Genomic Studies in Non-Alcoholic Fatty Liver Disease

Yang-Lin Liu

A thesis submitted to the University of Newcastle for the Degree of Doctor of Philosophy

Institute of Cellular Medicine

Newcastle University

May 2017

Abstract

Non-alcoholic fatty liver disease (NAFLD) encompasses a spectrum that spans simple steatosis, through steatohepatitis (NASH) to fibrosis and ultimately cirrhosis. NAFLD is characterised by substantial inter-patient variation in rate of progression and disease outcome: whilst up to 25% of the general population are at risk of progressive disease, only a minority experience associated liver-related morbidity. Inter-patient genetic variation and environment determine severity and progression of NAFLD. This thesis reports a series of studies examining the association of genetic variations in two genes *patatin-like phospholipase domain-containing 3* (*PNPLA3,* rs738409 c.444 C>G, p.I148M) and *transmembrane 6 superfamily member 2*, (*TM6SF2*, rs58542926 c.449 C>T, p.E167K) with severity of NAFLD and risk of NAFLD-associated hepatocellular carcinoma (HCC).

Addressing first the role of *PNPLA3*, I demonstrate that the rs738409 variant is associated with steatosis, steatohepatitis and fibrosis in the largest histologically characterised NAFLD cohort of European-Caucasian descent (n=1,005) studied to date. Subsequently, adopting a case-control analyses in a cohort of 100 consecutive Northern European Caucasian patients with NAFLD-associated HCC arising and a cohort of patients with histologically characterised NAFLD, I demonstrate that carriage of the rs738409 minor (G) allele is significantly associated with increased risk of developing NAFLD-associated HCC, independent of potential confounding factors including gender, age at diagnosis, presence of advanced fibrosis/cirrhosis, T2DM and BMI.

During my studies, a genome-wide association study identified a SNP in *TM6SF2* as a modifier of hepatic triglyceride accumulation measured by MR Spectroscopy. It was therefore pertinent to determine whether this variant also affected risk of steatohepatitis or fibrosis in NAFLD. Using the aforementioned cohorts, I demonstrate for the first time that, in addition to its association with steatosis, the rs58542926 SNP is significantly associated with stage of fibrosis in NAFLD. In contrast to *PNPLA3* however, no association with NAFLD-HCC was found.

In conclusion, the current thesis confirms the association of *PNPLA3* with NAFLD severity and provides new evidence of its association with HCC risk. In addition, it

demonstrates for the first time that *TM6SF2* is associated with NAFLD-fibrosis severity. These studies provide important new insights into NAFLD pathogenesis and mandate further functional study.

Acknowledgements

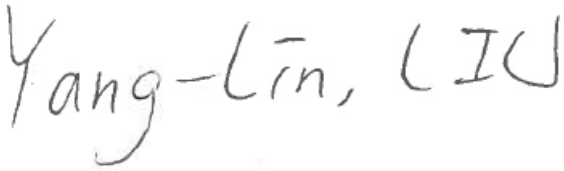
I would like to show my gratitude to my supervisors Professor Quentin M. Anstee, Professor Ann K. Daly, and Professor Christopher Day for giving me the opportunity to work in their lab and for their full guidance and support throughout this journey. I am grateful for the support and help of my colleagues: Julian Leathart, Julia Patch, Jeremy Palmer, Mohammad Alshabeeb, Tom Chamberlain, Salah Abohelaika, Salwani Bakar, Wipaporn Phatvej, Olivier Govaere, and Emma Scott.

I am also grateful for the support of my dearest friends through this journey: Yvonne Lai, Shirley Ho, Tawei Wang, Axilleas Floudas, Eirini Giannoudaki, Evie Mallini, Karolien Jordens, Michael Jin, and Chiao-En Peter Wu.

This thesis would not have been possible without the support of my supervisory team and most importantly, my beloved family. I thank my father and my sisters for their love, encouragement and many years of support. I owe my deepest gratitude to my late mother for her endless love, support, patience and understanding throughout this time. This work is dedicated to her.

Declaration of Originality

I hereby certify that the work described in this thesis is entirely my own, except where specifically stated otherwise.



Yang-Lin Liu

Table of Contents

Publications

Liu, Y.L., Patman, G.L., Leathart, J.B., Piguet, A.C., Burt, A.D., Dufour, J.F., Day, C.P., Daly, A.K., Reeves, H.L. and Anstee, Q.M. (2014) 'Carriage of the *PNPLA3* rs738409 C>G polymorphism confers an increased risk of non-alcoholic fatty liver disease associated hepatocellular carcinoma', *J Hepatol*, 61(1), pp. 75-81.

Liu, Y.L., Reeves, H.L., Burt, A.D., Tiniakos, D., McPherson, S., Leathart, J.B., Allison, M.E., Alexander, G.J., Piguet, A.C., Anty, R., Donaldson, P., Aithal, G.P., Francque, S., Van Gaal, L., Clement, K., Ratziu, V., Dufour, J.F., Day, C.P., Daly, A.K. and Anstee, Q.M. (2014) *'TM6SF2* rs58542926 influences hepatic fibrosis progression in patients with non-alcoholic fatty liver disease', *Nat Commun*, 5, p. 4309.

Liu, Y.L., Day, C.P. and Anstee, Q.M. (2016) 'Can genetic influence in non-alcoholic fatty liver disease be ignored?', *Clinical Dilemmas in Non-Alcoholic Fatty Liver Disease* (eds R. Williams and S. D. Taylor-Robinson).

Liu, W., Anstee, Q.M., Wang, X., Gawrieh, S., Gamazon, E.R., Athinarayanan, S., **Liu, Y.L.**, Darlay, R., Cordell, H.J., Daly, A.K., Day, C.P. and Chalasani, N. (2016) 'Transcriptional regulation of *PNPLA3* and its impact on susceptibility to nonalcoholic fatty liver disease (NAFLD) in humans', *Aging (Albany NY)*, 9(1), pp. 26-40.

Anty, R., **Liu, Y.L.**, Canivet, C.M., Iannelli, A., Patouraux, S., Ben-amor, I., Schneck, A.-S., Saint-Paul, M.C., Gugenheim, J., Day C.P., Daly A.K., Tran, A., Gual, P. and Q.M., A. (2016) 'The *FNDC5* rs3480 is protective on the steatosis and fibrosis in patients with NAFLD', submitted to publication.

Abbreviations

·OH	Hydroxyl radical
1H-MRS	Proton magnetic resonance spectroscopy
1O2	Singlet oxygen
ABCC	ATP-binding cassette, sub-family C (CFTR/MRP), member 2
ACC	Acetyl-coenzyme A carboxylase
AFP	Alpha-fetoprotein
AGTR1	Type-1 angiotensin II receptor
Akt	Protein kinase B
ALT	Alanine transaminase
AMPK	Adenosine monophosphate-activated protein kinase
ApoB	Apolipoprotein B
APOC3	Apolipoprotein C-III
APOE	Apolipoprotein E
AST	Aspartate transaminase
ATF6	Activating transcription factor 6
ATG16L1	Autophagy related 16 like 1
ATGL	Adipose triglyceride lipase
ATP	Adenosine triphosphate
AUROC	Area under the receiver operator characteristics curve
BMI	Body mass index
BSA	Bovine serum albumin
CAPS	3-(Cyclohexylamino)-1-propanesulfonic acid
CARD15	Known as NOD2, nucleotide binding oligomerization domain containing 2
CD14	Monocyte differentiation antigen CD14
CE	Cholesterol ester
ChREBP	Carbohydrate response element-binding protein
CHUK	Conserved helix-loop-helix ubiquitous kinase
CILP2	Cartilage intermediate layer protein 2
CKD	Chronic kidney disease
CPN1	Carboxypeptidase N subunit 1
CPT-I	Carnitine palmitoyl transferase I
CPT-II	Carnitine palmitoyl transferase II

CT	X-ray computed tomography
CVD	Cardiovascular disease
DAG	Diacylglycerol
DEN	Diethylnitrosamine
DEPC	Diethylpyrocarbonate
DM	Diabetes mellitus
DMEM F-12	Dulbecco's modified Eagle's medium-F12
DNL	de novo lipogenesis
ECM	Extracellular matrix
eQTL	Expression quantitative trait
ER	Endoplasmic reticulum
ERGIC	ER-Golgi intermediate compartments
ERLIN1	ER lipid raft associated 1
ETC	Electron transport chain
FADH2	Flavin adenine dinucleotide
FAS	Fatty acid synthase
FDFT1	Farnesyl diphosphate farnesyl transferase 1
FDR	False discovery rate
FFAs	Free fatty acids
FOXO1	Folkhead box protein O1
G6Pase	Glucose-6-phosphatase
GCKR	Glucokinase regulatory protein
GCL	Glutamate cysteine ligase
GCLC	Glutamate cysteine ligase, catalytic subunit
GCLM	Glutamate cysteine ligase, modifier subunit
GPx	Glutathione peroxidase
GS	Glycogen synthase
GSH	Glutathione
GSK3	Glycogen synthase kinase-3
GSSG	Glutathione disulphide
GULT	Glucose transporter
GWAS	Genome-wide association studies
GWAS	Genome-wide association studies
H2O2	Hydrogen peroxide

HapMap	Haplotype map
HbA1c	Haemoglobin A1c
HCC	Hepatocellular carcinoma
HCV	Hepatitis C virus
HFE	Hereditary hemochromatosis protein
HIV	Human immunodeficiency virus infection
HMG-CoA	Hydroxymethylglutaryl-CoA
HNE	4-hydroxy-2-nonenal
HOMA-IR	Homeostatic model assessment
HSCs	Hepatic stellate cells
HSD17B13	Hydroxysteroid 17-beta dehydrogenase 13
HSL	Hormone-sensitive lipase
HTGC	Hepatic triglyceride content
IL23R	Interleukin 23 receptor
IL-6	Interleukin-6
IL-8	Interlukine-8
IMT	Intima-media thickness
IR	Insulin resistance
IRE-1	Inositol-requiring enzyme-1
IRS	Insulin receptor substrates
JAK	Janus kinases
KCNJ11	Potassium voltage-gated channel subfamily J member 11
KLF6	Kruppel-like factor 6
LCFAs	Long chain fatty acids
LD	Linkage disequilibrium
LDL-C	Low-density lipoprotein-cholesterol
LPIN1	Phophatidate phosphatase LPIN1
L-PK	Liver-type pyruvate kinase
LXR	Liver X receptor
LXR-RXR	Liver X receptor–retinoid X receptor
LYPLAL1	Lysophospholipase-like protein 1
MAF	Minor allele frequency
MAPK10	Mitogen-activated protein kinase 10
MAT	Methionine adenosyl transferase

MDA	Malondialdehyde
MRI	Magnetic resonance imaging
mRNA	Messenger RNA
mtDNA	Mitochondrial DNA
mTOR	Mammalian/Mechanistic target of rapamycin
MTTP	Microsomal triglyceride transfer protein
NADH	Nicotinamide adenine dinucleotide
NAFLD	Non-alcoholic fatty liver disease
NAS	NAFLD activity score
NASH	Non-alcoholic steatohepatitis
NASH CRN	NASH Clinical Research Network
NCAN	Neurocan
NEAA	Non-essential amino acid
NECP: ATPIII	National Cholesterol Education Program's Adult Treatment Panel III
NEFAs	Nonesterifed fatty acids
NF-κB	Nuclear factor kappa-light-chain-enhancer of activated B-cells
NO·	Nitric oxide
NO2·	Nnitric dioxide
NR1I2	Nuclear receptor subfamily 1 group I member 2
O2-	Superoxide anion
OONO-	Peroxynitrite
OR	Odds ratio
OXPHOS	Oxidative phosphorylation
PAR	Population attributable risk
PBS	Phosphate-buffered saline
PBX4	PBX homeobox 4
PC	Pyruvate carboxylase
PEMT	Phosphatidylethanolamine N-methyltransferase
PEPCK	Phosphoenolpyruvate carboxykinase
PERK	Protein kinase R-like ER kinase
PI3K	Phosphoinositide-3 kinase
PKC	Protein kinase C
PKCε	Protein kinase C isoform in the liver
PNPLA3	Patatin-like phospholipase domain-containing 3

PPARG	Peroxisome proliferator activated receptor gamma
PPAR-α	Peroxisome-proliferator-activated receptor α
PPP1R3B	Protein phosphatase 1 regulatory subunit 3B
PUFAs	Polyunsaturated fatty acids
PVDF	Polyvinylidene fluoride
RFLP	Restriction fragment length polymorphism
RO·	Alkoxyl radical
ROO·	Peroxyl radical
SAF	Steatosis, activity and fibrosis
SAMe	S-adenosylmethionine
SAMM50	Sorting and assembly machinery component
SFAs	Saturated fatty acids
SLC27A5	Very long-chain acyl-CoA synthetase
SLC2A1	Solute carrier family 2, facilitated glucose transporter member 1
SNPs	Single nucleotide polymorphisms
SOD	Superoxide dismutase
SREBP-1c	Sterol regulatory element-binding protein-1c
STAT	Signal transducers and activators of transcription
SUGP1	SURP and G-patch domain containing 1
T2DM	Type 2 diabetes mellitus
TAE	Tris-acetate-EDTA
TC	Total cholesterol
TCF7L2	Transcription factor 7-like 2
TEMED	Tetramethylethylenediamine
TG	Triglycerides
TGFβ	Transforming growth factor beta
TLR4	Toll-like receptor 4
TM6SF2	Transmembrane 6 superfamily member 2
TNF	Tumour necrosis factor
TNFα	Tumour necrosis factor-alpha
TRIB1	Tribbles pseudokinase 1
UPCs	Uncoupling proteins
UPR	Unfolded protein response
VLDLs	Very low-density lipoproteins

WHO	World Health Organization

List of Figures

List of Tables

Chapter 1. Introduction

1.1 NAFLD

Non-alcoholic fatty liver disease (NAFLD) is the most common cause of liver dysfunction worldwide (Angulo, 2007; Yilmaz, 2012). It is estimated that one third of the population in Western countries is afflicted by NAFLD (Day, 2010), and the prevalence is increasing owing to its close association with metabolic syndrome (central obesity, insulin resistance/type 2 diabetes mellitus (T2DM), dyslipidaemia, hypertension and thus with cardiovascular disease (CVD)) (Sanyal and American Gastroenterological, 2002; Anstee *et al.*, 2013b). NAFLD represents a spectrum of liver disease that ranges from steatosis to more progressive forms of non-alcoholic steatohepatitis (NASH), fibrosis, cirrhosis, and ultimately hepatocellular carcinoma (HCC) and liver failure without a history of excess alcohol consumption (Anstee *et al.*, 2011b).

Observations on fatty degeneration in the liver was first described by Thomas Addison in 1836 (Addison, 1836). A “fatty liver hepatitis” was first described by Adler and Schaffner more than three decades ago, discovering that the histopathological patterns of non-alcoholic patients resembled those of alcoholics (Adler and Schaffner, 1979). However, it was Ludwig that originally coined the term non-alcoholic steatohepatitis ‘’NASH’’ one year later (Ludwig *et al.*, 1980) whilst describing a small group of 20 patients with fatty liver disease of unknown cause. The majority of these patients were obese and had comorbidities including hyperlipidaemia and diabetes. Their liver biopsy specimens were characterized by lobular hepatitis, focal necrosis with mixed inflammatory infiltrates and Mallory bodies, the most common histological feature of alcohol-induced liver disease (Ludwig *et al.*, 1980). In line with the resemblance of histopathology between the two diseases, several other terms as pseudo-alcoholic liver disease, alcohol-like hepatitis, diabetic hepatitis, non-alcoholic Laennec’s disease and steatonecrosis had been adopted to refer to this entity prior to the umbrella term NAFLD which was first introduced in 1986 (Schaffner and Thaler, 1986; Sheth *et al.*, 1997). Ever since then, both clinical and research interest in NAFLD has increased, with investigations encompassing heritability, diagnosis and natural history, the association with the metabolic syndrome, underlying mechanisms, and

developing possible pharmaceutical treatments (Dietrich and Hellerbrand, 2014).

1.2 Epidemiology

The prevalence of NAFLD has risen rapidly due to prevalent lifestyle habits of diets enriched in fat and carbohydrates and sedentary behaviour, which have changed profoundly over the past few decades (Angulo, 2007; Ratziu *et al.*, 2010). However, accurate estimates of NAFLD/NASH prevalence remain challenging as the numbers vary greatly by the population studied (different ethnicities, genders, and comorbidities), and the sensitivity of the diagnostic methods employed (radiology or histology). Nevertheless, studies performed so far provide some information on the global trend of the NAFLD epidemic. A recent meta-analysis that included twenty-one population-based studies (Europe: 6, East Asia: 7, South Asia: 5, and 3 in Middle East) has demonstrated that the pooled worldwide prevalence of NAFLD is 24.24% and was significantly correlated to economic status (Zhu *et al.*, 2015). When taking regions/ethnicity into account, the numbers changed slightly: Western countries have a higher incidence rate of NAFLD with 20-30% of the population affected (Day, 2010; Masarone *et al.*, 2014) compared to a rate of 5-18% in Asia (Masarone *et al.*, 2014). Using proton magnetic resonance spectroscopy (^{1}H-MRS) to assess hepatic steatosis in a multi-ethnic population from The Dallas Heart Study, 33.6% of 2349 US individuals were found to have NAFLD. Once again, an ethnicity effect was seen: the frequency of hepatic steatosis showed a distinct variation between Hispanics (45%), Caucasians (33%), and African Americans (17%) (Browning *et al.*, 2004b; Szczepaniak *et al.*, 2005). As liver biopsy remains the diagnostic gold standard to reliably assess the degree of severity of NAFLD, histological studies in apparently healthy, prospective living liver donors indicate that the prevalence of NAFLD was 12–18% in Europe (Browning *et al.*, 2004b; Nadalin *et al.*, 2005) and 27–38% in the USA (Ryan *et al.*, 2002; Browning *et al.*, 2004b; Tran *et al.*, 2006). Estimates increase further when populations with known risk factors are targeted. For instance, NAFLD was found in 91% of obese patients (body mass index (BMI) ≥30 kg/m^{2}), 67% of overweight

(>25-<30) and 25% of normal individuals (Bellentani *et al.*, 2004) reported by the European DIONYSOS study. Additionally, the overall prevalence of NAFLD is much greater in patients with T2DM with an incidence of 40-70% (Argo and Caldwell, 2009).

1.3 Diagnosis, staging and grading of NAFLD

1.3.1 Diagnosis

The most common symptoms reported by NAFLD patients are fatigue, malaise and an uncomfortable feeling or fullness in the right upper abdomen. Patients should have a history of excess alcohol intake (<20g/day for women; <30g/day for men) excluded together with alternative diagnoses including chronic viral hepatitis (hepatitis B and hepatitis C), autoimmune liver diseases, hereditary hemochromatosis, α1-antitrypsin deficiency, Wilson's disease and drug induced liver injury (Anstee *et al.*, 2011a). Obesity and insulin resistance or other features of the metabolic syndrome are the most common comorbidities found in NAFLD patients. Liver function tests in NAFLD patients are commonly within the normal range or exhibit only moderate elevations of aspartate transaminase (AST) and alanine transaminase (ALT). Generally, the AST/ALT ratio is less than one, but increases in the presence of fibrosis. Although hepatic steatosis can be demonstrated by ultrasound and other non-invasive radiological diagnostic approaches (computerised tomography (CT), proton magnetic resonance spectroscopy (^{1}H-MRS) and magnetic resonance imaging), a solid histopathological diagnosis of NAFLD (and the presence of NASH) can only be confirmed by the golden diagnostic method - liver biopsy which effectively documents disease stage and grade (Angulo, 2007; Wieckowska and Feldstein, 2008; Burt *et al.*, 2015).

