.

The natural progression of NAFLD follows from steatosis to NASH to cirrhosis and finally to hepatocellular carcinoma. Some patients develop hepatocellular carcinoma without underlying cirrhosis and others fail to progress along the full spectrum of the disease. In the general population, approximately 25% of adults have hepatic steatosis. Of those patients, approximately one in three will develop NASH. In patients with NASH, 10-29% will go on to develop cirrhosis and 4-27% will develop hepatocellular carcinoma. (Figure adapted from Turchinovich et al. Frontiers in Physiology 2018) [62]

Patients with NAFLD have a 34%-69% increased risk of mortality over 15 years compared with the general population. Cardiovascular disease is the number one cause of death in NASH patients with hepatic complications accounting for approximately 5% of deaths [56]. In addition to cardiovascular disease, NASH patients have an increased risk of bowel and breast cancer. Despite this extremely high morbidity and mortality, there are currently no FDA approved pharmaceuticals for the treatment of NAFLD or NASH. At present, the only available treatment for this disease is lifestyle modification such as weight loss and exercise. A reduction in body weight by 5-10% has been shown to result in reversion of NASH and fibrosis. However, lifestyle modifications are difficult to achieve and maintain resulting in limited success of this treatment in the clinical setting [58].

1.3.2 Pathogenesis of NAFLD

The pathogenesis of NAFLD generally involves several individual insults that culminate in severe disease including severe lipid accumulation, insulin resistance, oxidative stress, and mitochondrial dysfunction leading to lipotoxicity and inflammation. Scientists have adopted the multiple-hit hypothesis to describe the pathogenesis of NAFLD, illustrated

in Figure 1.11. This hypothesis asserts that NAFLD progression requires a combination of multiple genetic, external, and intracellular events [63]. Some patients are predisposed to developing NAFLD due to genetic and epigenetic factors. In these patients, several environmental factors such as insulin resistance and oxidative stress lead to a state of chronic hepatic inflammation due to the activation of several different hepatocellular damage pathways. For this reason, the pathogenesis of NAFLD is extremely complex and the mechanism of disease progression is unique to each patient.
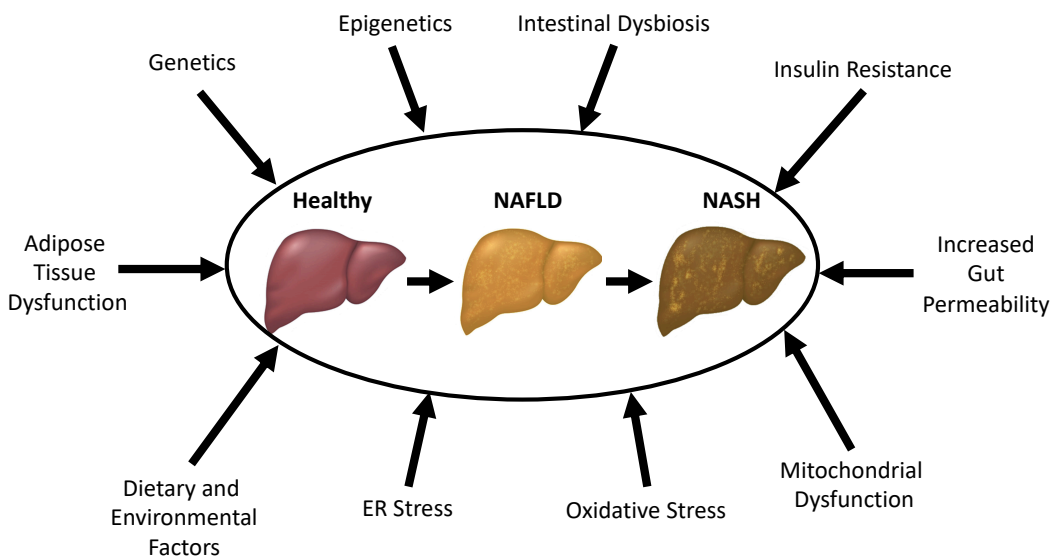
Figure 1.11 Multiple hit hypothesis of NAFLD progression.

The development and progression of NAFLD is mediated by several genetic, environmental, and intracellular factors. For example, genetic and epigenetic factors may make a person more susceptible to environmental insults such as diet which leads to insulin resistance, ER stress, and mitochondrial dysfunction which all contribute toward the inflammatory state and progression toward NASH. Other factors that may contribute to NAFLD progression are intestinal dysbiosis, increased gut permeability, oxidative stress, and adipose tissue dysfunction. Not all factors are necessary or sufficient for NAFLD progression to NASH making the progression extremely complex and unique to each patient.

The mechanisms that facilitate disease progression are still under investigation. Our understanding of the disease mechanisms has been severely limited by the lack of pre-clinical models that reflect the full spectrum of human disease. Mouse models using high-fat diets, methionine and choline deficient diets, and atherogenic diets have failed to reproduce the histology, pathophysiology, and/or the metabolic abnormalities seen in humans [64]. In order to fully elucidate the mechanisms of human disease and develop

drugs for its treatment, better pre-clinical models must be developed. In the following section, I will describe in general terms the current understanding of the pathogenesis of NAFLD from simple steatosis through NASH and fibrosis.

1.3.2.1 Lipid Accumulation

Hepatic fat accumulation or steatosis occurs when lipid acquisition outpaces lipid disposal in the liver. Steatosis can be divided into two categories: microsteatosis and macrosteatosis. Microsteatosis is defined as the accumulation of innumerable small lipid droplets with the nucleus remaining in its original location. Macrosteatosis occurs when larger lipid droplets displace the nucleus from its original location. In NAFLD, steatosis is generally centred around zone 3 hepatocytes, but this distribution can change as the disease progresses [56].

Several pathways are involved in the delicate balance of hepatic lipid homeostasis including uptake of circulating fatty acids, de novo lipogenesis, fatty acid oxidation, and VLDL secretion. Altered functionality of any of these pathways results in steatosis. Environmental factors such as diet and type 2 diabetes may also contribute to the development and severity of steatosis through insulin resistance as well as increased flux of fatty acids or their precursors to the liver.

Fatty acids used to synthesize hepatic triglycerides are derived from three sources. Approximately 15% of hepatic fat is derived from the diet, 26% is derived from de novo lipogenesis, and 59% is derived from adipose tissue lipolysis [65]. Therefore, the majority of hepatic TG is derived from fatty acids acquired from circulation. Hepatic lipid uptake is dependent on several fatty acid transporters (FATP and CD36) while their transport within the cell relies upon the availability of FABP1. All three of these transport proteins are upregulated in the early stages of NAFLD. This likely represents a compensatory mechanism to prevent hypertriglyceridemia as well as lipotoxicity within the liver. The transport proteins facilitate the incorporation of cytotoxic free fatty acids into metabolically inert triglycerides to prevent cellular injury. However, in later stages of the disease, this increased lipid flux could contribute to lipotoxicity and disease progression [64].

De novo lipogenesis is abnormally elevated in NAFLD patients and these patients fail to properly nutritionally control this process. This failure to regulate de novo lipogenesis is a central feature contributing to steatosis in NAFLD. NAFLD patients have higher expression of the master transcription factors that control de novo lipogenesis: SREBP1c and ChREBP [66]. This leads to the upregulation of downstream targets such as FAS and ACC and increased de novo lipogenesis of saturated fatty acids. The increased production of these cytotoxic fatty acids may contribute to lipotoxicity and disease progression.

One mechanism for disposal of excess fatty acids in the liver is oxidation. Fatty acid oxidation is increased in NAFLD patients as an adaptive mechanism to reduce the lipid burden in hepatocytes and the risk of lipotoxicity. This oxidation generally occurs in the mitochondria or peroxisomes to generate energy from the stored lipids. However, when there is a lipid overload within the hepatocyte, ω-oxidation by cytochromes is activated. ω-oxidation generates significantly more reactive oxygen species (ROS), oxidative stress, and toxic by-products. The increased fatty acid oxidation in NAFLD leads to high production of ROS that overwhelms the antioxidant capacity of the cell. This leads to oxidative stress and mitochondrial dysfunction which are precursors to lipotoxicity. So, although fatty acid oxidation is upregulated in NAFLD as a protective mechanism, this upregulation may ultimately contribute to disease progression [67, 68].

The major pathway for exporting fatty acids from the liver is VLDL secretion. These water-soluble particles are packed with triglycerides, cholesterol, phospholipids, and apolipoproteins before being secreted into circulation. In early stages of NAFLD, VLDL secretion is increased in an attempt to decrease the lipid burden of hepatocytes [69, 70]. However, as the disease progresses, ER stress results in the degradation of APOB-100 and reduced secretion of VLDLs [71, 72]. This diminished export results in steatosis, lipotoxicity, and disease progression.

1.3.2.2 Selective Insulin Resistance

Insulin resistance is a key mediator of hepatic fat accumulation, especially in NAFLD patients with metabolic syndrome. Insulin resistance occurs when insulin is unable to suppress gluconeogenesis but retains its ability to stimulate lipid production and secretion

Samantha G. Tilson - January 2021

[73-75]. The increased de novo production of glucose further stimulates de novo lipogenesis and metabolic dysfunction. The ultimate result of insulin resistance is hyperglycaemia and hypertriglyceridemia. Insulin exerts its effects on de novo lipogenesis by upregulating SREBP1c. SREBP1c then further exacerbates insulin resistance by increasing the accumulation of harmful lipid species such as diglycerides. Cytoplasmic diglycerides induce the expression and membrane translocation of protein kinase C which inhibits insulin receptor signalling [76-78]. The conversion of diglycerides to triglycerides prevents their deleterious effect on insulin signalling. Therefore, hepatocytes may be using steatosis as a protective mechanism against accumulation of diglycerides. An overview of this process can be found in Figure 1.12.
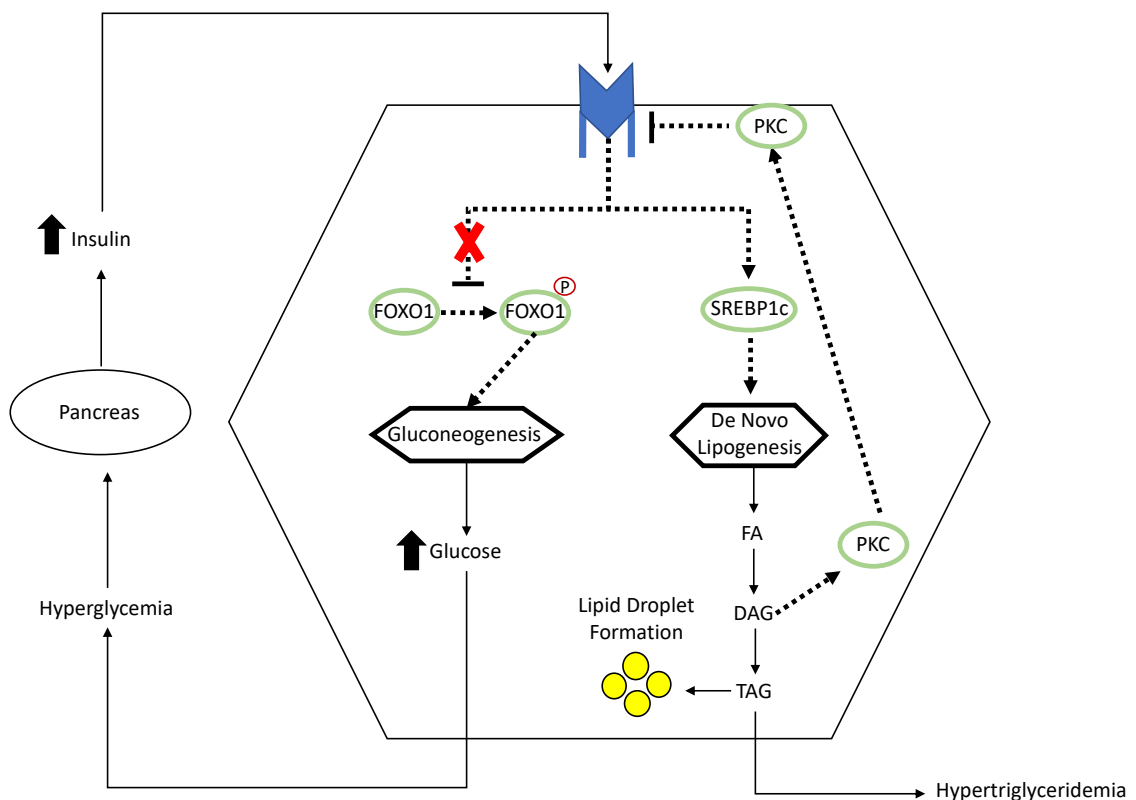
Figure 1.12 Selective insulin resistance.

Hepatic insulin resistance occurs when insulin loses the ability to suppress gluconeogenesis but retains its ability to induce de novo lipogenesis. In insulin resistant states, AKT signalling is disrupted so insulin loses the ability to suppress the activation of FOXO1 leading to constitutively active gluconeogenesis. This high hepatic glucose production leads to hyperglycaemia which in turn causes increased production of insulin in the pancreas and hyperinsulinemia. This hyperinsulinemia leads to upregulation of SREBP1c and downstream de novo lipogenesis genes which causes hepatic steatosis and hypertriglyceridemia. In

addition, high levels of de novo lipogenesis lead to accumulation of toxic intermediates such as diglycerides that further exacerbate insulin resistance through irregular protein kinase C signalling.

1.3.2.3 Lipotoxicity

Lipotoxicity is apoptosis caused by the harmful effects of high concentrations of lipids and lipid derivatives. Lipotoxicity may occur as a result of increased fatty acid oxidation and oxidative stress, alterations to cellular membrane and phospholipid composition, altered cellular cholesterol concentration, disturbed ceramide signalling, or direct toxicity of free fatty acids. These various mechanisms may trigger apoptosis via the intrinsic or extrinsic pathways. The intrinsic pathway of lipoapoptosis is triggered by intracellular ER stress and mitochondrial dysfunction whereas the extrinsic pathway is activated by the binding of death ligands FAS and TRAIL to receptors on the cell membrane. A schematic of the intrinsic lipotoxicity pathway can be found in Figure 1.13. Following activation of apoptosis, both the intrinsic and extrinsic pathways lead to mitochondrial permeabilization and activation of effector caspases. This lipoapoptosis is often accompanied by the release of proinflammatory cytokines which contribute to the progression from NAFLD to NASH [64, 79, 80].

The primary mediator of the intrinsic lipoapoptosis pathway is JNK activation. Activated JNK interacts with CHOP to upregulate the expression of PUMA [81]. PUMA and JNK then activate the pro-apoptotic protein BAX which is necessary for mitochondrial permeabilization. This permeabilization results in the release of cytochrome c and activation of the apoptosis effector caspases 3 and 7 [79]. JNK can be activated through several different cell stress pathways including ER stress, oxidative stress, and mitochondrial dysfunction. ER stress pathways are adaptive mechanisms to overcome stress stimuli such as accumulation of misfolded proteins or toxic lipid derivatives. However, under conditions of prolonged ER stress, apoptosis is triggered [80]. Oxidative stress occurs when the generation of ROS overcomes the antioxidant defences of the cell. This type of cell stress is caused in NASH by enhanced β- and ω-oxidation and CYP2E1 induction. This oxidative stress as well as increased mitochondrial permeability as a result of lipid accumulation results in mitochondrial dysfunction.
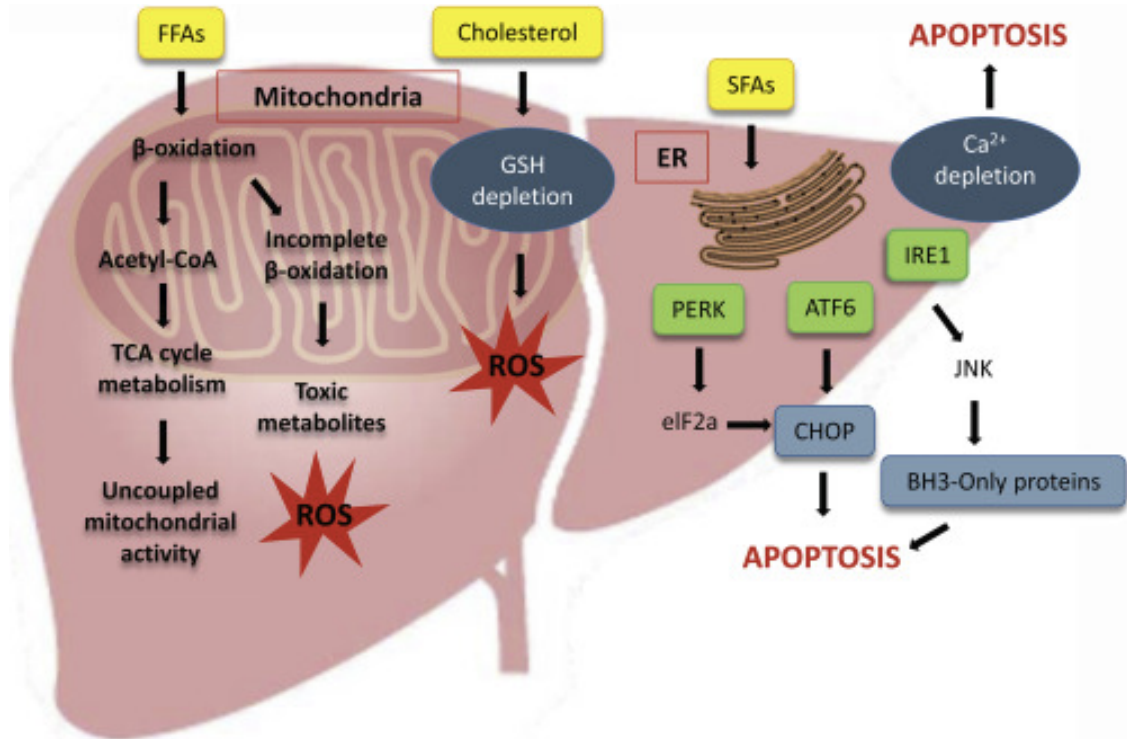
Figure 1.13 Mechanisms of lipotoxicity.

Lipids may induce toxicity through a variety of mechanisms including mitochondrial dysfunction, oxidative stress, and ER stress which ultimately lead to the activation of apoptosis. Excessive levels of free fatty acids may lead to mitochondrial dysfunction by uncoupling the TCA cycle from mitochondrial respiration or cause oxidative stress through incomplete β-oxidation. Cholesterol contributes to oxidative stress by depleting the hepatocyte's glutathione levels. SFAs cause ER stress by activating PERK, ATF6, and IRE1 which cause apoptosis through the activation of CHOP and JNK as well as disruption to calcium homeostasis. (Adapted from Svegliati-Baroni et al. Free Radical Biology and Medicine 2019) [82]

Several lipid species are considered lipotoxic including free fatty acids, especially saturated fatty acids, lysophosphatidyl choline (LPC), ceramides, and free cholesterol. Saturated fatty acids are increased in the hepatocytes of NAFLD patients due to increased fatty acid uptake and de novo lipogenesis. These fatty acids lead to lipotoxicity through several different mechanisms. Saturated fatty acids such as palmitic acid lead to the synthesis and accumulation of other cytotoxic lipid by-products such as ceramides and LPC [83]. Additionally, palmitic acid increases the synthesis of saturated phospholipids and glycerolipids [84-86]. When these saturated lipids are incorporated into plasma membranes, it triggers the activation of ER stress pathways which leads to JNK activation and ultimately lipoapoptosis. LPC is the phospholipid mediator of saturated fatty acid lipotoxicity in NASH. The production of this lipid causes a depletion of phosphatidyl

choline from membranes which results in loss of membrane integrity, extracellular release of lipotoxic lipids, inflammation, and apoptosis. Additionally, the accumulation of LPC independently induces ER stress [87].

Ceramides are cell membrane components that are biologically active in oxidative stress and inflammation. Accumulation of ceramides can induce ER stress by perturbing calcium homeostasis as well as promoting the release of ROS through its interaction with TNFα. The accumulation of ceramides is associated with insulin resistance and correlates with increased concentrations of proinflammatory cytokines [88-90].

Free cholesterol concentrations increase progressively from NAFLD to NASH. This cholesterol accumulates in the mitochondria and causes increased permeability and mitochondrial dysfunction. Additionally, cholesterol stimulates the activation of Kupffer cells and HSCs causing progression of disease toward inflammation and fibrosis [79, 81].

1.3.2.4 Inflammation and Fibrosis

Hepatocyte injury from lipotoxicity causes the release of damage associated molecular patterns (DAMPs) and proinflammatory cytokines that recruit and activate Kupffer cells and HSCs which cause disease progression from simple steatosis through inflammation and fibrosis. There are three inflammatory pathways in NAFLD: the JNK-AP-1 pathway, the NF-κB pathway, and the inflammasome pathway. Each of these pathways results in the production of proinflammatory cytokines such as IL1β, IL6, and TNFα. The accumulation of these proinflammatory cytokines leads to histological changes such as Kupffer cell activation, neutrophil chemotaxis, activation of HSCs, and production of Mallory bodies (inclusions of damaged cytoskeletal components in hepatocytes) [63]. The chronic inflammatory state of NASH increases the risk for fibrosis. The rate of fibrosis progression is higher in patients with obesity, diabetes, hypertension, and a high degree of inflammation found on biopsy [56]. Both inflammation and fibrosis have a degree of reversibility with weight loss and exercise [58]. However, some patients with severe inflammation and fibrosis will go on to develop cirrhosis and hepatic decompensation.

1.3.3 Genetics of NAFLD

Until recently, NAFLD was considered to merely be a consequence of the metabolic syndrome. However, recent studies have demonstrated a strong genetic component to the disease. The prevalence of hepatic steatosis varies among different ethnic groups. Hispanic Americans have the highest prevalence of hepatic steatosis at 45%, followed by European Americans at 33% and African Americans at 24% [91]. Severity of the disease also mirrors these ethnic differences with Hispanic Americans being at highest risk for severe NASH while African Americans appear to be protected from this disease. These ethnic differences are independent of adiposity, insulin resistance, and socioeconomic factors. Given these observations, the heritability of NAFLD is estimated to be approximately 52% while fibrosis is approximately 50% heritable among NAFLD patients. 78% of this heritability is shared between steatosis and fibrosis [92].

In order to identify candidate genes that are associated with the development and progression of NAFLD, several genome wide association studies (GWAS) were undertaken [91, 93, 94]. These GWAS identified single nucleotide polymorphisms (SNP), single base pair changes that are common in more than 1% of the population [95]. The first GWAS for NAFLD was performed by Romeo et al. on the Dallas Heart Study cohort. They identified a SNP (rs738409) in the gene coding for patatin-like phospholipase domain-containing protein 3 (*PNPLA3*). This SNP results in a missense mutation that replaces methionine for isoleucine at amino acid position 148. This variant has the strongest association with NAFLD and has been shown to increase risk for developing the full spectrum of NAFLD from simple steatosis through cirrhosis [91, 92, 95].

Several other SNPs have also been identified, though with lower effect sizes than *PNPLA3*, including *TM6SF2, MBOAT7*, and *GCKR*. The E167K (rs58542926) variant in TM6SF2 is the second most commonly associated variant with NAFLD. TM6SF2 is involved in the enrichment of triglycerides to APOB-100 for VLDL secretion and the E167K variant is loss of function. This results in retention of hepatic triglycerides which simultaneously increases a patient's risk for liver disease and reduces the risk of cardiovascular disease. The rs641738 SNP in *MBOAT7* increases the risk of NAFLD, inflammation, fibrosis, and hepatocellular carcinoma. The P446L (rs1260326) SNP in

GCKR is a loss of function mutation that results in constitutively active hepatic glucose uptake which stimulates de novo lipogenesis. Additionally, SNPs have been identified in several genes linked to insulin resistance (*ENPP1, IRS1*, and *TRIB1*). Some SNPs have also been found to be protective against NAFLD including *HSD17B13, LPIN1, UCP2, MnSOD*, and *MERTK* [92, 96]. A summary of SNPs that impact the NAFLD phenotype can be found in Figure 1.14. The effect size of the deleterious genetic variants is increased with increasing BMI indicating that genetic factors interact with environmental factors to determine disease severity [92, 97]. These genetic variants are also shared between other forms of chronic liver disease such as alcoholic fatty liver disease and viral hepatitis. This indicates that the steatosis in these diseases may share common pathogenic mechanisms.
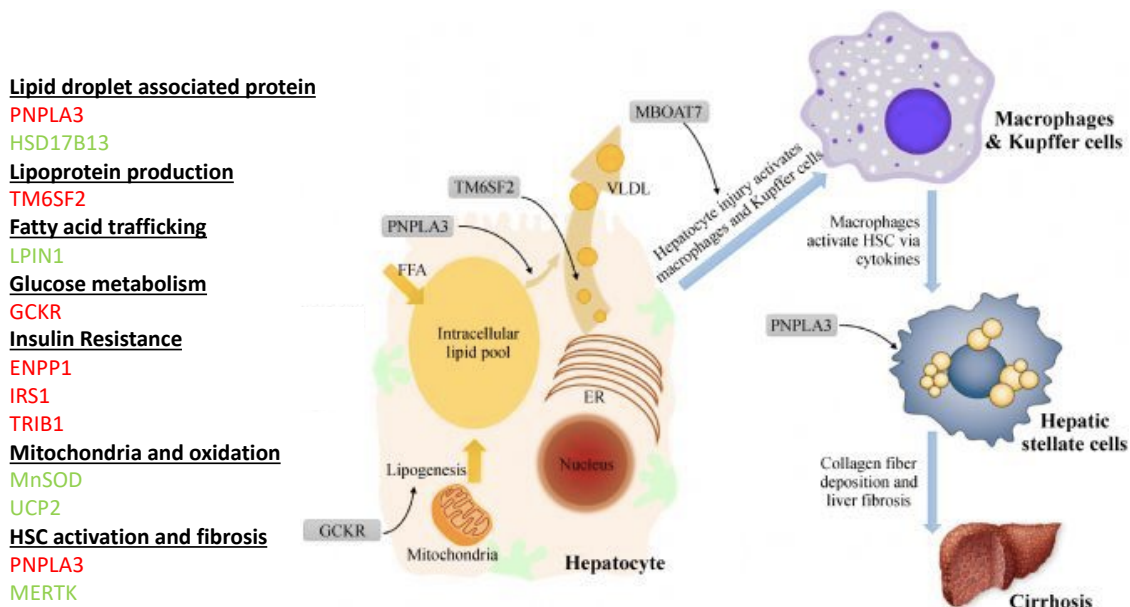
Figure 1.14 SNPs that modulate NAFLD phenotype.

List of genetic variants that have an impact on NAFLD progression and the cellular processes that they affect. Genes shown in red represent genetic variants that lead to worsening disease phenotype while genes shown in green represent genetic variants that are protective against NAFLD progression. The four genetic variants with the strongest effect sizes are *GCKR, MBOAT7, TM6SF2*, and *PNPLA3*. GCKR causes constitutively active de novo lipogenesis which causes hepatic steatosis. MBOAT7 plays a role in activation of the inflammatory response though its function remains to be fully elucidated. TM6SF2 causes hepatic steatosis by interfering with the secretion of VLDL. Finally, PNPLA3 plays a role in preventing the mobility of fatty acids from lipid droplets as well as activating HSCs to their pro-fibrotic state. (Figure adapted from Danford et al. The Journal of Biomedical Research 2018) [98]

In addition to the identified genetic risk of NAFLD, there is a growing appreciation for the role of epigenetics in disease progression. There is a general tendency toward demethylation of hepatic DNA in patients with NAFLD [99, 100]. The epigenetic regulation of genes involved in methylation, inflammation, and fibrogenesis is stage dependent. For example, *PPAR* genes are hypermethylated while fibrotic genes like *TGFβ* and *PDGFα* are hypomethylated in NAFLD patients with progressive disease [99]. This is evidence for the hypothesis that epigenetics may play a key role in the progression from NAFLD to NASH. Additionally, micro RNAs have also been linked to NAFLD pathogenesis. The most common micro RNA in the liver is miR-122 accounting for approximately 70% of hepatic mRNAs. It is downregulated in NAFLD patients leading to increased lipogenesis and activation of fibrosis through induction of HIF1α, MAPK1, and NIK [101-104]. However, it remains unclear if these epigenetic changes are causative agents in the disease progression from NAFLD to NASH or if they are mere by products of this progression.

Despite the growing appreciation for the genetic and epigenetic contributions to NAFLD pathogenesis the mechanistic role of the identified genetic variants remains largely obscure. This is due in large part to the lack of an appropriate pre-clinical model that recapitulates the full spectrum of human disease. This is especially true for the genetic variant with the highest effect size: *PNPLA3*. Despite the identification of the I148M variant over a decade ago, the mechanistic role of this protein both in healthy livers and NAFLD has yet to be fully elucidated. In the following section, I will summarize the state of knowledge about this gene and its contribution to NAFLD pathogenesis

1.4 PNPLA3 Gene

1.4.1 PNPLA3 Gene

PNPLA3, also known as adiponutrin or calcium-independent phospholipase A2ε, is a 52.8 kDa protein with 481 amino acids. The gene sequence is 9 exons long (2805 base pairs) and is located on the long arm of chromosome 22 at position 13.31 [105]. This protein is a member of the PNPLA family which has 9 members that demonstrate non-specific lipid acyl hydrolase activity [106-108]. The members of this family are soluble

proteins that are tightly associated with membranes and lipid droplets [109]. PNPLA3 shares the closest homology to PNPLA2 also known as ATGL [110].

1.4.1.1 Expression

In humans, PNPLA3 is expressed in several tissues including the liver, retina, skin, and adipose tissue. PNPLA3 is expressed at its highest levels in the liver and retina. Within the liver, PNPLA3 expression levels are higher in HSCs than hepatocytes [111-115].

1.4.1.2 Structure

The structure of PNPLA3 has yet to be fully resolved. However, several aspects of its structure have been elucidated using secondary structure prediction and homology modelling. The homology modelling of PNPLA3 resembles the crystal structure of the heartleaf horse nettle patatin. This includes a consensus serine lipase motif (Gly-X-Ser-X-Gly) that lies between a beta strand and an alpha helix. Like other members of the patatin family, PNPLA3 contains a catalytic dyad (Ser-Asp) as opposed to a catalytic triad that is common in most lipases. The patatin fold brings the catalytic serine (Ser47) in close proximity to Asp166 at the edge of a hydrophobic substrate binding groove to create the catalytic dyad [109, 116]. Controversy remains surrounding the presence of transmembrane domains within the protein. Studies of the secondary structure predicted 4 transmembrane domains that strongly associated with the ER and lipid droplets; however, tertiary structure modelling indicates that all alpha helices form a globular structure that does not span the membrane [105, 109]. Regardless, PNPLA3 is tightly associated with membranes, especially lipid droplet membranes.

1.4.1.3 Function

Considerable debate still remains surrounding the function of PNPLA3. *In vitro* studies have shown that the enzyme has triglyceride lipase activity against glycerolipids (triglycerides, diglycerides, and monoglycerides) as well as acylglycerol transacylase activity, and lysophosphatidic acid acyl transferase (LPAAT) activity in hepatocytes [117-121]. It has also been hypothesized that PNPLA3 may play a role in VLDL secretion [122]. Given this diverse range of functions, it has been difficult to ascertain the role of PNPLA3 in lipid metabolism. The LPAAT activity suggests a role in lipid anabolism while the lipase activity suggests a role in lipid catabolism. However, when the wild type

enzyme (derived from the reference allele) is overexpressed in mouse or human cells, there is no appreciable change in the intracellular triglyceride content [109, 119, 123]. Therefore, overexpression of PNPLA3 leads to neither a net increase nor decrease in triglyceride content which calls into question its role in either lipid anabolism or catabolism.

Recently, a new hypothesis has emerged to resolve these seemingly opposing results. Since PNPLA3 demonstrates both triglyceride lipase activity and acyl transferase activities, it has been hypothesized that PNPLA3 plays an active role in lipid droplet remodelling. Specifically, PNPLA3 seems to catalyse the transfer of polyunsaturated fatty acids (PUFA) from triglycerides to phospholipids [124-127]. In this way, PNPLA3 does not alter the total amount of triglycerides in the cell, rather the enzyme plays a role in maintaining the dynamic properties of lipid droplets by selective sorting of fatty acids between storage and membrane lipids. By sorting fatty acids between triglycerides and phospholipids, PNPLA3 plays a crucial role in determining the availability of fatty acids throughout the cell for signalling and other functional purposes. Additionally, by remodelling the phospholipid composition of the lipid droplet membrane, PNPLA3 may contribute to the recruitment of other lipid droplet proteins to further remodel the lipid content of the cell [128].

Additionally, in HSCs, PNPLA3 shows retinyl palmitate hydrolase activity which facilitates the release of retinol from HSCs [112, 129]. Due to its role in retinol homeostasis in these cells, it has been hypothesized that PNPLA3 contributes to HSC activation. This role has been confirmed by the upregulation of PNPLA3 expression in response to pro-fibrotic stimuli such as TGFβ and *in vitro* culture techniques. Overexpression of the reference PNPLA3 enzyme in HSCs revealed that upregulation of PNPLA3 leads to a reduced secretion of pro-fibrotic enzymes such as MMP2, TIMP1, and TIMP2 in a process that is mediated by retinoids [113, 114]. Therefore, PNPLA3-mediated retinol release is a protective mechanism against fibrosis.

1.4.1.4 Transcriptional and Nutritional Regulation
PNPLA3 is tightly controlled by nutritional status. In the fasted state, PNPLA3 expression is downregulated to near zero levels. Upon refeeding, especially a high carbohydrate diet,

PNPLA3 is quickly upregulated by insulin and glucose concentrations [111, 130, 131]. The pattern of induction of PNPLA3 is similar to proteins involved in glycolysis and lipogenesis. In humans, *PNPLA3* is transcriptionally controlled by SREBP1c whereas in mice, it is under the control of both SREBP1c and ChREBP [130]. Therefore, the upregulation of PNPLA3 in humans in response to a high carbohydrate diet is mediated entirely by insulin whereas the upregulation in mice is dependent on both glucose and insulin.

In addition to its transcriptional regulation, PNPLA3 is also regulated post-translationally by fatty acids. Fatty acids, especially saturated and monounsaturated fatty acids, inhibit the degradation of PNPLA3 protein. Addition of oleic acid extends the half-life of the enzyme by over four-fold. Since the products of SREBP1c-stimulated de novo lipogenesis are saturated and monounsaturated fatty acids, PNPLA3 appears to be controlled by a feed forward loop which allows for efficient upregulation of protein levels following feeding [111]. SREBP1c upregulates the transcription of *PNPLA3* and de novo lipogenesis genes, the products of this de novo lipogenesis then retard the degradation of PNPLA3 protein which causes the efficient postprandial rebound in PNPLA3 expression. This process is summarized in Figure 1.15.
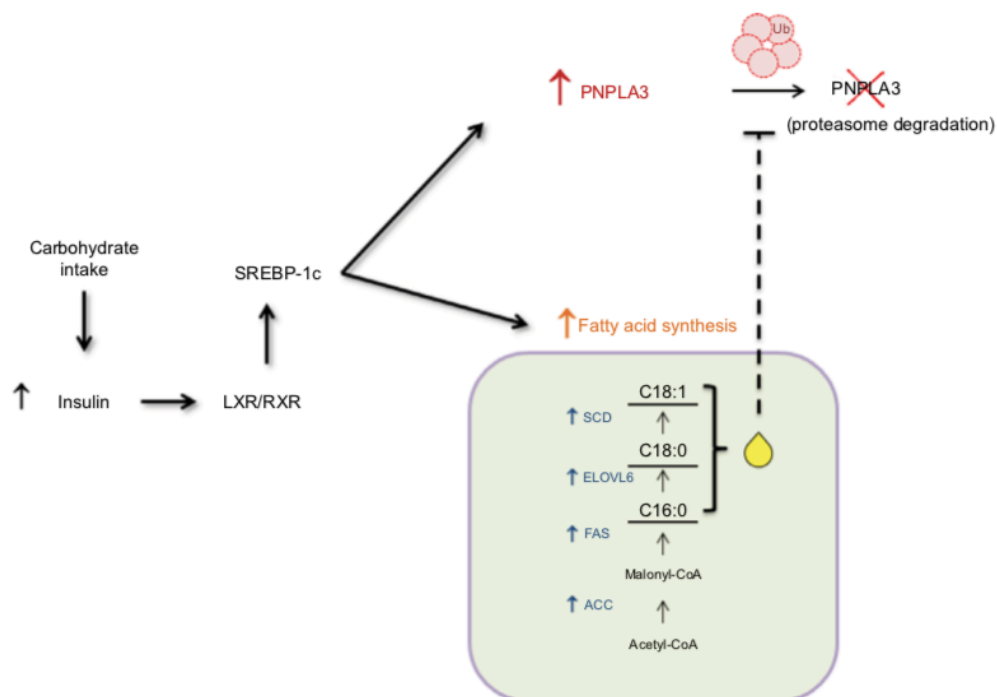
Figure 1.15 Nutritional regulation of PNPLA3.

Samantha G. Tilson - January 2021

Carbohydrate consumption leads to increased secretion of insulin from the pancreas. The binding of insulin to its receptor on the surface of hepatocytes leads to the activation of LXR/RXR which upregulates the expression of SREBP1c. SREBP1c then induces the expression of both PNPLA3 and de novo lipogenesis genes. The end-products of de novo lipogenesis, long chain fatty acids, prevent the proteasomal degradation of PNPLA3 thus creating a feed forward loop that allows for efficient post-prandial rebound of PNPLA3 expression. (Figure adapted from Bruschi et al. Hepatic Medicine: Evidence and Research 2017) [110]

PNPLA3 expression is not upregulated by a high fat diet despite induction of hepatic steatosis. In fact, the inclusion of lipids in a high carbohydrate diet fed to mice prevented the carbohydrate mediated induction of PNPLA3 in a dose dependent manner. The expression of PNPLA3 was especially reduced by the addition of PUFAs regardless of glucose concentration. These PUFAs reduce the nuclear abundance of SREBP1c by suppressing its gene transcription and enhancing mRNA decay [131]. The reduced expression of SREBP1c prevents the postprandial upregulation of *PNPLA3* transcription. Thus, PNPLA3 expression is not induced by all lipogenic stimuli and the presence of this enzyme is not necessary for the formation of steatosis.

In HSCs, *PNPLA3* is transcriptionally controlled by intracellular retinoid concentrations. When there is a high concentration of retinol, PNPLA3 expression is downregulated to allow the storage of retinyl palmitate in lipid droplets. When cells are deficient in retinol, PNPLA3 is upregulated to catalyse the hydrolysis of retinyl palmitate and the release of retinol. This reduction in lipid droplet size was also accomplished by supplementing HSCs with insulin indicating that the SREBP1c control of PNPLA3 is intact in HSCs [112]. In addition to nutritional regulation, PNPLA3 is also regulated by pro-fibrotic stimuli such as TGFβ [113].

1.4.2 I148M Variant
In 2009, Romeo et al. were the first to identify the rs738409 SNP in *PNPLA3* which causes a substitution of guanine for cytosine resulting in an amino acid change from isoleucine to methionine at position 148 (I148M) [91]. The authors found that the I148M variant (PNPLA3 risk allele) was strongly associated with high intrahepatic fat content ($P = 5.9 \times 10^{-10}$) as well as liver inflammation ($P = 3.7 \times 10^{-4}$). In this study, the minor allele frequency (MAF) varied between ethnic groups; Hispanics had the highest frequency of the risk allele (MAF = 0.49), followed by European Americans (MAF =

0.23), and African Americans had the lowest frequency of the risk allele (MAF = 0.17). The frequency of the risk allele in these ethnic populations mirrors the prevalence of hepatic steatosis in these groups which is estimated to be 45% in Hispanics, 33% in European Americans, and 24% in African Americans [54]. Since its initial discovery, the I148M variant has been associated with higher risk of developing the full spectrum of NAFLD from steatosis through inflammation, fibrosis, cirrhosis, and hepatocellular carcinoma [132-135]. Most recently, the correlation between the rs738409 SNP and NAFLD was confirmed in another GWAS (P = 1.70 x 10^{-20}). This study found that the rs738409 SNP was correlated with both pediatric (P = 9.92 x 10^{-6}) and adult (P = 9.73 x 10^{-15}) NAFLD as well as increased disease severity measured by the NAFLD Activity Score (P = 3.94 x 10^{-8}) [136]. The I148M variant confers increased risk of NAFLD development and progression independent of metabolic dysfunction [137, 138]. In addition to conferring increased risk of NAFLD, PNPLA3 I148M has been linked to worsening phenotypes in many other chronic liver diseases such as alcoholic fatty liver disease (ALD), hepatitis C virus (HCV), hepatitis B virus (HBV), hereditary hemochromatosis, and Wilson's Disease [139-144]. It should be noted that NAFLD and other chronic liver diseases are complex, multifactorial diseases that cannot be accounted for using standard inheritance models. Patients that are heterozygous for the PNPLA3 risk allele are at elevated risk of developing NAFLD and the addition of a second risk allele increases this probability in an additive manner [132]. However, the presence of this risk allele does not invariably lead to disease onset or progression; a patient's genetic predisposition interacts with environmental factors to determine disease severity.

1.4.2.1 Clinical Implications

The increased clinical risk of the I148M variant has been robustly confirmed in over 50 clinical studies [145]. Patients with the I148M variant tend to have 73% higher lipid content, 3.2 fold greater risk of high necro-inflammatory scores, 3.5 fold higher risk of NASH, 3.2 fold greater risk of developing fibrosis, and 3 fold increased risk of developing hepatocellular carcinoma [56, 57, 95, 133]. PNPLA3 I148M is associated with a higher risk of more severe NAFLD in several ethnicities (Hispanics, Argentinians, Germans, Italians, Japanese, Taiwanese, African Americans, Americans, Malaysians, etc.) [91, 137, 146-152]. The frequency of the risk allele in different ethnicities mirrors the relative

prevalence of NAFLD in those ancestry groups indicating that *PNPLA3* genotype is the major genetic determinant of NAFLD risk. The global prevalence of the PNPLA3 I148M variant is shown in Figure 1.16. It is estimated that the I148M variant accounts for 72% of the heritability of NAFLD [91]. The I148M variant is associated with increased fat content without altered glucose homeostasis or lipoprotein metabolism indicating that hepatic steatosis and insulin resistance are not inextricably linked. Additionally, the variant increases the risk of developing NASH and fibrosis independent of steatosis severity indicating that PNPLA3 I148M may differentially affect hepatic cell types involved in lipid metabolism and fibrosis [132, 145].

However, it is important to note that while NAFLD may develop in patients carrying the I148M variant in the absence of metabolic disease or insulin resistance, the presence of these comorbid conditions may exacerbate the NAFLD disease phenotype in these patients. It has been shown that the effect size of the PNPLA3 risk allele is increased in patients with a higher BMI and may be altered by dietary factors [97, 153, 154]. Additionally, studies show that insulin resistance in patients carrying the I148M variant amplifies the risk of developing steatosis. The reverse condition is also true. By improving the insulin response in patients carrying the risk allele, their genetic predisposition can be overcome leading to a reduction in steatosis and related disease factors [155]. Similarly, carriers of the I148M variant respond disproportionately well to lifestyle modification. The same amount of weight loss in these patients resulted in a larger reduction in liver fat and enzymes than non-carriers [156-158]. Thus, additional studies are needed to extricate the relationship between the I148M variant in PNPLA3 and the metabolic syndrome.
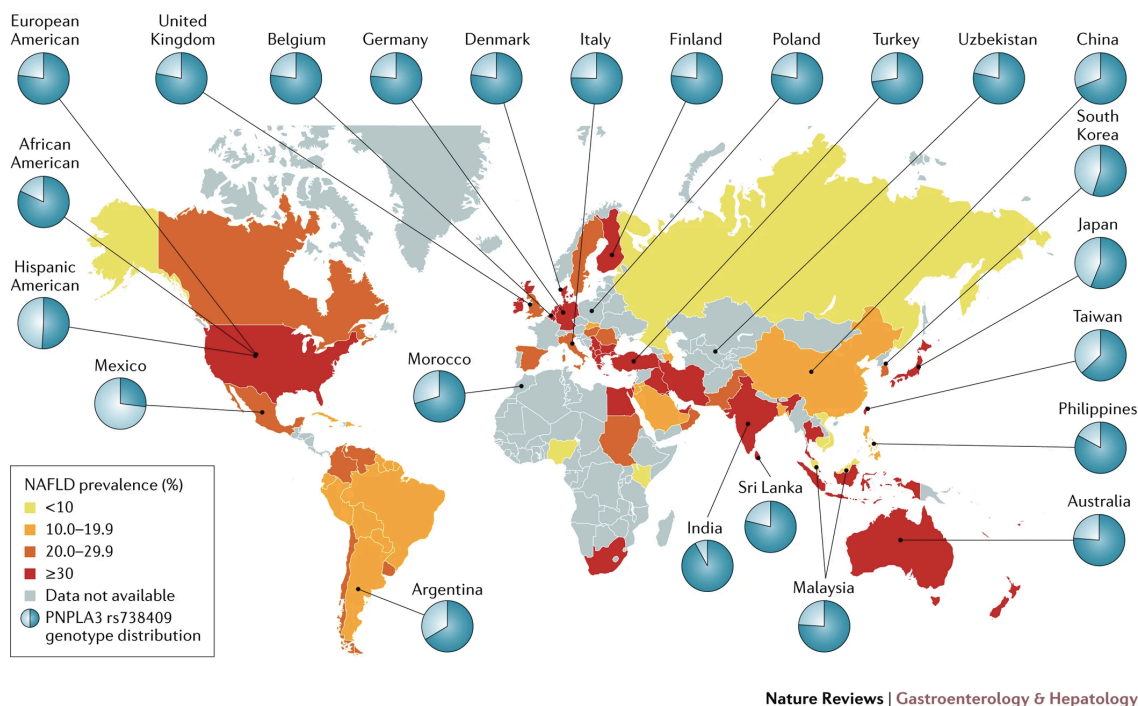
Nature Reviews | Gastroenterology & Hepatology

Figure 1.16 Global prevalence of PNPLA3 I148M variant.

The general prevalence of NAFLD in each country is denoted by the colour of the country with red being the most prevalent (>30%) and yellow being the least prevalent (<10%). The prevalence of the I148M variant in each country is denoted as the light blue section of the pie chart. The highest prevalence of the variant is found in Mexican and Hispanic American populations while the lowest prevalence is found in Indian populations. (Figure adapted from Younossi et al. Nature Reviews Gastroenterology & Hepatology 2017) [159]

In addition to NAFLD, PNPLA3 I148M has been linked to a worse prognosis in several other chronic liver diseases indicating a common genetic component to chronic liver diseases. In patients with ALD, the I148M variant has been linked to increased risk of cirrhosis, shorter time to hepatic decompensation, and increased mortality independent of alcohol consumption [139, 142]. In patients with chronic HCV infection, the I148M variant is associated with increased steatosis and fibrosis risk [141]. Whereas in chronic HBV infection, the variant is associated with worsening steatosis and inflammation but not necessarily fibrosis [144]. Additionally, I148M leads to increased steatosis and advanced fibrosis in patients with hereditary hemochromatosis [143]. This variant also confers and increased risk of developing hepatocellular carcinoma regardless of aetiology, although this correlation is stronger in metabolic liver diseases (NAFLD and ALD) than viral diseases [110].

1.4.2.2 Loss of Function or Gain of Function?

Given the fierce debate within the scientific community about the function of the PNPLA3 protein in a healthy liver, the role of the I148M variant in disease remains hotly contested. Several different hypotheses have been postulated to explain the pathogenesis of this variant, each of which is dependent on a different proposed function of the enzyme and the model system used.

There is general consensus that amino acid position 148 falls within the active site of the enzyme and that the substitution of isoleucine for the much bulkier methionine restricts the ability of substrates to bind to the active site of the enzyme [109]. This substitution results in an incomplete loss of lipase activity for both glycerolipids and retinyl esters [109, 112]. Additionally, the I148M variant was shown to decrease VLDL secretion *in vitro* [122]. Therefore, initial studies indicated that the pathogenic variant leads to reduced enzymatic activity which in turn causes steatosis through the reduced hydrolysis and secretion of triglycerides. However, when PNPLA3 was either overexpressed in or knocked-out of mouse livers, there was no effect on hepatic triglyceride content [160, 161]. This simultaneously argued against the role for PNPLA3 in triglyceride hydrolysis and the I148M variant being a loss of function variant.

Given that PNPLA3 also exhibited LPAAT activity *in vitro*, studies were undertaken to investigate the effect of the I148M variant on this function. The I148M variant appeared to increase the LPAAT activity *in vitro* which led to the hypothesis that the I148M variant is a gain of function that causes hepatic steatosis by upregulating triglyceride synthesis [118]. However, these results failed to be replicated in mouse models [162].

When the human or mouse I148M variant was overexpressed in mouse livers, a steatotic phenotype was observed. A similar phenotype was observed in mice overexpressing an obligate loss of function variant where the catalytic serine was mutated (S47A) [123, 163]. Both the I148M and S47A variants caused a significant increase in hepatic triglyceride content, in part, by evading degradation. These variants caused accumulation of I148M PNPLA3 proteins on the surface of lipid droplets which were hypothesized to reduce mobility of lipids from the lipid droplet [162]. Additionally, the high concentration of I148M PNPLA3 on the surface of lipid droplets has the potential to restrict access and

functionality of other lipid droplet proteins by displacing these proteins from the droplet surface or sequestering key cofactors [164]. Given that the full PNPLA3 knock-out produced no phenotype but the I148M variant had the same phenotype as the loss of function S47A variant, it was hypothesized that this variant may be a dominant-negative variant. Additional studies are needed to parse the pathogenic mechanism of action of the I148M variant.

1.4.2.3 Pathogenesis

Given the lack of clarity on the enzymatic activity and effect of the I148M variant, the full pathogenic mechanism of PNPLA3-induced NAFLD has yet to be elucidated. However, from what is known, the pathogenesis of the PNPLA3 I148M variant can be divided into two parts: steatosis and fibrogenesis.

The I148M variant disturbs the lipid homeostasis in hepatocytes by interfering with triglyceride hydrolysis, lipid anabolism, lipid droplet remodelling, and/or lipid mobility and secretion [109, 118, 122, 126, 162, 165]. The mechanism by which this variant interferes with lipid metabolism has yet to be fully elucidated. This metabolic dysfunction leads to triglyceride accumulation and steatosis which characterizes the early stages of NAFLD. It is unknown if this perturbed metabolism results in lipotoxicity or release of proinflammatory mediators that contribute to disease progression.

The progression from simple steatosis to NASH in patients with the I148M variant appears to be mediated by the effect of the variant on HSCs. PNPLA3 is known to play a role in activating HSCs to a profibrotic phenotype and the I148M variant appears to augment this function. HSCs with the I148M variant express significantly higher levels of the proinflammatory cytokines CCL2, CCL5, GM-CSF, CXCL8, CXCL1, and TGFβ which results in increased macrophage chemotaxis. HSCs harbouring the disease allele also have increased proliferation and migration rate, indicating that these cells are more prone to activation from their quiescent state [113, 114]. The role of the I148M variant in HSC activation is many-fold. In its role remodelling lipid droplets, PNPLA3 I148M alters the fatty acid composition of lipid droplets within HSCs leading to reduced retinol stores and increased concentrations of monounsaturated fatty acids and PUFAs. The

accumulation of these fatty acid species leads to the phosphorylation of JNK. Phospho-JNK then inactivates PPARγ. The inactivation of PPARγ has two consequences: decreased synthesis of oxysterols and increased expression of the pro-inflammatory and pro-fibrotic protein AP-1. The reduced synthesis of oxysterols prevents activation of LXR which causes the accumulation of free and total cholesterol in HSCs which further amplifies the pro-fibrotic response in these cells [113, 166]. An overview of this pathogenic mechanism can be found in Figure 1.17. The activation of HSCs by the I148M variant results in a profound inflammatory response as well as progressive fibrosis in patients carrying the disease allele.
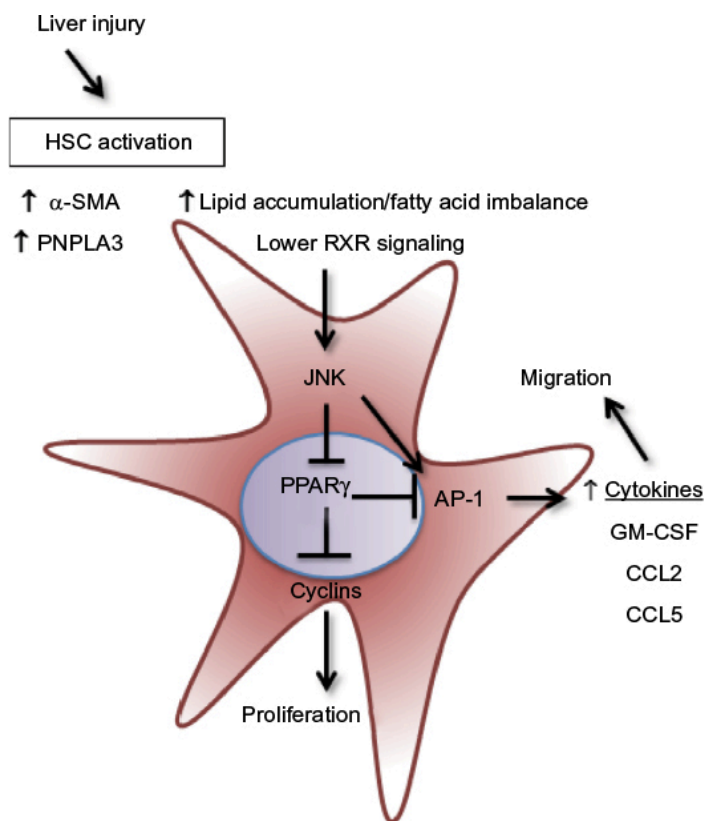
Figure 1.17 Pathogenic mechanism of I148M variant in HSCs.

Activation of HSCs leads to the upregulation of PNPLA3 and α-SMA. This leads to the accumulation of irregular fatty acid moieties such as MUFAs and sterols which lower RXR signalling and activate JNK. JNK simultaneously activates the pro-fibrotic gene AP-1 and inhibits PPARγ signalling. This has the downstream effect of increasing proliferation, migration, and secretion of pro-inflammatory cytokines such as GM-CSF, CCL2, and CCL5. (Figure adapted from Bruschi et al. Hepatic Medicine: Evidence and Research 2017) [110]

1.4.3 Limitations of Mouse Models for PNPLA3 Pathogenesis

The study of PNPLA3 and its effect on NAFLD has been extremely hindered by the lack of an appropriate human model of the disease. PNPLA3 function and mechanistic studies on the I148M variant have been performed almost exclusively in mice which has added a significant amount of confusion to the elucidation of this protein's function. We assert that the mouse is an inappropriate model to study PNPLA3 for three reasons. First, the homology between the two proteins is quite low. Second, the expression pattern of PNPLA3 is starkly different between the two species. Third, the transcriptional control of the protein differs between mice and humans indicating that the two proteins may function differently in each species.

1.4.3.1 Low homology

The protein sequence homology between mouse Pnpla3 and human PNPLA3 is approximately 68% [115]. This homology is extremely low when compared to the inter-species homology of other lipases which tend to have homologies greater than 90%. For reference, the sequence homology between human and mouse PNPLA3 is only approximately 20% higher than that sequence homology between human PNPLA3 and human PNPLA2 (46%). The human PNPLA3 protein is 97 amino acids longer than the mouse protein on the C-terminal end. In these additional 97 amino acids, the human protein contains 2 vesicle targeting motifs [167]. Given the low homology and large structural differences between human PNPLA3 and mouse Pnpla3, it cannot be guaranteed that the two proteins have identical functions in the two species. It is important to note that the majority of structural differences between human PNPLA3 and mouse Pnpla3 occur near the C-terminal end while the active site of the enzyme as well as exon 3 where the I148M variant resides remain largely conserved. The PNPLA3 protein sequence homology between humans and mice is shown in Figure 1.18.

```
PLPLA3_HUMAN  MYDAERGWSLSFAGCGFLGFYHVGATRCLSEHAPHLLRDARMLFGASAGALHCVGVLSGI  60
PLPLA3_MOUSE  MYDPERRWSLSFAGCGFLGFYHVGATLCLSERAPHLLRDARTFFGCSAGALHAVTFVCSL  60
              *** ** ****************** ****:********* :**.******.* .:..:

PLPLA3_HUMAN  PLEQTLQVLSDLVRKARSRNIGIFHPSFNLSKFLRQGLCKCLPANVHQLISGKIGISLTR  120
PLPLA3_MOUSE  PLGRIMEILMDLVRKARSRNIGTLHPFFNINKCIRDGLQESLPDNVHQVISGKVHISLTR  120
              ** : ::* ************ :** **:.* :*:** :.** ****:****: *****

PLPLA3_HUMAN  VSDGENVLVSDFRSKDEVVDALVCSCFIPFYSGLIPPSFRGVRYVDGGVSDNVPFIDAKT  180
PLPLA3_MOUSE  VSDGENVLVSEFHSKDEVVDALVCSCFIPLFSGLIPPSFRGERYVDGGVSDNVPVLDAKT  180
              **********:*:*****************::********** ************.:****

PLPLA3_HUMAN  TITVSPFYGEYDICPKVKSTNFLHVDITKLSLRLCTGNLYLLSRAFVPPDLKVLGEICLR  240
PLPLA3_MOUSE  TITVSPFYGEHDICPKVKSTNFFHVNITNLSLRLCTGNLQLLTRALFPSDVKVMGELCYQ  240
              **********:***********:**:**:********** **:**:.* *:**:**:* :

PLPLA3_HUMAN  GYLDAFRFLEEKGICNRPQPGLKSSSEGMDPEVAMPSWANMSLDSSPESAALAVRLEGDE  300
PLPLA3_MOUSE  GYLDAFRFLEENGICNGPQRSLSLSLVAP--E---------------AC----LENGKL  278
              ***********:**** ** .*. *  .   *              :.      :*.

PLPLA3_HUMAN  LLDHLRLSILPWDESILDTLSPRLATALSEEMKDKGGYMSKICNLLPIRIMSYVMLPCTL  360
PLPLA3_MOUSE  VGDKVPVSLCFTDENIWETLSPELSTALSEAIKDREGYLSKVCNLLPVRILSYIMLPCSL  338
              : *:: :*:    **.* :****.*:**** :**: **:**:*****:**:**:****:*

PLPLA3_HUMAN  PVESAIAIVQRLVTWLPDMPDDVLWLQWVTSQVFTRVLMCLLPASRSQMPVSSQQASPCT  420
PLPLA3_MOUSE  PVESAIAAVHRLVTWLPDIQDDIQWLQWATSQVCARMTMCLLPSTRSRASKDDHRMLKHG  398
              ******* *:*********: **: **** **** :*: *****::**:  ...::

PLPLA3_HUMAN  PEQDWPCWTPCSPKGCPAETKAEATPRSILRSSLNFFLGNKVPAGAEGLSTFPSFSLEKSL  481
PLPLA3_MOUSE  ---H--HPSPHKPQGNSAGL-----------------------------------------  413
                 .     :* .*:*  *
```

Figure 1.18 PNPLA3 protein sequence homology between humans and mice.

Amino acid sequence alignment of human and mouse PNPLA3 using UniProt. Identical amino acids are shown in white while non-identical amino acids are highlighted in yellow. Highly similar amino acids are denoted with (:), somewhat similar amino acids are denoted with (.), and entirely dissimilar amino acids are denoted with a blank space. The two proteins are extremely different, especially in the C-terminal end. The sequence homology as calculated by UniProt was only 57.6%. This low sequence homology makes the mouse protein an imperfect model to study the function and pathogenicity of human PNPLA3.

1.4.3.2 Differential tissue expression

The expression patterns of PNPLA3 are quite different between humans and mice. In humans, PNPLA3 is expressed at its highest levels in the retina, liver, skin, and adipose tissue. It is expressed at 10-fold higher levels in the liver than adipose tissue. Within the liver, it is expressed at significantly higher levels in HSCs than hepatocytes [111, 112]. Alternately, in mice, Pnpla3 is only expressed at appreciable levels in adipose tissue. The expression of Pnpla3 is 1000-fold higher in adipose tissue than liver tissue [105, 117, 119, 161]. Within mouse livers, Pnpla3 has the opposite expression pattern of humans with

higher expression levels in hepatocytes than HSCs [111]. This differential tissue expression is summarized in Figure 1.19.
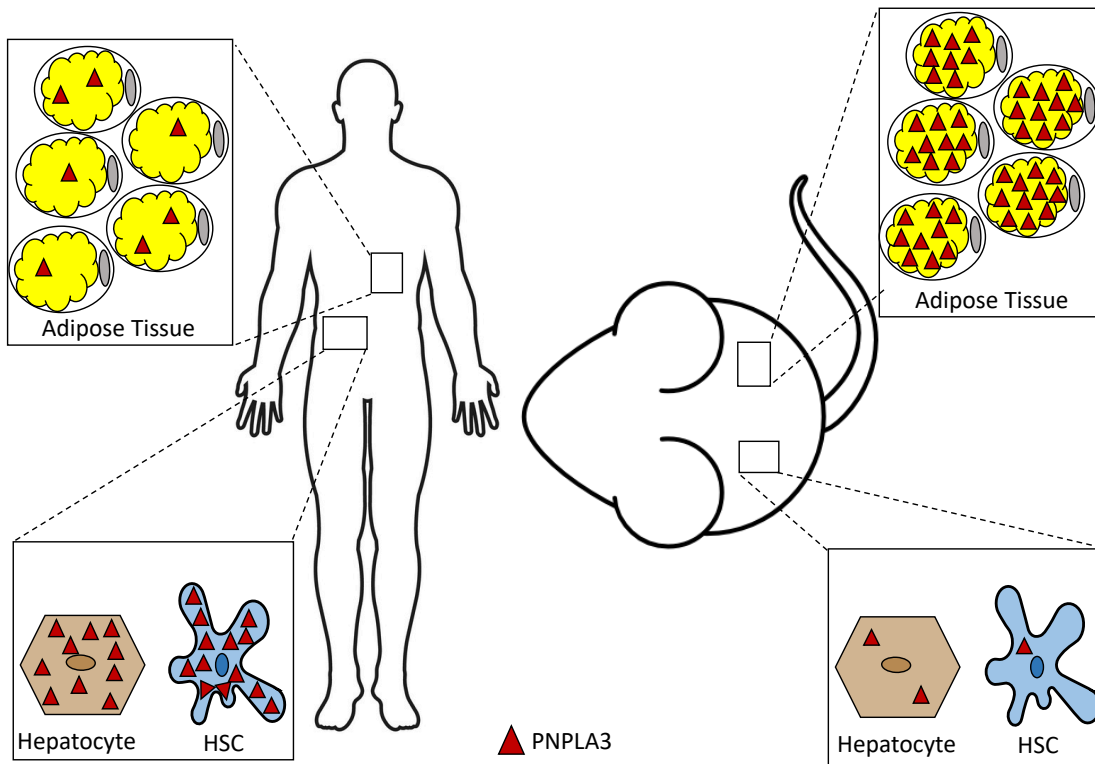
Figure 1.19 Differential tissue expression of PNPLA3 between humans and mice.

In humans, there is a very low expression of PNPLA3 in adipose tissue and the expression of PNPLA3 is 10-fold higher in the liver. In human livers, PNPLA3 is expressed at its highest levels in HSCs followed by hepatocytes. In mice, Pnpla3 is expressed at a 1000-fold higher level in adipose tissue than the liver. In mouse liver, Pnpla3 is expressed at higher levels in hepatocytes than HSCs. This differential tissue expression implies that PNPLA3 may play a different functional role in humans and mice.

Given that knocking-in the I148M variant to adipose tissue in mice has no effect while knocking the variant into liver tissue results in marked hepatic steatosis, it is clear that Pnpla3 exerts its detrimental effects on metabolism in the liver, not adipose tissue [123]. Since Pnpla3 is endogenously expressed at extremely low levels in mouse livers, it is unlikely that this protein plays the same role in hepatocyte metabolism that it does in human livers. Given this differential expression pattern, the murine model may not be the most appropriate model to study the function of PNPLA3 in hepatocytes and the I148M variant's effect on NAFLD disease progression.

1.4.3.3 Differential Transcriptional and Nutritional Control

The transcriptional regulation of human and mouse PNPLA3 is also different. In mice, *Pnpla3* transcription is regulated by both Srebp1c and Chrebp while human *PNPLA3* is controlled by SPREBP1c alone [111, 130]. This differential transcriptional control may indicate that human and mouse PNPLA3 possess different functions that require unique control mechanisms. To support this notion, mice are unable to nutritionally control the transcription of human *PNPLA3* when it is knocked-in to mouse livers [111].

The murine model of Pnpla3 differs in its sequence, expression pattern, and transcriptional control making it an inappropriate model to elucidate the mechanism of PNPLA3. In order to understand the function of human PNPLA3 in healthy livers as well as determine if the I148M variant is a loss of function or gain of function variant, a more appropriate model system is needed. This will require the development of a human model that endogenously expresses PNPLA3 at physiological levels. The human models that currently exist utilize either non-hepatic cell types or hepatocellular carcinoma cell lines. Often these cell lines do not express PNPLA3 at endogenous levels or at all so the protein must be overexpressed which may cause artefacts. I propose the use of human induced pluripotent stem cells and CRISPR/CAS9 to create a non-cancerous, hepatic, human model to study PNPLA3.

1.5 Human Induced Pluripotent Stem Cells

1.5.1 Introduction to hiPSCs

The first induced pluripotent stem cell (iPSC) line was established in 2006 by Takahashi and Yamanaka. They found that the overexpression of just four transcription factors, Oct4, Sox2, Klf4, and Cmyc, reprogrammed adult somatic cells to their pluripotent state. These iPSCs possessed a pluripotent morphology with a high nucleus to cytoplasm ratio and they formed densely packed colonies. Additionally, they had an extremely high proliferation rate as well as expression of key pluripotency genes [168, 169]. The original experiments were performed on mouse embryonic fibroblast cells but these results were quickly replicated in several somatic cell types from numerous species, indicating that this reprogramming method is robust and reproducible regardless of donor cell type or species [169-172].

The reprogrammed iPSCs possessed the key characteristics of pluripotent stem cells, namely they had self-renewal capacity and the ability to differentiate into cells from all three germ layers. These characteristics made them ideal candidates for disease modelling. These cells were capable of replicating indefinitely without undergoing growth arrest or developing chromosomal abnormalities. This overcomes one of the major limitations of using primary cells for disease modelling by providing a nearly endless supply of cells to work with. When given the correct stimuli, iPSCs can be differentiated into any cell type in the body by mimicking the embryonic development of a particular cell lineage. Therefore, multiple cells types can be generated in a single genetic background to model the complex intercellular interplay of many diseases. Human induced pluripotent stem cells (hiPSCs) can be generated quickly and economically from patient biopsies. This makes them suitable for use in personalized, regenerative medicine [168, 173]. They can be used to study both genetic diseases as well as the effect that genetic background has on a disease phenotype [174-176]. Importantly, these cells overcome the ethical questions that plagued their predecessors, human embryonic stem cells. Since hiPSCs are derived from consenting, adult, somatic cell samples rather than human embryos, they are not subject to the same funding restrictions and regulations [177, 178]. These hiPSCs can be generated from nearly any type of somatic cell which eliminates the need for expensive or invasive procedures to obtain a specific cell type. In addition, these cells can be grown *in vitro* for longer periods of time while primary adult cells often undergo spontaneous loss of function after being placed in culture [179]. All of these features make hiPSCs extremely useful tools for *in vitro* disease modelling. A summary of the potential uses for hiPSCs can be found in Figure 1.20.
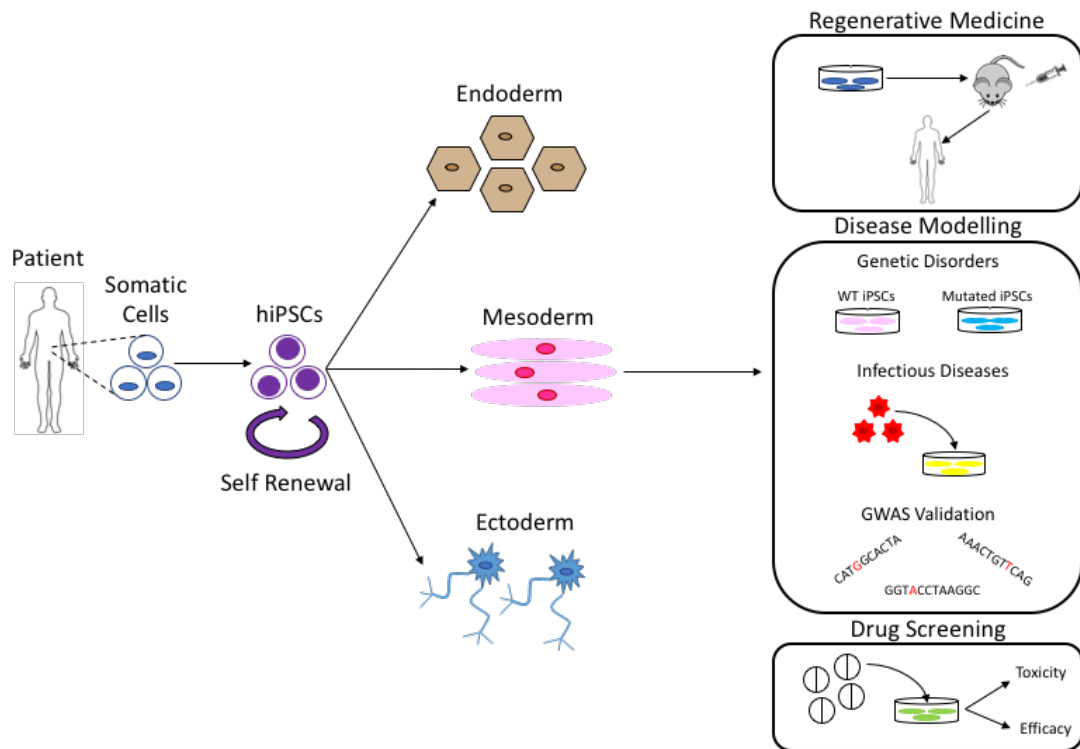
Figure 1.20 Potential of hiPSCs.

Schematic representation of the use of hiPSCs in disease modelling. hiPSCs are generated from human somatic cells. These hiPSCs possess the ability to self-renew indefinitely as well as differentiate into the three germ layers: endoderm, mesoderm, and ectoderm. Following differentiation, these cells can be used for regenerative medicine, modelling of genetic or infectious diseases, validation of GWAS hits, and drug screening. (Figure adapted from Yiangou et al. Cell Stem Cell 2018) [180]

1.5.2 Use of hiPSCs in Modelling Liver Disease

The *in vitro* modelling of liver diseases has been severely limited by the lack of access to primary tissue samples. These primary cells are extremely difficult to obtain due to the invasive nature of biopsies and the lack of availability of healthy donor tissues. Additionally, primary cells, especially hepatocytes, quickly lose their functional characteristics during *in vitro* culture. Primary human hepatocytes only remain differentiated and functional up to 7 days in 2D *in vitro* culture and thus have limited usefulness in disease modelling. Recent advancements in 3D culture as well as the development of co-culture systems have increased the viability and functionality of primary human hepatocyte cultures from days to weeks or months [181]. However, these hepatocytes do not replicate *in vitro* and therefore it remains difficult to obtain enough cells for disease modelling purposes.

