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