1.3.1.1 *Steatosis*

The hallmark and initiation stage of NAFLD is hepatic steatosis, characterised by the accumulation of triglycerides (TG) affecting > 5% hepatocytes. Steatosis in NAFLD usually is presented as a mixture of macrovesicular (large) and microvesicular (small) droplet vacuoles within hepatocytes; true

microvesicular steatosis is rarely detected (Yeh and Brunt, 2014). Although hepatic steatosis is generally considered to be benign or self-limited, it may progress towards NASH in some patients. A widely accepted semiquantitative method for steatosis assessment is based on the percentage of hepatocytes involvement: four different degrees of steatosis are reported, comprising <5% steatosis, 5-33%, 34-66%, and over 66% (Kleiner *et al.*, 2005).

1.3.1.2 *NASH*

NASH, the more progressive form of NAFLD, is characterised by specific histological features in addition to steatosis. These include hepatocellular injury (hepatocellular ballooning degeneration) and lobular inflammation, with or without fibrosis. These key lesions usually occur in acinar zone 3, however there may be loss of acinar localization in the presence of progressive fibrosis and parenchymal remodelling (Yeh and Brunt, 2014). Hepatocellular ballooning is characterized by a swollen shape of hepatocytes, usually enlarged and lightly stained cytoplasm, in which Mallory-Denk Bodies are frequently found on routine histology with cellular diameter > 30 µm. The loss of cytoplasmic staining of keratins 8 and 18 by immunohistochemistry has been proposed to be the key feature in identification of ballooned hepatocytes (Guy *et al.*, 2012). Other helpful, but not necessary, morphological features for diagnosis of NASH include micrograulomas and acinar lipograulomas, megamitochondria, apoptotic bodies, and pericellular and perivenular fibrosis (Neuschwander-Tetri and Caldwell, 2003; Sanyal *et al*., 2011). Lobular inflammatory infiltrates are usually composed of lymphocytes (mainly T cells), eosinophils, macrophages, and less frequently neutrophils. Mild, chronic and mixed portal inflammation may also be present at this stage, and is suggested to be correlated with advanced severity of NASH and serological indicator of insulin resistance (IR), homeostatic model assessment (HOMA-IR) in both adult and paediatric cases (Brunt *et al.*, 2009; Harmon *et al.*, 2011; Smith, 2013). Notably, there is a blurred gap between steatosis and NASH depending on whether NASH is destined to occur after steatosis or both forms are discrete entities (Cohen *et al.*, 2011; Yilmaz, 2012). Around 10-29% of

NASH patients could progress to a more advanced stage of cirrhosis in a 10-year period (Argo and Caldwell, 2009).

1.3.1.3 *Fibrosis and cirrhosis*

Hepatic fibrosis is a result of repeated wound-healing in response to chronic liver injury, occurring in most types of chronic liver diseases (Albanis and Friedman, 2001; Bataller and Brenner, 2005). The predominant concept of liver fibrosis is the imbalance between increased generation and reduced degradation of extracellular matrix (ECM) proteins under a persistent liver injury. The excessive accumulation of ECM proteins results in the transformation of hepatic architecture by substituting hepatocytes for ECM, including fibrillar collagens. Collagens (I, III and IV), fibronectin, undulin, elastin, laminin, hyaluronan, and proteoglycans were found to increase 6 times more than the normal level in advanced fibrotic-stage (Bataller and Brenner, 2005). Hepatic stellate cells (HSCs) are the major ECM producing source in prolonged liver injury (Gabele *et al*., 2003). HSCs are located in the space of Disse (between hepatocytes and sinusoidal endothelial cells) and are responsible for vitamin A storage. HSCs are activated by inflammatory insults to the liver, undergoing a phenotype transformation into myofibroblast-like cells which lay down ECM. HSC maintaining methods (activated HSC model on plastic culture dishes) from rodent and human livers were studied and established in the 1980s to provide a useful tool for researchers to investigate the essential role of HSCs in liver fibrosis (Otto and Veech, 1980; Friedman *et al*., 1992).

Mild fibrosis (either in zone 3 perisinusoidal or portal) can be seen in NASH patients and it can further progress to bridging fibrosis and consequent cirrhosis if inflammatory insults are sustained (Brunt *et al*., 1999). Patients with progressive fibrosis over a period of 15-20 years could develop cirrhosis (Bataller and Brenner, 2005). Cirrhosis, the key risk factor for development of hepatocellular carcinoma, is encountered when normal hepatic lobules are replaced by fibrotic nodules thus causing disruption of the hepatic architecture and loss of liver function (Bircher, 1999).

1.3.1.4 *Hepatocellular carcinoma (HCC)*

HCC is the fifth most frequently diagnosed cancer worldwide and the third most common cause of cancer mortality (Baffy *et al*., 2012). The progression to HCC is most frequently associated with cirrhosis secondary to chronic liver infections such as hepatitis B or hepatitis C viruses or by alcoholic injury, while genetically defined diseases such as hemochromatosis are associated with a lesser proportion (Bosch *et al*., 2005; El-Serag and Rudolph, 2007). The tumour marker blood test alpha-fetoprotein (AFP) is often used to detect liver cancer (60-70% of HCC patients are found AFP positive). However, patients with elevated AFP should be screened by radiological examinations (abdominal ultrasound, X-ray computed tomography (CT), or magnetic resonance imaging (MRI)). If serological AFP together with radiological evidence is highly significant toward to a diagnosis of liver cancer, the liver biopsy is then not warranted. The histopathological hallmark of HCC is the resemblance both in its cytology and plate-like morphology, however, this aspect is beyond the scope of this chapter and has been summarized elsewhere (Paradis, 2013). HCC has been linked to NAFLD in a large number of epidemiological studies, illustrating several common NAFLD risk factors (overweight, obesity, insulin resistance, and diabetes mellitus) also associated with HCC morbidity (Baffy *et al*., 2012). Though the accurate prevalence of NAFLD-related HCC is unclear, it is suggested that 4 to 27% of cirrhotic NAFLD patients could develop HCC (Cohen *et al.*, 2011) while HCC is rarer in non-cirrhotic NAFLD patients (Baffy *et al.*, 2012).

1.3.2 Staging and grading of NAFLD

A semi-quantitative scoring system – the NAFLD Activity Score (NAS) developed by NASH Clinical Research Network (NASH CRN) Pathology Committee, sponsored by National Institute of Diabetes and Digestive and Kidney Diseases, is currently the most widely used approach for grading disease severity of NAFLD and stage of fibrosis in clinical trials (Kleiner *et al.*, 2005). It is a revision of the original NAFLD grading system proposed by Brunt et al., illustrating the NASH grades according to the combinations mixed by steatosis, ballooning and inflammation (Brunt *et al.*, 1999). The NAS is differ

from the original one as it is derived from the sum of the discrete scores of steatosis (0-3), hepatocellular ballooning (0-2) and lobular (acinar) inflammation (0-3), therefore provides a NAS range from 0 to 8. CRN also issued a 5-tier method for staging of fibrosis (0-4) (Table 1.1) (Kleiner *et al.*, 2005).

Although NAS has been well validated in current clinical and medical practice, its sensitivity and specificity for a threshold NAS ≥ 5 as a histological diagnosis of NASH was only 57% and 95%, respectively (Hjelkrem *et al.*, 2011). A further report on a study performed in a cohort of 976 individuals demonstrated that only 75% of the biopsies with 'definite NASH' reached this threshold whilst 28% with borderline NASH, and 7% of them without NASH had NAS ≥ 5 (Brunt *et al.*, 2011). Caution therefore should be taken when applying NAS for diagnostic purposes as the threshold value does not always reflect the actual disease severity. This discrepancy may be best explained by inter-observer-related variation among pathologists and the different histopathological criteria adopted (Machado *et al.*, 2006; Younossi *et al.*, 2011).

To address this, a simplified steatosis (S), activity (A) and fibrosis (F) (SAF) score has recently been designed by the European Fatty Liver Inhibition of Progression (FLIP) consortium, in an attempt to improve inter-observer variability (Bedossa and Consortium, 2014). The key difference in SAF scoring evaluation from NASH CRN score is that the degree of steatosis is separated from those of inflammation/ballooning to prove three separate measures of NAFLD severity. Basically, the FLIP algorithm encompasses semi-quantitative scoring of the key features: steatosis (0-3), activity (0-4) as the sum for hepatocellular ballooning and lobular inflammation, and fibrosis (0-4). Hence, NAFLD activity can be histologically divided into 2 categories; mild (A <2 and/or F <2) and significant (A >2 and/or F >2). In a study with 679 obese patients, the threshold of A >2 correctly identified all patients with NASH, and none of the patients below this threshold had NASH (Bedossa and Consortium, 2014). Comparison of the SAF Score and the NAFLD Kleiner Score for the histological grading and staging of NAFLD/NASH is shown in Table 1.1.

Table 1.1: Comparison of the FLIP SAF Score and the NAFLD CRN Score for the histological grading and staging of NAFLD/NASH.

SAF Score (Bedossa and Consortium, 2014)			Kleiner Score (Kleiner *et al.*, 2005)		
Histological Feature	**Category Definition**		**Histological Feature**	**Category Definition**	
Steatosis	0	<5%	**Steatosis**	0	<5%
	1	5-33%		1	5-33%
	2	34-66%		2	34-66%
	3	>66%		3	>66%
(S) Steatosis Score 0-3			***PLUS***		
Hepatocyte Ballooning	0	None	**Hepatocyte Ballooning**	0	None
	1	Clusters of hepatocytes with rounded shape and pale cytoplasm		1	Few
	2	Same as grade 1 with enlarged hepatocytes (>2x normal size)		2	Many
PLUS			***PLUS***		
Inflammation	0	None	**Inflammation**	0	None
	1	< 2 foci per 20x field		1	1–2 foci per x20 field
	2	> 2 foci per 20x field		2	2–4 foci per x20 field
				3	>4 foci per x20 field
(A) Total = Activity Score 0-4			**(NAS) Total = NAFLD Activity Score 0-8**		
Fibrosis	0	No fibrosis	**Fibrosis**	0	No fibrosis
	1a	Zone 3 mild perisinusoidal fibrosis		1a	Zone 3 mild perisinusoidal fibrosis
	1b	Zone 3 moderate perisinusoidal fibrosis		1b	Zone 3 moderate perisinusoidal fibrosis
	1c	Periportal/portal fibrosis only		1c	Periportal/portal fibrosis only
	2	Zone 3 plus portal/periportal fibrosis		2	Zone 3 plus portal/periportal fibrosis
	3	Bridging fibrosis		3	Bridging fibrosis
	4	Cirrhosis		4	Cirrhosis
(F) Fibrosis Stage 0-4			**Fibrosis Stage 0-4**		

Table modified from (Dyson *et al.*, 2014) with permission.

1.4 Natural history of NAFLD

A US-based long-term study with 23-year length revealed that patients with NAFLD had higher mortality compared to those in the general population, and the outcome of a lower survival rate was associated with age (hazard ratio per decade 2.2, 95%CI [1.7-2.7], *P*<0.0001), impaired fasting glucose (hazard ratio 2.6, 95%CI [1.3-5.2], *P*=0.005), and cirrhosis diagnosed at baseline (hazard ratio 3.1, 95%CI [1.2-7.8], *P*=0.02) (Adams *et al.*, 2005). Other studies on long-term mortality in NAFLD patients during 15 year follow-up showed 26% death rate, and this increased profoundly to 34-69% compared to the general population sharing the same age and gender (Angulo, 2013). In general, the majority of NAFLD patients exhibit only simple steatosis which is considered to be benign and self-limited, whilst a small fraction of individuals progress to NASH, fibrosis, cirrhosis and HCC. However, recent study using serial biopsy data (a median interval of 6.6 years) has challenged this dogma, showing that 44% of the NAFLD patients with 'pure fatty liver' graduated to NASH, and 37% had fibrosis progression (22% of them were at advanced stage) (McPherson *et al.*, 2015). Another systematic study of follow-up biopsies over a mean of 3.7 years documented that 64% of NAFLD patients progressed to NASH from steatosis, while 24% developed advanced fibrosis (Pais *et al.*, 2013). The summary of disease progression and natural history of NAFLD is displayed in Figure 1.1.

Poor disease prognosis and reduced survival is predicted once patients develop NASH. A Swedish study investigated the survival and cause of death within a cohort of 129 NAFLD patients with a mean follow-up period of 13.7 years; mortality was not increased in patients with steatosis, patients with NASH however exhibited a greater than 10-fold increased risk of liver-induced causes (2.8% versus 0.2%) and twofold higher risk from CVD-related death (15.5% versus 7.5%) compared with a matched reference population (Ekstedt *et al.*, 2006). Estimates of disease progression from patients with NASH to fibrosis vary between 27-53% by several studies using paired serial histological data within 3-6 years (Fassio *et al.*, 2004; Hui *et al.*, 2005; Wong *et al.*, 2010). Correspondingly, a recent meta-analysis pooled 11 studies with a total of 411 histologically confirmed NAFLD patients and with over 2145.5 person-years of follow-up evaluation; this study discovered that the annual fibrosis progression rate in patients with steatosis only at index biopsy was 0.07 stages (95%CI [0.02-0.11

stages]) and it was 0.14 stages in those with NASH (95%CI [0.07-0.21 stages]). Namely, patients with steatosis only progressed to fibrosis stage one on average in 14.3 years while the average was 7.1 years for those diagnosed with NASH at baseline (Singh *et al.*, 2015). Another study suggested that an overall annual rate of fibrosis progression from steatosis and NASH patients at baseline were 0.067 and 0.08 stages, respectively (McPherson *et al.*, 2015). Notably, although fibrosis progression is generally slow, rapid progression from simple steatosis or NASH (both at F0) to F3-4 does occur in a small set of patients over a mean follow-up period of 5.9 years (Singh *et al.*, 2015).

The increased risk of mortality in patients with NASH compared with those with steatosis could be explained by the greater incidence and the higher severity of fibrosis found in NASH patients as the presence of fibrosis is the key histological determinant of long-term prognosis (Younossi *et al.*, 2011; Chan *et al.*, 2014; Singh *et al.*, 2015). Supporting evidence from a study with 209 NAFLD patients over a median of 12.1 years demonstrated that NASH only correlated with liver-related mortality in the presence of advanced fibrosis; only F3 portal fibrosis was independently associated with liver mortality (hazard ratio 5.68, 95%CI [1.5-21.5]) when those with histological features of NASH were selected (Younossi *et al.*, 2011). Additional study also supports this assertion; a longitudinal study with a median follow-up period of 12.6 years recruited 619 NAFLD patients at medical centres in USA, Europe and Thailand, investigating the long-term prognostic relevance of histological features and analysing the overall mortality, liver transplantation and liver-related events as outcomes (Angulo *et al.*, 2015). The key finding was that histologic feature of fibrosis, but no other features of steatohepatitis, was independently associated with death, liver transplantation and liver-induced events included fibrosis stage 1 (hazard ratio 1.88, 95%CI [1.28-2.77]), stage 2 (hazard ratio 2.89, 95%CI [1.93-4.33]), stage 3 (hazard ratio 3.76, 95%CI [2.40-5.89]) and stage 4 (hazard ratio 10.9, 95%CI [6.06-19.62]).

Approximately 10-25% of NASH patients progress to extensive fibrosis and cirrhosis (McCullough, 2004; Onnerhag *et al.*, 2014; Goh and McCullough, 2016). The natural history of cirrhosis is characterised by two phases termed ‘compensated’ and ‘decompensated’ with the status of portal pressure as the watershed. The decompensated phase is subject to a series of liver complications once portal

hypertension develops, including varices, portal hypertensive gastrointestinal bleeding, ascites, jaundice, hepatic encephalopathy, and ultimately to HCC (D'Amico *et al.*, 2006). Accumulating data has provided some information on the mortality of cirrhosis due to NASH (Hui *et al.*, 2003; Sanyal *et al.*, 2006; Yatsuji *et al.*, 2009; Bhala *et al.*, 2011). An Australian study compared the rates of liver complications and survival between 23 patients with NASH-associated cirrhosis and those with hepatitis C virus (HCV)-related disease over 7 years follow-up; liver failure was the main cause of morbidity and mortality in NASH-related cirrhosis, but with a lower risk of HCC development despite a similar survival was seen in both groups (Hui *et al.*, 2003). Conversely, a separate study over 10-year period (152 NASH-cirrhosis patients vs. 150 HCV-cirrhosis) reached different conclusions; compensated cirrhosis due to NASH had a lower mortality rate and lower incidence for development of ascites, hyperbilirubinemia, and HCC against HCV-cirrhosis. However, patients with NASH-cirrhosis had a higher risk of CVD-related mortality (Sanyal *et al.*, 2006). Nonetheless, one Japanese study with 5-year follow-up using a relatively small cohort of 68 NASH cases demonstrated that similar rates of complications of cirrhosis (ascites, varices, hepatic encephalopathy and HCC) were found in both cohorts of NASH and HCV related cirrhosis (Yatsuji *et al.*, 2009). The discrepancy between the above studies may be explained by ethnic differences and variation in study design. Further information on the natural history of NAFLD/NASH-associated cirrhosis could be obtained from data on cryptogenic cirrhosis as there is growing recognition that NAFLD/NASH may be responsible for a great proportion of cryptogenic cirrhosis since metabolic syndrome is also the most common feature in those patients (Powell *et al.*, 1990; Bugianesi *et al.*, 2002). Moreover, as a frequent cause of cirrhosis, NASH is projected to be the leading indication for liver transplantation in the US by 2020 (Ratziu *et al.*, 2010; Sanyal *et al.*, 2011).

As progression to advanced fibrosis and cirrhosis in NASH patients has become a global concern, this also naturally extends to the risk for development of HCC (Adams *et al.*, 2005; Rafiq *et al.*, 2009; Satapathy and Sanyal, 2015). Available data suggests that the prevalence of HCC in NAFLD and NASH patients is estimated to be 0.5% and 2.8%, respectively (Starley *et al.*, 2010). A recent Surveillance, Epidemiology, and End Results (SEER) database study in United States demonstrated that individuals affected by NAFLD had 2.6-fold higher risk for HCC,

with a 9% annual increase over 6-year period (Younossi *et al.*, 2015). These estimates are forecast to be greatly increased in parallel to the NAFLD epidemic (Baffy *et al.*, 2012). In general, HCV and alcohol currently account for the majority of HCC underlying causes. However, a US-based study with 4,406 HCC patients recruited between 2002-2008 challenged this assertion, providing evidence that NAFLD/NASH was the most common aetiological factor among those cases (59%), followed by T2DM (36%) and HCV (22%) (Sanyal *et al.*, 2010). The presence of cirrhosis is the key risk factor for HCC and a surveillance study from Japan indicates that 80% of HCC patients are cirrhotic irrespective of aetiology (Hashimoto *et al.*, 2009). The cumulative incidence of HCC in patients with NASH-related cirrhosis has been suggested in one systemic review to range from 2.4% (over 7 years) to 12.8% (over 3 years) (White *et al.*, 2012), whilst another study reported the annually cumulative incidence of HCC was 2.6% in NASH-associated cirrhosis compared to 4% for those with HCV-related cirrhosis (Ascha *et al.*, 2010).

Yet, a growing number of case reports suggest that cirrhosis is not a necessary determinant for HCC, especially in NAFLD patients. In a US prospective study, the absence of cirrhosis was found to be 54% in all NAFLD-HCC cases while there was only 22% with absence in the HCV-cirrhosis group (Sanyal *et al.*, 2010). Another German study examined the prevalence of HCC with different aetiologies and revealed that 41.7% of NAFLD/NASH-related HCC patients were non-cirrhotic (Ertle *et al.*, 2011). In a study involving 1,500 U.S veterans over 6-year period, non-cirrhotic HCC was mainly due to NAFLD and patients with either NAFLD (unadjusted odds ratio 5.4; 95% CI [3.4-8.5]) or metabolic syndrome (unadjusted odds ratio, 5.0; 95% CI [3.1-7.8]) had a greater than 5-fold risk to present with HCC in the absence of cirrhosis, compared with the cohort of HCV-related HCC (Mittal *et al.*, 2016).