To overcome these limitations, hiPSCs have been used to create hepatocyte-like cells (HLCs). hiPSCs undergo a directed differentiation that mimics the embryonic development of hepatocytes *in vivo* [182-184]. The cells are first differentiated into definitive endoderm before being specified into anterior foregut, hepatic endoderm, bipotent hepatoblasts, and ultimately "mature" HLCs. An overview of this differentiation protocol can be found in Figure 1.21. These HLCs express key markers of mature hepatocytes including HNF4α, HNF6, CEBPα, PROX1, GATA4, ALB, A1AT, and CYP3A4. Additionally, they demonstrate several hepatocellular functions including albumin secretion, glycogen synthesis, urea production, LDL uptake, and limited CYP activity [184]. However, these hepatocytes are not fully mature and more closely resemble the phenotype of foetal hepatocytes because, in addition to expressing adult hepatocyte markers such as ALB and A1AT, they express foetal markers such as alpha fetoprotein (AFP) and foetal isoforms of drug metabolizing enzymes such as CYP3A7 [185]. This is due in large part to the lack of knowledge regarding the signalling pathways and factors that control hepatocyte maturation *in vivo*. The liver is unique in that functional maturation occurs progressively after birth which makes these mechanisms difficult to study.

hiPSCs	Endoderm	Hepatic Specification	Hepatocyte Maturation
Day 0	Day 1-3	Day 4-8	Day 9-30
E8	CDM/PVA	RPMI	Hepatozyme
	100ng/mL Activin A 80ng/mL FGF2 10ng/mL BMP4 10μM LY2944002 3μM CHIR99021	50ng/mL Activin A	50ng/mL HGF 20ng/mL OSM
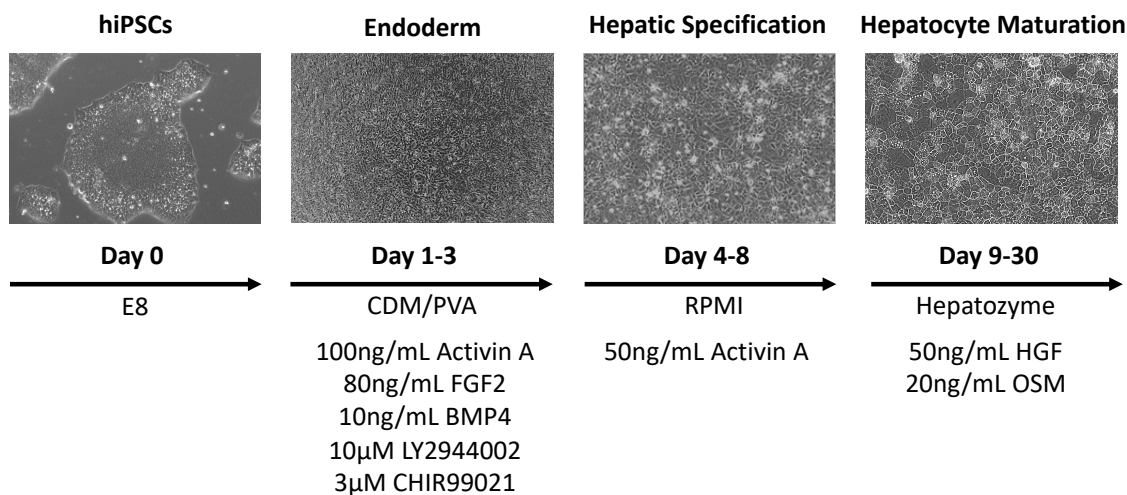
Figure 1.21 Overview of the HLC differentiation protocol.

The HLC differentiation protocol mimics embryonic development from pluripotency through endoderm, hepatic specification, and ultimately hepatocyte maturation. This process takes 30 days and is entirely chemically defined. hiPSCs are maintained in E8 medium until differentiation. Endoderm specification takes place over the first three days of differentiation and is achieved using a combination of several cytokines and small molecule inhibitors (Activin A, FGF2, BMP4, LY2944002, and CHIR99021). This is

followed by hepatic specification which is achieved over 4 days with Activin A signalling. Finally, hepatocyte maturation is the longest phase which takes place over days 9-30 with the supplementation of HGF and OSM.

Despite this functional immaturity, HLCs have been used extensively to model several different liver diseases *in vitro*. They have been used to study inherited metabolic disorders such as α-1-anti-trypsin deficiency, familial transthyretin amyloidosis, glycogen storage disease type 1a, Wilson's disease, and familial hypercholesterolemia [186-190]. In addition to using HLCs to gain mechanistic insights into disease progression, these cells have also been used as a platform for drug screening and validation for treatment of these diseases. In addition to these relatively rare genetic disease, HLCs have been used to model chronic liver diseases such as viral hepatitis and NAFLD [191-193]. Complex chronic liver diseases such as NAFLD are more difficult to model given the intricate inter-cellular interactions necessary for disease pathogenesis. In addition to HLCs, differentiation protocols have been published for several other hepatic cell types including cholangiocytes and HSCs [186, 194]. Several studies have been published that take advantage of the differentiation capacity of hiPSCs to create co-culture systems to model NAFLD *in vitro* and *in vivo [193, 195]*.

1.5.3 Limitations of hiPSCs

As with any model system, there are several limitations in using hiPSCs for disease modelling. These limitations include variation in differentiation capacity between hiPSC lines, limited maturity and functionality as well as short duration of culture for differentiated cells derived from hiPSCs. When hiPSCs are reprogrammed from somatic cells, these hiPSCs tend to retain an epigenetic signature that resembles their parental cell that is resistant to the reprogramming process. This "epigenetic memory" may cause cells to be more amenable to differentiation into the lineage of their parental cell type. This memory depends on several factors including the donor cell type, the reprogramming method used, and the culture conditions. Despite this major drawback, the epigenetic memory of hiPSCs is diluted by long term culture which reduces this inter-line variability [196]. Additionally, recent studies suggest that epigenetic memory is a rather minor contributor to inter-line variability. Primarily, this variation is driven by differences in genetic background between individuals [196-199]. Thus, it is important to include

isogenic hiPSC lines derived from multiple donors in experimental designs in order to overcome this limitation.

The major limitation of using hiPSCs for disease modelling, especially in metabolic organs such as the liver, is their functional immaturity. These cells more closely resemble foetal cells than adult cells due to the lack of understanding of hepatocellular development in the post-natal period [173, 180]. Given that these cells are not functionally mature, they may not faithfully recapitulate all aspects of adult disease. Additionally, hiPSC derived disease models have a short lifespan in culture [188]. Though the half-life of these cultures is significantly longer than primary cells, it remains on the order of weeks rather than months or years. This severely limits our ability to model chronic diseases that develop over long time scales. These limitations have been mitigated by recent developments in 3D culture technologies. Placing hiPSC derived hepatocytes into 3D culture has been shown to improve both their maturity and functionality by increasing the expression of mature hepatocyte markers as well as ameliorating their metabolic functionality [200, 201]. Not only do these 3D culture systems augment the maturity of hiPSC-derived hepatocytes, they also prolong their viability and functionality for much longer than traditional 2D culture systems. The use of 3D culture matrices also allows for co-cultures of hepatocytes and other non-parenchymal liver cells to be constructed. These co-culture systems have been shown to spontaneously self-assemble into physiologically relevant micro-structures with meaningful intercellular connections and communication [193, 195]. Co-culture of multiple hepatic cell types further improves metabolic functionality and extends culture time [202]. Unpublished data from our lab shows that co-culture of parenchymal and non-parenchymal liver cells in collagen matrices can maintain viability and functionality for up to 1 year of *in vitro* culture (data not shown). Similar advancements have been made in 3D and co-culture technologies for primary liver cells as well. Primary human hepatocytes cultured in 3D with or without other non-parenchymal liver cells can be cultured for long periods of time and faithfully recapitulate hepatic functions *in vitro [203-205]*. However, these primary cells remain extremely difficult to obtain in high quantities making them difficult to use for disease modelling and toxicity assay purposes. In addition, it is difficult to obtain several hepatic cell types from the same donor which may introduce unnecessary error into disease models utilizing

the co-culture of primary hepatic cells. Therefore, despite these minor limitations, hiPSCs remain the most ideal candidate to model complex diseases *in vitro*.

1.6 CRISPR/CAS9

CRISPR/CAS9 is an affordable, efficient, and simple method to achieve targeted genome editing. Originally described as an innate immune system in bacteria, this technology has since been utilized to precisely edit genomes in a variety of cell types and species. Unlike its predecessors zinc finger-nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs), CRISPR/CAS9 relies on simple Watson and Crick base pairing rules rather than complex protein engineering [206, 207]. The CAS9 protein is guided to a specific point in the DNA sequence of a cell by a single guide RNA (sgRNA) where it causes a double stranded break in the DNA and activates one of two DNA repair pathways.

1.6.1 History of CRISPR/CAS9

CRISPR stands for clustered regularly interspaced palindromic repeats named after a series of repeated DNA sequences interspaced with short sequences derived from plasmid or viral origins within the genome of bacteria. These CRISPR regions serve as a type of memory of past invasions. Bacteria cleave a short sequence of the invading DNA and incorporate it into the CRISPR array. These sequences can then be transcribed into RNAs used to guide the cleavage of foreign DNA. The bacteria uses these antisense RNAs in combination with CRISPR associated (CAS) proteins as an adaptive immune system to interfere with virus proliferation [207]. The role of CRISPR/CAS9 in bacterial adaptive immunity is summarized in Figure 1.22. Given the simplicity and specificity of this system, researchers exploited the CRISPR/CAS9 system for the purposes of genome editing.
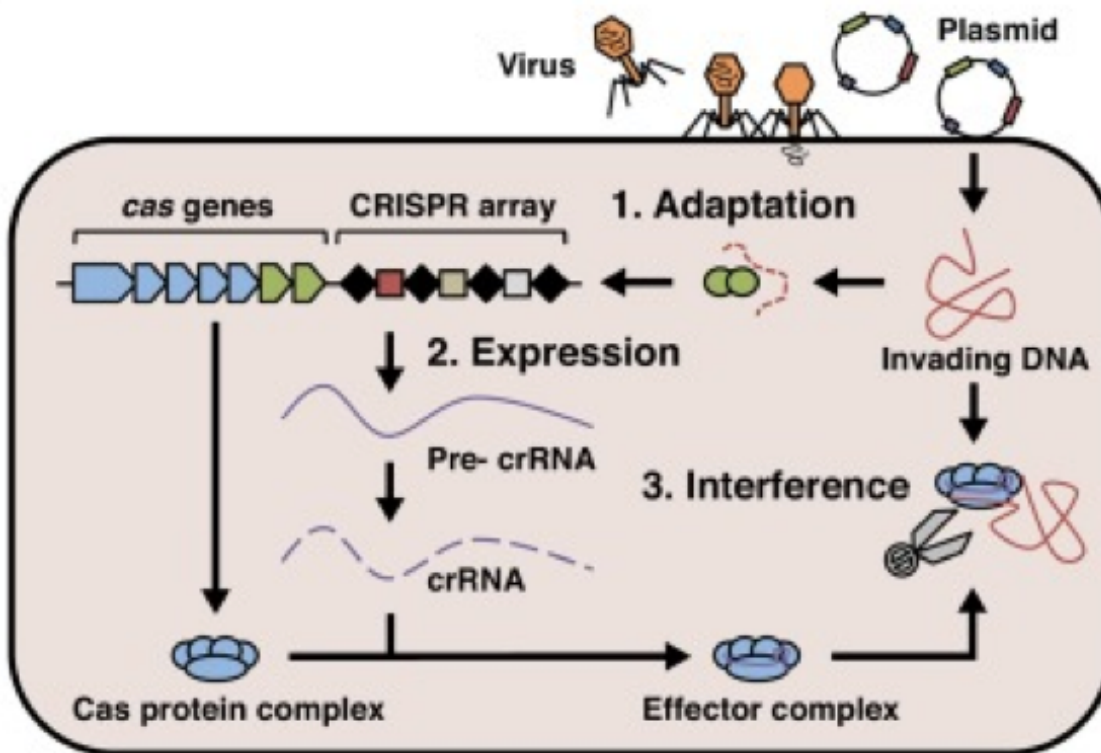
Figure 1.22 CRISPR/CAS9 adaptive immune system.

The CRISPR/CAS9 adaptive immunity can be divided into three steps. The first step is adaptation which involves incorporating a portion of invading DNA into a CRISPR array. In step two, expression, CRISPR arrays are transcribed and processed into crRNAs. In step 3, the crRNA in complex with CAS proteins binds to and degrades the invading DNA. (Figure adapted from Patterson et al. Current Opinion in Microbiology 2017) [208]

1.6.2 Mechanism of CRISPR/CAS9

CRISPR/CAS9 requires two functional partners: sgRNA and the CAS9 protein. The CAS9 protein is derived from Streptococcus pyogenes but can efficiently cleave DNA from any species. The sgRNA is designed to be complementary to a 20-nucleotide sequence in the genome. The target sequence must be adjacent to a protospacer adjacent motif (PAM). This PAM is usually NGG and is required for target site recognition by the CAS9 protein. Therefore, in theory, any DNA sequence that contains N_{20}-NGG could be used as a target site. However, additional aspects of the DNA landscape must be considered in order to design a highly specific and efficient sgRNA such as the repetitiveness of the DNA region of interest as well as the distance of the PAM site from the editing site. Following binding of the sgRNA to the target site, CAS9 generates a

double stranded break in the DNA. This double stranded break can be repaired by two mechanisms: non-homologous end joining (NHEJ) or homology directed repair (HDR). When a template sequence is not provided, the double stranded break is repaired by NHEJ which is error-prone and often results in short indels. These indels can cause non-sense mutations that generate a knock-out of a certain gene. In order to specifically edit a gene locus, a homologous repair template such as an ssODN can be provided to direct the cell toward HDR. Using HDR, the cell will replace the target site in a template-dependent manner [209, 210]. In this way, CRISPR/CAS9 can be used to either knock-out a gene of interest or knock-in a specific change to the DNA sequence. The mechanism of CRISPR/CAS9 genome editing is illustrated in Figure 1.23. CRISPR/CAS9 unlocked the potential for genome editing to be used in conjunction with hiPSCs for modelling genetic diseases *in vitro*.
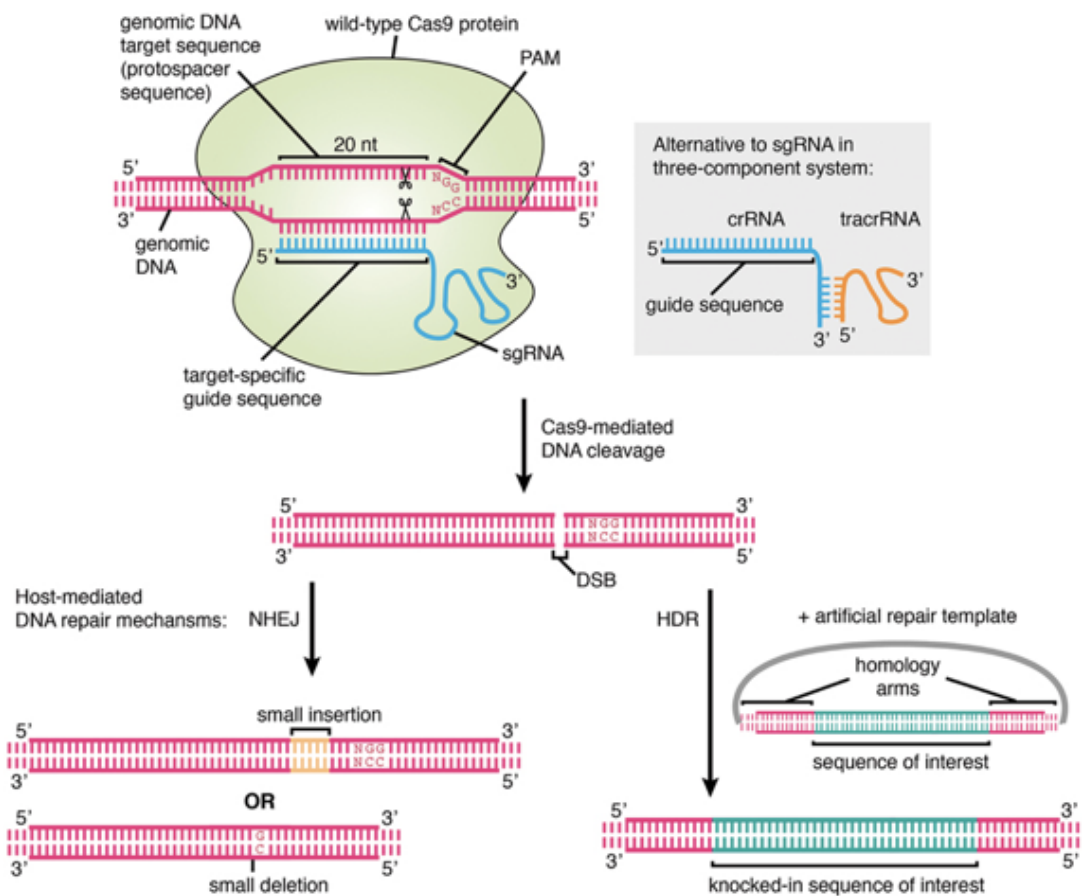
Figure 1.23 CRISPR/CAS9 genome editing mechanism.

The gRNA is designed against a 20 base pair genomic DNA target sequence that is adjacent to a PAM. This gRNA guides the CAS9 protein to the sequence of interest where it creates a double stranded break in the DNA. This double stranded break can then be repaired through NHEJ which is error prone and often results

in short indels or HDR which uses an exogenously supplied homologous sequence to specifically change the genomic sequence in a predictable manner. (Figure adapted from Agrotis et al. Frontiers in Genetics 2015) [211]

1.6.3 Limitations of CRISPR/CAS9

There are several limitations to the use of CRISPR/CAS9. The editing efficiency of CRISPR/CAS9 is inconsistent and often not very high. This editing efficiency is dependent on several factors including cell type, culture conditions, and sgRNA efficiency. This low editing efficiency can make the process of genome editing incredibly time consuming and expensive. Not all sgRNAs are equally efficient and it is difficult to determine the efficiency of any given gRNA without testing it. Several online search tools have been designed in an attempt to overcome this pitfall. These design tools are used to predict the activity of an sgRNA design *in silico*. However, despite the implementation of these design tools, several sgRNAs must still be tested experimentally to identify the sequences with the highest activity [206]. These experimental tests can be costly and time consuming.

The most concerning limitation of CRISPR/CAS9 is its off-target effects. Despite the specificity that is inherent to the design of sgRNAs, these sgRNAs may still hybridize with DNA sequences that have base-pair mismatches. This occasionally causes cleavage at off target sites. These off-target effects are often difficult to identify because they may not occur in coding regions and whole genome sequencing is not feasible for all cell lines that are generated [207, 209]. In order to mitigate these effects, web-based tools have been designed to identify potential off target effects by comparing sequence compositions. In addition, it is common practice to use "control" cell lines that have been exposed to the CRISPR/CAS9 gene editing process without editing at the site of interest. These cell lines are called untargeted control lines and would likely be subject to similar off target effects as their genetically edited brethren.

Despite these limitations, CRISPR/CAS9 remains the best genome editing tool available to biological researchers today. Several labs are working to improve the efficiency and specificity of CRISPR/CAS9 genome editing. Recent advances in base editing technology allows for the efficient and precise introduction of point mutations without

initiating a double stranded break in the DNA. These base editors directly convert one nucleotide into another and thus can be used in non-dividing cells without inducing off target effects [212]. The potential for the CRISPR/CAS9 editing technology and its derivatives for use in disease modelling as well as cell therapy and personalized medicine will only improve in the years to come.

1.7 Open Questions and Project Aims

NAFLD/NASH is one of the largest public health burdens facing the developed world. The genetic underpinnings of this disease have only recently been elucidated. The major genetic contributor to NAFLD/NASH is the I148M variant in PNPLA3. Despite over a decade of dedicated research into this enzyme and its pathogenic functions, very little is known about PNPLA3. PNPLA3 appears to play a key role in lipid homeostasis in several hepatic cell types but its function has yet to be fully clarified. Additionally, the I148M variant is a major determinant in NALFD onset and progression but the pathogenic mechanism of this variant remains incompletely understood. The study of PNPLA3 and its effect on NAFLD has been severely hindered by the lack of a human model of this disease. In order to overcome this limitation, we chose to take advantage of the unique properties of hiPSCs and CRISPR/CAS9 to develop a human model of PNPLA3-induced NAFLD.

The specific aims of this PhD project are:

1) Use CRISPR/CAS9 to create hiPSC lines with either a full knock-out of the PNPLA3 protein or with the I148M variant knocked in

2) Determine if the I148M variant represents a loss of function or gain of function variant in human cells

3) Investigate the effect of PNPLA3 knock-out and the I148M variant on lipid metabolism and gene transcription

4) Gain mechanistic insight into the pathophysiology of the PNPLA3 I148M variant in NAFLD

2 MATERIALS AND METHODS

2.1 Genome Editing

2.1.1 Design

The guide RNA (sgRNA) sequences were designed using the Zhang Lab MIT CRISPR Design (http://crispr.mit.edu) and E-Crisp (http://www.e-crisp.org/E-CRISP/index.html) online design tools. The sgRNAs were designed to target exon 3 where the I148M variant resides. sgRNAs were chosen to maximize specificity and minimize off-target effects. The first set of sgRNAs were designed to target hiPSC lines that were homozygous for the risk allele. The second set of the sgRNAs were designed to target hiPSC lines that were heterozygous for the risk allele. Their sequences are as follows: sgRNA1 Forward: CACCGCTGTAGAAGGGGATGAAGC, sgRNA1 Reverse: AAACGCTTCATCCCCT TCTACAGC, sgRNA2 Forward: CACCGGGATAAGGCCACTGTAGAA, sgRNA2 Reverse: AAACTTCTACAGTGGCCTTATCCC The single-stranded donor oligonucleotides (ssODN) were then designed to contain the variant of interest as well as three silent mutations to prevent re-editing. The sequence of the ssODN in the sense direction was tctctcctttgctttcacagGCTTTGGTATGTTCTTGTTTCATGCCCTTATATA GTGGCCTTATCCCTCCTTCCTTCAGAGGCGTGgtaa. The sequence of the ssODN in the anti-sense direction was ttacCACGCCTCTGAAGGAAGGAGGGATAAGGCCAC TATATAAGGGCATGAAACAAGAACATACCAAAGCctgtgaaagcaaaggagaga. The sense ssODN was used with sgRNA1 which targeted the sense strand of the DNA while the anti-sense ssODN was used with sgRNA2 which targeted the anti-sense strand of the DNA.

2.1.2 Molecular Cloning

The sgRNA was then cloned into the pSpCas9(BB)-2A-Puro (PX459) (Addgene Plasmid #48139) expression vector [213]. To start, the forward and reverse oligos were annealed together to create a double stranded molecule with BbsI overhangs. A 10 µL reaction containing 100 µM of sgRNA top and bottom oligos and nuclease free water was annealed at 95°C for 5 minutes followed by a ramp down to 25°C at 5°C/min. The PX459 plasmid was then digested with BbsI. The 20 µL digestion reaction contained 2 µL of 10X digestion buffer (Thermo Fisher), 1 µL of the PX459 plasmid (stock concentration: 1

µg/µL), 1 µL of BbsI (Thermo Fisher), and 16 µL of nuclease free water. The digestion reaction was incubated at 37°C for 1 hour. The annealed oligonucleotides were then ligated into the expression vector. The ligation components (1.5 µL T4 DNA ligase (Promega), 2.5 µL 10X T4 ligase buffer (Promega), and 1 µL annealed sgRNA oligos) were added directly to the digest and incubated at 37°C for 1 hour. Following ligation, 2 µL of the ligation reaction was transformed into 25 µL of Alpha select Gold competent cells (Bioline) according to the manufacturers protocol. The transformed cells were plated onto LB Ampicillin agar plates (10 g LB powder (Formedium), 7.5 g agar (Sigma), 400 mL distilled water, and 40 mg Ampicillin (Sigma) and incubated at 37°C overnight. The following day, 2-4 colonies were picked and miniprepped according to the manufacturer's instructions. The miniprepped DNA was sent for sequencing using the U6-Forward primer, GAGGGCCTATTTCCCATGATCC. The sequenced plasmid was then purified from 35-50 mL of LB broth (10 g LB and 400 mL distilled water) culture using the high yield protocol of the Plasmid Plus Midiprep Kit (Qiagen). The purified plasmid was eluted using 50 µL of sterile water for embryo transfer and the DNA concentration was measured.

2.1.3 Nucleofection

The sgRNA expression construct was then nucleofected into hiPSCs using the Lonza P3 Primary Cell 4D-Nucleofector X Kit. 24 hours prior to nucleofection, cell medium was changed to TeSR E8 (Stem Cell Technologies) without penicillin/streptomycin (Life Technologies) supplemented with 10 µM Y-27632 (Stem Cell Technologies). On the day of nucleofection, cells were dissociated into small colonies of 3-4 cells using Accutase (Thermo Fisher) for 5 minutes at 37°C and counted. Two million cells were then resuspended in P3 nucleofection solution with 10 µg plasmid DNA and 200 pmol ssODN before being transferred to a cuvette. The cuvette was then placed in the nucleofector and pulsed using the CA 137 program. Cells were allowed to recover for 5 minutes at 37°C before the addition of culture medium supplemented with Y-27632. The cells were allowed to recover for an additional 5 minutes before being plated onto two vitronectin coated p100 plates.

2.1.4 Selection

The cells were incubated for 24 hours before starting puromycin selection. 1 μg/mL of Puromycin (Sigma) was added to the cells and selection was allowed to continue for 48 hours. Following this selection, the cells were allowed to recover, and clonal colonies were picked approximately 14-21 days after selection.

2.1.5 Genotyping

DNA from the selected colonies was extracted using the Quickextract kit (CamBio). Cells were resuspended in 50 μL of Quickextract solution followed by vigorous pipetting. The solution was incubated at 65°C for 6 minutes, vortexed thoroughly, and incubated for an additional 2 minutes at 98°C. This DNA was used for PCR amplification. The PCR master mix included 1-4 μL genomic DNA, 2 μL 10X Titanium Taq buffer (Clontech), 0.4 μL of 10 mM dNTP's (Promega), 0.4 μL of Titanium Taq Polymerase (Clontech), 1 μL each of forward (ggagcaaggagaggaagttg) and reverse (cgggtagcctggaaatagg) genotyping primers (stock concentration: 10 μM), and nuclease-free water for a total volume of 20 μL. The mixture was heated for 2 minutes at 95°C before cycling 35 times (98°C for 30 seconds, melting temperature for 30 seconds, 68°C for 1 minute) and finally heated for 10 minutes at 68°C. The PCR product was checked using an agarose gel before being purified with the Exosap-it reagent (Affymetrix). 2 μL of Exosap-it reagent was added to 5 μL of PCR product and incubated at 37°C for 15 minutes followed by 80°C for 15 minutes. The samples were then sent to GATC Biotech for genotyping.

2.2 Cell Culture

2.2.1 hiPSC Culture

The hiPSC lines, FSPS13B and A1ATDR/R, were obtained from the hiPSC Core Facility, Cambridge Biomedical Research Centre [185, 190, 197]. Both hiPSC lines were derived from male donors. The FSPS13B line was homozygous for the reference PNPLA3 allele while the A1ATDR/R line was heterozygous for the risk PNPLA3 allele. Neither line harboured the risk allele for other genetic variants that play a role in NAFLD onset and progression including TM6SF2 and GCKR. These lines have been extensively characterized and do not contain any know genetic abnormalities. Both lines have been

shown to have a high propensity for differentiating into hepatocytes. The cells were maintained on vitronectin (Stem Cell Technologies) coated plates in Essential 8 medium (Gibco). A list of media and supplementation components for cell culture can be found in Table 2.1. Cells were maintained at 37°C and 5% CO_2 and split every 5-7 days using 0.5 mM EDTA (Quality Biological Inc.).

2.2.2 Differentiation Protocol

Tissue culture plastic was coated with 0.1% porcine gelatin (Sigma Aldrich) in water for embryo transfer (Sigma Aldrich) for 1 hour. The coating was then overlaid with MEF medium for 24 hours prior to cell seeding. For differentiation, hiPSCs were split into single cell suspension using accutase (Thermo Fisher) for 5 minutes at 37°C. The cells were then seeded at a concentration of 50,000 cells/cm^2 in Essential 8 medium supplemented with 10 µM Y-27632 (Stem Cell Technologies). Cells were allowed to recover for 48 hours before beginning the differentiation protocol. On day 1 of differentiation, cells were differentiated in CDM/PVA supplemented with 100 ng/mL Activin (Dr. Marko Hyvönen, University of Cambridge, 80 ng/mL FGF2 (Dr. Marko Hyvönen, University of Cambridge), 10 ng/mL BMP4 (R&D), 10 µM LY294002 (Promega), and 3 µM CHIR99021 (Tocris). On day 2, the CHIR99021 was removed. On day 3, the medium was replaced with RPMI supplemented with 100 ng/mL Activin, and 80 ng/mL FGF2. For days 4-8 FGF2 supplement was removed and the medium was replaced every 24 hours. Beginning on day 9, the medium was changed to Hepatozyme supplemented with 20 ng/mL Oncostatin M (Stem Cell Technologies), and 50 ng/mL HGF (Peprotech). For days 9-30 the medium was changed every 48 hours. For the duration of the differentiation protocol, cells were cultured at 37˚C, 5% CO_2, and 5% O_2. This differentiation protocol was adapted from Hannan et.al [184].

Table 2.1 Cell Culture Media Composition

Medium	Use	Components	Concentration	Supplier
E8	hiPSC Maintenance	Essential 8 medium	-	Gibco
		Penicillin/Streptomycin	1%	Life Technologies
MEF	Plate Coating for Differentiation	Advanced DMEM/F12, GlutaMAX	-	Gibco
		FBS	10%	Sigma
		L-glutamine	1%	Life Technologies
		Penicillin/Streptomycin	1%	Life Technologies
		β-Mercaptoethanol	0.0007%	Sigma Aldrich
CDM/PVA	Differentiation of hiPSCs toward HLCs	F-12	-	Gibco
		IMDM	-	Gibco
		Chemically Defined Lipid Concentrate	1%	Thermo Fisher
		1-Thio Glycerol	0.004%	Sigma Aldrich
		Insulin	7 μg/mL	Sigma Aldrich
		Transferrin	15 μg/mL	Sigma Aldrich
		Penicillin/Streptomycin	1%	Life Technologies
		Polyvinyl Alcohol	0.5g	Sigma Aldrich
RPMI	Differentiation of hiPSCs toward HLCs	RPMI	-	Gibco
		B27 Serum-Free Supplement	2%	Life Technologies
		MEM Non-Essential Amino Acids Solution	1%	Gibco
		Penicillin/Streptomycin	1%	Life Technologies
Hepatozyme	Differentiation of hiPSCs toward HLCs	Hepatozyme	-	Life Technologies
		MEM Non-Essential Amino Acids Solution	2%	Gibco
		Chemically Defined Lipid Concentrate	2%	Thermo
		L-glutamine	1%	Life Technologies
		Insulin	14 μg/mL	Sigma Aldrich
		Transferrin	30 μg/mL	Sigma Aldrich
		Penicillin/Streptomycin	1%	Life Technologies
3D Culture Medium	Long term maintenance in 3D culture	William's E	-	Invitrogen
		Nicotinamide	10 mM	Sigma Aldrich
		Hepes	20 mM	Sigma Aldrich
		Sodium Carbonate	17 mM	Sigma Aldrich
		Sodium Pyruvate	6.3 mM	Invitrogen
		2-Phospho-L-Ascorbic Acid Trisodium Salt	0.2 mM	Sigma Aldrich
		Glucose	14 mM	Thermo
		ITS+ Premix Universal Culture Supplement	1%	Corning
		Dexamethasone	0.1 μM	R&D
		L-glutamine	1%	Life Technologies
		Penicillin/Streptomycin	1%	Life Technologies

2.2.3 3D Cell Culture

Cells designated for 3D culture followed the above hepatocyte differentiation protocol until day 23. On day 23, the cells were dissociated with cell dissociation buffer (Thermo Fisher) for 12 minutes at 37˚C, 5% CO_2, and 5% O_2. Following incubation, cells were removed from the plate and manually broken into small clumps using a serological pipette. Cell count was estimated using the Luna Automated Cell Counter. Cells were then resuspended at 2.0×10^6 cells/mL in growth factor reduced matrigel basement membrane matrix (Corning). The matrigel was dispersed in 24 well plates at 50 µL/well and the plates were stored upside-down at 37˚C, 5% CO_2, and 5% O_2 until the matrigel had completely solidified. The matrigel was then overlaid for 24 hours with 3D culture medium supplemented with 20 ng/mL Oncostatin M, and 50 ng/mL HGF. Following the 24-hour recovery period, experimentation could begin including treatment with fatty acids. Cells could be maintained in 3D culture at 37˚C, 5% CO_2, and 5% O_2 for up to 2 weeks changing the medium every 48 hours.

2.2.4 Lipid Treatment

Oleic acid (Sigma Aldrich) and palmitic acid (sigma Aldrich) were reconstituted to 100 mM in 95% Ethanol. To prepare fatty acid medium, fatty acids are first conjugated to BSA (Sigma Aldrich) in order to facilitate uptake into the cells. Fatty acids were warmed at 60˚C for 30 minutes. Fatty acids were then mixed with 12.5% BSA and heated at 60˚C for an additional 10 minutes. Control medium consisted of 8% heated BSA in 3D culture medium while OA and PA medium consisted of 3D culture medium supplemented with 250 µM of BSA-conjugated oleic acid and palmitic acid respectively. It is important to note that ethanol was not added to the control medium as a vehicle control. This was a mistake in the experimental design that should be rectified in any future experiments. However, I believe that the effect of this oversight will be minor as the OA and PA medium contain only 0.25% ethanol. The cells were maintained in their respective medium for 7 days at 37˚C, 5% CO_2, and 5% O_2, changing the medium every 48 hours for the duration.

2.2.5 Triglyceride Blocking Assay

Cells were differentiated and placed into 3D cell culture as described above. The cells were then treated for 1 week with PA medium or PA medium supplemented with 0.5µM

Triacsin C (R&D) or 10 μM T863 (Sigma Aldrich) to block triglyceride formation. Following 1 week of treatment, cells were assessed for viability and lipid accumulation.

2.2.6 Ethanol Toxicity Assay
Cells were differentiated and placed into 3D cell culture as described above. The cells were then treated for 48 hours with either control or OA medium to induce steatosis. Following this treatment, "untreated cells" continued to be treated with their respective medium for the duration of the experiment. The cells treated with the toxic insult were cultured in their respective medium supplemented with 100 mM ethanol for 48 hours. For the final 24 hours of the experiment, the ethanol treated cells were treated with their respective medium supplemented with 100 mM ethanol and 20 ng/mL of TNFα. TNFα was used to enhance the ethanol toxicity.

2.2.7 Acetaminophen Toxicity Assay
Cells were differentiated and placed into 3D cell culture as described above. The cells were then treated for 48 hours with either control or OA medium to induce steatosis. Following this treatment, the cells were cultured in their respective medium supplemented with 25 mM acetaminophen (Sigma) for 48 hours.

2.2.8 Iron Toxicity Assay
Cells were differentiated and placed into 3D cell culture as described above. The cells were then treated for 48 hours with either control or OA medium to induce steatosis. Following this treatment, the cells were cultured in their respective medium supplemented with 200 μM ferric ammonium citrate (Sigma) for 24 hours.

2.2.9 Cytochrome P450 Assay
CYP3A4 activity was measured using the P450-Glo CYP3A4 Assay System (Promega). This assay was performed in Hepatozyme medium because the 3D culture medium exhibited a high background luminescence. The Matrigel-embedded cells were incubated with 3 μM luciferin-IPA in Hepatozyme for 4 hours at 37˚C, 5% CO_2, and 5% O_2. The medium was then collected and mixed 1:1 with luciferin detection reagent and allowed to equilibrate for 20 minutes at room temperature. The luminescence was measured using

the SpectraMax I3x monochromator with an integration time of 5 seconds/well. Each assay was performed in technical triplicate. The luminescence measurements of the technical replicates were averaged and then normalized to cell number. Since hepatocytes did not replicate in 3D culture, the number of cells seeded in the Matrigel at the beginning of the 3D culture was used for normalization purposes (100,000 cells/well).

2.2.10 Presto Blue Viability Assay

Cell viability was measured using the presto blue viability reagent (Thermo Fisher). Matrigel-embedded cells were incubated for 4 hours at 37˚C, 5% CO_2, and 5% O_2 in a solution of 10% Presto Blue/3D culture medium. Following incubation, the medium was collected and the fluorescence in the excitation/emission range of 560/590 nm was measured. Each assay was performed in technical triplicate. The fluorescence measurements of the technical replicates were averaged and then normalized to cell number. Since hepatocytes did not replicate in 3D culture, the number of cells seeded in the Matrigel at the beginning of the 3D culture was used for normalization purposes (100,000 cells/well).

2.3 Molecular Biology Techniques

2.3.1 qPCR

Total RNA was extracted using the RNeasy Micro Kit (Qiagen). Depending on the total RNA yield, 50ng-500ng of total RNA was reverse-transcribed using Superscript II Reverse Transcriptase (Invitrogen). For reverse transcription, RNA for each sample was first diluted to a total volume of 11.8 µL using nuclease free water. This RNA mixture was then denatured by adding 0.5 µL of Random Primers and 1 µL of dNTPs and incubating the mixture at 65°C for 5 minutes. The samples were snap cooled on ice before beginning the reverse transcription reaction. The reverse transcription master mix contained 4 µL of 5X 1st strand buffer, 2 µL of 0.1M DTT, 0.5 µL of RNase OUT, and 0.2 µL of Superscript II. The master mix (6.7 µL) was added to each denatured RNA sample and mixed thoroughly by pipetting. The solutions was then incubated at 25°C for 10 minutes, 42°C for 50 minutes, and 70°C for 15 minutes. The resulting cDNA was then diluted to a final concentration of 0.83 ng/µL with nuclease free water for use in qPCR. The qPCR reaction mixture for each well was prepared using 5 µL of 2X Kapa Sybr Fast

Low Rox (Kapa Biosystems), 0.125 µL each of forward and reverse primers, 2.05 µL of nuclease free water, and 2.5 µL of cDNA for a total rection volume of 10 µL. Each reaction was performed in technical duplicate. The mixture was denatured at 95°C for 10 minutes. The mixture was then heated for 30 seconds at 95°C, cooled to 60°C for 30 seconds, and reheated to 72°C for 30 seconds. This cycle was repeated 40 times before the mixture was maintained at 72°C for 10 minutes. Following amplification, the threshold for all genes was set to 0.1. The genes were normalized to the geometric mean of the housekeeping genes ubiquitin (UBC) and porphobilinogen deaminase (PBGD). Primer sequences can be found in Table 2.2.

Table 2.2 qPCR Primer Sequences

Gene Name	Forward Sequence	Reverse Sequence
PNPLA3	AGTCGTGGATGCCTTGGTATG	CGGTGATGGTTGTTTTGGCA
A1AT	CCACCGCCATCTTCTTCCTGCCTGA	GAGCTTCAGGGGTGCCTCCTCTG
ALB	CCTTTGGCACAATGAAGTGGGTAACC	CAGCAGTCAGCCATTTCACCATAG
HNF4a	CATGGCCAAGATTGACAACCT	TTCCCATATGTTCCTGCATCAG
TTR	ATGGCTTCTCATCGTCTGCT	TGTCATCAGCAGCCTTTCTG
BIP	TGTTCAACCAATTATCAGCAA	TTCTGCTGTATCCTCTTCACCAGT
GADD34	CGCCCAGAAACCCCTACTCAT	TCGGAGAAGCGCACCTTTCT
CHOP	GGAACCTGAGGAGAGAGTGTTC	GCTTTCAGGTGTGGTGATGT
PERK	TCCAGAGATTGAGACTGCGTG	TAATGACCTTTTCTTCCCTGCTCC
UBC	ATTTGGGTCGCGGTTCTTG	TGCCTTGACATTCTCGATGGT
PBGD	GGAGCCATGTCTGGTAACGG	CCACGCGAATCACTCTCATCT

2.3.2 Immunofluorescence Microscopy

Cells grown in 2D were rinsed with dPBS with calcium chloride and magnesium chloride (Gibco) before being fixed for 20 minutes in a cold 4% paraformaldehyde (ChemCruz) solution. Following fixation, cells were washed 3 times for 5 minutes in dPBS and stored at 4°C until staining. Cells were blocked and permeabilized in a 10% Donkey Serum (Jackson ImmunoResearch)/0.1% Triton-X (Sigma) for 30 minutes at room temperature. The cells were then incubated overnight at 4°C in primary antibody (1:100 in 1% Donkey Serum/0.1% Triton-X). Following incubation, the cells were rinsed 3 times for 5 minutes

in dPBS. The cells were then incubated with the secondary antibody (1:1000 in 1%Donkey Serum/0.1% Triton-X) for 1 hour at room temperature. Following incubation, the cells were washed 3 times for 5 minutes in dPBS. In order to stain the nucleus, cells were incubated with Hoeschst (1:5,000 in dPBS; Thermo Scientific) for 5 minutes at room temperature. The samples were rinsed twice with dPBS and stored at 4°C until imaging. A list of antibodies used can be found in Table 2.3.

Table 2.3 Antibodies

Antibody Name	Primary Antibody	Secondary Antibody
PNPLA3	R&D AF5208	Thermo Scientific A-11015
ALB	Bethyl A80-229A	Thermo Scientific A-11055
HNF4a	Abcam ab92378	Thermo Scientific A-10042

In order to visualize the lipid droplets within the cells, 3D Matrigel-embedded cells were stained with Bodipy (Thermo Scientific). Cells in the 3D matrigel system were first rinsed with dPBS before being fixed for 20 minutes at room temperature in a warm 4% paraformaldehyde solution to reduce the loss of the Matrigel. Following fixation, cells were washed 3 times for 5 minutes in dPBS and stored at 4˚C until staining. The cells were incubated in a 1 µg/mL solution of Bodipy for 30 minutes at room temperature. Following the incubation, the cells were washed twice with dPBS. Nuclei were stained using Hoescht (1:5000 in dPBS) for 30 minutes at room temperature. Following staining, the cells were washed twice with dPBS and stored at 4˚C until imaging.

2.3.3 Flow Cytometry

Flow cytometry was used to quantitatively measure the amount of lipid accumulation within fatty acid-treated cells. Due to technical difficulties dissociating cells grown in 3D, all flow cytometry experiments were performed on cells grown in 2D. Cells were differentiated into hepatocyte-like cells in 2D as described above. Beginning on day 23, the cells were treated with Control, OA, or PA medium for 1 week, changing the medium every 48 hours. Following this treatment, the cells were dissociated into single cells using

TryplE Express (Gibco) for 20 minutes and 37˚C. The dissociated cells were then collected and fixed in 4% paraformaldehyde at 4˚C for 15 minutes. Cells were rinsed and stored at 4˚C in dPBS/1% BSA until staining. Cells were incubated for 30 minutes at room temperature with 1 µg Bodipy in dPBS/1%BSA. Following staining, cells were rinsed twice with dPBS/1% BSA before being analysed by flow cytometry.

2.4 RNA Sequencing

2.4.1 Library Preparation

Total RNA was extracted using the RNeasy Micro Kit. The quantity of RNA was measured using the Quant-iT RiboGreen RNA Assay Kit (Thermo Scientific). The quality of RNA was then measured using the Bioanalyzer 2100 (Agilent). Library preparation was performed using the NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (New England Biosystems) according to the manufacturer's protocol. Twelve libraries were multiplexed per lane and single end, 50 base pair sequencing was performed on the HiSeq 2500 (Illumina).

2.4.2 Bioinformatics Analysis

The initial quality control, alignment to the human reference genome (GRCh37/hg19), and read counting was performed using the CCBR RNA-Seq Pipeliner. Differential expression analysis was performed using DESeq2 1.26.0. An adjusted P value of less than 0.05 was considered as significantly differentially expressed. Principal component analysis was performed on normalized, vsd transformed, and batch corrected counts. Heatmaps were acquired using the ClustVis web tool and gene ontology analysis was performed using the DAVID bioinformatics resources 6.8 functional annotation tool.

2.5 Lipidomics

Cells were removed from Matrigel using Cell Recovery Solution (Corning) for 30 minutes at 4˚C. The cells were then washed twice with dPBS to remove any remaining Matrigel before lipid extraction. A Folch lipid extraction was performed to extract total lipids from the cells. The cells were homogenized in a 2:1 chorloform:methanol solution. The solution was centrifuged for 5 minutes at 20,000 g. The organic phase was recovered

for analysis. The extraction was performed twice on each sample and the organic phase was from each extraction was combined and sent for analysis via direct injection liquid chromatography mass spectrometry. Each experiment was performed in technical quintuplicate. The lipidomic data was then analysed using the MetaboAnalyst web-based analytical pipeline for high-throughput metabolomic studies.

2.6 Statistical Analysis

All graphical values are shown as mean ± standard deviation. Statistical analyses were performed using GraphPad Prism 8 and the R statistical environment. Ordinary one-way ANOVAs with Dunnett's multiple comparisons tests were performed to test statistical significance between means. P-values less than 0.05 were considered to be statistically significant.

3 INTRODUCTION OF MUTATIONS INTO THE PNPLA3 GENE DOES NOT AFFECT HIPSC'S ABILITY TO DIFFERENTIATE INTO HEPATOCYTES

3.1   Introduction

CRISPR/CAS9 is a powerful tool for genome editing that can be used to modify the genome of nearly any cell type. While the potential for use of hiPSCs in disease modelling is nearly endless, the combination of CRISPR/CAS9 and hiPSCs for the purposes of modelling genetic disorders is a booming area of research[214-216]. hiPSCs are easily genetically manipulated by CRISPR/CAS9 due to their high proliferative capacity and clonogenicity [217]. The high proliferation rate characteristic of pluripotent cells improves the efficiency of transfection and CRISPR/CAS9 gene editing by allowing more exogenous genetic material to enter the nucleus during mitotic events. Additionally, hiPSCs are highly clonogenic meaning that they are capable of producing an entire colony of genetically identical clones from a single starting cell. Since CRISPR/CAS9 is not 100% efficient, cells must be evaluated to determine if they have been successfully targeted and/or edited. This is usually accomplished using some combination of antibiotic resistance, expression of a fluorescent marker, and/or genetic sequencing. Given their clonogenic potential, hiPSCs can be placed into single cell suspension following transfection and seeded in such a way that allows distinct, clonogenic colonies to grow from a single cell. These clonogenic colonies can then be collected and sequenced to evaluate the editing efficiency of CRISPR/CAS9. This unique characteristic of hiPSCs prevents the necessity of complex, costly techniques such as fluorescence-activated cell sorting (FACS) to separate edited cells before sequencing. hiPSCs also appear to be rather impervious to CRISPR/CAS9-mediated off target effects when using highly specific sgRNAs [218, 219]. This improved editing efficiency as well as the increased availability make hiPSCs more practical models for genome editing than primary cells.

CRISPR/CAS9 edited hiPSCs are ideal candidates for disease modelling of genetic disorders for several reasons. Several isogenic lines can be generated and differentiated into the cell type(s) of interest. These isogenic lines act as internal controls to dissociate the effects of the specific mutation of interest, genetic background, and off-target effects. Given the nature of CRISPR/CAS9 off-target effects, it is unlikely that the same off-target editing events would occur in multiple lines. Therefore, if isogenic lines are exhibiting different phenotypes, it can be assumed that the observed phenotype is a result of an off-target effect rather than the mutation of interest [220]. Additionally, CRISPR/CAS9 can

be used to easily generate isogenic, edited lines from multiple genetic backgrounds. These lines can be used to robustly confirm the pathogenic mechanism of a genetic mutation as well as investigate the interactions between a pathogenic mutation and genetic background. The most unique advantage of using hiPSCs for genetic editing is their ability to differentiate into multiple cell types of interest. Genetically edited hiPSCs can be differentiated into multiple disease-relevant cell types to dissect the unique effect of the mutation of interest on different cell types. Multiple cell types could then be cultured together in a co-culture system to model intercellular interactions and the effect of genotype on these interactions. This allows for a more complete and nuanced understanding of how genetic variants or mutations contribute to a complex disease phenotype.

For these reasons, we chose to use CRISPR/CAS9 to either completely knock-out the PNPLA3 protein (PNPLA3^{KO}) or to knock-in the I148M variant (PNPLA3^{I148M}). In order to confirm the robustness of our findings, we performed CRISPR/CAS9 targeting in two well-characterized hiPSC lines: FSPS13B and A1ATDR/R [185, 190, 197]. Both lines were genetically normal and had the capacity to differentiate into hepatocytes. We chose to target exon 3 of the *PNPLA3* gene as this is where the I148M variant resides. Additionally, by targeting exon 3, we increased the likelihood of knocking-out the PNPLA3 protein. Since exon three is shared by all PNPLA3 isoforms and it is positioned early in the coding region of the protein which increases the likelihood of producing a true knock-out rather than a truncated protein. sgRNAs were designed in silico using online resources to both optimize their efficiency and minimize their off-target effects. The sgRNAs were then cloned into the PX459 CRISPR/CAS9 editing plasmid (Figure 3.1) The ssODN was designed to introduce the variant or interest as well as several silent mutations. These silent mutations are used to prevent CRISPR/CAS9 from re-editing an allele that has been properly targeted.
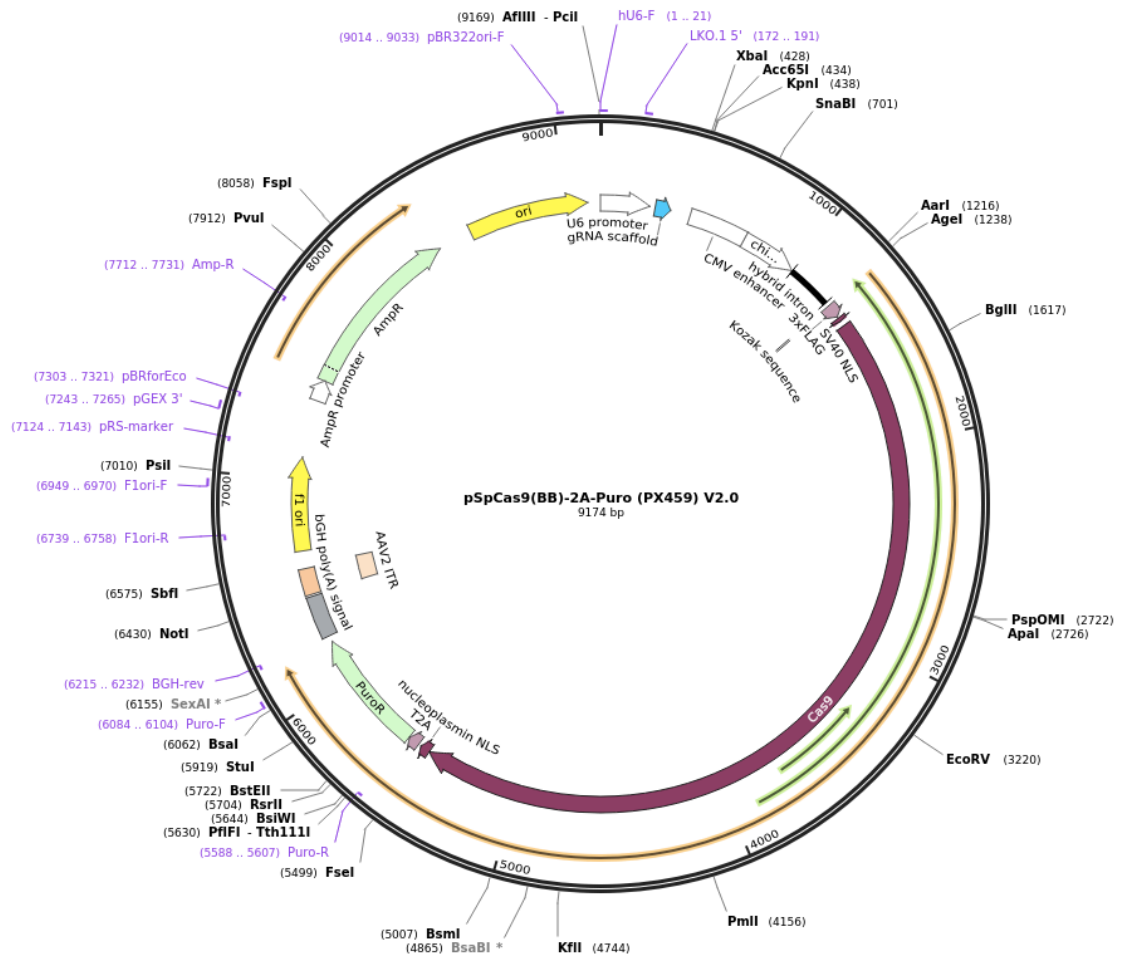
Figure 3.1 Gene editing plasmid map.

Plasmid map of the PX459 plasmid used for CRISPR/CAS9 gene editing.

Nucleofection was used to transfect the editing plasmid and ssODN into the hiPSCs. This is an electroporation technique that uses specific voltages and reagents to transfer nucleic acids directly into the nucleus of cells. Compared to other transfection techniques that rely on mitotic events to transport the donor plasmids into the nucleus, nucleofection improves the editing efficiency of CRISPR/CAS9 by ensuring maximum influx of the plasmid into each cell regardless of mitotic status. hiPSCs also demonstrate higher viability when subjected to nucleofection compared to other electroporation techniques.

The editing plasmid was only transiently expressed in the cells rather than stably incorporating into the genome. This technique reduces the exposure time of hiPSCs to

CAS9 which simultaneously prevents re-editing events and reduces off-target editing events. Re-editing events increase the risk for unwanted mutations to arise in the area of interest and reduce the likelihood of producing homozygous edited clones. Long exposure times of CAS9 have been linked to increased off target effects due to the promiscuous nature of CAS9 [209, 221]. By controlling for these variables using transient expression, we were able to improve the editing specificity of CRISPR/CAS9 in our cells. However, the short half-life of the plasmid does reduce the editing efficiency somewhat. This burden can be overcome by designing an extremely efficient sgRNA for targeting. The transient expression of the editing plasmid required only a short selection period with puromycin, approximately 48 hours. Following selection, cells were allowed to clonally expand for two to three weeks before colonies were picked and sequenced to identify edited clones of interest and their untargeted counterparts.

In this chapter, we used CRISPR/CAS9 to generate PNPLA3^{UC}, PNPLA3^{I148M}, and PNPLA3^{KO} lines in two different genetic backgrounds. Following CRISPR/CAS9 editing, lines with the appropriate genotype were identified by sequencing and the expression of PNPLA3 in each of these lines was extensively characterized. The characterized lines were then differentiated into hepatocyte like cells (HLCs) to ensure that CRISPR/CAS9 editing did not negatively affect differentiation capacity. Following differentiation, the effect of 3D culture on HLC maturity and functionality was evaluated.

3.2   PNPLA3 was successfully edited using CRISPR/CAS9

The two hiPSC lines were sequenced before CRISPR/CAS9 targeting and it was determined that FSPS13B was homozygous for the reference allele while A1ATDR/R was naturally heterozygous for the I148M variant. These initial sequences informed our CRISPR/CAS9 targeting strategy. For FSPS13B, we chose a target sequence that included amino acid position 148. The chosen sgRNA was predicted to have high editing efficiency given the proximity of the double stranded break to the editing position. However, in practice, the editing efficiency of this sgRNA was quite low. Of the 650 clones that were picked and sent for sequencing, only two were homozygous for the I148M variant which equates to an editing efficiency of approximately 0.3%. Given the

heterozygosity of the A1ATDR/R hiPSC line, a different sgRNA had to be chosen so that both alleles could be targeted and edited in order to generate a homozygous knock-out. This sgRNA was designed to target the sequence directly adjacent to amino acid position 148 in order to maintain a relatively high editing efficiency. The actual editing efficiency of this sgRNA was significantly higher than the first. Of the 189 clones that were screened, one was homozygous for the I148M variant which equates to an editing efficiency of approximately 0.5%. It should be noted that the low editing efficiency of these sgRNAs may be due to a suboptimal targeting protocol as opposed to a lack of specificity. In order to increase the editing efficiency, it would be advisable to directly deliver purified Cas9 protein into hiPSCs as opposed to relying on the Cas9 expression vector [222, 223]. Following targeting and sequencing, several isogenic lines were identified and used for further experiments. For the FSPS13B cell line, seven clones were chosen for further experiments: two untargeted control lines (PNPLA3^{UC}), two I148M knock-in lines (PNPLA3^{I148M}), and three knock-out lines (PNPLA3^{KO}). PNPLA3^{UC} lines were defined as cells that had undergone CRISPR/CAS9 targeting but their genetic sequence remained unchanged. These lines functioned as both reference allele controls as well as an internal control for potential off-target effects due to CRISPR/CAS9 targeting. Each of the PNPLA3^{I148M} clones were homozygous for the cytosine to guanine substitution that characterizes the I148M variant as well as three silent mutations. Each of the PNPLA3^{KO} clones contained a slightly different nonsense mutation. PNPLA3^{KO} clone 1 had a very large deletion of greater than 130 base pairs while PNPLA3^{KO} clones 2 and 3 had a single base pair deletion and insertion respectively. The selected clones were homozygous for their respective indels. The homozygous clones were chosen in order to avoid any complications pertaining to haplosufficiency. The targeting strategy as well as the sequences for each of the chosen clones can be found in Figure 3.2.
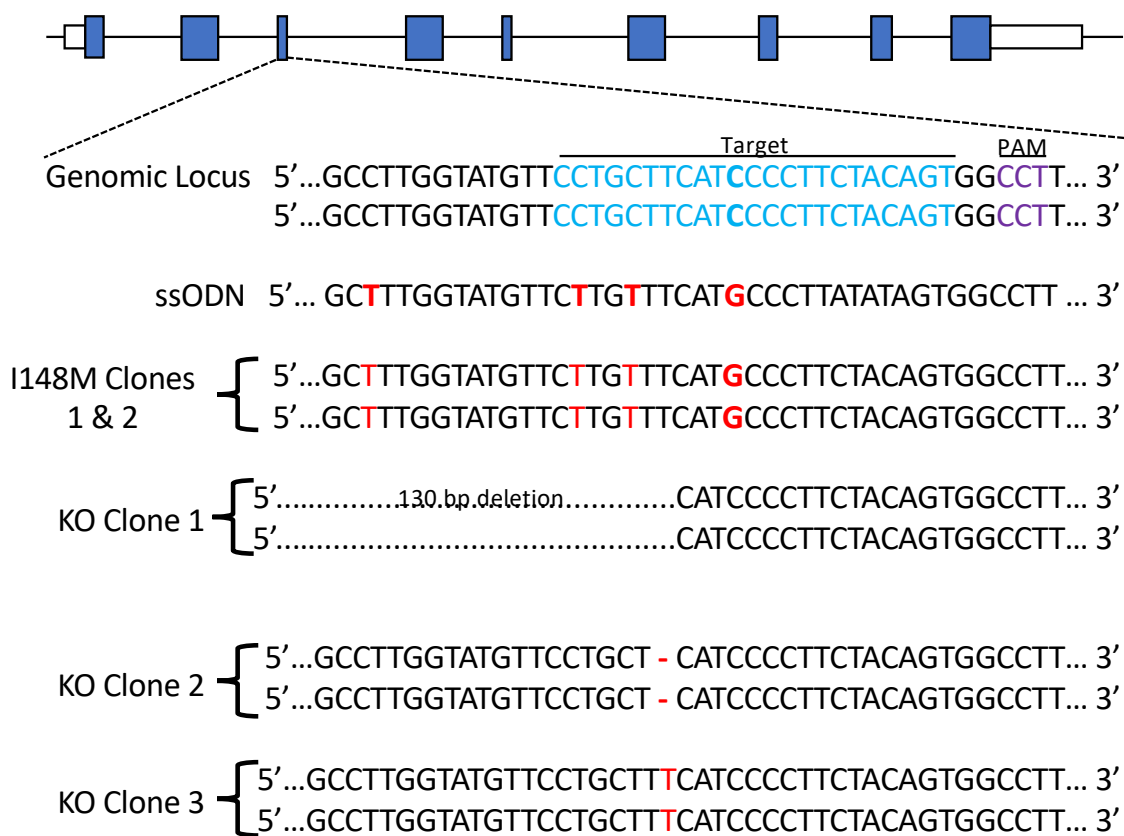
Figure 3.2 CRISPR/CAS9 targeting strategy for FSPS13B.

Schematic demonstrating the CRISPR/CAS9 targeting strategy at exon 3 of the *PNPLA3* gene locus for the FSPS13B cell lines. The sgRNA target sequence is shown in blue while the PAM is shown in purple. The sequences of the ssOND, PNPLA3^{I148M}, and PNPLA3^{KO} clones are shown below the targeting schematic. Any mutations in these sequences are highlighted in red.

For the A1ATDR/R validation cell line, five lines were chosen: two PNPLA3^{UC} lines, one PNPLA3^{I148M} line, and two PNPLA3^{KO} lines. The PNPLA3^{I148M} line was homozygous for the I148M variant as well as four silent mutations while both PNPLA3^{KO} lines were homozygous for their respective indels. The targeting strategy as well as the sequences for each of the clones of interest can be found in Figure 3.3.

Samantha G. Tilson - January 2021
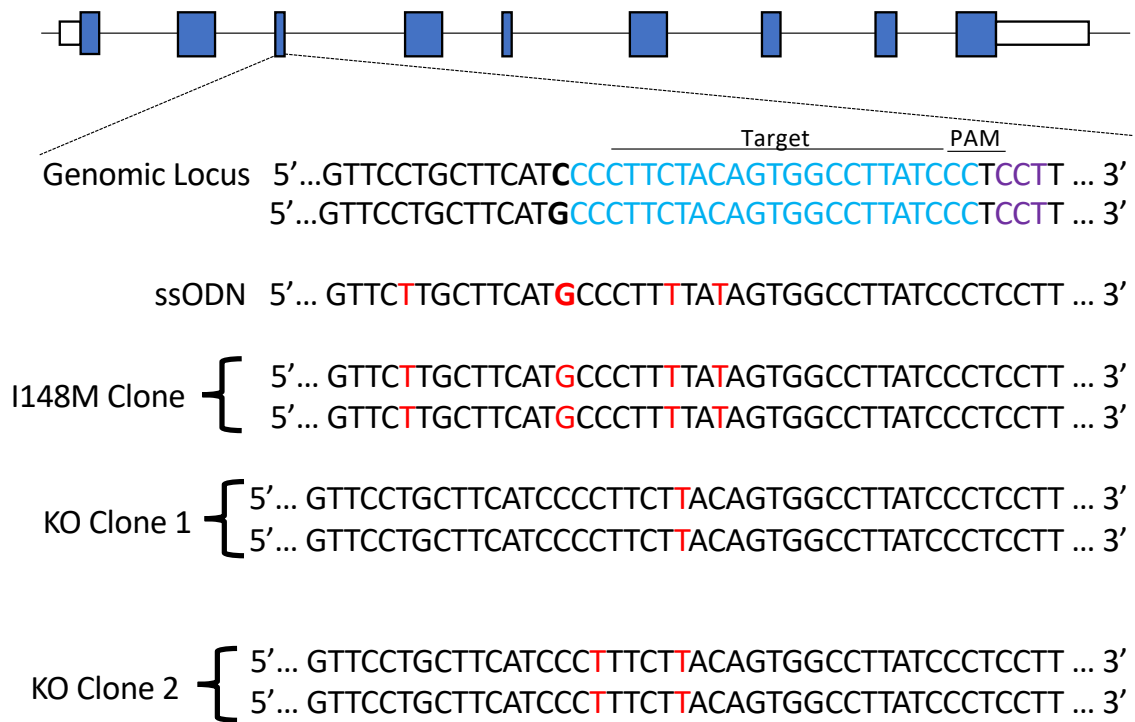
Figure 3.3 CRISPR/CAS9 targeting strategy for A1ATDR/R.

Schematic demonstrating the CRISPR/CAS9 targeting strategy at exon 3 of the *PNPLA3* gene locus for the A1ATDR/R cell lines. The sgRNA target sequence is shown in blue while the PAM is shown in purple. The sequences of the ssOND, PNPLA3[I148M], and PNPLA3[KO] clones are shown below the targeting schematic. Any mutations in these sequences are highlighted in red.

Following the targeting and initial identification of clones, we sought to quantify the expression of PNPLA3 in each of these clones. We measured the mRNA expression of *PNPLA3* using qPCR (Figure 3.4). Each of the clones were pooled into their respective genotype group for these analyses. We found that in both genetic backgrounds, the PNPLA3[UC] and PNPLA3[I148M] lines had high expression of *PNPLA3* mRNA while the PNPLA3[KO] lines had near zero expression of the mRNA, likely due to nonsense mediated RNA degradation. In the FSPS13B line, the PNPLA3[I148M] lines had significantly lower expression of *PNPLA3* mRNA than PNPLA3[UC] lines. It is unclear why this would be the case, but the same trend was not observed in the A1ATDR/R line which implies that it may be an idiosyncrasy that is unique to the FSPS13B line. Regardless, the PNPLA3[I148M] line still expressed higher levels of PNPLA3 mRNA than the PNPLA3[KO] line so we chose to move forward with the characterization of these lines.
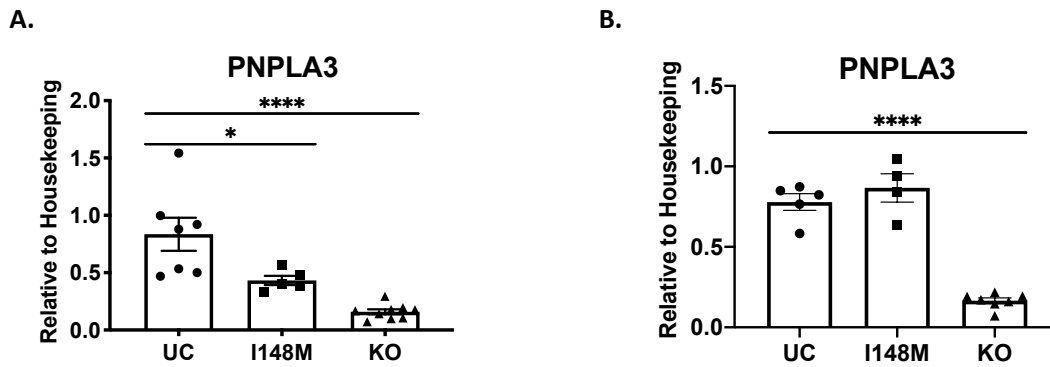
Figure 3.4 Expression of PNPLA3 mRNA.

PNPLA3 mRNA expression in HLCs of each genotype at the exon-exon junction of exons 2 and 3 for the **A.** FSPS13B lines (PNPLA3^{UC}: n = 2 clones and 5 independent experiments; PNPLA3^{I148M}: n = 2 clones and 4 independent experiments; PNPLA3^{KO}: n = 3 clones and 3 independent experiments) and **B.** A1ATDR/R lines (PNPLA3^{UC}: n = 2 clones and 4 independent experiments; PNPLA3^{I148M}: n = 1 clones and 4 independent experiments; PNPLA3^{KO}: n = 2 clones and 4 independent experiments). PNPLA3^{UC} cells express high levels of PNPLA3 mRNA while PNPLA3^{KO} cells express near zero levels of *PNPLA3* mRNA (greater than five-fold lower than PNPLA3^{UC}). PNPLA3^{I148M} cells express high levels of *PNPLA3* mRNA, though in the case of the FSPS13B background, the expression was slightly lower than that of PNPLA3^{UC} cells. Ordinary one-way ANOVAs with Dunnett's multiple comparisons tests were performed to test statistical significance between means. Error bars represent SEM

We next sought to evaluate the protein expression of PNPLA3 in each of the chosen lines. Initially, we attempted to quantify the protein expression of PNPLA3 using western blot. However, our efforts were hindered by the lack of a reliable antibody for quantifying human PNPLA3 protein levels via western blot. Studies of PNPLA3 in human cells are almost invariably conducted using overexpression systems. When PNPLA3 is expressed at endogenous levels, these antibodies are no longer specific or reliable in our hands. We tested four commercially available human PNPLA3 antibodies (Abcam ab69170, Abcam ab81874, Novus NBP2-34094, and R&D AF5208). In addition, we requested permission from Helen Hobbs to test the human PNPLA3 antibody designed by her lab. Unfortunately, none of these antibodies gave us results that were reliable and/or reproducible. Therefore, we chose instead to measure PNPLA3 protein expression using immunofluorescence. While this method is not quantitative, it was the only reliable tool available to us at the time.

PNPLA3 is a membrane bound protein that localizes to the plasma membrane, ER membranes, and lipid droplet membranes. However, due to technical difficulties, we were unable to permeabilize the cells and visualize the internal membranes such as lipid droplet membranes. Permeabilization resulted in a diffuse, non-specific staining pattern which made imaging difficult and unreliable. Some cytoplasmic staining is visible despite the lack of permeabilization. Figures 3.5 and 3.6 show the membrane expression of PNPLA3 in FSPS13B and A1ATDR/R, respectively. In general, the PNPLA3^{UC} and PNPLA3^{I148M} lines expressed appreciable amounts of PNPLA3 protein in both genetic backgrounds while the PNPLA3^{KO} lines did not. The A1ATDR/R lines appear to have a slightly lower basal expression of PNPLA3 compared to FSPS13B. In both genetic backgrounds, the PNPLA3^{UC} and PNPLA3^{I148M} lines appear to express similar levels of PNPLA3 protein, even in the FSPS13B line where the PNPLA3^{I148M} line had significantly lower mRNA expression than the PNPLA3^{UC} line. This result appears to contradict the evidence in mice that shows PNPLA3 I148M protein accumulates on the surface of lipid droplets at 40-fold higher levels than I148I PNPLA3 by evading ubiquitination [162, 163]. Given that we were unable to visualize the internal lipid droplet membranes due to the lack of a reliable, specific human PNPLA3 antibody for intracellular staining, we cannot dismiss the possibility that the I148M clones have higher intracellular expression of PNPLA3 than their UC counterparts.

Despite the high basal expression in the PNPLA3^{UC} and PNPLA3^{I148M} FSPS13B lines, there was no detectable expression of PNPLA3 protein in the PNPLA3^{KO} lines. This lack of protein expression confirms that CRISPR/CAS9 was successfully used to knock-out the PNPLA3 gene in all three of the chosen FSPS13B PNPLA3^{KO} lines.
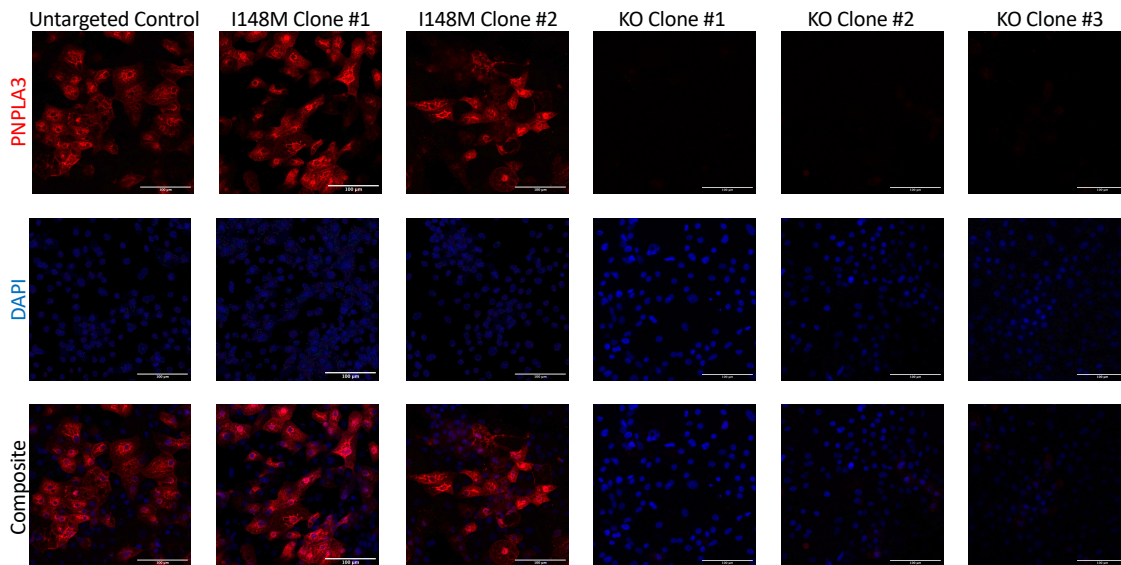
Figure 3.5 PNPLA3 protein expression in FSPS13B lines.

Protein expression of PNPLA3 in FSPS13B lines following differentiation to HLCs. Expectedly, PNPLA3^{UC} and PNPLA3^{I148M} cells expressed high levels of PNPLA3 protein while PNPLA3^{KO} cells did not express any detectable levels of PNPLA3 protein. These results indicate that a full knock-out of PNPLA3 was achieved in these cells. This experiment was performed 5 times and the best representative images are represented in this figure.

In the A1ATDR/R PNPLA3^{KO} lines, there was a small amount of PNPLA3 protein staining detected. However, this staining appears to be background noise because the staining pattern differed from that of the PNPLA3^{UC} and PNPLA3^{I148M} lines. In the PNPLA3^{UC} and PNPLA3^{I148M} lines, the PNPLA3 expression localized in the cytosol around the nucleus; however, in the PNPLA3^{KO} lines, there was no such specificity. It is possible that this signal is a result of autofluorescence or non-specific staining. Given that the mRNA expression of *PNPLA3* in these cells is extremely low and that the protein staining was unconvincing, we concluded that PNPLA3 was successfully knocked out in the A1ATDR/R line as well.
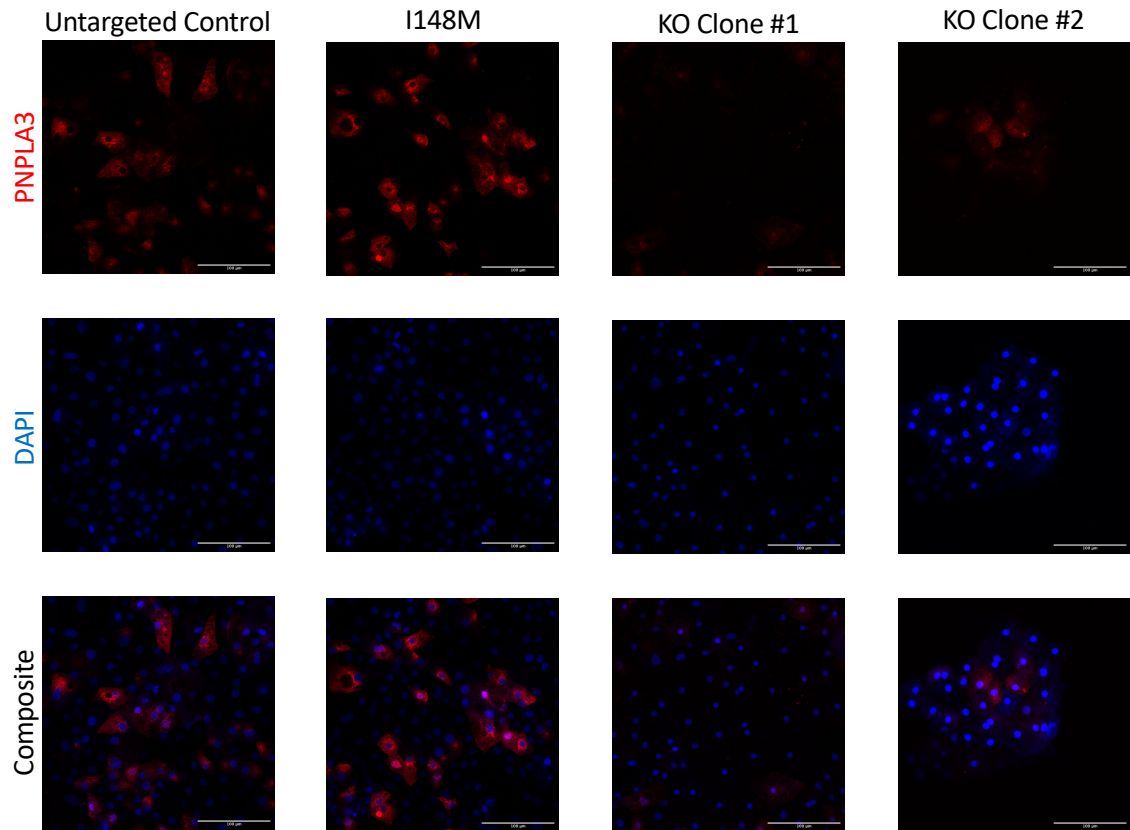
Figure 3.6 PNPLA3 protein expression in A1ATDR/R lines.

Protein expression of PNPLA3 in A1ATDR/R lines following differentiation to HLCs. Expectedly, PNPLA3^{UC} and PNPLA3^{I148M} cells expressed high levels of PNPLA3 protein. PNPLA3^{KO} cells had a small amount of staining that was determined to be non-specific. This experiment was performed 1 time and the best representative images are represented in this figure.

Overall, CRISPR/CAS9 was successfully used to generate isogenic PNPLA3^{UC}, PNPLA3^{I148M}, and PNPLA3^{KO} lines in two different genetic backgrounds. All chosen lines were homozygous for their respective genotypes. Sanger sequencing revealed that the PNPLA3^{I148M} lines were homozygous for the risk allele (C -> G at amino acid position 148) and three silent mutations designed to prevent re-editing of CRISPR/CAS9 at this locus. These silent mutations should have no effect on the structure or function of the protein. The PNPLA3^{UC} and PNPLA3^{I148M} lines both expressed high levels of PNPLA3 mRNA and protein, though the level of expression varied slightly between genetic backgrounds. The PNPLA3^{KO} lines expressed very low levels of *PNPLA3* mRNA and undetectable levels of PNPLA3 protein. Given these low levels of expression at both the mRNA and protein levels, it can be concluded that PNPLA3 was fully knocked out in

these lines. Given their isogenic nature and the similarities in their PNPLA3 expression pattern, each of the individual lines were combined with the other lines in their respective genotypes for further analyses. The FSPS13B line was used to perform most experiments and the A1ATDR/R line was used to verify the key findings of this work.

3.3 All edited hiPSC lines retain similar differentiation capacity toward hepatocytes

In order to confirm that changes to PNPLA3 and/or the CRISPR/CAS9 gene editing process did not affect the hiPSC's ability to differentiate toward the hepatic lineage, all lines were differentiated into HLCs and their maturity and functionality were assessed. The hiPSCs were differentiated into HLCs using a well-established differentiation protocol that mimics the embryonic development of hepatocytes in vivo [184]. This protocol is entirely chemically defined and produces functional hepatocytes within 30 days. The PNPLA3^{UC}, PNPLA3^{I148M}, and PNPLA3^{KO} lines in both genetic backgrounds were subjected to this differentiation protocol and their differentiation capacity was assessed in several ways. The expression of key hepatocyte markers was quantified at both the mRNA and protein levels. Additionally, their metabolic functionality was evaluated using their basal CYP3A4 activity.

Several key hepatocyte markers were used to evaluate the differentiation efficiency and HLC maturity including albumin (ALB), α-1-anti-trypsin (A1AT), hepatocyte nuclear factor 4α (HNF4a), and transthyretin (TTR). The mRNA expression of these markers was measured at day 30 of differentiation using qPCR (Figure 3.7). In order to confirm that alterations to PNPLA3 do not affect differentiation capacity, the expression levels of each of these markers in PNPLA3^{I148M} and PNPLA3^{KO} lines were compared to the expression level in PNPLA3^{UC} lines. This analysis was performed on the FSPS13B and A1ATDR/R lines separately. All of the evaluated lines successfully differentiated into HLCs as evidenced by their high expression of the hepatocyte markers *ALB, A1AT, HNF4a*, and *TTR*. The expression of these markers indicates that the CRISPR/CAS9 targeting process did not affect the differentiation capacity of the edited hiPSCs. Regardless of genetic background, the PNPLA3^{UC}, PNPLA3^{I148M}, and PNPLA3^{KO} lines expressed similar

levels of the four markers. This demonstrates that alterations to the PNPLA3 gene does not affect the cell's ability to differentiate into hepatocytes.
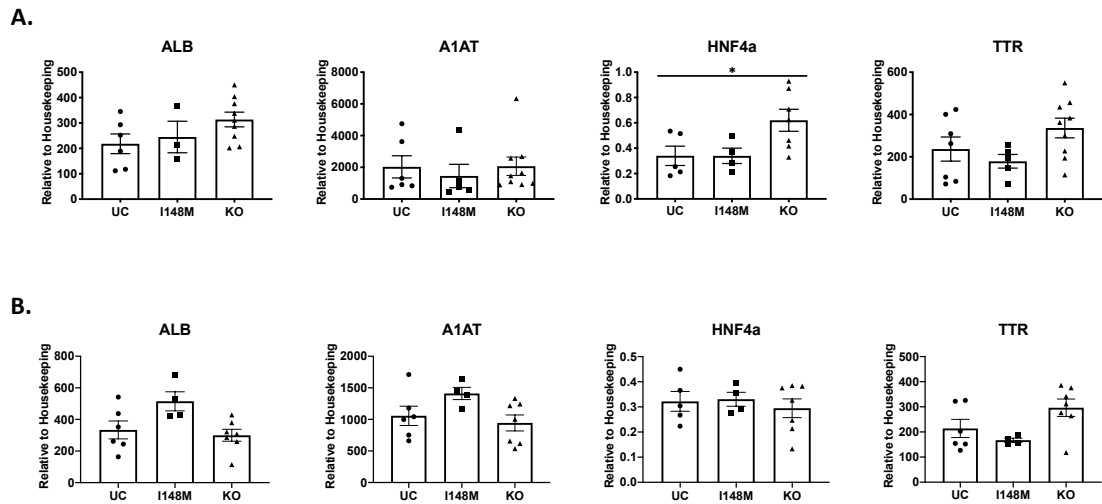
Figure 3.7 mRNA expression of hepatocyte markers.

mRNA expression of hepatocyte functionality markers *A1AT, ALB, HNF4a*, and *TTR* following differentiation of **A.** FSPS13B lines (PNPLA3^{UC}: n = 2 clones and 4 independent experiments; PNPLA3^{I148M}: n = 2 clones and 3 independent experiments; PNPLA3^{KO}: n = 3 clones and 3 independent experiments) and **B.** A1ATDR/R lines (PNPLA3^{UC}: n = 2 clones and 4 independent experiments; PNPLA3^{I148M}: n = 1 clone and 4 independent experiments; PNPLA3^{KO}: n = 2 clones and 4 independent experiments) into HLCs. High expression of these markers indicates a successful differentiation toward the hepatocyte lineage. Levels of mRNA expression for each marker are similar between the three genotypes indicating that all genotypes possess a similar capacity to differentiate into HLCs. Ordinary one-way ANOVAs with Dunnett's multiple comparisons tests were performed to test statistical significance between means. Error bars represent SEM.

The FSPS13B PNPLA3^{KO} line expressed significantly higher levels of *HNF4a* than the PNPLA3^{UC} line. Given that HNF4a plays an important role in lipid metabolism, it could be hypothesized that HNF4a and PNPLA3 are functionally connected and loss of PNPLA3 leads to the upregulation of HNF4a through a compensatory mechanism. However, the A1ATDR/R line did not demonstrate a similar up-regulation, so it is more likely that the higher expression of *HNF4a* in the PNPLA3^{KO} line of FSPS13B is an artefact of the genetic background rather than a true effect. There are slight variations in

the expression levels of each of the markers between the two genetic backgrounds, but no significant differences were observed.
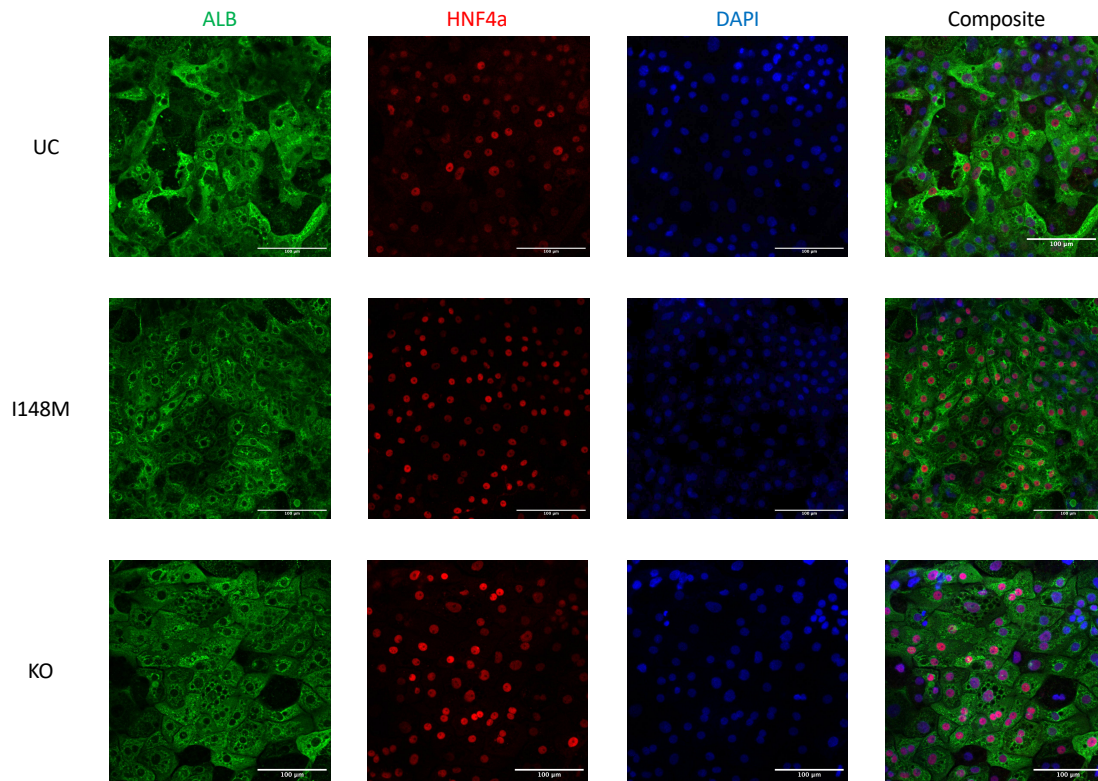
Figure 3.8 Protein expression of hepatocyte markers in FSPS13B.

Representative images of ALB and HNF4a protein staining in differentiated FSPS13B HLCs shows similar expression of mature hepatocyte markers between all three genotypes. This experiment was performed 4 times and the best representative images are represented in this figure.

In order to confirm that the hepatocyte markers were homogenously expressed on the protein level as well as the mRNA level, the expression of ALB and HNF4a was visualized using immunofluorescent microscopy. Representative micrographs for each genotype can be found in Figures 3.8 and 3.9 for FSPS13B and A1ATDR/R, respectively. All evaluated lines demonstrated homogenous expression of both ALB and HNF4a. Within each genetic background, there were no clear differences in expression of the markers between the three genotypes. These similarities offer further evidence that the CRISPR/CAS9 gene editing process did not affect the ability of the hiPSCs to differentiate into hepatocytes regardless of genotype.
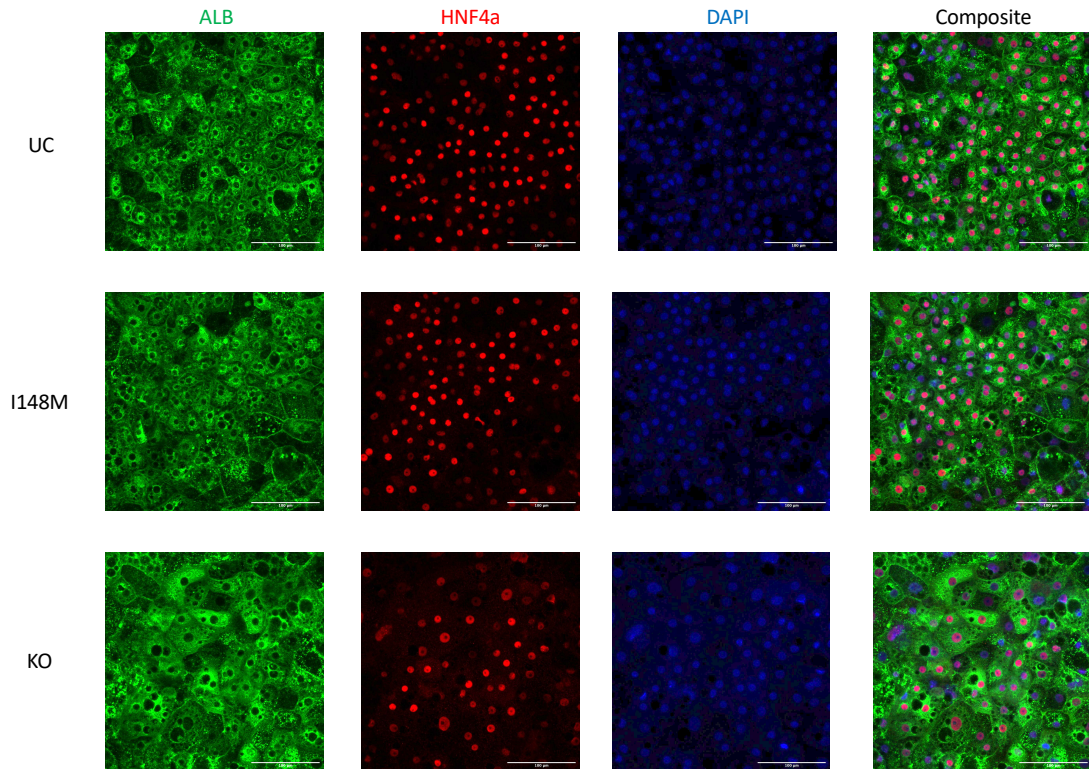
Figure 3.9 Protein expression of hepatocyte markers in A1ATDR/R.

Representative images of ALB and HNF4a protein staining in differentiated A1ATDR/R HLCs shows similar expression of mature hepatocyte markers between all three genotypes. This experiment was performed 2 times and the best representative images are represented in this figure.

We next evaluated the metabolic functionality of these cells. CYP3A4 is an extremely important protein that functions in several detoxification processes within hepatocytes. It is the most abundantly expressed member of the CYP family and it is key to the functionality of hepatocytes. CYP3A4 is present at high levels in adult hepatocytes while CYP3A7 is expressed more highly in foetal hepatocytes. Therefore, CYP3A4 activity can be used not only to assess the functionality of HLCs but also their maturity. In both the FSPS13B and A1ATDR/R lines, the basal CYP3A4 activity of the PNPLA3[UC], PNPLA3[I148M], and PNPLA3[KO] lines was measured. There was significant variability in the CYP3A4 activity between differentiations in all three genotypes. This variability is inherent to the differentiation protocol which does not generate fully functional hepatocytes. However, despite this large variability, no statistically significant differences were observed in basal CYP3A4 activity between the three genotypes in either genetic

background. This indicates, once again, that CRISPR/CAS9 targeting and PNPLA3 genotype do not affect differentiation capacity of hiPSCs toward HLCs.

A.

B.
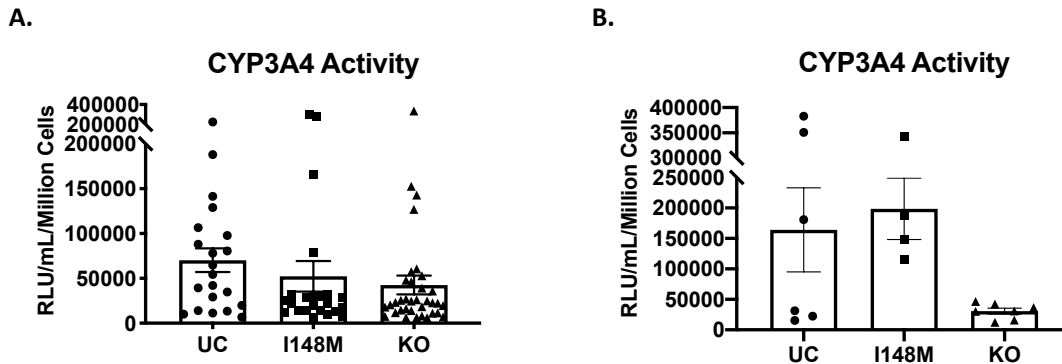
Figure 3.10 Basal CYP3A4 Activity for FSPS13B and A1ATDR/R.

Basal CYP3A4 activity in HLCs of each genotype from the **A.** FSPS13B (PNPLA3^{UC}: n = 2 clones and 11 independent experiments; PNPLA3^{I148M}: n = 2 clones and 15 independent experiments; PNPLA3^{KO}: n = 3 clones and 15 independent experiments) and **B.** A1ATDR/R (PNPLA3^{UC}: n = 2 clones and 4 independent experiments; PNPLA3^{I148M}: n = 1 clones and 4 independent experiments; PNPLA3^{KO}: n = 2 clones and 4 independent experiments) backgrounds. The basal CYP3A4 activity was not statistically different between genotypes indicating that all genotypes differentiate into HLCs with similar functionality and metabolic activity. Ordinary one-way ANOVAs with Dunnett's multiple comparisons tests were performed to test statistical significance between means. Error bars represent SEM.

In both genetic backgrounds, the PNPLA3^{KO} line trended toward lower CYP3A4 activity that the PNPLA3^{UC} or PNPLA3^{I148M} lines; however, this difference did not reach statistical significance. This trend is seen most clearly in the A1ATDR/R background. This phenomenon could be attributed to several things. First, significantly fewer differentiations were used in this line compared to FSPS13B. It is possible that with increased replicates, the difference would either reach statistical significance or equilibrate as seen in the FSPS13B line. Alternately, this difference could be biologically relevant. PNPLA3 has been linked to metabolism though its exact role has yet to be fully elucidated. Alterations in PNPLA3 (I148M variant) have also been connected to diseases of failed detoxification (ALD, drug-induced liver injury, etc.) where CYP3A4 plays a key role [139, 224, 225]. Therefore, it is conceivable that loss of PNPLA3 could impair drug metabolism pathways in undetermined ways. Given the high expression of other hepatocyte markers in the PNPLA3^{KO} line of this genetic background, it is unlikely that

differences in basal CYP3A4 activity are a reflection of reduced functionality; rather, these differences are likely a result of low statistical power, high variability, and/or phenotypic effect of PNPLA3 loss.

The PNPLA3^{UC}, PNPLA3^{I148M}, and PNPLA3^{KO} lines from FSPS13B and A1ATDR/R were differentiated into HLCs and their maturity and functionality were assessed by mRNA and protein expression of key hepatocyte markers as well as measurement of the basal CYP3A4 activity. In both genetic backgrounds, the three genotypes did not significantly differ in their expression of mature hepatocyte markers on either an mRNA or a protein level. Additionally, all lines demonstrated metabolic functionality as measured through their basal CYP3A4 activity. No statistical differences were observed in the basal CYP3A4 activity between genotypes in either genetic background, though the PNPLA3^{KO} lines trended toward lower CYP3A4 activity. This difference may be a result of low statistical power, high variability between differentiations, and or a biologically relevant phenomenon. Additional studies will be needed to determine the statistical and biological significance of this finding. Overall, these findings indicate that mutations in PNPLA3 do not affect the differentiation capacity of hiPSCs into HLCs. Any phenotypic differences between lines can now be attributed to the *PNPLA3* genotypes rather than differences in the maturity of differentiated cells.

3.4   3D culture improves maturity and functionality of HLCs

It has been shown that differentiated HLCs have a functional maturity of foetal cells; however, the maturity and functionality of these cells can be improved when placed in a 3D culture system [200, 226]. HLCs grown in 3D culture have been shown to express higher levels of mature hepatocyte markers as well as drug metabolizing enzymes. Their increased metabolic activity makes them a more appropriate tool for measuring drug toxicity as well as modelling drug induced liver injury [201, 202]. It is hypothesized that the 3D matrix allows these cells to establish proper polarization which allows them to improve their metabolic functionality. This paradigm is not only true for hiPSC-derived hepatocytes, but also primary human hepatocytes cultured *in vitro*. Various strategies for 3D culture have been explored including liver spheroid cultures, micropatterned co-cultures, and liver-on-a chip platforms [227]. One of the most common strategies, primary

human hepatocytes are aggregated into spheroids about 200 microns in diameter [203, 228]. These cells demonstrate mature hepatocyte functionality such as bile and albumin secretion as well as high expression of drug metabolizing enzymes [203, 204]. Primary human hepatocytes cultured in spheroids have increased metabolic activity and are more sensitive to hepatotoxins compared to their 2D counterparts [205]. This increased functionality can be preserved *in vitro* over a much longer timespan than traditional 2D culture methods [203, 204]. In addition to increased functionality and prolonged culture time, 3D culture methods also allow primary human hepatocytes to be co-cultured with other parenchyma and non-parenchymal liver cells. Initial studies of these co-cultures reveal that the presence of non-parenchymal liver cells further increases the functionality of primary human hepatocytes and prolongs their viability *in vitro* for several weeks [203, 229]. These 3D cultures provide the opportunity to create more physiological *in vitro* systems in order to gain valuable insight into hepatic development and disease mechanisms.

Given these findings and the importance of proper metabolic functionality to NAFLD disease modelling, we decided to test the effects of 3D culture on the maturity of our cells. Given that all the genetically edited lines differentiated to a similar extent in 2D culture, we chose to test the effects of 3D culture on PNPLA3^{UC} and PNPLA3^{KO} cells for our pilot experiments. Several 3D matrices were tested to determine which matrix supported the highest viability of the HLCs (data not shown). The chosen matrix was Matrigel Basement Membrane. While using a collagen-based matrix would have been ideal for modelling the human liver microenvironment, the protocol for splitting the HLCs and embedding them into the collagen scaffold was extensive and it was difficult to maintain high cell viability over the long duration. No such problems were observed in the Matrigel-based 3D culture system, so we chose to move forward with Matrigel as the 3D matrix for our future experiments. It remains possible that this technical difficulty was specific to the FSPS13B cell line and further experiments should be undertaken to improve the physiological relevance of this system by optimizing the 3D matrix. Since Matrigel is not chemically defined, the growth factor reduced version was used in order to reduce the risk of unwanted signalling molecules leading to the de-differentiation or trans-differentiation of the HLCs.

For 3D culture experiments, HLCs were first differentiated in 2D before being dissociated into small clumps and placed into the Matrigel matrix. The HLCs were placed in 3D culture in small clumps in order to maximize their viability and functionality. Figure 3.11 shows representative micrographs of PNPLA3^{UC} and PNPLA3^{KO} cells in 2D and 3D culture. In 2D culture, these cells adopt a classical polygonal morphology of hepatocytes. In 3D, the polygonal morphology of the HLCs is less clear because the cells form very tight spherical clusters. These clusters are dense throughout and do not contain a lumen. Once placed in 3D culture, the cells do not expand and can be maintained in 3D culture for 2-3 weeks or until the Matrigel begins to dissolve. These characteristics were consistent across genotypes.
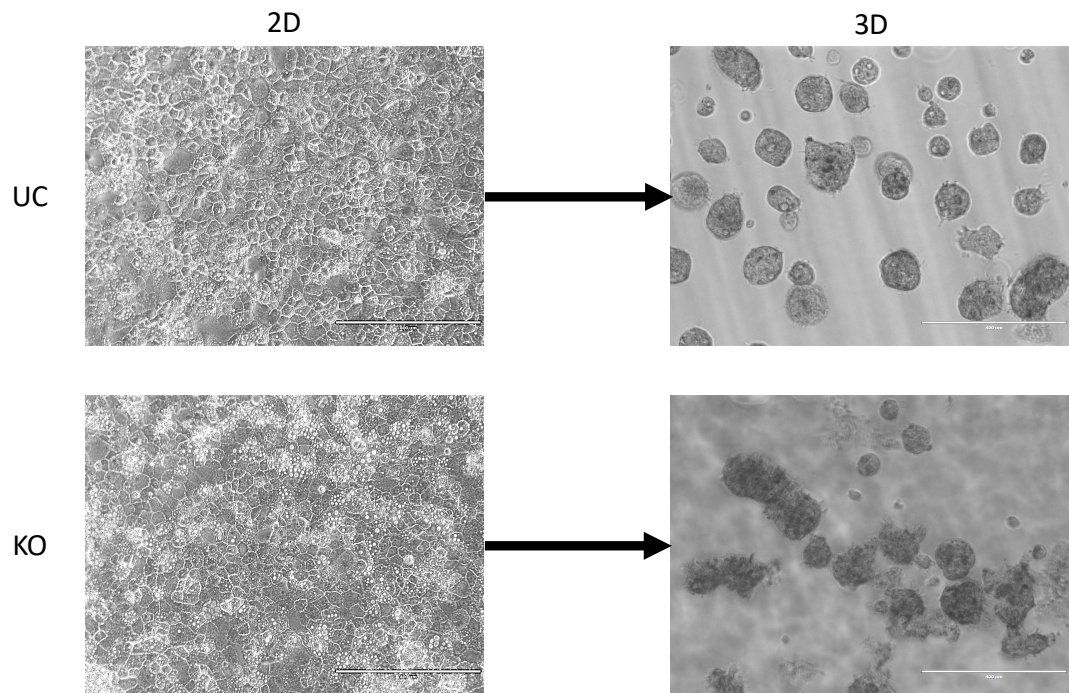
Figure 3.11 Representative micrographs of 2D and 3D cultured HLCs.

Representative micrographs of PNPLA3^{UC} and PNPLA3^{KO} HLCs cultured in 2D or 3D on day 30 of differentiation. The 2D HLCs of both genotypes have a polygonal morphology similar to hepatocytes. In 3D, the HLCs of both genotypes form small, dense, spherical clusters of cells that do not expand with additional time in culture.

Matrigel is not chemically defined and contains several growth factors including TGFβ, EGF, and FGF that could affect the differentiation status of the HLCs. In order to ensure that 3D culture in this matrix did not negatively affect the differentiation of HLCs, we

examined the expression of the hepatocyte markers used to characterize the 2D HLCs. In Figure 3.12a, the mRNA expression of *A1AT, ALB, HNF4a*, and *TTR* was quantified for PNPLA3^{UC} and PNPLA3^{KO} cells grown in either 2D or 3D. For each genotype, the mRNA expression was normalized to cells grown in 2D. This analysis indicated that 3D culture leads to increased expression of hepatocyte markers. 3D culture had the most profound effect on *A1AT, TTR,* and *PNPLA3*, leading to a more than four-fold increase in expression. *ALB* expression increased two- to three-fold following placement in 3D culture. *HNF4a* was more impervious to the effects of 3D culture with a modest two-fold increase in expression that did not reach statistical significance. These data indicate that Matrigel does not negatively affect the differentiation status in our HLCs, in fact, 3D culture may lead to enhanced maturity as has been previously demonstrated. This trend held true for both PNPLA3^{UC} and PNPLA3^{KO} cells. Therefore, the signalling environment of the Matrigel does not differentially affect the two genotypes. To confirm these results on the protein level, ALB and HNF4a protein expression were evaluated using immunofluorescent microscopy (Figure 3.12b). Both PNPLA3^{UC} and PNPLA3^{KO} cells showed homogeneous expression of ALB and HNF4a in 3D culture confirming that this culture system does not negatively affect the differentiation of HLCs.

A.
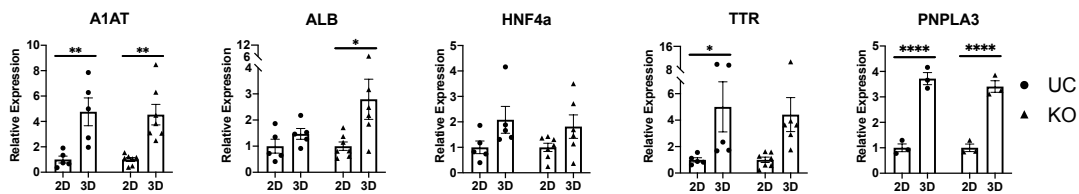
B.
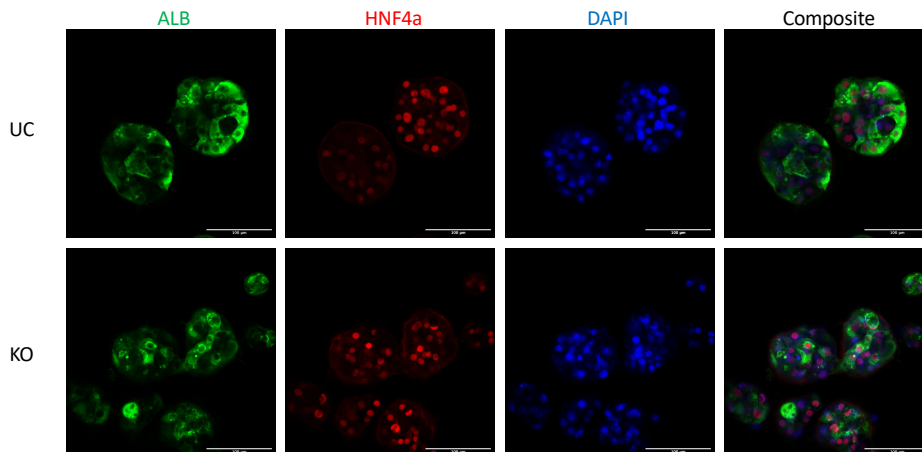
Figure 3.12 Comparison of differentiation marker expression between 2D and 3D cultured HLCs.

A. mRNA expression of hepatocyte markers *A1AT, ALB, HNF4a, TTR,* and *PNPLA3* for PNPLA3^{UC} and PNPLA3^{KO} cells cultured in 2D or 3D conditions following 30 days of differentiation (PNPLA3^{UC}: n = 3 clones and 3 independent experiments; PNPLA3^{KO}: n = 3 clones and 4 independent experiments). All values are normalized to the 2D conditions. 3D culture increases expression of all five hepatocyte markers in both genotypes indicating that 3D culture may improve maturity of HLCs. Ordinary one-way ANOVAs with Dunnett's multiple comparisons tests were performed to test statistical significance between means. Error bars represent SEM. **B.** Representative images of ALB and HNF4a protein staining in HLCs cultured in 3D conditions shows that HLCs maintain homogenous expression of hepatocyte markers in these culture conditions. This experiment was performed 2 times and the best representative images are represented in this figure.

PNPLA3^{UC} and PNPLA3^{KO} cells were differentiated in 2D before being placed in 3D culture in Matrigel. The 3D cultured HLCs formed small, dense, spherical clumps of cells that did not expand and could be cultured for an additional 2-3 weeks following embedment. The 3D cultured cells maintained their expression of key hepatocyte markers at the mRNA and protein levels. The 3D culture enhanced the mRNA expression of these markers in both PNPLA3^{UC} and PNPLA3^{KO} cells. In addition to the mature hepatocyte markers, 3D culture also increased the expression of *PNPLA3* which may exacerbate the phenotypic differences between PNPLA3^{UC}, PNPLA3^{I148M}, and PNPLA3^{KO} HLCs. These results indicate that 3D culture of HLCs in Matrigel does not negatively affect the differentiation status of these cells. Given these results and previously published work from our lab that indicates 3D culture improves the maturity and metabolic functionality of HLCs, we chose to use 3D culture for all disease modelling experiments moving forward [200].

3.5   Conclusions

In this chapter, we used CRISPR/CAS9 to generate isogenic PNPLA3^{UC}, PNPLA3^{I148M}, and PNPLA3^{KO} lines in two different genetic backgrounds. We demonstrated that PNPLA3^{UC} and PNPLA3^{I148M} cells expressed high levels of PNPLA3 while the PNPLA3^{KO} cells did not express detectable levels of PNPLA3 indicating a complete knock-out of the gene. These lines were then differentiated into HLCs to ensure that neither the editing process nor the *PNPLA3* genotype affected the differentiation capacity of the hiPSCs. PNPLA3^{UC}, PNPLA3^{I148M}, and PNPLA3^{KO} cells expressed similarly high levels of hepatocyte markers on the mRNA and protein level indicating that all three

genotypes retained their differentiation potential. There were no statistical differences in metabolic functionality as measured by basal CYP3A4 activity. We then tested the effects of 3D culture on HLC maturity. PNPLA3^{UC} and PNPLA3^{KO} cells were placed in 3D culture in Matrigel. This 3D culture system resulted in increased expression of hepatocyte markers indicating that Matrigel does not negatively affect the differentiation status of HLCs and this 3D culture may even promote enhanced maturity. Given that all three genotypes express the expected levels of PNPLA3 and demonstrate similar differentiation capacity toward HLCs, these lines were deemed suitable for disease modelling of PNPLA3-induced NAFLD.

4 ALTERATIONS TO PNPLA3 IN HEPATOCYTES LEADS TO INCREASED LIPID ACCUMULATION AND ALTERED RESPONSE TO LIPID-INDUCED STRESS

4.1 Introduction

At its core, NAFLD is a disease of disrupted fatty acid homeostasis. When the concentration of incoming fatty acids overwhelms a hepatocyte's ability to properly store or metabolize those fatty acids, steatosis and eventually lipotoxicity result. Fatty acids are biologically active metabolites that have many functions ranging from energy source to signalling molecule [230]. Free fatty acids have several possible fates upon entering a hepatocyte. These fatty acids can be exported to other tissues such as adipose tissue via VLDL secretion, esterified into triglycerides for storage, or metabolized to produce energy through β-oxidation.

The first line of defence a hepatocyte has against increased fatty acid influx is fatty acid efflux. When the lipid concentration in a hepatocyte gets too high, the cell upregulates VLDL secretion in order to re-establish homeostasis. The capacity for fatty acid efflux via VLDL secretion is extremely high, but the process is saturable. It has also been shown that hepatocellular stress such as that caused by NASH reduces the VLDL export capacity by reducing the production of APOB-100 [64]. When the concentration of fatty acids exceeds the export capacity of the hepatocyte, the remaining fatty acids must either be stored or metabolized, each of which has a differential effect on the viability of the cell.

The esterification of free fatty acids into triglycerides is generally viewed as a cytoprotective process. Triglycerides are metabolically inert molecules that can be stored in lipid droplets indefinitely with little or no effect on the health of a hepatocyte. When faced with a large burden of metabolically harmful free fatty acids, hepatocytes shunt these fatty acids into triglyceride storage [35, 231]. If the increased lipid flux continues for a prolonged period of time, steatosis develops. For this reason, there remains significant debate in the research community about whether steatosis is a pathogenic step in NAFLD progression or if it is merely an epiphenomenon that results from hepatocytes' attempt to buffer themselves from increased fatty acid flux [232, 233].

Alternately, free fatty acids that are metabolized by β-oxidation or other metabolic pathways can cause significant damage to the cell. The process of β-oxidation generates ROS. At normal physiological levels, these ROS are neutralized by the antioxidant

defence systems in the cell. However, when faced with pathological levels of free fatty acids, β-oxidation and other metabolic processes are starkly upregulated to address the increased lipid burden. This upregulation of metabolism causes increased ROS production which overwhelms the antioxidant capacity of hepatocytes. The un-neutralized ROS then wreak havoc on the cell by causing lipid peroxidation, ER stress, and mitochondrial dysfunction. If this insult persists and cannot be controlled by cell defence pathways such as the unfolded protein response, the hepatocyte will succumb to lipotoxicity [64, 79, 80]. This lipoapoptosis results in the release of pro-inflammatory and pro-fibrotic mediators that facilitate the progression from simple steatosis to NASH.

The PNPLA3 protein has been hypothesized to play a role in each of the aforementioned aspects of fatty acid homeostasis. It is believed that the pathogenic effects of the I148M variant result from disruption to at least one of these processes. It has been demonstrated previously that patients homozygous for the risk allele have lower circulating levels of lipoproteins and *in vitro* studies have suggested that this phenomenon is a result of reduced VLDL export in these patients [122]. It is possible that in its role as a lipid droplet remodelling protein, PNPLA3 may sequester triglycerides in lipid droplets and thereby prevent their export in VLDL which in turn exacerbates steatosis. PNPLA3 has also been shown to have LPAAT activity and actively participate in the formation of triglycerides. The I148M variant is hypothesized to increase this function thereby causing steatosis through increased generation of triglycerides [118]. However, the veracity of these claims has recently been called into question due to concerns regarding the experimental approach and the inability to confirm these findings *in vivo* [111]. Finally, the most well-established role for PNPLA3 is its lipolysis activity. Lipolysis is one of the first steps in the β-oxidation of triglycerides thus connecting PNPLA3 to each fatty acid disposal pathway within hepatocytes. The I148M variant causes a loss of lipolytic activity which causes increased triglyceride accumulation and steatosis [109, 121, 164]. Each of these proposed functional roles for PNPLA3 result in the sequestration of neutral lipids rather than increased metabolic activity which can result in hepatocellular damage. This sequestration of triglycerides by reduced VLDL export, increased triglyceride formation, or decreased triglyceride lipolysis offers mechanistic insight into how the I148M variant

induces steatosis; however, these mechanisms fail to explain the progressive nature of PNPLA3-induced NAFLD.

Given the biological significance of free fatty acids in the pathogenesis of NAFLD and the role of PNPLA3 in the fate of hepatic free fatty acids, we chose to use palmitic acid and oleic acid in our *in vitro* model of PNPLA3-induced NAFLD. Palmitic acid is a saturated fatty acid (SFA) while oleic acid is a monounsaturated fatty acid (MUFA). SFAs are metabolically destructive lipid species because they are primed for entry to several metabolic pathways that generate ROS including β-oxidation and are used to synthesize other damaging lipid species such as LPC and ceramides [80, 81]. Additionally, when high concentrations of SFAs are incorporated into phospholipid membranes, the fluidity of the membranes is reduced resulting in cell stress and apoptosis. Alternately, MUFAs are almost exclusively incorporated into triglycerides rather than participating in metabolic pathways. Cells treated with high concentrations of MUFAs develop steatosis without suffering any loss of viability [234]. Given their differential roles in lipid metabolism, we chose to use oleic acid and palmitic acid to model two aspects of NAFLD disease progression: steatosis and lipotoxicity, respectively. In order to model NAFLD *in vitro*, we sought to use a physiological concentration of free fatty acids in our system. However, clinical data on the reference ranges of plasma fatty acid concentrations are extremely limited. The concentration of free fatty acids in the serum varies extensively and can be influenced by the feeding/fasting cycle as well as conditions such as obesity and type 2 diabetes. In addition, it is difficult to quantify the concentration of free fatty acids that hepatocytes are directly exposed to as collecting blood samples from the portal vein is quite invasive. A recent profile of plasma fatty acid concentrations in healthy adults found that the concentration of palmitic acid varied from 0.3 to 4.1 mM while the concentration of oleic acid varied from 0.03 to 3.2 mM [235]. Similarly, studies have found that NAFLD patients have elevated plasma levels of both palmitic and oleic acid [236]. Given this clinical data, as well as previous *in vitro* studies using the aforementioned fatty acids, we chose to use a fatty acid concentration of 250 μM for our experiments. In order to facilitate fatty acid uptake by our HLCs, the free fatty acids were first conjugated to BSA at a molar ratio of 1.66:1 FFA to BSA.

In this chapter, we generated an *in vitro* model to better understand the role of PNPLA3 in NAFLD development and progression. First, PNPLA3^{UC}, PNPLA3^{I148M}, and PNPLA3^{KO} cells were differentiated into HLCs and placed in 3D culture. The lines were then treated with either control media or media supplemented with 250 μM oleic acid or 250 μM palmitic acid to induce a steatotic or lipotoxic phenotype, respectively. The cells were treated with the free fatty acids for up to one week and the phenotypic differences between lines were assessed. We began by quantifying the lipid accumulation in each of the lines using immunofluorescence microscopy and flow cytometry. We then assessed the reaction of each genotype to lipid induced stress by measuring viability and expression of ER stress markers. Finally, we examined the question of whether the I148M variant is a loss of function or gain of function variant using transcriptomic analyses.

4.2   PNPLA3 edited cells accumulate more lipids than untargeted control cells

The I148M variant of PNPLA3 has been robustly correlated with increased hepatic lipid accumulation [91, 123, 134]. In order to assess the effect of *PNPLA3* genotype on lipid accumulation in our system, we began by differentiating PNPLA3^{UC}, PNPLA3^{I148M}, and PNPLA3^{KO}, cells into HLCs and placing them in 3D culture. The HLCs were then treated for one week with control media or media supplemented with fatty acids. The media used for these experiments had a high concentration of glucose and its derivatives in order to stimulate de novo lipogenesis and increase lipid accumulation to the fullest extent possible. Following the one week of treatment, lipid accumulation was assessed qualitatively using immunofluorescence microscopy and quantitatively using flow cytometry. We used Bodipy to stain the neutral lipids in the cells and visualize the lipid droplets. We performed each of these experiments on both the FSPS13B and A1ADR/R cell lines.

Based upon the characteristics of the fatty acids, we expected that, in PNPLA3^{UC} cells, the oleic acid would induce profound steatosis while the palmitic acid would result in only a small amount of lipid accumulation due to its cytotoxicity. Based upon the clinical evidence, we hypothesized that PNPLA3^{I148M} cells would accumulate more lipid droplets than PNPLA3^{UC} cells and be more susceptible to lipid-induced stress. Due to the lack of

phenotype seen in mice with a complete knock-out of the Pnpla3 protein, we hypothesized that the PNPLA3^{KO} cells would mimic the PNPLA3^{UC} cells phenotypically.

Figures 4.1 shows representative micrographs of the PNPLA3^{UC}, PNPLA3^{I148M}, and PNPLA3^{KO} lines treated with control medium, oleic acid, or palmitic acid in the FSPS13B background. These images show that PNPLA3^{UC} cells behaved exactly as hypothesized. These cells accumulated a small number of lipid droplets under basal conditions. The oleic acid induced a stark steatotic phenotype with nearly all of the cytoplasmic space being occupied by microstetotic lipid droplets. The palmitic acid treated cells accumulated a very small amount of lipid droplets, much more closely resembling the control treated cells rather than the oleic acid treated cells. As expected, the PNPLA3^{I148M} cells accumulated more lipid droplets in all treatment groups than their PNPLA3^{UC} counterparts. Interestingly, the PNPLA3^{I148M} cells were extremely steatotic when treated with both oleic acid and palmitic acid. Strikingly, contrary to our original hypothesis, the PNPLA3^{KO} cells accumulated the most lipid droplets of all the genotypes regardless of treatment type. The PNPLA3^{KO} line had the highest basal level of lipid accumulation and followed the same phenotypic pattern of response as PNPLA3^{I148M} cells to oleic acid and palmitic acid treatment. Notably, the lipid droplets in PNPLA3^{KO} cells appeared to be qualitatively larger than PNPLA3^{UC} cells. This phenomenon has been noted before in *in vitro* overexpression models of the PNPLA3 I148M variant [109].
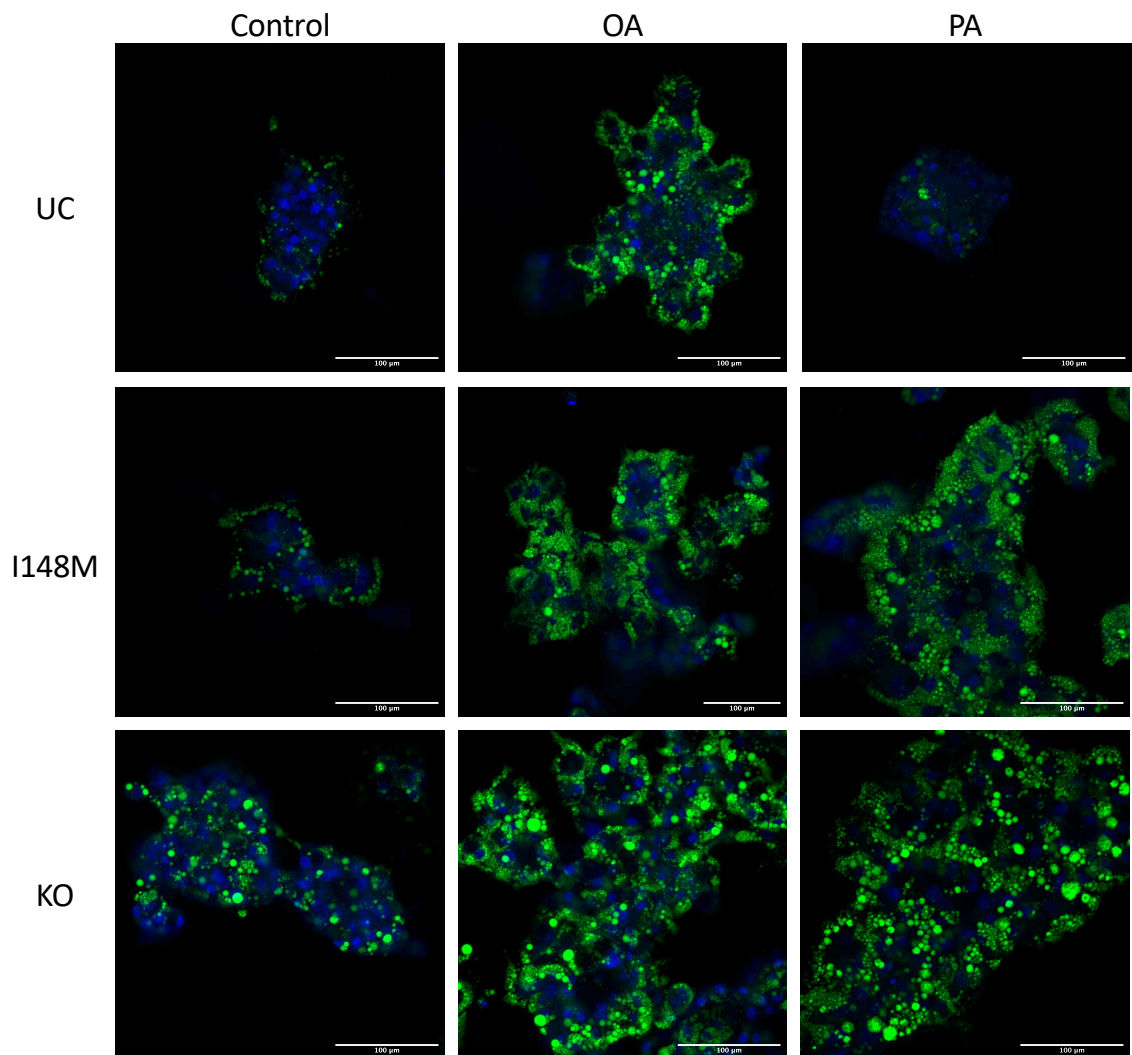
Figure 4.1 Lipid droplet accumulation in FSPS13B lines.

Representative images of bodipy stained lipid droplets in cells from the three genotypes. Regardless of treatment type, progressively more lipids were accumulated in PNPLA3^{I148M} cells and PNPLA3^{KO} cells than PNPLA3^{UC} cells. This experiment was performed 6 times and the best representative images are represented in this figure.

Both the PNPLA3^{I148M} and PNPLA3^{KO} lines accumulated more lipid droplets than PNPLA3^{UC} cells and failed to metabolically differentiate between SFAs and MUFAs when esterifying triglycerides and forming lipid droplets. This failure to differentiate between the metabolic roles of the two classes of fatty acids leads to increased triglyceride storage and steatosis. These results, in stark contrast with the results from murine knock-out experiments, indicate that the loss of PNPLA3 protein causes enhanced steatosis in

human hepatocytes. These data argue that the pathogenic mechanism of the I148M variant is facilitated through a loss of function.
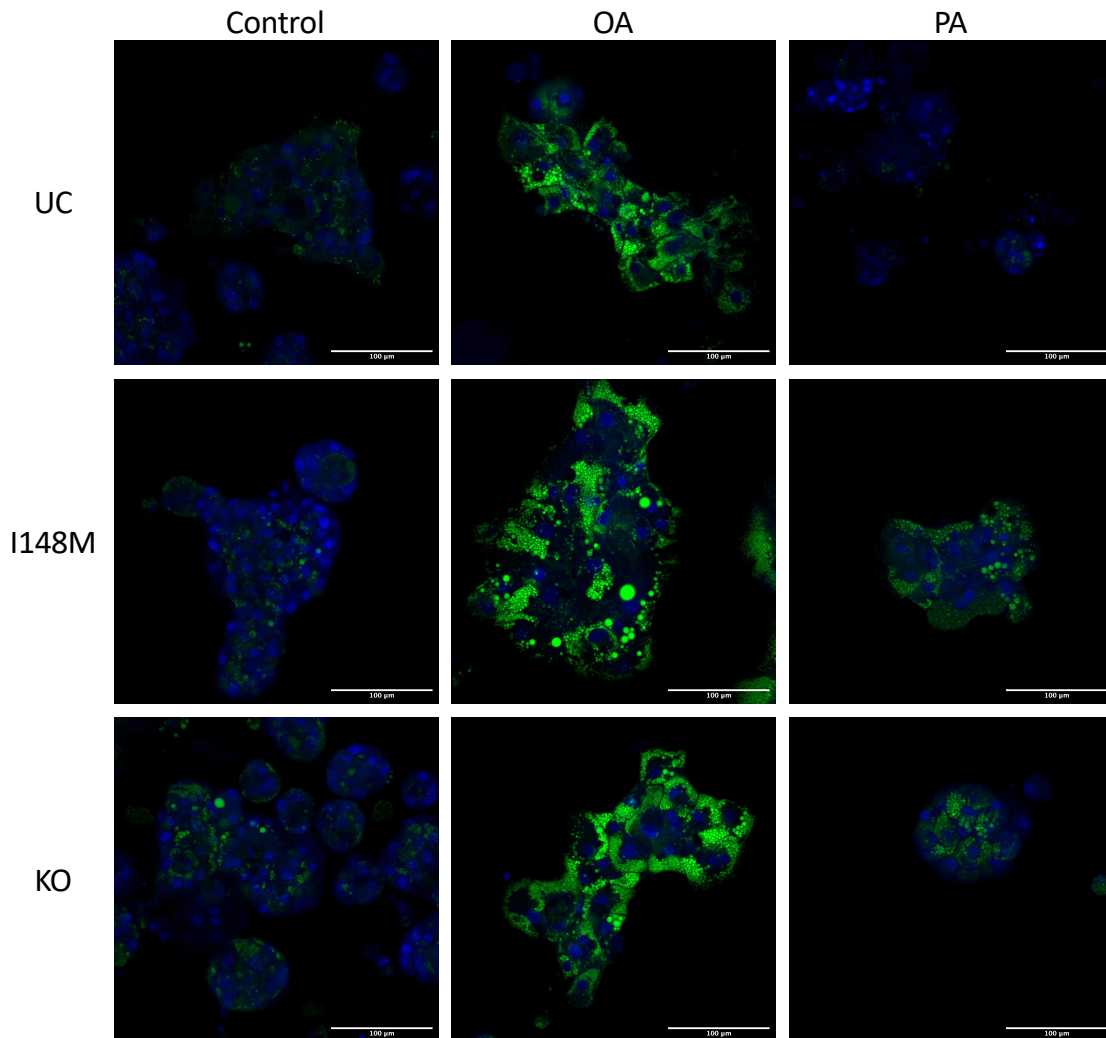
Figure 4.2 Lipid droplet accumulation in A1ATDR/R lines.

Representative images of bodipy stained lipid droplets in cells from the three genotypes. Regardless of treatment type, progressively more lipids were accumulated in PNPLA3^{I148M} cells and PNPLA3^{KO} cells than PNPLA3^{UC} cells. This experiment was performed 1 time and the best representative images are represented in this figure.

In order to confirm these findings, we performed the same experiment in the A1ATDR/R lines (Figure 4.2). At basal levels, the A1ATDR/R lines appeared to accumulate fewer lipid droplets than their FSPS13B counterparts, regardless of genotype. However, the phenotypic pattern of lipid accumulation in this genetic background was consistent with the results from the previous experiment for all three genotypes. The PNPLA3^{UC} cells

had a very low amount of basal lipid accumulation, exhibited a steatotic phenotype when treated with oleic acid, and nearly undetectable levels of lipid accumulation when treated with palmitic acid. Both the PNPLA3^{I148M} and PNPLA3^{KO} lines had higher basal lipid accumulation and steatotic phenotypes when treated with both oleic acid and palmitic acid. In this line, the palmitic acid appears to induce slightly less lipid accumulation than oleic acid indicating that this genetic background may be slightly more proficient at differentiating between the saturation status of different fatty acids in the absence of PNPLA3. Additionally, the PNPLA3^{I148M} cells accumulated significantly larger lipid droplets than the PNPLA3^{UC} and PNPLA3^{KO} lines in this genetic background. Overall, the data from the A1ATDR/R lines confirms the findings in the FSPS13B line. Namely, loss of PNPLA3 function causes steatosis and increased lipid droplet size. This steatosis is facilitated, at least in part, by the failure to differentiate between SFAs and MUFAs when partitioning fatty acids into triglyceride storage.

In order to confirm these qualitative observations, we used flow cytometry to quantify the differences in lipid accumulation between genotypes. Unfortunately, we faced technical difficulties in removing the cells from the 3D matrix and dissociating them while maintaining a high cell viability. Therefore, we chose to perform flow cytometric analyses on cells treated with fatty acids in 2D. These analyses confirmed the microscopy results. As before, the FSPS13B lines had higher basal lipid accumulation than the A1ATDR/R lines. In both the FSPS13B and A1ATDR/R lines, there was a stepwise increase in lipid accumulation from PNPLA3^{UC} cells to PNPLA3^{I148M} cells to PNPLA3^{KO} cells in all three treatment groups (Figure 4.3). This analysis revealed that the PNPLA3^{I148M} cells possess an intermediate phenotype between PNPLA3^{UC} cells and PNPLA3^{KO} cells as would be expected for an incomplete loss of function variant. In the A1ATDR/R lines, the PNPLA3^{I148M} and PNPLA3^{KO} lines accumulated significantly more lipids than the PNPLA3^{UC} cells in all three treatment groups. In the FSPS13B line, PNPLA3^{KO} cells accumulated significantly more lipid droplets that PNPLA3^{UC} lines in all treatment groups while the differences between PNPLA3^{I148M} and PNPLA3^{UC} only reached statistical significance in the PA treatment group. Indicating once again that saturated fatty acids induce the most profound phenotype when PNPLA3 protein function is lost.
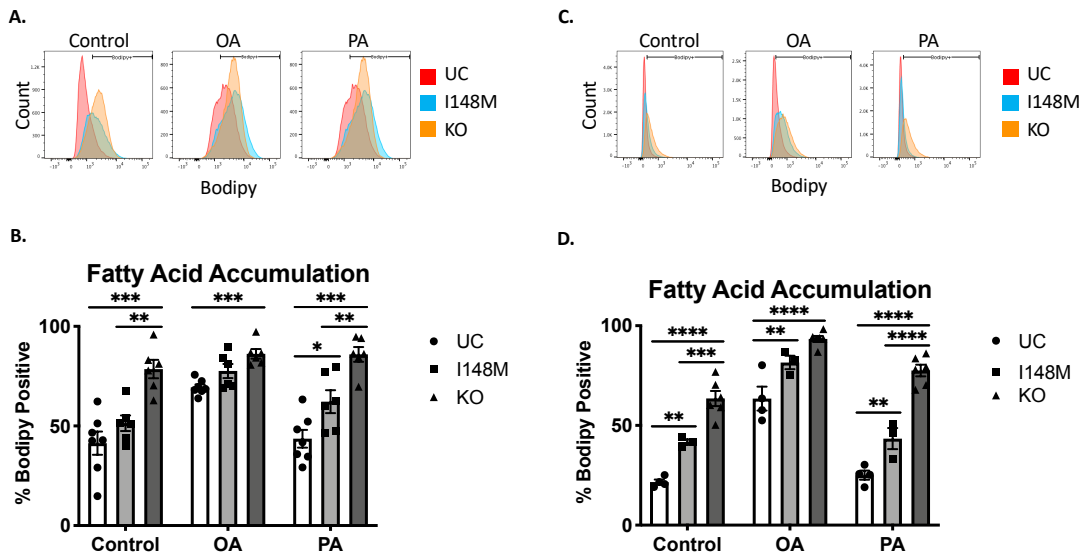
Figure 4.3 Flow cytometric analysis of lipid accumulation.

A. Representative flow cytometric analysis of lipid accumulation via bodipy staining in FSPS13B cell lines. **B.** Quantification of the percentage of bodipy positive cells in A (PNPLA3^{UC}: n = 2 clones and 5 independent experiments; PNPLA3^{I148M}: n = 2 clones and 5 independent experiments; PNPLA3^{KO}: n = 3 clones and 2 independent experiments). There was a stepwise increase in lipid accumulation from PNPLA3^{UC} to PNPLA3^{I148M} to PNPLA3^{KO} cells. **C.** Representative flow cytometric analysis of lipid accumulation via bodipy staining in A1ATDR/R cell lines. **D.** Quantification of the percentage of bodipy positive cells in C (PNPLA3^{UC}: n = 2 clones and 4 independent experiments; PNPLA3^{I148M}: n = 1 clones and 3 independent experiments; PNPLA3^{KO}: n = 2 clones and 4 independent experiments). These results are consistent with the FSPS13B line as well as the disease phenotype. Ordinary one-way ANOVAs with Dunnett's multiple comparisons tests were performed to test statistical significance between means. Error bars represent SEM.

Overall, we found that PNPLA3^{I148M} and PNPLA3^{KO} cells accumulate more, larger lipid droplets than PNPLA3^{UC} cells. The PNPLA3^{I148M} cells appear to have an intermediate phenotype between PNPLA3^{UC} and PNPLA3^{KO} cells indicating that the I148M variant may cause steatosis through a loss of enzymatic function. Additionally, loss of PNPLA3 function results in increased steatosis regardless of treatment type. PNPLA3^{I148M} and PNPLA3^{KO} cells appear to lack the ability to differentiate between SFAs and MUFAs when partitioning fatty acids into triglyceride storage. Given that metabolism of SFAs has profound effects on the health of hepatocytes, this lack of discrimination could have profound effects on the metabolism and viability of these cells.

Samantha G. Tilson - January 2021

4.3 PNPLA3^{KO} cells are resistant to palmitic acid-induced lipotoxicity

Given the unexpected steatotic effect of palmitic acid in the PNPLA3^{I148M} and PNPLA3^{KO} cell lines, we wanted to further evaluate how PNPLA3 edited cells may react differently to lipid induced stress. We began by differentiating and treating the cells in the same manner as the previous experiments. Following the one week of lipid treatment, the viability of the cells was assessed using the Presto Blue cell viability assay. Additionally, we examined the expression of ER stress markers after 48 hours of lipid treatment in each genotype. We chose the 48-hour timepoint because at that time, cells had accumulated large amounts of lipid droplets but there was not yet a significant effect on viability (data not shown). Therefore, we assumed this timepoint would be ideal for evaluating the reaction of each genotype to lipid induced stress.

The results of the viability assessment for both the FSPS13B lines and the A1ATDR/R lines are shown in Figure 4.4. The viability of the cells was normalized to the control-treated cells in each of the respective genotypes. In the FSPS13B background, the PNPLA3^{UC} line performed exactly as expected. When treated with oleic acid, there was minimal effect on viability, but palmitic acid treatment induced a stark reduction in cell viability to approximately 20% of control-treated cells. This loss of viability was severely blunted in the PNPLA3^{I148M} and PNPLA3^{KO} lines. In the PNPLA3^{KO} line, palmitic acid treatment caused very little cell death with approximately 80% of cells remaining viable following treatment. Similarly, in the PNPLA3^{I148M} line, approximately 60% of cells remained viable following palmitic acid treatment. This line experienced much more variability in viability between experiments. The susceptibility of PNPLA3^{I148M} lines to palmitic acid-induced lipotoxicity may be dependent on the level of enzymatic activity that is retained in the I148M protein.

These results were confirmed in the A1ATDR/R line. Generally, the A1ATDR/R line was more resistant to palmitic acid-induced lipotoxicity. This is likely due to the genotype of the A1ATDR/R line which is naturally heterozygous for the I148M variant. For this reason, the PNPLA3^{UC} line did see a significant reduction in viability following palmitic acid treatment but this reduction was equivalent to that seen in the PNPLA3^{I148M} cells. In

both PNPLA3^{UC} and PNPLA3^{I148M} lines, the cell viability was reduced to approximately 60% of control-treated cells. The PNPLA3^{KO} line remained resistant to this palmitic acid-induced lipotoxicity. This result calls into question whether the A1ATDR/R line was the appropriate confirmation line to choose for such experiments. Patients that are heterozygous for the PNPLA3 risk allele are at increased risk of developing NAFLD albeit less risk than their homozygous counterparts. Thus, even the presence of one risk allele may alter the metabolic phenotype of hepatocytes. The heterozygous nature of the A1ATDR/R line makes it difficult to determine if phenotypic differences between FS13B and A1ATDR/R are driven by the presence of one risk allele or some other difference in their genetic backgrounds. However, since the I148M variant exerts its pathogenicity in an additive manner, we do believe that some insight may be gained from experiments with the A1ATDR/R lines. In each of the experiments that we performed, the FS13B and A1ATDR/R lines mirrored one another showing similar trends between PNPLA3^{UC}, PNPLA3^{I148M}, and PNPLA3^{KO} HLCs with only minor differences in statistical significance in a few experiments. Therefore, despite the heterozygous genotype of the PNPLA3^{UC} cells in this genetic background, we determined that the A1ATDR/R lines were sufficient to use as a confirmatory cell line for the purposes of this thesis. However, in the future, additional efforts should be undertaken in order to confirm these results in another hiPSC line that is homozygous for the PNPLA3 reference allele.
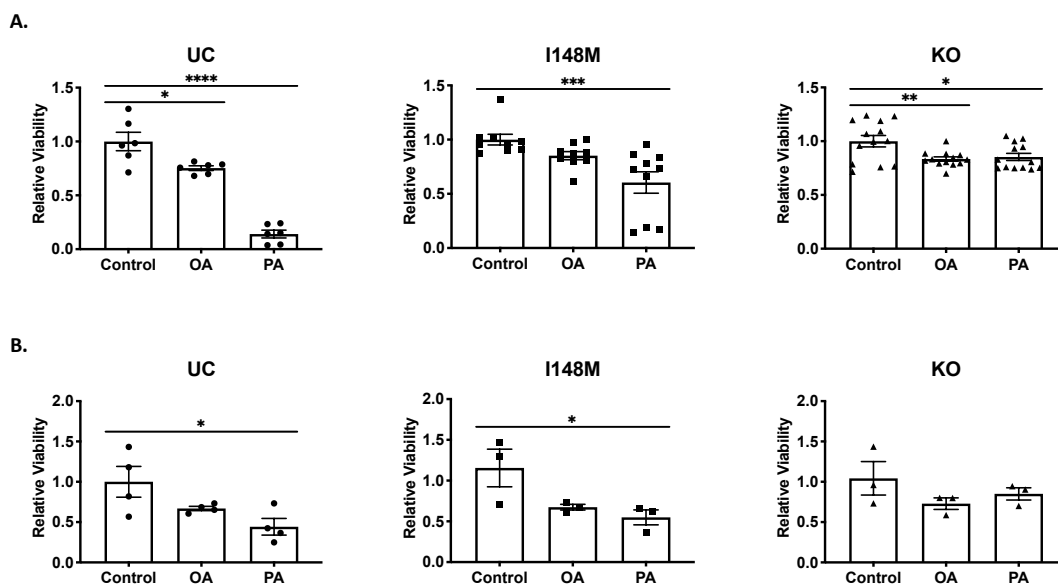
Figure 4.4 Viability of fatty acid treated cells.

Relative viability of FFA treated cells to control cells within each genotype for the **A.** FSPS13B genetic background (PNPLA3^{UC}: n = 2 clones and 3 independent experiments; PNPLA3^{I148M}: n = 2 clones and 7 independent experiments; PNPLA3^{KO}: n = 3 clones and 5 independent experiments) and the **B.** A1ATDR/R genetic background (PNPLA3^{UC}: n = 2 clones and 3 independent experiments; PNPLA3^{I148M}: n = 1 clones and 4 independent experiments; PNPLA3^{KO}: n = 2 clones and 3 independent experiments). PNPLA3^{UC} cells are susceptible to palmitic acid lipotoxicity while PNPLA3^{KO} cells are resistant to the lipotoxic insult of palmitic acid. Ordinary one-way ANOVAs with Dunnett's multiple comparisons tests were performed to test statistical significance between means. Error bars represent SEM.

These data offer an explanation for the increased steatosis observed in PNPLA3^{I148M} and PNPLA3^{KO} lines treated with palmitic acid. In these cell types, palmitic acid appears to be partitioned to triglyceride storage rather than metabolic pathways which results in a decreased lipotoxicity and increased steatosis. Once again, the PNPLA3^{KO} cells demonstrated the most profound phenotype while the PNPLA3^{I148M} cells had an intermediate phenotype between PNPLA3^{UC} and PNPLA3^{KO} cells. There was a similar viability pattern in cells that were heterozygous or homozygous for the I148M variant. This may tentatively indicate that a single risk allele is sufficient to reduce the lipotoxicity of palmitic acid in hepatocytes; however, additional studies are needed to assess this fully.

Previous studies have demonstrated that treatment with palmitic acid leads to an accumulation of di-saturated glycerolipids in the ER membrane [84-86]. The decreased membrane fluidity triggers activation of the unfolded protein response via to IRE1α and PERK pathways. Saturated ER membranes have altered morphology and integrity which disturbs ER proteostasis. The activation of PERK signalling causes the phosphorylation of eIF2α which leads to reduced protein synthesis in an effort to re-establish homeostasis [80]. If the lipotoxic stress persists and the unfolded protein response is unable to re-establish homeostasis, the integrated stress response, regulated by PERK and eIF2α, leads to the transcription of CHOP and GADD34 and the induction of apoptosis [79]. In patients with NAFLD and NASH, PERK expression is starkly upregulated. In addition, the expression of BIP, an ER chaperone protein that plays a role in regulating ER stress, is elevated in patients with NASH. There are conflicting reports regarding the expression level of other downstream activators such as CHOP and GADD34 in these patients [79].