This specific characteristic (the absence of cirrhosis in progression of NAFLD-HCC) could be explained by the well-known NAFLD-associated risk factors, obesity and metabolic syndrome (Hardy *et al.*, 2016). A large American prospective study with more than 900,000 adults demonstrated that men with a BMI of 35 kg/m^2 or above had 4.5 times higher risk of dying from liver cancer while the risk was 1.68 higher in women when compared to individuals with normal BMI (Calle *et al.*, 2003). One meta-analysis also concluded that the summary relative risks for HCC were 1.17 for overweight and 1.89 for obese subjects (Larsson and Wolk, 2007). Substantial data

suggests that T2DM promotes the development and progression of HCC (El-Serag *et al.*, 2004; Davila *et al.*, 2005; El-Serag *et al.*, 2006; Welzel *et al.*, 2011). Supporting evidence from a systematic review of 26 studies published during the last decade, 13 case-control studies and another 13 cohort studies, shows that the presence of diabetes was associated with a relatively consistent 2.5-fold increase in HCC risk in different populations and geographic areas (El-Serag *et al.*, 2006). Since NAFLD is the major hepatic manifestation of obesity, impaired glucose tolerance and insulin resistance, T2DM and other associated metabolic conditions, this once again supports the notion that NAFLD/NASH would be the leading future cause of HCC and liver transplantation (Marrero *et al.*, 2002; Starley *et al.*, 2010).

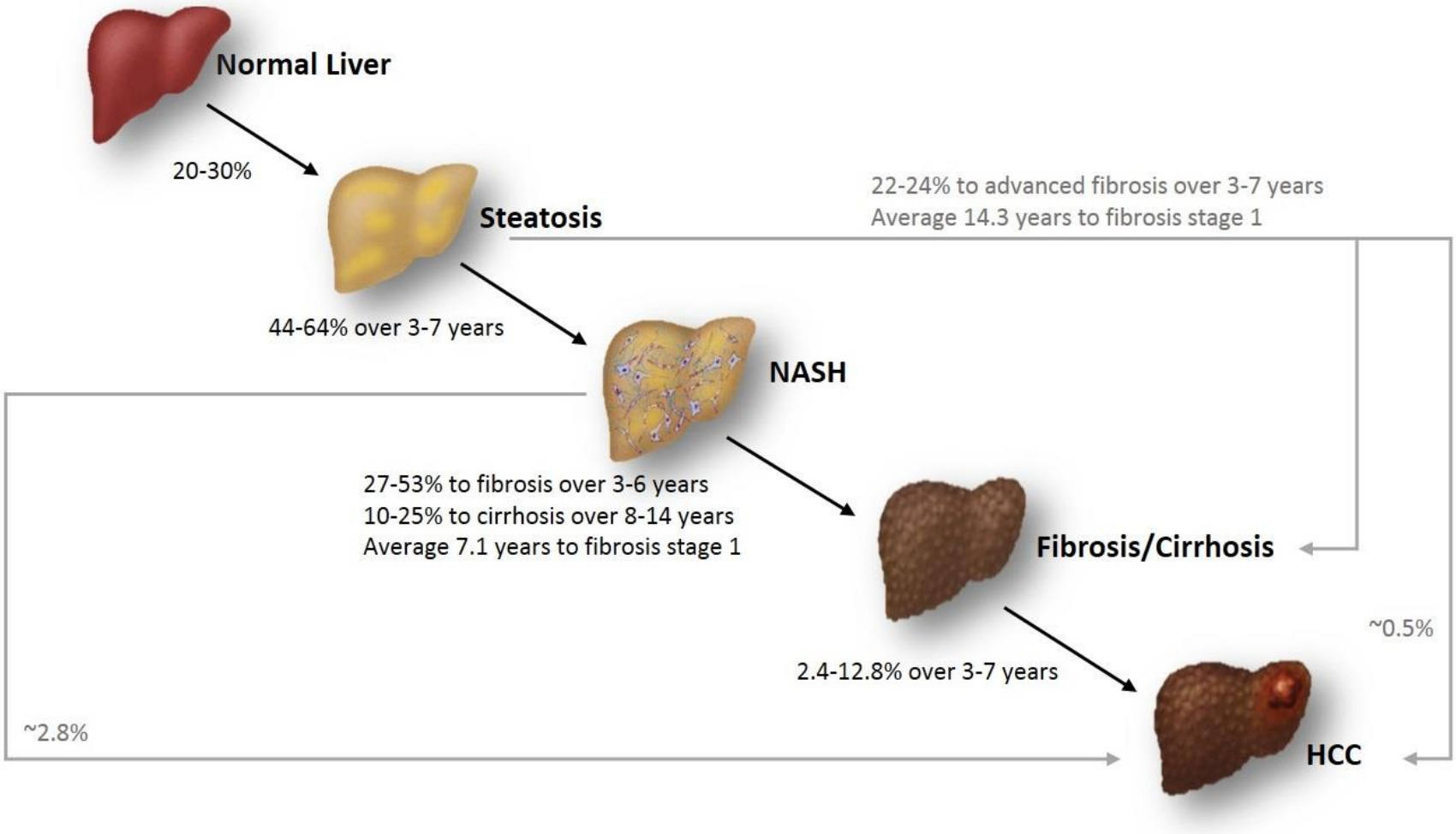


Figure 1.1: Summary of the progression and natural history of NAFLD. Adapted and modified from (Cohen *et al.*, 2011; Goh and McCullough, 2016).

1.5 Risk factors

NAFLD is best considered as a paradigm of complex genetic disease trait, since susceptibility appears to involve a combination of inter-patient genetic variation and environmental factors. With respect to environmental risk factors contributing to NAFLD, high calorific diet and sedentary life-style are well-known risk factors (Day, 2006). Older age has also been linked to NAFLD (Chen *et al.*, 2007a; Frith *et al.*, 2009). One Chinese case-control study investigated 4,226 adults aged above 60 years compared with 3,145 randomly selected younger controls and found that the prevalence of NAFLD was slightly higher in the elderly (26.7%) than in non-elderly (22.8%) (Wang *et al.*, 2013). It is noteworthy that whether age has its true effect on NAFLD/NASH or this is owing to a cumulative end results of other risk components (metabolic syndrome and fibrosis for example) and longer duration of disease progression in these patients (Vacca *et al.*, 2015). Available data investigating gender difference in NAFLD remains conflicting. Women accounted for 60%-83% of NAFLD/NASH diagnosis in several early clinical studies (Ludwig *et al.*, 1980; Powell *et al.*, 1990; Angulo *et al.*, 1999; Caldwell *et al.*, 1999; Matteoni *et al.*, 1999). In further studies on women, NAFLD has recently been reported to associate with late menopausal transition and postmenopausal stages (Ryu et al., 2015b), miscarriage and induced abortion (Liu *et al.*, 2013), but an inverse association was found between age at menarche and NAFLD in a Korean study involved 76,415 middle-aged women (Ryu *et al.*, 2015a). Although some of these findings suggest a particular female predisposition, several other reports suggest a male predominance in NAFLD (Bacon *et al.*, 1994; Sanyal and American Gastroenterological, 2002; Browning *et al.*, 2004b; Williams *et al.*, 2011).

It is globally acknowledged that NAFLD is strongly correlated with the metabolic syndrome, especially obesity and T2DM (Bian and Ma, 2012; Ortiz-Lopez *et al.*, 2012) and this section mainly focuses on this aspect. The metabolic syndrome, also known as the X syndrome, the insulin resistance syndrome, and the deadly quartet, was recognized at least eight decades ago (Cameron *et al.*, 2004; Eckel *et al.*, 2005). Although the definitions of metabolic syndrome slightly differ between the criteria established either by World Health Organization (WHO) (Alberti and Zimmet, 1998), European Group for the Study of Insulin Resistance (Balkau and Charles, 1999), and the National Cholesterol Education Program's Adult Treatment Panel III

(NECP: ATPIII) (Expert Panel on Detection and Treatment of High Blood Cholesterol in, 2001), the agreed essential elements are glucose intolerance, central obesity, dyslipidaemia and hypertension. The criterion of 'glucose intolerance' also involves impaired glucose tolerance, impaired fasting glycaemia, IR and T2DM (Eckel *et al.*, 2005). These key components of metabolic syndrome have been proposed as potential risk factors in developing diabetes, CVD, stroke, chronic kidney disease (CKD), and NAFLD (Eckel *et al.*, 2005; Anstee *et al.*, 2013b). A proportion greater than 90% of the NAFLD patients have at least one component of the metabolic syndrome (Marchesini et al., 2003). While environmental factors in NAFLD are well-established, the underlying mechanisms of genetic factors remain unclear. Current understanding on the genetic basis of NAFLD is described in detail in the next section 1.7.

1.5.1 NAFLD and obesity

Obesity is defined by BMI ≥ 30 kg/m^2 and BMI ≥ 40 kg/m^2 is defined as morbid obesity (Kubik *et al.*, 2013). NAFLD was first recognized as a clinical entity based on the studies discovering the presence of fatty liver, variable degrees of inflammation and fibrosis in morbidly obese individuals (Payne *et al.*, 1963; Kern *et al.*, 1973; Catlin, 1976). Since then, accumulating data confirmed that obesity is a major risk factor in the development of NAFLD (Nasrallah et al., 1981; Braillon et al., 1985; Angulo et al., 1999; Matteoni et al., 1999). As described in section 1.2, the prevalence of NAFLD increased greatly in obese patients (Bellentani *et al.*, 2004). One Italian study reported that 75.8% of obese patients had NAFLD compared to only 16% of those with normal BMI and without metabolic syndrome (Bellentani *et al.*, 2000). Overall, the prevalence of NAFLD is estimated at between 74-91% in obese subjects (Angulo, 2002; Abrams *et al.*, 2004; Bellentani *et al.*, 2004) while NASH is thought to be present in 25-30% of an obese population and >35% when coexisting with T2DM (Silverman *et al.*, 1989; Musso *et al.*, 2011; Smith and Adams, 2011). Visceral adiposity (central obesity), not overall obesity, is significantly associated with NASH (Kral *et al.*, 1993; Omagari *et al.*, 2002; Thomas *et al.*, 2005; Farrell and Larter, 2006), especially for 'non-obese' NAFLD (Ha *et al.*, 2015). On the other hand, bariatric surgery does ameliorate disease severity in NAFLD (Angulo, 2006; Mummadi *et al.*, 2008). One meta-analysis pooled 15 studies which

investigated the effect of bariatric surgery on NAFLD and revealed that 91.6% of patients had improvement or resolution in steatosis (95%CI [82.4%-97.6%]), 81.3% in NASH (95%CI [61.9%-94.9%]), and 65.5% in fibrosis (95%CI [38.2%-88.1%]); the completely disappearance of NASH was found in 69.5% (95%CI [42.2%-90.8%]) (Mummadi *et al.*, 2008). A recent study with 1-year of follow-up also showed that nearly 85% of patients had reversal of NASH and reduced pathologic features of the disease after bariatric surgery (Lassailly *et al.*, 2015).

1.5.2 NAFLD and type 2 diabetes mellitus

NAFLD is closely associated with IR and T2DM, it however remains unclear whether it is NAFLD that induces T2DM or *vice versa*. NAFLD patients have an average of 2 fold increased risk of incident T2DM as reported in numerous studies (Fan *et al.*, 2007; Shibata *et al.*, 2007; Fraser *et al.*, 2009; Yamada *et al.*, 2010; Sung *et al.*, 2012). These studies however performed a NAFLD diagnosis by ultrasonography or serum levels of liver enzymes which has limitations due to intraobserver variability and the fact that 80-85% of patients with T2DM and NAFLD are negative for increased serum levels of liver enzymes (Targher *et al.*, 2007a; Williams *et al.*, 2011; Williamson *et al.*, 2011). Only one available data has so far demonstrated the association between histologically diagnosed NAFLD and the risk of incident T2DM. In a cohort of 129 histologically-proven NAFLD patients who were followed up for a mean period of 13.7 years, Ekstedt *et al.* showed that the prevalence of previously known T2DM was 8.5% at baseline, but a striking proportion of 78% of these individuals developed either T2DM (58%) or impaired glucose tolerance at the end of the study (Ekstedt *et al.*, 2006).

When considering NAFLD in patients with established T2DM, the prevalence of NAFLD is thought to be over 70% (Targher *et al.*, 2007a; Leite *et al.*, 2009). Both a personal and family history of IR and diabetes mellitus (DM) increases predisposition to NASH and fibrosis (Loomba *et al.*, 2012). One study reinforced the notion that NAFLD is an independent risk factor for T2DM after adjustment for other elements of metabolic syndrome (Musso *et al.*, 2011). In addition, patients with both T2DM and NAFLD often have poorer glycaemic control compared to those with T2DM only (Jimba *et al.*, 2005; Williamson *et al.*, 2011). The presence of T2DM itself is a strong

risk factor for liver disease as its long-term prognostic significance includes NASH, cirrhosis and HCC (de Marco *et al.*, 1999; Bugianesi *et al.*, 2002; El-Serag *et al.*, 2004; Bugianesi, 2007; Anstee *et al.*, 2013b).

1.5.3 NAFLD, cardiovascular disease and chronic kidney disease

The recognition that NAFLD may stimulate the progression and development of CVD has already been a heated topic as common features of NAFLD (i.e. hypertension, atherogenic dyslipidaemia, dysglycaemia, IR and obesity) are risk factors for CVD (Marchesini *et al.*, 2003). Numerous reports showed that NAFLD has been linked to markers of subclinical atherosclerosis (i.e. impaired flow-mediated vasodilation) that are dependent on endothelium, increased arterial stiffness, and increased carotid artery intima-media thickness (IMT) (Anstee *et al.*, 2013b; Oni *et al.*, 2013). One meta-analysis with seven cross-sectional studies (involving a total of 3,497 subjects) confirmed that NAFLD patients diagnosed with ultrasonography had a significantly increased prevalence of carotid artery IMT and carotid plaques (Sookoian and Pirola, 2011). In a recent prospective study of 465 consecutive patients with ischaemic heart disease diagnosed by coronary angiography, the prevalence of fatty liver was 64.7% and fatty liver was associated with more severe coronary artery disease (Wong *et al.*, 2011). The intimate correlation between CVD and NAFLD could be further demonstrated by several retrospective studies (NAFLD diagnostic approach as imaging or biopsy) examining the natural history of NAFLD patients within a reasonably long duration; CVD is a very common cause of death among these patients (Adams *et al.*, 2005; Ekstedt *et al.*, 2006; Targher *et al.*, 2007b; Rafiq *et al.*, 2009; Soderberg *et al.*, 2010; Treeprasertsuk *et al.*, 2012; Zhou *et al.*, 2012). Two of the studies with biopsy-proven NAFLD also reported that patients with NASH, not those with simple steatosis, possess an increased risk of death from CVD compared with the reference population (Ekstedt *et al.*, 2006; Soderberg *et al.*, 2010).

There is growing awareness regarding the interaction between NAFLD and CKD as it is a microvascular diabetic complication in people with T2DM which is also a common feature in NAFLD (Targher *et al.*, 2011). The presence of CKD is usually defined as estimated glomerular filtration rate of <60 mL/min/1.73 m^2 and/or microalbuminuria and/or overt proteinuria. In a study with 2,103 Type 2 diabetic

patients, those who coexisted with NAFLD (ultrasonography) had almost 2-fold increased prevalence of CKD (OR 1.87, 95CI% [1.3-4.1]) and proliferative/laser-treated retinopathy (OR 1.75, 95CI% [1.1-3.7]) than in those without NAFLD (after adjustment for age, gender, BMI, waist circumference, blood pressure level, hypertension, duration of T2DM, haemoglobin A_{1c} (HbA_{1c}), lipids, smoking status and medications use) (Targher *et al.*, 2008). An increased incidence of CKD in NAFLD patients also has been reported in several other cross-sectional studies (Hwang *et al.*, 2010; Arase *et al.*, 2011; Yasui *et al.*, 2011).

Although there is accumulating data concerning the close correlation between NAFLD and CVD/CKD, study bias should be noted; factors such as ethnicity, population studied (hospital-based or community-based), an adequate adjustment with potential confounders (i.e. IR, obesity, T2DM, hypertension and other traditional and non-traditional risk factors for CVD) and the diagnostic method used is needed. Whether NAFLD is simply a risk marker which coexists in patients of these two disease entities or is an independent risk factors remains to be fully elucidated. However, data to date is sufficient to consider that NAFLD patients are at high-risk for CVD/CKD events. More comprehensive information regarding current knowledge on the relationship between NAFLD and CVD/CKD can be found in two other review articles (Anstee *et al.*, 2013b; Vanni *et al.*, 2015).

1.6 Pathogenesis

The preeminent hypothesis for NAFLD pathogenesis was proposed by Day and James in 1998, suggesting that NAFLD might be driven in a “two-hit” fashion. The initial stage of NAFLD, steatosis, begins with an imbalance between lipid acquisition and removal in the liver, followed by steatohepatitis resulted from a complex system of inflammatory cascade which promotes lipotoxicity, oxidative stress and further induction of subsequent stages of NAFLD (Day and James, 1998). However, accumulating evidence during the past decade suggests that NAFLD may be contributed from “multiple hits”, a combination of several biochemical and immunological effects, rather than a simple “two-hit” manner. It is generally accepted that initiation of NAFLD is dependent on development of obesity and IR/T2DM, therefore this section mainly focuses on involvement of lipotoxicity, damage caused

by overwhelming oxidative stress, impaired metabolic homeostasis, cellular dysfunction (Hardy *et al.*, 2016) and how these effects are governed by insulin and nutrient-sensing transcription factors.

1.6.1 Lipid metabolism in NAFLD

The liver plays a crucial role in the metabolism of carbohydrate, lipid and protein. Hepatic dysfunction usually is associated with systemic metabolic imbalances, and it is universally agreed that an imbalance in the acquisition, delivery and removal of long chain fatty acids (LCFAs) and triglycerides (TGs) is fundamental to NAFLD.

1.6.1.1 *Free fatty acids supply*

Free fatty acids (FFAs) are the key elements forming TGs, which can be further stored in adipocytes or be hydrolysed into three FFAs when the body requires higher energy demand. Liver FFAs are sourced from dietary TGs intake, enhanced *de novo* lipogenesis (DNL), and excess fatty acids influx via adipose tissue lipolysis (Vacca *et al.*, 2015). In NAFLD cases, it has been reported that 59% of TG in the liver is derived from circulating FFAs, also known as nonesterifed fatty acids (NEFAs), 26% from DNL, and the remaining 15% from the diet (Donnelly *et al.*, 2005). Acetyl-coenzyme A carboxylase (ACC) and fatty acid synthase (FAS) are the predominant enzymes that catalyse FFA synthesis in the liver, while acetyl-coenzyme A (acetyl-CoA) and malonyl-coenzyme A (malonyl-CoA) are the essential metabolic intermediates in DNL (Mendez-Sanchez *et al.*, 2007) since increased DNL is found in NAFLD patients (Adams *et al.*, 2005). DNL is strongly regulated by nutritional level of insulin and glucose, via two transcription factors - carbohydrate response element-binding protein (ChREBP) and sterol regulatory element-binding protein-1c (SREBP-1c), respectively. SREBP-1c, a membrane bound transcription factor to the endoplasmic reticulum (ER) (Horton *et al.*, 2002), is regulated by insulin, saturated fatty acids (SFAs) and a nuclear receptor, liver X receptor (LXR), to promote fatty acid synthesis under hyperglycaemia. Conversely, glucagon, polyunsaturated fatty acids (PUFAs) inhibit its expression (Vacca *et al.*, 2015). Mice with a SREBP-1c deletion showed remarkable reduction of insulin-mediated lipogenic gene expression (Shimano *et al.*, 1999). Similar results were reported in LXRα-null mice with

decreased SREBP-1c and reduced lipogenesis (Repa *et al.*, 2000). In hyperglycaemia, excess carbohydrates are converted to FFAs/TGs in the liver via glycolytic enzymes such as glucokinase, liver-type pyruvate kinase (L-PK), and the previously mentioned enzymes of DNL - ACC and FAS (Towle *et al.*, 1997). ChREBP is expressed abundantly in lipogenic organs: liver, brown and white adipose tissue, small intestine, kidney, and muscle (Yamashita *et al.*, 2001). Genes that are regulated by ChREBP are involved in glycolysis (L-PK), the NADPH supply system (e.g. glucose-6-phosphate dehydrogenase, transketolase, and malic enzyme), gluconeogenesis, and lipogenesis (ACC, FAS) (Iizuka *et al.*, 2004; Ma *et al.*, 2006). Inhibition of ChREBP *in vivo* resulted in suppression of lipogenic genes/enzymes expression in the liver, and amelioration of hepatic steatosis, insulin signalling and glucose intolerance (Iizuka *et al.*, 2004; Dentin *et al.*, 2006). Taken together, SREBP-1c and ChREBP are activated by insulin and glucose during hyperglycaemia, inducing glycolysis (to release more acetyl-Co A as substrate for FFA synthesis) and lipogenesis, thus further worsening steatosis.