In order to further explore the differential response to lipid-induced stress, we examined markers of ER stress and lipotoxicity. We performed these experiments on the FSPS13B cell lines due to the profound differences between the PNPLA3^{UC} and PNPLA3^{KO} lines. Following 48 hours of lipid treatment, we measured mRNA expression of *BIP, GADD34, CHOP*, and *PERK* which are all markers of ER stress and the unfolded protein response (Figure 4.5). In all three genotypes, there was little to no induction of these markers by either control or oleic acid treatment. However, significant differences were observed between genotypes in the palmitic acid treated group. In PNPLA3^{UC} cells, palmitic acid treatment strongly induced all of the cell stress markers, as anticipated. However, the PNPLA3^{I148M} and PNPLA3^{KO} lines largely failed to upregulate these markers. PNPLA3^{KO} cells expressed significantly lower levels of *BIP, GADD34*, and *CHOP* than PNPLA3^{UC} cells while PNPLA3^{I148M} cells expressed significantly lower levels of *BIP, CHOP*, and *PERK*. No significant differences were observed between PNPLA3^{I148M} and PNPLA3^{KO} cells. The failure of these cells to upregulate cell stress markers correlates nicely with the viability data indicating that loss of PNPLA3 function reduces the lipotoxic effects of palmitic acid.
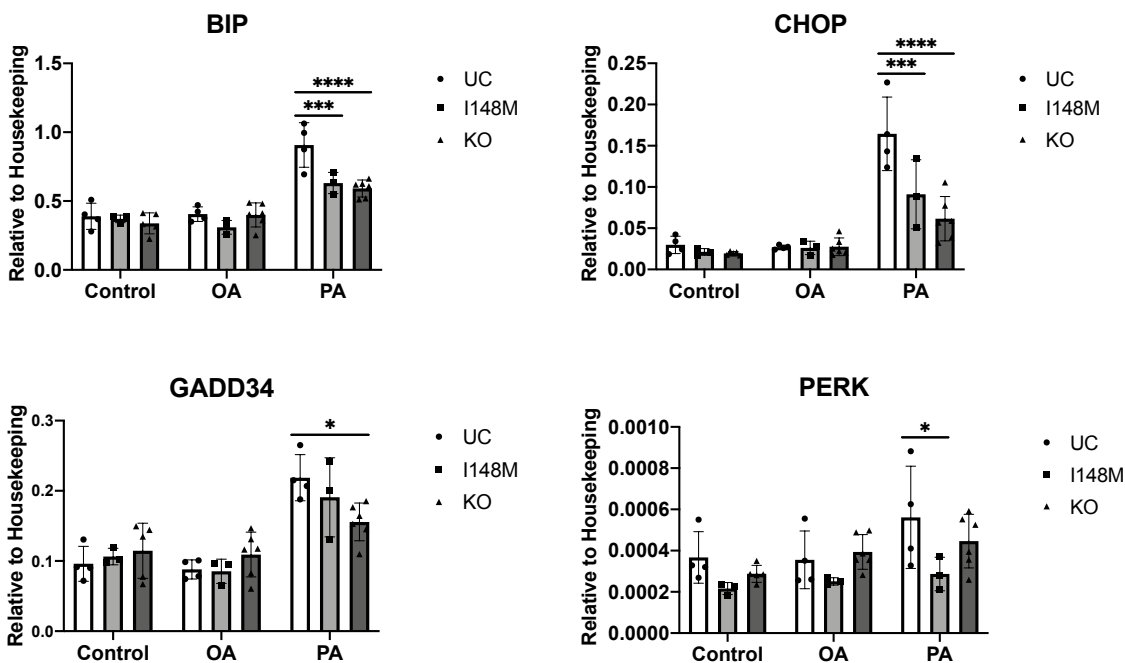
Figure 4.5 Expression of ER stress markers in response to lipid-induced stress.

mRNA expression of ER stress markers *BIP, GADD34, CHOP*, and *PERK* following 48 hours of FFA treatment in cell lines from the FSPS13B genetic background (PNPLA3^{UC}: n = 2 clones and 2 independent

experiments; PNPLA3^{I148M}: n = 2 clones and 2 independent experiments; PNPLA3^{KO}: n = 3 clones and 2 independent experiments). In PNPLA3^{UC} cells, ER stress markers are induced by palmitic acid treatment. PNPLA3^{I148M} and PNPLA3^{KO} cells are resistant to the induction of ER stress markers following palmitic acid treatment. Ordinary one-way ANOVAs with Dunnett's multiple comparisons tests were performed to test statistical significance between means. Error bars represent SEM.

As with the lipid accumulation experiments, analysis of the phenotypic effects of lipid induced stress on these lines revealed stark genotypic differences. We found that the PNPLA3^{UC} lines upregulated ER stress markers and had significantly lower viability when treated with palmitic acid. These results are consistent with lipotoxicity caused by treatment with an abundance of SFAs. Surprisingly, the PNPLA3^{I148M} and PNPLA3^{KO} lines were resistant to this lipotoxicity. These cells largely failed to upregulate markers of cell stress and the palmitic acid treatment did not markedly decreased viability, especially in the PNPLA3^{KO} lines. These results were initially perplexing given the clinical correlation between the I148M variant and severe NASH. We had expected that the cell lines carrying the I148M variant would not only accumulate more lipids, but this increased lipid storage would lead to worsening lipotoxicity that could be used as an *in vitro* surrogate for the severe liver damage seen in NASH patients who carry the I148M variant. However, our results show the opposite trend. We found that loss of PNPLA3 function either through the I148M variant or a complete knock-out of the PNPLA3 protein was cytoprotective against lipid induced stress.

This cytoprotective effect could be a result of several factors. First, these results could indicate that the metabolism of the HLCs is not mature enough to properly model the lipotoxic effects of saturated fatty acids on hepatocytes. However, this seems unlikely given that the PNPLA3^{UC} cells were susceptible to this lipotoxic stressor. Second, it is possible that this cytoprotective effect is a result of off-target effects from the CRISPR/CAS9 editing process. This possibility is equally unlikely given that we used several isogenic lines for each genotype and observed the same trend in two different genetic backgrounds. Third, these results argue that the increased steatosis seen in carriers of the I148M variant may be an epiphenomenon that does not play a role in disease progression. The I148M variant has been clinically correlated to increased inflammation and fibrosis; however, carriers of the risk allele rarely have markers of hepatocellular

damage such as hepatocyte ballooning [132, 151]. Since PNPLA3 is expressed highly in human HSCs and the functionality of these cells is also influenced by the I148M variant, it is possible that the I148M variant has divergent pathological effects on hepatocytes and stellate cells. In hepatocytes, the I148M variant interferes with lipid metabolism causing steatosis which may ultimately be harmless to the cells. In stellate cells, the I148M variant influences the activation state of the HSCs causing increased inflammation and fibrosis which leads to progression from simple steatosis to NASH [113, 114]. Fourth, it is possible that this altered metabolic homeostasis could have detrimental effects on other metabolic pathways of the liver such as drug metabolism. We could hypothesize that the dysregulated metabolism of I148M carriers sensitizes them to other forms of hepatotoxicity. Further investigations are needed to contextualize this seemingly discordant finding in the pathogenesis of human disease.

4.4 The intermediate phenotype of PNPLA3^{I148M} indicates that the variant may be loss of function

The status of the I148M variant as a loss of function or gain of function variant is still a matter of great contention in the scientific community. Several *in vitro* studies claim that the I148M variant causes a loss of enzymatic function, but these results were not corroborated in murine models where Pnpla3 knock-out did not produce any phenotypic changes [109, 118, 122, 160, 161, 165]. We contend that the lack of clarity on this point is due in large part to the lack of an appropriate disease model. The majority of *in vitro* studies were performed on purified enzymes which is a good first step but fails to parse the role of this enzyme in the metabolic activity of the whole cell. Alternately, murine models are not an appropriate model for PNPLA3 pathogenesis due to their low protein homology and extremely low basal expression of Pnpla3 in the liver.

Since our model uses human cells that express an endogenous level of PNPLA3 (data not shown), we believe that this model would be ideal to answer the question of whether the I148M variant in PNPLA3 is a loss of function or gain of function variant. In our initial studies of the PNPLA3^{UC}, PNPLA3^{I148M}, and PNPLA3^{KO} lines, we found that PNPLA3^{I148M} cells had an intermediate phenotype between PNPLA3^{UC} and PNPLA3^{KO} lines. Given these results, we hypothesized that the I148M variant is a loss of function

variant in human hepatocytes. In order to further examine this question, we performed RNA sequencing to examine the transcriptomic differences between the three genotypes. We differentiated the three lines into HLCs, placed them in 3D culture, and treated them with either control, oleic acid, or palmitic acid medium for 24 hours. In order to examine the signalling pathways that influence the phenotypic differences between lines, we chose a time point that allowed ample time for signalling cascades to initiate a cellular response but before any phenotypic differences were observed. The RNA sequencing was performed on the FSPS13B genetic background. At least two lines from each genotype and two differentiations were analysed in order to account for batch effects caused by these variables. Following the sequencing, we performed comparative analyses to examine the transcriptomic differences between PNPLA3^{UC}, PNPLA3^{I148M}, and PNPLA3^{KO} cells when subjected to various lipids.

Figure 4.6a shows the principle component analysis (PCA) plots for the three genotypes treated with control, oleic acid, and palmitic acid medium, respectively. We found that each of the genotypes clustered far away from the others. Genotype was responsible for 60-70% of the variability in the PCA plot. These data indicate that genotype profoundly influences the transcriptome of these cells regardless of treatment. Additionally, each of the genotypes clustered together regardless of cell line and differentiation batch effects. These batch effects were responsible for only 10-20% of the variability in the PCA plot. Thus, the isogenic lines produced robust, reproducible results.

In all three treatment groups, the PNPLA3^{I148M} cells clustered between the PNPLA3^{UC} and PNPLA3^{KO} lines. The PNPLA3^{I148M} lines showed the widest intra-genotype differences, sometimes clustering closer to either PNPLA3^{UC} cells or PNPLA3^{KO} cells. These results are consistent with data from the previous section which showed high intra-genotype variability in the viability of PNPLA3^{I148M} cells when treated with palmitic acid. It appears that different PNPLA3^{I148M} cell lines and/or differentiations cause the cells to behave more like PNPLA3^{UC} or PNPLA3^{KO} lines, respectively. We hypothesize that these differences are caused by different levels of PNPLA3 functionality. This intermediate clustering indicates that the I148M variant is a loss of function variant in human hepatocytes.

In order to further confirm these findings, we compiled heatmaps to visualize the differences in expression pattern between the three genotypes (Figure 4.6b). The heatmap included the top 500 differentially expressed genes between the PNPLA3^{UC} and PNPLA3^{KO} cells for all three treatment groups. Since there were few intra-genotypic differences observed in the PCA plots, we used the average expression of each genotype in the heatmaps for the sake of clarity. These heatmaps clearly demonstrate that the PNPLA3^{UC} and PNPLA3^{KO} lines have divergent expression patterns in all three treatment groups while the PNPLA3^{I148M} cells appear to once again represent an intermediate between the two extremes. These data confirm that in our system, the I148M variant represents a loss of function variant.
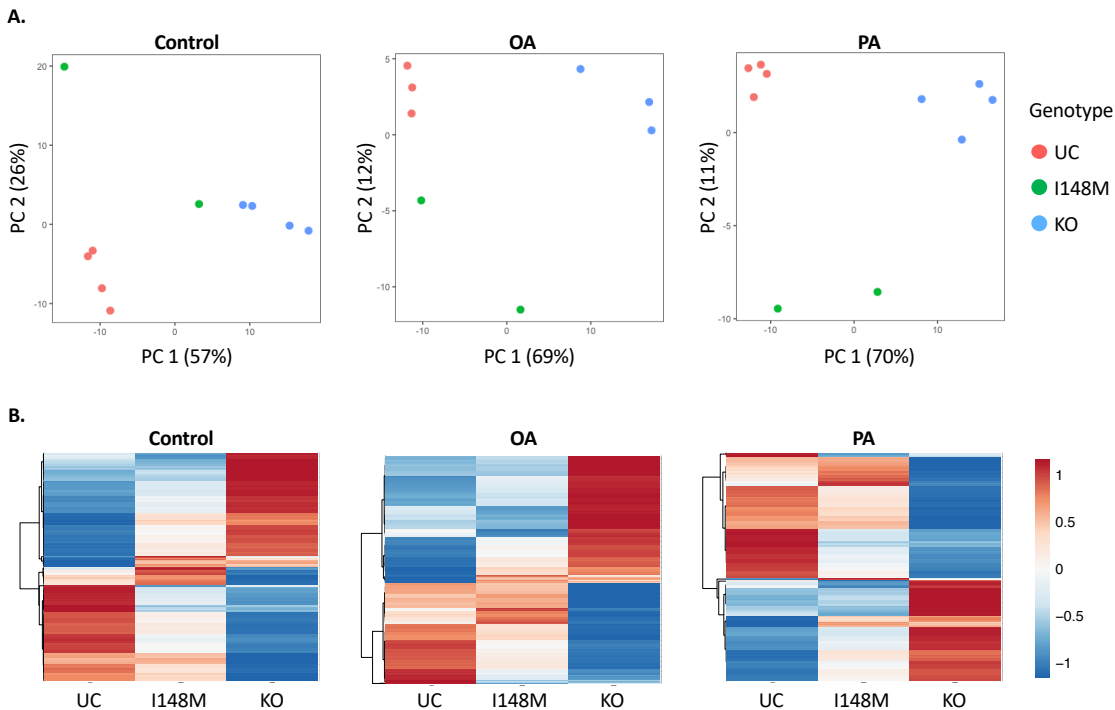
Figure 4.6 Comparative transcriptomic analysis of the three genotypes.

A. PCA plots comparing the three genotypes following the respective treatments (PNPLA3^{UC}: n = 2 clones and 2 independent experiments; PNPLA3^{I148M}: n = 1 clones and 2 independent experiments; PNPLA3^{KO}: n = 2 clones and 2 independent experiments). Regardless of treatment, PNPLA3^{UC} and PNPLA3^{KO} cells clustered far apart from one another while PNPLA3^{I148M} cells are positioned between the two indicating that it represents an intermediate phenotype. **B.** Heatmaps of the top 500 differentially expressed genes between PNPLA3^{UC} and PNPLA3^{KO} cells for each treatment. This once again shows that PNPLA3^{I148M}

cells have a transcriptomic expression profile that is intermediate between those of PNPLA3^{UC} cells and PNPLA3^{KO} cells.

Overall, we found that PNPLA3^{I148M} cells represent a transcriptomic intermediate between PNPLA3^{UC} and PNPLA3^{KO} cells. The PNPLA3^{I148M} cells clustered halfway between the PNPLA3^{UC} and PNPLA3^{KO} cells in principle component analyses of the three genotypes. The transcriptome of these cells varied slightly between differentiations and clustered closer to PNPLA3^{UC} and PNPLA3^{KO} cells, accordingly. We believe that these variations may be a result of fluctuating levels of PNPLA3 functionality between differentiations and cell lines. Additionally, the transcriptomic pattern of the PNPLA3^{I148M} cells was intermediate between the PNPLA3^{UC} and PNPLA3^{KO} cells as visualized by the heatmaps of the top 500 differentially expressed genes. These data robustly confirm the hypothesis that the I148M variant is a loss of function variant in human hepatocytes.

4.5 Conclusions

In this chapter, we used the genetically edited PNPLA3 lines to create an *in vitro* model of NAFLD to investigate the effect of PNPLA3 genotype on lipid accumulation and response to lipid-induced stress. We first differentiated the PNPLA3^{UC}, PNPLA3^{I148M}, and PNPLA3^{KO} lines into HLCs, placed them in 3D culture, and treated them with control medium or medium supplemented with oleic acid or palmitic acid to induce a steatotic or lipotoxic phenotype, respectively.

Our analyses indicate that PNPLA3^{I148M} and PNPLA3^{KO} lines accumulate more lipid droplets than PNPLA3^{UC} cells in all treatment groups. This phenotype was especially stark in the palmitic acid treated cells. Given this increased lipid accumulation in PNPLA3 edited cells subjected to palmitic acid treatment, we next investigated several phenotypic markers of cell stress. We found that while PNPLA3^{UC} HLCs upregulated the expression of ER stress markers and demonstrated a profound loss of viability when treated with palmitic acid, the PNPLA3^{I148M} and PNPLA3^{KO} HLCs failed to mount such a response to the lipotoxic stimuli. Thus, contrary to the expected result, loss of PNPLA3 function appears to be cytoprotective. In our system, it appears that PNPLA3 edited cells

fail to properly distinguish between SFAs and MUFAs. The PNPLA3 edited cells seem to preferentially shunt SFAs into triglyceride storage rather than other metabolic pathways. This has the dual effect of increasing steatosis in these cells while simultaneously protecting them from the cytotoxic effects of SFAs. Additional studies are needed to contextualize this finding in the pathogenic role of the I148M variant in human disease.

The initial phenotypic data from our model indicated that the I148M variant was most likely a loss of function variant given that the PNPLA3^{I148M} cells trended with the PNPLA3^{KO} cells in all analyses. In order to confirm this, we performed RNA sequencing. We found that the PNPLA3^{I148M} were transcriptomically intermediate between the PNPLA3^{UC} and PNPLA3^{KO} lines. These data confirmed that the I148M variant represents a loss of function variant in our model. The loss of PNPLA3 functionality either through the I148M variant or the complete knock-out profoundly affects the lipid metabolism of the HLCs causing steatosis and an altered response the lipid-induced stress. Though it has been hypothesized for over a decade, based upon clinical and *in vitro* data, our study offers conclusive evidence that the I148M variant is a loss of function variant in human hepatocytes. This *in vitro*, human hepatic system expresses PNPLA3 at an endogenous level, making it potentially invaluable for future studies into the pathogenesis of PNPLA3-induced NAFLD.

5 LIPID METABOLISM IS DISRUPTED BY ALTERATIONS TO PNPLA3

5.1   Introduction

Lipidomics is defined as the large-scale study of lipid pathways and networks and is a powerful tool for understanding lipid metabolism. This technique uses mass spectrometry to quantify the thousands of lipid species within a cell at a given time. This data can then be used to deduce the effect of physiological perturbations on the interactions and dynamics of lipid species. By examining the relative abundance of lipid species between experimental groups, one can gain insight into the metabolic pathways that may be affected by changes in genotype or treatment type. Lipidomics, in conjunction with other "omics" studies such as transcriptomics, offers the opportunity to gain mechanistic insight into pathologies that affect lipid metabolism. Additionally, lipidomics studies have been undertaken in the clinical setting for several pathologies known to affect lipid metabolism in a pathogenic way, including obesity, NAFLD, diabetes, etc [124, 237-239]. Given this huge influx for data from the clinic, lipidomics analyses of experimental systems can be used to examine the authenticity of a given system by comparing the results to published results from human patients. Lipidomics offers the unique opportunity to contextualize an *in vitro* system or murine model in terms of the human physiology and disease, making it an important tool to confirm the veracity of results discovered in these systems.

Due to the acute clinical interest in the PNPLA3 I148M variant and the lack of clarity on its mechanism of action, several clinical lipidomic studies have been undertaken in an attempt to parse the pathogenic mechanisms involved in PNPLA3-induced NAFLD [124, 126, 165, 237, 240]. These studies have revealed that PNPLA3 I148M carriers have vastly different lipidomic profiles than non-carriers of the risk allele with either a healthy metabolic profile or NAFLD caused by obesity rather than the variant. There are two main lipidomic changes in carriers of the risk allele: (1) reduced hepatic retention of lipotoxic and insulin resistance-inducing intermediates and (2) inversion of the PUFA concentration profile in triglycerides and phospholipids [126, 165].

Patients with obesity-associated NAFLD have increased concentrations of SFAs and SFA-containing triglycerides, ceramides, and diglycerides in their livers [237]. These lipid species are implicated in insulin resistance as well as lipotoxicity. Thus, the increased concentrations of these toxic intermediates in the livers of NAFLD patients

offers insight into the mechanisms of disease progression. While the absence of these lipids in the livers of I148M carriers dissociates potential mechanism of PNPLA3-induced NAFLD from that of obesity-induced NAFLD.

I148M carriers have reduced accumulation of triglycerides containing SFAs in their hepatocytes. This is due in large part to a reduction in de novo lipogenesis in these patients as well as increased secretion of SFA-containing triglycerides in VLDL. I148M carriers tend to have lower circulating concentrations of VLDL but they are uniquely enriched for SFAs [126, 165, 241]. This reduced hepatic SFA concentration also results in reduced de novo synthesis of ceramides. Ceramides are critical cell membrane components that play an important role in signalling as well. These lipids have been implicated in insulin resistance, oxidative stress, inflammation, and lipotoxicity. Ceramides are the primary mediators of insulin resistance in the liver. They activate protein kinase C which in turn activates lipogenic and lipid uptake processes. Simultaneously, ceramides impair AKT-mediated regulation of hepatic glucose production. Ceramides also play a key role in lipotoxicity. The accumulation of C16 ceramide derivatives perturbs calcium homeostasis which induces ER stress, the unfolded protein response, and ultimately cell death. Additionally, ceramides have been shown to interact with TNFα to promote the production of ROS, hepatic inflammation, and apoptosis. Thus, the reduction of ceramide concentrations has the dual benefit of reducing liver injury and improving insulin response [64, 79, 80]. Further, I148M carriers also have reduced hepatic concentrations of another key inducer of insulin resistance: diglycerides [124, 137]. Diglycerides are implicated in insulin resistance by stimulating protein kinase C to inhibit both PI3K and AKT signalling pathways. The reduced concentration of hepatic ceramides and diglycerides corresponds to the lack of metabolic syndrome often observed in patients with the I148M variant [57, 124, 137, 237, 242].

This discovery also demonstrates that PNPLA3-induced NAFLD has a unique pathogenic mechanism that is completely dissociated from that of obesity-induced NAFLD. In patients with obesity induced NAFLD, insulin resistance and increased de novo lipogenesis lead to the accumulation of lipid intermediates that exacerbate insulin resistance and lead to lipotoxicity and inflammation. Meanwhile, patients with PNPLA3-

induced NAFLD have reduced de novo lipogenesis and are more likely to escape metabolic dysfunction caused by insulin resistance. Thus, disease progression in these patients must be mediated by an, as yet, undiscovered perturbation to lipid metabolism.

PNPLA3 has been hypothesized to play in important role in lipid droplet remodelling by removing PUFAs from triglycerides and transferring them to phospholipids [127]. The I148M variant is believed to result in a loss of this remodelling capacity. In support of this hypothesis, patients with the I148M variant have increased hepatic retention of PUFA-containing triglycerides with a concomitant reduction in PUFA-containing phospholipids [165]. This pattern of PUFA accumulation may have profound effects on the lipid metabolism of these cells. First, the accumulation of PUFA in hepatic triglycerides leads to the inhibition of SREBP1c which causes downregulation of de novo lipogenesis [243, 244]. The reduction in de novo lipogenesis prevents the accumulation of toxic intermediates such as SFAs, ceramides, and diglycerides. By preventing the accumulation of these intermediates, carriers of the I148M variant are less likely to develop insulin resistance. The lack of insulin resistance as well as the presence of PUFAs per se reduces the apoptotic and inflammatory effect of lipid overload in hepatocytes [245-247]. Therefore, the lipid profile of steatotic livers from I148M patients is functionally distinct from those of obese patients. Since the increased steatosis in these patients cannot be accounted for by increased de novo lipogenesis as in obese patients, this triglyceride accumulation must result from a different mechanism.

The reduction of PUFA concentration in the phospholipids that comprise the lipid droplet membranes may offer some insight into this phenomenon. Lipid droplets are extremely dynamic organelles that maintain a diverse number of specific lipid droplet proteins based upon the temporal needs of the cell. The lipid composition of the membranes surrounding these organelles plays an important role in recruiting the appropriate proteins to the surface of the lipid droplet in order to maintain the balance between triglyceride synthesis and hydrolysis at any given time [248]. By altering the phospholipid composition of these membranes, PNPLA3 may be influencing the protein composition on the surface of the lipid droplet thereby disrupting the homeostasis between triglyceride anabolism and catabolism or restricting the mobility of these lipid species which results in the

sequestration of triglycerides in lipid droplets [128, 143]. Additional studies are needed to understand the mechanistic role of altered patterns of PUFA accumulation in the pathogenesis of PNPLA3-induced NAFLD.

In this chapter, we sought to harness the power of lipidomics to better understand the mechanism of lipid accumulation and escape from lipid induced stress by PNPLA3 edited cells. In addition, we used published lipidomic profiles from patients who are homozygous carriers of the I148M variant to validate lipidomic data from our *in vitro* system. The PNPLA3^{UC}, PNPLA3^{I148M}, and PNPLA3^{KO} lines were differentiated into HLCs, placed into 3D culture, and treated with either control, oleic acid, or palmitic acid medium for one week. Following the one-week treatment, intracellular lipids were extracted and analysed using mass spectrometry We then compared the differences in lipidomic profiles between the PNPLA3 edited and untargeted control lines to gain mechanistic insight into the phenotypes of these lines. We examined the triglyceride profiles of each of the PNPLA3 lines to identify lipid species that may play a role in lipid accumulation as well as reduced lipotoxicity of SFAs. In addition, we compared the lipidomic profile of the PNPLA3^{I148M} cells to patient data from I148M carriers in order to verify the results in our system in the context of human disease. We also tested the hypothesis that blocking triglyceride synthesis would re-sensitize PNPLA3-edited cells from both the FSPS13B and A1ATDR/R backgrounds to palmitic acid-induced lipotoxicity.

5.2   PNPLA3 edited lines have altered lipidomic profiles

In the previous chapter, we showed that alterations to the *PNPLA3* genotype resulted in profound effects on the lipid metabolism of HLCs. In order to more deeply analyse the differences in lipid metabolism caused by alterations to PNPLA3, we undertook lipidomic analyses of these cells. HLCs from each of the genotypes were treated with control or fatty acid medium for one week in order to maximize the steatotic phenotype of the cells and amplify the differences in lipid composition between genotypes. All lipidomic analyses were performed on lines derived from the FSPS13B background. For maximum clarity, the PNPLA3^{I148M} and PNPLA3^{KO} cells were compared individually to the PNPLA3^{UC} cells. This was done because the PNPLA3^{I148M} and PNPLA3^{KO} cells were

analysed at different times, in different mass spectrometry batches. The batch correction was attempted but it drastically altered the data sets as only 30% of the lipid species analysed were present in both batches. Consequently, we determined that it would be best to analyse the two genotypes separately with respect to the PNPLA3^{UC} cells in the same batch. Therefore, it is not possible to make direct, quantitative comparisons between PNPLA3^{I148M} and PNPLA3^{KO} cells. The analysis of the mass spectrometry data was performed using the publicly available online analysis pipeline Metaboanalyst [249, 250]. Given the observed phenotypes of each line and the working definition of the I148M variant as a loss of function variant, we hypothesized that the lipidomic profiles of the PNPLA3^{KO} and PNPLA3^{I148M} cells would be similar.

The overall lipidomic differences between PNPLA3^{KO} and PNPLA3^{UC} cells and between PNPLA3^{I148M} and PNPLA3^{UC} cells can be found in Figures 5.1 and 5.2, respectively. The PCA plots and heatmaps comparing the PNPLA3 edited cells to the untargeted control cells demonstrate that although fatty acid treatment alters lipidomic profiles significantly, the main determinant in lipidomic differences is genotype. There is relatively clear distinction between fatty acid treatments in PNPLA3^{UC} cells; however, this distinction is less clear in PNPLA3^{KO} and PNPLA3^{I148M} lines. These results track with the previous data showing that PNPLA3 edited cells fail to properly distinguish between SFAs and MUFAs when partitioning fatty acids for metabolism and storage. When examining the top 25 differentially abundant lipid species for PNPLA3^{KO} and PNPLA3^{I148M} cells compared to PNPLA3^{UC} cells, it is clear that the largest lipidomic differences are driven by triglycerides which is in accordance with previous studies that have shown that the I148M variant causes increased accumulation of hepatic triglycerides [126, 134, 150, 165].

In order to more precisely examine the lipidomic differences between genotypes, we separated the analyses by treatment type. Figure 5.3 shows the results of these analyses for the PNPLA3^{KO} line while Figure 5.4 shows the results for the PNPLA3^{I148M} line. These results further illustrate the marked lipidomic differences between PNPLA3^{KO} and PNPLA3^{UC} cells. Regardless of treatment type, the PNPLA3^{KO} cells cluster completely separately from PNPLA3^{UC} cells in PCA analyses. There is significantly more overlap

between PNPLA3^{I148M} and PNPLA3^{UC} cells in the same analysis. This indicates that the PNPLA3^{I148M} cells may once again represent an intermediate phenotype between PNPLA3^{UC} and PNPLA3^{KO} cells. Further, the number of differentially abundant lipid species between PNPLA3^{KO} HLCs and PNPLA3^{I148M} HLCs compared to their respective PNPLA3^{UC} HLCs supports the conclusions drawn from the PCA plots. The number of differentially abundant lipid species in PNPLA3^{KO} cells treated with control, oleic acid, and palmitic acid were 70, 71, and 59 respectively. There was a total of 788 lipid species included in the analysis therefore, 8.9%, 9%, and 7.5% of all lipid species were differentially expressed in PNPLA3^{KO} cells treated with control, oleic acid, and palmitic acid, respectively. The number of differentially abundant lipid species in PNPLA3^{I148M} cells treated with control, oleic acid, and palmitic acid were 77, 58, and 38, respectively. There were 910 lipid species included in this analysis. Thus, 8.4%, 6.3%, and 4.2% of the lipid species were differentially expressed in PNPLA3^{I148M} cells treated with control, oleic acid, and palmitic acid, respectively. This analysis indicates that alterations to PNPLA3 profoundly affects the lipidomic profile of HLCs and that the complete loss of PNPLA3 has a larger effect than the I148M variant, in accordance with previous results.
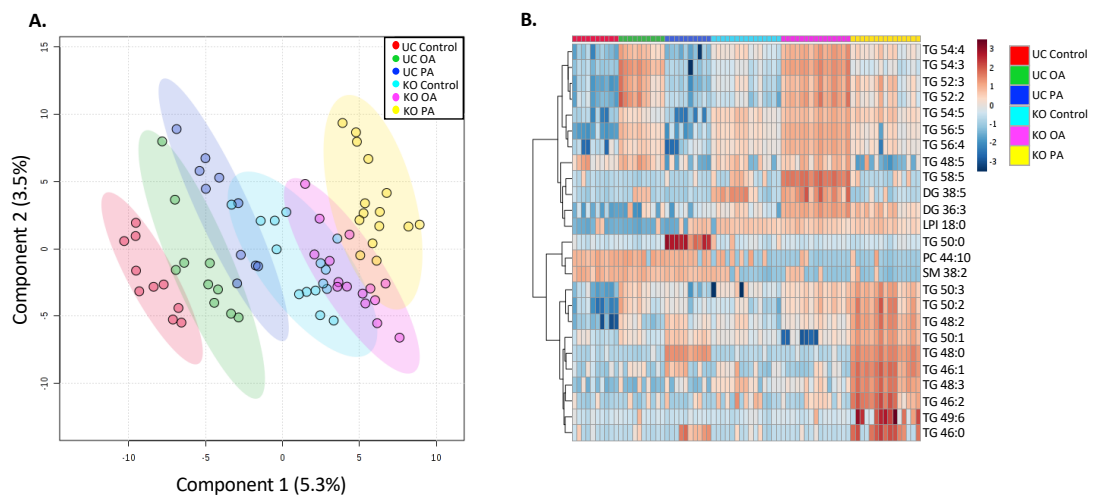
Figure 5.1 Lipidomic overview comparing PNPLA3^{KO} and PNPLA3^{UC} cells.

A. PLSDA plot of the 6 experimental groups (PNPLA3^{UC}: n = 2 clones (with 5 technical replicates) and 1 independent experiment; PNPLA3^{KO}: n = 3 clones (with 5 technical replicates) and 1 independent experiment). The two genotypes cluster separately with some differences attributable to the FFA treatment within each genotype. **B.** Heatmap of the top 25 differentially abundant lipid species in all 6 experimental groups. The lipid species most responsible for lipidomic differences between groups is triglycerides.
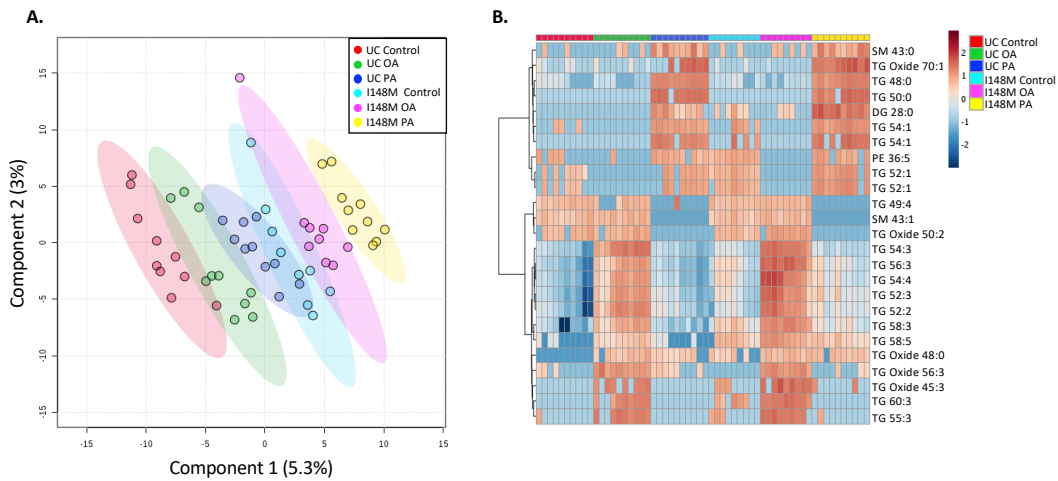
Figure 5.2 Lipidomic overview comparing PNPLA3^{I148M} and PNPLA3^{UC} cells.

A. PLSDA plot of the 6 experimental groups (PNPLA3^{UC}: n = 2 clones (with 5 technical replicates) and 1 independent experiment; PNPLA3^{I148M}: n = 2 clones (with 5 technical replicates) and 1 independent experiment). The two genotypes largely cluster separately with some differences attributable to the FFA treatment within each genotype. **B.** Heatmap of the top 25 differentially abundant lipid species in all 6 experimental groups. Triglycerides and their resultant oxides are the lipid species most responsible for differentiating the lipidomic profiles of the respective groups.
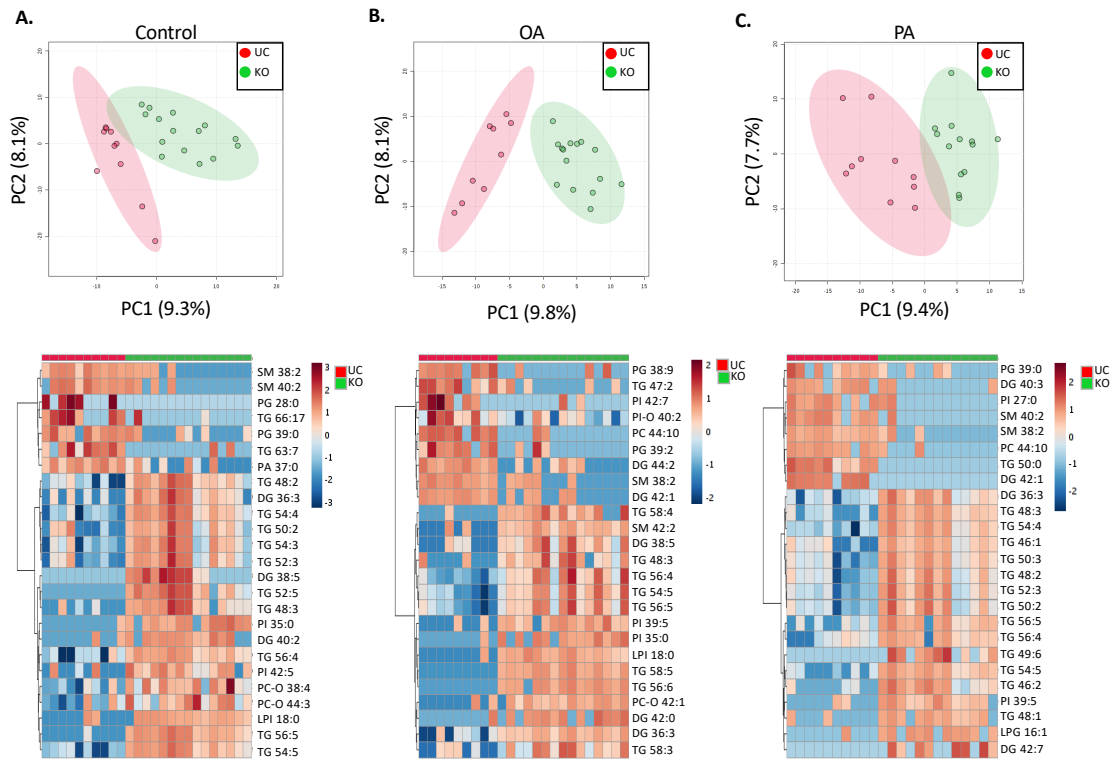
Figure 5.3 Lipidomic comparison of PNPLA3^{KO} and PNPLA3^{UC} cells separated by treatment.

PCA plots and heatmaps of the top 25 differentially abundant lipid species comparing the effect of media treatment on genotypic differences (PNPLA3[UC]: n = 2 clones (with 5 technical replicates) and 1 independent experiment; PNPLA3[KO]: n = 3 clones (with 5 technical replicates) and 1 independent experiment). Regardless of treatment, PNPLA3[UC] and PNPLA3[KO] cells clustered far apart from one another. Once again, triglycerides were a major determinant in driving lipidomic differences between the groups. **A.** PNPLA3[UC] and PNPLA3[KO] cells treated with control media. **B.** PNPLA3[UC] and PNPLA3[KO] cells treated with oleic acid. **C.** PNPLA3[UC] and PNPLA3[KO] cells treated with palmitic acid.

Upon closer examination of the lipid species comprising the top 25 differentially abundant lipid species in each treatment group for the two examined genotypes, several interesting trends appeared. In general, the PNPLA3-edited lines had a higher abundance of most examined lipid species. This trend was most clear when examining the triglyceride species. Interestingly, there were a few classes of lipids that were downregulated in the PNPLA3-edited cells. Of the few lipid species that were downregulated in PNPLA3-edited cells, diglycerides and sphingomyelins were the most common. Sphingomyelins are membrane components that are derived from phosphocholines and ceramides [251]. Therefore, a reduced accumulation of both diglycerides and sphingomyelins is consistent with the known lipidomic signatures of I148M patients.
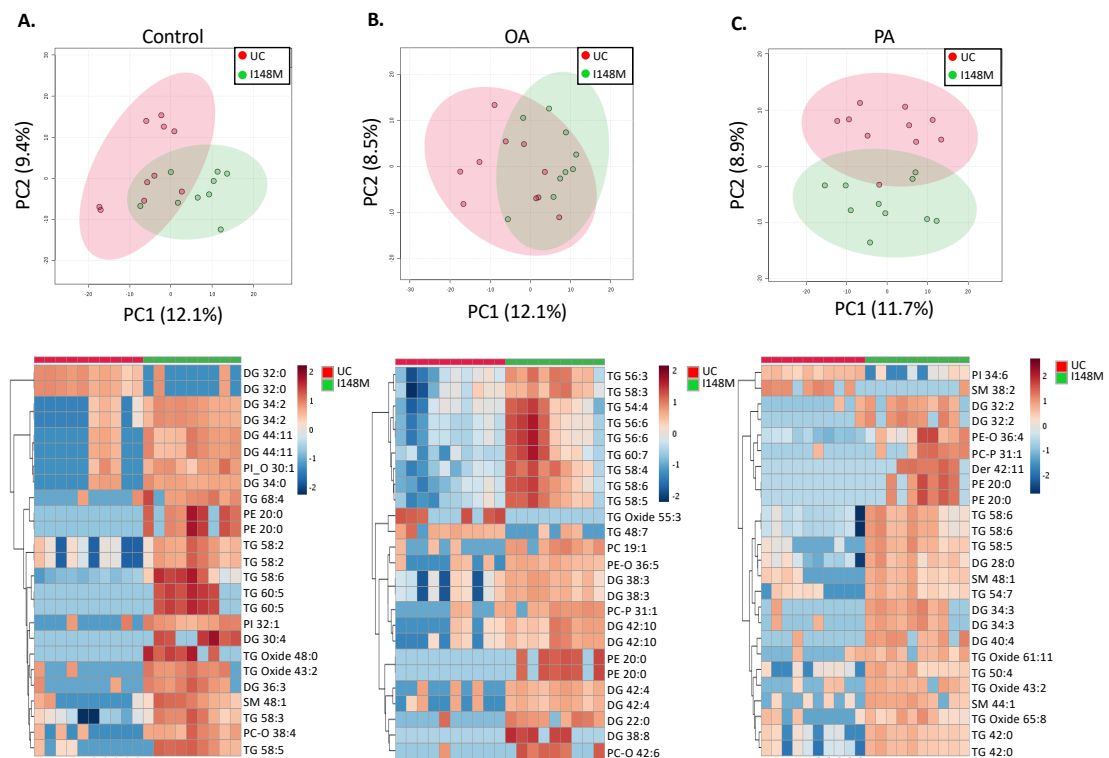
Figure 5.4 Lipidomic comparison of PNPLA3[I148M] and PNPLA3[UC] cells separated by treatment.

PCA plots and heatmaps of the top 25 differentially abundant lipid species comparing the effect of media treatment on genotypic differences (PNPLA3^{UC}: n = 2 clones (with 5 technical replicates) and 1 independent experiment; PNPLA3^{I148M}: n = 2 clones (with 5 technical replicates) and 1 independent experiment). There is some overlap between PNPLA3^{UC} and PNPLA3^{I148M} cells that varies depending on treatment type. Palmitic acid treatment appears to create the most lipidomic differences between the groups. Triglycerides as well as diglycerides were major determinants in driving lipidomic differences between the groups. **A.** PNPLA3^{UC} and PNPLA3^{I148M} cells treated with control medium. **B.** PNPLA3^{UC} and PNPLA3^{I148M} cells treated with oleic acid. **C.** PNPLA3^{UC} and PNPLA3^{I148M} cells treated with palmitic acid.

Given that the most differentially abundant lipid species in all treatment groups, regardless of genotype, were triglycerides, we decided to examine the triglyceride profile of the cells in more depth. We hypothesized that triglycerides are the key lipid mediators of PNPLA3 genotype on lipid metabolism, and the PNPLA3^{KO} and PNPLA3^{I148M} cells would preferentially accumulate triglycerides containing PUFAs. Additionally, given the results in Chapter 4, we hypothesized that the PNPLA3-edited cells would accumulate more triglycerides containing SFAs when treated with palmitic acid. The implications of these hypotheses will be discussed in more detail in the following sections. Heatmaps of the significantly differentially abundant triglyceride species between PNPLA3^{KO} and PNPLA3^{UC} cells in each treatment group can be found in Figure 5.6 while the corresponding heatmaps comparing PNPLA3^{I148M} cells and PNPLA3^{UC} cells can be found in Figure 5.7. An explanatory diagram to assist with the interpretation of these triglyceride heatmaps can be found in Figure 5.5.
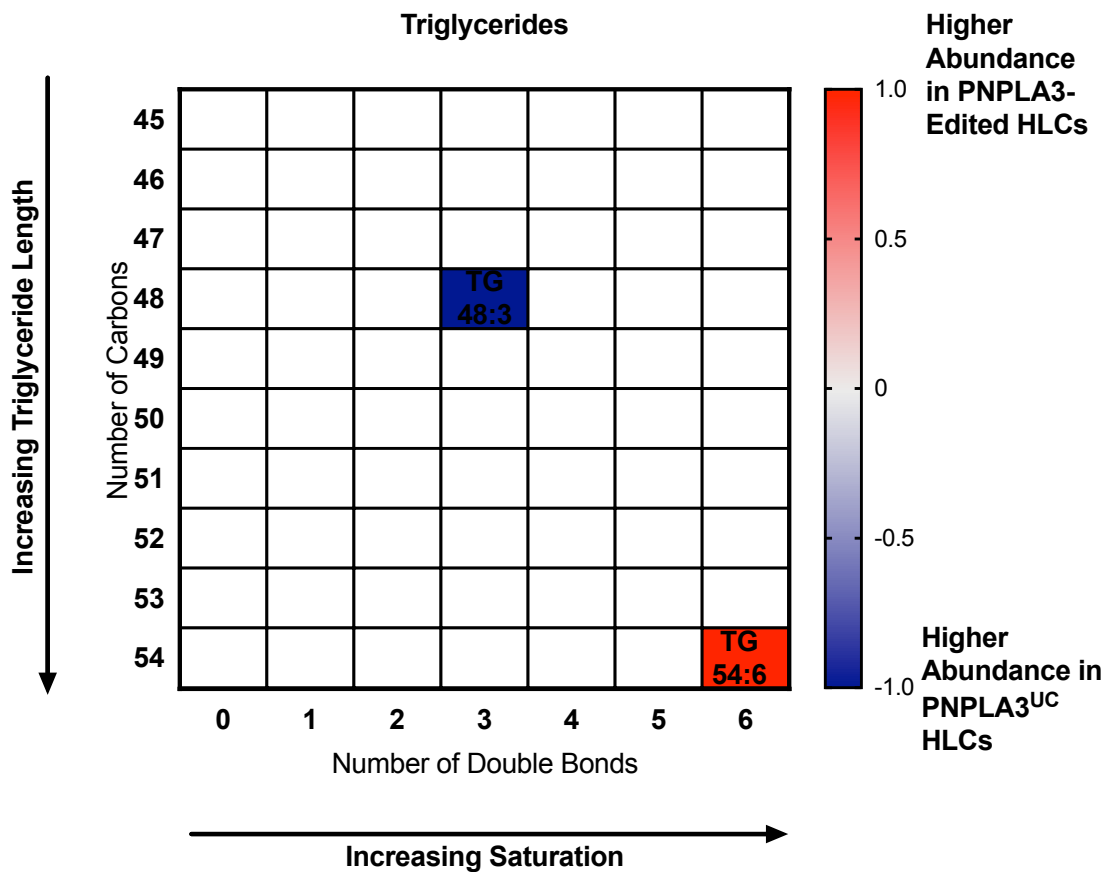
Figure 5.5 Explanatory diagram to assist with the interpretation of Figures 5.6 and 5.7.

This diagram represents a heatmap that shows differentially abundant triglyceride species between PNPLA3-edited and PNPLA3[UC] HLCs. The Y-axis shows the number of carbons in the triglyceride so as one moves down the graph, the length of the triglyceride represented increases. The X-axis shows the number of double bonds in the triglyceride so as one moves across the graph, the saturation of the triglyceride represented increases. Any triglyceride species with 0-2 double bonds contains at least one saturated fatty acid while any species with 4 or more double bonds contains at least one polyunsatured fatty acid. Each box in the graph represents a single triglyceride. If the box is coloured white, the triglyceride is not differentially abundant between the two groups. If the box is coloured red, the triglyceride is more highly abundant in the PNPLA3-edited HLCs. If the box is coloured blue, the triglyceride is more highly abundant in the PNPLA3[UC] HLCs. Therefore, in this example, TG 48:3 is more abundant in the PNPLA3[UC] HLCs while TG 54:6 is more abundant in the PNPLA3-edited cells.
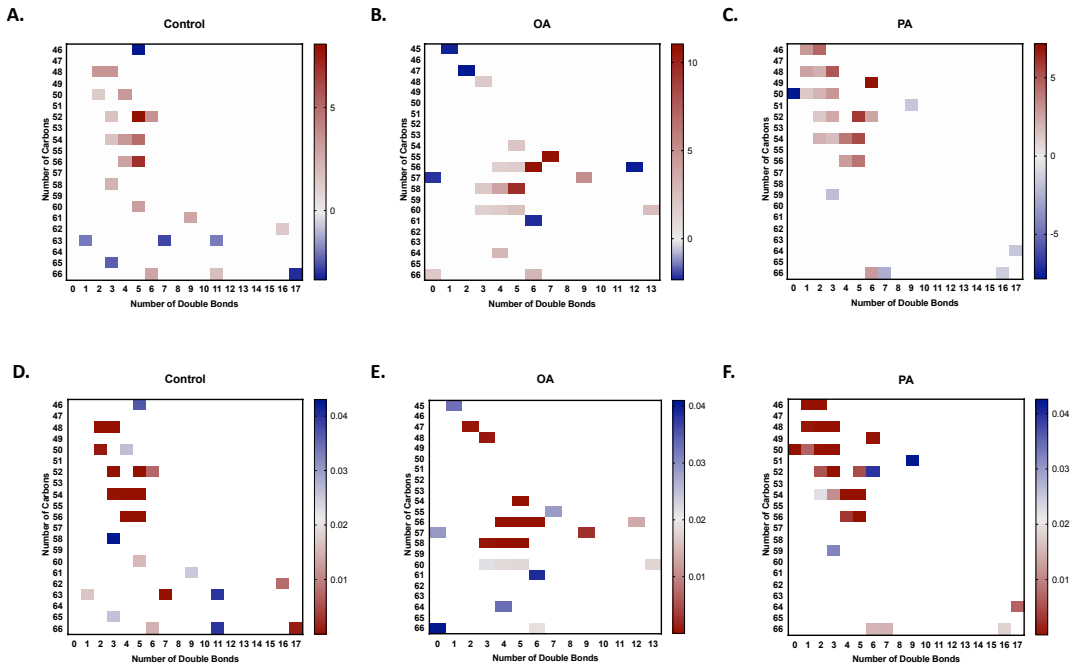
Figure 5.6 Heatmaps comparing the triglyceride profiles of PNPLA3^{KO} and PNPLA3^{UC} cells.

Heatmaps showing the (**A.-C.**) fold change difference in abundance and the (**D.-F.**) level of significance of differentially abundant triglycerides between PNPLA3^{UC} and PNPLA3^{KO} cells following the three treatments. Regardless of treatment type, the PNPLA3^{KO} cells accumulated more triglycerides containing PUFAs. In the palmitic acid treatment group, the heatmaps indicate that PNPLA3^{KO} cells may be preferentially incorporating SFAs into triglycerides as a plurality of the triglycerides that are differentially abundant between the two groups possess less than 3 double bonds.
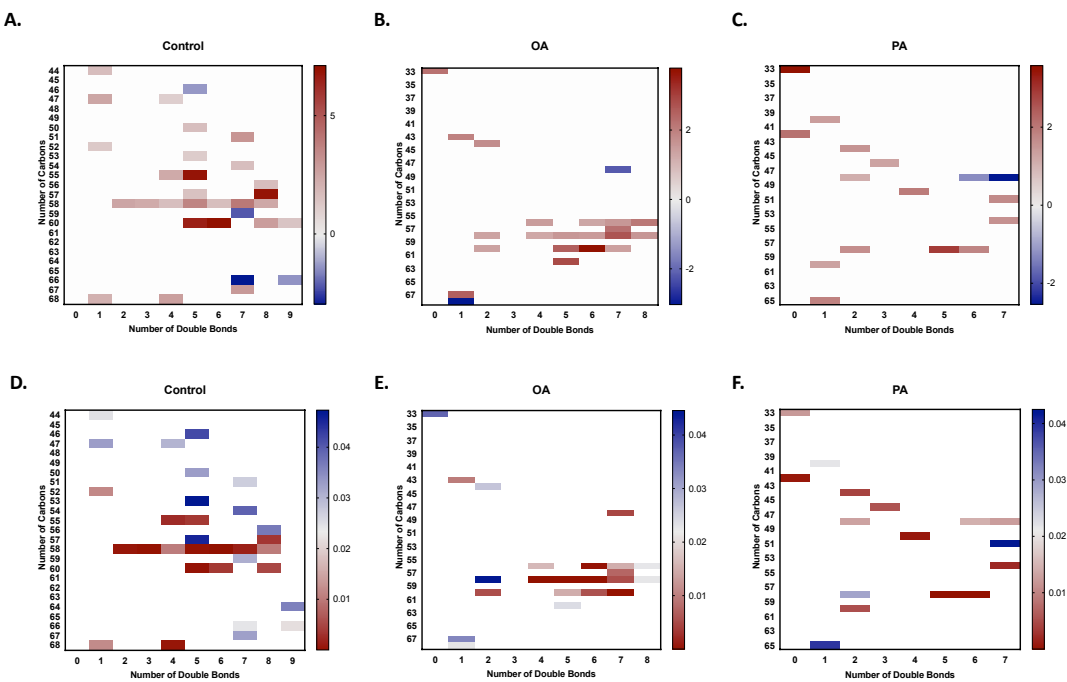
Figure 5.7 Heatmaps comparing the triglyceride profiles of PNPLA3^{I148M} and PNPLA3^{UC} cells.

Samantha G. Tilson - January 2021

Heatmaps showing the (**A.-C.**) fold change difference in abundance and the (**D.-F.**) level of significance of differentially abundant triglycerides between PNPLA3^{UC} and PNPLA3^{I148M} cells following the three treatments. Similar to the trends seen in PNPLA3^{KO} cells, the PNPLA3^{I148M} cells preferentially accumulated triglycerides containing PUFAs and when treated with palmitic acid, the PNPLA3^{I148M} cells had a higher abundance of triglycerides comprised of SFAs.

5.2.1 PNPLA3 edited lines preferentially accumulate triglycerides that contain polyunsaturated fatty acid chains

The hepatic triglycerides of carriers of the PNPLA3 I148M variant are enriched for PUFAs. PUFAs have numerous beneficial effects on lipid metabolism and have even been used as a treatment for NAFLD [252, 253]. High concentrations of PUFAs inhibit the transcription and activation of SREBP1c. PUFAs simultaneously inhibit the activation of LXRα which prevents the transcriptional upregulation of *SREBP1c* in response to insulin and increase the expression of INSIG1 protein which prevents nuclear translocation of SREBP1c [254, 255]. This reduced expression and functionality of SREBP1c leads to a downregulation of de novo lipogenesis and reduction of steatosis. Additionally, n-3 long-chain PUFAs are necessary for lipid removal from hepatocytes [81]. In addition to their role in reducing steatosis, PUFAs prevent the ER stress, apoptosis, and inflammation caused by SFAs. PUFAs accomplish this by suppressing the production of IL8 which in turn inhibits JNK and NFκB activation. This reduced JNK activity decreases ER stress markers, prevents lipoapoptosis, and reduces the release of proinflammatory mediators [81, 245-247]. Therefore, the increased relative abundance of these lipid species in the triglycerides of PNPLA3^{I148M} and PNPLA3^{KO} cells could offer insight into the resistance of these cells to palmitic acid-induced lipotoxicity.

In accordance with the clinical data, PNPLA3^{I148M} cells were enriched with triglycerides containing four to nine double bonds indicating that at least one PUFA was incorporated into the triglyceride [124, 165, 237]. A similar trend was observed in PNPLA3^{KO} cells. The PUFA species generally contain triglycerides with between 50 and 60 carbons. This is consistent with the hypothesized lipid droplet remodelling action of PNPLA3 which has a preference for catalysing the transfer of long and very long chain fatty acids from triglycerides to phospholipids [127]. The enrichment for PUFA species in triglycerides

remained consistent regardless of treatment type indicating that the role for PNPLA3 lipid droplet remodelling is independent of lipid stimuli.

The lipidomics data is also consistent with the hypothesis that the I148M variant is loss of function because the lipidomic profiles of the PNPLA3^{I148M} and PNPLA3^{KO} cells were quite similar. However, unlike the other phenotypic analyses of these cells, the PNPLA3^{I148M} cells had a more pronounced phenotype than the PNPLA3^{KO} cells. It is unclear why this would be the case. It is possible that this is a result of inter-experimental variability given that the PNPLA3^{I148M} and PNPLA3^{KO} cells were run in different mass spectrometry batches with their respective PNPLA3^{UC} controls. Alternately, this could offer evidence for the PNPLA3 I148M variant being a dominant negative variant. In this case, we use dominant negative to mean that the presence of mutated protein further exacerbates the negative effects of the loss of function variant. Previous studies have shown that I148M Pnpla3 accumulates on lipid droplets due to decreased ubiquitination [162, 163]. The presence of the I148M PNPLA3 protein may further exacerbate the negative effect of PNPLA3 loss of function on lipid droplet remodelling by further restricting the mobility of lipids or access of lipid droplet proteins. However, we did not observe the same upregulation of I148M PNPLA3 protein in our system which argues against this hypothesis. While the possibility that the I148M variant acts as a dominant negative variant remains feasible, the preponderance of evidence in our system continues to argue that the I148M variant exerts its pathogenic effects via a simple loss of enzymatic function.

	KI-Control	KI-OA	KI-PA	PNPLA3- Hyysalo Circulating	PNPLA3- Luukkonen Liver Biopsy
52:4				X	
52:6					X
54:4				X	
54:5				X	X
54:6				X	X
54:7	X		X		X
55:5	X				X
56:4		X			X
56:5				X	X
56:6		X		X	
56:7		X			X
56:8	X	X			X
58:4	X	X			X
58:5	X	X	X		X
58:6	X	X	X		X
58:7	X	X			X
58:8	X	X		X	X
58:9				X	X
60:8	X				X
60:9					X

Figure 5.8 Lipidomic comparison of PNPLA3^{I148M} cells to human patients.

Table showing upregulated triglyceride species of PNPLA3 I148M carriers in circulating samples (Hyysalo et al) and liver biopsies (Luukkonen et al) compared to PNPLA3^{I148M} lipidomic samples. The PNPLA3^{I148M} species shown in this table are triglyceride species that were differentially upregulated in PNPLA3^{I148M} cells compared to their PNPLA3^{UC} counterparts. The PNPLA3^{I148M} lipidomic profile more closely resembles the profile of liver biopsies from PNPLA3 I148M patients than the circulating lipidomic profile. The significant overlap (12/17 species) between our system and patient samples indicates the biological relevance of this system for modelling PNPLA3-induced NAFLD.

In order to validate the results from our *in vitro* system, we compared the lipidomic profile of the PNPLA3^{I148M} cells with lipidomic profiles from patients homozygous for the I148M risk allele. We used published data from Hyysalo, et al. and Luukkonen, et al. to compare the profile of our cells to patients' lipidomic profiles obtained from circulation and liver biopsy, respectively [165, 237]. Given that triglycerides were the lipid species with the largest effect on the lipidomic profile of our cells, we confined our analysis to triglycerides. The two publications found 20 triglyceride species that were upregulated in carriers of the I148M risk allele. A table of these species and their differential expression

pattern in the circulating and liver biopsy lipidome samples as well as our PNPLA3^{I148M} cells can be found in Figure 5.7. We found that of the 20 triglyceride species that were upregulated in I148M carriers, 12 were similarly upregulated in PNPLA3^{I148M} cells. As expected, the PNPLA3^{I148M} lipidome more closely resembled the sample from the liver biopsy than circulation. Of the 12 triglycerides that were upregulated in the PNPLA3^{I148M} cells, 10 were exclusively found in the lipidomic samples from patient liver biopsies. This result offers impressive clinical validation of our system. Consistent with clinical data, we have shown that the PNPLA3^{I148M} cells accumulate more lipids than their PNPLA3^{UC} counterparts and the lipidomic makeup of these prolific lipid droplets closely resemble the profile seen in the livers of patients with the I148M variant. These data demonstrate that our genetically edited HLCs recapitulate key metabolic aspects of the clinical phenotype seen in human patients and may be reliable surrogates for the effects of the I148M variant on human disease.

5.2.2 PNPLA3 edited lines incorporate lipotoxic saturated fatty acids into triglycerides

The increased concentration of triglycerides containing PUFAs could be one explanation for the decreased palmitic acid-induced lipotoxicity experienced by PNPLA3^{I148M} and PNPLA3^{KO} cells. However, there are fewer PUFA-containing triglyceride species in palmitic acid treated cells than control or oleic acid treated cells of the same genotype indicating that this is likely not the only mechanism by which PNPLA3 edited cells are avoiding lipotoxicity. Triglycerides in palmitic acid treated cells of both genotypes skew toward saturation with several species containing between zero and two double bonds. This increased triglyceride saturation in palmitic acid treated cells presents another mechanism by which PNPLA3^{I148M} and PNPLA3^{KO} cells may exhibit resistance to lipotoxicity.

SFAs have long been implicated in metabolic stress of hepatocytes due to their tendency to overwhelm β-oxidation antioxidant systems as well as triggering ER stress through hyper-saturation of membrane phospholipids. Studies have shown that metabolism of SFAs causes the accumulation of toxic intermediates and these toxic intermediates are

the culprits in lipotoxicity. SFAs stimulate their own metabolism by upregulating lipid metabolic pathways such as β-oxidation through the upregulation of PPARα expression [79-81]. Alternately, MUFAs, such as oleic acid, activate PPARγ and downstream triglyceride synthesis pathways. Therefore, somewhat paradoxically, MUFAs are not toxic to the cell because they are not metabolized to toxic intermediates but rather stored in metabolically inert triglycerides [234]. Experimental conditions that promote triglyceride synthesis such as co-supplementation of MUFAs reduces the lipotoxicity of SFAs. Therefore, if palmitic acid can be diverted into triglyceride synthesis rather than other metabolic pathways, its lipotoxic effect is neutralized [35, 231, 234].

Following this line of logic, the increased incorporation of SFAs into triglycerides in the PNPLA3[I148M] and PNPLA3[KO] lines treated with palmitic acid offers yet another mechanistic explanation for the reduced lipotoxic effect of palmitic acid on these cells. It appears that PNPLA3 may play a role in partitioning fatty acids between lipid storage and lipid metabolism. Loss of PNPLA3 function, either through the I148M variant or a complete knock-out, may result in a malfunction of this system. This failure to partition fatty acids has the dual effect of increasing lipid accumulation and paradoxically reducing the susceptibility of cells to lipotoxicity.

5.3 Blocking triglyceride synthesis re-sensitizes PNPLA3 edited cells to palmitic acid-induced lipotoxicity

In order to test the hypothesis that PNPLA3 edited cells escape SFA-induced lipotoxicity by rerouting them from metabolic pathways to triglyceride storage, we used small molecule inhibitors to block triglyceride formation in the cells. We hypothesized that blocking triglyceride formation would have no effect on PNPLA3[UC] cells since they are already susceptible to palmitic acid-induced lipotoxicity. However, in PNPLA3[I148M] and PNPLA3[KO] cells, we hypothesized that blocking triglyceride formation would re-sensitize these cells to lipotoxicity. We chose two different small molecule inhibitors that blocked triglyceride synthesis at the top and bottom of the enzyme cascade, Triacsin C and T863, respectively. Triacsin C inhibits GPAT which is the first enzyme in the cascade while T863 inhibits DGAT1, the last enzyme in the cascade. A schematic of triglyceride formation and the impact of each of the inhibitors can be found in Figure 5.8. The

PNPLA3^{UC}, PNPLA3^{I148M}, and PNPLA3^{KO} cells were differentiated into HLCs, placed
in 3D culture, and treated with either palmitic acid or palmitic acid supplemented with
each inhibitor for one week. Following the treatment, viability and lipid droplet
accumulation were quantified. These experiments were performed on cells from both the
FSPS13B background and the A1ATDR/R background for confirmation.
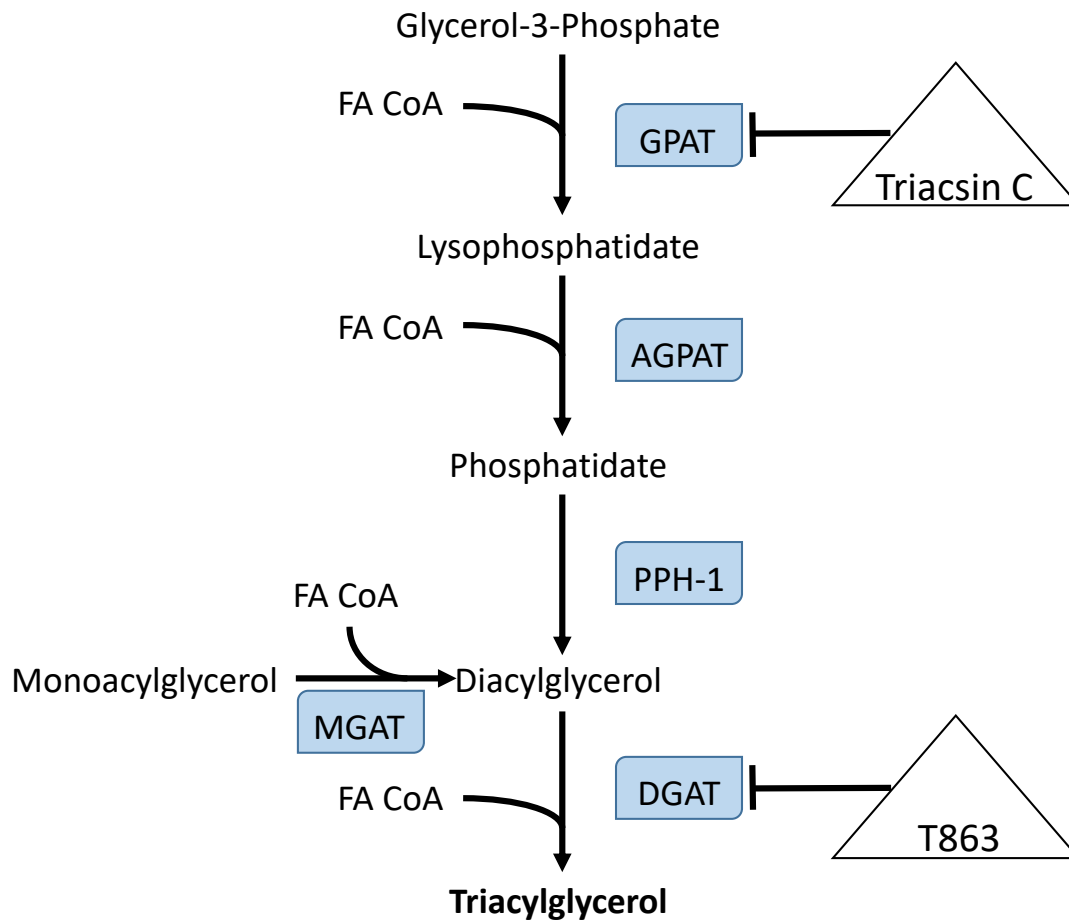
Figure 5.9 Schematic of triglyceride inhibition experiment.

Schematic of the triglyceride formation enzyme cascade denoting that Triacsin C inhibits GPAT at the top
of the cascade while T863 inhibits DGAT1 at the bottom of the cascade.

Figures 5.9 and 5.10 show the results of the triglyceride blocking experiment in the
FSPS13B and A1ATDR/R lines, respectively. Figure 5.9b shows that both Triacsin C and
T863 reduced lipid droplet formation in all three genotypes. However, the inhibitors
failed to completely inhibit triglyceride formation at the concentrations used in this

experiment. Based upon these representative images, it appears that T863 may have a stronger inhibitory effect on lipid droplet formation than Triacsin C but this qualitative observation would need to be confirmed with quantitative measurements to ensure its accuracy.
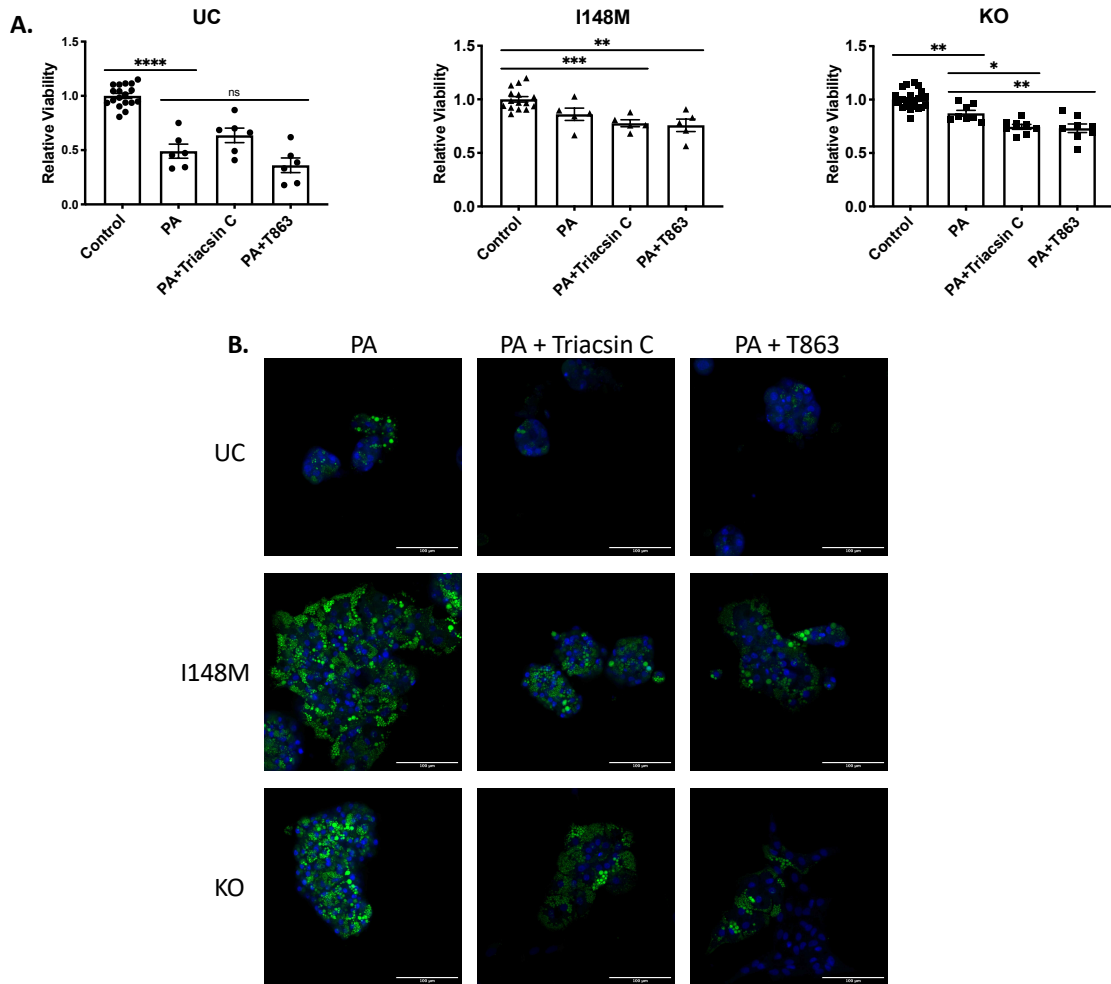
Figure 5.10 Triglyceride Blocking Assay in FSPS13B.

A. Relative viability of cells treated with palmitic acid alone or palmitic acid plus each inhibitor compared to control treated cells within each genotype (PNPLA3[UC]: n = 2 clones and 3 independent experiments; PNPLA3[I148M]: n = 2 clones and 3 independent experiments; PNPLA3[KO]: n = 3 clones and 3 independent experiments). For PNPLA3[UC] cells, treatment with palmitic acid resulted in a significant loss of viability but treatment with the triglyceride inhibitors did not result in additional loss of viability. While treatment with both inhibitors resulted in a significant reduction in viability for PNPLA3[KO] cells. PNPLA3[I148M] cells showed an intermediate phenotype with a slight reduction in viability when treated with the triglyceride inhibitors that did not reach significance. Ordinary one-way ANOVAs with Dunnett's multiple comparisons tests were performed to test statistical significance between means. Error bars represent SEM. **B.** Representative images of bodipy lipid staining following 1 week of treatment with palmitic acid and palmitic acid supplemented with the respective inhibitors. Lipid accumulation was inhibited by both

Induced Pluripotent Stem Cell Derived Liver Model for the Study of PNPLA3-Associated Non-Alcoholic Fatty Liver Disease

inhibitors in all three genotypes. This experiment was performed 3 times and the best representative images are represented in this figure.