1.6.1.2 *Free fatty acids removal*

Clearance of hepatic TG is accomplished by the only disposal pathway – through the formation of very low-density lipoproteins (VLDLs) which can be further secreted from the liver. Apolipoprotein B-100 (apoB-100) is required for VLDL assembly, and microsomal triglyceride transfer protein (MTTP) is required for incorporating TG. In general, TG, cholesterol ester (CE) and phospholipid are first transferred to rough ER to be incorporated with newly formed apo B, known as the formation of the primordial particle. Further in the lumen of the rough ER, TG-enriched globules produced by the smooth ER are incorporated into the primordial particle to form the mature VLDL. Once VLDL is assembled, it is then transported to the Golgi and released from the cell as secretory vesicles. This secretory pathway is entirely assisted by apo B-100 (Mason, 1998). The secretion of VLDL is inhibited by SREBP-1c through decreasing expression of MTTP (Sato *et al.*, 1999).

1.6.1.3 *Free fatty acids oxidation*

FFAs can be utilised for energy production, the most efficient mechanism during energy deficiency. FFAs oxidation may be performed in three unique locations: 1) β-oxidation in the mitochondria, 2) in the peroxisomes, and 3) ω-oxidation in the ER of hepatocytes (Koek *et al.*, 2011). Under normal physiological conditions, β-oxidation of FFA in the mitochondria is the main source of short, medium, and long chain fatty acids (LCFAs) (Anstee and Goldin, 2006). Oxidation of FFA yields acetyl-CoA which is an important substrate for entering the citric acid cycle to generate electrons. These electrons further pass to oxygen via mitochondrial 'respiratory chain' (electron transport chain, ETC) and providing the energy for adenosine triphosphate (ATP) synthesis by oxidative phosphorylation (OXPHOS) (Lehninger *et al.*, 2000). Short- and medium-chain FFAs (below 12 carbons length) can simply diffuse across the mitochondrial membrane for β-oxidation. However, LCFAs (chain length with 14 or more carbons) are activated by acyl-CoA-synthetase in the cytosol as acyl-CoA, which then be catalyzed and be transported into the mitochondrial matrix in 3 steps (Serviddio *et al.*, 2011):

1. Acyl-CoA is catalyzed by carnitine palmitoyl transferase I (CPT-I) as fatty acyl-carnitine at the outer mitochondrial membrane, then passing through the intermembrane space.
2. The inner membrane delivery is accomplished by acyl-carnitine translocase.
3. At the inner face of the inner mitochondrial membrane, the reconversion of acyl-CoA is catalyzed and released by carnitine palmitoyl transferase II (CPT-II) into the matrix along with free carnitine.

Notably, CPT-I can be inhibited by malonyl-CoA, the key intermediate of DNL, and insulin. In general, hepatic lipogenesis is activated with carbohydrate feeding, resulting in elevated levels of insulin and malonyl-CoA; the expression of CPT-I is therefore suppressed and LCFAs are not oxidised but instead esterified, principally into TG (McGarry and Brown, 1997).

1.6.2 Oxidative stress

Oxidative stress, the essential underlying mechanisms of NAFLD, occurs when reactive oxygen species (ROS) exceed the production of protective antioxidants. Oxidation of FFAs is the primary source of ROS. ROS are constituted by oxygen-

centered radicals and oxygen-centered nonradicals. Superoxide anion (O_2^-), hydroxyl radical (·OH), alkoxyl radical (RO·), and peroxyl radical (ROO·) comprised the former group, while the latter one includes hydrogen peroxide (H_2O_2) and singlet oxygen ($^1O_2$). Other reactive species are nitrogen species such as nitric oxide (NO·), nitric dioxide (NO_2·), and peroxynitrite ($OONO^-$) (Halliwell *et al.*, 1995; Simon *et al.*, 2000). The nature of oxygen is readily to accept electrons, forming oxygen radicals which react with proteins, FFAs and DNA causing biological and physiological damage. The hydroxyl radical, nitric oxide radical and superoxide anion are the most dangerous ones among ROS group. Although the dominant source of cellular ROS is β-oxidation in the mitochondria, alternative pathways of β-oxidation in the peroxisomes and ω-oxidation in the ER are also ROS donors while oxidative capacity in mitochondria becomes overwhelmed in the event of FFAs overloading (Reddy, 2001). Upregulation in β-oxidation of LCFAs within peroxisomes and ω-oxidation in the ER has been reported in NASH (Robertson *et al.*, 2001). This could also be explained by an accumulation of malonyl-CoA in hyperglycaemia causing reduced mitochondrial β-oxidation due to inhibition of CPT-I (Donnelly *et al.*, 2005; Dentin *et al.*, 2006).

Once ROS are elevated, peroxisome-proliferator-activated receptor α (PPAR-α) is then activated for compensation. PPAR-α is expressed predominantly in the liver, and it belongs to the nuclear receptor subfamily of ligand-activated transcription factors, the key regulator of long chain and PUFAs catabolism (Reddy, 2001). Although the major function of PPAR-α is to increase FFA utilization during fasting for gluconeogenesis (Le May *et al.*, 2000; Atherton *et al.*, 2009), it is one of the key regulators that drives an adaptive response to FFA overloading. PPAR-α is involved in several aspects: modulation of FFA uptake, lipid trafficking, promotion of β-oxidation, and inhibition of DNL (Vacca *et al.*, 2015). PPAR-α regulates proteins involved in mitochondrial, peroxisomal and microsomal oxidation with the latter two preferentially induced by PPAR-α (Anstee and Goldin, 2006).

1.6.2.1 *FFA oxidation in mitochondria, peroxisomes and microsomes*

Two major physiological and chemical events occur in the mitochondria: fat metabolism and energy production. Each β-oxidation generates reducing equivalents

of nicotinamide adenine dinucleotide (NADH, reduced form) and flavin adenine dinucleotide ($FADH_2$), which serve as substrate and donate their electrons to the ETC. The ETC, a collection of mitochondrial membrane protein complexes embedded in the inner membrane (Pessayre *et al.*, 2001), is the main source of ROS and the dominant site of ATP synthesis in eukaryotic organelles. Under physiological conditions, electron transport throughout the respiratory chain matches with the pumping of protons (from matrix to intermembrane space) via Complexes I, III and IV. A proton gradient is then generated during this process, creating an electrochemical potential difference across the inner membrane. By this potential force, ATP is ultimately synthesized at Complex V (ATP synthase) (Pessayre *et al.*, 2001; Serviddio *et al.*, 2011).

Normally, a large fraction of electrons participating in ETC is safely dissipated via cytochrome c oxidase with oxygen and water (Pessayre, 2007). Two regions of ETC, Complex I and III, however are the major sources of ROS by production of the superoxide anion. Estimates of the fraction of oxygen that is diverted into ROS production are between 0.15 to 4% (Brand, 2010; Koek *et al.*, 2011; Serviddio *et al.*, 2011).

A rodent study revealed that 50% of FFA β –oxidation takes place in the peroxisomes (Latruffe *et al.*, 2000). However, peroxisomal β-oxidation only accounts for a minority of total FFA oxidation in human body under normal status. Peroxisomal proliferation is only augmented when the capacity of oxidation in the mitochondria is overwhelmed. Hydrogen peroxide, generated during peroxisomal β-oxidation, can be easily transformed into reactive hydroxyl radicals, thus contributing more ROS production. As previously mentioned, peroxisomal β-oxidation is also regulated by PPAR-α as it encodes acetyl-CoA oxidase, the rate-limiting enzyme for the oxidation (Benzie, 1996). The microsomal ω-oxidation is mediated by cytochrome P450 4A (CYP4A) and cytochrome P450 2A1 (CYP2E1). Both P450s produce ROS during ω-oxidation, further exacerbating oxidative stress within hepatocytes (Anstee and Goldin, 2006; Koek *et al.*, 2011).

1.6.2.2 *Antioxidants*

Human body develops a defence system of anti-oxidants to compensate oxidative stress. This protective system is performed by scavenging free radicals and reactive metabolites. The major enzymatic anti-oxidants include superoxide dismutase (SOD), catalase, and glutathione peroxidase (GPx). There are three different forms and localizations of SOD: 1) SOD1: copper/zinc-containing SOD, localized to cytoplasm (Cu/Zn-SOD); 2) SOD2: manganese-containing SOD, located in mitochondrial matrix (MnSOD), 3) SOD3: presented in extracellular space (EC-SOD). SOD mediates the conversion of superoxide anion radicals into hydrogen peroxide. This is followed by the detoxification of hydrogen peroxide into H_2O which is accomplished by catalase and GPx (Pessayre *et al.*, 2001). SOD2 has been suggested to be associated with NAFLD due to its mitochondrial localization; NASH-related fibrotic patients show an increased frequency of a common non-synonymous polymorphism in *SOD2* (C47T; rs4880) which is associated with a reduction in SOD2 mitochondrial targeting and activity (Al-Serri *et al.*, 2012).

Glutathione (GSH), S-adenosylmethionine (SAMe), and vitamins are major sources of non-enzymatic anti-oxidants. GSH is the dominant anti-oxidant in the liver with two forms: thiol-induced glutathione (GSH) and oxidized state of glutathione (glutathione disulphide, GSSG) (Chen *et al.*, 2007b; Koek *et al.*, 2011). The rate of GSH synthesis is mediated by cysteine, with assistance of two subunits of glutamate cysteine ligase (GCL): a catalytic form (GCLC) and a modifier (GCLM) (Meister and Anderson, 1983). *GCLC* knock-out mice resulted in depletion of GSH, leading to promotion of steatosis, mitochondrial and hepatic injuries has been reported (Chen *et al.*, 2007b).

Deficiency of GSH could be due to SAMe shortage as it is a precursor for GSH synthesis. SAMe is synthesized from methionine and catalysed by methionine adenosyl transferase (MAT) (Mato *et al.*, 2002). MAT is encoded by *MAT1A* and *MAT2A*. Deletion of *MAT1A* leads to liver injury due to a profound reduction of SAMe expression (Lu *et al.*, 2001). Additionally, *SAMe* knock-out mice had a tendency of developing hepatic steatosis and NASH (Wortham *et al.*, 2008). Other anti-oxidants such as vitamin C and E are also major free radical scavengers (Valko *et al.*, 2004).

1.6.2.3 *Multiple insults driven by ROS*

As previously described, excess accumulation of fat eventually lead to ROS overproduction. Although ROS have a short half-life, they are able to oxidise nuclear and mitochondrial DNA (mtDNA), causing mtDNA lesions and somatic point mutations (Pessayre, 2007). A comprehensive pathway of cellular injury can therefore be activated by the liberation of pro-inflammatory, necro-inflammatory or pro-fibrotic cytokines due to nuclear and mtDNA damage, phospholipid membrane disruption by lipid peroxidation, and mitochondrial dysfunction (Browning and Horton, 2004). Impaired mtDNA-encoded polypeptides (e.g. cytochrome c oxidase) involved in ETC can be affected by mtDNA lesions further aggravating ROS effects by blocking electron flow along through the ETC and disrupting hepatic ATP homeostasis (Pessayre, 2007). Reduced ATP synthesis due to impaired ETC has been demonstrated in both rodent models and NASH patients (Cortez-Pinto *et al.*, 1999; Serviddio *et al.*, 2008). Lipid peroxidation of PUFAs generates toxic intermediates of aldehyde such as 4-hydroxy-2-nonenal (HNE) and malondialdehyde (MDA). These by-products possess longer half-lives than ROS and so can damage more distant intracellular targets, causing cell death (Levene and Goldin, 2012).

ROS is a well-known inducer of nuclear factor kappa-light-chain-enhancer of activated B-cells (NF-κB), a pro-inflammatory signalling pathway which can also be activated by tumour necrosis factor-alpha (TNFα) (Serviddio *et al.*, 2011). NF-κB promotes survival of hepatic myofibroblasts and hepatic fibrogenesis (Oakley *et al.*, 2009). TNFα is secreted by enlarged adipocytes (Kern *et al.*, 1995) and fat-laden hepatocytes (Crespo *et al.*, 2001; Feldstein *et al.*, 2004), triggering permeability of the mitochondrial membranes by the reaction with its receptor on hepatocytes. Upregulation of cascade-8, Bid and Bax are mediated by TNFα, triggering hepatocyte apoptosis by releasing cytochrome C (Pessayre and Fromenty, 2005). Other key inflammatory mediators such as transforming growth factor beta (TGFβ) and interleukin-8 (IL-8) are also activated by oxidative stress, worsening hepatic inflammation and promoting fibrogenesis by activation of HSCs (Bataller and Brenner, 2005; Levene and Goldin, 2012).

When ER stress is overwhelmed due to the presence of excess FFAs, misfolded or unfolded proteins (also known as uncoupling proteins, UPCs) may ensue as ER is an intracellular membranous network where the majority of secreted and membrane

proteins are folded. An activation called the unfolded protein response (UPR) is then triggered by the aggregation of UPCs. This response helps re-establish normal homeostasis within the cell by triggering cell cycle arrest, folding catalysts, induction of ER-localized chaperone proteins, and ER-associated protein degradation. The UPR can lead to the induction of autophagy if required (Yang *et al.*, 2010). Apoptosis can also be triggered if this response fails by activation of stress-sensor proteins which include activating transcription factor 6 (ATF6), inositol-requiring enzyme-1 (IRE-1), and protein kinase R-like ER kinase (PERK) (Puri *et al.*, 2008; Tilg and Moschen, 2010).

1.6.3 Lipotoxicity and insulin resistance in NAFLD

Lipotoxicity is induced by fat overloading. The phenomenon - 'adipose tissue remodeling' refers to a phenotypic alteration of adipocytes to compensate for excess nutrition and fat within the cell; adipocyte enlargement results in the release of several cytokines. These cytokines are principally adiponectin, leptin, TNFα, and Interleukin-6 (IL-6). Although adiponectin has anti-inflammation, anti-fibrosis, anti-angiogenesis, sensitization of insulin, and tumor-growth limiting properties (Dalamaga *et al.*, 2012; Karagozian *et al.*, 2014), its availability is diminished in obesity, IR, and NAFLD (Pajvani *et al.*, 2003; Targher *et al.*, 2004; Pagano *et al.*, 2005; Levene and Goldin, 2012). Adiponectin mediates FFA and glucose metabolism by activating adenosine monophosphate-activated protein kinase (AMPK), a metabolic regulator. AMPK was first identified as a kinase which phosphorylates and inhibits ACC (rate-limiting enzyme for DNL) and hydroxymethylglutaryl-CoA synthase (HMG-CoA) reductase (rate-limiting enzyme for DNL cholesterol synthesis). Overall, activation of AMPK stimulates FFA oxidation, inhibits FFA synthesis and esterification, enhances insulin sensitivity, stimulates glucose uptake in muscle, negatively regulates hepatic glucose production, and diminishes pro-inflammatory activity (Luo *et al.*, 2010). AMPK has been suggested to inhibit lipogenesis by negatively regulating SREBP-1c (Zhou *et al.*, 2001). Adiponectin promotes lipogenesis not only through activation of AMPK, but also via PPAR-γ (Anstee and Goldin, 2006; Levene and Goldin, 2012). Hepatic stellate cells (HSCs) are also regulated by adiponectin which has anti-fibrotic effects including inhibition of a pro-inflammatory pathway involving NF-κB, decreased transforming

growth factor beta (TGFβ) -induced profibrogenic gene expression, and increased caspase-mediated apoptosis. The apoptosis effect is via activation of AMPK which downregulates the mammalian target of rapamycin, mTOR (Ding *et al.*, 2005).

Conversely, leptin possesses totally opposite functions to adiponectin, pro-inflammatory and pro-fibrotic. Leptin deficient (*ob/ob*) mice have reduced fibrogenesis (Leclercq *et al.*, 2002). Leptin is predominantly secreted from adipocytes, but can also be produced by HSCs (Bian and Ma, 2012). Leptin also targets sinusoidal endothelial and Kupffer cells, inducing the expression of TGFβ (Wang *et al.*, 2009). Activation of HSCs is modulated by leptin as its functional receptors are located on stellate cells. Levels of leptin appear low in quiescent HSCs and increase after activation (Ding *et al.*, 2005). Leptin can induce inflammation and fibrosis via HSC activation, including stimulation of proliferation, activation of NADPH and ROS, and increase in collagen 1 and tissue inhibitor of metalloproteinase 1 expression (Marra, 2007). Leptin can ultimately promote carcinogenesis via Janus kinase/signal transducers and activators of transcription (JAK/STAT) and phosphoinositide-3 kinase/protein kinase B/mTOR (PI3K/Akt/mTOR) signalling pathways (Saxena *et al.*, 2007).

TNFα, a potent activator of pro-inflammation and oncogenesis, was found to be upregulated in both obese human and mice studies (Anstee and Goldin, 2006; Park *et al.*, 2010). Overexpression of TNFα disrupts insulin-mediated suppression of hormone-sensitive lipase (HSL, the enzyme involved in TG hydrolysis) by inhibited phosphorylation of insulin receptor substrates (IRS-1 & IRS-2), hence more FFA is released (Anstee and Goldin, 2006; Sugimoto and Takei, 2011). Available data also indicates that dietary or genetic obesity promoted the growth of diethyl nitrosamine (DEN)-induced liver tumours in mice and this effect was dependent on TNFα and IL-6 mediated oncogenic signalling pathways via activation of STAT3 (Park *et al.*, 2010).

Insulin plays an essential role in glucose and lipid metabolism (Schreuder *et al.*, 2008). Insulin is released from β-cells in the islets of Langerhans, and further binds to its receptor, insulin receptor substrate proteins (IRS-1 and IRS-2) to transmit the signal. IRS-1 act as the initiator in glucose metabolism via stimulating PI3K/Akt pathway, resulting in recruitment of glucose transporter (GULT) and the suppression

of hepatic glucose production. Whereas IRS-2 is responsible for lipid metabolism as a regulator of DNL via SREBP-1c (Schreuder *et al.*, 2008). Insulin resistance (IR) is a condition when the signalling pathway has been disrupted by several insults, leading to an impaired cellular responses to insulin. NAFLD patients commonly have both hepatic and adipose tissue IR, together with diminished whole-body insulin sensitivity (Marchesini *et al.*, 2001; Seppala-Lindroos *et al.*, 2002; Bugianesi *et al.*, 2005). Whether IR is a result of lipotoxicity due to excess FFA flux to the liver or *vice versa* however remains unclear. The potential molecular mechanism of lipid-induced IR has been discussed thoroughly in a recent review article (Perry *et al.*, 2014). In summary, increased liver fat content (especially diacylglycerol, DAG) cause a key alteration – the translocation of the primary novel protein kinase C (PKC) isoform in the liver, PKCε. Following PKCε translocation (from cytosol to the plasma membrane), the enzyme binds to and inhibits the activity of the intracellular kinase domain of insulin receptors (Perry *et al.*, 2014), disrupting insulin signalling. This results in decreased IRS-2 phosphorylation, decreased IRS-2-induced PI3K activity, impaired phosphorylation of Akt, and suppression of glycogen synthase kinase-3 (GSK3) phosphorylation. The overall effect is a reduction of insulin-mediated glycogen synthesis via diminished activity of glycogen synthase (GS). The other way for insulin suppression of glucose production is the inhibition of gluconeogenesis in the liver; insulin decreases the expression of gluconeogenic enzymes by phosphorylation and nuclear exclusion of the folkhead box protein O1 (FOXO1) and its downstream targets. Impaired Akt activity promotes FOXO1 translocation to the nucleus (due to reduced phosphorylation of this transcriptional regulator) thus inducing the expression of key gluconeogenic proteins, principally pyruvate carboxylase (PC), phosphoenolpyruvate carboxykinase (PEPCK), and glucose-6-phosphatase (G6Pase) (Perry *et al.*, 2014).