Similar to previous experiments, we found that treating PNPLA3[UC] cells with palmitic acid resulted in a significant reduction in viability while treatment with the inhibitors did not lead to increased cell death. The PNPLA3[KO] cells had a small but significant decrease in viability when treated with palmitic acid. Treatment with both triglyceride inhibitors resulted in significantly higher cell death than palmitic acid treatment alone. Similarly, the addition of both Triacsin C and T863 caused a significant decrease in viability in PNPLA3[I148M] cells compared to control.
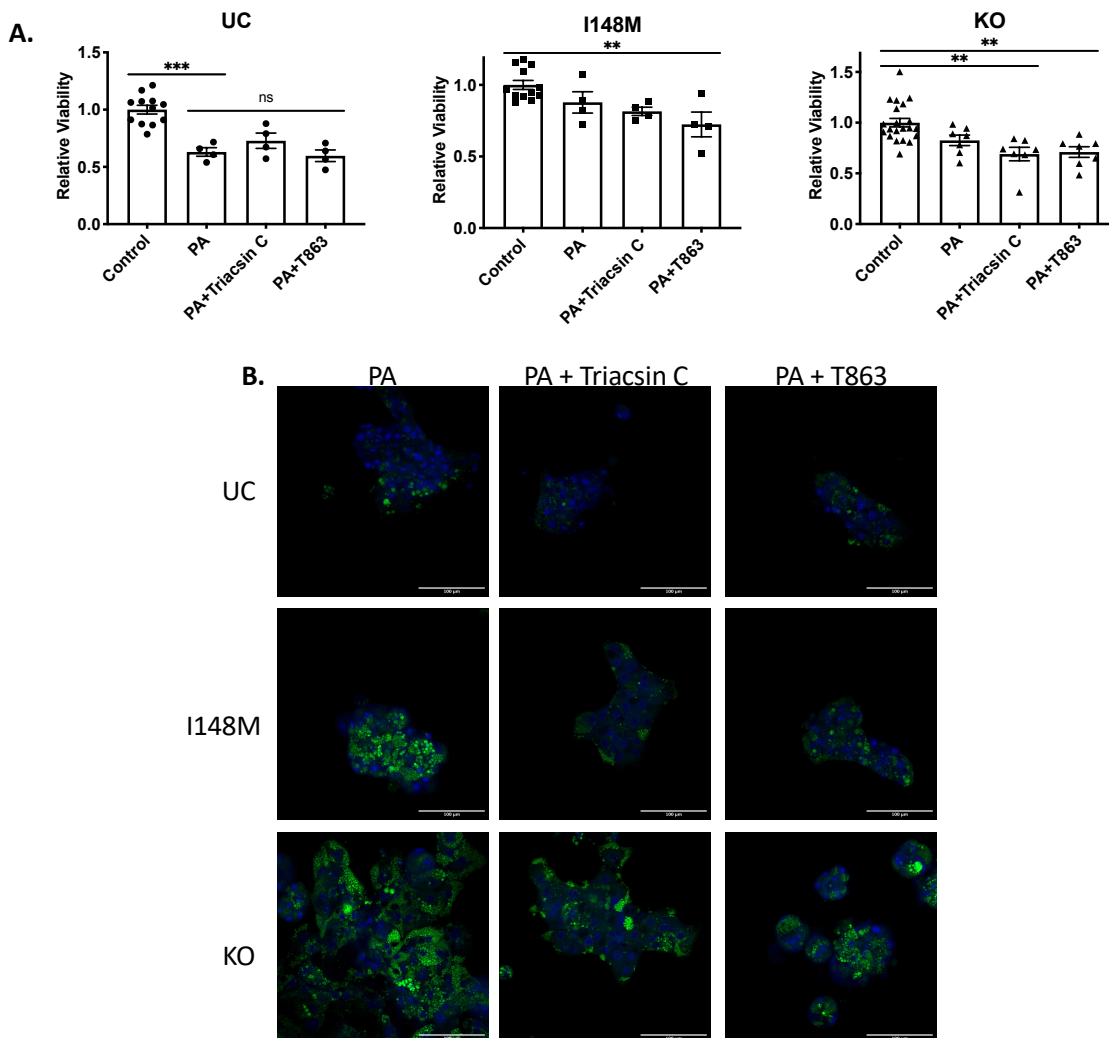
Figure 5.11 Triglyceride Blocking Assay in A1ATDR/R.

A. Relative viability of palmitic acid and palmitic acid + inhibitor treated cells to control treated cells within each genotype (PNPLA3[UC]: n = 2 clones and 3 independent experiments; PNPLA3[I148M]: n = 1 clone and 4

independent experiments; PNPLA3^{KO}: n = 2 clones and 4 independent experiments). The triglyceride inhibitors have less of an effect on this genetic background, but the trends remain consistent. In PNPLA3^{UC} cells, triglyceride inhibitors have no effect on further reducing the viability of palmitic acid treated cells. In PNPLA3^{I148M} and PNPLA3^{KO} cells, addition of the triglyceride inhibitors results in a slight decrease in viability when compared to palmitic acid treated cells alone. Ordinary one-way ANOVAs with Dunnett's multiple comparisons tests were performed to test statistical significance between means. Error bars represent SEM. **B.** Representative images of bodipy lipid staining following 1 week of treatment with palmitic acid and palmitic acid supplemented with the respective inhibitors. Treatment with these inhibitors resulted in a slight decrease in lipid accumulation in all three genotypes but to a lesser extent than the other genetic background. This experiment was performed 2 times and the best representative images are represented in this figure.

These results were confirmed in the A1ATDR/R genetic background. Figure 5.10b shows that the A1ATDR/R lines were more resistant to triglyceride inhibition as the cells treated with each inhibitor still had a significant amount of lipid accumulation. Accordingly, the effects on viability were less pronounced that in the FSPS13B line. The PNPLA3^{UC} line in this genetic background behaved nearly identically to FSPS13B. The viability of these cells was significantly reduced by palmitic acid treatment and treatment with the triglyceride inhibitors had no further effect on the viability of these cells. In the PNPLA3^{I148M} line, only treatment with T863 caused a significant reduction in viability. While treatment with both inhibitors caused increased cell death in the PNPLA3^{KO} cells compared to control. These results indicate that inhibiting triglyceride synthesis in PNPLA3 edited cells does re-sensitize them to palmitic acid-induced lipotoxicity.

These results support the hypothesis that mis-partitioning SFAs into storage rather than other metabolic pathways protects these cells from lipoapoptosis. However, the differences in viability between groups, though statistically significant, were small. Therefore, it is unlikely that shunting SFAs into triglycerides is the only protective mechanism at work in these cells. It remains possible that the increased concentration of PUFAs in the triglycerides of these cells and/or another unknown mechanism plays a role in PNPLA3^{I148M} and PNPLA3^{KO} cell's escape from palmitic acid induced lipotoxicity. Additional studies will be needed to fully understand the mechanisms at play that allow PNPLA3 edited cells to resist PA-induced lipotoxicity.

5.4   Conclusions

In this chapter, we used lipidomic analyses to understand the differences in lipid metabolism between PNPLA3^{UC} cells and PNPLA3^{KO} and PNPLA3^{I148M} cells respectively. We used the lipidome of these cells to gain mechanistic insight into how PNPLA3^{I148M} and PNPLA3^{KO} cells accumulate more lipid droplets and escape palmitic-acid induced lipotoxicity. We found that PNPLA3^{I148M} and PNPLA3^{KO} cells both had starkly different lipidomic profiles than PNPLA3^{UC} cells and that the PNPLA3-edited cells had similar lipidomic profiles to one another. Both PNPLA3^{I148M} and PNPLA3^{KO} cells preferentially sequestered triglycerides that contained PUFAs. These data were consistent with the hypothesis that PNPLA3 is a lipid droplet remodelling protein that catalyses the transfer of PUFAs between triglycerides and phospholipids. Given that both the PNPLA3^{I148M} and PNPLA3^{KO} cells both had increased accumulation of PUFA-containing triglycerides, these data remain consistent with the assumption that the I148M is a loss of function variant.

In addition, we compared the lipidomic profile of PNPLA3^{I148M} cells to the circulating and hepatic lipidomic profiles of patients with the I148M variant to validate our system. There was a significant overlap between our system and the clinical samples. The lipidomic profile of the PNPLA3^{I148M} closely resembles that of patient liver biopsy samples indicating that our cells have a competent lipid metabolism that recapitulates the phenotype observed in patients homozygous for the risk allele. These data indicate that our system has clinical relevance and may be used to gain mechanistic insight into the pathogenesis of the I148M variant.

We then used the lipidomic profiles of our cells to gain mechanistic insight into why PNPLA3-edited cells are resistant to palmitic-acid induced lipotoxicity. The lipidome of PNPLA3^{I148M} and PNPLA3^{KO} cells indicate that there is a dual protective effect of PUFA accumulation in triglycerides as well as preferential incorporation of SFAs into triglycerides. The presence of PUFAs in the triglycerides of these cells may prevent JNK activation and downregulate ER stress markers and inflammation while incorporation of SFAs into triglycerides may prevent the accumulation of toxic intermediates that facilitate lipotoxicity. We tested the latter hypothesis by blocking triglyceride formation in the cells

in an attempt to re-sensitized PNPLA3-edited cells to palmitic acid-induced lipotoxicity. We found that blocking triglyceride formation in PNPLA3^{I148M} and PNPLA3^{KO} cells led to a significant decrease in viability indicating that these cells are escaping palmitic acid-induced lipotoxicity, at least in part, by shuttling SFAs into metabolically inert triglycerides rather than other metabolic processes. Though the decrease in viability was statistically significant, the effect was extremely small. This small effect size may be due to an incomplete blockage of triglyceride formation at the concentrations used. Alternately, it is likely that triglyceride formation is just one mechanism by which the PNPLA3-edited cells escape palmitic acid induced lipotoxicity. Additional studies are needed to fully establish the molecular mechanism of this phenomenon. Overall, the lipidomic profile of PNPLA3-edited cells indicate that our *in vitro* system recapitulates key aspects of the human disease, offers further evidence to support the hypothesis that the I148M variant is loss of function, and sheds insight into the mechanism that prevents lipotoxicity in PNPLA3-edited cells.

6 MECHANISTIC INSIGHTS INTO HOW THE I148M VARIANT IN PNPLA3 CONTRIBUTES TO NAFLD PATHOGENESIS

6.1   Introduction

In our investigations thus far, we have found evidence that reduced PNPLA3 enzymatic activity leads to a reduction in lipid catabolism in favour of lipid anabolism. Loss of PNPLA3 function either through the I148M variant or knock-out of the PNPLA3 gene leads to increased lipid accumulation in the form of neutral triglycerides. This increased lipid accumulation appears to be cytoprotective against SFAs. Based upon the lipidomics data in the previous chapter, we showed that PNPLA3^{I148M} and PNPLA3^{KO} cells are able to escape palmitic acid-induced lipotoxicity by preferentially incorporating the SFAs into triglycerides rather than other metabolic processes such as oxidation. Accordingly, the PNPLA3-edited cells appear to process SFAs in a manner more similar to MUFAs such as oleic acid, causing a steatotic rather than a lipotoxic phenotype.

The lipidomics data further demonstrated a downregulation in lipid droplet remodelling which caused the accumulation of PUFAs in the triglycerides of PNPLA3-edited cells. The reduced lipid droplet remodelling capacity of these cells could have several, major implications for overall lipid metabolism. The content of the lipid droplets as well as the lipid composition of the organelle's membrane both play a role in determining the protein composition of the lipid droplet membrane [248, 256]. The failure of PNPLA3-edited cells to remodel the triglyceride and phospholipid pools in lipid droplets may result in an altered protein composition on the lipid droplet membranes in these cells. This altered protein composition may interfere with lipid homeostasis as well as intracellular signalling pathways. PNPLA3-edited cells sequester large quantities of fatty acids and other lipid species in their lipid droplets resulting in less availability of fatty acids for other metabolic processes. Additionally, this sequestration interferes with intracellular lipid signalling processes. Lipids are the major endogenous ligands that activate or repress transcription factors involved in metabolism and other cellular functions. Sequestration of lipids in lipid droplets may prevent activation of lipid catabolic pathways by failing to activate PPARα and downstream oxidation processes [257, 258]. Failure to properly regulate lipid metabolism may result in short term benefits such as escape from SFA-induce lipotoxicity; however, this dysregulated metabolism may have long term negative impacts on the health of the hepatocyte. Long-term dysregulation of lipid metabolism has

the potential to contribute to the inflammatory state as well as negatively impact other metabolic pathways such as drug metabolism or detoxification mechanisms in these cells.

In addition to maintaining lipid homeostasis in the body, the liver plays an important role in drug detoxification. It is estimated that the liver is responsible for clearing up to 60% of prescribed drugs. The purpose of drug metabolism is to facilitate the elimination of endogenous and exogenous molecules from the body by converting lipophilic compounds to hydrophilic products. This metabolism can be divided into four phases. Phase zero enzymes, usually belonging to the solute carrier (SLC) transporter family, are uptake transporters that facilitate the import of drugs into the hepatocyte. Phase one enzymes modify the drugs to become more water soluble, most commonly by oxygenating the drug. Enzymes that fall into the phase one category include cytochrome P450 (CYP) enzymes as well as several classes of dehydrogenases. The expression of these phase one enzymes is under the control of several transcription factors including CAR and PXR and downregulation of these enzymes generally leads to increased hepatotoxicity. The phase two enzymes are responsible for conjugating phase one products to hydrophilic compounds. Glucuronidation, facilitated by uridine diphospho-glucuronosytransferases (UGTs), sulfation, carried out by sulfotransferases (SULTs), and glutathionylation, performed by glutathione s-transferases (GSTs), are the most common conjugation reactions. Following detoxification, the hydrophilic products are exported by phase three enzymes. These efflux transporters are generally members of the ATP binding cassette (ABC) transporter family also known as the multidrug resistant proteins [50, 259-261].

There is significant overlap between lipid metabolism and drug metabolism in hepatocytes. These two metabolic processes share many of the same transporters, detoxifying enzymes, and transcriptional regulators. Both lipids and xenobiotic compounds can be imported and exported from hepatocytes using the ABC and SLC transporters. Since both lipids and most xenobiotics are hydrophobic in their un-metabolized state, it is efficient for hepatocytes to use the same transporters for both classes of molecules. The most prolific class of phase I drug metabolizing enzymes (DMEs) is the CYP family of enzymes. Humans express over 50 different CYP isoforms with different substrate specificity and functionality. CYPs are responsible for

metabolizing up to 75% of all prescribed xenobiotics as well as endogenous lipids such as PUFAs, retinol, bile acids, and steroids [51]. Thus, CYP enzymes are responsible for metabolizing both exogenous drugs and endogenous lipid species.

Given the overlap in the enzymes used to metabolize both drugs and lipids, there should be some overlap in the transcriptional regulation of these two pathways. Several transcription factors have been implicated in activation of both lipid and drug metabolism pathways. Notably, HNF4α is the master regulator of hepatocyte function and plays a key role in regulating gene expression of genes involved in both drug metabolism and lipid homeostasis. HNF4α activity is modulated by fatty acids and this transcription factor has significant crosstalk with other nuclear receptors active in lipid and drug metabolism including PPARα, PXR, and CAR. PPARα has been shown to activate both lipid β-oxidation pathways as well as upregulate the expression of CYP3A4 which is responsible for metabolizing up to 50% of commercially available drugs [50, 259, 262]. PPARα has also been implicated in modulating the expression of PXR and CAR, two major transcription factors responsible for activating DMEs. In addition to its role in drug metabolism, PXR also plays a role in controlling triglyceride homeostasis.

The crosstalk between lipid and drug metabolism pathways is further exemplified by the influence of NAFLD on drug metabolism. Patients with NAFLD are more susceptible to severe drug induced liver injury and hepatotoxicity. This hepatotoxicity is due to a combination of increased oxidative stress and reduced expression of DMEs. Excess fat accumulation in the liver results in a global downregulation of all phases of drug metabolism [262-265]. In order to protect hepatocytes from accumulation of toxic drugs, the phase zero transporters are downregulated to decrease drug uptake into hepatocytes. Once drugs enter the hepatocyte, both phase one and two detoxifying enzymes tend to be downregulated in NAFLD. NASH causes the downregulation of nearly all drug metabolizing CYP enzymes including CYP3A4, CYP2C9, CYP2A6, CYP2B6, and CYP2D6. The exception to this trend is CYP2E1 which is consistently elevated in the livers of NAFLD patients [264, 266, 267]. The upregulation of CYP2E1 leads to the exhaustion of glutathione reserves which causes oxidative stress and liver damage in these patients. Similarly, phase two DMEs are largely downregulated as well. UGT2B7,

SULT1A2, and GST activity were all downregulated in the livers of NASH patients [263, 268]. Together these data indicate that dysregulation of lipid metabolism, as in NASH, causes a reduction in the drug metabolism capacity of hepatocytes. This reduced drug metabolism leads to increased risk of hepatotoxicity and drug-induced liver injury.

Due to its role in lipid metabolism and the pathogenic role for the I148M variant in lipid accumulation, it is possible that PNPLA3 may also be implicated in drug metabolism. Indeed, PNPLA3 has been clinically correlated with several liver diseases in which detoxification of endogenous and exogenous compounds plays a key role. PNPLA3 has been strongly correlated with worse prognosis in patients with ALD, hereditary hemochromatosis, and cancer patients undergoing chemotherapy [139, 143, 224, 225]. The I148M variant has been linked to elevated serum levels of aminotransferases, increased steatosis, fibrosis progression, and cirrhosis development in these patients. These clinical data offer strong evidence that PNPLA3 may play a role in drug metabolism as well as lipid metabolism.

In this chapter, we sought to explore the mechanism by which the PNPLA3 I148M variant interferes with lipid and drug metabolism. We began by analysing the transcriptomic data collected from the PNPLA3^{UC}, PNPLA3^{I148M}, and PNPLA3^{KO} cells treated with and without fatty acid supplementation. We used this transcriptomic data to understand the pathways that were perturbed by changes in the PNPLA3 genotype with a particular focus on lipid and drug metabolism pathways. Following this analysis, we used our *in vitro* model system to validate these findings and test the hypothesis that loss of PNPLA3 function may have a deleterious effect on drug metabolism. In order to test this hypothesis, we examined the effect of ethanol, iron, and acetaminophen toxicity on the viability of PNPLA3^{UC}, PNPLA3^{I148M}, and PNPLA3^{KO} cells.

6.2   PNPLA3 edited cells downregulate metabolic pathways following lipid exposure

As demonstrated in both patient cohorts and our *in vitro* system, alterations to PNPLA3 result in stark phenotypic and lipidomic differences. These differences are likely due to fundamental alterations to existing metabolic processes; however, over a decade of

research dedicated to PNPLA3 has failed to uncover the mechanism of pathogenesis of the I148M variant in NAFLD. We believe that our system offers a unique opportunity to understand the transcriptomic differences between the three genotypes as well as gain insight into how the different genotypes respond to lipid induced stress. Using the RNA sequencing data described in Chapter 4 of this thesis, we sought to understand how changes to the PNPLA3 genotype altered gene expression and cellular pathways in an effort to gain mechanistic insight into the pathology of the I148M variant. As a reminder, the RNA sequencing was performed on PNPLA3^{UC}, PNPLA3^{I148M}, and PNPLA3^{KO} cells that had been differentiated into HLCs, placed into 3D culture, and treated for 24 hours with either control, oleic acid, or palmitic acid supplemented medium. Following sequencing, we performed principle component analysis, differential gene expression analysis, and gene ontology analysis to examine the differential transcriptomic response of the three genotypes to lipid induced stress. Given the phenotypic differences that we previously observed, we hypothesized that PNPLA3-edited cells would have similar transcriptomic profiles regardless of lipid treatment and that these cells would downregulate pathways involved in lipid catabolism and cell stress.

We began by performing a gene ontology enrichment analysis on the top 500 differentially expressed genes in order to identify pathways that are up or downregulated in PNPLA3-edited cells compare to untargeted control. We found in Chapter 4 that PNPLA3^{KO} cells exhibited the most extreme phenotype and displayed the most profound differences from PNPLA3^{UC} cells. Given that PNPLA3^{I148M} cells tended to be intermediate between PNPLA3^{UC} and PNPLA3^{KO} cells in all analyses, we chose to perform our gene ontology analyses using solely PNPLA3^{UC} and PNPLA3^{KO} cells in order to maximize the effect size. Since we have proven in our system that the I148M variant is a loss of function variant and exhibits an intermediate phenotype between untargeted control and knock-out genotypes, we believe that the results of these analyses can be used to infer how loss of PNPLA3 function via the I148M variant would affect hepatocyte function.
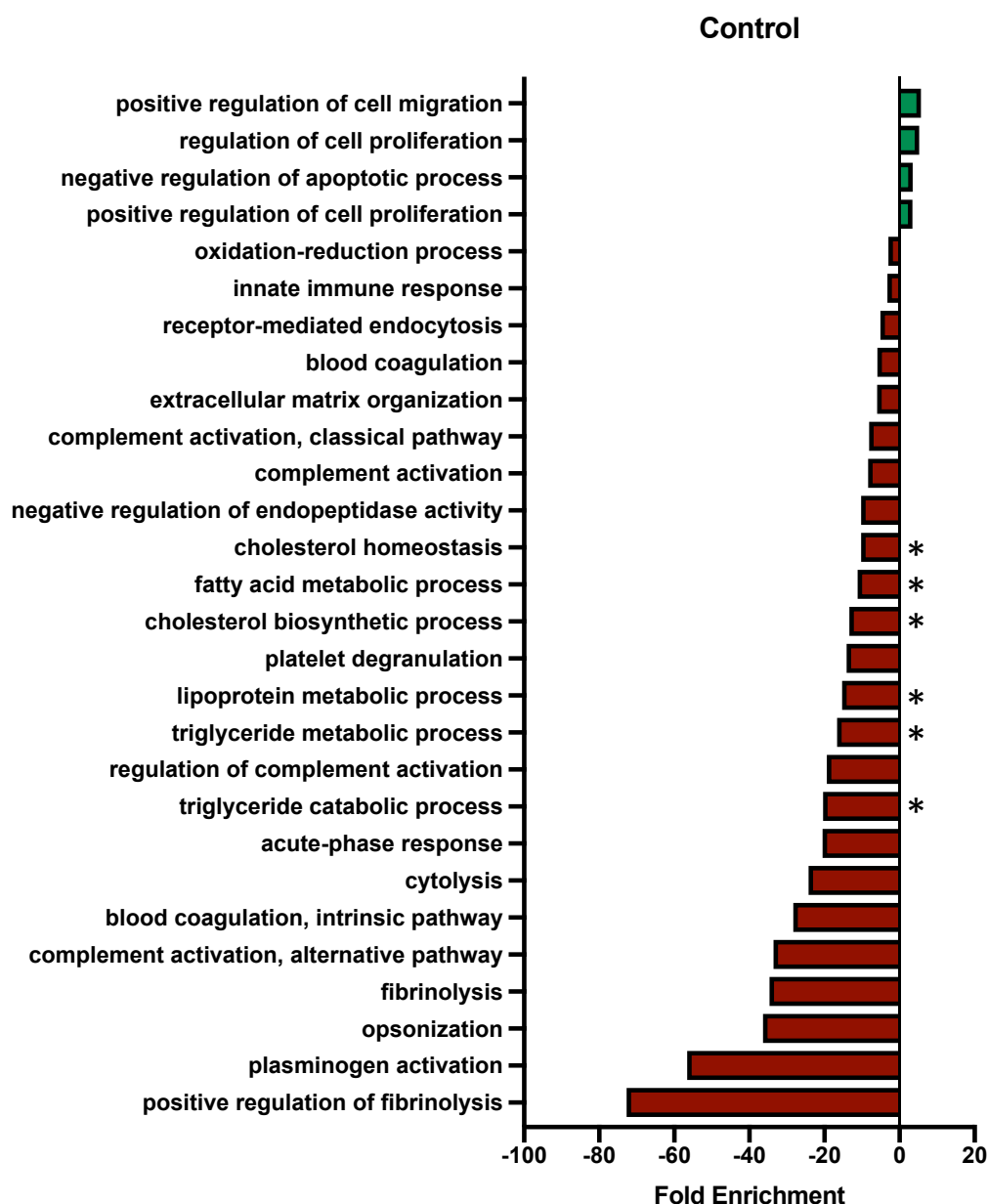
Figure 6.1 Gene ontology enrichment analysis for PNPLA3^{KO} and PNPLA3^{UC} cells treated with control medium.

Gene ontology enrichment analysis comparing PNPLA3^{UC} and PNPLA3^{KO} cells treated with control medium. Statistically significantly upregulated pathways are shown in green and downregulated pathways are shown in red. Metabolic processes are marked with a star. These results indicate that PNPLA3^{KO} cells downregulate several metabolic pathways compared to PNPLA3^{UC} cells regardless of treatment type.
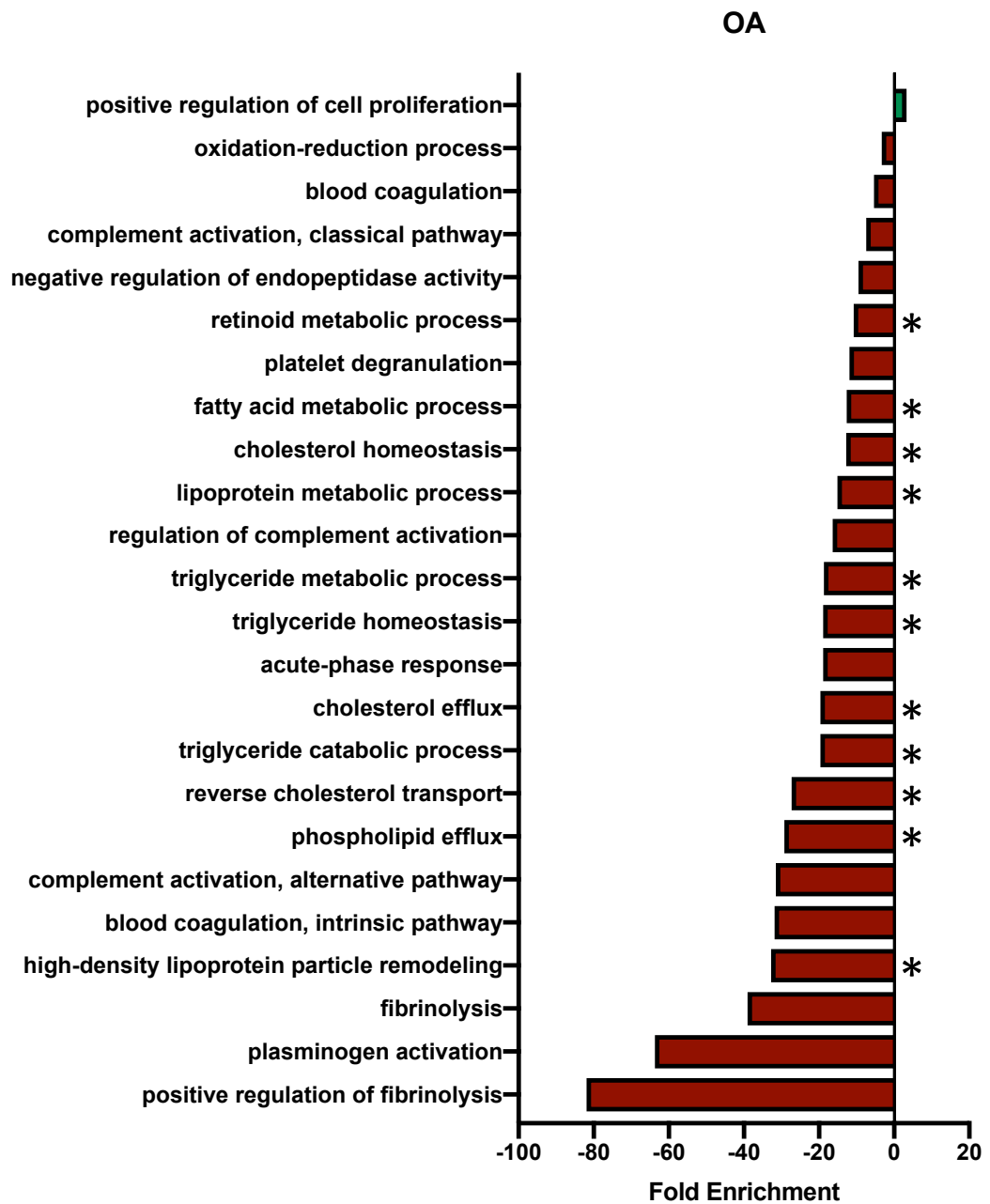
Figure 6.2 Gene ontology enrichment analysis for PNPLA3^{KO} and PNPLA3^{UC} cells treated with oleic acid.

Gene ontology enrichment analysis comparing PNPLA3^{UC} and PNPLA3^{KO} cells treated with oleic acid. Statistically significantly upregulated pathways are shown in green and downregulated pathways are shown in red. Metabolic processes are marked with a star. These results indicate that PNPLA3^{KO} cells downregulate several metabolic pathways compared to PNPLA3^{UC} cells regardless of treatment type.

PA
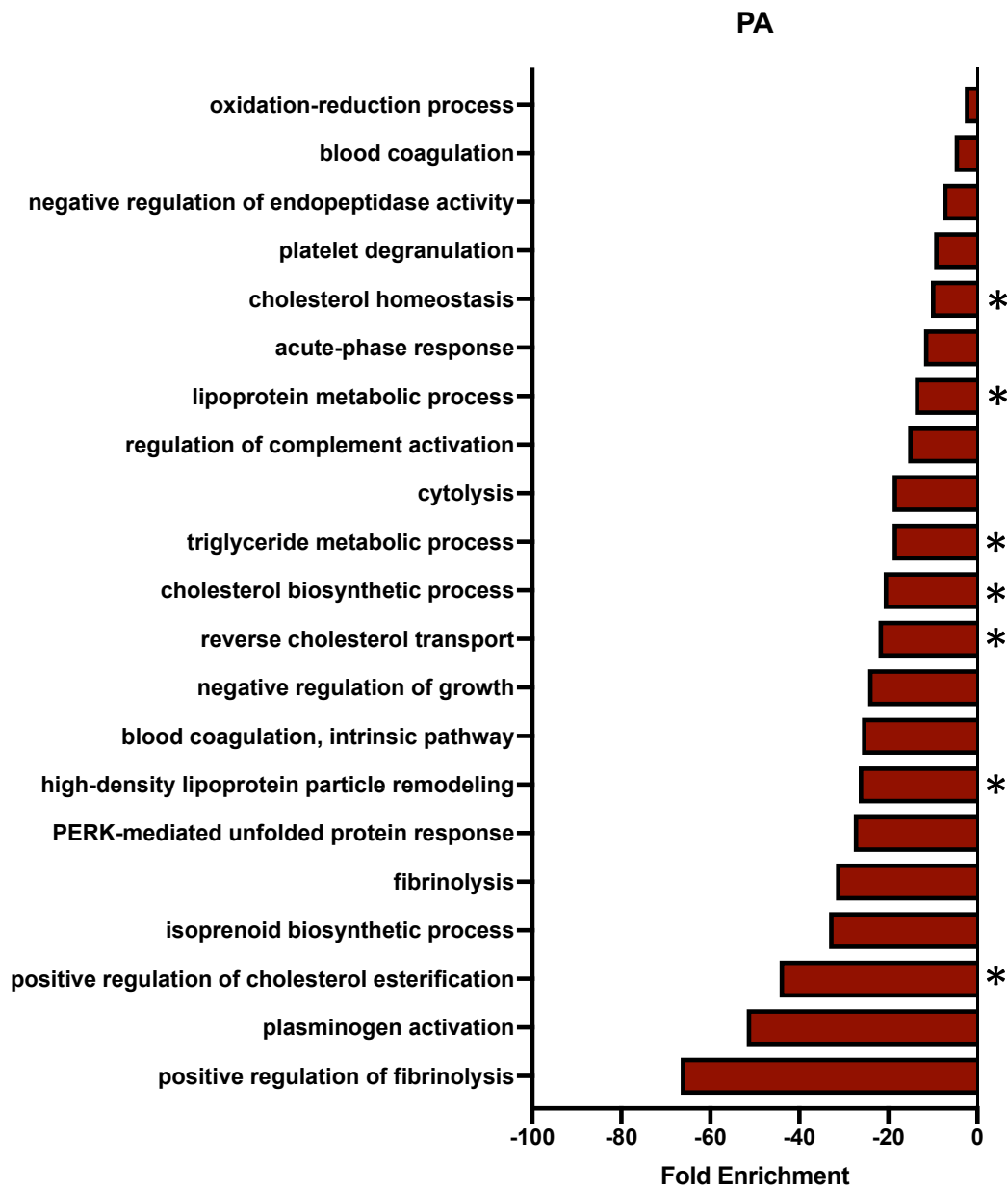
Figure 6.3 Gene ontology enrichment analysis for PNPLA3^{KO} and PNPLA3^{UC} cells treated with palmitic acid.

Gene ontology enrichment analysis comparing PNPLA3^{UC} and PNPLA3^{KO} cells treated with palmitic acid. Statistically significantly upregulated pathways are shown in green and downregulated pathways are shown in red. Metabolic processes are marked with a star. These results indicate that PNPLA3^{KO} cells downregulate several metabolic pathways compared to PNPLA3^{UC} cells regardless of treatment type.

The results of the gene ontology analysis can be found in Figures 6.1-6.3 where upregulated pathways are shown in green and downregulated pathways are shown in red.

We found that, regardless of treatment type, the vast majority of differentially enriched pathways were downregulated in PNPLA3^{KO} cells compared to PNPLA3^{UC} cells. The small number of upregulated pathways related to cell proliferation and migration, but the enrichment of these pathways was relatively small. In general, the downregulated pathways involved metabolism, innate immunity, inflammation, and cell-stress responses. The affected metabolic pathways generally related to the homeostasis of triglycerides, cholesterol, and lipoproteins though pathways involved in the metabolism of other lipid species such as retinoids and fatty acids were also represented. This general downregulation of lipid metabolism is in accordance with our previous results which show that PNPLA3-edited cells favour triglyceride formation over upregulation of metabolic processes in response to lipid stressors.

Consistent with the reduced cytotoxicity of palmitic acid in PNPLA3^{KO} cells, we found that cellular pathways responsible for responding to cell stress were downregulated in palmitic acid treated PNPLA3^{KO} cells compared to PNPLA3^{UC} cells. Specifically, the PERK-mediated unfolded protein response pathway was downregulated in these cells. This data further supports our hypothesis that PNPLA3-edited cells escape palmitic acid-induced lipotoxicity by downregulating lipid metabolic pathways in favour of triglyceride formation. Though consistent with our previous experiments, these data remain difficult to reconcile with the worsening disease phenotype seen in carriers of the PNPLA3 I148M variant.

In addition, several pathways involved in innate immunity and inflammation were downregulated in the PNPLA3^{KO} cells compared to PNPLA3^{UC} cells. This result was surprising given that the I148M variant is associated with worsening disease and high necro-inflammatory scores. However, the innate immune response system that was most profoundly affected in the PNPLA3-edited cells was the complement system. The complement system is an innate immune process that plays a significant role in defending the host from pathogens. However, recent evidence suggests that it also plays a significant role in metabolic homeostasis [269]. The complement factor C3 plays a significant role in regulating hepatic steatosis. Mice deficient in C3 have marked hepatic steatosis due to increased triglyceride accumulation and reduced fatty acid oxidation [270]. Therefore,

the reduced expression of genes relating to the complement system, including C3, in PNPLA3^{KO} cells could be contributing to the observed phenotype of increased lipid accumulation and decreased lipid metabolism in these cells.

However, progression from NAFLD to NASH has been linked to increased hepatic deposition of activated complement factors C3 and C9 [271]. Therefore, the reduction in expression of genes involved in the complement system as well as other inflammatory and cell stress pathways remains antithetical to the role of the I148M variant in worsening disease. There are several explanations for this discrepancy. It is possible that our *in vitro* system does not allow enough time for an inflammatory response to be mounted in response to metabolic dysregulation. NAFLD is a disease that requires decades to develop and progress from simple steatosis through end-stage liver disease. Thus, the full spectrum of the disease is unlikely to be captured in our system after just 24 hours to one week of lipid treatment. It remains possible that longer treatment durations with lipids or prolonged metabolic dysregulation could lead to the upregulation of pathways involved in the inflammatory response. Alternately, these data could indicate that the inflammatory stimuli for disease progression in patients with the I148M variant may originate from cells other than hepatocytes. NAFLD is a multicellular disease that requires complex intercellular signalling between all hepatic cell types. Given that HSCs have been implicated in the inflammatory response in NAFLD and the I148M variant exacerbates the inflammatory phenotype of these cells, it is possible that the I148M variant exerts its inflammatory phenotype through HSCs rather than hepatocytes [113, 166]. Additionally, it remains possible that hepatocytes carrying the risk allele may secrete factors that propagate an immune response in Kupffer cells or other liver resident immune cells.

Given that the majority of differentially enriched pathways between PNPLA3^{KO} and PNPLA3^{UC} cells were downregulated, we wanted to understand the nature of this downregulation. The gene ontology enrichment analysis only examines relative gene expression between PNPLA3^{KO} and PNPLA3^{UC} cells. Based upon this analysis, it is difficult to understand whether PNPLA3^{KO} cells truly downregulate these pathways or if they fail to properly upregulate these pathways in response to lipid supplementation. In

order to assess this, we performed principle component analysis comparing the three treatments, control, oleic acid, and palmitic acid, within each genotype (Figure 6.4).
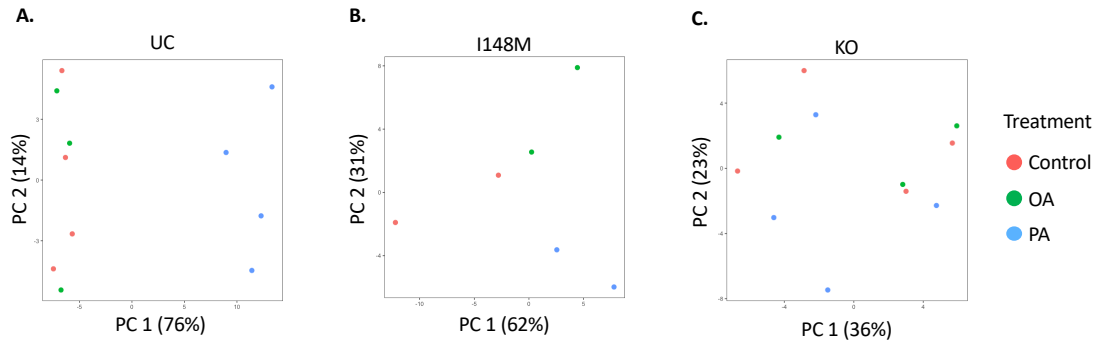
Figure 6.4 PCA plot of treatments by genotype.

PCA plot comparing transcriptomic differences between the three treatments within **A.** PNPLA3^{UC} cells, **B.** PNPLA3^{I148M} cells, and **C.** PNPLA3^{KO} cells (PNPLA3^{UC}: n = 2 clones and 2 independent experiments; PNPLA3^{I148M}: n = 1 clone and 2 independent experiments; PNPLA3^{KO}: n = 2 clones and 2 independent experiments). In PNPLA3^{UC} cells, the control and oleic acid treated groups cluster together while the palmitic acid treated group clusters far apart. However, treatment type does not seem to profoundly affect the transcriptomic profile of PNPLA3^{I148M} or PNPLA3^{KO} cells which indicates that these cells may not be activating the proper metabolic pathways when presented with a lipid stressor.

We found that the PNPLA3^{UC} cells demonstrated significant transcriptomic differences between treatments. The PNPLA3^{UC} cells treated with control and oleic acid medium clustered together while the palmitic acid treated cells clustered separately. The variability between palmitic acid treated cells and the other treatments accounted for 76% of the variability in this analysis. This observation is supported by the number of differentially expressed genes between the treatment groups. In PNPLA3^{UC} cells treated with oleic acid, there were only 31 differentially expressed genes compared to control treated cells. However, when PNPLA3^{UC} cells were treated with palmitic acid, there were 1,826 differentially expressed genes. These profound transcriptomic differences are in accordance with the stark lipotoxic phenotype observed in palmitic acid treated cells of this genotype. Alternatively, the transcriptomic profiles of PNPLA3^{I148M} and PNPLA3^{KO} cells failed to separate based upon treatment alone. In both treatments, there were very few differentially expressed genes in PNPLA3^{I148M} (OA: 40; PA: 93) and PNPLA3^{KO} (OA:27; PA: 75) cells compared to control treated cells in each genotype. These data indicate that PNPLA3-edited cells fail to upregulate lipid metabolic processes when

presented with an influx of fatty acids. As suspected the PNPLA3^{I148M} and PNPLA3^{KO} cells fail to differentiate between MUFAs and SFAs and relegate them both to the same metabolic fate: esterification into triglycerides. In lieu of the appropriate metabolic response to each class of fatty acid, PNPLA3 edited cells divert the exogenous fatty acids into triglyceride storage which results in increased steatosis and reduced SFA-induced lipotoxicity.

There is significant overlap between lipid metabolic processes, especially those for cholesterol transport/metabolism, and drug metabolism. Given the lower expression of genes involved in lipid metabolism in the PNPLA3^{KO} cell compared to PNPLA3^{UC} cells, we wanted to examine if *PNPLA3* genotype has an effect on drug metabolism as well. We hypothesized that PNPLA3-edited cells may downregulate their drug metabolism concomitantly with lipid metabolism. Since downregulation of DMEs has been linked to worsening disease phenotype in NAFLD patients, we believe that this could represent the link between the I148M variant and worsening disease phenotypes. Tables 6.1 through 6.3 show that PNPLA3^{KO} cells treated with control medium had lower expression of genes involved in all four phases of drug metabolism compared to PNPLA3^{UC} cells treated with the same medium.

Enzyme	Type	Phase	LFC	FDR
ADH1A	Alcohol Dehydrogenase	Phase 1	-6.2256	4.92E-11
ADH1B	Alcohol Dehydrogenase	Phase 1	-5.3291409	1.19E-05
ADH5	Alcohol Dehydrogenase	Phase 1	-0.283764	0.03026105
ADH6	Alcohol Dehydrogenase	Phase 1	-4.0456059	6.30E-13
ALDH1A1	Aldehyde Dehydrogenase	Phase 1	-0.680466	1.38E-06
ALDH1A3	Aldehyde Dehydrogenase	Phase 1	1.10719819	5.02E-09
ALDH1L2	Aldehyde Dehydrogenase	Phase 1	-1.8362438	0.04998308
ALDH4A1	Aldehyde Dehydrogenase	Phase 1	-3.6080842	1.60E-09
ALDH5A1	Aldehyde Dehydrogenase	Phase 1	-0.7758045	0.00010471
ALDH6A1	Aldehyde Dehydrogenase	Phase 1	-0.7463205	0.00066543
ALDH7A1	Aldehyde Dehydrogenase	Phase 1	-0.7553237	2.03E-06
ALDH8A1	Aldehyde Dehydrogenase	Phase 1	-1.8004226	1.11E-10
ALDH9A1	Aldehyde Dehydrogenase	Phase 1	-0.3717261	0.00319375
CYP19A1	Cytochrome P450	Phase 1	-3.1671584	2.73E-29

CYP1A1	Cytochrome P450	Phase 1	-0.806655	0.09241302
CYP1B1	Cytochrome P450	Phase 1	-0.5165573	0.00252969
CYP2B7P	Cytochrome P450	Phase 1	0.87391532	0.00574331
CYP2C8	Cytochrome P450	Phase 1	-3.7677259	1.04E-16
CYP2C9	Cytochrome P450	Phase 1	-1.3121207	0.05168297
CYP2J2	Cytochrome P450	Phase 1	-1.3578995	0.0001618
CYP39A1	Cytochrome P450	Phase 1	-1.9488667	5.65E-09
CYP3A7	Cytochrome P450	Phase 1	-6.7157076	1.61E-19
CYP4A11	Cytochrome P450	Phase 1	-3.717564	1.37E-06
CYP4V2	Cytochrome P450	Phase 1	-1.0090765	2.10E-07
CYP7A1	Cytochrome P450	Phase 1	-1.9578987	0.00156614
CYP8B1	Cytochrome P450	Phase 1	-1.0605197	5.93E-07

Table 6.1 List of differentially expressed phase one DMEs.

Phase one DMEs differentially expressed between PNPLA3^{UC} and PNPLA3^{KO} cells. The enzymes are divided by the enzyme family they belong to and the log fold change (LFC) and false discovery rate (FDR) are listed for each differentially expressed enzyme.

Table 6.1 shows the significantly differentially expressed phase one DME genes between PNPLA3^{KO} and PNPLA3^{UC} genes. Expression of *ADH1*, the gene responsible for oxidizing ethanol into acetaldehyde, had a more than five-fold lower expression in PNPLA3^{KO} cells compared to PNPLA3^{UC} cells. Since the I148M variant is implicated in more severe ALD phenotypes, this finding is extremely interesting. Additionally, we saw a stark downregulation of nearly all of the *CYP* family genes that were differentially expressed in our system. Of note, the *CYP3A7* gene showed a greater than six-fold decrease in expression in PNPLA3^{KO} cells. CYP3A7 is the foetal isoform of CYP3A4 which is the most prolific phase one DME responsible for nearly 50% of all drug metabolism in hepatocytes [259]. Given that the HLCs have a phenotype more closely resembling foetal cells than adult cells, CYP3A7 likely assumes the role of CYP3A4 in our system. The profound downregulation of this gene in PNPLA3^{KO} cells indicates that these cells have severely impaired drug metabolism. This data is also consistent with the data from Chapter 3 showing PNPLA3^{KO} cells in the A1ATDR/R background have lower basal CYP3A4 activity, though this difference did not reach statistical significance. Based

upon these data, it appears that PNPLA3-edited cells have impaired drug metabolism, and this is a direct result of their genotype rather than altered differentiation capacity.

We next examined the expression of phase two DMEs in Table 6.2. We examined differentially expressed genes belonging to three classes of phase two DMEs: *UGTs, SULTs*, and *GSTs*. We found that, in general, *UGTs* and *SULTs* were downregulated in PNPLA3^{KO} cells compared to PNPLA3^{UC} cells. However, the relative expression difference in these enzymes was much smaller than seen in the phase one enzymes. Interestingly, of the four *GST* enzymes that were differentially expressed, three were upregulated in PNPLA3^{KO} cells compared to PNPLA3^{UC} cells. This would indicate that PNPLA3^{KO} cell may be more resistant to oxidative stress due to their high capacity to conjugate oxidized molecules to the antioxidant glutathione. Despite this anomaly, the general downregulation of key phase two DMEs supports the hypothesis that loss of PNPLA3 function leads to a reduced functionality of drug metabolism in hepatocytes.

Enzyme	Type	Phase	LFC	FDR
GSTA1	Glutathione S-Transferase	Phase 2	1.22859997	1.27E-07
GSTA4	Glutathione S-Transferase	Phase 2	-0.5785737	0.00102473
GSTM4	Glutathione S-Transferase	Phase 2	0.48543491	0.03224733
GSTP1	Glutathione S-Transferase	Phase 2	0.26516017	0.09518873
SULT1A1	Sulfotransferase	Phase 2	0.64732791	0.00050337
SULT1C2	Sulfotransferase	Phase 2	-0.7774563	0.00808405
SULT1E1	Sulfotransferase	Phase 2	-1.5830545	0.00160952
UGT2A3	Uridine Diphospho Glucuronosyltransferase	Phase 2	0.61115064	0.06711183
UGT2B10	Uridine Diphospho Glucuronosyltransferase	Phase 2	-2.0847303	7.98E-06
UGT2B17	Uridine Diphospho Glucuronosyltransferase	Phase 2	1.14170542	0.03227578
UGT2B28	Uridine Diphospho Glucuronosyltransferase	Phase 2	-1.0059826	0.05619861
UGT2B4	Uridine Diphospho Glucuronosyltransferase	Phase 2	-2.0253223	2.00E-15
UGT2B7	Uridine Diphospho	Phase 2	0.57236177	0.00148831

	Glucuronosyltransferase			
UGT3A1	Uridine Diphospho Glucuronosyltransferase	Phase 2	-0.7236335	0.08331915
UGT8	Uridine Diphospho Glucuronosyltransferase	Phase 2	0.68452334	0.06832374

Table 6.2 List of differentially expressed phase two DMEs.

Phase two DMEs differentially expressed between PNPLA3^{UC} and PNPLA3^{KO} cells. The enzymes are divided by the enzyme family they belong to and the log fold change (LFC) and false discovery rate (FDR) are listed for each differentially expressed enzyme.

Finally, we examine the expression of transporters responsible for the influx (phase zero DMEs) of drugs and the efflux (phase three DMEs) of drug metabolites in Table 6.3. Once again there was a general down regulation of both phase zero and phase three DMEs in PNPLA3^{KO} cells compared to PNPLA3^{UC} cells. While decreasing the expression of drug importers is a cytoprotective process to prevent the import of unmetabolizable toxins into hepatocytes, the downregulation of exporters could increase hepatocellular damage by increasing the exposure time of hepatocytes to toxic substances.

Enzyme	Type	Phase	LFC	FDR
SLC10A1	Solute Carrier Transporter Family	Phase 0	-0.8274799	0.01063994
SLC22A1	Solute Carrier Transporter Family	Phase 0	-2.5726521	1.44E-05
SLC22A7	Solute Carrier Transporter Family	Phase 0	-1.6193629	7.40E-18
SLCO1B1	Solute Carrier Organic Anion Transporter Family	Phase 0	-5.4283323	5.78E-08
ABCB1	ATP Binding Cassette Transporter	Phase 3	-1.4563573	0.01747606
ABCG2	ATP Binding Cassette Transporter	Phase 3	0.62407264	0.03966208
ABCG8	ATP Binding Cassette Transporter	Phase 3	-0.4650829	0.05316166

Table 6.3 List of differentially expressed drug transporters (phase zero/phase three DMEs).

Phase zero and phase three DMEs differentially expressed between PNPLA3^{UC} and PNPLA3^{KO} cells. The enzymes are divided by the enzyme family they belong to and the log fold change (LFC) and false discovery rate (FDR) are listed for each differentially expressed enzyme.

In all, the RNA sequencing analysis indicates that PNPLA3-edited cells globally downregulate both lipid and drug metabolic pathways. Gene ontology enrichment analyses indicated that PNPLA3^{KO} cells downregulate metabolic pathways involved in

triglyceride, fatty acid, cholesterol, and lipoprotein homeostasis compared to PNPLA3^{UC} cells. This lower expression of genes involved in metabolic processes is maintained even after treatment with lipid stimuli such as oleic acid or palmitic acid. The PNPLA3^{I148M} and PNPLA3^{KO} cells did not exhibit differential transcriptomic responses to oleic acid and palmitic acid unlike the PNPLA3^{UC} cells. Therefore, these cells have incompetent metabolic responses regardless of external stimuli supporting the evidence from previous chapters that PNPLA3-edited cells accumulate more lipid droplets and fail to properly respond to lipid-induced stress. This metabolic dysregulation has the potential to affect the long-term health of these hepatocytes; however, given the short duration of treatment in our system, we failed to observe any maladaptive effects on our cells.

PNPLA3^{KO} cells also have lower expression of genes involved in innate immunity, inflammation, and cell stress. This downregulation is difficult to reconcile with the worsening disease phenotype and higher necro-inflammatory scores of patients with the I148M variant. The innate immunity pathway that is most closely linked to alterations in *PNPLA3* genotype is the complement pathway which has been implicated in triglyceride homeostasis as well. Therefore, in our cells, perturbations in the complement system are likely linked to its role in metabolism rather than the innate immune response.

Given the global downregulation of metabolic pathways, we hypothesized that drug metabolism may also be negatively affected in PNPLA3-edited cells which offers an explanation for the higher disease burden in carriers of the risk allele. We found that several DMEs from all phases of drug metabolism were differentially expressed in our system. In accordance with the trends seen in lipid metabolism, all phases of drug metabolism were downregulated in PNPLA3^{KO} cells compared to PNPLA3^{UC} cells. This trend was especially pronounced in the phase one DMEs that are the first line of defence against toxic drugs and lipid species. Given these findings, we believe that the I148M variant, at least in part, exerts its deleterious effects on NAFLD progression by interfering with the detoxification of exogenous xenobiotics as well as endogenous lipid species.

6.3 PNPLA3 edited cells may be more susceptible to other forms of hepatotoxicity

In addition to its role in metabolic diseases such as NAFLD, the I148M variant in PNPLA3 has also been linked to diseases of hepatic detoxification such as alcoholic liver disease, hemochromatosis, and chemotherapy-induced liver injury [139, 143, 224, 225]. Given this clinical data and the global downregulation of DMEs in PNPLA3^{KO} cells, we hypothesized that this incompetent drug metabolism makes PNPLA3-edited cells more susceptible to hepatotoxic insults. Analogously, it can be hypothesized that carriers of the I148M variant have more liver damage than non-carriers due to their inability to detoxify common hepatotoxic metabolites such as ethanol. This hypothesis presents two distinct pathways for increased steatosis and hepatic injury contributing to NASH progression in these patients. This notion is supported by data from our system that increased steatosis in PNPLA3-edited cells does not contribute to lipotoxic cell injury or disease progression. In order to test this hypothesis, we subjected the cells in our system to three common hepatotoxic insults and assessed the differential response of the *PNPLA3* genotypes to these insults. We differentiated PNPLA3^{UC}, PNPLA3^{I148M}, and PNPLA3^{KO} cells from the FSPS13B background into HLCs and placed them into 3D culture. In order to assess the effect of steatosis on toxicity, we then treated the HLCs with either control medium or medium supplemented with oleic acid for 48 hours. Following fatty acid treatment, the HLCs were then subjected to treatment with ethanol, iron, or acetaminophen (paracetamol) to induce toxicity. The extent of toxicity was then measured using viability analyses. We chose these three toxic compounds because they have been previously verified in *in vitro* culture systems in our lab. They are toxins that humans encounter in sublethal doses on a regular basis, and PNPLA3 has been connected to diseases caused by failure to detoxify both ethanol and iron, namely ALD and hemochromatosis. We hypothesized that PNPLA3^{I148M} and PNPLA3^{KO} cells treated with the three toxic insults would have lower viability than PNPLA3^{UC} cells. Since NAFLD has been proven to increase the risk of hepatotoxicity and drug induced liver injury, we also hypothesized that higher steatosis in oleic acid treated cells would increase the toxicity of these insults in all three genotypes.

Figure 6.5a shows a schematic of the ethanol toxicity treatment. Differentiated HLCs of all three genotypes were first treated for 48 hours with either control medium or oleic acid medium to induce steatosis. These lipid-loaded cells were then treated for 48 hours with 100 mM ethanol. Cells grown *in vitro* are less susceptible to ethanol toxicity, so the ethanol treatment was followed by treatment with 20 mM TNFα for 24 hours to exacerbate toxicity [272]. Figure 6.5b shows the effect of this treatment on the viability of HLCs from the three genotypes. This dose of ethanol was relatively non-toxic to PNPLA3^{UC} cells and had very little effect on the viability of control or oleic acid treated cells of this genotype. However, the ethanol and TNFα treatment resulted in significantly lower viability of PNPLA3^{KO} cells compared to PNPLA3^{UC} cells. The PNPLA3^{I148M} cells had lower viability than their PNPLA3^{UC} counterparts but this difference did not reach statistical significance. This result verifies the RNA sequencing data shown in Table 6.1 which shows that the major enzymes involved in ethanol metabolism, *ADH1A, ADH1B*, and *ALDH1A1*, are all significantly downregulated in PNPLA3^{KO} cells compared to PNPLA3^{UC} cells. Due to the lower expression of these detoxifying enzymes, even ethanol concentrations that were relatively nontoxic to PNPLA3^{UC} cells were quite harmful to the PNPLA3^{KO} cells. Given these data, NALFD patients with the I148M variant who consume even very small amounts of alcohol may be at risk of significant liver damage which contributes to the progression of the disease toward NASH. Interestingly, we found that induction of steatosis using oleic acid did not exacerbate the ethanol toxicity in any of the genotypes. These data support the hypothesis that steatosis and the progressive liver damage seen in NASH are functionally separate processes in carriers of the I148M variant.
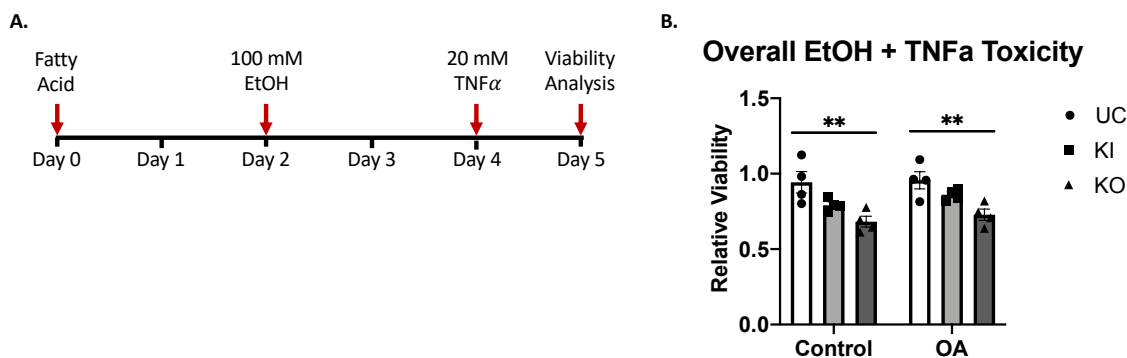
Figure 6.5 Ethanol Toxicity.

Samantha G. Tilson - January 2021

A. Timeline of ethanol and TNFα treatment to induce toxicity. **B.** Relative viability of ethanol + TNFα treated cells to untreated cells within each genotype (PNPLA3^{UC}: n = 2 clones and 2 independent experiments; PNPLA3^{I148M}: n = 2 clones and 2 independent experiments; PNPLA3^{KO}: n = 2 clones and 2 independent experiments)

. Given the sublethal dose of ethanol, PNPLA3^{UC} cell maintained high viability following treatment. Conversely, PNPLA3^{KO} cells showed an acute toxicity when treated with even a sublethal dose of ethanol. PNPLA3^{I148M} cells were intermediate between PNPLA3^{UC} and PNPLA3^{KO} cells. This trend was consistent in both control and oleic acid treated cells indicating that this effect may be independent of steatosis. Ordinary one-way ANOVAs with Dunnett's multiple comparisons tests were performed to test statistical significance between means. Error bars represent SEM.

In addition to NAFLD and ALD, the I148M variant in PNPLA3 has been linked to worsening disease phenotypes in patients with hereditary hemochromatosis. Therefore, we hypothesized that PNPLA3-edited cells would be more susceptible to iron-induced toxicity. Figure 6.6a shows the timeline of iron treatment. As before, HLCs of each genotype were treated for 48 hours with either control or oleic acid medium followed by a 24-hour treatment with 200 μM ferric ammonium citrate (FAC). The viability analysis in Figure 6.6b indicates that this iron treatment was non-toxic in all three genotypes. This lack of toxicity is likely due to impaired iron uptake in our system. The 3D culture medium used in these treatments is chemically defined and does not contain foetal bovine serum; additionally, the concentration of transferrin is quite low. Thus, there was likely very little import of the FAC into the HLCs. However, there was a slight decrease in viability in all three genotypes pre-treated with oleic acid which indicates that at least small concentrations of FAC were entering the cells and that steatosis sensitizes these cells to ferroptosis. Unlike the ethanol treatment, no inter-genotypic differences were observed in the viability of iron treated cells. These data indicate that iron alone may not be specifically toxic to carriers of the I148M variant. The role of the I148M variant in inherited hemochromatosis may rely on the interaction between the I148M variant in PNPLA3 and the C282Y mutation in the HFE gene that causes hereditary hemochromatosis [143]. Thus, our system may not be the ideal platform to examine the role of the PNPLA3 I148M variant in iron toxicity.

A.

Fatty Acid → Day 0

200 µM FAC → Day 2

Viability Analysis → Day 3

Day 0 — Day 1 — Day 2 — Day 3

B.

Iron Toxicity

Relative Viability

1.5
1.0
0.5
0.0

Control   OA

- UT
- KI
- KO
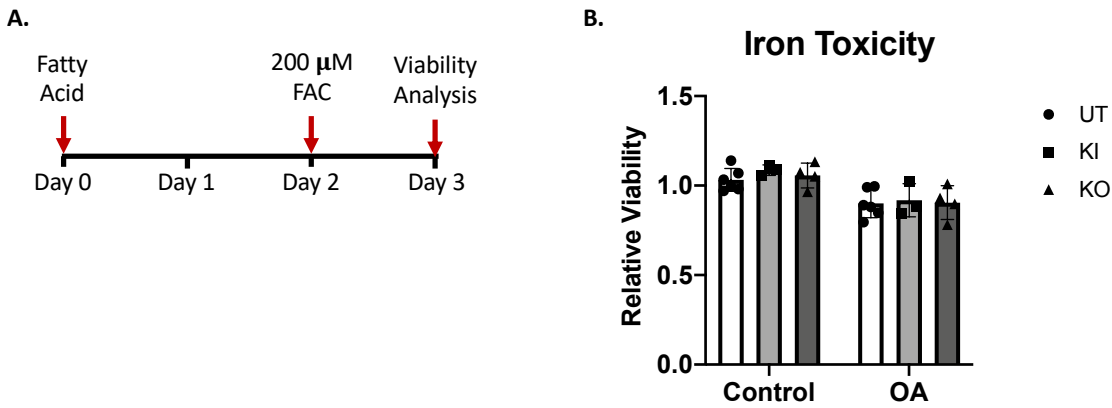
Figure 6.6 Iron Toxicity.

A. Timeline of ferric ammonium citrate (iron) treatment. **B.** Relative viability of FAC treated cells to untreated cells within each genotype (PNPLA3^{UC}: n = 2 clones and 3 independent experiments; PNPLA3^{I148M}: n = 2 clones and 3 independent experiments; PNPLA3^{KO}: n = 2 clones and 3 independent experiments). FAC induced very little toxicity and there was no notable difference in viability between PNPLA3 genotypes. Ordinary one-way ANOVAs with Dunnett's multiple comparisons tests were performed to test statistical significance between means. Error bars represent SEM.

We next examined the effect of PNPLA3 genotype on the toxicity of the common over-the-counter drug acetaminophen, also known as paracetamol. The I148M variant has not been associated with increased acetaminophen toxicity clinically, but we chose to use this drug as a surrogate for drug toxicity in our system. Figure 6.7a shows the treatment schematic for the acetaminophen toxicity experiment. We treated the HLCs of each genotype with either control or oleic acid medium for 48 hours followed by treatment with 25 mM acetaminophen for 48 hours. Figure 6.7b shows that acetaminophen treatment induced strong toxicity in all three genotypes. There was no significant difference in viability between the three genotypes or between control or oleic acid treated cells of the same genotype. To ensure that the high toxicity of this dose was not obscuring intergenotypic differences, we tested a sublethal dose (12.5 mM acetaminophen) and saw a similar trend (data not shown). These results, while initially contradictory to the ethanol toxicity results, are supported by the RNA sequencing data. The majority of acetaminophen is metabolized through glucuronidation and sulfation; however, at high concentrations, these detoxification pathways are overwhelmed, and acetaminophen is instead metabolized through a combination of oxidation by CYP2E1 and glutathione

conjugation by GST enzymes [273]. The RNA sequencing data indicates that while the PNPLA3^{KO} cells have lower expression of both *UGT* and *SULT* enzymes, these cells express higher levels of GST enzymes. There was no significant difference in *CYP2E1* expression between PNPLA3^{KO} and PNPLA3^{UC} cells. Based upon these data, it is conceivable that PNPLA3^{KO} cells may have a lower capacity for detoxification of acetaminophen through glucurondation and sulfation but their increased glutathione transferase capabilities compensate for these deficiencies. Thus, the overall viability of the cells remained unchanged, but the process of detoxification may differ between PNPLA3-edited cells and their untargeted control counterparts. Additional experiments are needed to examine the veracity of this hypothesis.
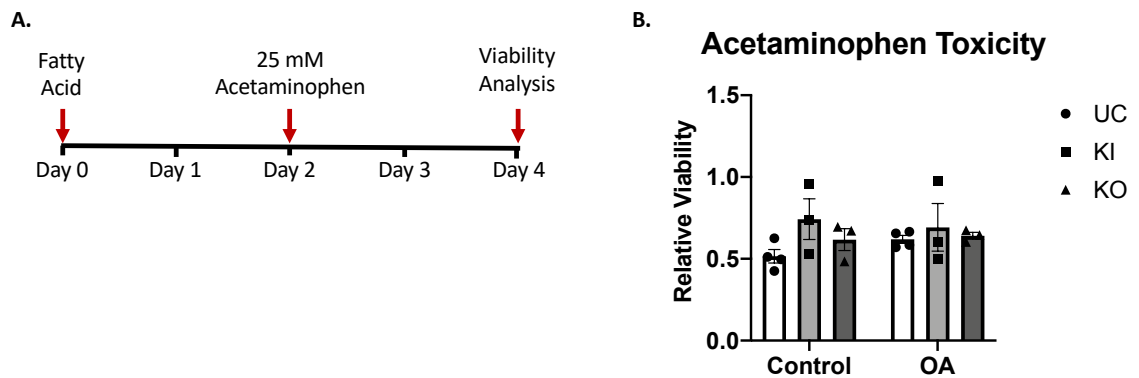
Figure 6.7 Acetaminophen toxicity.

A. Timeline of acetaminophen treatment to induce toxicity. **B.** Relative viability of acetaminophen treated cells to untreated cells within each genotype (PNPLA3^{UC}: n = 2 clones and 2 independent experiments; PNPLA3^{I148M}: n = 2 clones and 2 independent experiments; PNPLA3^{KO}: n = 2 clones and 2 independent experiments). Acetaminophen caused acute toxicity in all cells. There were no significant differences in viability between PNPLA3 genotypes or treatment groups. Ordinary one-way ANOVAs with Dunnett's multiple comparisons tests were performed to test statistical significance between means. Error bars represent SEM.

We used three *in vitro* toxicity assays to examine the hypothesis that downregulation of DMEs in PNPLA3-edited cells leads to increased susceptibility to hepatotoxicity in these cells. Additionally, we sought to determine if steatosis was a contributing factor to this hepatotoxicity by pre-treating the cells subjected to toxic insult with oleic acid to induce lipid accumulation in these cells. Induction of toxicity with ethanol and TNFα, FAC, and acetaminophen had differential results in our system. Treatment with ethanol and TNFα

induced a significantly higher toxicity in PNPLA3^{KO} cells than PNPLA3^{UC} cells while PNPLA3^{I148M} cells demonstrated an intermediate phenotype. This effect was seen at ethanol concentrations that were sublethal to PNPLA3^{UC} cells. These results are consistent with a role for the PNPLA3 I148M variant in exacerbating ALD. However, treatment with iron and acetaminophen failed to produce inter-genotypic differences in toxicity. There was no difference between control and oleic acid treated cells in any of the toxicity challenges indicating that susceptibility to hepatotoxicity is dissociated from the presence of steatosis. The data from these three toxicity treatments are inconsistent making it difficult to conclude if loss of PNPLA3 activity contributes to hepatocellular damage by increasing the susceptibility of these cells to common hepatotoxins. Additional studies will be needed to fully understand the role that PNPLA3 plays in drug metabolism. That being said, the results from the RNA sequencing as well as the preliminary results from the ethanol toxicity experiment still offer validity to this hypothesis.

6.4   Conclusions

In this chapter, we sought to explore the mechanisms by which changes to *PNPLA3* genotype affect susceptibility to hepatic injury. We examined the transcriptomic differences between PNPLA3^{UC}, PNPLA3^{I148M}, and PNPLA3^{KO} cells using principle component analyses, differential gene expression analyses, and gene ontology enrichment analyses. We found that PNPLA3^{KO} cells had a general reduction in expression of metabolic genes compared to PNPLA3^{UC} cells, especially those involved in the homeostasis of triglycerides, fatty acids, cholesterol, and lipoproteins. This trend was independent of fatty acid treatment and offers insight into how PNPLA3-edited cells accumulate more lipid droplets than their PNPLA3^{UC} counterparts. Principle component analyses indicated that, unlike PNPLA3^{UC} cells, PNPLA3-edited cells failed to differentiate between control, oleic acid, and palmitic acid treatment in their transcriptional profile. This data offers further credence to the hypothesis that PNPLA3-edited cells fail to metabolically differentiate between SFAs and MUFAs. The global downregulation of metabolism in PNPLA3^{KO} cells compared to PNPLA3^{UC} cells extended beyond lipid metabolism to drug metabolism as well. DMEs from every phase of the detoxification cascade were downregulated in PNPLA3^{KO} cells compared to

PNPLA3^{UC} cells. These data led us to hypothesize that alterations to PNPLA3 may sensitize HLCs to other forms of hepatotoxicity beyond lipotoxicity. We subjected PNPLA3^{UC}, PNPLA3^{I148M}, and PNPLA3^{KO} cells to treatment with ethanol and TNFα, iron, and acetaminophen to induce cytotoxicity. We found that PNPLA3^{KO} cells were more susceptible to ethanol-induced toxicity compared to PNPLA3^{UC} cells, even at low concentrations that were sublethal to the PNPLA3^{UC} cells. However, we failed to replicate these findings in the iron and acetaminophen toxicity assays. Additional studies are needed to determine if carriers of the I148M variant are more susceptible to other forms of hepatotoxicity beyond ethanol. In all, we found that loss of PNPLA3 expression leads to a global downregulation in metabolism. This metabolic dysregulation could have adverse consequences on the long-term health of the hepatocyte either by inducing toxicity or inflammation. Additional toxicity studies as well as long-term culture studies will be needed to fully elucidate the consequences of reduced PNPLA3 expression and metabolic activity on these cells.

7 DISCUSSION

Samantha G. Tilson - January 2021

7.1 Summary of Findings

Although the I148M variant in PNPLA3 was identified as a major risk factor for NAFLD progression over a decade ago, our understanding of the function of PNPLA3 in health and disease remains largely obscure. The controversy surrounding the function and pathogenesis of PNPLA3 stems largely from the lack of an appropriate human model that expresses endogenous levels of the protein.

In this thesis, we used CRISPR/CAS9 and hiPSCs to create a novel, *in vitro*, human model to study the effects of *PNPLA3* genotype on NAFLD development and progression. We used CRISPR/CAS9 to create genetically edited hiPSC lines that were homozygous for the reference allele (PNPLA3^{UC}), homozygous for the risk allele (PNPLA3^{I148M}), or had a complete knock-out of the PNPLA3 gene (PNPLA3^{KO}) in two different genetic backgrounds (FSPS13B and A1ATDR/R). Following the generation and validation of these lines, we differentiated each line into HLCs using an established protocol that mimics the embryonic development of hepatocytes. We demonstrated that PNPLA3 genotype does not affect the capacity of hiPSCs to differentiate into HLCs. We also examined the effect of 3D culture on these HLCs and found that this culture method improved their maturity. Therefore, for all NAFLD modelling experiments, differentiated HLCs of each genotype were cultured in this 3D format.

In order to examine the effect of *PNPLA3* genotype on lipid metabolism, we treated the lines with control medium, to model a normal diet, or medium supplemented with oleic acid or palmitic acid to model the NAFLD disease phenotypes of steatosis and lipotoxicity, respectively. We found that regardless of treatment, there was a stepwise increased in lipid accumulation from PNPLA3^{UC} cells to PNPLA3^{I148M} cells to PNPLA3^{KO} cells. In addition to this increased steatosis, PNPLA3-edited cells were also resistant to lipid-induced stress. PNPLA3^{I148M} cells and PNPLA3^{KO} cells failed to upregulate ER stress markers in response to palmitic acid treatment and these PNPLA3-edited cells were resistant to the lipotoxic effects of this SFA. This resistance to

lipotoxicity gave way to a steatotic phenotype that closely resembled oleic acid treated cells.

In order to elucidate the mechanisms behind this increased steatosis and resistance to lipotoxicity, we used lipidomic and transcriptomic data from each of the lines. We found that PNPLA3[I148M] and PNPLA3[KO] cells had vastly different lipidomic profiles from PNPLA3[UC] cells. The PNPLA3-edited cells accumulated higher concentrations of triglycerides and these triglycerides were preferentially composed of PUFAs and SFAs. These lipidomic findings offered insight into the functionality of PNPLA3 as well as the mechanism by which these cells escape from SFA-induced lipotoxicity. PNPLA3 has been hypothesized to have lipid droplet remodelling functionality. Specifically, it is hypothesized that PNPLA3 catalyses the transfer of PUFAs from triglycerides to phospholipids and the I148M variant results in the loss of this functionality. The accumulation of PUFA species in the triglycerides of PNPLA3-edited cells in our system supports this hypothesis. Additionally, the accumulation of SFAs and PUFAs in triglycerides has been proven to be cytoprotective against SFA-induced lipotoxicity, offering insight into the phenotypic differences between genotypes. In order to test the hypothesis that PNPLA3-edited cells escape SFA-induced lipotoxicity by diverting SFAs from metabolic pathways into triglyceride storage, we used small molecule inhibitors to block triglyceride formation in our system. In support of this hypothesis, we found that blocking triglyceride formation in our system re-sensitized PNPLA3-edited cells to SFA-induced lipotoxicity.

Given these data, we hypothesized that loss of PNPLA3 function results in the downregulation of lipid catabolism pathways in favour of triglyceride formation. RNA sequencing of the three genotypes revealed a global downregulation of lipid metabolism pathways involved in the homeostasis of fatty acids, triglycerides, cholesterol, and phospholipids. This analysis also revealed that, perhaps in part due to their metabolic dysregulation, PNPLA3-edited cells fail to metabolically differentiate between oleic acid and palmitic acid. This finding offers further evidence in support of the hypothesis that PNPLA3-edited cells avoid lipotoxicity by downregulating metabolic pathways and preferentially storing all fatty acid species as triglycerides.

In addition to the downregulation of lipid metabolic pathways, PNPLA3-edited cells also downregulated the expression of DMEs in every phase of the drug metabolism cascade. Given this global downregulation and the clinical evidence that the I148M variant plays a role in progression of diseases of drug metabolism such as ALD, we hypothesized that loss of PNPLA3 function makes hepatocytes more susceptible to hepatotoxicity. In support of this hypothesis, we found that PNPLA3^{KO} cells were more susceptible to ethanol-induced toxicity than PNPLA3^{UC} cells, even at concentration that was sublethal for the PNPLA3^{UC} cells. However, we failed to replicate this increased lethality in two other models of hepatotoxicity, iron and acetaminophen. Therefore, additional studies are needed to elucidate the role of PNPLA3 in drug metabolism and hepatotoxicity.

In nearly every experiment that was performed, the PNPLA3^{I148M} cells represented an intermediate phenotype between PNPLA3^{UC} and PNPLA3^{KO} cells. These results demonstrate that the I148M variant in PNPLA3 is a loss of function variant. The most compelling evidence for this finding is the RNA sequencing data that showed in both principle component analysis and heatmap analysis that PNPLA3^{I148M} cells represented a near perfect transcriptomic intermediate between PNPLA3^{UC} and PNPLA3^{KO} cells. This finding represents the first conclusive evidence from a human experimental system that the I148M variant is a loss of function variant.

Our findings are summarized in Figure 7.1. This study represents the first step in fully elucidating the role of PNPLA3 and its deficiencies in the pathophysiology of NAFLD. Our system has the advantage of being derived from human cells and demonstrating normal PNPLA3 expression patterns and levels. The use of CRISPR/CAS9 allowed us to create isogenic lines of three genotypes to not only explore the mechanistic effects of altered *PNPLA3* genotype on cell function but also to answer the most controversial question in the field: is the I148M variant a loss of function or gain of function variant? The use of hiPSCs to create this model offers the opportunity to explore the effects of *PNPLA3* genotype on the functionality of several different hepatic cell types that play a role in NAFLD development and progression. The work in this thesis has offered new insight into the functionality of PNPLA3 and its pathophysiology in NAFLD. Our system was used to validate clinical findings and identify the hypotheses from previous

publications that are most applicable to a human model of PNPLA3-induced NAFLD. However, a great deal about the function of PNPLA3 in healthy tissue and disease remains obscure, so additional studies are needed. We believe that the *in vitro* system developed in this thesis represents the ideal platform in which to elucidate the full pathophysiology of the PNPLA3 I148M variant in NAFLD development and progression.
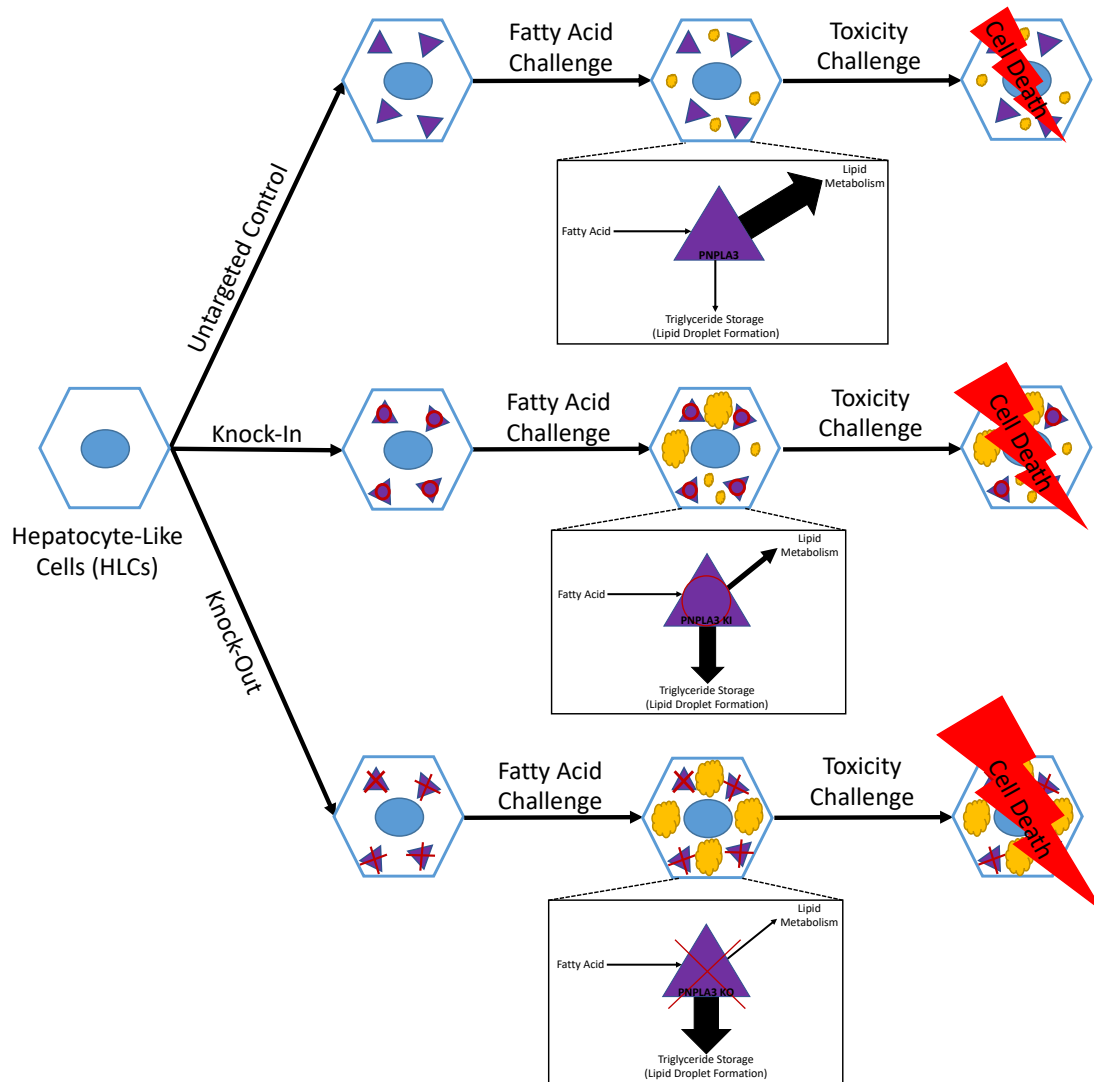
Figure 7.1 Summary of findings.

The three *PNPLA3* genotypes exhibited differential phenotypes when subjected to either a fatty acid or a toxicity challenge. When presented with a fatty acid challenge, PNPLA3[UC] cells upregulate lipid metabolism pathways and few of the fatty acids are esterified into triglycerides. PNPLA3[KO] cells demonstrate the opposite phenotype downregulating lipid metabolism and storing the majority of fatty acids in triglycerides in lipid droplets. This metabolic dysfunction causes PNPLA3[KO] cells to be more susceptible to toxic injury than PNPLA3[UC] cells. In both insult systems, PNPLA3[I148M] cells have an intermediate phenotype which indicates that the I148M variant is a loss of function variant.

7.2 PNPLA3 is a Lipid Droplet Remodelling Protein

Our data supports the hypothesis that PNPLA3 functions as a lipid droplet remodelling protein and that I148M variant is a loss of function variant. It is clear from the lipidomics data that loss of PNPLA3 function results in the sequestration of certain fatty acid moieties, namely PUFAs and SFAs, in triglycerides. This finding in combination with the RNA sequencing data that indicated a global downregulation of lipid catabolism pathways indicates that the lack of lipid droplet remodelling in PNPLA3-edited cells causes the sequestration of fatty acids in metabolically inert triglycerides.

Lipid droplet remodelling is defined as the act of altering the lipid content within the lipid droplet or the composition of the organelle's membrane in order to influence the protein composition of the droplet's surface. Lipid droplets are extremely dynamic structures and the proteins that coat the surfaces of these droplets vary based upon the metabolic needs of the cell. In the fed state, lipid droplets are remodelled to allow for expansion of triglyceride storage and budding. Whereas in the fasted state, lipid droplets are remodelled to recruit lipases and other enzymes to assist in hydrolysis of triglycerides and transport of free fatty acids to metabolic pathways for energy utilization [31, 33, 248]. Loss of PNPLA3 appears to reduce lipid droplet remodelling especially with respect to transfer of PUFAs from triglycerides to phospholipids. By failing to properly regulate the phospholipid composition of the lipid droplet membrane, the protein composition of the droplet's surface may be altered which may prevent the cell from balancing the rates of triglyceride synthesis and lipolysis [128, 163, 164]. Inhibition of lipid droplet remodelling capacity could result in the sequestration of fatty acid moieties in lipid droplets by restricting the hydrolysis of triglycerides. In this way, loss of lipid droplet remodelling function in PNPLA3-edited cells also explains the downregulation of lipid catabolism genes. By requisitioning fatty acids to triglycerides in lipid droplets, these molecules are restricted from entering other metabolic processes. The diminished supply of lipid precursors leads to a global downregulation of metabolic pathways.

The role of PNPLA3 in lipid droplet remodelling has been well established in both clinical and experimental models [125-127, 165]. This is compelling evidence that validates our model as an appropriate system to study the function of PNPLA3 in health and disease.

The function of PNPLA3 has long been controversial because the protein appears to have a wide range of enzymatic functions; PNPLA3 has been hypothesized to function as a lipase, an acyl transferase, and a facilitator of VLDL secretion [109, 118, 121, 122]. Lipid droplet remodelling is a hypothesis that allows us to reconcile this diverse functionality. This hypothesis contends that PNPLA3 catalyses the lipolysis of triglycerides in order to transfer PUFAs to phospholipids. Therefore, PNPLA3 must utilize both its lipase and acyl transferase capacities to facilitate this lipid droplet remodelling in a PUFA direction. Additionally, by controlling the lipid droplet membrane composition, PNPLA3 influences the formation and secretion of VLDLs. Therefore, the findings from our study are in accordance with established findings on the function of PNPLA3 and offer clarity on the role of this enzyme in lipid homeostasis.

7.3   I148M Variant is Loss of Function

Extensive debate remains about the role of the I148M variant as a loss of function or a gain of function variant. This controversy is compounded by the lack of consensus on the function of the protein. Generally, *in vitro* studies using purified PNPLA3 protein indicate that the I148M variant causes a stark reduction in triglyceride lipase and acylglycerol transacylase activities as well as reduced VLDL secretion [109, 121, 122]. These data strongly suggest that the I148M variant is a loss of function variant. However, when the Pnpla3 protein was knocked out in mice, there was no NAFLD-like phenotype observed [160, 161]. Only when the I148M variant was over-expressed and/or knocked-in to the mouse liver did steatosis occur [123, 163]. Leading many to assert that the I148M variant in PNPLA3 was in fact a gain of function variant. While others still claim that the I148M variant is a dominant negative variant due to the accumulation of I148M murine Pnpla3 on the surfaces of lipid droplets [162, 164]. We assert, based upon the results of our model as well as the preponderance of clinical data, that the I148M variant is a simple loss of function variant.

In our experimental model, the phenotype of the PNPLA3^I148M cells unfailingly trended in the same direction as the PNPLA3^KO cells. Both of the PNPLA3-edited lines accumulated more lipid droplets than PNPLA3^UC cells regardless of treatment type, were resistant to palmitic acid-induced ER stress and lipotoxicity, preferentially accumulated

triglycerides containing PUFAs and SFAs, failed to metabolically differentiate between oleic acid and palmitic acid, downregulated lipid metabolic pathways in spite of lipid stimuli, and showed increased susceptibility to ethanol-induced toxicity. Additionally, the transcriptomic analysis of the three genotypes indicates that PNPLA3^{I148M} cells have a gene expression profile that is functionally midway between PNPLA3^{UC} and PNPLA3^{KO} cells. It is clear from these experiments that the I148M variant does not result in a complete loss of PNPLA3 function. We found that the phenotype of PNPLA3^{I148M} cells varies slightly between cell line, genetic background, and HLC differentiation, in some instances more closely resembling either the PNPLA3^{UC} or PNPLA3^{KO} cells. We believe that these variations can be attributed to variable functionality of the PNPLA3 protein. When the PNPLA3 protein has higher functionality, due to a variety of factors such as genetic background and maturity, it more closely resembles the PNPLA3^{UC} phenotype and vice versa. This result is not entirely unexpected since *in vitro* studies of the purified PNPLA3 protein have shown that the I148M variant retains some PNPLA3 functionality. The variant does not disrupt the catalytic serine and thus retains the theoretical ability to catalyse the hydrolysis of glycerolipids. However, the bulkier side chain of methionine reduces access of the substrate to the active site of the enzyme leading to a reduced, but incomplete loss of enzymatic activity. In other words, the V_{max} of the I148M PNPLA3 enzyme is reduced by 100-fold while the K_m remains unchanged [120, 121]. Since the I148M variant retains the enzymatic capacity to catalyse lipolysis, it is possible that the I148M variant may cause a hypomorphic-type phenotype rather than a complete deletion. Thus, the I148M variant has a phenotype that is intermediate between the reference allele and a complete knock-out. These data offer strong evidence that the I148M variant disrupts the functionality of the PNPLA3 enzyme and causes disease via a simple loss of function mechanism in hepatocytes. This loss of function contributes to NAFLD by inducing steatosis through impaired lipid droplet remodelling that results in downregulation of lipid metabolic pathways and increased susceptibility to certain hepatotoxins such as ethanol.

Our data is supported by clinical findings about PNPLA3 and the I148M variant in patients with NAFLD. The original GWAS study by Romeo et al. found that patients with the I148M variant have increased hepatic fat content. In addition, the authors identified three patients with mutations in the *PNPLA3* gene that were likely to be null. These

patients with null *PNPLA3* alleles were also more likely to have very high levels of hepatic triglyceride content [91]. This indicates that complete loss of PNPLA3 function results in a phenotype closely resembling the I148M variant with increased hepatic steatosis.

As seen in our system, carriers of the PNPLA3 I148M variant retain PUFAs in triglycerides in their livers. Carriers of the risk allele have lower circulating levels of PUFA-containing TAGs while liver biopsies reveal that these patients have increased hepatic levels of TAGs containing 3-11 double bonds. This finding was confirmed in an *in vitro* system using human A431 cells with either a complete knock-out of the PNPLA3 protein or with the I148M variant knocked-in. The A431 I148M and knock-out cells had increased lipid droplet formation and preferential sequestration of PUFAs to triglycerides as well as impaired transfer of PUFAs from triglycerides to phospholipids [124, 165]. These data indicate that in human cells, the I148M variant metabolically resembles a complete knock-out of the gene. The metabolomic phenotype of human PNPLA3 I148M cells also closely resembles that of obligate loss of function Pnpla3 S47A mice. Pnpla3 S47A mice have no enzymatic activity due to mutations in the active site of the enzyme. In mice, the S47A variant results in increased levels of long chain and very long chain PUFAs in hepatic triglycerides while the I148M variant causes the opposite trend [127]. These data indicate that the human I148M variant results in a phenotype that closely resembles both human and mouse loss of function, arguing not only that the I148M variant is a loss of function variant but that the mouse is not an appropriate model to study the PNPLA3 I148M variant.

Finally, those who believe that the I148M variant is a gain of function variant claim that reducing the levels of PNPLA3 expression may improve NAFLD in patients. However, studies of *PNPLA3* SNPs that modulate transcription levels of the protein do not support this hypothesis. The minor allele of the *PNPLA3* SNP rs139051 causes reduced mRNA expression levels of *PNPLA3*; however, this lower expression is not significantly associated with improvement in steatosis, steatohepatitis, or fibrosis. This indicates that the pathogenicity of the I148M variant is independent of transcriptional variability [274]. One study did identify the rs2294918 SNP in *PNPLA3*, which is associated with lower

expression of PNPLA3 and a slight improvement in NAFLD-associated traits in carriers of the I148M variant. However, this study also found that in reducing the expression of the I148I PNPLA3 enzyme, the rs2294918 SNP also reduced the protective effect of the reference allele against NAFLD development and progression [275]. This implies that reduced expression of the PNPLA3 protein predisposes patients to NAFLD offering further evidence that reduced PNPLA3 expression and/or functionality is pathogenic. In all, clinical evidence and *in vitro* studies on human cells indicate that loss of PNPLA3 either through null mutations, reduced expression levels, or the I148M variant contributes to NAFLD development.

Though the clinical data strongly supports the hypothesis that the I148M variant is a loss of function variant, the data from murine models has introduced substantial confusion on this topic. Current consensus in the murine research community is that the I148M variant in PNPLA3 is not a simple loss of function variant but requires the presence of a non-functional protein [162, 164]. This notion is based on the finding that complete knock-out of Pnpla3 in mice does not result in a NAFLD-like phenotype [160, 161]. However, this finding fails to account for the substantial sequence divergence of mouse and human PNPLA3 genes and differences in tissue expression of PNPLA3 between mice and humans. Pnpla3 is endogenously expressed at extremely low levels in mouse livers and high levels in adipose tissue [105, 111, 119]. However, it has been shown that Pnpla3 must be (over)expressed in liver tissue for a phenotype to be observed [123]. Fundamentally, the differences in tissue localization as well as the distinct nutritional regulation between the mouse and human PNPLA3 suggests that this protein may have a different function in these two species. Therefore, we contend that mice are not a suitable model system in which to study the function of PNPLA3 and the pathogenicity of the I148M variant. Despite these profound interspecies differences, we contend that data from the murine models does support the hypothesis that the I148M variant is a loss of function variant, or at the very least, a dominant negative variant. The data from mouse models shows that overexpression of a catalytically dead version of Pnpla3 results in a similar phenotype as overexpression of the I148M variant [109, 163, 165]. These data indicate that the I148M variant is equivalent to loss of Pnpla3 catalytic function. Murine studies contend that steatosis in I148M and S47A mice is caused by accumulation of

mutant protein on the surface of lipid droplets which would make the I148M variant a dominant negative variant [162, 164].

The data from our model does not contradict the dominant negative hypothesis wherein the presence of mutated protein exacerbates the negative phenotype caused by the loss of protein function. We did not observe increased expression of I148M protein in our system, but we did not perform intracellular staining, so it remains possible that this accumulation also occurs in our system. Additionally, the lipidomic data from our model did indicate that PNPLA3[I148M] cells accumulate more PUFA-containing triglyceride species than even PNPLA3[KO] cells. This data suggests that the presence of the I148M protein may exacerbate this phenotype by further restricting access of lipid droplet proteins or reducing the mobility of lipids from lipid droplets. Thus, it is possible that the presence of the I148M protein can augment the negative effect of loss of PNPLA3 catalytic function. However, the other experiments in our system indicate that complete loss of PNPLA3 function through genetic elimination causes a more extreme disease phenotype that the I148M variant. Thus, in contrast to murine models, lack of PNPLA3 protein is sufficient to cause a steatotic phenotype which implies that the main pathological effects of the I148M variant are due to a simple loss of catalytic function.

7.4 Our Model and Human Disease

NAFLD can be divided into two related but pathologically distinct features, steatosis and hepatocellular damage leading to inflammation and fibrosis. The I148M variant in PNPLA3 has been conclusively linked to both aspects of NAFLD [91, 132, 133, 135]. In our system, we have been able to clearly demonstrate the role that loss of PNPLA3 function plays in the increased accumulation of intracellular lipid droplet in hepatocytes. However, the other data from our system, namely escape from palmitic acid-induced lipotoxicity and the downregulation of innate immune response and inflammation pathways, are difficult to reconcile with the role of the I148M variant in NAFLD progression. Our data seems to argue that the I148M variant in hepatocytes may cause hepatic steatosis, but the variant may be protective against the hepatocellular damage which is a hallmark of NAFLD progression. We assert that there are several hypotheses that could help explain this seemingly paradoxical result. First, we hypothesize that the

global downregulation of cellular metabolism in hepatocytes with the I148M variant may make these cells more susceptible to hepatotoxicity. Data from our system supports this hypothesis; however, additional studies are needed to confirm its validity. Second, we hypothesize that the decreased lipotoxicity observed in our system may be the *in vitro* manifestation of intact insulin signalling. Finally, we hypothesize that hepatocytes and stellate cells are differentially affected by the I148M variant with hepatocytes manifesting the steatosis and HSCs causing hepatocellular damage, inflammation, and fibrosis. Each of these hypotheses will be discussed at length in the following sections.

7.4.1 PNPLA3 Edited Cells are More Susceptible to Other Forms of Hepatotoxicity

NAFLD is an extremely complex, multifactorial disease and the temporal order of insults has long been up for debate. For decades, the two-hit hypothesis drove NAFLD research asserting that steatosis was the "first hit" that sensitized the liver to a "second hit" of oxidative stress or mitochondrial dysfunction that caused inflammation [276, 277]. Recently, researchers have begun to discount this hypothesis as too simplistic. Nothing in biology can be explained by a simple two-step process. Researchers now accept the slightly more complex, though largely similar, multi-hit hypothesis [63, 278]. Simply, this hypothesis asserts that genetics, epigenetics, environmental factors, and a fair amount of bad luck combine to increase the susceptibility of a patient to progressive disease. The multi-hit hypothesis acknowledges that liver disease does not happen in a vacuum. The liver is constantly bombarded with insults as it is one of the body's first defences against toxic substances whether they originate from endogenous or exogenous sources. It has been well established that patients with NAFLD are more susceptible to drug induced liver injury because the hepatocytes of these individuals are damaged and preoccupied with lipotoxic insults from overwhelming lipid accumulation within the liver [262, 265, 267]. In the spirit of the multi-hit hypothesis, we chose to examine the effect of *PNPLA3* genotype and steatosis on drug metabolism in our system.

The I148M variant has been extensively correlated with increased disease severity for nearly every known chronic liver disease from NAFLD to ALD to viral hepatitis to hereditary hemochromatosis [91, 135, 139, 141, 144]. In addition, this disease variant has also been linked to drug induced liver damage in paediatric patients undergoing

chemotherapy for acute lymphoblastic leukaemia [224, 225]. The contribution of this variant to the pathogenesis of this diverse list of diseases indicates that PNPLA3 plays a role in more than just steatosis progression. We hypothesize that the metabolic dysregulation caused by loss of PNPLA3 function may sensitize hepatocytes to other forms of hepatotoxicity. Therefore, the I148M variant causes progressive liver damage not through lipid-induced hepatocellular damage but rather through reduced detoxification capacity.

Our data, in large part, supports this hypothesis. We found that in addition to the global downregulation of metabolic pathways related to lipid metabolism, there was a similar downregulation in genes responsible for drug detoxification in PNPLA3-edited cells. Reduced expression of DMEs has the potential to cause accumulation of toxic compounds in hepatocytes and impair the ability of these cells to metabolize or excrete said compounds. This accumulation then leads to oxidative stress, mitochondrial dysfunction, and ultimately cell death. Patients with advanced NAFLD/NASH have reduced expression of DMEs and it is believed that this reduction is a contributing factor to the increased susceptibility to drug induced liver injury [262, 263, 267, 268]. Given this global downregulation of DMEs and the clinical evidence that this reduction contributes to liver injury, we compared the effect of *PNPLA3* genotype on the toxicity of three common hepatotoxins, alcohol, iron, and acetaminophen, in our system. We found that PNPLA3-edited cells were more susceptible to ethanol-induced toxicity; however, a similar trend was not present in the iron and acetaminophen toxicity assays. There are several potential explanations for this discrepancy. It is possible that the drug metabolism in our cells is not functionally mature enough to test this hypothesis reliably. The HLCs in our system are immature and phenotypically resemble foetal hepatocytes. Therefore, it is possible that differences in drug metabolism are not captured in our *in vitro* system due to this functional immaturity. Alternately, it is possible that iron and acetaminophen detoxification are not affected by *PNPLA3* genotype, but the metabolism of other drugs is affected. The increased toxicity of ethanol in PNPLA3-edited cells supports this hypothesis. We propose additional toxicity studies to fully elucidate the role of *PNPLA3* genotype in drug metabolism. For example, it would be prudent to examine the toxicity of the chemotherapeutic drugs that have been linked to increased toxicity in patients with

the I148M variant. Additionally, more mechanistic studies should be performed to understand how PNPLA3 genotype influences ethanol toxicity.

Data from our *in vitro* system indicates that loss of PNPLA3 function reduces the capacity of hepatocytes to metabolize drugs and other toxic metabolites. This reduced metabolic capacity increases the susceptibility of these cells to hepatotoxicity and drug induced liver injury. This increased hepatotoxicity does appear to be specific to certain drugs such as ethanol. Thus, we contend that carriers of I148M variant have a global downregulation of lipid metabolism that makes them less susceptible to lipotoxicity; however, these patients also have reduced drug metabolism capacity which makes them more susceptible to other forms of hepatotoxicity. Additional studies are needed to fully analyse the accuracy of this hypothesis. However, if this hypothesis is proven true, it could have major implications for the clinical management of patients. In addition to diet and exercise to decrease the NAFLD disease burden, the alcohol consumption and drug cocktail used to treat comorbidities would need to be carefully monitored.

7.4.2 PNPLA3 Causes NAFLD Without Insulin Resistance

Selective insulin resistance is a hallmark of NAFLD caused by metabolic disease. Insulin resistance in the liver causes uncontrolled de novo lipogenesis and hepatic glucose production. The increased concentration of glucose and free fatty acids derived from gluconeogenesis and de novo lipogenesis, respectively, leads to glucotoxicity and lipotoxicity in hepatocytes [64]. The metabolic dysregulation in insulin resistant hepatocytes leads to oxidative stress and mitochondrial dysfunction which causes inflammation through the activation of JNK and NF-κB. This inflammation then further exacerbates insulin resistance causing a vicious circle of liver damage inducing more liver damage [79-81]. However, despite being linked to increased disease severity, NAFLD patients carrying the PNPLA3 I148M variant can develop severe NAFLD in the absence of insulin resistance [138]. The mechanism by which these patients can progress toward severe disease without underlying insulin resistance remains a mystery; however, the lack of insulin resistance in I148M carriers indicates that the pathophysiology of PNPLA3-induced NAFLD is functionally distinct from obesity-induced NAFLD. Thus, it is not altogether unsurprising that PNPLA3-edited cells behave differently from genetically normal cells when presented with a lipid stressor.

A major effector of insulin signalling on lipid and glucose metabolism is the AKT signalling cascade. At high insulin levels, AKT is activated through phosphorylation. AKT phosphorylation then leads to upregulation of SREBP and downstream de novo lipogenesis genes as well as inhibition of gluconeogenesis via FOXO1 phosphorylation [21]. Since AKT is the major inhibitor of FOXO1 and downstream hepatic glucose production, disruption to this signalling cascade results in constitutively active gluconeogenesis [64]. However, SREBP is controlled by LXR as well as AKT so even in the absence of proper AKT signalling, insulin still maintains the ability to activate de novo lipogenesis [28]. Previous studies have shown that treatment with palmitic acid causes reduced insulin-mediated phosphorylation of AKT. However, by upregulating ChREBP or SCD1, cells increased the ratio of MUFAs to SFAs. This had the dual effect of increasing the incorporation of SFAs into triglycerides while restoring AKT signalling and insulin sensitivity. The enhanced insulin sensitivity was accompanied by decreased mitochondrial β-oxidation as well as reduced inflammation despite increased steatosis [29, 30]. Thus, insulin sensitivity can be maintained through increased incorporation of SFAs into triglycerides which restores AKT signalling and results in reduced β-oxidation, lipotoxicity, and inflammation. Therefore, we hypothesize that the reduced lipotoxicity seen in PNPLA3-edited cells treated with palmitic acid is the *in vitro* manifestation of intact insulin signalling.

The lipidomic data from PNPLA3-edited cells indicates that loss of PNPLA3 activity causes increased incorporation of SFAs into triglycerides as well as retention of PUFAs in triglycerides. Both of these observations align with clinical evidence showing that I148M carriers often maintain insulin sensitivity, have increased hepatic retention of PUFAs, and have reduced expression of SREBP1c and downstream de novo lipogenesis [138, 165, 241]. We believe that the incorporation of these two fatty acid moieties into triglycerides maintains insulin sensitivity in the HLCs. The increased incorporation of SFAs into triglycerides may prevent the deleterious effect of these fatty acids on insulin mediated AKT phosphorylation while the high concentration of PUFAs in the triglycerides of these cells leads to a downregulation of SREBP1c and downstream de novo lipogenesis genes. Therefore, it is possible that the reduced lipotoxicity of palmitic

acid in PNPLA3-edited cells is merely an epiphenomenon caused by the effort of these cells to maintain insulin sensitivity. This hypothesis is supported by the downregulation of metabolic as well as inflammatory pathways. By preventing insulin resistance, loss of PNPLA3 function dissociates steatosis from hepatocellular damage and inflammation. Additional studies are needed to understand the mechanism by which the I148M variant maintains insulin sensitivity.

7.4.3 PNPLA3 Differentially Affects Hepatocytes and Hepatic Stellate Cells to Manifest Different Aspects of the Disease

The intact insulin signalling in some PNPLA3 I148M carriers, however, is at odds with the increased NAFLD severity and higher necro-inflammatory scores of these patients. Inflammation and disease progression in NAFLD patients are usually inversely proportional to insulin sensitivity [56, 57, 276]. Given the clinical evidence as well as the data from our system, it appears that increased steatosis and increased hepatocellular damage are functionally dissociated in the livers of patients with the I148M variant in PNPLA3. The downregulation of metabolic pathways that generally lead to lipotoxicity as well as reduced inflammatory gene expression in PNPLA3-edited cells argue against a direct role of the I148M variant in causing hepatotoxicity, despite increased steatosis. However, the I148M variant has been linked to HSC activation as well as increased inflammation and fibrogenesis [113, 114, 132, 135]. Therefore, we hypothesized that the I148M variant affects hepatocytes and HSCs differently, causing each cell type to manifest a different aspect of the disease. We hypothesize that the I148M variant causes a downregulation of metabolism in hepatocytes which results in increased steatosis as well as reduced lipotoxicity. Meanwhile, the I148M variant causes activation of HSCs which are responsible for the inflammation, fibrosis, and general disease progression.

The data from our *in vitro* system is consistent with this hypothesis. We found that loss of PNPLA3 function, either through the I148M variant or a complete PNPLA3 protein knock-out, led to increased steatosis. The increased lipid accumulation that characterized the PNPLA3-edited cells was caused by concomitant upregulation of triglyceride synthesis and downregulation of lipid catabolic pathways. The increased steatosis in these cells did not appear to harm the HLCs in any way. In fact, the increased steatosis appeared to protect the HLCs from ER stress and lipotoxicity caused by palmitic acid. These

findings support the growing body of evidence showing that steatosis is not necessarily cytotoxic and likely represents an epiphenomenon in NAFLD pathogenesis [35, 117, 279, 280]. In addition to the findings of our system, clinical evidence also supports this hypothesis. Clinical studies have correlated the I148M variant in PNPLA3 to steatosis, inflammation, fibrosis, cirrhosis, and hepatocellular carcinoma in numerous aetiologies of liver disease. The variant is less often correlated with markers of hepatocyte damage such as hepatocellular ballooning and the presence of Mallory bodies [132, 151]. These clinical data indicate that the liver damage in carriers of the I148M variant may be independent of hepatocyte damage.

The I48M variant in PNPLA3 confers increased risk for fibrosis independent of other markers of disease severity [132, 135]. Therefore, we contend that HSCs are responsible for liver damage in carriers of the I148M variant independent of steatosis in hepatocytes. PNPLA3 participates in HSC activation by promoting the extracellular release of retinol from HSCs, the first step in activation [112]. HSCs with the I148M variant retain less retinol than their I148I counterparts which causes higher expression of profibrotic genes [114]. Additionally, HSCs expressing the I148M variant have higher proliferation rate and chemotaxis, both hallmarks of activated HSCs. HSCs with the I148M variant also secrete proinflammatory markers such as CCL2, CCL5, and TGFβ which causes increased recruitment of inflammatory cells like macrophages [113]. Given these data, it is reasonable to hypothesize that HSCs control the inflammatory and fibrotic response in PNPLA3-induced NAFLD. This hypothesis offers insight into how the I148M variant increases the risk of severe liver damage in several different aetiologies of liver disease despite divergent pathogenesis. By conferring a proinflammatory and profibrotic phenotype to HSCs, the I148M variant contributes to progressive liver damage independent of other disease processes.

We believe that this hypothesis offers the most reasonable explanation for the role of the I148M variant in chronic liver diseases such as NAFLD given the data from our system and the existing clinical and experimental evidence. Patients with the I148M variant have increased hepatic steatosis and the triglycerides stored in the livers of these patients are enriched for PUFAs. The high concentration of PUFAs as well as the insulin sensitivity

in the livers of these patients are cytoprotective from lipid-induced stress, as confirmed in our *in vitro* system. Though the global downregulation of lipid metabolism that results in triglyceride accumulation may be cytoprotective against lipotoxicity, it is accompanied by a concomitant downregulation of xenobiotic metabolism which may reduce the detoxification capacity of hepatocytes and result in increased sensitivity to other hepatotoxins such as ethanol. Additionally, the predilection of the HSCs in carriers of the I148M variant toward a proinflammatory, profibrotic phenotype causes inflammation, hepatocellular damage, and progressive fibrosis in these patients. It is possible that the steatosis in I148M carriers is an epiphenomenon or that the increased steatosis in the hepatocytes sensitizes these cells to hepatocellular damage mediated by activated HSCs. In order to fully elucidate the pathogenic mechanism of the PNPLA3 I148M variant, the cell to cell interactions between hepatocytes and HSCs carrying the risk allele must be fully characterized.

7.5   Limitations of Our Model

Our system is the first of its kind to model PNPLA3-associated NAFLD. By utilizing CRISPR/CAS9 and hiPSCs, we were able to create an *in vitro* system to analyse the effect of PNPLA3 genotype on HLCs and their response to lipid induced stress. Our system is the first to model the I148M variant at endogenous levels in a fully human system. However, despite the numerous unique advantages, our system still has quite a few limitations.

First, the differentiated HLCs are not fully mature and more closely resemble foetal hepatocytes than adult hepatocytes. Therefore, it is possible that the metabolic phenotype of these cells may not be fully representative of the adult phenotype or disease. However, the HLCs in our system did respond to lipid supplementation in a predictable manner. The PNPLA3^{UC} cells became steatotic when treated with the MUFA oleic acid and suffered from lipotoxicity when treated with the SFA palmitic acid. Thus, we are reasonably confident that the HLCs in our system have a functional lipid metabolism. Additionally, the HLCs were susceptible to drug induced toxicity when treated with both ethanol and acetaminophen indicating that the drug metabolism of these cells was intact. So even though these cells are not fully functionally mature, they are metabolically active

and represent a reasonable model system to study the role of PNPLA3 in NAFLD progression.

Second, the *in vitro* nature of this system lends a certain degree of artificiality to the model. In our system we used two different fatty acid species to model different aspects of NAFLD. However, *in vivo*, hepatocytes would not encounter a single moiety of fatty acid. Instead, hepatocytes would be subjected to a complex mixture of several different fatty acids belonging to each of the different fatty acid classes, SFAs, MUFAs, and PUFAs. It is possible that this complex, dynamic nutritional environment would alter the phenotype of PNPLA3-edited cells. Additionally, we examined only the effect of fatty acid supplementation in our system. Other nutritional factors such as fructose have been shown to significantly affect NAFLD development and progression [55, 63, 81]. In order to confirm the phenotype of PNPLA3-edited cells, additional studies are needed to examine the effects of complex nutritional signalling on steatosis and cell stress pathways in these cells. Beyond the lack of nutritional diversity, the HLCs in our *in vitro* system are not exposed to the complex network of cytokines and adipokines that influence the metabolic behaviour of hepatocytes *in vivo*. NAFLD is a multi-system disease that involves significant cross talk between several metabolic organs including the liver, adipose tissue, the gut, and the pancreas [56, 151]. Hepatocytes in the proper signalling environment may exhibit an entirely different phenotype that we are unable to capture in our *in vitro* system. However, the relative simplicity of this system does allow us to dissect the role of the I148M variant in hepatocytes without introducing confounding factors that may mask the mechanism of action. In this way, the simplicity may offer unique advantages as well.

Third, this thesis only examined the effect of *PNPLA3* genotype on the phenotype of hepatocytes. NAFLD is a complex metabolic disease that involves intercellular signalling between several hepatic cell types. PNPLA3 is known to be expressed in several cell types that are relevant to NAFLD including hepatocytes, stellate cells, and macrophages. Our system offers the opportunity to dissect the role of PNPLA3 pathogenesis on hepatocytes from the other hepatic cells in which it is expressed. As the main metabolic cell of the liver, hepatocytes are the site of steatosis which is a hallmark of NAFLD. In this work,

we demonstrated that loss of PNPLA3 function is a direct cause of the increased steatosis found in the livers of I148M carriers. However, we also demonstrated that loss of PNPLA3 is protective against lipotoxicity which is counterintuitive to the data linking I148M to more severe disease phenotypes, though this phenomenon may be unique to our *in vitro* challenge model and not fully reflective of *in vivo* response. It is possible that unique signals are produced by hepatocytes carrying the I148M variant that leads to increased recruitment or activation of macrophages and stellate cells. This nuance would not be captured in our model which has only a single cell type. Alternately, the I148M variant could uniquely influence the phenotypes of other hepatic cell types toward a proinflammatory or profibrotic state leading to increased liver damage and disease progression. In order to fully understand the pathogenesis of this disease and the contribution of PNPLA3, additional studies are needed. Recent advances in 3D organoid culture with many cell types may provide additional opportunities to further elucidate the mechanism of PNPLA3 in NAFLD.

8 FUTURE DIRECTIONS AND CONCLUSIONS

8.1 Future Directions

In this thesis, we developed an *in vitro* culture system using CRISPR/CAS9 and hiPSCs to study the effect of *PNPLA3* genotype on NAFLD development in hepatocytes. We were able to replicate several key clinical findings including increased steatosis in PNPLA3-edited cells as well as the unique lipidomic signature of hepatic retention of PUFAs in triglycerides. Additionally, we confirmed for the first time in a fully human model system that the I148M variant is a loss of function variant that causes steatosis by downregulating metabolic pathways likely through reduced lipid droplet remodelling capabilities. The findings from this thesis also raised several questions about (1) how loss of PNPLA3 function leads to the global downregulation of lipid metabolism, (2) why loss of PNPLA3 function appears to be cytoprotective against lipotoxicity in hepatocytes, (3) how increased steatosis in PNPLA3-edited cells may lead to the increased liver injury that characterizes PNPLA3-induced NAFLD, (4) what role the PNPLA3 I148M variant plays in the phenotype of other hepatic cell types, and (5) how these other cell types may contribute to the pathogenesis of NAFLD. The work described in this thesis provides a first step in understanding the pathogenesis of PNPLA3-induced NAFLD; however, additional studies are needed to examine the mechanism by which the I148M variant in PNPLA3 causes increased steatosis and liver damage to induce NAFLD development and progression.

8.1.1 In-Depth Analysis of Signalling Pathways Influenced by PNPLA3

The RNA sequencing of the three genotypes of HLCs revealed an interesting pattern of downregulated genes involved in lipid homeostasis. This analysis in addition to the increased steatosis and reduced lipotoxicity in PNPLA3-edited cells suggests that loss of PNPLA3 function results in downregulation of lipid catabolic genes in favour of triglyceride formation. However, additional studies are needed to identify that signalling cascades that are directly affected by loss of PNPLA3. Very little is known about the role of PNPLA3 in the larger network of lipid metabolism and this *in vitro* system could be used to analyse the downstream genes that are influenced by *PNPLA3* genotype. Using the PNPLA3^{UC} and PNPLA3^{KO} lines, it may be possible to parse which signalling networks PNPLA3 participates in and which genes are directly affected by PNPLA3 functionality.

In addition, this system can be used to understand how carriers of the PNPLA3 I148M variant may be able to maintain insulin sensitivity in the context of obesogenic stimuli. It is possible that the inability of PNPLA3-edited cells to differentiate between saturated and unsaturated fatty acids are involved in this process. SFAs and the downstream cell stress responses have been directly implicated in the development and exacerbation of insulin resistance [81, 281]. In order to explore this possibility, we would need to first confirm that the insulin signalling pathway is intact and functional in our HLCs. Following such confirmation, insulin sensitivity assays could be used to analyse how PNPLA3^{UC} and PNPLA3-edited cells are differentially affected by oleic acid and palmitic acid supplementation. In this way, the *in vitro* system could easily be used to test the hypothesis that escape from palmitic acid-induced lipotoxicity is the *in vitro* manifestation of insulin sensitivity.

The relative simplicity of this *in vitro* system would be a major asset in the mechanistic analysis of the role of PNPLA3 in lipid metabolism and insulin signalling. Since the I148M variant has been proven to be a loss of function variant in our system, the PNPLA3^{KO} cells can be used to cleanly dissect the role of PNPLA3 in hepatocytes. Additionally, the supplementation of a single fatty acid to the *in vitro* system will allow us to understand with fine granularity how the saturation status of a lipid affects PNPLA3 and its downstream targets.

8.1.2 Proteomic Analysis of Lipid Droplets

We hypothesized based upon our lipidomics data and the available clinical data that the I148M variant causes a loss of lipid droplet remodelling function. Specifically, the loss of PNPLA3 function prevents the transfer of PUFAs from triglycerides to phospholipids causing a hepatic retention of these fatty acids. We hypothesized that this lack of lipid droplet remodelling affected the access of other lipid droplet proteins to the fatty acids contained in lipid droplets. Thus, causing a reduced flux of lipid precursors to metabolic pathways and resulting in the downregulation of the affected pathways. In order to test this hypothesis, we would need to perform proteomic analysis on the lipid droplets of PNPLA3^{UC} and PNPLA3-edited cells. This proteomic analysis would allow us to

determine if the reduced lipid droplet remodelling capacity of PNPLA3-edited cells alters the protein coating of lipid droplets. Further analyses of the differences in lipid droplet proteins would allow us to determine if loss of PNPLA3 reduces metabolic processes through an indirect mechanism by reducing lipid transport out of lipid droplets or through a more direct mechanism.

8.1.3 Long-Term Culture to Model Chronic Nature of NAFLD

Our analyses have demonstrated that the loss of PNPLA3 function has a profound effect on the transcriptomic and lipidomic profiles of HLCs. Loss of PNPLA3 function causes global downregulation of metabolic pathways and improper responses to lipid-induced stress. We hypothesize that this metabolic dysregulation will have a negative impact on the long-term-health of the cell. Since NAFLD is a chronic disease that takes decades to develop and progress toward end-stage liver disease, it is likely that our model, which has the longevity of only one week, did not capture the long-term effects of metabolic dysregulation. Additional studies are needed to examine how the global downregulation of lipid metabolism affects ER stress and inflammatory pathways over time. In order to accomplish these studies, the *in vitro* system would need to be adjusted to allow for longer term culture. Once the longevity of the culture can be extended, time-course analyses should be undertaken to analyse the effect of *PNPLA3* genotype on viability, lipotoxicity, and inflammation of palmitic acid treated cells to understand how metabolic dysregulation may contribute to long-term hepatocellular damage and disease progression.

8.1.4 Additional Drug Toxicity Assays

Our preliminary results indicate that loss of PNPLA3 function may increase the susceptibility of hepatocytes to drug-induced cytotoxicity. Clinical evidence supports this hypothesis as carriers of the I148M variant are more susceptible to liver damage induced by ethanol and chemotherapeutic drugs. In our system, we found that PNPLA3-edited cells downregulated the expression of DMEs and were more susceptible to ethanol-induced toxicity. However, the PNPLA3-edited cells were not more susceptible to toxicity induced by iron overload or acetaminophen overdose. Thus, additional studies are needed in order to confirm or refute this hypothesis as a potential mechanism for increased liver injury in patients with the I148M variant.

It would be important to confirm the increased toxicity of ethanol in PNPLA3-edited cells as well as explore the mechanism by which ethanol causes toxicity in the context of PNPLA3 loss of function. Ethanol is consumed at non-harmful doses by a large portion of the population but if the results in this thesis are confirmed, even small doses of ethanol could be extremely deleterious to patients carrying the I148M variant. Thus, this finding could have significant clinical implications for the management of NAFLD patients with the I148M variant. In addition, we propose testing the toxicity of several other drugs that have the potential to cause hepatocellular damage in patients with the I148M variant. The toxicity of drugs that have been connected to hepatotoxicity in I148M carriers clinically as well as drugs that are specifically metabolized by phase one DMEs that were specifically downregulated in the PNPLA3-edited HLCs should be assessed.

8.1.5 Co-Culture Model for a More Physiologically Relevant Model of NAFLD

NAFLD induced by the I148M variant in PNPLA3 has implicated several cell types in the pathogenesis of the disease. Hepatocytes are the main metabolic hub in the liver and the site of steatosis. HSCs are the hepatic cell type responsible for fibrosis and these cells also play an important role in coordinating the inflammatory response in NAFLD. Finally, macrophages coordinate with immune cells and HSCs to control the proinflammatory environment that causes disease progression and characterizes NASH. The two most commonly associated hepatic cell types with PNPLA3-induced NAFLD are hepatocytes and HSCs; however, preliminary studies in our lab suggest that PNPLA3 is also expressed in macrophages and thus could influence the inflammatory aspect of the disease (data not shown). The interplay of each of these cell types is important in the pathogenesis of NAFLD and therefore, all three of these cell types would be necessary to properly model NAFLD progression *in vitro*. We propose expanding on the findings in this thesis by creating a more physiologically relevant co-culture model of PNPLA3-induced NAFLD. The three genotypes of hiPSCs could be differentiated into hepatocytes, to model the metabolic aspects of NAFLD, HSCs, to model fibrosis, and macrophages, to model inflammation before combining the three hepatic cell types together into a single co-

culture system to model PNPLA3-induced NAFLD *in vitro*. A schematic of the proposed co-culture model is shown in Figure 8.1.
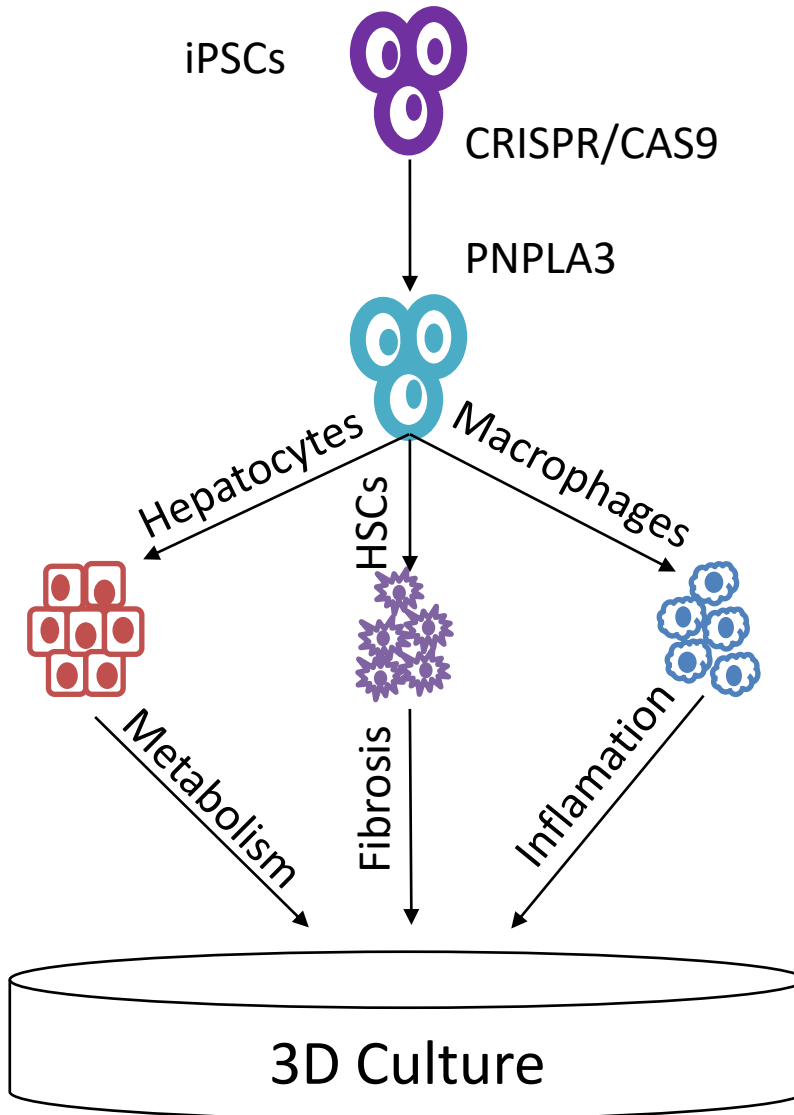
Figure 8.1 Co-Culture Model Schematic.

Normal hiPSCs were edited using CRISPR/CAS9 in order to generate PNPLA3^{UC}, PNPLA3^{I148M}, and PNPLA3^{KO} hiPSC lines. These lines can then be differentiated into the three hepatic cell types that are most relevant for NAFLD development and progression: hepatocytes to represent the metabolic aspects of the disease, HSCs to model fibrosis, and macrophages to model inflammation. Following differentiation, the three cell types can be combined into a 3D co-culture to create an *in vitro* physiological model of PNPLA3-induced NAFLD.

Well-characterized differentiation protocols exist for each of the proposed hepatic cell types making the creation of this model readily feasible [184, 194, 282]. This system would capture the complex cellular interactions and intercellular signalling that cause progression from simple steatosis through NASH and fibrosis. In addition, since each of the cell types would be differentiated from the same hiPSC line, the system would be derived from a single genetic background which eliminates potential confounding variables from the model. This co-culture system shows the true advantage of using hiPSCs and CRISPR/CAS9 to study the effects of *PNPLA3* genotype on liver disease as multiple hepatic cell types can be derived from a single line. This physiologically relevant 3D co-culture model would be the first of its kind that simulates the effect of genotype on a complex chronic liver disease. A model of this type would largely reduce the need to use the less appropriate murine model to understand the function of PNPLA3 and the contribution of the I148M variant to NAFLD development and progression.

8.1.6 CRISPRi Screen for Mechanistic Insight

As extensively discussed throughout this thesis, the function of PNPLA3 in healthy tissues and NAFLD pathogenesis is subject to great debate in the scientific community. Over a decade of research has failed to fully characterize this protein despite its huge impact on public health. We propose the use of our *in vitro* system as well as the CRISPRi screening technology to help elucidate the mechanism by which PNPLA3 influences metabolism as well as identify potential druggable targets to reduce NAFLD risk in patients carrying the I148M variant. A CRISPRi screen performed on PNPLA3^{UC} cells would allow us to identify genes that contribute to the increased steatosis and reduced lipotoxicity that characterize PNPLA3-edited cells. The PNPLA3^{I148M} and PNPLA3^{KO} cells could be used as a unique internal control to dissect incidental findings from genes truly associated with PNPLA3. The CRISPRi system offers the opportunity to quickly and easily analyse which metabolic pathways PNPLA3 may play a role in as well as downstream genes that are repressed by loss of PNPLA3 function which are responsible for the phenotypic behaviour of PNPLA3-edited cells. Alternately, the CRISPRi screen could be used on PNPLA3^{KO} cells to identify potential drug targets that improve the disease phenotype. This CRISPRi screen would offer novel insight into treatment options that are unique to carriers of the I148M variant in PNPLA3.

8.1.7 Humanized Mouse Model

As discussed, *in vitro* disease models have an inherent amount of artificiality due to the difficulty in modelling the complex, multi-system crosstalk that happens *in vivo*. In order to overcome this challenge, we propose the development of humanized mouse models using cells derived from the PNPLA3^{UC}, PNPLA3^{I148M}, and PNPLA3^{KO} hiPSC lines. As described for the proposed *in vitro* co-culture model, hiPSCs from each of the three genotypes could be differentiated into relevant hepatic cell types before being injected into the livers of humanized mice for engraftment [283, 284]. Following engraftment, the mice could be fed with a high fat diet to cause a NAFLD phenotype before being sacrificed and having their livers collected for histology. This humanized mouse model would take advantage of the physiological *in vivo* environment without the confounding factor of differential PNPLA3 expression and functionality between humans and mice. This model would allow us to study the effect of organ to organ crosstalk on the development and progression of NAFLD for each of the *PNPLA3* genotypes. Additionally, the humanized mouse model represents the ideal model system to test the long-term effects of PNPLA3-induced metabolic dysregulation on NAFLD phenotype because mice can live for several months as opposed to weeks as with *in vitro* models. This humanized mouse model would be a massive contribution to the field and would represent the most physiologically relevant model of human PNPLA3-induced NAFLD to date.

8.2   Conclusions

According to the University of Cambridge guidelines for the award of a PhD Degree, the candidate must demonstrate that his or her dissertation "represents a significant contribution to learning, for example through the discovery of new knowledge, the connection of previously unrelated facts, the development of new theory, or the revision of older views". To this end, this dissertation has accomplished such criteria through the following

- Developing an *in vitro* human model to elucidate the role of PNPLA3 in NAFLD development and progression

- Conclusively demonstrating that the I148M variant in PNPLA3 is a loss of function variant that causes steatosis in hepatocytes

- Discovering that cells lacking PNPLA3 functionality are resistant to lipotoxicity induced by SFAs due to increased incorporation of these fatty acids into triglycerides

- Confirming the role of PNPLA3 in remodelling phospholipids in lipid droplets in a PUFA direction

- Demonstrating that loss of PNPLA3 function results in a global downregulation of genes involved in the metabolism of lipids and drugs

- Discovering that loss of PNPLA3 function causes increased susceptibility to hepatotoxins such as ethanol

Collectively, the results presented in this thesis have advanced both the knowledge and tools available to the field and will significantly contribute to several exciting future developments in the years to come.

10 APPENDICES

Figure 10.1 Sanger Sequence of FSPS13B Untargeted Control Clone #1
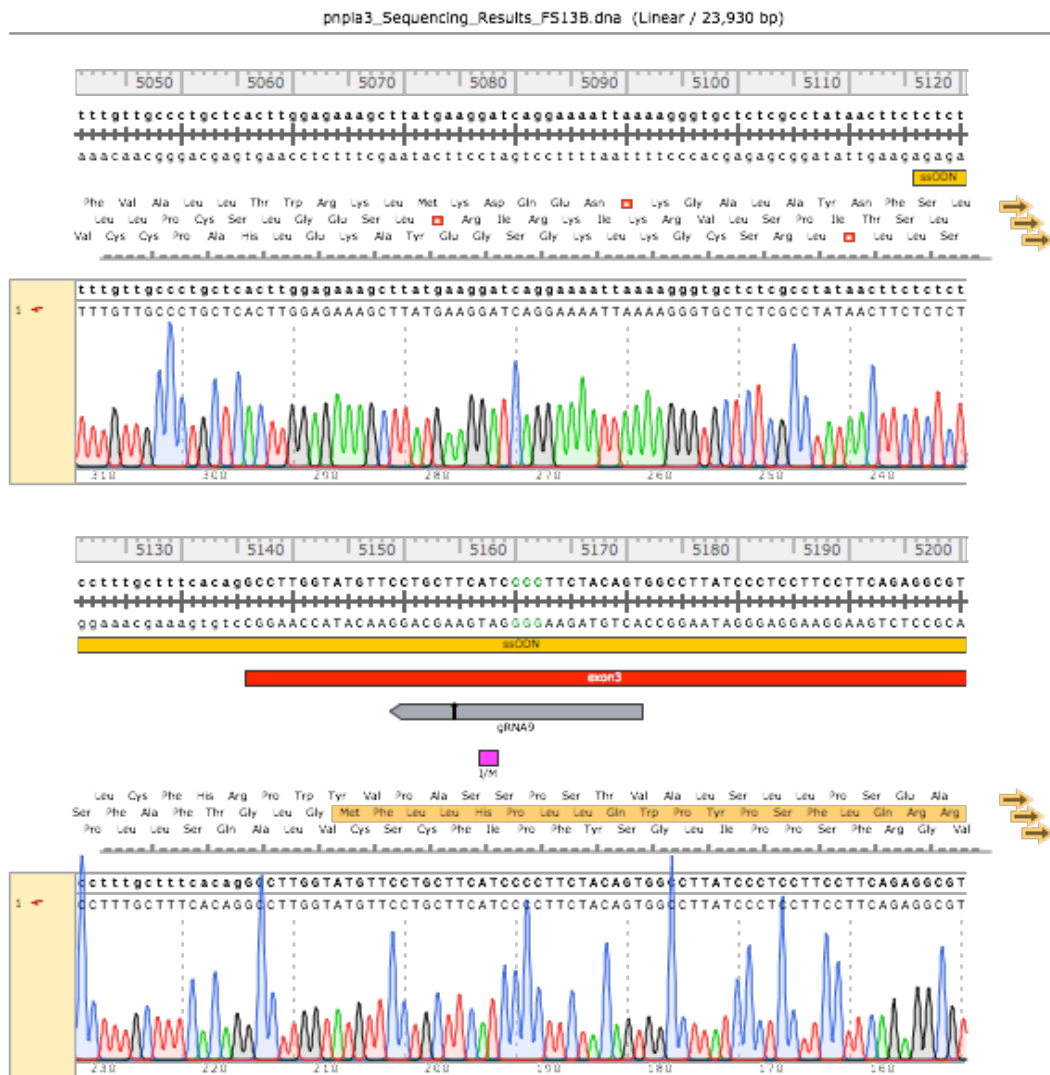
Figure 10.2 Sanger Sequence of FSPS13B Untargeted Control Clone #2

pnpla3_Sequencing_Results_FS13B.dna (Linear / 23,930 bp)
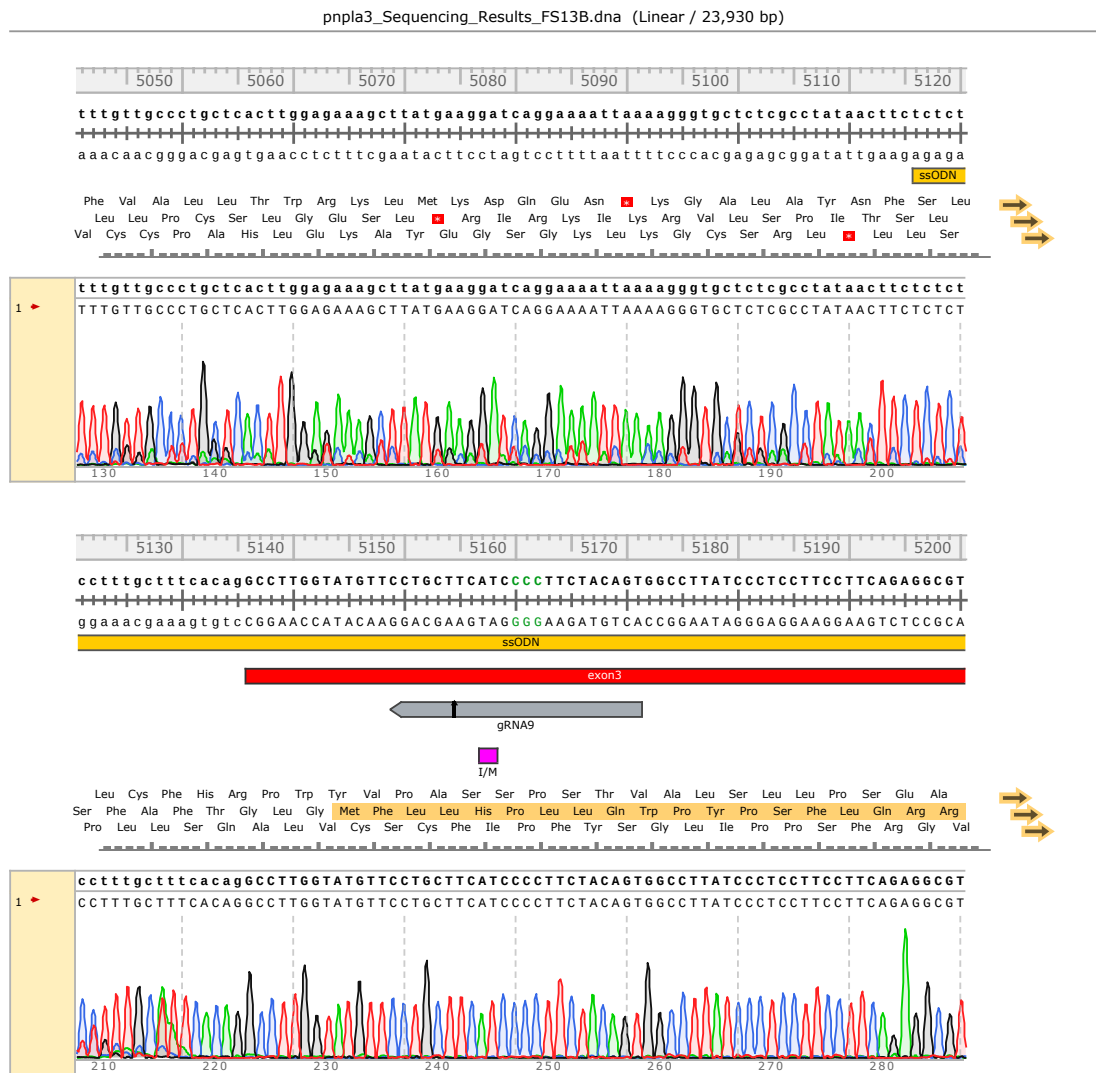
Figure 10.3 Sanger Sequence of FSPS13B I148M Clone #1

Samantha G. Tilson - January 2021
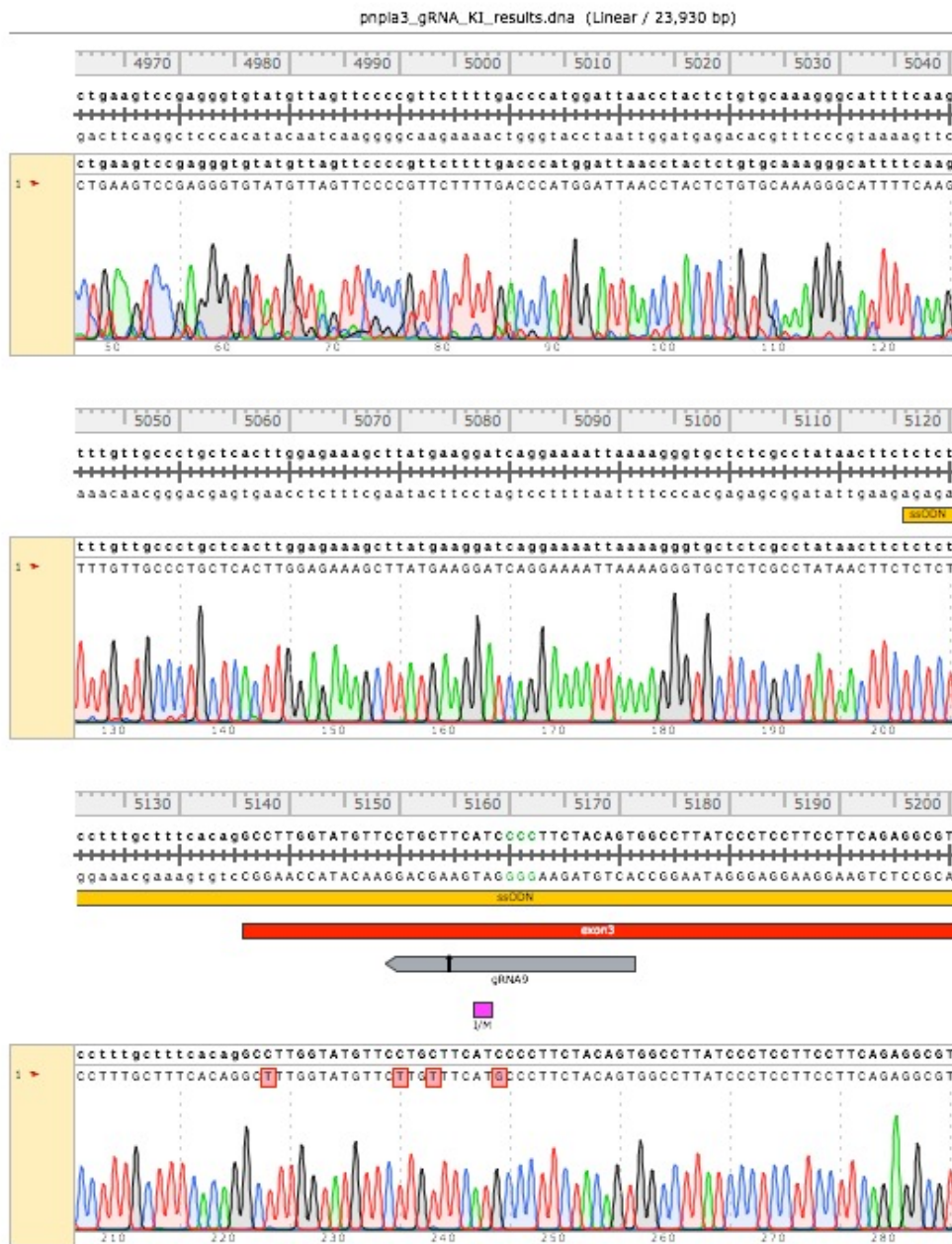
Figure 10.4 Sanger Sequence of FSPS13B I148M Clone #2
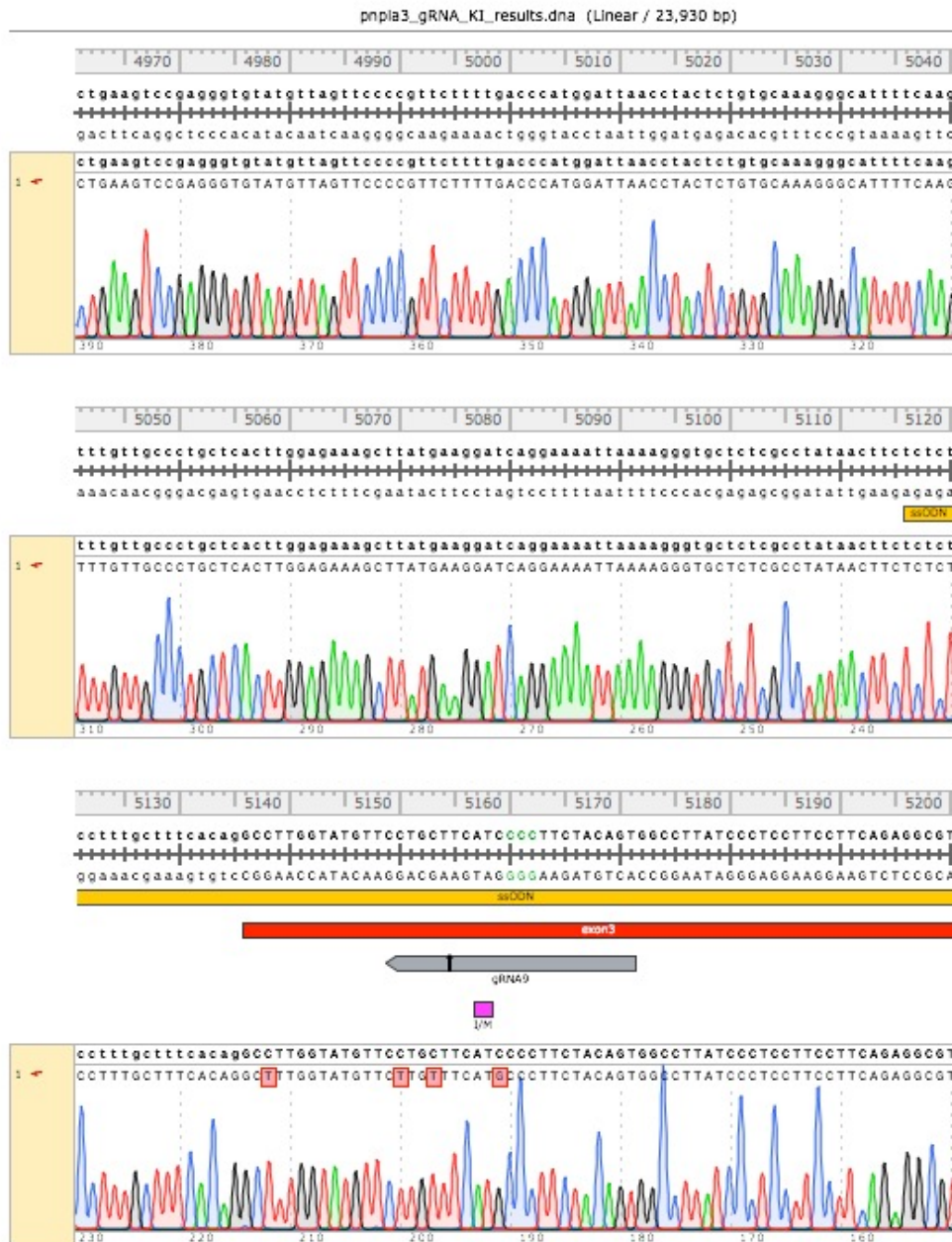
Figure 10.5 Sanger Sequence of FSPS13B Knock-Out Clone #1

pnpla3_Sequencing_Results_FS13B.dna (Linear / 23,930 bp)
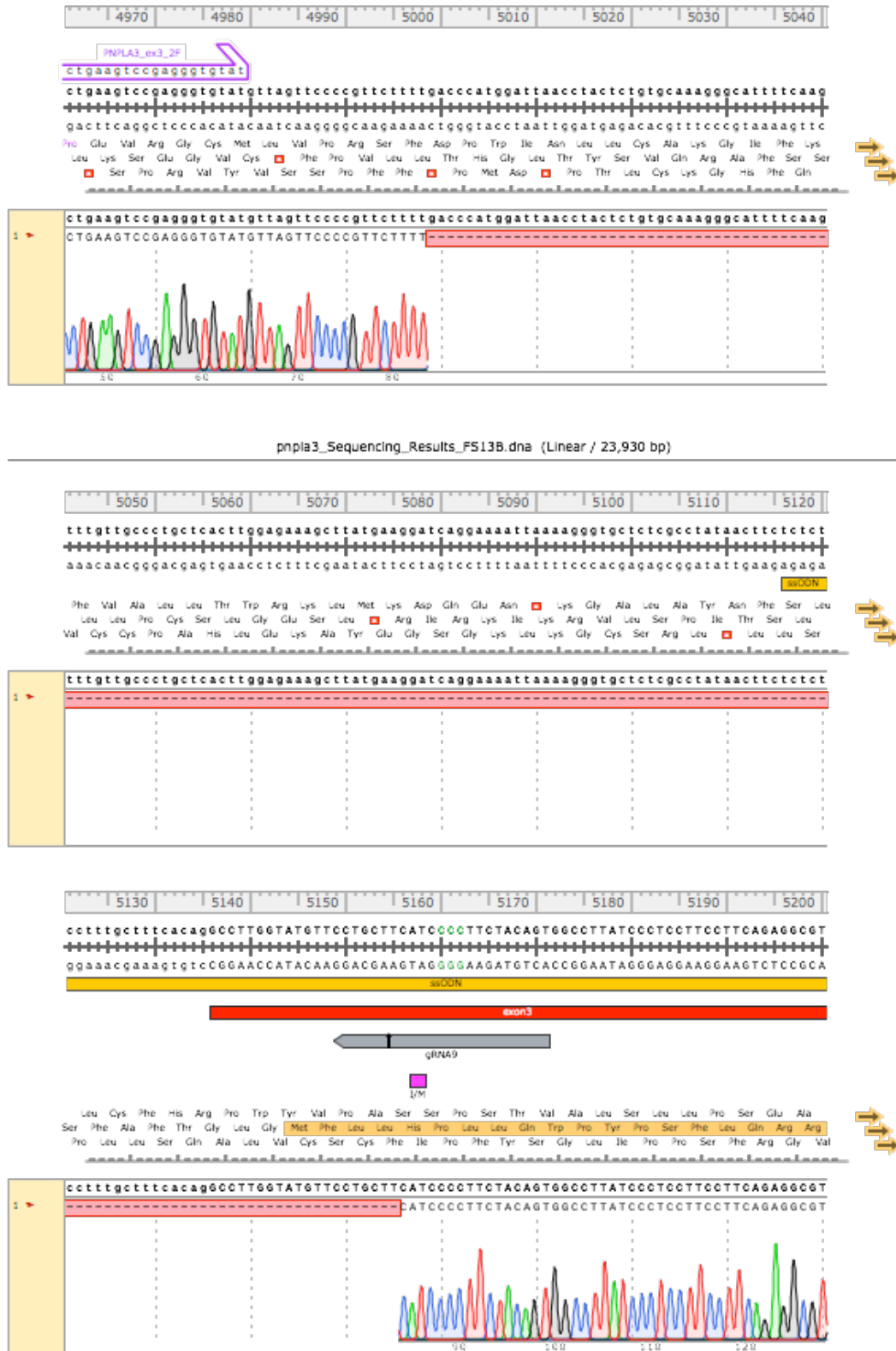
Figure 10.6 Sanger Sequence of FSPS13B Knock-Out Clone #2

pnpla3_Sequencing_Results_FS13B.dna (Linear / 23,930 bp)