Insulin resistance in NAFLD could be driven by multiple insults including direct toxic effects of FFAs, inflammation, oxidative injury and ER stress associated with obesity. These factors may impair the insulin signalling pathway, consequently leading to hyperinsulinemia resulting in increased glucose synthesis and lipogenesis (Saltiel and Kahn, 2001). Expression of SREBP-1c and ChREBP thus be activated once increased circulating levels of glucose and insulin occur, initiating a vicious cycle from promotion of glucose synthesis and lipogenesis to lipotoxicity. In addition,

adipose expansion, adipocytokine imbalance, immune dysregulation, ectopic fat deposition, lipotoxicity, IR, and hyperinsulinemia have all been suggested to be the fundamental risk factors that promote HCC development in NAFLD (Baffy *et al.*, 2012; Karagozian *et al.*, 2014).

1.7 Genetics of NAFLD

NAFLD is now considered as the product of multi-directional interactions between intrinsic factors (genetic, epigenetic and age related) and a range of extrinsic (environmental) influences (including dietary/nutritional factors, intestinal flora/microbiome, xenobiotics as well as activity/behavioural factors). Intestinal flora/microbiome and xenobiotics in NAFLD is beyond the scope of this thesis, but have been reviewed in depth by (Leung *et al.*, 2016) and (Naik *et al.*, 2013), respectively. Identification of genetic factors involved in susceptibility to NAFLD not only sheds light onto the background of disease pathogenesis, but also allows a non-invasive strategy that could help prevent/control disease initiation/progression and more importantly may develop new pharmacological treatments that exploit this knowledge.

1.7.1 Evidence for a heritable component to NAFLD

As has already been discussed, NAFLD is best considered a complex disease trait (Anstee and Day, 2013). Unlike Mendelian disorders (i.e. cystic fibrosis and sickle-cell disease) when rare and highly penetrant single-gene mutations are necessary and sufficient to cause disease, complex diseases traits are attributable to multiple genetic modifiers (where disease susceptibility in a population is determined by multiple variants as each is insufficient in isolation to cause disease) (Hirschhorn and Gajdos, 2011). Thus, a complex disease arises from a contribution of individual genetic variants in the presence of a permissive environment. In the case of NAFLD, the degree of heritability (fraction of the variation in the disease attributable to genetic causes) is estimated to be 26%-27% for radiologically measured hepatic triglyceride content (HTGC) (Speliotes *et al.*, 2011). Established evidence comes from familial disease clustering and inter-ethnic variation studies further supporting that there is a heritable component to NAFLD.

- **Familial aggregation**: a study comparing HTGC in parents and siblings of obese children with NAFLD to those without discovered that steatosis was significantly more common in the siblings and parents of children than those without NAFLD (59% and 78% vs. 17% and 37% respectively) (Schwimmer *et al.*, 2009). Additionally, MRI-measured hepatic fat fraction was correlated more closely with BMI in the families of children with NAFLD than those without (Schwimmer *et al.*, 2009). Further evidence from a Finnish study with a large cohort of twins discovered that intra-pair correlations of ALT and fasting serum levels were much higher in monozygotic than dizygotic twins (Makkonen *et al.*, 2009). Another study performed a cross-sectional analysis of a cohort of well-characterized twins in Southern California, reporting that hepatic steatosis and fibrosis were correlated with monozygotic but not dizygotic twins (Loomba *et al.*, 2015). Studies in adults with NAFLD showed co-existence of NASH and/or cryptogenic cirrhosis in 7 out of 8 kindreds (Struben *et al.*, 2000). Another familial study found that 18% of NASH patients had an affected first-degree relative (Willner *et al.*, 2001).

- **Inter-ethnic variations in susceptibility to NAFLD**: the prevalence of NAFLD and NAFLD-associated 'cryptogenic cirrhosis' varied substantially between ethnicities in the presence of similar incidences of T2DM (Browning *et al.*, 2004a; Bambha *et al.*, 2012). Further evidence from the Dallas Heart Study in a multi-ethnic population (n=2349) reported that NAFLD was present in 45% of Hispanics, 33% of whites and 24% of blacks, independent of BMI, IR, ethanol ingestion, or use of medication (Browning *et al.*, 2004b). It is difficult to make a precise interpretation of these data because of the differences in some metabolic risk factors and socioeconomic characteristics associated with NAFLD in these ethnic groups. As discussed below, differences in population prevalence of one genetic variant, rs738409 in the gene called *patatin-like phospholipase domain-containing 3* (*PNPLA3*), which contributed to this variability, accounted for up to 72% of ethnic differences in the Dallas Heart Study cohort (Romeo *et al.*, 2008).

1.7.2 Approaches to identify genetic risk factors for NAFLD

Genetic approach such as linkage analysis has been very successful in identifying causative variants for many rare single-gene disorders (Risch and Merikangas, 1996). However, linkage analysis has limited value for detection of common genetic risk factors associated with complex diseases due to: 1) the principle of linkage analysis is based on chromosomal crossover during meiosis, therefore it is restricted in familial cluster; 2) limitations for discovering weakly penetrant alleles; 3) low feasibility to obtain large families with complex diseases (Risch and Merikangas, 1996; Hirschhorn and Gajdos, 2011). These issues however can be addressed by candidate gene case-control association studies, performing direct association within large unrelated population (considered as a single family with unknown pedigree) between phenotype (disease) and variants in targeted genes with a priori biological support (Witte, 2010). Nonetheless, other causal genes that may influence a disease could be easily missed as it only focuses on particular genes. Logically, it would be inappropriate to adopt candidate gene studies to test genetic basis of any disease in the absence of a biological hypothesis (Hirschhorn and Daly, 2005). Up to date, candidate-gene studies have only identified a small number of genes that are reproducibly associated with common diseases; this may owe to the small sample-size used in the analyses or insufficient background knowledge of the genes targeted (Hirschhorn, 2009). After the establishment of final draft of human genome sequence, the International HapMap (haplotype map) Project, revolutionized the study of complex diseases (International HapMap, 2003; International HapMap, 2005). The HapMap Project was to determine the patterns of common sequence variation in human genome, providing sequence variants and their frequencies and correlations. Millions of common variants (single nucleotide polymorphisms, SNPs) were tagged by using the techniques of linkage disequilibrium (LD) and haplotype, and were further served as proxies for group of other SNPs that are nearby, thus eventually enabled the development of SNP genotyping arrays (International HapMap, 2005). These tools allow the majority of common (minor allele frequency [MAF]>5%) variability in human genome (approximately 1 million of different SNPs) to be captured by genotyping methods of cases and controls in genome-wide association studies (GWAS) (International HapMap, 2005). GWAS have the ability to discover variants with modest effects across the whole genome without particular biological evidence, improving the weakness of linkage analysis and candidate-gene

studies. But the loci identified are usually fairly novel, requiring further replications and functional experiments to ascertain their significance and to elucidate their biological/pathogenic role in disease mechanisms (Witte, 2010). A few years after the HapMap was established, GWAS in seven different complex diseases were reported (Wellcome Trust Case Control, 2007). Additional GWAS data relating to other important diseases, including cancer, soon became available (Manolio *et al.*, 2008).

1.7.3 Identified genetic risk factors for NAFLD by GWAS

Several genes have been identified to influence susceptibility to NAFLD by GWAS Table 1.2, and their suggested effects have been independently confirmed in either GWAS or candidate-gene association studies. Among all, two loci *PNPLA3* and *TM6SF2* genes on chromosome 22 and chromosome 19 deserve particular attention and so will be discussed in detail in section 1.7.3.6 and section 1.7.3.7, respectively.

1.7.3.1 ***Identified genetic risk factors by radiological-based GWAS***

The first GWAS-like association study in NAFLD was published in 2008 (Romeo *et al.*, 2008), this study partly covered the whole genome involving only genotyping for nonsynonymous variants. This study examined 9229 non-synonymous polymorphisms in a multi-ethnic population (Hispanic, African-American and European ancestry) from the Dallas Heart Study (Victor *et al.*, 2004), 2051 patients with ^{1}H-MRS-measured HTGC were recruited. Romeo *et al.* provided preliminary evidence that a SNP (rs738409) located in *PNPLA3* is significantly associated with increased HTGC (P=5.9x 10^{-10}) and elevated level of ALT (P=3.7x 10^{-4}). This index SNP is a non-synonymous transversion from cytosine to guanine, resulting in an amino acid substitution from isoleucine to methionine at residue 148 (I148M). The MAF of this polymorphism was most common in Hispanics (0.49), the group with higher prevalence of aggressive NAFLD and increased levels of ALT and AST (P=3.7x10^{-4}), and was lower in European ancestry (0.23) and African-Americans (0.17). A gene dosage effect of PNPLA3-I148M has been reported with a stepwise increase in HTGC with increasing carriage of the minor allele, and the homozygous 148M had twofold elevated HTGC compared to the wild-type carriers (Romeo *et al.*,

2008). Conversely, the other SNP also located in *PNPLA3* (rs6006460, p.S453I) had opposite ethnic distribution; African-Americans (0.104), European ancestry (0.003), and Hispanics (0.008), and was strongly linked to reduced HTGC. Taken together, these two SNPs accounted for 72% of the ethnic differences in the population studied, indicating that *PNPLA3* may be responsible for the ethnic variation in susceptibility to NAFLD (Romeo *et al.*, 2008).

Three years later, a two-stage study with a large US-based population (n=7,176) confirmed the association between *PNPLA3* genotype and fatty liver (Speliotes *et al.*, 2011). At stage 1 with computer-based imputation, Speliotes *et al.* performed a meta-analysis from four studies (Age/Gene/Environment Susceptibility –Reykjavik Study, Old Order Amish Study, Family Heart Study, and Framingham Heart Study). In this analysis, 2.4 million SNPs were examined and 45 loci were identified to be greatly associated with CT-measured HTGC ($P<10^{-3}$ was adopted as a permissive significant threshold). These loci were subsequently tested in stage 2 using a candidate-gene approach in a separate cohort with histology-proven NAFLD (n=592). To summarize, this study validated that carriage of *PNPLA3* rs738409 polymorphism was a major genetic modifier of NAFLD and also identified other potential variant risk factors; these included *neurocan* (*NCAN*, rs2228603), *glucokinase regulatory protein* (*GCKR*, rs780094), *protein phosphatase 1 regulatory subunit 3B* (*PPP1R3B*, rs4240624), and *lysophospholipase-like protein 1* (*LYPLAL1*, rs12137855). The above SNPs were significantly associated with both CT and histological confirmed steatosis, lobular inflammation and/or fibrosis, with the exception of *PPP1R3B* (rs4240624) as its effect did not remain significant in histologic NAFLD. Nonetheless, another SNP near *PPP1R3B* (rs2126259), in complete linkage disequilibrium with rs4240624 (LD=1, r^2=0.8), has been suggested to be associated with CT-measured steatosis in a later GWAS (Feitosa *et al.*, 2013).

1.7.3.2 *Identified genetic risk factors for NAFLD by histological-based GWAS*

In terms of histologically based GWAS for NAFLD, the first one was performed by Chalasani *et al.* on a limited cohort of 236 female NAFLD patients (Chalasani *et al.*, 2010). Instead of reporting associations with each stage of NAS (namely the individual stages of steatosis, ballooning degeneration, portal/lobular inflammation,

and fibrosis), they reported that the overall NAS (as primary outcome of interest) was significantly associated with a variant in a gene on chromosome 8 - *FDFT1* (*farnesyl diphosphate farnesyl transferase 1*, rs2645424) after adjustment for age, BMI, waist/hip ratio, diabetic status, and HbA1c level, (P=6.8x10^{-7}). Although FDFT1 has been suggested to have a role in biosynthesis of cholesterol, this effect was not replicated in a later study using a cohort of 340 NAFLD cases (Ballestri *et al.*, 2011). *PNPLA3* was not associated with any NAFLD phenotype in the GWAS published by Chalasani *et al*.; the key rs738409 SNP was not captured in their study and the other 5 SNPs in complete LD with rs738409 also failed to show significance. Additional loci were associated with secondary outcomes; a SNP on chromosome 7 was linked to the degree of fibrosis (rs343062, P=2.7x10^{-8}) and a variant on chromosome 10 with lobular inflammation (rs1227756, P=2.0x10^{-7}). Notably, caution should be taken in interpreting these results as type I or type II error may have occurred due to limited statistical power within this small cohort. Further independent validations are warranted.

Up to date, there have been two histological-based GWAS reported in East Asian populations (Kawaguchi *et al.*, 2012; Kitamoto *et al.*, 2013). Kawaguchi *et al*. recruited 529 biopsy-diagnosed NAFLD patients according to Matteoni classification and 932 population controls; 484,751 SNPs were genotyped (Kawaguchi *et al.*, 2012). The second study by Kitamoto *et al.* was a two-stage study; 261,540 SNPs were initially genotyped in 392 NAFLD patients and 932 population controls, and the findings were subsequently validated in a separate cohort of 172 NAFLD versus 1,012 controls (Kitamoto *et al.*, 2013). Both studies clearly confirmed the association between NAFLD and the SNPs flanking *PNPLA3* (with rs738409 having strongest signal) in this population, but no novel findings were identified.

1.7.3.3 *Identified genetic factors for NAFLD by clinical chemistry-based GWAS*

Although liver biochemistry is insufficient for NAFLD diagnosis and results in reports of a 3-12% lower disease prevalence than studies based on radiology or histology (Musso *et al.*, 2011), there are two established GWAS identifying genes that influence plasma levels of liver enzymes, principally ALT (Yuan *et al.*, 2008; Chambers *et al.*, 2011). The first one was reported by Yuan *et al.* and included 3

discovery and 3 replication groups with a total number of 12,419 patients. Two loci were identified to be associated with rasied serum ALT level, including 22q13.31 where *PNPLA3* is found (rs738409, rs22949158, rs2076211, rs2281135, rs2073081 were all genome-wide significant) and 10q24.2 spanning three genes *CPN1-ERLIN1-CHUK* (*carboxypeptidase N subunit 1-ER lipid raft associated 1-conserved helix-loop-helix ubiquitous kinase*; rs11597390, rs11591741, rs11597086) (Yuan *et al.*, 2008). The latter association has since been confirmed with CT-measured steatosis and ALT levels in another GWAS-correlated meta-analysis using ~2.5 million imputed SNPs with 9 variants genome-wide significant (rs2862954, rs1408579, rs10883451, rs11597086, rs11591741, rs17729876, rs17668255, rs17668357, rs12784396) (Feitosa *et al.*, 2013). A second GWAS on liver enzymes with 61,089 individuals enrolled identified 42 novel loci that associated with serum liver transaminase; 4 of them were associated with raised ALT, the major *PNPLA3* rs738409, rs2954021 near *TRIB1* (*tribbles pseudokinase 1*), rs10883437 near *CPN1* (*carboxypeptidase N subunit 1*), and rs6834314 near *HSD17B13* (*hydroxysteroid 17-beta dehydrogenase 13*) and *MAPK10* (*mitogen-activated protein kinase 10*) (Chambers *et al.*, 2011).

1.7.3.4 *Novel genetic risk factors for NAFLD reported in recent GWAS*

One recent GWAS published by the same group that originally reported the *PNPLA3* association (Kozlitina *et al.*, 2014) deserves particular attention as they discovered another novel SNP involved in the pathogenic mechanisms of NAFLD, also suggesting that this SNP is the causative variant within the well-known region (19p13.11) which has been associated with lipid levels (Kathiresan *et al.*, 2008; Teslovich *et al.*, 2010). In addition to the well-established association between *PNPLA3* and NAFLD, Kozlitina *et al*. determined that a non-synonymous SNP in a gene of unknown function called *TM6SF2* (rs58542926 c.449 C>T, p.Glu167Lys (E167K)), *transmembrane 6 superfamily member 2* on chromosome 19 (19p13.11), was strongly associated with ^{1}H-MRS quantified HTGC based on genotyping with a genome-wide exome chip (Kozlitina *et al.*, 2014).

The region (19p13.11) that the *TM6SF2* rs58542926 SNP is located on also contains several genes including *NCAN, CILP2* (*cartilage intermediate layer protein 2*), and

PBX4 (*PBX homeobox 4*) which have been reported in a previous study (Speliotes *et al.*, 2011). This region had also been associated with variations in plasma cholesterol, triglyceride and low-density lipoprotein levels in the previous studies (Kathiresan *et al.*, 2008; Teslovich *et al.*, 2010). Initial research focussed on *NCAN* as the lead candidate gene for this association, which was replicated for steatosis, inflammation and fibrosis. However, *NCAN* lacked biologically plausible evidence of a functional role in NAFLD (Anstee and Day, 2013). The use of a genome-wide exome-chip genotyping approach in 2,736 individuals, combined with detailed association analysis conditioning on previously identified variants across the 19p13.11 region, determined that the causative non-synonymous variant affecting HTGC is in strong linkage disequilibrium (D'=0.926, r^2=0.798) with the previously identified *NCAN* variant (rs2228603) but actually lies within the neighbouring gene, *TM6SF2* (rs58542926) (Kozlitina *et al.*, 2014). Most importantly, the effect of the *NCAN* variant was completely abrogated when conditioning on the *TM6SF2* variant whilst the reverse did not occur, suggesting that the association with HTGC at this locus is driven by *TM6SF2* rs58542926. Homozygous carriage of the *TM6SF2* rs58542926 minor (T) allele was shown to be associated with a modest but statistically significant increase in ^{1}H-MRS measured HTGC from 5.86±0.25% in CC homozygotes to 15.04±2.23% in TT homozygotes (Kozlitina *et al.*, 2014). While carriage of the *TM6SF2* rs58542926 minor (T) allele was associated with increased HTGC, another group coincidentally demonstrated that carriage of the common (C) allele of the *TM6SF2* rs58542926 was significantly associated with increased circulating cholesterol levels and greater risk for developing cardiovascular disease (CVD) (Holmen *et al.*, 2014). Another newly published GWAS enrolled 2,300 extremely obese individuals with liver biopsy data, once again validating the strongest association of *PNPLA3* and suggesting a locus (rs10401969) in the gene *SURP* and *G-patch domain containing 1* (SUGP1) near *TM6SF2* was significantly linked to hepatic steatosis (DiStefano *et al.*, 2015).

1.7.3.5 *Additional genetic risk factors for NAFLD identified in candidate gene studies*

Other than *PNPLA3* and *TM6SF2*, only a few loci associated with NAFLD have been independently validated and can be considered of proven importance. These include: *GCKR,* a key regulator of glucokinase activity that controls glucose

metabolism (Speliotes *et al.*, 2011); *SOD2,* which affects intracellular resistance to oxidative stress (Al-Serri *et al.*, 2012); *phosphatidylethanolamine N-methyltransferase* (*PEMT*), which catalyses the conversion of phosphatidylethanolamine to phosphatidylcholine and so is needed for normal hepatic VLDL secretion (Dong *et al.*, 2007); and *kruppel-like factor 6* (*KLF6*), a transcription factor that is highly expressed by activated stellate cells soon after injury (Ratziu *et al.*, 1998; Miele *et al.*, 2008).