Induced Pluripotent Stem Cell Derived Liver Model for the Study of PNPLA3-Associated Non-Alcoholic Fatty Liver Disease
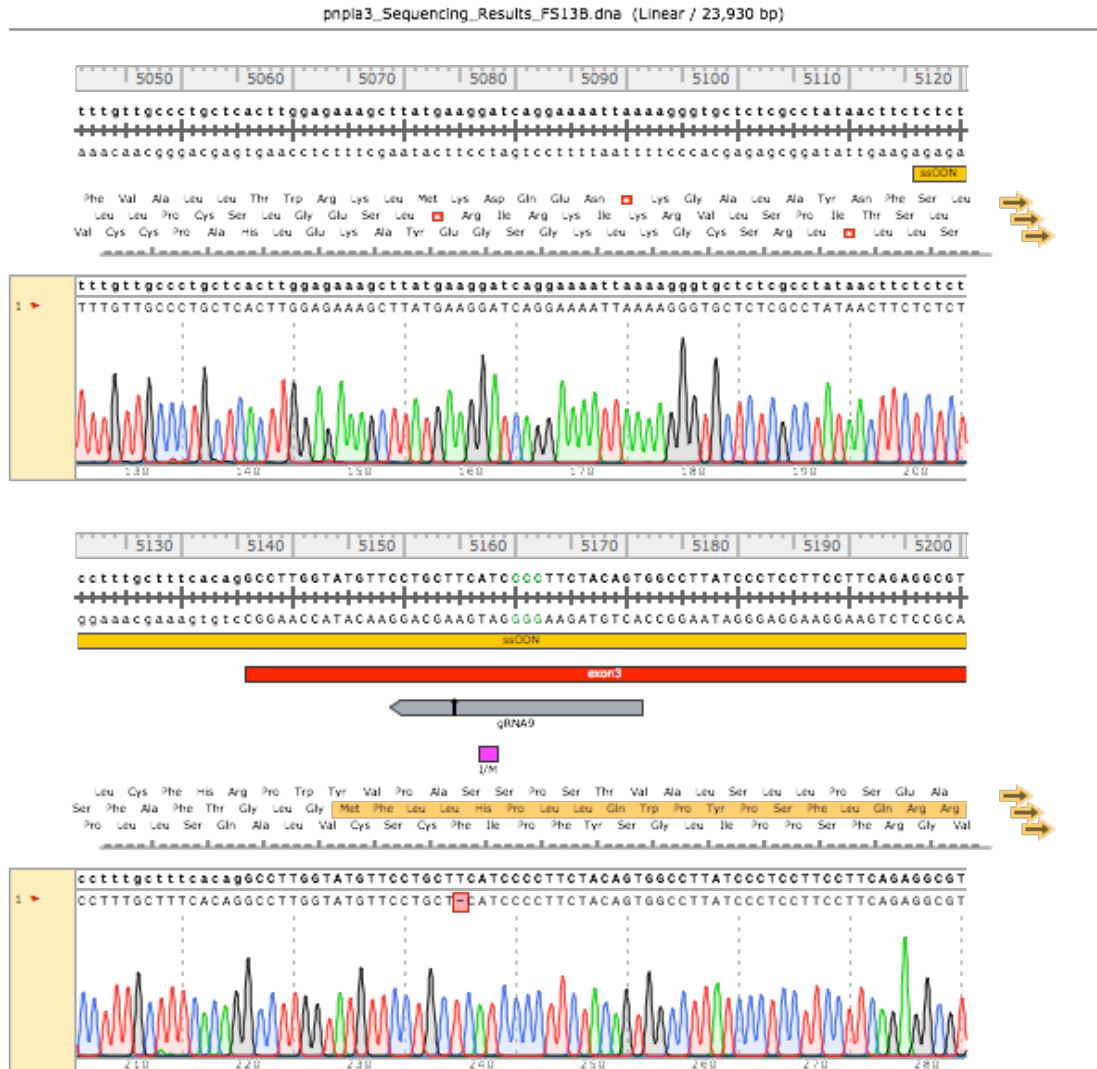
Figure 10.7 Sanger Sequence of FSPS13B Knock-Out Clone #3
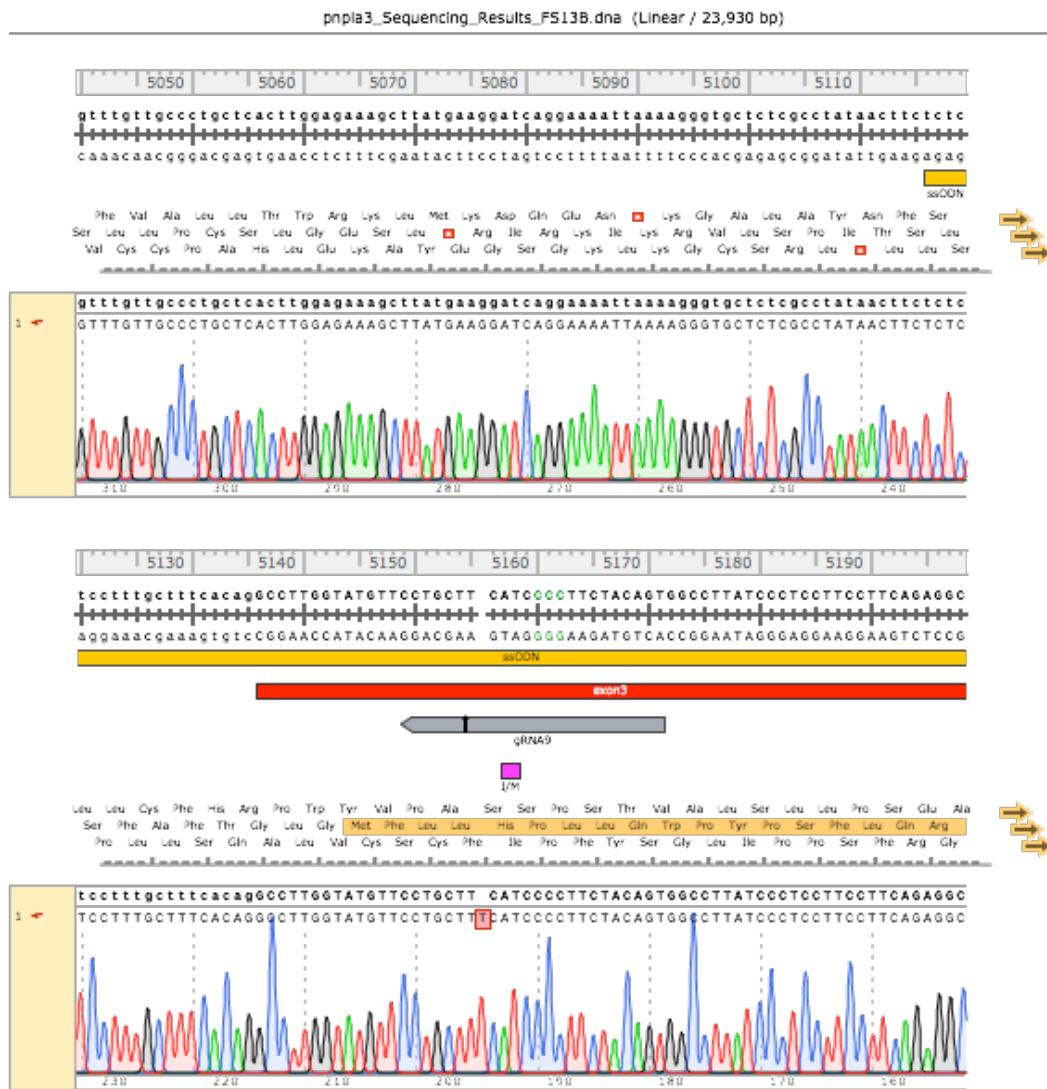
Figure 10.8 Sanger Sequence of A1ATDR/R Untargeted Control Clone #1

pnpla3_Sequencing_Results_BobC (1).dna (Linear / 23,930 bp)
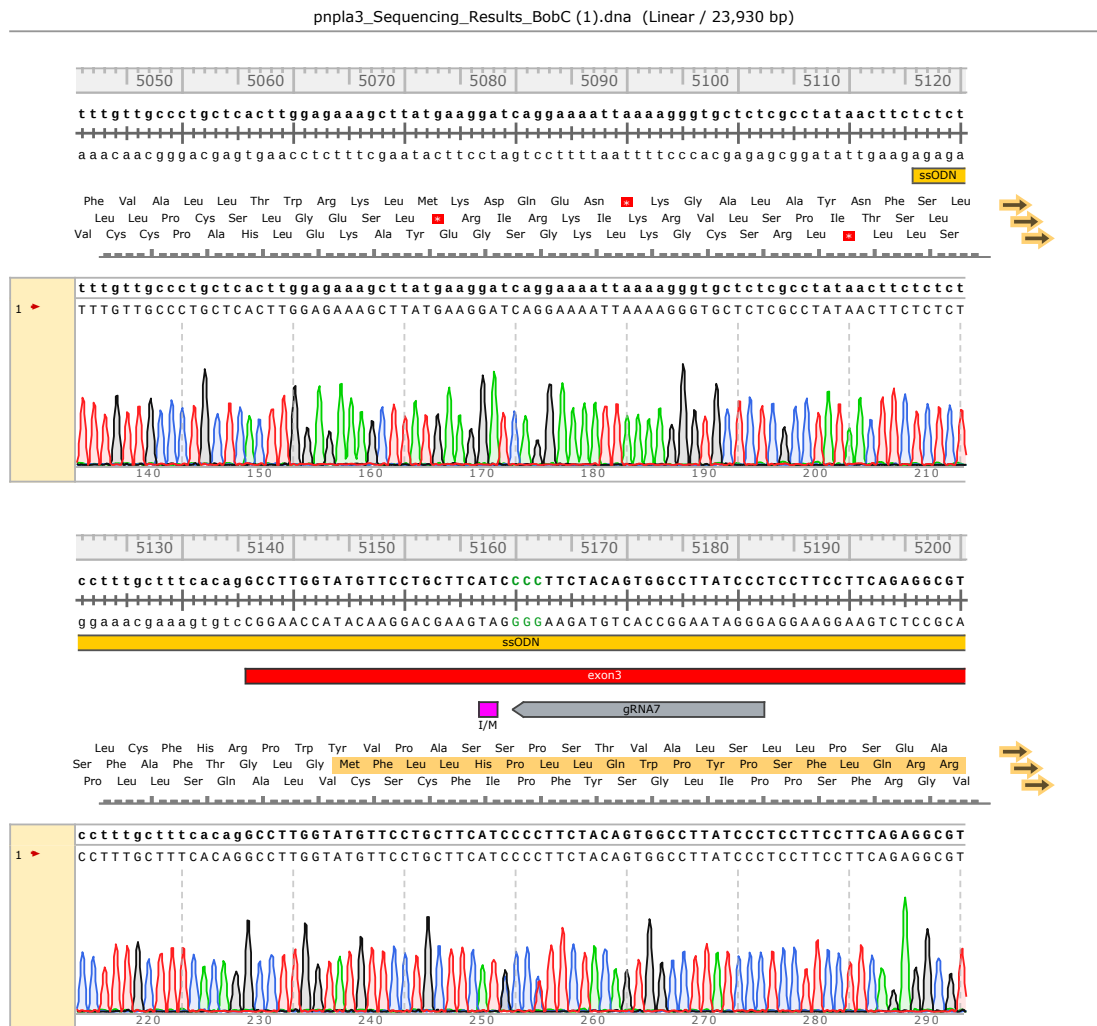
Figure 10.9 Sanger Sequence of A1ATDR/R Untargeted Control Clone #2
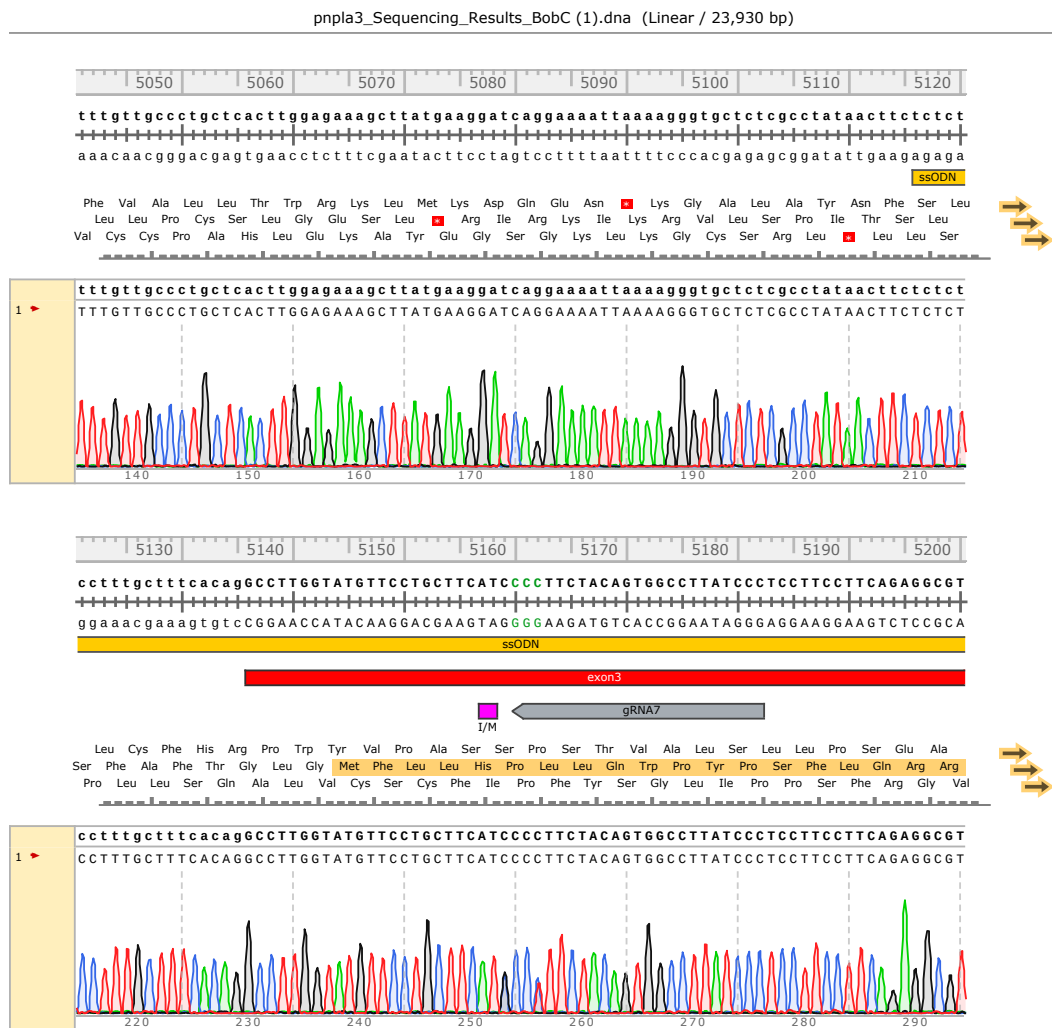
Figure 10.10 Sanger Sequence of A1ATDR/R I148M Clone #1
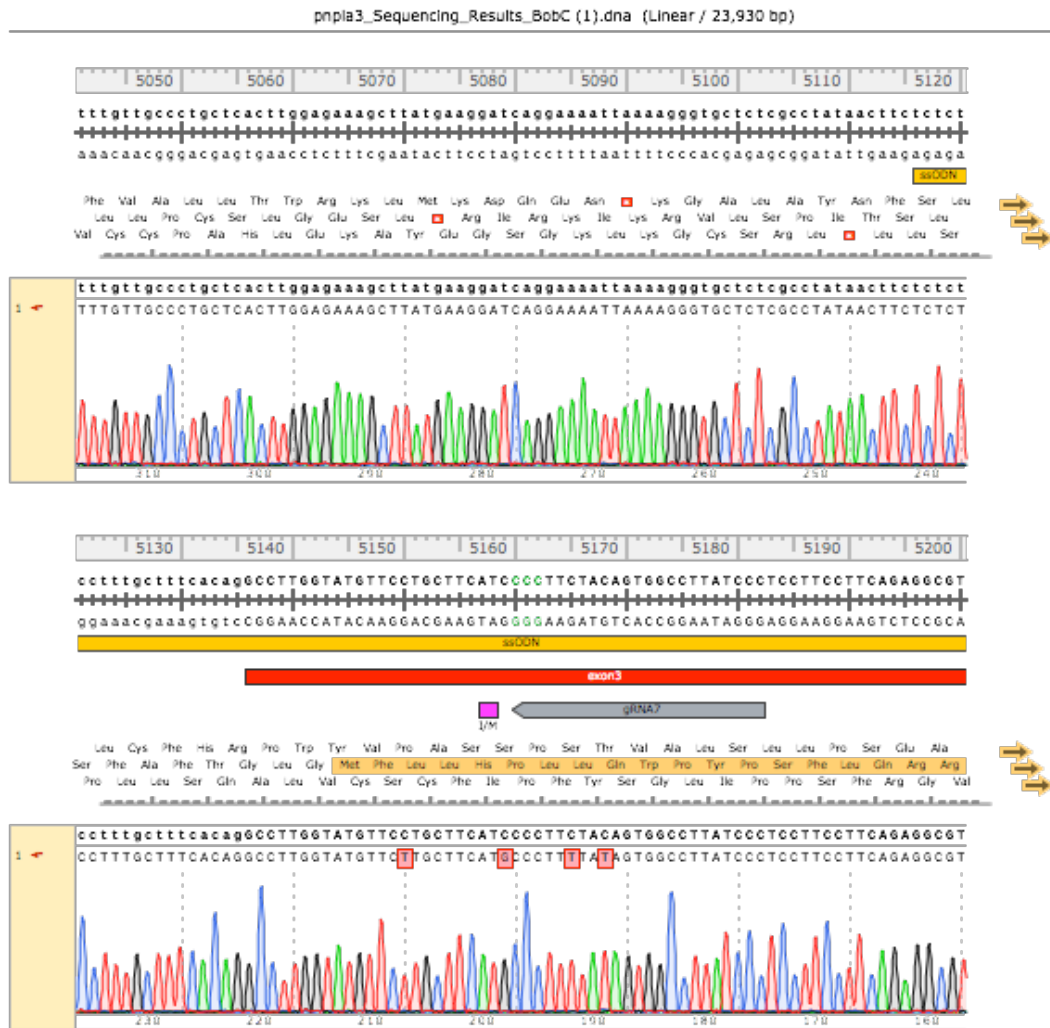
Figure 10.11 Sanger Sequence of A1ATDR/R Knock-Out Clone #1

Figure 10.12 Sanger Sequence of A1ATDR/R Knock-Out Clone #2
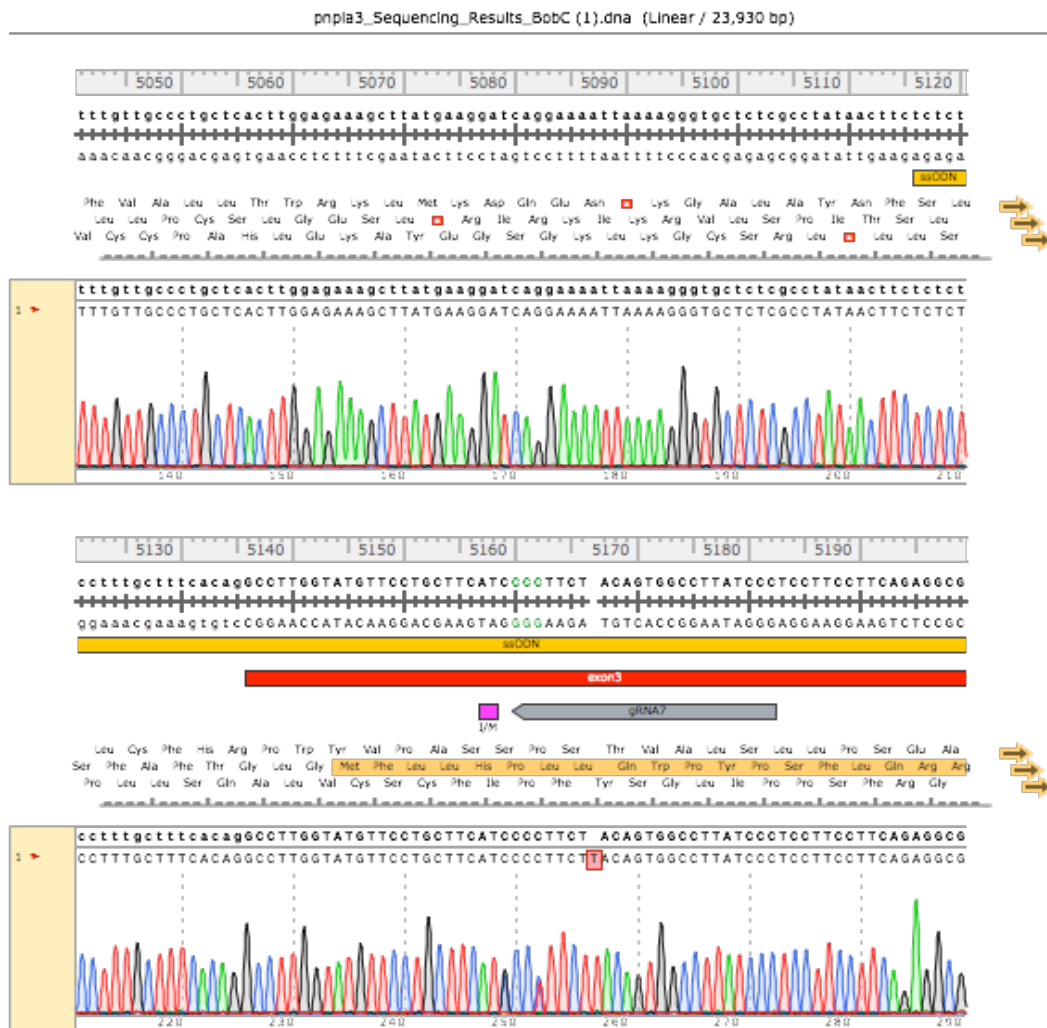
Table 10.1 List of differentially expressed genes between PNPLA3^{KO} and PNPLA3^{UC} Cells

Control			OA			PA		
Gene	LFC	FDR	Gene	LFC	FDR	Gene	LFC	FDR
TSPYL5	-7.981	2.8E-09	TSPYL5	-9.695	3.2E-04	TSPYL5	-9.754	5.2E-09
LRRC61	-7.960	2.1E-11	APCS	-7.667	1.5E-06	LRRC61	-8.391	5.0E-13
CRYAA	-7.106	6.8E-07	ACAN	-7.592	3.3E-02	APCS	-8.240	8.1E-11
CRP	-7.081	2.5E-11	LRRC61	-7.520	2.8E-07	PRSS3	-7.094	2.9E-07
AC026412.1	-6.938	4.4E-07		-6.993	3.0E-05	ZFP42	-6.975	1.1E-07
OAS3	-6.922	5.2E-08	CRP	-6.802	1.2E-07	CRP	-6.791	2.4E-08
CYP3A7	-6.716	1.6E-19	ZFP42	-6.733	1.1E-04	DES	-6.152	2.6E-07
APCS	-6.560	4.0E-05	OAS3	-6.481	2.4E-02	AC026412.1	-6.074	2.0E-05
PRSS3	-6.548	1.9E-06	IQCA1	-6.224	4.7E-04	AC116050.1	-5.873	1.8E-03
DES	-6.534	2.1E-07	DES	-6.117	4.7E-04	CYP3A7	-5.759	1.6E-08
CRYBA1	-6.425	4.8E-05	ADH1A	-5.964	1.3E-05	TXNRD2	-5.738	2.5E-06
IQCA1	-6.281	1.7E-05	LINC02701	-5.791	6.4E-03	ADH1A	-5.335	1.3E-08
CRABP1	-6.237	2.3E-05	CYP3A7	-5.773	9.9E-05	KCNJ5	-5.022	6.7E-10
ADH1A	-6.226	4.9E-11	PRSS3	-5.577	4.0E-05	BAAT	-4.712	4.2E-08
AC116050.1	-6.081	2.5E-05	PON3	-5.478	1.9E-02	CFHR2	-4.470	1.5E-18
CD248	-5.652	5.6E-06	CD248	-5.438	1.9E-04	CFHR1	-4.290	1.5E-53
AP000808.1	-5.488	4.8E-05	DCN	-5.422	1.7E-04	ANO9	-4.227	3.8E-04
SLCO1B1	-5.428	5.8E-08	CRABP1	-5.195	2.2E-03	ARHGAP40	-4.206	1.7E-02
ARHGAP40	-4.967	7.2E-05	BAAT	-5.114	4.7E-03	SLCO1B1	-4.052	2.9E-06
BAAT	-4.913	7.7E-09	TXNRD2	-4.938	2.4E-03	C9	-4.039	3.4E-20
ANO9	-4.752	3.3E-06	CFHR2	-4.837	5.9E-17	F9	-4.022	4.1E-37
SLC38A4	-4.730	4.1E-36	MFAP4	-4.740	1.3E-04	LOXL4	-4.016	8.3E-09
TXNRD2	-4.689	5.3E-04	POSTN	-4.737	3.2E-02	CYP2C8	-4.010	4.2E-17
THY1	-4.503	2.5E-04	PCDH18	-4.679	8.1E-03	ASPG	-3.959	2.0E-03
KCNJ5	-4.479	7.0E-07	ADH6	-4.574	7.2E-08	AP000808.1	-3.908	1.2E-03
DCN	-4.474	2.1E-08	BOC	-4.438	2.5E-04	PLAT	-3.902	1.3E-04
PDPN	-4.416	4.7E-05	THY1	-4.299	4.9E-02	ALDH4A1	-3.861	3.3E-09
ASPG	-4.404	5.3E-11	KCNJ5	-4.193	2.9E-08	ARG1	-3.777	7.7E-76
PLAT	-4.241	3.2E-06	DLK1	-4.170	9.0E-04	ADH6	-3.737	5.6E-11
CFHR2	-4.160	3.0E-20	PLAT	-4.161	2.1E-04	RARRES2	-3.710	3.0E-11
ATP2B2	-4.051	2.2E-09	LOXL4	-4.061	2.8E-03	G6PC	-3.581	8.6E-09
ADH6	-4.046	6.3E-13	CYP2C8	-4.009	1.9E-11	PKP3	-3.517	1.4E-05
G6PC	-3.974	5.0E-21	RARRES2	-4.002	1.1E-25	MX2	-3.511	3.9E-04

Samantha G. Tilson - January 2021

MX2	-3.947	1.6E-04	F9	-3.945	3.3E-06	SLC38A4	-3.435	1.7E-10
LOXL4	-3.923	5.4E-07	G6PC	-3.904	8.3E-04	ACVRL1	-3.346	8.2E-07
CFHR1	-3.878	1.5E-42	CFHR1	-3.853	5.6E-29	DCN	-3.287	1.2E-03
INHBE	-3.858	5.0E-38	ALDH4A1	-3.850	4.9E-06	ADAMTS2	-3.252	3.2E-04
PTX3	-3.831	5.6E-07	C9	-3.828	9.0E-12	ATP2B2	-3.184	2.9E-06
SVEP1	-3.786	1.5E-02	NPTX1	-3.810	4.1E-02	C11orf96	-3.058	1.6E-17
CYP2C8	-3.768	1.0E-16	ARHGAP40	-3.806	8.4E-03	AMN	-3.043	3.5E-07
POSTN	-3.758	8.1E-03	MX2	-3.683	1.4E-02	POSTN	-3.029	3.6E-02
C9	-3.732	1.6E-05	PKP3	-3.660	1.1E-02	CYP19A1	-3.028	7.0E-22
GALNT17	-3.659	3.5E-02	PTX3	-3.605	9.8E-05	ITIH3	-3.015	1.0E-07
ALDH4A1	-3.608	1.6E-09	ARG1	-3.566	5.5E-50	ALOX15B	-3.002	1.6E-03
RARRES2	-3.583	7.6E-35	CYP4A11	-3.532	4.0E-06	ACAN	-2.992	2.3E-02
COL11A1	-3.542	1.8E-03	COL3A1	-3.524	4.1E-03	PRG4	-2.952	1.0E-32
CSPG4	-3.500	2.7E-03	SLC38A4	-3.470	1.5E-22	HRG	-2.926	5.0E-13
SULF1	-3.483	2.7E-02	SOD3	-3.374	1.2E-04	SVEP1	-2.914	4.9E-03
ANTXR1	-3.462	1.1E-05	ITIH3	-3.353	6.2E-06	SLC22A9	-2.888	1.8E-06
ARG1	-3.458	2.0E-66	MMP9	-3.254	3.3E-03	SLC17A9	-2.882	1.4E-02
F9	-3.428	1.1E-23	F11	-3.101	3.6E-10	PTX3	-2.874	5.7E-06
COL3A1	-3.375	8.8E-05	PRG4	-3.081	9.4E-18	SCARA3	-2.846	3.1E-08
HRG	-3.312	6.5E-23	TRPV2	-3.015	4.7E-04	SOD3	-2.844	3.4E-07
COL1A2	-3.290	3.2E-02	SLC22A9	-2.964	4.8E-17	ARC	-2.828	6.8E-10
GPM6B	-3.263	6.5E-05	AMN	-2.935	1.1E-03	DLK1	-2.791	1.1E-02
C4BPA	-3.253	3.0E-22	SCARA3	-2.922	3.2E-07	INHBE	-2.756	1.0E-12
PRG4	-3.241	4.0E-19	ANTXR1	-2.878	5.1E-03	UCP2	-2.745	6.0E-09
ITIH3	-3.221	1.1E-10	SVEP1	-2.868	4.8E-03	IGDCC3	-2.688	6.4E-03
CYP19A1	-3.167	2.7E-29	CYP19A1	-2.805	3.3E-12	KLKB1	-2.684	4.8E-20
DLK1	-3.157	2.1E-03	SCUBE3	-2.764	2.6E-02	TRIML2	-2.663	1.2E-03
PKP3	-3.127	4.8E-03	ADAMTS2	-2.760	2.6E-02	NTS	-2.654	1.1E-06
F11	-3.122	1.5E-16	ACOT12	-2.742	1.2E-12	HHEX	-2.618	1.4E-17
ACVRL1	-3.076	1.5E-02	KLKB1	-2.733	2.7E-14	APBB1IP	-2.575	1.5E-07
AMN	-3.034	6.4E-06	TTPA	-2.706	1.5E-08	GLYATL1	-2.528	6.8E-09
NTS	-2.928	6.0E-07	NTS	-2.705	3.8E-03	CDO1	-2.526	1.7E-15
MMP2	-2.802	2.7E-03	C4BPA	-2.692	2.9E-10	PCDHGA10	-2.523	4.5E-05
SLC22A9	-2.793	3.1E-15	IAPP	-2.658	2.4E-05	TNFAIP2	-2.485	1.3E-05
SCARA3	-2.779	1.5E-11	LUM	-2.656	2.2E-07	F13B	-2.484	5.9E-17
SLC17A9	-2.752	6.4E-03	GLYATL1	-2.565	1.8E-06	F11	-2.479	3.4E-10
SNED1	-2.714	1.5E-04	LYZ	-2.526	2.1E-02	C4BPA	-2.458	7.6E-14
SLC17A2	-2.689	9.8E-20	LIPC	-2.505	2.8E-33	NR4A3	-2.420	1.2E-06
KLKB1	-2.676	9.9E-25	CDO1	-2.495	2.8E-13	MMP10	-2.409	1.3E-21
CDO1	-2.676	3.4E-19	EMILIN1	-2.482	1.4E-05	COL3A1	-2.389	3.1E-02

EMILIN1	-2.664	9.5E-05	HHEX	-2.469	1.2E-14	MMP2	-2.381	7.0E-03
TTPA	-2.664	2.3E-14	GC	-2.444	3.4E-06	VNN3	-2.352	4.7E-11
GLYATL1	-2.624	3.4E-09	F13B	-2.399	1.4E-10	CFHR3	-2.342	4.6E-14
ALOX15B	-2.615	1.7E-03	EPO	-2.386	3.4E-13	MT1H	-2.340	9.8E-04
COL9A3	-2.601	1.7E-02	CPB2	-2.379	2.1E-04	IGFBP1	-2.245	7.8E-09
EPO	-2.598	9.6E-30	HRG	-2.350	1.0E-05	LIPC	-2.241	2.4E-27
CACNA1E	-2.542	5.6E-11	ACSM2A	-2.322	7.2E-07	GC	-2.217	7.8E-08
MRC2	-2.540	9.1E-07	KNG1	-2.253	5.6E-29	ANTXR1	-2.213	1.8E-03
DPPA4	-2.539	4.5E-06	ITIH1	-2.188	5.8E-10	UGT2B4	-2.178	1.1E-10
LIPC	-2.520	2.4E-48	PAH	-2.188	8.6E-20	MRC2	-2.177	3.3E-09
OXTR	-2.498	3.5E-02	UGT2B4	-2.145	1.8E-08	PAH	-2.176	3.8E-20
LRRC32	-2.468	2.0E-02	INHBE	-2.132	1.3E-03	ITIH1	-2.130	1.5E-35
PALMD	-2.461	2.5E-06	ITIH2	-2.117	7.8E-30	C8B	-2.128	1.3E-38
SOD3	-2.420	1.3E-06	C1S	-2.099	1.2E-07	NR4A1	-2.116	6.0E-20
CPXM1	-2.398	4.6E-06	LGALS1	-2.089	4.6E-03	TTPA	-2.110	1.7E-08
CFHR3	-2.360	3.8E-13	ALDH8A1	-2.067	5.5E-09	SERPINA6	-2.098	6.8E-09
PCDHGA10	-2.353	1.8E-04	CACNA1E	-2.055	1.5E-04	LUM	-2.092	1.3E-10
SLC25A18	-2.349	2.3E-16	C8B	-2.047	1.2E-28	CGREF1	-2.092	1.6E-08
HHEX	-2.336	1.9E-15	SERPINI1	-2.021	8.6E-17	COL11A1	-2.066	3.2E-02
ITIH1	-2.297	3.0E-31	C8A	-1.999	4.1E-12	KNG1	-2.020	7.5E-13
MXRA8	-2.286	3.8E-03	ACSM2B	-1.980	1.9E-11	C8A	-2.007	2.5E-26
TNFAIP2	-2.272	3.2E-02	AADAC	-1.973	3.5E-05	CPB2	-1.993	7.5E-17
C1S	-2.270	9.2E-16	APOA5	-1.959	4.6E-14	LGALS1	-1.960	9.7E-05
GC	-2.240	9.1E-09	SERPINA6	-1.940	6.1E-20	GSDMD	-1.943	2.0E-02
CPB2	-2.205	2.5E-22	AKR1B10	-1.927	3.8E-03	SLC17A2	-1.939	8.4E-18
PAH	-2.190	1.0E-16	APOH	-1.907	3.7E-27	APOH	-1.933	3.1E-32
SPON2	-2.186	4.8E-02	TFR2	-1.895	9.3E-05	C1S	-1.892	1.8E-08
KNG1	-2.182	5.8E-33	C1R	-1.887	7.9E-04	ITIH2	-1.892	1.1E-26
GSDMD	-2.170	3.0E-02	SLC17A2	-1.869	2.0E-08	TRIM4	-1.856	3.2E-08
ACOT12	-2.123	1.1E-15	AMBP	-1.868	3.5E-08	S100A9	-1.854	5.5E-08
C8B	-2.111	1.2E-62	NIBAN1	-1.863	9.6E-09	AMBP	-1.845	5.9E-33
AC080128.1	-2.077	4.1E-03	ORM2	-1.829	5.4E-10	SPON2	-1.839	2.6E-02
SERPINA6	-2.068	2.5E-41	SLC25A18	-1.819	3.6E-05	EPO	-1.816	4.7E-10
ITIH2	-2.067	1.1E-21	SDR16C5	-1.796	1.7E-05	HSPA6	-1.811	2.7E-04
CXCL17	-2.045	4.8E-16	GMPR	-1.789	1.0E-05	AL355075.4	-1.792	1.5E-03
APOA5	-2.026	5.7E-15	TRIM4	-1.763	1.1E-04	C1R	-1.790	1.3E-05
UGT2B4	-2.025	2.0E-15	GFRA1	-1.757	2.4E-02	ACSM2B	-1.775	6.8E-10
LYZ	-1.967	1.7E-02	S100A9	-1.717	2.5E-06	BCL2A1	-1.759	1.9E-07
C8A	-1.961	7.3E-27	AKR1D1	-1.713	9.4E-18	FOSB	-1.759	8.8E-11
F13B	-1.960	1.0E-08	FBP1	-1.711	1.3E-08	CXCL12	-1.758	4.1E-02

CYP39A1	-1.949	5.6E-09	NT5E	-1.708	4.9E-06	PCK1	-1.753	1.5E-03
AKR1D1	-1.863	1.8E-14	TFF2	-1.702	2.2E-05	FBP1	-1.715	1.3E-13
MT1H	-1.860	4.7E-02	MT1L	-1.647	9.1E-12	ZNF331	-1.714	4.2E-16
GFRA1	-1.849	6.5E-12	GALNT5	-1.640	3.6E-06	LYZ	-1.710	9.1E-03
ACSM2A	-1.826	1.8E-05	OTC	-1.602	3.9E-05	DNAJA4	-1.694	3.9E-02
ALDH8A1	-1.800	1.1E-10	F12	-1.540	2.0E-02	TFF2	-1.687	2.3E-08
C1R	-1.797	1.0E-05	CFH	-1.528	1.1E-05	F12	-1.672	8.9E-14
CXCL12	-1.766	1.7E-02	SLC22A7	-1.500	1.4E-10	PCSK9	-1.660	8.6E-32
ORM2	-1.762	5.8E-16	ACSL4	-1.498	2.1E-37	DNAJB4	-1.654	2.6E-16
SERPINI1	-1.757	5.2E-18	SERPINC1	-1.465	1.7E-18	ACOT12	-1.653	3.6E-08
APOH	-1.748	1.8E-29	MMP10	-1.463	1.8E-04	AADAC	-1.650	8.0E-07
LUM	-1.725	2.6E-06	TM4SF4	-1.453	8.1E-03	CTH	-1.623	1.3E-21
MCAM	-1.720	4.8E-02	ATF5	-1.429	2.3E-02	ZNF165	-1.616	1.3E-16
S100A9	-1.716	2.0E-08	CPN1	-1.426	6.1E-03	MYH4	-1.590	3.0E-21
SLC38A3	-1.703	4.7E-22	SERPINA11	-1.422	3.5E-11	DDIT3	-1.583	1.3E-21
AMBP	-1.692	4.2E-39	SERPINA10	-1.418	3.3E-12	ORM2	-1.560	8.5E-10
ACSM2B	-1.669	2.2E-12	ORM1	-1.417	1.6E-13	OTC	-1.550	1.8E-05
GALNT5	-1.624	4.7E-14	MAT1A	-1.411	8.1E-26	AGR2	-1.542	6.0E-06
SLC22A7	-1.619	7.4E-18	CROT	-1.411	6.9E-05	S100A6	-1.534	5.6E-03
FAM129A	-1.616	5.4E-13	SERPING1	-1.410	1.8E-03	GALNT5	-1.517	4.5E-07
C3	-1.608	2.7E-14	C3	-1.392	9.5E-04	VASN	-1.516	1.6E-08
AADAC	-1.602	2.1E-11	APOA2	-1.385	2.5E-15	AKR1B10	-1.502	6.0E-03
TRIM4	-1.598	9.8E-06	IGFBP1	-1.385	1.3E-11	PTGS2	-1.500	1.2E-05
FBN1	-1.568	2.7E-02	NNMT	-1.378	2.6E-02	ALDH8A1	-1.497	3.0E-08
FBP1	-1.550	1.4E-13	AGR2	-1.351	9.2E-07	AKR1D1	-1.465	2.2E-13
NPW	-1.550	2.9E-05		-1.344	5.7E-11	CHAC1	-1.451	3.1E-08
CSF3R	-1.544	2.8E-09	AFM	-1.316	2.3E-03	MT1L	-1.442	2.2E-18
CFH	-1.519	2.8E-14	LRG1	-1.312	7.3E-09	GFRA1	-1.442	4.1E-03
F12	-1.466	6.8E-11	LBP	-1.310	2.9E-04	AKR1C2	-1.438	2.3E-08
COL5A2	-1.455	9.8E-04	MT1A	-1.307	2.3E-07	HBEGF	-1.426	2.4E-04
SERPINA11	-1.443	9.1E-18	METTL7A	-1.306	5.5E-13	C3	-1.419	3.2E-07
SERPINC1	-1.430	1.4E-19	FABP1	-1.297	1.1E-03	ANXA1	-1.418	1.1E-03
ACSL4	-1.423	2.3E-31	SGK2	-1.297	6.0E-06	CXCL17	-1.417	1.7E-02
MAT1A	-1.419	1.4E-25	NDUFAF1	-1.297	3.3E-05	SLC25A18	-1.405	2.1E-05
SERPING1	-1.417	4.7E-12	MBL2	-1.283	8.3E-10	ATF3	-1.401	8.2E-04
ITPR2	-1.402	2.1E-15	ITPR2	-1.277	1.8E-08	SERPINA11	-1.399	3.4E-11
TFF2	-1.387	2.0E-16	HMGCS2	-1.272	7.9E-11	SERPINC1	-1.391	2.9E-18
PNPLA3	-1.378	4.4E-07	CXCL17	-1.271	3.2E-06	SERPINI1	-1.378	4.6E-19
LRG1	-1.366	3.0E-21	ACSL1	-1.270	8.9E-16	SLC22A7	-1.377	1.1E-15
NT5E	-1.359	1.0E-03	F5	-1.263	3.5E-08	APOA5	-1.372	3.4E-10

MYH4	-1.339	1.8E-17	HAO2	-1.263	5.8E-07	PNPO	-1.370	8.4E-05
ATF5	-1.328	3.0E-03	DHODH	-1.246	4.9E-06	SERPING1	-1.363	1.2E-05
F5	-1.322	4.1E-08	OLFML3	-1.235	3.9E-06	PMAIP1	-1.354	1.6E-11
CROT	-1.317	1.7E-06	CPM	-1.230	1.5E-02	CPN1	-1.353	3.0E-03
IGFBP1	-1.316	5.8E-05	ANGPTL3	-1.229	3.9E-12	ORM1	-1.349	6.7E-06
OLFML3	-1.288	8.1E-06	CSF3R	-1.222	7.7E-06	MT1G	-1.347	1.8E-02
LBP	-1.281	1.1E-16	COL5A2	-1.221	2.9E-02	KATNBL1	-1.345	5.6E-15
MTSS1L	-1.275	7.4E-18	CBR4	-1.221	1.0E-07	HSPH1	-1.336	4.1E-13
PDGFRB	-1.268	2.6E-02	FGL1	-1.217	1.3E-20	CKS2	-1.336	2.2E-12
MBNL3	-1.265	5.4E-14	AKR1C1	-1.209	1.4E-08	ELOVL5	-1.321	1.2E-20
PCK1	-1.264	1.1E-03	SLC38A3	-1.199	1.8E-10	TOR4A	-1.314	1.2E-03
PNPO	-1.257	1.0E-03	ACAT1	-1.199	9.1E-16	AKR1C1	-1.305	8.1E-12
PEG10	-1.241	2.2E-13	PCK1	-1.191	8.7E-04	NT5E	-1.289	2.8E-03
PCSK9	-1.227	6.3E-19	FGG	-1.187	2.2E-17	ACSL4	-1.288	1.4E-29
SERPINA10	-1.224	4.0E-11	NDRG2	-1.168	4.9E-07	COL1A1	-1.271	1.6E-02
CCDC80	-1.210	4.3E-02	MYLK	-1.165	4.5E-02	RMRP	-1.258	1.7E-02
NDRG2	-1.190	7.0E-13	MBNL3	-1.157	5.4E-13	CFH	-1.252	1.4E-15
MT1L	-1.185	2.1E-07	ENPP2	-1.152	2.2E-15	PNPLA3	-1.251	2.6E-13
ACAT1	-1.180	4.0E-29	AOX1	-1.121	2.4E-09	CSF3R	-1.250	7.2E-08
MAP1A	-1.173	1.2E-09	MT1B	-1.109	2.3E-04	ITPR2	-1.244	4.6E-13
METTL7A	-1.167	4.3E-16	STEAP1	-1.107	7.7E-07	LRG1	-1.204	1.3E-12
ACSS2	-1.165	1.3E-21	FGA	-1.099	5.1E-17	LIPG	-1.200	1.6E-25
FGA	-1.163	1.4E-39	PEG10	-1.099	2.4E-09	CYTOR	-1.199	2.9E-10
HAO2	-1.160	5.9E-10	ANG	-1.097	3.8E-13	TM4SF4	-1.192	1.9E-04
IGF2	-1.134	3.2E-13	MTSS2	-1.096	5.9E-12	SERPINA10	-1.190	3.8E-17
MATN2	-1.134	3.9E-02	TTC39C	-1.094	7.0E-12	ACAT1	-1.188	2.1E-24
AC115619.1	-1.128	7.3E-10	HADHB	-1.092	9.1E-12	SELENOK	-1.185	3.3E-16
ITGA1	-1.114	9.2E-06	VSIG1	-1.091	4.4E-02	OLFML3	-1.183	2.1E-13
ANG	-1.108	7.0E-14	PCCB	-1.083	2.8E-08	ZFAND2A	-1.165	1.8E-16
ALB	-1.105	1.8E-02	MPC1	-1.073	4.6E-07	ATF5	-1.164	1.7E-02
FGG	-1.104	5.8E-31	PNPLA3	-1.072	7.2E-07	MTSS1L	-1.152	1.1E-14
PC	-1.080	8.8E-16	TIMP1	-1.066	2.5E-02	APOA2	-1.149	3.1E-17
CLMN	-1.073	2.8E-02	HPX	-1.059	1.1E-11	MAT1A	-1.149	2.4E-27
MBL2	-1.072	1.1E-10	S100P	-1.058	1.3E-04	SLC38A3	-1.144	4.8E-11
ORM1	-1.071	1.1E-07	TM4SF5	-1.032	1.2E-04	NNMT	-1.134	3.7E-03
CYP8B1	-1.061	5.9E-07	FGB	-1.021	2.7E-14	AFM	-1.117	2.7E-04
AOX1	-1.047	4.2E-11	NADK2	-1.007	4.0E-10	SNAI2	-1.106	2.9E-06
ENPP2	-1.045	2.1E-16	ITGA1	-1.007	6.9E-06	NOCT	-1.105	7.8E-07
VSIG1	-1.044	1.6E-07	PC	-1.003	4.2E-12	MT1A	-1.104	1.2E-07
LSS	-1.043	1.7E-18	GPX3	-0.995	1.9E-11	GDF15	-1.101	6.1E-09

TNFRSF10C	-1.031	1.5E-03	HSD17B4	-0.981	2.7E-14	LBP	-1.096	3.5E-10
AMOT	-1.024	2.7E-02	CXXC5	-0.980	3.1E-04	BAG3	-1.092	2.9E-12
PIK3AP1	-1.007	5.0E-09	APOC1	-0.971	2.2E-10	BTG2	-1.090	5.8E-15
MT1A	-0.999	1.4E-06	CYB5A	-0.966	4.1E-07	HAO2	-1.062	3.9E-12
NADK2	-0.993	3.2E-25	CA2	-0.963	2.2E-05	AGTR1	-1.044	4.5E-19
ACSL1	-0.992	1.7E-13	ATP11C	-0.937	4.9E-04	ANGPTL3	-1.038	8.0E-12
AFM	-0.990	1.8E-05	HP	-0.925	6.5E-11	METTL7A	-1.031	4.2E-06
C6	-0.988	1.0E-09	DPYD	-0.909	1.2E-06	CCL20	-1.030	3.1E-07
FGL1	-0.977	9.5E-18	GPC3	-0.896	5.2E-15	COL5A2	-1.025	1.1E-06
HPX	-0.974	1.7E-16	ACSS2	-0.881	2.0E-08	AC115619.1	-1.014	5.5E-09
DPYD	-0.965	3.5E-06	PHYH	-0.875	2.0E-05	DNAJB9	-1.009	1.3E-18
HMGCS2	-0.954	5.4E-09	PLG	-0.859	1.9E-11	HERPUD1	-1.007	6.8E-18
ATP11C	-0.932	9.2E-06	LIPG	-0.851	1.6E-07	MT1B	-1.006	4.0E-04
PLG	-0.931	1.4E-16	LSS	-0.786	8.9E-10	TIMP1	-1.000	1.6E-10
COL2A1	-0.930	3.6E-02	CLDN2	-0.769	6.0E-08	MVD	-0.993	3.3E-13
ANGPTL3	-0.929	4.5E-09	IGF2	-0.764	1.9E-06	ANG	-0.992	7.6E-14
FGB	-0.927	4.9E-27	MVD	-0.733	1.3E-04	ACSS2	-0.989	2.3E-09
DPYS	-0.926	3.8E-06	GMDS	-0.732	1.5E-04	F5	-0.988	2.7E-07
TIMP1	-0.889	8.8E-07	PRXL2A	-0.710	2.9E-06	ENPP2	-0.984	2.5E-17
AKR1C1	-0.873	8.4E-05	APOC3	-0.685	2.8E-04	HSPB8	-0.977	2.7E-06
GPX3	-0.871	3.0E-12	IDI1	-0.677	9.5E-04	APOC1	-0.973	1.3E-13
P2RY6	-0.870	8.0E-04	TFF1	-0.551	5.0E-04	WEE1	-0.970	9.6E-11
APOA2	-0.868	2.8E-05	SCD	-0.520	6.8E-03	ACSL1	-0.956	3.5E-10
IGSF1	-0.867	2.9E-11	APOB	-0.489	5.5E-03	FGL1	-0.952	1.5E-18
SCD	-0.865	1.2E-10	APOA4	-0.443	1.7E-02	GPX3	-0.945	2.8E-15
HSD17B4	-0.861	2.9E-24	GPRC5B	0.414	5.6E-03	VSIG1	-0.945	2.8E-02
DHCR7	-0.855	1.5E-08	HLA-B	0.476	3.5E-02	AC117402.1	-0.942	2.5E-02
TM4SF4	-0.852	1.2E-03	LAMA5	0.519	6.8E-03	IDI1	-0.940	2.9E-11
GPC3	-0.851	2.9E-18	TSKU	0.559	7.6E-04	MTHFD2	-0.937	2.0E-06
SLCO4C1	-0.831	2.3E-07	BASP1	0.567	1.9E-03	RTN4RL2	-0.934	1.2E-05
SLC10A1	-0.827	1.1E-02	GABARAPL1	0.568	2.1E-04	GPC3	-0.932	1.2E-25
APOB	-0.822	3.0E-07	PROM1	0.579	8.8E-05	FTCD	-0.932	4.5E-06
TMEM176B	-0.821	4.8E-03	AQP10	0.579	1.2E-03	DPYS	-0.930	6.8E-03
ENPP1	-0.816	3.0E-03	INHBB	0.584	3.9E-02	APOL1	-0.925	3.7E-02
ACLY	-0.809	2.5E-07	TNS1	0.588	3.3E-03	CEP85	-0.925	9.8E-11
CP	-0.808	1.4E-03	MAT2A	0.597	2.6E-02	LSS	-0.925	1.2E-19
APOL1	-0.807	2.3E-02	KLF10	0.599	9.3E-04	DUSP5	-0.923	1.8E-12
HP	-0.783	1.8E-13	SGK1	0.604	7.6E-04	RRAD	-0.919	1.6E-02
MT-CYB	-0.770	6.5E-08	MFSD10	0.609	9.4E-03	HSPA5	-0.911	2.0E-14
SIPA1L2	-0.743	1.1E-03	RAB25	0.624	2.2E-02	MAFF	-0.911	5.1E-11

CD36	-0.738	1.3E-03	CYP2S1	0.626	3.3E-02	PC	-0.899	1.0E-19
C5	-0.728	1.7E-09	NETO2	0.627	2.5E-02	PEG10	-0.892	3.5E-10
CPS1	-0.727	9.9E-04	FRAS1	0.635	3.3E-02	DNAJB1	-0.891	5.0E-12
MPC1	-0.682	8.3E-03	GRK3	0.636	1.3E-03	IGF2	-0.880	2.6E-16
RBP4	-0.637	2.3E-04	CEBPD	0.637	6.8E-03	FGG	-0.875	2.8E-13
FASN	-0.636	5.3E-04	EGLN3	0.641	3.9E-04	HPX	-0.875	2.8E-13
MT-ATP6	-0.594	1.1E-03	TGFB1	0.644	4.0E-06	FGA	-0.858	5.1E-15
IDI1	-0.592	1.1E-03	ZFP36	0.649	2.9E-04	CPM	-0.852	2.6E-02
FABP1	-0.576	3.4E-02	STK39	0.655	1.7E-04	ITGA1	-0.849	1.1E-05
HMGCR	-0.574	9.3E-04	RESF1	0.655	1.1E-02	MPC1	-0.817	8.9E-06
FADS2	-0.523	1.8E-03	GALNT12	0.663	4.6E-02	HMGCS2	-0.796	2.3E-08
AHSG	-0.523	7.1E-03	SEMA4B	0.663	1.9E-03	DHCR7	-0.790	4.3E-10
FADS1	-0.519	1.6E-03	PLCD3	0.667	5.6E-06	CXCL5	-0.787	2.2E-02
APOC3	-0.460	4.9E-02	NBEAL2	0.668	4.5E-03	PLG	-0.783	1.0E-11
SERPINA1	-0.427	1.9E-04	NEDD9	0.668	3.2E-04	FABP1	-0.781	1.6E-02
FN1	-0.386	3.3E-02	SEMA6A	0.671	3.8E-02	LDLR	-0.774	4.5E-07
MAPK4	0.559	4.8E-02	FAM83F	0.675	1.2E-03	FDFT1	-0.772	1.2E-10
RAB31	0.587	7.7E-03	FYN	0.678	3.0E-02	HMGCS1	-0.772	1.0E-06
INHBB	0.617	1.4E-02	NR5A2	0.685	6.0E-03	FGB	-0.770	3.8E-17
MAT2A	0.638	1.8E-02	TET3	0.686	2.6E-02	FDPS	-0.756	2.1E-09
RRM2	0.665	3.2E-02	ST6GALNAC2	0.700	2.6E-02	MAP1B	-0.744	3.5E-05
EGLN3	0.673	3.2E-05	NFKBIA	0.707	4.8E-04	HMGCR	-0.744	4.9E-06
NR5A2	0.680	9.9E-03	TOP3A	0.710	3.4E-03	UCHL1	-0.738	3.0E-05
CTBP2	0.690	3.3E-02	TMEM132A	0.717	9.9E-03	INSIG1	-0.735	6.7E-05
TRIM32	0.700	3.2E-02	SLC7A6	0.718	3.2E-03	C6	-0.722	2.6E-05
HSPG2	0.709	1.2E-03	PRDM1	0.722	1.5E-02	APOC3	-0.720	1.8E-07
F3	0.716	1.3E-04	MMP1	0.722	6.9E-04	SIPA1L2	-0.715	4.1E-04
DAB1	0.719	3.7E-02	TLCD3A	0.724	2.2E-03	CSRP1	-0.712	2.7E-03
JAG2	0.733	4.1E-02	MAX	0.724	1.0E-03	CLU	-0.708	9.5E-10
LAMB3	0.747	1.0E-03	GCNT3	0.725	5.6E-03	FASN	-0.708	1.9E-02
MALRD1	0.752	3.3E-02	PTK2B	0.732	1.2E-03	DUSP1	-0.703	1.2E-08
FAM151A	0.759	9.4E-04	SASH1	0.734	4.9E-02	STARD4	-0.701	4.6E-04
ACSL5	0.764	7.3E-04	VWCE	0.739	9.3E-06	SCD	-0.699	7.7E-10
SIGIRR	0.777	9.3E-03	FLRT3	0.746	3.7E-09	HP	-0.685	3.1E-06
NUAK2	0.787	3.1E-02	SOX9	0.748	1.3E-04	CP	-0.631	5.1E-04
IRF1	0.796	1.4E-04	PLAGL2	0.750	4.8E-04	RGS16	-0.566	2.9E-02
AQP10	0.809	3.7E-06	SEC14L2	0.752	3.9E-04	FADS1	-0.556	9.0E-04
FER1L4	0.813	2.0E-04	LAMC1	0.753	1.7E-07	SLC51A	-0.554	9.0E-04
PRR15	0.814	1.8E-02	NFKBIZ	0.754	9.5E-05	FADS2	-0.529	1.6E-05
ENG	0.827	7.7E-03	ALDH1A3	0.754	7.6E-03	MT-RNR1	-0.518	1.2E-04

SLC6A8	0.827	9.4E-06	IRF1	0.755	4.2E-03	APOA4	-0.406	2.2E-03
COL4A2	0.828	4.5E-03	SLC26A2	0.756	2.1E-04	DGAT2	0.460	2.7E-02
FMO5	0.832	1.8E-03	DCHS1	0.762	2.8E-02	GPRC5B	0.466	6.7E-05
SGK1	0.833	4.9E-15	SLC7A5	0.763	1.9E-05	TRIM9	0.554	2.3E-02
EMP1	0.841	1.9E-02	MYO1A	0.764	3.7E-02	PIEZO2	0.611	5.4E-03
UBALD2	0.845	2.1E-05	GRB10	0.764	2.5E-06	TOP3A	0.621	4.3E-05
RASD1	0.851	3.1E-02	SLC36A1	0.769	1.6E-05	TMEM132A	0.657	1.3E-03
LGALS2	0.857	3.6E-02	SEMA6D	0.771	1.5E-04	FAM174B	0.680	9.8E-04
LRRC1	0.862	7.3E-05	SLC7A1	0.778	7.0E-04	B3GNT3	0.700	1.6E-02
GBP1	0.865	1.5E-03	DPP7	0.779	8.9E-08	MTTP	0.710	3.3E-05
DUOXA2	0.871	4.9E-07	EGFR	0.784	1.7E-05	TAGLN	0.710	6.7E-06
CDK1	0.878	1.9E-03	MXD1	0.785	3.6E-05	STK39	0.719	5.5E-10
IGSF9	0.880	3.1E-03	MICAL3	0.792	3.7E-04	DUOXA2	0.735	1.2E-09
RGS2	0.883	5.0E-03	CYFIP2	0.792	2.3E-03	DPP7	0.740	5.0E-15
PDGFRA	0.885	9.2E-03	DUSP4	0.794	3.6E-06	CORO2A	0.750	1.6E-05
ELOVL7	0.886	2.7E-02	HIF3A	0.800	1.1E-04	TNS1	0.754	6.4E-05
PTK2B	0.887	2.2E-08	HIVEP2	0.807	2.9E-04	EGLN3	0.757	1.6E-10
PITPNC1	0.889	2.6E-06	SLC37A1	0.811	5.5E-03	TGFB1	0.761	3.7E-12
HLA-A	0.891	1.0E-06	PRR15	0.814	6.2E-03	WNT5A	0.762	1.2E-02
LMO4	0.891	2.2E-07	LMO4	0.819	3.3E-04	NETO2	0.769	3.1E-05
FZD4	0.894	3.7E-04	CDC42EP3	0.821	3.2E-02	NPPB	0.775	3.8E-04
FAM83F	0.895	1.0E-09	TMEM200A	0.825	2.1E-02	SLC1A1	0.781	2.2E-07
UPP1	0.903	1.8E-08	GOLGA8A	0.826	4.4E-02	PLCD3	0.788	4.3E-15
DUSP4	0.906	3.6E-09	MALT1	0.826	1.6E-03	ID1	0.791	4.4E-04
DPP7	0.908	5.3E-13	MASP1	0.828	6.3E-06	SGK1	0.800	1.1E-06
NPNT	0.910	4.4E-09	RASL11A	0.830	3.1E-03	GRB10	0.808	8.7E-12
SOAT2	0.912	3.1E-03	IRS1	0.831	8.2E-05	SEMA4B	0.810	3.5E-09
ARG2	0.915	8.4E-06	OAT	0.833	6.4E-06	JAG2	0.816	3.3E-03
DPYSL3	0.916	3.5E-06	MYC	0.835	6.3E-04	SEC14L2	0.816	1.8E-07
MYC	0.920	6.9E-05	ARHGAP45	0.837	2.6E-03	CES2	0.819	4.6E-07
ID1	0.920	5.5E-05	MIR22HG	0.842	1.4E-03	LAMC1	0.832	1.0E-12
SOCS2	0.922	7.3E-09	C6orf132	0.843	5.3E-04	ADGRD1	0.833	5.2E-10
CD3EAP	0.924	1.5E-03	DUOXA2	0.843	1.7E-06	TRIM32	0.838	1.5E-03
RPP25	0.925	1.6E-02	CLDN7	0.845	1.5E-06	FAM83F	0.841	1.3E-07
SOX9	0.926	3.3E-10	ABTB2	0.846	2.8E-05	ITGA2	0.843	1.9E-11
GCNT3	0.927	3.2E-04	CTSK	0.852	3.6E-02	MAPK4	0.863	4.2E-08
GIPC2	0.936	6.8E-05	KCNK5	0.858	3.0E-07	HSPG2	0.863	8.5E-10
B3GNT8	0.939	6.3E-05	LOX	0.862	3.6E-08	FAM43A	0.864	4.1E-07
HEPH	0.954	9.1E-05	FER1L4	0.863	6.7E-04	SEMA6D	0.867	9.5E-12
OAT	0.955	1.6E-07	BCOR	0.863	4.2E-04	LMO4	0.872	3.2E-08

AMOTL2	0.961	2.5E-07	FOSL2	0.865	1.1E-09	SOCS2	0.873	1.6E-04
TFRC	0.964	3.8E-06	ITPKA	0.867	2.1E-08	CTSV	0.878	3.6E-02
ZNF703	0.982	6.6E-07	SELENOP	0.873	4.5E-02	ITPR3	0.882	1.4E-08
ITGA2	0.983	1.3E-05	MYO18B	0.874	1.8E-05	PRR15	0.884	3.2E-06
SPHK1	0.986	3.8E-07	DDR1	0.880	5.0E-04	C11orf86	0.888	1.9E-02
CPA2	0.987	1.1E-02	HSPG2	0.883	9.3E-05	TET1	0.889	2.0E-03
ITPKA	0.991	1.3E-07	PIK3C2B	0.885	1.1E-04	INHBB	0.890	4.4E-07
SNHG15	1.006	9.6E-03	PIM1	0.886	2.8E-03	PCSK5	0.895	8.9E-12
MTCL1	1.008	5.9E-06	KIRREL1	0.887	1.0E-02	TYMP	0.897	6.5E-03
DNASE1	1.025	8.0E-04	ITGA2	0.888	5.0E-07	MTCL1	0.902	2.6E-09
AL392172.1	1.028	7.6E-03	IGSF9	0.890	1.6E-02	SDK2	0.906	8.6E-03
PTAFR	1.030	5.1E-03	ZNF703	0.905	1.1E-04	MYO1A	0.908	2.6E-03
RARRES1	1.030	5.0E-03	PDE9A	0.908	9.9E-04	GRAMD4	0.910	2.2E-06
MMP1	1.030	2.0E-04	GSTA1	0.909	3.4E-05	PDE9A	0.912	1.1E-05
LYPD1	1.035	1.3E-02	SDK2	0.910	2.9E-02	SLC16A10	0.913	2.4E-05
RASL11A	1.047	1.4E-04	NCR3LG1	0.912	1.4E-02	HLA-A	0.915	1.2E-04
CCND2	1.048	2.1E-07	GPC4	0.913	6.8E-07	ICK	0.917	5.2E-10
BMP8B	1.051	5.6E-06	ARG2	0.914	1.3E-03	PIK3C2B	0.920	4.6E-05
GRAMD1B	1.058	3.4E-05	OBSCN	0.914	9.8E-05	SPHK1	0.922	2.3E-05
SLC15A1	1.062	8.9E-08	PITPNC1	0.922	7.9E-04	RIN2	0.922	3.6E-03
NFKBIA	1.064	1.2E-07	SOAT2	0.924	3.3E-03	TRAM2	0.927	1.6E-15
CLDN7	1.067	1.0E-09	SORL1	0.929	1.2E-03	CLMP	0.928	5.4E-06
ISG20	1.072	3.5E-05	PPP1R26	0.933	3.8E-03	SLC37A1	0.931	1.0E-05
CRYBG2	1.082	8.7E-06	TFRC	0.934	7.5E-05	PTK2B	0.932	9.3E-11
KCNMB3	1.086	1.1E-02	ADGRD1	0.935	7.7E-07	SLC15A1	0.932	2.1E-15
PPP1R14A	1.089	1.1E-02	SLC6A8	0.936	6.0E-06	SLC17A4	0.933	9.3E-03
CTSK	1.095	1.1E-07	METRNL	0.937	1.4E-02	SLC26A2	0.933	2.7E-10
PIM1	1.098	1.4E-10	FLNC	0.943	5.1E-14	LGALS3	0.933	5.5E-03
IHH	1.099	2.0E-06	B3GNT8	0.944	1.7E-03	FAM151A	0.943	2.4E-06
SLC30A2	1.104	4.5E-02	PCSK5	0.949	3.5E-09	DEPDC7	0.955	7.3E-10
ALDH1A3	1.107	5.0E-09	ITPR3	0.954	6.3E-08	METRNL	0.960	2.5E-04
GSDMB	1.107	4.6E-03	TIFA	0.966	2.2E-02	MYO18B	0.962	1.5E-09
SELENOP	1.109	3.5E-06	SLC25A29	0.968	7.8E-06	OAT	0.963	3.1E-09
SHH	1.111	4.3E-05	PPARGC1B	0.969	4.6E-03	GBP1	0.966	3.9E-05
PLEKHG5	1.117	5.0E-05	FADS6	0.973	1.9E-02	GIPC2	0.975	1.4E-06
KCNG1	1.124	2.4E-03	INPP5F	0.979	7.1E-05	TNFSF10	0.983	5.3E-06
TMEM163	1.125	4.6E-02	CEACAM19	0.982	2.6E-02	PAQR7	1.010	2.0E-08
TIMP2	1.133	5.6E-06	KCNJ13	0.987	2.0E-02	SLC6A8	1.013	7.3E-08
PDE9A	1.137	9.2E-07	MAPK4	1.001	4.6E-07	CPA2	1.014	4.1E-03
CABLES1	1.139	5.5E-03	SLC15A1	1.003	4.9E-08	TIFA	1.016	6.8E-05