The rs780094 SNP in *GCKR* is in strong LD with a functional nonsynonymous SNP (rs1260326, encoding Pro446Leu), and has been associated with hepatic TAG accumulation due to consistent increase in glucokinase activity and glucose uptake caused by impaired GCKR ability (Beer *et al.*, 2009). The role of *GCKR* in NAFLD has also been independently validated across several ethnic groups (Palmer *et al.*, 2013; Tan *et al.*, 2014). The rs4880 SNP in *SOD2* has been linked to advanced fibrosis in Japanese (Namikawa *et al.*, 2004) and European (Al-Serri *et al.*, 2012) cohorts. Carriers of the common (C) allele of rs4880 have more effective MnSOD mitochondrial transport and are thus potentially protected from cell damage compared to those are homozygous mutant (TT). Two studies reported that carriers of the *PEMT* variant (rs7946, p.Val175Met) had an increased susceptibility to NAFLD (Song *et al.*, 2005; Dong *et al.*, 2007). The *KLF6*–IVS1 –27G>A (rs3750861) SNP has been associated with milder NAFLD-related hepatic fibrosis in three separate European cohorts (Miele *et al.*, 2008). Carriage of the *KLF6*-IVS1 –27A polymorphism generates more of the KLF6-SV1 alternative spice isoform, lowering hepatic insulin resistance and blood glucose levels. This may in part be mediated through *GCKR* and glucokinase activity (Bechmann *et al.*, 2012).

Other genetic modifiers that have been identified by GWAS or candidate-gene association studies, involving in different aspects of NAFLD pathogenic mechanisms are summarised in Table 1.3.

Table 1.2: Genetic risk factors identified by GWAS, adapted and slightly modified from (Liu *et al.*, 2016b).

Genetic risk factor	Population (sample size, n)	Methodology (Numbers of SNPs)	Reference
Chr 22: *PNPLA3* (rs738409)	United States, mixed ethnicity (n=2,051) (B: 1,032; W: 636; H: 383)	^{1}H-MRS Steatosis (9,229)	(Romeo *et al.*, 2008)
Chr 22: *PNPLA3* (rs738409, rs22949158, rs2076211, rs2281135, rs2073081) ***CPN1-ERLIN1-CHUK*** **(rs11597390, rs11591741, rs11597086)**	European, mixed ethnicity (n=12,419) (3 discovery and 3 replication groups)	Clinical Biochemistry, ALT (not applicable)	(Yuan *et al.*, 2008)
FDFT1 **(rs2645424),** ***COL13A1*** **(rs1227756)** ***EFCAB4B*** **(rs887304),** ***PZP*** **(rs6487679)** **Chromosome 7 (rs343062)**	United States, European Caucasian with female only (n=236)	Histology (324,623)	(Chalasani *et al.*, 2010)
Chr 22: *PNPLA3* (rs738409) **Chr 19: *NCAN* (rs2228603)** ***GCKR*** **(rs780094),** ***LYPLAL1*** **(rs12137855),** ***PPP1R3B*** **(rs4240624)**	Meta-analysis of previous studies from United States & Europe (n=7,176)	CT-measured steatosis with histological 'candidate gene' validation set (Range 329k-618k before imputation)	(Speliotes *et al.*, 2011)
Chr 22: *PNPLA3* (rs738409) ***TRIB1*** **(rs2954021)** **Loci near *HSD17B13* and *MAPK10* (rs6834314)** ***CPN1*** **(rs10883437)**	European, mixed ethnicity (n=61,089)	Clinical Biochemistry, ALT (~2.6 million)	(Chambers *et al.*, 2011)
Chr 22: *PNPLA3* (rs738409)	Japanese (n=529)	Histology (484,751)	(Kawaguchi *et al.*, 2012)
Chr 22: *PNPLA3* (rs738409)	Japanese (n=392)	Histology (261,540)	(Kitamoto *et al.*, 2013)
Chr 22: *PNPLA3* (rs738409), ***PPP1R3B*** **(rs2126259),** ***ERLIN1–CHUK–CWF19L1*** **gene cluster (9 SNPs in two haplotype blocks)**	United States (n=2,705)	CT-measured steatosis (~2.5 million)	(Feitosa *et al.*, 2013)
Chr 22: *PNPLA3* (rs738409) **Chr 19: *TM6SF2* (rs58542926)**	United States, mixed ethnicity (n=2,736) (B: 1,324; W: 882; H: 467; O:63)	^{1}H-MRS Steatosis (138,374, exome)	(Kozlitina *et al.*, 2014)
PNPLA3 **and** ***SUGP1*** **(neighbouring gene to *TM6SF2*)**	United States, mixed ethnicity (n=2,300)	Histology	(DiStefano *et al.*, 2015)

Table 1.3: Additional genetic risk factors for NAFLD identified in candidate gene studies.

Gene	Protein	Comments
Glucose metabolism and insulin resistance		
ENPP1*; *IRS1	Ectonucleotide pyrophosphatase/phosphodiesterase family member 1; Insulin receptor substrate 1	Functional variants in *ENPP1/PC-1* and *IRS1* impair insulin receptor signalling and promote insulin resistance (McGettrick *et al.*, 2005; Grarup *et al.*, 2006). In 702 biopsy-proven NAFLD cases, carriage of nonsynonymous SNPs in *ENPP1* (rs1044498, encoding Lys121Gln) and *IRS1* (rs1801278, encoding Gln972Arg) reduced AKT activation, promoted insulin resistance and were independently associated with greater fibrosis (Dongiovanni *et al.*, 2010b). A second smaller (underpowered) study on *ENPP1* did not find a significant effect (Carulli *et al.*, 2009).
GCKR	Glucokinase regulatory protein	*GCKR* SNP rs780094 is in strong LD with a functional nonsynonymous SNP (rs1260326, encoding Pro446Leu) and has been associated with hepatic TAG accumulation in several studies (Chambers *et al.*, 2011; Speliotes *et al.*, 2011).
SLC2A1	Solute carrier family 2, facilitated glucose transporter member 1	A study examining 3,072 SNPs across 92 candidate genes identified variants in *SLC2A1* associated with NAFLD, independent of insulin resistance or T2DM (Petta *et al.*, 2014). Downregulation of SLC2A1 *in vitro* promoted lipid accumulation and increased oxidative stress, potentially linking the key pathogenic features of NAFLD: oxidative injury and increased lipid storage (Vazquez-Chantada *et al.*, 2013).
TCF7L2	Transcription factor 7-like 2	A role for *TCF7L2*, which has a key role in Wnt signalling and has been implicated in T2DM, has been reported in NAFLD (Musso *et al.*, 2009).
PPARG	Peroxisome proliferator-activated receptor γ	A loss-of-function SNP (rs1805192, encoding Pro12Ala) impairs transcriptional activation and affects insulin sensitivity (Tonjes *et al.*, 2006). Carriage of haplotypes including the Pro12Ala allele were associated with progressive NAFLD (Gawrieh *et al.*, 2012), but two studies found no association (Dongiovanni *et al.*, 2010a; Rey *et al.*, 2010).
Steatosis		
Hepatic lipid import or synthesis		
SLC27A5	Very long-chain acyl-CoA synthetase	Two principal isoforms of FATPs are expressed in the liver, *SLC27A2* (also known as *FATP2*) and *SLC27A5* (also known as *FATP5*) (Hirsch *et al.*, 1998). Silencing *Slc27a5* reverses diet-induced NAFLD and improves hyperglycaemia in mice (Doege *et al.*, 2008). Carriage of the *SLC27A5* rs56225452 promoter region polymorphism has been associated with higher ALT levels, and greater postprandial insulin and triglyceride levels (Doege *et al.*, 2008). In patients with histologically proven NAFLD, the effect of BMI on degree of steatosis differed with *SLC27A5* genotype (Auinger *et al.*, 2010).

LPIN1	Phophatidate phosphatase LPIN1	LPIN1 is required for adipogenesis and the normal metabolic flux between adipose tissue and liver, where it also acts as an inducible transcriptional co-activator to regulate fatty-acid metabolism (Reue and Zhang, 2008; Reue, 2009). Variants have been associated with multiple components of the metabolic syndrome (Reue and Zhang, 2008; Wiedmann *et al.*, 2008). Although a large case–control study found no association with T2DM, obesity or related traits in 17,538 individuals (Burgdorf *et al.*, 2010), a meta-analysis in 8,504 individuals found that the *LPIN1* rs13412852 [T] allele was associated with lower BMI and insulin levels (Fawcett *et al.*, 2008). This same polymorphism was under-represented in paediatric (but not adult) NAFLD with a suggestion of less severe liver damage (Valenti *et al.*, 2012).
Hepatic lipid export or oxidation in steatosis		
PNPLA3	Patatin-like phospholipase domain-containing 3	The nonsynonymous 617C>G nucleotide transversion mutation SNP (rs738409, encoding Ile148Met) has been consistently associated with steatosis, steatohepatitis and hepatic fibrosis; however, function remains incompletely understood (Romeo *et al.*, 2008; Valenti *et al.*, 2010a).
NR1I2	Nuclear receptor subfamily 1 group I member 2 (also known as pregnane X receptor)	*NR1I2* encodes a transcription factor that regulates hepatic detoxification (Zhang *et al.*, 2008) and acts through CD36 (fatty-acidtranslocase) and various lipogenic enzymes to control lipid metabolism (Zhou *et al.*, 2006). *Nr1i2*-deficient mice develop steatosis (Zhou *et al.*, 2006). Two SNPs (rs7643645 and rs2461823) were associated with NAFLD and were also a predictor of disease severity (Sookoian *et al.*, 2010).
PPARA	Peroxisome proliferator-activated receptor α	PPAR-α is a molecular sensor for long-chain fatty acids, eicosanoids and fibrates; activated by increased hepatocyte fatty-acid load, it limits TAG accumulation by increasing fatty-acid oxidation (Kim *et al.*, 2003). Carriage of a nonsynonymous SNP (rs1800234, encoding Val227Ala) increases activity, and was associated with NAFLD despite reduced BMI (Yamakawa-Kobayashi *et al.*, 2002; Chen *et al.*, 2008). A loss-of-function polymorphism (rs1800206, encoding Leu162Val) was not associated with NAFLD (Dongiovanni *et al.*, 2010a).
PEMT	Phosphatidylethanolamine *N*-methyltransferase	Two studies have reported an association between NAFLD and a nonsynonymous *PEMT* exon 8 590G>A transversion (rs7946, encoding Val175Met) (Song *et al.*, 2005; Dong *et al.*, 2007).
MTTP	Microsomal triglyceride transfer protein	MTTP mediates hepatic synthesis and secretion of VLDL. Abetalipoproteinaemia (OMIM#200100) results from a loss-of-function frameshift mutation in *MTTP;* however, whereas this mutation causes severe hepatic TAG accumulation, steatohepatitis and fibrosis are infrequent (Lonardo *et al.*, 2006). A promoter region transversion (–493G>T; rs1800591), predisposed to steatosis and NASH (Bernard *et al.*, 2000; Namikawa *et al.*, 2004) in a small cohort, but a larger study in 131 patients found no association (Oliveira *et al.*, 2010).
APOC3	Apolipoprotein C-III	Two promoter region SNPs –455T>C; rs2854116 and –482C>T; rs2854117) that increased steatosis were reported in small (*n* = 95 and 163) cohorts of Asian-Indian and non-Asian ethnicity (Petersen *et al.*, 2010). To date, studies together examining >4,000 individuals have been unable to replicate these findings (Kozlitina *et al.*,

		2011; Sentinelli *et al.*, 2011; Valenti *et al.*, 2011a; Verrijken *et al.*, 2013).
APOE	Apolipoprotein E	ApoE is a plasma protein involved in lipid transport and metabolism (Utermann *et al.*, 1977). Three alleles (ε2, ε3 and ε4) determine three isoforms (ApoE2, ApoE3 and ApoE4) resulting in six ApoE genotypes (E2/2, E3/3, E4/4, E2/3, E2/4, E3/4). Overall homozygosity for the ε2 allele in one study was associated with dyslipidaemia, but not NAFLD (Demirag *et al.*, 2007). In a subgroup of nonobese individuals, the ε2 allele and the E2/3 genotype were more prevalent in controls, suggesting this allele might be protective. (Demirag *et al.*, 2007). Consistent with this result, the ApoE3/3 genotype was associated with NASH in a Turkish cohort, whereas ApoE3/4 was protective (Sazci *et al.*, 2008).
Steatohepatitis		
Oxidative stress		
HFE	Hereditary hemochromatosis protein	Hepatic iron accumulation promotes oxidative stress. Two studies, examining 177 patients, reported carriage of an *HFE* polymorphism (rs1800562, encoding Cys282Tyr) that was associated with more severe steatohepatitis and advanced fibrosis (George *et al.*, 1998; Nelson *et al.*, 2007). However, three other studies (Bugianesi *et al.*, 2004; Raszeja-Wyszomirska *et al.*, 2010; Valenti *et al.*, 2010c)have not shown increased carriage of either the Cys282Tyr or His63Asp (rs1799945) mutations. Meta-analyses have also provided conflicting results, with the latest finding no evidence of an effect (Ellervik *et al.*, 2007; Hernaez *et al.*, 2011).
GCLC; GCLM	Glutamate-cystein ligase catalytic unit; glutamate-cystein ligase regulatory unit	Glutamate-cysteine ligase (γ-glutamyl cysteine synthetase) is the rate-limiting step in glutathione synthesis; absence of Gclc causes steatosis and liver failure in mice (Buch *et al.*, 2007). A study of 131 patients with NAFLD found the *GCLC* promoter region polymorphism (–129C>T, rs17883901) was associated with steatohepatitis compared with simple steatosis (Oliveira *et al.*, 2010).
ABCC2	ATP-binding cassette, sub-family C (CFTR/MRP), member 2	Association studies support a role for *ABCC2* (also known as *MRP2*), which facilitates terminal excretion and detoxification of endogenous and xenobiotic organic anions, including lipid peroxidation products (Sookoian *et al.*, 2009a).
SOD2	Superoxide dismutase [Mn], mitochondrial	Carriage of the nonsynonymous SNP (rs4880, encoding Ala16Val) has been associated with advanced hepatic fibrosis in NAFLD in both Japanese (Namikawa *et al.*, 2004) and European (Al-Serri *et al.*, 2012) cohorts.
Endotoxin response		
TLR4	Toll-like receptor 4	Study of a spontaneous *Tlr4* null mutation in C3H/J mice has established the contribution of TLR4/endotoxin to NAFLD pathogenesis in the laboratory (Spruss *et al.*, 2009). *TLR4* polymorphisms (rs4986791 and rs4986790) influence hepatitis-C-related fibrosis (Huang *et al.*, 2007; Guo *et al.*, 2009), but no association with NAFLD and either *TLR4* or *NOD2* (bacterial cell wall peptidoglycan receptor) variants has been found.

CD14	Monocyte differentiation antigen CD14	CD14 is a lipopolysaccharide receptor expressed on monocytes, macrophages and neutrophils that enhances TLR4 endotoxin signaling. An association with a promoter-region polymorphism (–159C>T, rs2569190) that increases *CD14* expression has been reported (Baldini *et al.*, 1999; Day *et al.*, 2006).
Cytokines		
TNF	Tumour necrosis factor	A *TNF* (–238G>A, rs361525) promoter polymorphism has been associated with NASH (Valenti *et al.*, 2002; Tokushige *et al.*, 2007), suggesting a primary role in the transition from steatosis to steatohepatitis; a separate study, found that two other promoter region polymorphisms (–1031T>C, rs1799964 and –863C>A, rs1800630) were more common in NASH than steatosis, but were no more common in NAFLD than a control population (Tokushige *et al.*, 2007).
IL6	Interleukin 6	An *IL6* promoter region polymorphism (–174G>C, rs1800795) has been associated with NASH (Carulli *et al.*, 2009).
Fibrosis		
AGTR1	Type-1 angiotensin II receptor	Two studies have linked the *ATGR1* rs3772622 SNP with grade of steatohepatitis and stage of fibrosis, with the most recent study also suggesting an interaction with *PNPLA3* genotype (Yoneda *et al.*, 2009; Zain *et al.*, 2013).
KLF6	Kruppel-like factor 6	The *KLF6*–IVS1 –27G>A (rs3750861) SNP has been associated with milder NAFLD-related hepatic fibrosis in three separate European cohorts (Miele *et al.*, 2008).
MBOAT7/ LPIAT1	*Membrane bound O-acyltransferase domain containing 7 (also known as Lysophosphatidylinositol acyltransferase-1)*	A recent study with two large cohort sets (Dalles Heart Study, n=3854 and Liver Biopsy Cross-Sectional Cohort, n=1,149) reported the rs641738 SNP in *MBOAT7* was associated with increased risk in developing steatosis, severe necro-inflammation and advanced fibrosis (Mancina *et al.*, 2016). The effect of rs641738 variant on NAFLD/NASH development may be mediated by alterations in the hepatic phosphatidylinositol acyl-chain remodeling (Mancina *et al.*, 2016).

Adapted and slightly modified from (Liu *et al.*, 2016b).

1.7.3.6 *Biological relevance of PNPLA3*

Ever since the association of *PNPLA3* with NAFLD was identified by the first GWAS-like study in NAFLD (Romeo *et al.*, 2008), this association has been validated across multiple patient cohorts. The link between *PNPLA3* and NAFLD was established in radiologically phenotyped cohorts however such studies are unable to assess the presence of features such as NASH or stage of liver fibrosis that can only be detected histologically. This deficiency reflects the difficulty in assembling sufficiently large patient cohorts that have undergone invasive testing by liver biopsy to provide a well-powered fully phenotyped cohort and means that the majority of GWAS studies are radiologically based and so 'phenotype limited' to detecting variations in HTGC. Candidate-gene disease association studies in histologically characterised patient cohorts have therefore been necessary to establish whether genetic variants identified by GWAS, such as *PNPLA3*, do indeed influence disease progression towards the more clinically relevant phenotypes.

The *PNPLA3* rs738409 variant has been shown to increase the risk of NASH and advanced liver fibrosis, independent of potential confounding factors such as age, BMI or T2DM (Valenti *et al.*, 2010a; Anstee and Day, 2013). This variant has also been associated with increased HCC risk in alcohol-related liver disease (Nischalke *et al.*, 2011; Trepo *et al.*, 2012; Guyot *et al.*, 2013; Trepo *et al.*, 2014) and, more variably, in chronic viral hepatitis (Falleti *et al.*, 2011; Nischalke *et al.*, 2011; Valenti *et al.*, 2011b; Guyot *et al.*, 2013; Trepo *et al.*, 2014). Data has also been presented showing an association with HCC in morbidly obese patients (Burza *et al.*, 2012) and a mixed-aetiology cohort (Hassan *et al.*, 2013). Although it may be hypothesised that these latter associations are related to underlying NAFLD, prior to the studies within this thesis, no studies had specifically addressed the effect of *PNPLA3* rs738409 C>G carriage on HCC risk in a NAFLD cohort.

As a strong genetic risk factor for NAFLD, the biological function of PNPLA3 has been extensively studied. The *PNPLA3* gene encodes a 481 amino acid protein (also referred to as adiponutrin) that belongs to a family of lipid hydrolases (patatin-like phospholipase domain containing proteins, PNPLA1-9) (Park *et al.*, 1983). Among the human PNPLA proteins, PNPLA3 is structurally related to *adipose triglyceride lipase* (*ATGL,* also known as *PNPLA2*), the major triglyceride hydrolase in adipose tissue (Zimmermann *et al.*, 2004; Kienesberger *et al.*, 2009). Hence, PNPLA3 was

first thought to possess a similar function as ATGL. However, in contrast to the excessive TG accumulation that occurs following *ATGL* loss in mice, *PNPLA3* knock-out mice do not show increased hepatic TG content (Zimmermann *et al.*, 2004). Similarly, reduced TG content was not seen when the *PNPLA3* (I148) wild-type variant was overexpressed whereas this did occur with *ATGL* overexpression (He *et al.*, 2010). These studies both suggest that the function of PNPLA3 differs substantially from ATGL. Evidence established so far suggests that there are some differences in tissue distribution of PNPLA3 expression between humans and mice which has made functional study of PNPLA3 more difficult. Amongst metabolically active organs, PNPLA3 is mainly expressed in the liver of humans (Wilson *et al.*, 2006) whilst adipose tissue expression dominates in mice (Lake *et al.*, 2005). The expression of *PNPLA3* is controlled in a nutrition-dependent manner: on fasting, expression is down-regulated whilst consumption of a high-carbohydrate diet up-regulates expression (Dubuquoy *et al.*, 2011). *PNPLA3* is regulated by *ChREBP* under hyperglycaemia resulting in increased hepatic glycolysis and lipogenesis (Rae-Whitcombe *et al.*, 2010; Perttila *et al.*, 2012b). Insulin also controls postprandial *PNPLA3* expression through liver *X receptor–retinoid X receptor* (*LXR-RXR*) and *SREBP-1c* that also promotes *de novo* lipogenesis (Shimomura *et al.*, 1999a; Huang *et al.*, 2010).

Despite the plethora of genetic evidence for an effect, the physiological role of PNPLA3 and how it is perturbed by carriage of the rs738409 (I148M) variant has remained elusive. Although some progress has been made, the data is at times conflicting and there remains some debate as to whether the variant is a true loss-of-function mutation or not. *In vitro* studies with both wild-type and mutant isoforms of recombinant adiponutrin indicate that PNPLA3 mediates hydrolysis of acylglycerols and that maximal hydrolytic activity is towards any of the three major glycerolipids (triacylglycerol, diacylglycerol and monoacylglycerol) as substrates, with a strong preference for oleic acid as the acyl moiety (He *et al.*, 2010; Huang *et al.*, 2011). Compared to the PNPLA3 wild-type protein, the I148M variant possesses substantially reduced enzymatic activity but there does not appear to be any change in substrate affinity (He *et al.*, 2010; Huang *et al.*, 2011). These findings are consistent with the results of stable isotope tracer studies in overweight or obese men and *in vitro* studies which show that carriage of the rs738409 (I148M) variant

reduces VLDL secretion, an effect attributed to failure to mobilize TG from intracellular lipid droplets (Pirazzi *et al.*, 2012). Arguing against a simple loss-of-function effect however, *PNPLA3* knockout mice do not develop steatosis (Basantani *et al.*, 2011). Indeed, others have suggested that the I148M substitution acts as a gain-of-function mutation with the variant form possessing some lysophosphatidic acid acetyltransferase activity, leading to an increase of TG synthesis (Kumari *et al.*, 2012). Supporting this, murine overexpression of wild-type PNPLA3 does not induce steatosis but overexpression of the variant (I148M) form does (Li *et al.*, 2012). A recent study with PNPLA3 I148M knockin mice showed that HTGC level was normal on a chow diet but increased 2 to 3-fold compared to wild-type littermates with high sucrose diet. This increased fat was accompanied by a 40-fold increase in PNPLA3 on hepatic lipid droplets, without increase in hepatic *PNPLA3* messenger RNA (mRNA). Similar results were observed when inactivating the catalytic dyad of PNPLA3 by substituting the catalytic serine with alanine (S47A), suggesting that the catalytically inactive PNPLA3 protein is necessitate to drive NAFLD pathogenesis instead of a complete absence of PNPLA3 activity (Smagris *et al.*, 2015).Taken together, the currently available data would suggest that the PNPLA3 I148M variant alters TG remodelling in lipid droplets within hepatocytes (Li *et al.*, 2012; Ruhanen *et al.*, 2014).

Beyond steatosis, the underlying mechanisms through which *PNPLA3* influences progression to NASH and hepatic fibrosis still remained uncertain. Recent data suggests that *PNPLA3* may have a role in retinol metabolism, acting as retinyl-palmitate hydrolase in human HSCs which are dominant players in fibrogenesis (Pirazzi *et al.*, 2014). Overexpressed wild-type PNPLA3 in HSCs resulted in a substantial reduction of lipid droplets, an effect that was lost with I148M. However, how this alters HSC activation and affects collagen deposition and fibrosis is not known however it does suggest that PNPLA3 may have specific roles in different cell types and metabolic conditions which each contribute to the progression of NAFLD from steatosis to fibrosis. Thus, challenges remain in which cell types to target and what kinds of experimental conditions should be applied in order to investigate the comprehensive role of *PNPLA3* in NAFLD.

1.7.3.7 *Biological aspects of TM6SF2*

Reflecting the recent nature of the discoveries mentioned in section 1.7.3.4, little is currently known about the biological function of TM6SF2. It was first described as a multi-pass membrane protein in 2000 (Carim-Todd *et al.*, 2000), and was later discovered to be highly expressed in liver, kidney and intestines (Kozlitina *et al.*, 2014; Surakka *et al.*, 2015). *In vitro* study using confocal microscopy, GFP-tagged TM6SF2 was mainly localized to the endoplasmic reticulum (ER) and ER-Golgi intermediate compartments (ERGIC) (Mahdessian *et al.*, 2014). In the same study, TM6SF2 siRNA inhibition led to a reduction of lipid secretion (triglycerides (TG) and Apolipoprotein B (ApoB)) and an increase in the number and overall size of lipid droplets, which represented as a manifestation of cellular triglyceride accumulation. On the contrary, overexpression of TM6SF2 caused a decrease in the lipid droplet contents (Mahdessian *et al.*, 2014). Prediction of mouse knockout phenotype by analysis of coexpressed gene profiles based on Mouse Genome Informatics (MGI) Database, TM6SF2 is linked to abnormal lipids levels (decreased total cholesterol (TC), low-density lipoprotein-cholesterol (LDL-C) and VLDL) and may act as lipid transporter and interact with proteins involved in intestinal absorption (Surakka *et al.*, 2015). *In vivo* studies may provide more clues for above statements; adenovirus-mediated short hairpin RNA knockdown of *Tm6sf2* in mice has been shown to increase hepatic triglyceride content and reduce plasma cholesterol and VLDL secretion, suggesting that *TM6SF2* activity is necessary for normal VLDL secretion and that impaired *TM6SF2* function causally contributes to NAFLD (Kozlitina *et al.*, 2014). On the other hand, transient overexpression of *Tm6sf2* in C57BL/6J mice resulted in a significant increase of serum TC, LDL-C and TG. These *in vivo* (Holmen et al., 2014; Kozlitina et al., 2014) and *in vitr*o (Holmen *et al.*, 2014; Kozlitina *et al.*, 2014; Mahdessian *et al.*, 2014) functional studies indicate that *TM6SF2* has effects on fat retention in the liver, lipid efflux and alteration in the number and average size of lipid droplets. However, the genetic and biological studies previously mentioned only focused on relevance to hepatic steatosis. Whether the effect of *TM6SF2* is limited to steatosis or has broader clinical relevance, affecting susceptibility to both steatohepatitis and fibrosis as has already been shown for *PNPLA3,* remains undetermined. Further research is needed to establish the precise function of the TM6SF2 protein and to determine how this knowledge can be exploited clinically.

1.8 Aims

NAFLD is best considered a complex disease trait in which subtle inter-patient genetic variations and environmental factors interact to determine disease phenotype and progression. Recent technical advances have led to the identification of important genetic modifiers, in particular non-synonymous gene variants in *PNPLA3* (I148M) and *TM6SF2* (E167K) that are associated with NAFLD. The aims of this thesis are to identify and validate candidate genes/SNPs based on the aforementioned literature that confer progression of NAFLD to its progressive stages, principally the *PNPLA3* and *TM6SF2* genes.

The work of this thesis can be divided into two parts: case-control association studies (the majority of the thesis) and functional studies:

- Case-control association studies:
    - To validate the *PNPLA3* association in the FLIP GWAS cohort (n=1,005) to confirm its relevance to disease spectrum of NAFLD (steatosis, steatohepatitis and fibrosis)
    - To validate and expand the relevance of *TM6SF2* to NAFLD severity
    - To investigate the relevance of *PNPLA3* and *TM6SF2* to NAFLD-related HCC
- With the available genetic data, pilot studies investigating the function of *PNPLA3* were subsequently conducted:
    - To explore biological differences caused by the amino-acid substitution *in vitro*, mainly focused on the investigation of the correlation between PNPLA3-I148M and lipid metabolism in different cultured cells

Chapter 2. Materials and Methods

2.1 Patients

The patients involved in this project can be divided into two cohorts: 1) the patients recruited from Newcastle, UK and 2) the patients from a FP7 project (FLIP, Fatty liver: Inhibition of Progression) funded by European Commission. These samples were collected at NAFLD clinics and bariatric surgery clinics across Europe.

- Newcastle-based patients: Patients with biopsy-proven NAFLD were recruited from the Freeman hospital NAFLD clinic.
- FLIP-based patients: Patients were recruited from hepatology clinics at several European specialist centres: the Freeman Hospital, Newcastle-upon-Tyne, UK; Addenbrooke's Hospital, Cambridge, UK; Queen's Medical Centre, Nottingham, UK; Inselspital Hospital, Bern, Switzerland; Pitié-Salpêtrière Hospital, Paris, France; and Antwerp University Hospital, Belgium.

All the necessary ethical approvals were obtained for this project and all participants gave informed consent. In all cases, alternative diagnoses were excluded, including excess alcohol intake (alcohol consumption < 20g/day for women; < 30g/day for men), chronic viral hepatitis (hepatitis B and hepatitis C), autoimmune liver diseases, hereditary hemochromatosis, α1-antitrypsin deficiency, Wilson's disease and drug-induced liver injury. Clinical and laboratory data were collected at the time of diagnosis including basic anthropometrics so that body mass index (BMI) could be calculated, and relevant co-morbidity including the presence of type 2 diabetes mellitus (fasting glucose ≥7.1 mmol/L [≥128 mg/dl] or treatment with anti-diabetic drugs) and evidence of underlying cirrhosis was recorded.

2.2 Liver biopsy

Liver biopsy was performed under radiological guidance. Specimens (at least 1.6 cm length and 1.5 mm thick) were fixed in 10% neutral formalin for evaluation and embedded in paraffin for histological examination. Tissue sections were stained with hematoxylin and eosin, impregnated with silver for visualizing reticulin framework and stained with Sirius Red Fast Green for visualizing collagen. Liver biopsies for Newcastle samples were reviewed by two expert liver pathologists: Alastair Burt and Dina Tiniakos. Liver biopsies for FLIP samples were performed by the collaborators

at each participating centre. Liver histology scoring was performed according to Kleiner *et al*. (Kleiner *et al.*, 2005) and this provided a NAFLD activity score (NAS) which covers the degree of steatosis (0-3), hepatocyte ballooning (0-2) and lobular inflammation (0-3), and a separate fibrosis stage (0-4). A recently released scoring system, the validated semi-quantitative SAF score, is similar to the NAS score but separates the degree of steatosis from those of inflammation/ballooning to prove three separate measures of NAFLD severity. On this system, NAFLD is classified into the degree of steatosis (0-3), activity of steatohepatitis (A0-4) and stage of fibrosis (F0-4) (Bedossa and Consortium, 2014). This SAF score system was also considered in this study. Comparison of the SAF Score and the NAFLD Kleiner Score for the histological grading and staging of NAFLD/NASH is shown in Table 1.1.

2.3 Laboratory methods

2.3.1 Materials

Materials and methods described in this chapter is relevant to several different chapters. More specific methods sections are provided in Chapters 3, 4, 5 and 6. Suppliers of materials commonly used in this work are shown in Table 2.1 or otherwise stated in the text. The composition of stock solutions commonly used is listed in Table 2.2.

Table 2.1: List of suppliers and addresses.

Suppliers	Address
Applied Biosystems	California, USA
Bioline	London, UK
Bio-Rad	Hemel Hempstead, UK
Fisher scientific	Loughborough, UK
Greiner Bio-One	Stonehouse, UK
Lonza	Slough, UK
New England Biolabs (NEB)	Hitchin, UK
QIAGEN	Crawley, UK
Sigma Aldrich	Gillingham, UK
Thermo Scientific	Loughborough, UK

Table 2.2: Compositions of commonly used stock solutions.

Solution	Composition
10X TAE (Tris-acetate-EDTA) buffer per litre	10 mM Tris-base 48.4 g/L 17.4 M Glacial acetic acid 11.4 ml/L 10 mM EDTA 3.7 g/L
DNA loading buffer	0.25% (w/v) bromophenol blue 0.25% (w/v) xylene cyanol 30 % glycerol
2X protein sample loading buffer	0.125 M Tris-HCl pH 6.8 4% SDS 20% glycerol 0.004% bromophenol blue 10% β-mercaptoethanol

2.3.2 DNA extraction from blood samples

Venous blood was collected from each patient and DNA was prepared from peripheral blood lymphocytes using a perchlorate-chloroform isolation method (Daly *et al.*, 1996). Briefly, 35 ml lysis buffer (10 mM Tris-HCl (pH 8.0), 320 mM Sucrose, 5 mM Magnesium Chloride, 1% Triton X-100) was added to 5 ml venous blood in a 50 ml polypropylene centrifuge tube. After mixing, the tube was centrifuged at 3000G for 10 min. The supernatant was discarded and the cell pellet was re-suspended in 2 ml of solution B (400 mM Tris-HCl (pH 8.0) 60 mM EDTA, 150 mM NaCl, 1% SDS). 500 µl of sodium perchlorate (5 M) was added and the sample was mixed at room temperature for 15 min before incubating in a preheated hot block at 65^{o}C for 30 min. Next, 2 ml chloroform was added and the sample was further mixed for 10 min at room temperature. Having been centrifuged at 1400G for 10 min, the upper and clear DNA-containing phase was transferred to a new 15 ml polypropylene tube. Cold ethanol (7 ml) was added to the aqueous phase and the tube was gently inverted until the DNA precipitated. The DNA was spooled by using a soft plastic sterile loop and allowed to air dry for 20 min. DNA was then re-suspended in 200 µl of 5 mM Tris-HCl solution (pH 8) followed by the incubation at 60^{o}C overnight. Samples were quantitated and quality assessed by absorbance measurements at 260 and 280 nm. Some DNA preparation on Newcastle samples was performed by Julian Leathart and Julia Patch (Newcastle University, ICM). DNA isolation from FLIP samples was performed locally by the collaborators.

2.3.3 Polymerase chain reaction (PCR)

Lyophilised primers were purchased from Sigma Aldrich (UK). Forward primers were designed approximately 180 base pairs upstream from the polymorphism of interest, while reverse primers were around 100 base pairs downstream. Primers were re-suspended in sterile water for stock purpose (200 µM). Further dilutions (25 µM) were performed during each work. Stock primers and working dilutions were stored at -20°C and 4°C, respectively. A total reaction volume of 25 µl of polymerase chain reaction was performed in a 200 µl sterile tube, reagents used including 0.1 mM dNTPs, 0.25 µM for both forward and reverse primers, 0.025 units of Taq DNA polymerase, 1X ThermoPol reaction buffer, and genomic DNA (typically 0.2 µg). Standard cycling conditions were used with varying annealing temperature according

to different nature of primers. 8 µl of PCR products were mixed with 3 µl of DNA loading buffer (Table 2.2) and viewed on 2% agarose gel (Table 2.4) at 50 mA for 30 min. Materials and suppliers used for PCR are listed in Table 2.3.

Table 2.3: Reagents used for PCR.

Name	Supplier	Concentration
100X dNTP mix solution	Bioline, London, UK	0.1 mM
100X forward and reverse primers	Sigma Aldrich, Gillingham, UK	25 µM
200X Taq DNA polymerase	New England Biolabs, Hitchin, UK	0.025 units
10X Thermopol Reaction Buffer	New England Biolabs, Hitchin, UK	1X
DNA	Samples listed in section 2.1	0.2 µg
Sterile water	Braun Medical Supplies, Inc. Melsungen, Germany	

Table 2.4: Composition of 2% agarose gel.

Composition	
2% Agarose gel	2 g Agarose powder 100 ml 1X TAE buffer 0.5 µg/ml ethidium bromide

2.3.4 Genotyping

Genotyping was performed without knowledge of clinical status or histology of patients using two methods: digestion of PCR products by restriction fragment length polymorphism (PCR-RFLP) and TaqMan SNP genotyping assay. The method used depended on the individual polymorphism being studied.

2.3.4.1 *Digestion of PCR products by restriction fragment length polymorphism (PCR-RFLP)*

Restriction fragment length polymorphism (RFLP) is a technique that genotypes for polymorphisms by using highly locus-specific restriction enzymes which digest one variant only. Following digestion, fragments are separated according to their size by gel electrophoresis. Digestion was performed in a total reaction volume of 22 µl, containing 17 µl PCR product, 3 units restriction enzyme (NEB, UK), and digestion buffer (NEB, UK). Samples were incubated at 37°C at least for 3 hours but usually overnight. Digested products (8 µl) were mixed with DNA loading buffer (3 µl) and electrophoresed on 2% agarose gel (Table 2.4) at 50 mA for an hour. Various sizes of DNA ladders were included for reference. Agarose gels were visualized and photographed on a transilluminator (GENi documentation system, Cambridge, UK).

2.3.4.2 *TaqMan SNP genotyping assay*

Genotyping was also performed by TaqMan SNP Genotyping Analysis (Applied Biosystems, USA). TaqMan SNP genotyping is a fast, simple and high-quality approach to genotyping a large number of samples. Each TaqMan assay was supplied as a solution (20X or 40X, 188 µl) containing unlabelled forward and reverse primers (1X final concentration was 900 nM) and two reporter probes (1X final concentration was 200 nM): VIC dye is linked to the 5´ end of the Allele 1 sequence while fluorescein amidite (FAM) dye is linked to the 5´ end of the Allele 2. The 3´ end of each probe is incorporated with minor groove binder (MGB) technology, which the manufacturers assert delivers superior allelic discrimination. A 2X TaqMan universal master mix (Applied Biosystems, USA) including AmpliTaq Gold® DNA polymerase, dNTPs and a passive internal reference based on proprietary ROX™ dye for the Applied Biosystems® real-time PCR instrument

(StepOne Real-Time PCR system) was also used. Materials and background TaqMan chemistry are displayed in Figure 2.1.

The PCR was performed using a 48-well reaction plate (Applied Biosystems, USA). Enough reaction mix was made for 50 samples to avoid volume loss in pipetting. 20X working solution (12.5 ul) of SNP Genotyping Assay (6.25 µl for 40X) were added to 273.5 µl 2X universal master mix. After briefly vortexing, 5 µl of the mixture were then transferred into each well of a 48-well plate already containing 5 µl of genomic DNA (5 ng/µl) diluted in sterile H_2O. Quality controls were also included in each plate (duplicate of negative, homozygous wild-type, homozygous mutant, and heterozygous controls). The plate was sealed and briefly centrifuged at 300 rpm for 1 min to avoid air bubbles, and further inserted into the StepOne Real-Time PCR machine. A standard 40 cycles of PCR was performed. Each PCR cycle was started with a hold mode at 95°C for 10 min, the denaturation process was run at 92°C for 15 sec, and stages of annealing and extension were performed at 60°C for 1 min.

Allelic discrimination plots were generated by the proprietary StepOne software. Well-distributed clusters were displayed on a XY axis chart based on the different fluorescence signals each genotype defined. A cluster located horizontally at the bottom end of X axis represented homozygosity of one allele (XX), whereas a cluster of homozygous of the other allele (YY) located vertically toward to Y axis. The cluster located in between indicated heterozygotes (XY) (Figure 2.2).