Samantha G. Tilson - January 2021

UGT2B17	1.142	3.2E-02	ENO2	1.004	4.6E-02	CKB	1.022	6.5E-18
SLC46A1	1.145	2.9E-02	SHH	1.008	1.4E-03	TSKU	1.023	5.5E-10
SLC7A7	1.151	6.7E-07	DAB1	1.009	5.8E-04	FADS6	1.025	1.0E-04
DUOX2	1.162	2.8E-18	SOCS2	1.010	1.7E-04	GPSM1	1.038	1.4E-05
HNF4G	1.165	2.7E-03	DPYSL3	1.020	5.9E-08	HNF4G	1.048	2.5E-05
CDKN1C	1.169	3.8E-02	ABHD17C	1.021	3.2E-05	PPARGC1B	1.059	1.7E-03
FADS6	1.193	1.0E-03	CPA2	1.040	2.5E-02	MYO15B	1.073	5.4E-04
MYO15B	1.198	9.7E-04	UNC5B	1.041	3.3E-02	TMEM200A	1.095	4.8E-06
IRS2	1.206	2.0E-06	MGAT3	1.062	1.3E-02	B3GNT8	1.095	3.1E-06
PDGFA	1.207	1.8E-09	DIO3OS	1.065	4.9E-02	DCHS1	1.097	1.7E-08
MAF	1.208	9.4E-06	PDGFA	1.072	1.1E-03	CACNA1D	1.101	2.4E-08
SLC37A2	1.211	1.4E-03	DYNC2H1	1.085	1.9E-02	PITPNC1	1.105	6.9E-09
EDN1	1.219	3.6E-02	BMP8B	1.085	2.9E-05	UGT8	1.111	2.6E-04
PMEPA1	1.220	2.1E-02	ENTPD2	1.087	3.6E-03	GPRC5A	1.117	1.4E-11
CKB	1.229	3.0E-11	CCN2	1.087	2.9E-07	ALDH1A3	1.120	3.3E-03
GSTA1	1.229	1.3E-07	MTCL1	1.092	5.9E-07	TMCC3	1.122	7.6E-04
PRR15L	1.229	1.7E-05	NRGN	1.096	2.6E-02	SELENOP	1.129	1.7E-04
SPOCK1	1.258	3.9E-07	USP2	1.098	2.5E-02	DPYSL3	1.150	2.9E-14
RARRES3	1.259	4.5E-04	SLC7A7	1.102	2.2E-02	RAI2	1.164	3.3E-02
METRNL	1.275	8.4E-07	GPSM1	1.104	9.8E-03	COL4A1	1.166	1.7E-09
ABHD17C	1.279	1.8E-10	CCND2	1.109	4.2E-09	TIMP2	1.171	3.3E-03
NFASC	1.283	4.5E-02	PTAFR	1.114	4.3E-02	CTSK	1.179	1.8E-07
FAM43A	1.297	7.5E-10	NPNT	1.121	4.2E-08	NPNT	1.192	1.7E-12
PHGR1	1.307	2.7E-02	CKB	1.131	3.0E-18	HEPH	1.204	4.5E-08
LRRC75A	1.311	5.7E-03	CACNA1D	1.144	1.3E-04	RARRES3	1.222	2.2E-02
PLAUR	1.316	4.1E-02	ATF3	1.147	4.0E-07	COLEC12	1.228	1.9E-05
PCYT1B	1.320	1.1E-03	HNF4G	1.149	4.7E-04	SOAT2	1.233	9.2E-08
KCNF1	1.324	4.5E-03	TMCC3	1.161	9.5E-04	LAMC2	1.233	5.6E-10
LAMC2	1.333	5.7E-04	C11orf86	1.176	4.8E-03	CCND2	1.252	3.2E-27
SHISA9	1.343	1.4E-03	PLEKHG5	1.181	5.6E-04	SPOCK1	1.255	3.6E-03
NRGN	1.345	9.7E-04	IRS2	1.188	5.8E-09	SLC7A7	1.257	1.4E-04
CTGF	1.346	7.2E-08	FMO5	1.194	1.2E-06	TBX15	1.273	1.7E-02
PRDM16	1.350	4.0E-03	CDKN1C	1.194	2.6E-02	PLAC8	1.285	6.9E-10
KCNG3	1.354	2.9E-06	DNASE1	1.195	3.5E-03	FMO5	1.291	1.9E-09
ABCG1	1.362	9.6E-04	SPOCK1	1.205	3.8E-03	COL4A2	1.306	4.5E-10
LRRC66	1.379	8.6E-03	FXYD3	1.212	3.7E-02	DUOX2	1.312	3.2E-27
CTSE	1.384	1.1E-02	BIRC3	1.214	1.3E-02	PTGER4	1.316	5.0E-05
KLF4	1.385	2.0E-17	SSTR1	1.217	6.5E-03	MYH14	1.320	2.2E-02
ANPEP	1.398	1.7E-04	TIAM2	1.219	4.5E-03	FXYD3	1.323	7.7E-05
SLC5A12	1.400	3.7E-05	OTUD3	1.222	1.3E-04	BMP8B	1.324	1.2E-09

NPPB	1.422	2.7E-03	PUS7L	1.232	9.0E-03	BCOR	1.338	1.1E-26
PLB1	1.429	6.0E-05	LAMC2	1.235	1.2E-02	XKRX	1.362	4.8E-07
ANXA13	1.434	9.3E-03	AMOTL2	1.238	9.2E-10	GUCY2C	1.380	1.2E-02
B3GALT5	1.439	1.5E-03	PLAC8	1.250	9.4E-05	CDKN1C	1.382	1.1E-04
SEMA3B	1.450	7.3E-04	IRAK4	1.254	7.1E-03	ENG	1.382	4.3E-06
HSD3B2	1.455	1.9E-05	IHH	1.275	1.1E-07	DNASE1	1.385	4.8E-06
GOLGA7B	1.457	1.8E-02	FAM43A	1.281	9.2E-05	MALRD1	1.394	8.7E-06
PIGR	1.459	1.1E-03	SEMA5B	1.290	1.5E-02	LYPD1	1.395	1.8E-03
PTGER4	1.464	7.0E-05	KLF4	1.303	8.7E-10	FZD4	1.400	2.1E-15
KCNJ13	1.472	6.2E-07	MAF	1.304	4.4E-07	HSD3B2	1.408	2.3E-08
COLEC12	1.475	1.0E-04	PDE10A	1.307	1.0E-02	SSTR1	1.408	5.9E-07
PLAC8	1.476	5.3E-13	PCYT1B	1.310	6.6E-03	LGALS2	1.410	2.1E-05
FXYD3	1.483	6.4E-04	C15orf48	1.322	4.4E-02	GREM2	1.420	2.8E-11
TMCC3	1.499	1.1E-05	LMNTD2-AS1	1.339	3.7E-02	GOLGA7B	1.422	1.2E-02
GYG2P1	1.506	2.7E-03	ERVH48-1	1.349	3.3E-02	CTSE	1.425	3.7E-03
TYMP	1.514	1.7E-09	MSX2	1.359	3.1E-02	PDGFA	1.429	2.4E-09
ALPI	1.515	1.3E-03	PRR15L	1.364	7.1E-05	ADAMTS12	1.430	5.3E-11
LCT	1.521	5.0E-03	FZD4	1.365	5.0E-11	ABHD17C	1.436	5.3E-12
CLIC5	1.547	3.8E-04	THBD	1.365	1.1E-19	GSTA1	1.446	1.2E-05
ICOSLG	1.548	5.4E-11	PRDM16	1.366	2.4E-03	GRAMD1B	1.466	3.3E-16
PDE10A	1.549	1.9E-05	KCNC3	1.394	6.4E-03	ANXA13	1.471	3.2E-03
TCN1	1.550	4.9E-04		1.394	9.3E-04	KCNG1	1.479	1.2E-04
HHLA2	1.551	8.5E-04	LRRC19	1.405	1.4E-02	SLC5A5	1.482	7.7E-04
RIMS4	1.569	9.2E-03	LINC02381	1.422	4.1E-02	PHYHIPL	1.537	4.1E-05
THBD	1.579	1.9E-16	GREM2	1.424	2.1E-05	PHGR1	1.544	6.8E-03
ADAMTS12	1.582	3.3E-10	ICOSLG	1.430	2.2E-05	SLC5A12	1.551	6.8E-10
CDX2	1.587	2.2E-05	NPPB	1.447	6.3E-20	SLC5A11	1.553	2.0E-05
CBLN1	1.597	1.3E-04	MN1	1.453	1.9E-05	MAF	1.559	3.6E-17
MYO7B	1.598	1.9E-07	CD55	1.475	4.0E-04	PRDM16	1.566	4.5E-05
LINC02381	1.613	9.4E-05	S100G	1.483	1.7E-03	MN1	1.576	1.3E-10
FAM3B	1.613	1.6E-05	SLC5A11	1.491	8.1E-03	IHH	1.576	6.4E-10
PRODH	1.615	5.4E-07	COL4A2	1.493	1.1E-02	KCNJ13	1.611	6.3E-07
EFHD1	1.624	1.8E-04	DUOX2	1.506	3.1E-33	CDH17	1.618	1.6E-02
LINC01108	1.635	4.2E-04	TMPRSS2	1.507	3.2E-02	SLC30A2	1.631	2.9E-04
SLC26A3	1.636	2.3E-02	XKRX	1.513	4.8E-05	PRODH	1.639	2.4E-09
XKRX	1.638	1.4E-11	ANPEP	1.532	1.7E-16	LRRC19	1.651	3.2E-05
C11orf86	1.641	9.4E-06	PLB1	1.534	1.2E-03	CBLN1	1.651	6.2E-05
TMPRSS2	1.641	2.1E-03	KCNG3	1.540	6.0E-04	ANPEP	1.652	4.8E-10
ADGRG7	1.649	5.4E-03	COLEC12	1.540	7.5E-05	PDE10A	1.654	4.7E-11
PLA2G5	1.651	4.5E-03	GRAMD1B	1.544	3.9E-11	HHLA2	1.659	7.0E-08

DLGAP3	1.660	1.3E-04	MYO7B	1.544	3.1E-02	PCYT1B	1.660	6.3E-09
CD244	1.665	9.3E-03	PITX1	1.548	7.0E-06	NRN1	1.706	1.2E-06
SLC5A11	1.676	5.5E-05	FAM3B	1.552	4.0E-03	CA4	1.707	2.5E-02
MN1	1.685	6.3E-10	CD3G	1.585	2.6E-13	GSDMB	1.746	6.4E-10
MSX2	1.694	6.3E-06	ERV3-1	1.588	1.5E-03	ABCG1	1.753	8.0E-07
MNX1	1.703	9.5E-03	GPRC5A	1.612	3.2E-04	SLC26A3	1.760	5.2E-03
SLC5A1	1.707	2.5E-05	ALPI	1.622	6.9E-03	GYG2P1	1.760	1.5E-03
ITGA8	1.721	1.4E-08	KCNG1	1.630	5.8E-03	ADGRG7	1.783	2.7E-04
TBX15	1.727	2.4E-03	SLC5A12	1.637	1.0E-03	LRRC75A	1.801	2.5E-10
C15orf48	1.740	6.3E-04	EDN1	1.658	4.6E-03	TMPRSS2	1.812	2.3E-04
FLT1	1.741	9.3E-05	ADAMTS12	1.661	3.0E-07	KCNMA1	1.828	7.6E-03
GPRC5A	1.758	2.0E-11	SLC26A3	1.661	3.2E-02	PLB1	1.836	5.1E-10
WNT11	1.774	3.6E-04	BICDL2	1.687	4.5E-02	C3orf85	1.868	8.0E-04
B4GALNT2	1.775	5.5E-04	SHISA9	1.750	3.7E-04	ITGA8	1.868	7.6E-14
VIPR1	1.783	3.2E-03	FLT1	1.763	4.0E-04	MELTF	1.869	4.2E-04
TM4SF20	1.784	1.9E-03	SATB2	1.778	2.6E-02	FLT1	1.870	5.6E-08
CD3G	1.799	3.9E-15	PRODH	1.784	2.3E-07	PITX1	1.875	1.4E-09
C6orf222	1.814	2.1E-04	ACE2	1.811	1.2E-11	PLA2G5	1.887	2.6E-03
NRN1	1.825	2.7E-03	MELTF	1.869	8.7E-03	ALPI	1.893	2.6E-05
PITX1	1.840	5.1E-09	CA8	1.922	3.0E-02	NRGN	1.909	1.0E-07
CA4	1.854	1.1E-05	ITGA8	1.928	5.9E-07	DLGAP3	1.923	6.2E-07
SP8	1.859	2.9E-02	CLIC5	1.932	3.2E-06	KCNG3	1.944	5.2E-16
RNF186	1.864	2.2E-02	B3GALT5	1.933	3.0E-04	C15orf48	1.950	1.0E-05
ACE2	1.901	2.1E-05	EFHD1	1.933	3.1E-03	CD3G	1.967	6.7E-16
ZNF469	1.928	4.7E-07	ABCG1	1.959	3.6E-05	MGAM2	1.969	5.0E-03
FABP2	2.033	4.3E-06	DPEP1	2.039	5.7E-04	PRR15L	1.972	1.4E-15
MELTF	2.061	5.6E-06	KCNMA1	2.040	2.8E-05	FABP2	1.975	1.1E-04
CA8	2.063	3.4E-04	DLGAP3	2.051	4.9E-05	CD244	2.001	1.8E-03
WDR72	2.163	3.1E-06	JPH3	2.090	4.2E-03	CLIC5	2.017	9.8E-04
AL121974.1	2.198	3.2E-02	VIPR1	2.118	5.7E-03	LCT	2.059	7.2E-06
S100G	2.268	3.6E-08	SLC5A1	2.280	1.5E-09	MYO7B	2.067	2.3E-06
TMEM200C	2.283	2.6E-16	MUC13	2.287	4.5E-03	WDR72	2.071	1.0E-03
MUC13	2.335	5.3E-05	TBX15	2.290	5.2E-05	TMEM200C	2.078	9.8E-11
FZD8	2.340	9.6E-12	MNX1	2.305	9.9E-03	ACE2	2.142	8.6E-23
GPA33	2.358	4.5E-02	ZNF469	2.347	6.2E-08	SLC5A1	2.159	9.8E-11
GJB4	2.388	4.8E-02	BNIP5	2.350	2.0E-04	ZNF469	2.212	3.6E-15
ZIC1	2.418	3.9E-05	GYG2P1	2.416	5.4E-03	PIGR	2.224	1.1E-07
GPR35	2.431	2.4E-03	PLA2G5	2.441	1.1E-02	FAM3B	2.249	5.2E-09
DPEP1	2.438	4.1E-06	TM4SF20	2.444	1.0E-03	CDX2	2.292	4.9E-06
IRX2	2.553	9.2E-06	TMEM200C	2.649	1.7E-08	DPEP1	2.325	8.2E-06

LINC00982	2.770	7.2E-04	ALOX5	2.683	3.9E-02	TM4SF20	2.379	1.3E-09
ZNF439	2.931	1.4E-02	GPR35	2.718	3.4E-02	FZD8	2.382	1.5E-21
REG1B	2.938	1.7E-02	WDR72	2.767	4.6E-11	MUC13	2.478	2.5E-07
URAD	2.956	1.7E-07	FZD8	2.794	3.7E-04	GPA33	2.599	2.5E-02
VWC2	3.009	8.3E-03	CASP5	2.932	1.5E-02	S100G	2.605	4.3E-13
PTPRT	3.212	4.4E-12	URAD	2.994	7.0E-06	VWC2	2.764	3.1E-02
GDF7	3.272	6.6E-07	BTNL8	3.198	1.3E-08	IRX2	2.780	9.4E-08
SLC2A5	3.376	4.0E-19	GNRH2	3.210	2.4E-09	URAD	2.919	1.0E-10
BTNL8	3.389	1.7E-11	SLC2A5	3.330	3.1E-19	LINC00982	2.971	1.8E-03
NKX2-5	3.787	2.6E-03	GDF7	3.514	5.6E-06	ZIC4	3.222	1.1E-04
DMBT1	3.968	3.6E-03	DMBT1	3.747	1.8E-02	GNRH2	3.268	3.9E-10
ZNF528	3.986	6.7E-03	REG3A	3.768	8.7E-03	SLC2A5	3.286	8.7E-10
CEACAM6	4.159	9.2E-06	ZIC1	3.822	7.6E-09	BTNL8	3.328	4.5E-14
BTNL3	4.164	2.2E-13	SLC9A3	3.937	5.6E-04	SLC9A3	3.330	8.9E-04
GNRH2	4.280	1.0E-16	GPA33	3.979	1.0E-03	PTPRT	3.615	5.2E-06
AC078881.1	4.345	9.4E-04	PTPRT	4.026	1.0E-16	VIPR2	3.840	1.4E-06
ZNF253	4.418	2.5E-03	ZIC5	4.035	2.8E-02	BTNL3	4.065	6.8E-13
HMX1	4.904	1.7E-03	BTNL3	4.836	3.3E-12	ZIC1	4.291	2.4E-13
TBX1	5.516	9.0E-09	HMX1	5.057	1.3E-02	ZNF253	4.360	2.0E-03
ADGRG5	5.599	1.5E-04	VIPR2	5.082	5.7E-04	GDF7	5.287	2.1E-12
VIPR2	6.232	2.7E-06	ZIC4	5.410	4.1E-04	HMX1	5.635	6.2E-05
ZIC4	6.702	2.0E-07	TBX1	6.533	9.4E-08	TBX1	5.837	1.2E-11
CHCHD2	8.158	1.5E-03	CHCHD2	9.293	7.8E-04	CHCHD2	9.695	1.1E-06

9 REFERENCES

1. Kang, L.I., W.M. Mars, and G.K. Michalopoulos, *Signals and cells involved in regulating liver regeneration.* Cells, 2012. **1**(4): p. 1261-92.

2. Allan Tsung, D.A.G., *Gross and Cellular Anatomy of the Liver*, in *Molecular Pathology of Liver Diseases*, S.P.S. Monga, Editor. 2011, Springer US: New York, NY.

3. Robinson, M.W., C. Harmon, and C. O'Farrelly, *Liver immunology and its role in inflammation and homeostasis.* Cell Mol Immunol, 2016. **13**(3): p. 267-76.

4. Stolz, Z.K.J.M.C.D.B., *Ultrastructure of the Hepatocyte*, in *Textbook of Hepatology: From Basic Science to Clinical Practice*, J.P.B.M. Juan Rodés MD, Andres T. Blei MD, Jürg Reichen MD, Mario Rizzetto MD,, Editor. 2007, Blackwell Publishing Ltd.

5. Sabine Colnot, C.P., *Liver Zonation*, in *Molecular Pathology of Liver Diseases*, S.P.S. Monga, Editor. 2011, Springer, US: New York, NY.

6. Alejandro Soto-Gutierrez, N.N.-A., Naoya Kobayashi, *Hepatocytes*
, in *Molecular Pathology of Liver Diseases*, S.P.S. Monga, Editor. 2011, Springer, US: New York, NY.

7. Treyer, A. and A. Musch, *Hepatocyte polarity.* Compr Physiol, 2013. **3**(1): p. 243-87.

8. Sato, K., et al., *Ductular Reaction in Liver Diseases: Pathological Mechanisms and Translational Significances.* Hepatology, 2019. **69**(1): p. 420-430.

9. Gouw, A.S., A.D. Clouston, and N.D. Theise, *Ductular reactions in human liver: diversity at the interface.* Hepatology, 2011. **54**(5): p. 1853-63.

10. Sun, T., S. Annunziato, and J.S. Tchorz, *Hepatic ductular reaction: a double-edged sword.* Aging (Albany NY), 2019. **11**(21): p. 9223-9224.

11. Dufour, J.F., *Biliary Epithelial Cells*, in *Textbook of Hepatology: From Basic Science to Clinical Practice*, J.P.B.M. Juan Rodés MD, Andres T. Blei MD, Jürg Reichen MD, Mario Rizzetto MD,, Editor. 2007: Blackwell Publishing Ltd.

12. Yoshiaki Mizuguchi, S.S., Kumiko Isse, John G. Lunz III, Anthony J. Demetris, *Biliary Epithelial Cells*, in *Molecular Pathology of Liver Diseases*, S.P.S. Monga, Editor. 2011, Springer, US: New York, NY.

13. Pinzani, M., *The Hepatic Stellate Cell*, in *Textbook of Hepatology: From Basic Science to Clinical Practice*, J.P.B.M. Juan Rodés MD, Andres T. Blei MD, Jürg Reichen MD, Mario Rizzetto MD,, Editor. 2007, Blackwell Publishing Ltd.

14. Blomhoff, R., et al., *Transport and storage of vitamin A.* Science, 1990. **250**(4979): p. 399-404.

15. Gandhi, C.R., *Stellate Cells*, in *Molecular Pathology of Liver Diseases*, S.P.S. Monga, Editor. 2011, Springer, US: New York, NY.

16. Gandhi, C.R., *Kupffer Cells*, in *Molecular Pathology of Liver Diseases*, S.P.S. Monga, Editor. 2011, Springer, US: New York, NY.

17. Jaeschke, H., *Kupffer Cells*, in *Textbook of Hepatology: From Basic Science to Clinical Practice*, J.P.B.M. Juan Rodés MD, Andres T. Blei MD, Jürg Reichen MD, Mario Rizzetto MD,, Editor. 2007, Blackwell Publishing Ltd.

18. Shah, D.S.V.H., *Liver Sinusoidal Endothelial Cells*, in *Textbook of Hepatology: From Basic Science to Clinical Practice*, J.P.B.M. Juan Rodés MD, Andres T. Blei MD, Jürg Reichen MD, Mario Rizzetto MD,, Editor. 2007, Blackwell Publishing Ltd.

19. Stolz, D.B., *Sinusoidal Endothelial Cells*, in *Molecular Pathology of Liver Diseases*, S.P.S. Monga, Editor. 2011, Springer, US: New York, NY.

20.   Dirk Raddatz, G.R., *Hepatic Carbohydrate Metabolism*, in *Molecular Pathology of Liver Diseases*, S.P.S. Monga, Editor. 2011, Springer, US: New York, NY.

21.   Titchenell, P.M., M.A. Lazar, and M.J. Birnbaum, *Unraveling the Regulation of Hepatic Metabolism by Insulin.* Trends Endocrinol Metab, 2017. **28**(7): p. 497-505.

22.   Gross, D.N., M. Wan, and M.J. Birnbaum, *The role of FOXO in the regulation of metabolism.* Curr Diab Rep, 2009. **9**(3): p. 208-14.

23.   Zhang, W., et al., *FoxO1 regulates multiple metabolic pathways in the liver: effects on gluconeogenic, glycolytic, and lipogenic gene expression.* J Biol Chem, 2006. **281**(15): p. 10105-17.

24.   Ramnanan, C.J., et al., *Physiologic action of glucagon on liver glucose metabolism.* Diabetes Obes Metab, 2011. **13 Suppl 1**: p. 118-25.

25.   Lin, H.V. and D. Accili, *Hormonal regulation of hepatic glucose production in health and disease.* Cell Metab, 2011. **14**(1): p. 9-19.

26.   Jiansheng Huang, J.B., Janardan K. Reddy, *Hepatic Lipid Metabolism*, in *Molecular Pathology of Liver Diseases*, S.P.S. Monga, Editor. 2011, Springer, US: New York, NY.

27.   Strable, M.S. and J.M. Ntambi, *Genetic control of de novo lipogenesis: role in diet-induced obesity.* Crit Rev Biochem Mol Biol, 2010. **45**(3): p. 199-214.

28.   Horton, J.D., J.L. Goldstein, and M.S. Brown, *SREBPs: activators of the complete program of cholesterol and fatty acid synthesis in the liver.* J Clin Invest, 2002. **109**(9): p. 1125-31.

29.   Filhoulaud, G., et al., *Novel insights into ChREBP regulation and function.* Trends Endocrinol Metab, 2013. **24**(5): p. 257-68.

30.   Benhamed, F., et al., *The lipogenic transcription factor ChREBP dissociates hepatic steatosis from insulin resistance in mice and humans.* J Clin Invest, 2012. **122**(6): p. 2176-94.

31.   Ducharme, N.A. and P.E. Bickel, *Lipid droplets in lipogenesis and lipolysis.* Endocrinology, 2008. **149**(3): p. 942-9.

32.   Murphy, S., S. Martin, and R.G. Parton, *Lipid droplet-organelle interactions; sharing the fats.* Biochim Biophys Acta, 2009. **1791**(6): p. 441-7.

33.   Walther, T.C. and R.V. Farese, Jr., *The life of lipid droplets.* Biochim Biophys Acta, 2009. **1791**(6): p. 459-66.

34.   Zehmer, J.K., et al., *A role for lipid droplets in inter-membrane lipid traffic.* Proteomics, 2009. **9**(4): p. 914-21.

35.   Listenberger, L.L., et al., *Triglyceride accumulation protects against fatty acid-induced lipotoxicity.* Proc Natl Acad Sci U S A, 2003. **100**(6): p. 3077-82.

36.   Davidson, N.O. and G.S. Shelness, *APOLIPOPROTEIN B: mRNA editing, lipoprotein assembly, and presecretory degradation.* Annu Rev Nutr, 2000. **20**: p. 169-93.

37.   Gibbons, G.F., et al., *Synthesis and function of hepatic very-low-density lipoprotein.* Biochem Soc Trans, 2004. **32**(Pt 1): p. 59-64.

38.   Shelness, G.S. and J.A. Sellers, *Very-low-density lipoprotein assembly and secretion.* Curr Opin Lipidol, 2001. **12**(2): p. 151-7.

39.   Rustaeus, S., et al., *Assembly of very low density lipoprotein: a two-step process of apolipoprotein B core lipidation.* J Nutr, 1999. **129**(2S Suppl): p. 463S-466S.

40.   Steinberg, S.J., et al., *Very long-chain acyl-CoA synthetases. Human "bubblegum" represents a new family of proteins capable of activating very long-chain fatty acids.* J Biol Chem, 2000. **275**(45): p. 35162-9.

41.   Reddy, J.K. and T. Hashimoto, *Peroxisomal beta-oxidation and peroxisome proliferator-activated receptor alpha: an adaptive metabolic system.* Annu Rev Nutr, 2001. **21**: p. 193-230.

42.   Laffel, L., *Ketone bodies: a review of physiology, pathophysiology and application of monitoring to diabetes.* Diabetes Metab Res Rev, 1999. **15**(6): p. 412-26.

43.   Bartlett, K. and S. Eaton, *Mitochondrial beta-oxidation.* Eur J Biochem, 2004. **271**(3): p. 462-9.

44.   Wanders, R.J., E.G. van Grunsven, and G.A. Jansen, *Lipid metabolism in peroxisomes: enzymology, functions and dysfunctions of the fatty acid alpha- and beta-oxidation systems in humans.* Biochem Soc Trans, 2000. **28**(2): p. 141-9.

45.   Desvergne, B. and W. Wahli, *Peroxisome proliferator-activated receptors: nuclear control of metabolism.* Endocr Rev, 1999. **20**(5): p. 649-88.

46.   Yu, S., S. Rao, and J.K. Reddy, *Peroxisome proliferator-activated receptors, fatty acid oxidation, steatohepatitis and hepatocarcinogenesis.* Curr Mol Med, 2003. **3**(6): p. 561-72.

47.   Hashimoto, T., et al., *Defect in peroxisome proliferator-activated receptor alpha-inducible fatty acid oxidation determines the severity of hepatic steatosis in response to fasting.* J Biol Chem, 2000. **275**(37): p. 28918-28.

48.   Kersten, S., et al., *Peroxisome proliferator-activated receptor alpha mediates the adaptive response to fasting.* J Clin Invest, 1999. **103**(11): p. 1489-98.

49.   Leone, T.C., C.J. Weinheimer, and D.P. Kelly, *A critical role for the peroxisome proliferator-activated receptor alpha (PPARalpha) in the cellular fasting response: the PPARalpha-null mouse as a model of fatty acid oxidation disorders.* Proc Natl Acad Sci U S A, 1999. **96**(13): p. 7473-8.

50.   Omiecinski, C.J., et al., *Xenobiotic metabolism, disposition, and regulation by receptors: from biochemical phenomenon to predictors of major toxicities.* Toxicol Sci, 2011. **120 Suppl 1**: p. S49-75.

51.   Hart, S.N., et al., *Three patterns of cytochrome P450 gene expression during liver maturation in mice.* Drug Metab Dispos, 2009. **37**(1): p. 116-21.

52.   Bellentani, S., *The epidemiology of non-alcoholic fatty liver disease.* Liver Int, 2017. **37 Suppl 1**: p. 81-84.

53.   Blachier, M., et al., *The burden of liver disease in Europe: a review of available epidemiological data.* J Hepatol, 2013. **58**(3): p. 593-608.

54.   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.

55.   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.

56.   Benedict, M. and X. Zhang, *Non-alcoholic fatty liver disease: An expanded review.* World J Hepatol, 2017. **9**(16): p. 715-732.

57.   Petaja, E.M. and H. Yki-Jarvinen, *Definitions of Normal Liver Fat and the Association of Insulin Sensitivity with Acquired and Genetic NAFLD-A Systematic Review.* Int J Mol Sci, 2016. **17**(5).

58.   Araujo, A.R., et al., *Global epidemiology of non-alcoholic fatty liver disease/non-alcoholic steatohepatitis: What we need in the future.* Liver Int, 2018. **38 Suppl 1**: p. 47-51.

59.   Younossi, Z., et al., *Global Perspectives on Nonalcoholic Fatty Liver Disease and Nonalcoholic Steatohepatitis.* Hepatology, 2019. **69**(6): p. 2672-2682.

60.   Parikh, N.D., et al., *Projected increase in obesity and non-alcoholic-steatohepatitis-related liver transplantation waitlist additions in the United States.* Hepatology, 2019. **70**(2): p. 487-495.

61.   Malik, S.M., et al., *Recurrent disease following liver transplantation for nonalcoholic steatohepatitis cirrhosis.* Liver Transpl, 2009. **15**(12): p. 1843-51.

62.   Turchinovich, A., et al., *Cell-Free Circulating Nucleic Acids as Early Biomarkers for NAFLD and NAFLD-Associated Disorders.* Front Physiol, 2018. **9**: p. 1256.

63.   Buzzetti, E., M. Pinzani, and E.A. Tsochatzis, *The multiple-hit pathogenesis of non-alcoholic fatty liver disease (NAFLD).* Metabolism, 2016. **65**(8): p. 1038-48.

64.   Ipsen, D.H., J. Lykkesfeldt, and P. Tveden-Nyborg, *Molecular mechanisms of hepatic lipid accumulation in non-alcoholic fatty liver disease.* Cell Mol Life Sci, 2018. **75**(18): p. 3313-3327.

65.   Donnelly, K.L., et al., *Sources of fatty acids stored in liver and secreted via lipoproteins in patients with nonalcoholic fatty liver disease.* J Clin Invest, 2005. **115**(5): p. 1343-51.

66.   Lambert, J.E., et al., *Increased de novo lipogenesis is a distinct characteristic of individuals with nonalcoholic fatty liver disease.* Gastroenterology, 2014. **146**(3): p. 726-35.

67.   Rao, M.S. and J.K. Reddy, *Peroxisomal beta-oxidation and steatohepatitis.* Semin Liver Dis, 2001. **21**(1): p. 43-55.

68.   Videla, L.A. and P. Pettinelli, *Misregulation of PPAR Functioning and Its Pathogenic Consequences Associated with Nonalcoholic Fatty Liver Disease in Human Obesity.* PPAR Res, 2012. **2012**: p. 107434.

69.   Fabbrini, E., et al., *Alterations in adipose tissue and hepatic lipid kinetics in obese men and women with nonalcoholic fatty liver disease.* Gastroenterology, 2008. **134**(2): p. 424-31.

70.   Adiels, M., et al., *Overproduction of large VLDL particles is driven by increased liver fat content in man.* Diabetologia, 2006. **49**(4): p. 755-65.

71.   Ota, T., C. Gayet, and H.N. Ginsberg, *Inhibition of apolipoprotein B100 secretion by lipid-induced hepatic endoplasmic reticulum stress in rodents.* J Clin Invest, 2008. **118**(1): p. 316-32.

72.   Zhang, X.Q., et al., *Role of endoplasmic reticulum stress in the pathogenesis of nonalcoholic fatty liver disease.* World J Gastroenterol, 2014. **20**(7): p. 1768-76.

73.   Kitade, H., et al., *Nonalcoholic Fatty Liver Disease and Insulin Resistance: New Insights and Potential New Treatments.* Nutrients, 2017. **9**(4).

74.   Tessari, P., et al., *Hepatic lipid metabolism and non-alcoholic fatty liver disease.* Nutr Metab Cardiovasc Dis, 2009. **19**(4): p. 291-302.

75.   Stefan, N., F. Schick, and H.U. Haring, *Ectopic fat in insulin resistance, dyslipidemia, and cardiometabolic disease.* N Engl J Med, 2014. **371**(23): p. 2236-7.

76.   Kumashiro, N., et al., *Cellular mechanism of insulin resistance in nonalcoholic fatty liver disease.* Proc Natl Acad Sci U S A, 2011. **108**(39): p. 16381-5.

77.   Perry, R.J., et al., *The role of hepatic lipids in hepatic insulin resistance and type 2 diabetes.* Nature, 2014. **510**(7503): p. 84-91.

78. Ter Horst, K.W., et al., *Hepatic Diacylglycerol-Associated Protein Kinase Cepsilon Translocation Links Hepatic Steatosis to Hepatic Insulin Resistance in Humans.* Cell Rep, 2017. **19**(10): p. 1997-2004.

79. Malhi, H. and G.J. Gores, *Molecular mechanisms of lipotoxicity in nonalcoholic fatty liver disease.* Semin Liver Dis, 2008. **28**(4): p. 360-9.

80. Song, M.J. and H. Malhi, *The unfolded protein response and hepatic lipid metabolism in non alcoholic fatty liver disease.* Pharmacol Ther, 2019. **203**: p. 107401.

81. Mota, M., et al., *Molecular mechanisms of lipotoxicity and glucotoxicity in nonalcoholic fatty liver disease.* Metabolism, 2016. **65**(8): p. 1049-61.

82. Svegliati-Baroni, G., et al., *Lipidomic biomarkers and mechanisms of lipotoxicity in non-alcoholic fatty liver disease.* Free Radic Biol Med, 2019. **144**: p. 293-309.

83. Hirsova, P., et al., *Lipotoxic lethal and sublethal stress signaling in hepatocytes: relevance to NASH pathogenesis.* J Lipid Res, 2016. **57**(10): p. 1758-1770.

84. Leamy, A.K., et al., *Enhanced synthesis of saturated phospholipids is associated with ER stress and lipotoxicity in palmitate treated hepatic cells.* J Lipid Res, 2014. **55**(7): p. 1478-88.

85. Ariyama, H., et al., *Decrease in membrane phospholipid unsaturation induces unfolded protein response.* J Biol Chem, 2010. **285**(29): p. 22027-35.

86. Volmer, R., K. van der Ploeg, and D. Ron, *Membrane lipid saturation activates endoplasmic reticulum unfolded protein response transducers through their transmembrane domains.* Proc Natl Acad Sci U S A, 2013. **110**(12): p. 4628-33.

87. Kakisaka, K., et al., *Mechanisms of lysophosphatidylcholine-induced hepatocyte lipoapoptosis.* Am J Physiol Gastrointest Liver Physiol, 2012. **302**(1): p. G77-84.

88. Pagadala, M., et al., *Role of ceramides in nonalcoholic fatty liver disease.* Trends Endocrinol Metab, 2012. **23**(8): p. 365-71.

89. Pettus, B.J., C.E. Chalfant, and Y.A. Hannun, *Ceramide in apoptosis: an overview and current perspectives.* Biochim Biophys Acta, 2002. **1585**(2-3): p. 114-25.

90. Schwabe, R.F. and D.A. Brenner, *Mechanisms of Liver Injury. I. TNF-alpha-induced liver injury: role of IKK, JNK, and ROS pathways.* Am J Physiol Gastrointest Liver Physiol, 2006. **290**(4): p. G583-9.

91. Romeo, S., et al., *Genetic variation in PNPLA3 confers susceptibility to nonalcoholic fatty liver disease.* Nat Genet, 2008. **40**(12): p. 1461-5.

92. Eslam, M., L. Valenti, and S. Romeo, *Genetics and epigenetics of NAFLD and NASH: Clinical impact.* J Hepatol, 2018. **68**(2): p. 268-279.

93. Chalasani, N., et al., *Genome-wide association study identifies variants associated with histologic features of nonalcoholic Fatty liver disease.* Gastroenterology, 2010. **139**(5): p. 1567-76, 1576 e1-6.

94. 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.

95. Hernaez, R., *Genetic factors associated with the presence and progression of nonalcoholic fatty liver disease: a narrative review.* Gastroenterol Hepatol, 2012. **35**(1): p. 32-41.

96. Abul-Husn, N.S., et al., *A Protein-Truncating HSD17B13 Variant and Protection from Chronic Liver Disease.* N Engl J Med, 2018. **378**(12): p. 1096-1106.

97. Stender, S., et al., *Adiposity amplifies the genetic risk of fatty liver disease conferred by multiple loci.* Nat Genet, 2017. **49**(6): p. 842-847.

98.   Danford, C.J., Z.M. Yao, and Z.G. Jiang, *Non-alcoholic fatty liver disease: a narrative review of genetics.* J Biomed Res, 2018. **32**(5): p. 389-400.

99.   Zeybel, M., et al., *Differential DNA methylation of genes involved in fibrosis progression in non-alcoholic fatty liver disease and alcoholic liver disease.* Clin Epigenetics, 2015. **7**: p. 25.

100.   Murphy, S.K., et al., *Relationship between methylome and transcriptome in patients with nonalcoholic fatty liver disease.* Gastroenterology, 2013. **145**(5): p. 1076-87.

101.   Gerhard, G.S. and J.K. DiStefano, *Micro RNAs in the development of non-alcoholic fatty liver disease.* World J Hepatol, 2015. **7**(2): p. 226-34.

102.   Csak, T., et al., *microRNA-122 regulates hypoxia-inducible factor-1 and vimentin in hepatocytes and correlates with fibrosis in diet-induced steatohepatitis.* Liver Int, 2015. **35**(2): p. 532-41.

103.   Cheung, O., et al., *Nonalcoholic steatohepatitis is associated with altered hepatic MicroRNA expression.* Hepatology, 2008. **48**(6): p. 1810-20.

104.   Lowey, B., et al., *Hepatitis C Virus Infection Induces Hepatic Expression of NF-kappaB-Inducing Kinase and Lipogenesis by Downregulating miR-122.* mBio, 2019. **10**(4).

105.   Baulande, S., et al., *Adiponutrin, a transmembrane protein corresponding to a novel dietary- and obesity-linked mRNA specifically expressed in the adipose lineage.* J Biol Chem, 2001. **276**(36): p. 33336-44.

106.   Wilson, P.A., et al., *Characterization of the human patatin-like phospholipase family.* J Lipid Res, 2006. **47**(9): p. 1940-9.

107.   Kienesberger, P.C., et al., *Mammalian patatin domain containing proteins: a family with diverse lipolytic activities involved in multiple biological functions.* J Lipid Res, 2009. **50 Suppl**: p. S63-8.

108.   Fuchs, C.D., T. Claudel, and M. Trauner, *Role of metabolic lipases and lipolytic metabolites in the pathogenesis of NAFLD.* Trends Endocrinol Metab, 2014. **25**(11): p. 576-85.

109.   He, S., et al., *A sequence variation (I148M) in PNPLA3 associated with nonalcoholic fatty liver disease disrupts triglyceride hydrolysis.* J Biol Chem, 2010. **285**(9): p. 6706-15.

110.   Bruschi, F.V., et al., *PNPLA3 expression and its impact on the liver: current perspectives.* Hepat Med, 2017. **9**: p. 55-66.

111.   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.

112.   Pirazzi, C., et al., *PNPLA3 has retinyl-palmitate lipase activity in human hepatic stellate cells.* Hum Mol Genet, 2014. **23**(15): p. 4077-85.

113.   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.

114.   Pingitore, P., et al., *PNPLA3 overexpression results in reduction of proteins predisposing to fibrosis.* Hum Mol Genet, 2016. **25**(23): p. 5212-5222.

115.   Basu Ray, S., *PNPLA3-I148M: a problem of plenty in non-alcoholic fatty liver disease.* Adipocyte, 2019. **8**(1): p. 201-208.

116.   Rydel, T.J., et al., *The crystal structure, mutagenesis, and activity studies reveal that patatin is a lipid acyl hydrolase with a Ser-Asp catalytic dyad.* Biochemistry, 2003. **42**(22): p. 6696-708.

117. Jenkins, C.M., et al., *Identification, cloning, expression, and purification of three novel human calcium-independent phospholipase A2 family members possessing triacylglycerol lipase and acylglycerol transacylase activities.* J Biol Chem, 2004. **279**(47): p. 48968-75.

118. Kumari, M., et al., *Adiponutrin functions as a nutritionally regulated lysophosphatidic acid acyltransferase.* Cell Metab, 2012. **15**(5): p. 691-702.

119. Lake, A.C., et al., *Expression, regulation, and triglyceride hydrolase activity of Adiponutrin family members.* J Lipid Res, 2005. **46**(11): p. 2477-87.

120. Pingitore, P., et al., *Recombinant PNPLA3 protein shows triglyceride hydrolase activity and its I148M mutation results in loss of function.* Biochim Biophys Acta, 2014. **1841**(4): p. 574-80.

121. 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.

122. Pirazzi, C., et al., *Patatin-like phospholipase domain-containing 3 (PNPLA3) I148M (rs738409) affects hepatic VLDL secretion in humans and in vitro.* J Hepatol, 2012. **57**(6): p. 1276-82.

123. Li, J.Z., et al., *Chronic overexpression of PNPLA3I148M in mouse liver causes hepatic steatosis.* J Clin Invest, 2012. **122**(11): p. 4130-44.

124. Luukkonen, P.K., et al., *Hepatic ceramides dissociate steatosis and insulin resistance in patients with non-alcoholic fatty liver disease.* J Hepatol, 2016. **64**(5): p. 1167-1175.

125. Ruhanen, H., et al., *PNPLA3 mediates hepatocyte triacylglycerol remodeling.* J Lipid Res, 2014. **55**(4): p. 739-46.

126. Peter, A., et al., *PNPLA3 variant I148M is associated with altered hepatic lipid composition in humans.* Diabetologia, 2014. **57**(10): p. 2103-7.

127. Mitsche, M.A., H.H. Hobbs, and J.C. Cohen, *Patatin-like phospholipase domain-containing protein 3 promotes transfer of essential fatty acids from triglycerides to phospholipids in hepatic lipid droplets.* J Biol Chem, 2018. **293**(18): p. 6958-6968.

128. Chamoun, Z., et al., *PNPLA3/adiponutrin functions in lipid droplet formation.* Biol Cell, 2013. **105**(5): p. 219-233.

129. 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.

130. Dubuquoy, C., et al., *Distinct regulation of adiponutrin/PNPLA3 gene expression by the transcription factors ChREBP and SREBP1c in mouse and human hepatocytes.* J Hepatol, 2011. **55**(1): p. 145-53.

131. Hao, L., et al., *Shifts in dietary carbohydrate-lipid exposure regulate expression of the non-alcoholic fatty liver disease-associated gene PNPLA3/adiponutrin in mouse liver and HepG2 human liver cells.* Metabolism, 2014. **63**(10): p. 1352-62.

132. Rotman, Y., et al., *The association of genetic variability in patatin-like phospholipase domain-containing protein 3 (PNPLA3) with histological severity of nonalcoholic fatty liver disease.* Hepatology, 2010. **52**(3): p. 894-903.

133. 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.

134.	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.

135.	Krawczyk, M., et al., *Variant adiponutrin (PNPLA3) represents a common fibrosis risk gene: non-invasive elastography-based study in chronic liver disease.* J Hepatol, 2011. **55**(2): p. 299-306.

136.	Namjou, B., et al., *GWAS and enrichment analyses of non-alcoholic fatty liver disease identify new trait-associated genes and pathways across eMERGE Network.* BMC Med, 2019. **17**(1): p. 135.

137.	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.

138.	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.

139.	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.

140.	Stattermayer, A.F., et al., *Hepatic steatosis in Wilson disease--Role of copper and PNPLA3 mutations.* J Hepatol, 2015. **63**(1): p. 156-63.

141.	Salameh, H., et al., *PNPLA3 polymorphism increases risk for and severity of chronic hepatitis C liver disease.* World J Hepatol, 2016. **8**(35): p. 1584-1592.

142.	Tian, C., et al., *Variant in PNPLA3 is associated with alcoholic liver disease.* Nat Genet, 2010. **42**(1): p. 21-3.

143.	Valenti, L., et al., *Patatin-like phospholipase domain containing-3 gene I148M polymorphism, steatosis, and liver damage in hereditary hemochromatosis.* World J Gastroenterol, 2012. **18**(22): p. 2813-20.

144.	Vigano, M., et al., *Patatin-like phospholipase domain-containing 3 I148M affects liver steatosis in patients with chronic hepatitis B.* Hepatology, 2013. **58**(4): p. 1245-52.

145.	Dongiovanni, P., et al., *PNPLA3 I148M polymorphism and progressive liver disease.* World J Gastroenterol, 2013. **19**(41): p. 6969-78.

146.	Cox, A.J., et al., *Association of PNPLA3 SNP rs738409 with liver density in African Americans with type 2 diabetes mellitus.* Diabetes Metab, 2011. **37**(5): p. 452-5.

147.	Hotta, K., et al., *Association of the rs738409 polymorphism in PNPLA3 with liver damage and the development of nonalcoholic fatty liver disease.* BMC Med Genet, 2010. **11**: p. 172.

148.	Lin, Y.C., et al., *A common variant in the PNPLA3 gene is a risk factor for non-alcoholic fatty liver disease in obese Taiwanese children.* J Pediatr, 2011. **158**(5): p. 740-4.

149.	Valenti, L., et al., *Homozygosity for the patatin-like phospholipase-3/adiponutrin I148M polymorphism influences liver fibrosis in patients with nonalcoholic fatty liver disease.* Hepatology, 2010. **51**(4): p. 1209-17.

150.	Romeo, S., et al., *Morbid obesity exposes the association between PNPLA3 I148M (rs738409) and indices of hepatic injury in individuals of European descent.* Int J Obes (Lond), 2010. **34**(1): p. 190-4.

151. Zain, S.M., et al., *A multi-ethnic study of a PNPLA3 gene variant and its association with disease severity in non-alcoholic fatty liver disease.* Hum Genet, 2012. **131**(7): p. 1145-52.

152. 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.

153. 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.

154. Nobili, V., et al., *Influence of dietary pattern, physical activity, and I148M PNPLA3 on steatosis severity in at-risk adolescents.* Genes Nutr, 2014. **9**(3): p. 392.

155. Barata, L., et al., *Insulin Resistance Exacerbates Genetic Predisposition to Nonalcoholic Fatty Liver Disease in Individuals Without Diabetes.* Hepatol Commun, 2019. **3**(7): p. 894-907.

156. Shen, J., et al., *PNPLA3 gene polymorphism and response to lifestyle modification in patients with nonalcoholic fatty liver disease.* J Gastroenterol Hepatol, 2015. **30**(1): p. 139-46.

157. Sevastianova, K., et al., *Genetic variation in PNPLA3 (adiponutrin) confers sensitivity to weight loss-induced decrease in liver fat in humans.* Am J Clin Nutr, 2011. **94**(1): p. 104-11.

158. Marzuillo, P., et al., *Weight loss allows the dissection of the interaction between abdominal fat and PNPLA3 (adiponutrin) in the liver damage of obese children.* J Hepatol, 2013. **59**(5): p. 1143-4.

159. Younossi, Z., et al., *Global burden of NAFLD and NASH: trends, predictions, risk factors and prevention.* Nat Rev Gastroenterol Hepatol, 2018. **15**(1): p. 11-20.

160. 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.

161. 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.

162. BasuRay, S., et al., *The PNPLA3 variant associated with fatty liver disease (I148M) accumulates on lipid droplets by evading ubiquitylation.* Hepatology, 2017. **66**(4): p. 1111-1124.

163. Smagris, E., et al., *Pnpla3I148M knockin mice accumulate PNPLA3 on lipid droplets and develop hepatic steatosis.* Hepatology, 2015. **61**(1): p. 108-18.

164. Wang, Y., et al., *PNPLA3, CGI-58, and Inhibition of Hepatic Triglyceride Hydrolysis in Mice.* Hepatology, 2019. **69**(6): p. 2427-2441.

165. Luukkonen, P.K., et al., *Human PNPLA3-I148M variant increases hepatic retention of polyunsaturated fatty acids.* JCI Insight, 2019. **4**(16).

166. Bruschi, F.V., et al., *PNPLA3 I148M Variant Impairs Liver X Receptor Signaling and Cholesterol Homeostasis in Human Hepatic Stellate Cells.* Hepatol Commun, 2019. **3**(9): p. 1191-1204.

167. Dutta, A.K., *Adiponutrin (PNPLA3) in liver fibrogenesis: Is unaltered HepG2 cell line a better model system compared to murine models?* Med Hypotheses, 2015. **85**(6): p. 736-9.

168. Yildirim, S., *Induced Pluripotent Stem Cells.* 2012, New York, NY: Springer.

169.   Takahashi, K. and S. Yamanaka, *Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors.* Cell, 2006. **126**(4): p. 663-76.

170.   Liu, H., et al., *Generation of induced pluripotent stem cells from adult rhesus monkey fibroblasts.* Cell Stem Cell, 2008. **3**(6): p. 587-90.

171.   Li, W., et al., *Generation of rat and human induced pluripotent stem cells by combining genetic reprogramming and chemical inhibitors.* Cell Stem Cell, 2009. **4**(1): p. 16-9.

172.   Park, I.H., et al., *Reprogramming of human somatic cells to pluripotency with defined factors.* Nature, 2008. **451**(7175): p. 141-6.

173.   Grandy, R., R.A. Tomaz, and L. Vallier, *Modeling Disease with Human Inducible Pluripotent Stem Cells.* Annu Rev Pathol, 2019. **14**: p. 449-468.

174.   Warren, C.R., et al., *Induced Pluripotent Stem Cell Differentiation Enables Functional Validation of GWAS Variants in Metabolic Disease.* Cell Stem Cell, 2017. **20**(4): p. 547-557 e7.

175.   DeBoever, C., et al., *Large-Scale Profiling Reveals the Influence of Genetic Variation on Gene Expression in Human Induced Pluripotent Stem Cells.* Cell Stem Cell, 2017. **20**(4): p. 533-546 e7.

176.   Pashos, E.E., et al., *Large, Diverse Population Cohorts of hiPSCs and Derived Hepatocyte-like Cells Reveal Functional Genetic Variation at Blood Lipid-Associated Loci.* Cell Stem Cell, 2017. **20**(4): p. 558-570 e10.

177.   Kimbrel, E.A. and R. Lanza, *Current status of pluripotent stem cells: moving the first therapies to the clinic.* Nat Rev Drug Discov, 2015. **14**(10): p. 681-92.

178.   Lo, B. and L. Parham, *Ethical issues in stem cell research.* Endocr Rev, 2009. **30**(3): p. 204-13.

179.   Benam, K.H., et al., *Engineered in vitro disease models.* Annu Rev Pathol, 2015. **10**: p. 195-262.

180.   Yiangou, L., et al., *Human Pluripotent Stem Cell-Derived Endoderm for Modeling Development and Clinical Applications.* Cell Stem Cell, 2018. **22**(4): p. 485-499.

181.   Kyffin, J.A., et al., *Impact of cell types and culture methods on the functionality of in vitro liver systems - A review of cell systems for hepatotoxicity assessment.* Toxicol In Vitro, 2018. **48**: p. 262-275.

182.   Roy-Chowdhury, N., et al., *Hepatocyte-like cells derived from induced pluripotent stem cells.* Hepatol Int, 2017. **11**(1): p. 54-69.

183.   Si-Tayeb, K., et al., *Highly efficient generation of human hepatocyte-like cells from induced pluripotent stem cells.* Hepatology, 2010. **51**(1): p. 297-305.

184.   Hannan, N.R., et al., *Production of hepatocyte-like cells from human pluripotent stem cells.* Nat Protoc, 2013. **8**(2): p. 430-7.

185.   Rashid, S.T., et al., *Modeling inherited metabolic disorders of the liver using human induced pluripotent stem cells.* J Clin Invest, 2010. **120**(9): p. 3127-36.

186.   Sampaziotis, F., et al., *Directed differentiation of human induced pluripotent stem cells into functional cholangiocyte-like cells.* Nat Protoc, 2017. **12**(4): p. 814-827.

187.   Cayo, M.A., et al., *JD induced pluripotent stem cell-derived hepatocytes faithfully recapitulate the pathophysiology of familial hypercholesterolemia.* Hepatology, 2012. **56**(6): p. 2163-71.

188.   Sampaziotis, F., C.P. Segeritz, and L. Vallier, *Potential of human induced pluripotent stem cells in studies of liver disease.* Hepatology, 2015. **62**(1): p. 303-11.

189.   Zhang, S., et al., *Rescue of ATP7B function in hepatocyte-like cells from Wilson's disease induced pluripotent stem cells using gene therapy or the chaperone drug curcumin.* Hum Mol Genet, 2011. **20**(16): p. 3176-87.

190.   Yusa, K., et al., *Targeted gene correction of alpha1-antitrypsin deficiency in induced pluripotent stem cells.* Nature, 2011. **478**(7369): p. 391-4.

191.   Carpentier, A., et al., *Engrafted human stem cell-derived hepatocytes establish an infectious HCV murine model.* J Clin Invest, 2014. **124**(11): p. 4953-64.

192.   Schwartz, R.E., et al., *Modeling hepatitis C virus infection using human induced pluripotent stem cells.* Proc Natl Acad Sci U S A, 2012. **109**(7): p. 2544-8.

193.   Ouchi, R., et al., *Modeling Steatohepatitis in Humans with Pluripotent Stem Cell-Derived Organoids.* Cell Metab, 2019. **30**(2): p. 374-384 e6.

194.   Coll, M., et al., *Generation of Hepatic Stellate Cells from Human Pluripotent Stem Cells Enables In Vitro Modeling of Liver Fibrosis.* Cell Stem Cell, 2018. **23**(1): p. 101-113 e7.

195.   Takebe, T., et al., *Vascularized and functional human liver from an iPSC-derived organ bud transplant.* Nature, 2013. **499**(7459): p. 481-4.

196.   Ortmann, D. and L. Vallier, *Variability of human pluripotent stem cell lines.* Curr Opin Genet Dev, 2017. **46**: p. 179-185.

197.   Kilpinen, H., et al., *Common genetic variation drives molecular heterogeneity in human iPSCs.* Nature, 2017. **546**(7658): p. 370-375.

198.   Cuomo, A.S.E., et al., *Single-cell RNA-sequencing of differentiating iPS cells reveals dynamic genetic effects on gene expression.* Nat Commun, 2020. **11**(1): p. 810.

199.   Rouhani, F., et al., *Genetic background drives transcriptional variation in human induced pluripotent stem cells.* PLoS Genet, 2014. **10**(6): p. e1004432.

200.   Gieseck, R.L., 3rd, et al., *Maturation of induced pluripotent stem cell derived hepatocytes by 3D-culture.* PLoS One, 2014. **9**(1): p. e86372.

201.   Kim, J.H., et al., *Enhanced Metabolizing Activity of Human ES Cell-Derived Hepatocytes Using a 3D Culture System with Repeated Exposures to Xenobiotics.* Toxicol Sci, 2015. **147**(1): p. 190-206.

202.   Ware, B.R., D.R. Berger, and S.R. Khetani, *Prediction of Drug-Induced Liver Injury in Micropatterned Co-cultures Containing iPSC-Derived Human Hepatocytes.* Toxicol Sci, 2015. **145**(2): p. 252-62.

203.   Bell, C.C., et al., *Characterization of primary human hepatocyte spheroids as a model system for drug-induced liver injury, liver function and disease.* Sci Rep, 2016. **6**: p. 25187.

204.   Vorrink, S.U., et al., *Endogenous and xenobiotic metabolic stability of primary human hepatocytes in long-term 3D spheroid cultures revealed by a combination of targeted and untargeted metabolomics.* FASEB J, 2017. **31**(6): p. 2696-2708.

205.   Bell, C.C., et al., *Comparison of Hepatic 2D Sandwich Cultures and 3D Spheroids for Long-term Toxicity Applications: A Multicenter Study.* Toxicol Sci, 2018. **162**(2): p. 655-666.

206.   Wang, H.X., et al., *CRISPR/Cas9-Based Genome Editing for Disease Modeling and Therapy: Challenges and Opportunities for Nonviral Delivery.* Chem Rev, 2017. **117**(15): p. 9874-9906.

207. Doudna, J.A. and E. Charpentier, *Genome editing. The new frontier of genome engineering with CRISPR-Cas9.* Science, 2014. **346**(6213): p. 1258096.

208. Patterson, A.G., M.S. Yevstigneyeva, and P.C. Fineran, *Regulation of CRISPR-Cas adaptive immune systems.* Curr Opin Microbiol, 2017. **37**: p. 1-7.

209. Sander, J.D. and J.K. Joung, *CRISPR-Cas systems for editing, regulating and targeting genomes.* Nat Biotechnol, 2014. **32**(4): p. 347-55.

210. Ma, Y., L. Zhang, and X. Huang, *Genome modification by CRISPR/Cas9.* FEBS J, 2014. **281**(23): p. 5186-93.

211. Agrotis, A. and R. Ketteler, *A new age in functional genomics using CRISPR/Cas9 in arrayed library screening.* Front Genet, 2015. **6**: p. 300.

212. Rees, H.A. and D.R. Liu, *Base editing: precision chemistry on the genome and transcriptome of living cells.* Nat Rev Genet, 2018. **19**(12): p. 770-788.

213. Ran, F.A., et al., *Genome engineering using the CRISPR-Cas9 system.* Nat Protoc, 2013. **8**(11): p. 2281-2308.

214. Bassett, A.R., *Editing the genome of hiPSC with CRISPR/Cas9: disease models.* Mamm Genome, 2017. **28**(7-8): p. 348-364.

215. Cardenas-Diaz, F.L., et al., *Modeling Monogenic Diabetes using Human ESCs Reveals Developmental and Metabolic Deficiencies Caused by Mutations in HNF1A.* Cell Stem Cell, 2019. **25**(2): p. 273-289 e5.

216. Zhang, Y., et al., *Patient iPSC-Derived Neurons for Disease Modeling of Frontotemporal Dementia with Mutation in CHMP2B.* Stem Cell Reports, 2017. **8**(3): p. 648-658.

217. Yumlu, S., et al., *Gene editing and clonal isolation of human induced pluripotent stem cells using CRISPR/Cas9.* Methods, 2017. **121-122**: p. 29-44.

218. Veres, A., et al., *Low incidence of off-target mutations in individual CRISPR-Cas9 and TALEN targeted human stem cell clones detected by whole-genome sequencing.* Cell Stem Cell, 2014. **15**(1): p. 27-30.

219. Suzuki, K., et al., *Targeted gene correction minimally impacts whole-genome mutational load in human-disease-specific induced pluripotent stem cell clones.* Cell Stem Cell, 2014. **15**(1): p. 31-6.

220. Ben Jehuda, R., Y. Shemer, and O. Binah, *Genome Editing in Induced Pluripotent Stem Cells using CRISPR/Cas9.* Stem Cell Rev Rep, 2018. **14**(3): p. 323-336.

221. Kimberland, M.L., et al., *Strategies for controlling CRISPR/Cas9 off-target effects and biological variations in mammalian genome editing experiments.* J Biotechnol, 2018. **284**: p. 91-101.

222. Mout, R., et al., *Direct Cytosolic Delivery of CRISPR/Cas9-Ribonucleoprotein for Efficient Gene Editing.* ACS Nano, 2017. **11**(3): p. 2452-2458.

223. Liang, X., et al., *Rapid and highly efficient mammalian cell engineering via Cas9 protein transfection.* J Biotechnol, 2015. **208**: p. 44-53.

224. Liu, Y., et al., *Genome-Wide Study Links PNPLA3 Variant With Elevated Hepatic Transaminase After Acute Lymphoblastic Leukemia Therapy.* Clin Pharmacol Ther, 2017. **102**(1): p. 131-140.

225. Gutierrez-Camino, A., I. Martin-Guerrero, and A. Garcia-Orad, *PNPLA3 rs738409 and Hepatotoxicity in Children With B-cell Acute Lymphoblastic Leukemia: A Validation Study in a Spanish Cohort.* Clin Pharmacol Ther, 2017. **102**(6): p. 906.

226. Ramasamy, T.S., et al., *Application of three-dimensional culture conditions to human embryonic stem cell-derived definitive endoderm cells enhances*

hepatocyte differentiation and functionality. Tissue Eng Part A, 2013. **19**(3-4): p. 360-7.

227. Lauschke, V.M., et al., *3D Primary Hepatocyte Culture Systems for Analyses of Liver Diseases, Drug Metabolism, and Toxicity: Emerging Culture Paradigms and Applications.* Biotechnol J, 2019. **14**(7): p. e1800347.

228. Foty, R., *A simple hanging drop cell culture protocol for generation of 3D spheroids.* J Vis Exp, 2011(51).

229. Baze, A., et al., *Three-Dimensional Spheroid Primary Human Hepatocytes in Monoculture and Coculture with Nonparenchymal Cells.* Tissue Eng Part C Methods, 2018. **24**(9): p. 534-545.

230. Eyster, K.M., *The membrane and lipids as integral participants in signal transduction: lipid signal transduction for the non-lipid biochemist.* Adv Physiol Educ, 2007. **31**(1): p. 5-16.

231. Mantzaris, M.D., E.V. Tsianos, and D. Galaris, *Interruption of triacylglycerol synthesis in the endoplasmic reticulum is the initiating event for saturated fatty acid-induced lipotoxicity in liver cells.* FEBS J, 2011. **278**(3): p. 519-30.

232. Pickhardt, P.J., et al., *Natural history of hepatic steatosis: observed outcomes for subsequent liver and cardiovascular complications.* AJR Am J Roentgenol, 2014. **202**(4): p. 752-8.

233. Yilmaz, Y., *Review article: is non-alcoholic fatty liver disease a spectrum, or are steatosis and non-alcoholic steatohepatitis distinct conditions?* Aliment Pharmacol Ther, 2012. **36**(9): p. 815-23.

234. Ricchi, M., et al., *Differential effect of oleic and palmitic acid on lipid accumulation and apoptosis in cultured hepatocytes.* J Gastroenterol Hepatol, 2009. **24**(5): p. 830-40.

235. Abdelmagid, S.A., et al., *Comprehensive profiling of plasma fatty acid concentrations in young healthy Canadian adults.* PLoS One, 2015. **10**(2): p. e0116195.

236. Feng, R., et al., *Free fatty acids profile among lean, overweight and obese non-alcoholic fatty liver disease patients: a case - control study.* Lipids Health Dis, 2017. **16**(1): p. 165.

237. Hyysalo, J., et al., *Circulating triacylglycerol signatures in nonalcoholic fatty liver disease associated with the I148M variant in PNPLA3 and with obesity.* Diabetes, 2014. **63**(1): p. 312-22.

238. Tofte, N., et al., *Lipidomic analysis reveals sphingomyelin and phosphatidylcholine species associated with renal impairment and all-cause mortality in type 1 diabetes.* Sci Rep, 2019. **9**(1): p. 16398.

239. Mousa, A., et al., *Lipidomic profiling reveals early-stage metabolic dysfunction in overweight or obese humans.* Biochim Biophys Acta Mol Cell Biol Lipids, 2019. **1864**(3): p. 335-343.

240. Min, H.K., et al., *Metabolic profiling reveals that PNPLA3 induces widespread effects on metabolism beyond triacylglycerol remodeling in Huh-7 hepatoma cells.* Am J Physiol Gastrointest Liver Physiol, 2014. **307**(1): p. G66-76.

241. Mancina, R.M., et al., *Paradoxical dissociation between hepatic fat content and de novo lipogenesis due to PNPLA3 sequence variant.* J Clin Endocrinol Metab, 2015. **100**(5): p. E821-5.

242. Franko, A., et al., *Dissociation of Fatty Liver and Insulin Resistance in I148M PNPLA3 Carriers: Differences in Diacylglycerol (DAG) FA18:1 Lipid Species as a Possible Explanation.* Nutrients, 2018. **10**(9).

243. Yahagi, N., et al., *A crucial role of sterol regulatory element-binding protein-1 in the regulation of lipogenic gene expression by polyunsaturated fatty acids.* J Biol Chem, 1999. **274**(50): p. 35840-4.

244. Clarke, S.D. and D.B. Jump, *Dietary polyunsaturated fatty acid regulation of gene transcription.* Annu Rev Nutr, 1994. **14**: p. 83-98.

245. Maruyama, H., et al., *Linoleate appears to protect against palmitate-induced inflammation in Huh7 cells.* Lipids Health Dis, 2014. **13**: p. 78.

246. Videla, L.A., et al., *Oxidative stress and depletion of hepatic long-chain polyunsaturated fatty acids may contribute to nonalcoholic fatty liver disease.* Free Radic Biol Med, 2004. **37**(9): p. 1499-507.

247. Zhang, Y., et al., *Effect of alpha-linolenic acid on endoplasmic reticulum stress-mediated apoptosis of palmitic acid lipotoxicity in primary rat hepatocytes.* Lipids Health Dis, 2011. **10**: p. 122.

248. Olzmann, J.A. and P. Carvalho, *Dynamics and functions of lipid droplets.* Nat Rev Mol Cell Biol, 2019. **20**(3): p. 137-155.

249. Xia, J., et al., *MetaboAnalyst: a web server for metabolomic data analysis and interpretation.* Nucleic Acids Res, 2009. **37**(Web Server issue): p. W652-60.

250. Xia, J. and D.S. Wishart, *Metabolomic data processing, analysis, and interpretation using MetaboAnalyst.* Curr Protoc Bioinformatics, 2011. **Chapter 14**: p. Unit 14 10.

251. Slotte, J.P. and B. Ramstedt, *The functional role of sphingomyelin in cell membranes.* European Journal of Lipid Science and Technology, 2007. **109**(10): p. 977-981.

252. Tanaka, N., et al., *Highly purified eicosapentaenoic acid treatment improves nonalcoholic steatohepatitis.* J Clin Gastroenterol, 2008. **42**(4): p. 413-8.

253. Nogueira, M.A., et al., *Omega-3 polyunsaturated fatty acids in treating non-alcoholic steatohepatitis: A randomized, double-blind, placebo-controlled trial.* Clin Nutr, 2016. **35**(3): p. 578-86.

254. Takeuchi, Y., et al., *Polyunsaturated fatty acids selectively suppress sterol regulatory element-binding protein-1 through proteolytic processing and autoloop regulatory circuit.* J Biol Chem, 2010. **285**(15): p. 11681-91.

255. Yoshikawa, T., et al., *Polyunsaturated fatty acids suppress sterol regulatory element-binding protein 1c promoter activity by inhibition of liver X receptor (LXR) binding to LXR response elements.* J Biol Chem, 2002. **277**(3): p. 1705-11.

256. Thiam, A.R. and M. Beller, *The why, when and how of lipid droplet diversity.* J Cell Sci, 2017. **130**(2): p. 315-324.

257. Nakamura, M.T., et al., *Mechanisms of regulation of gene expression by fatty acids.* Lipids, 2004. **39**(11): p. 1077-83.

258. Clarke, S.D., et al., *Peroxisome proliferator-activated receptors: a family of lipid-activated transcription factors.* Am J Clin Nutr, 1999. **70**(4): p. 566-71.

259. *Appendix: Drug Metabolizing Enzymes and Biotransformation Reactions*, in *ADME-Enabling Technologies in Drug Design and Development*. p. 545-565.

260. Almazroo, O.A., M.K. Miah, and R. Venkataramanan, *Drug Metabolism in the Liver.* Clin Liver Dis, 2017. **21**(1): p. 1-20.

261. Udayan Apte, P.K., *Detoxification Functions of the Liver*, in *Molecular Pathology of Liver Diseases*, S.P.S. Monga, Editor. 2011, Springer, US: New York, NY.

262. Cobbina, E. and F. Akhlaghi, *Non-alcoholic fatty liver disease (NAFLD) - pathogenesis, classification, and effect on drug metabolizing enzymes and transporters.* Drug Metab Rev, 2017. **49**(2): p. 197-211.

263. Lake, A.D., et al., *Analysis of global and absorption, distribution, metabolism, and elimination gene expression in the progressive stages of human nonalcoholic fatty liver disease.* Drug Metab Dispos, 2011. **39**(10): p. 1954-60.

264. Donato, M.T., et al., *Potential impact of steatosis on cytochrome P450 enzymes of human hepatocytes isolated from fatty liver grafts.* Drug Metab Dispos, 2006. **34**(9): p. 1556-62.

265. Massart, J., et al., *Role of nonalcoholic fatty liver disease as risk factor for drug-induced hepatotoxicity.* J Clin Transl Res, 2017. **3**(Suppl 1): p. 212-232.

266. Woolsey, S.J., et al., *CYP3A Activity and Expression in Nonalcoholic Fatty Liver Disease.* Drug Metab Dispos, 2015. **43**(10): p. 1484-90.

267. Fisher, C.D., et al., *Hepatic cytochrome P450 enzyme alterations in humans with progressive stages of nonalcoholic fatty liver disease.* Drug Metab Dispos, 2009. **37**(10): p. 2087-94.

268. Hardwick, R.N., et al., *Altered UDP-glucuronosyltransferase and sulfotransferase expression and function during progressive stages of human nonalcoholic fatty liver disease.* Drug Metab Dispos, 2013. **41**(3): p. 554-61.

269. Phieler, J., et al., *The role of the complement system in metabolic organs and metabolic diseases.* Semin Immunol, 2013. **25**(1): p. 47-53.

270. Bykov, I., et al., *Complement C3 contributes to ethanol-induced liver steatosis in mice.* Ann Med, 2006. **38**(4): p. 280-6.

271. Rensen, S.S., et al., *Activation of the complement system in human nonalcoholic fatty liver disease.* Hepatology, 2009. **50**(6): p. 1809-17.

272. Hong, F., et al., *Elevated interleukin-6 during ethanol consumption acts as a potential endogenous protective cytokine against ethanol-induced apoptosis in the liver: involvement of induction of Bcl-2 and Bcl-x(L) proteins.* Oncogene, 2002. **21**(1): p. 32-43.

273. Moyer, A.M., et al., *Acetaminophen-NAPQI hepatotoxicity: a cell line model system genome-wide association study.* Toxicol Sci, 2011. **120**(1): p. 33-41.

274. Liu, W., et al., *Transcriptional regulation of PNPLA3 and its impact on susceptibility to nonalcoholic fatty liver Disease (NAFLD) in humans.* Aging (Albany NY), 2016. **9**(1): p. 26-40.

275. Donati, B., et al., *The rs2294918 E434K variant modulates patatin-like phospholipase domain-containing 3 expression and liver damage.* Hepatology, 2016. **63**(3): p. 787-98.

276. Das, K. and P. Kar, *Non-alcoholic steatohepatitis.* J Assoc Physicians India, 2005. **53**: p. 195-9.

277. Wruck, W., et al., *Concise Review: Current Status and Future Directions on Research Related to Nonalcoholic Fatty Liver Disease.* Stem Cells, 2017. **35**(1): p. 89-96.

278. Peverill, W., L.W. Powell, and R. Skoien, *Evolving concepts in the pathogenesis of NASH: beyond steatosis and inflammation.* Int J Mol Sci, 2014. **15**(5): p. 8591-638.

279. Neuschwander-Tetri, B.A., *Nontriglyceride hepatic lipotoxicity: the new paradigm for the pathogenesis of NASH.* Curr Gastroenterol Rep, 2010. **12**(1): p. 49-56.

280.   Yamaguchi, K., et al., *Inhibiting triglyceride synthesis improves hepatic steatosis but exacerbates liver damage and fibrosis in obese mice with nonalcoholic steatohepatitis.* Hepatology, 2007. **45**(6): p. 1366-74.

281.   Luukkonen, P.K., et al., *Saturated Fat Is More Metabolically Harmful for the Human Liver Than Unsaturated Fat or Simple Sugars.* Diabetes Care, 2018. **41**(8): p. 1732-1739.

282.   Alasoo, K., et al., *Transcriptional profiling of macrophages derived from monocytes and iPS cells identifies a conserved response to LPS and novel alternative transcription.* Sci Rep, 2015. **5**: p. 12524.

283.   Grompe, M. and S. Strom, *Mice with human livers.* Gastroenterology, 2013. **145**(6): p. 1209-14.

284.   Bissig, K.D., et al., *P450-Humanized and Human Liver Chimeric Mouse Models for Studying Xenobiotic Metabolism and Toxicity.* Drug Metab Dispos, 2018. **46**(11): p. 1734-1744.