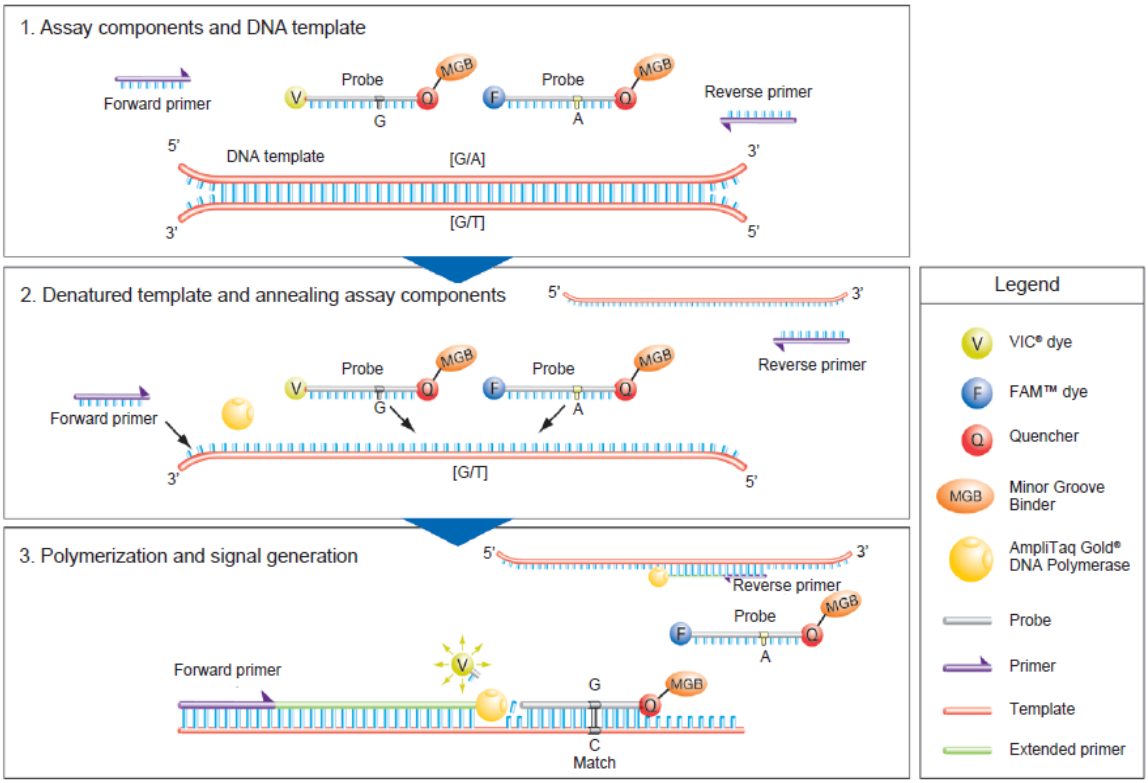


Figure 2.1: Allelic discrimination is achieved by the selective annealing of TaqMan® MGB probes extracted from

(https://products.appliedbiosystems.com/ab/en/US/adirect/abcmd=catNavigate2&catID=601283&tab=Literature).

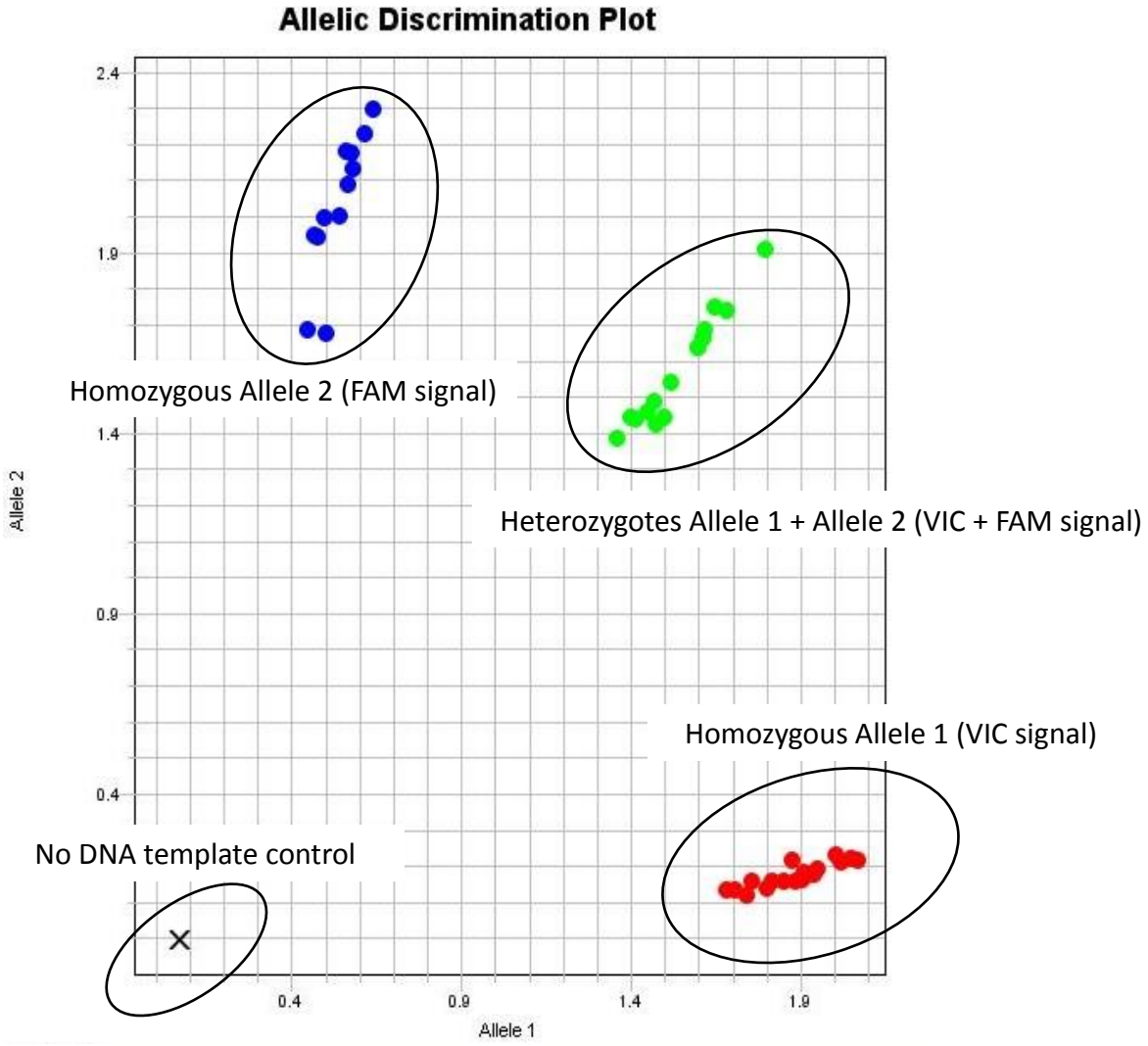


Figure 2.2: Example of an allelic discrimination plot of TaqMan SNP genotyping assay.

2.3.5 Statistical analysis

Statistical analyses were performed using SPSS version 22.0 (IBM, USA) and PLINK version 1.07 (Purcell *et al.*, 2007) (via the gPLINK version 2.050 GUI) to collate and analyze cohort phenotype data. Continuous variables were analyzed by independent two samples t-test and presented as means and standard deviations. Categorical variables were analysed by Chi-squared test unless otherwise stated. Multivariate logistic regression analysis was conducted incorporating biologically relevant covariates that were associated with risk of NAFLD development (age, gender, BMI and presence of T2DM) to test the genetic association. An additive genetic model best fitted the data and was reported. Results were expressed as beta β ± SEM for continuous parameters or odds ratio (OR) for categorical variables, with 95% confidence intervals (CI) as appropriate. Significance was taken as *P* value < 0.05 throughout.

2.4 Tissue culture

2.4.1 Cell culture conditions

All cell lines were cultured in 75cm^2 sterile filtered flasks (*Greiner Bio-One*, Stonehouse, UK), and were incubated at 37°C in an atmosphere with 5% humidified CO_2. Subculture was performed when cells reached to 80-90% confluence with cells normally were split into 1:3 ratio. All the solutions for cell culture were filter-sterilized by 0.22 μm filter (Merck Millipore, Hertfordshire, UK). When harvesting cultured cells, cells were washed by 1X PBS (Lonza, Slough, UK; Cat. No BE17-516F) once and were incubated with 1X trypsin for 10 min for cells detachment from flasks. All suspension was transferred into a 15 ml sterile tube and the sample was centrifuged twice (1st: fresh medium to deactivate trypsinization; 2nd: PBS, to wash the cells again) at 400G for 5 min. After centrifugation, cultured cell pellets were ready for further experimental purpose. For frozen stock purpose, cell pellets were re-suspended in freezing medium (5% (v/v) dimethyl sulfoxide (Sigma Aldrich) in filtered FBS), kept in freezing container (Nalgene, Thermo Scientific) and stored in liquid nitrogen.

2.4.2 DNA/RNA isolation from cultured cells

2.4.2.1 *DNA extraction*

DNA extraction was performed by using the DNeasy blood and tissue kit from QIAGEN, UK (Cat.no 69504) according to the manufacturer's instructions. Cells were detached by trypsin, then transferred into a 15 ml sterile tube and centrifuged twice at 400G for 5 min (1st: fresh medium to deactivate trypsinization; 2nd: PBS, to wash the cells again). The supernatant (PBS) of the final cell pellet was discarded and the pellet was re-suspended in 200 µl fresh PBS. The pellet was further mixed with 20 µl of proteinase K and 200 µl of Buffer AL, following a thorough mix by vortexing and an incubation at 56°C water bath for 10 min. The mixture was again thoroughly mixed with ethanol (96-100%) by vortexing.

In the following steps, the mixture underwent centrifugation four times. Firstly, the mixture was transferred into the DNeasy mini spin column in a 2 ml collection tube (provided in the kit) and was centrifuged at 6000G for 1 min. The flow-through and the collection tube were discarded.

The DNeasy mini spin column was placed in a new 2 ml collection tube, the Buffer AW1 (500 µl) was added to the mixture followed by the second centrifugation at 6,000G for 1 min. The flow-through and the bottom tube were discarded.

The third time, centrifugation was performed after adding 500 µl Buffer AW2 to the mixture and centrifuged at 20,000G to dry the DNeasy membrane, the flow-through and the bottom tube was discarded. For the last centrifugation, the DNeasy mini spin column was placed in a clean 1.5 ml microfuge tube (not provided in the kit) and mixed with 200 µl Buffer AE directly onto the DNeasy membrane. The mixture was further incubated at room temperature for 1 min followed by the centrifugation at 6,000G for 1 min (materials in Table 2.5). The final flow-through would be the DNA itself. The concentration and the purity of extracted DNA were measured by using absorbance 260/280 and 260/230 ratios provided in Nanodrop spectrophotometer software (version: 3.5.2). The DNA product was stored at 4°C.

2.4.2.2 *RNA isolation*

Cultured cell pellet was prepared as described in section 2.4.1. The pellet was re-suspended with 1 ml TRI reagent (Sigma Aldrich, Gillingham, UK) and the liquid was transferred into a sterile microfuge tube. The tube was allowed to stand at room temperature for 5 min. Chloroform (200 µl) was then added to the mixture which underwent vigorous shaking for 15 sec to ensure all reagents were mixed thoroughly. The tube was allowed to stand at room temperature for 10 min, followed by a centrifugation at 12,000G for 15 min at 4°C to separate the mixture into three phases: a red organic phase (containing protein), an interphase (containing DNA) and a colourless supernatant (RNA).

The upper aqueous phase (RNA containing) was carefully transferred into a new sterile microfuge tube, mixed with 500 µl isopropanol, and allowed to stand at room temperature for 5 min. The mixture was centrifuged at 18,000G for 10 min at 4°C to precipitate the RNA as a pellet on the side and bottom of the tube. The supernatant was discarded. The pellet was further washed by 1 ml of 75% ethanol and vortexed thoroughly, followed by a 5-min centrifugation at 8,500G at 4°C to yield the final RNA pellet. The RNA pellet was dissolved in 50 µl of RNAse-free water. Samples were quantitated and quality assessed by absorbance measurements at 260 and 280 nm using Nanodrop spectrophotometer software (version: 3.5.2). RNA was stored at -80°C for further use (materials used in Table 2.6).

2.4.2.3 *Reverse transcription and gene expression analysis*

The RNA was thawed on ice and a master mix of oligo dT primers (0.4 µg), dNTPs (10 mM each) and DEPC water (make the volume up to 8 µl) was made. RNA from the samples (1 µg) was then mixed with the above reagents (total volume of 10 µl) followed by a 10-minute incubation at 65^{o}C. After cooling on ice for 2 min, the mixture was further mixed with a second master mix containing 4 units RNAse inhibitor, 50 units of reverse transcriptase enzyme, 2 µl 10X reverse transcriptase reaction buffer, and DEPC water (total volume up to 10 µl). The RNA was reverse transcribed in a total reaction volume of 20 µl. The final mixture was incubated at 37^{o}C for 50 min and then at 70^{o}C for 15 min to inactivate any remaining enzyme. All reagents and suppliers are listed in Table 2.7. The cDNA was then used immediately for gene expression analysis or stored at -20^{o}C.

Quantitative analysis of gene expression was performed using TaqMan gene expression assays (Applied Biosystems). A 20X assay for gene of interest was delivered in a tube (250 µl) contains two unlabelled primers (1X final concentration is 900 nM) and a FAM-dye labelled on 5´ end of the MGB probe (250 nM, final concentration). Another 20X assay (250 µl) used for an endogenous housekeeping gene contains a VIC-dye labelled on 5´ end of the MGB probe (250 nM) and two unlabelled primers (150 nM, primer-limited). Each MGB probe is labelled with non-fluorescent quencher on 3´ end. 2X TaqMan universal master mix (described in section 2.3.4.2) and RNAse-free water were also applied.

A duplex gene expression analysis was performed using a 48-well plate. Each well contained a total reaction volume of 20 µl: 1 µl from the assay of target gene (FAM-dye labelled), 1 µl from the assay of housekeeping gene (VIC-dye labelled, primer-limited), 10 µl of 2X TaqMan universal master mix, 4 µl of RNAse-free water, and 4 µl of cDNA template (1-100 ng). Each sample was run in triplicate and negative control (no cDNA template) was included in each experiment. The plate was sealed and briefly centrifuged at 300 rpm for 1 min to avoid air bubbles, and was further inserted into the StepOne Real-Time PCR machine. PCR performance was described in section 2.3.4.2.

The comparative Ct (cycle threshold) method of PCR data analysis was generated by StepOne software version 2.1 (ABI-Biosystems, 1997). Basically, Ct values of different samples are directly normalised to an endogenous gene to generate ΔCt values. Next, ΔCt values of all samples are normalised to a reference sample (also known as calibrator sample) when performing relative quantification. The given equation is ΔΔCt = average ΔCt (sample of interest) – average ΔCt (reference sample).

Table 2.5: DNeasy Blood & Tissue Kit contents abstracted from the QIAGEN manual provided.

DNeasy Blood & Tissue Kit	(50)	(250)
Catalogue Number	69504	69506
DNeasy mini spin columns (colourless) in 2 ml collection tubes	50	250
Collection tubes (2 ml)	100	500
Buffer ATL	10 ml	50 ml
Buffer AL	12 ml	54 ml
Buffer AW1 (concentrate)	19 ml	95 ml
Buffer AW2 (concentrate)	13 ml	66 ml
Buffer AE	22 ml	2 x 60 ml
Proteinase K	1.25 ml	6 ml

Table 2.6: Materials used for RNA isolation.

Name	Supplier	Cat. No./Lot No.
Tri-reagent (100 ml)	Sigma Aldrich, Gillingham, UK	T9424
Chloroform (2.5 L)	Fisher Scientific, Loughborough, UK	Code: C/4920/17, Lot:1203452
Propan-2-ol (1 L)	Fisher Scientific, Loughborough, UK	Code: P/7500/15, Lot:1202163
Ethanol, absolute	Fisher Scientific, Loughborough, UK	Code: E/0650DF/17, Lot:1349671
RNAse-free water (1 L)	Fisher Scientific, Loughborough, UK	Code: 10245203, Lot:120135

Table 2.7: Reagents used for reverse transcription.

Name	Supplier	Concentration	Cat. No.
Oligo dT primers	QIAGEN, Crawley, UK	0.4 µg/µl	79237
dNTPs	QIAGEN, Crawley, UK	10 mM each	201901
DEPC water	Fisher Scientific, Loughborough, UK	-	BP5611
Reverse transcriptase Reaction Buffer	New England Biolabs, Hitchin, UK	10X	B0253S
RNase inhibitor	New England Biolabs, Hitchin, UK	40,000 units/ml	M0307S
Reverse transcriptase enzyme	New England Biolabs, Hitchin, UK	200,000 units/ml	M0253S

2.4.3 Western-blot analysis

The procedure of Western blotting consists of several steps: protein extraction, protein concentration determination, polyacrylamide gel electrophoresis, electrophoretic transfer of protein, blocking, probing and immunodetection of protein blots, and membrane staining. Materials used in these procedures are listed in Table 2.8.

2.4.3.1 ***Protein extraction***

A refrigerated centrifuge was precooled to 4^{o}C. A cultured cell pellet was prepared as described in section 2.4.1. The cell pellet was washed in 1X PBS (Lonza) and transferred into a microfuge tube followed by a 5-min centrifugation at 400G. The supernatant was discarded. The cell pellet was re-suspended in chilled CelLytic M buffer (Sigma Aldrich) with protease inhibitor (Roche, Welwyn Garden City, UK). The mixture was further vortexed moderately till fully dissolved and was left on ice for 10 min. After centrifugation was performed at 18,000G for 10 min at 4^{o}C to remove cell debris, the supernatant (the protein extracts) was transferred to a new microfuge tube without disturbing the pellet. The protein extracts were ready to be used for Western blotting or frozen at -80^{o}C long-term.

2.4.3.2 ***Protein concentration determination***

Concentrations of solubilized protein were determined using Coomassie Plus Assay Kit (Thermo Scientific, UK) based on the method of Bradford (Bradford, 1976). 10 µL of protein sample was added to 250 µl of Coomassie Plus Assay Reagent solution in each well of a single 96-well plate, mixed well on a plate shaker with 30 sec and incubated at room temperature for 10 min. The absorbance was measured at 595 nm. A standard curve was created using bovine serum albumin (BSA stock, 2 mg/ml) provided in the Coomassie Plus (Bradford) Assay Kit, diluted in PBS with working range from 0 to 2000 µg/ml protein. All standard curves and unknown samples were performed in triplicate.

2.4.3.3 *Polyacrylamide gel electrophoresis of protein samples*

Electrophoresis was carried out on a 12.5% polyacrylamide mini-gels containing 30% acrylamide, 0.378 M Tris/HCl (pH 8.8), 0.1% (w/v) SDS, 0.1% (w/v) ammonium persulphate and 0.05% TEMED. The separating gel was poured between glass plates (100 mm x 100 mm x 0.75 mm) to within 40 mm of the top. A few drops of butanol was applied on the top of the gel to ensure the gel was flatten. The gel was allowed to polymerise for 30 min. The butanol was gently removed by water. Next, the stacking gel solution (30% acrylamide, 0.126 M Tris/HCl (pH 6.8), 0.1% (w/v) SDS, 0.1% (w/v) ammonium persulphate and 0.05% TEMED) was added onto the polymerized separating gel, and a comb to form wells was inserted immediately.

All the samples were mixed with 2X or 4X sample loading buffer (Table 2.2) to reach a final dilution of 1X (62.5 mM Tris-HCl pH 6.8, 2% SDS, 10% glycerol, 0.002% bromophenol blue, and 5% β-mercaptoethanol) at an equivalent volume (20 µl). The mixtures were then heated at 95^{o}C for 5 min, and immediately applied to a 12.5% SDS-PAGE gel for size-fractionation. Electrophoresis was performed at 40 mA for 50 min in running buffer (25 mM Tris, 0.2 M glycine, 0.1% SDS).

2.4.3.4 *Electrophoretic blotting procedure of proteins separated by SDS-PAGE*

Electrophoretic transfer of protein from polyacrylamide gel to polyvinylidene fluoride (PVDF) membrane was carried out based on a method described by (Towbin *et al.*, 1979), using an Amersham Biosciences TE 22 Mighty Small Transphor. The physical assembly of electrophoretic blotting sandwich is shown in Figure 2.3. A sheet of PVDF (0.45 µm pore size) is pre-wetted with methanol briefly and soaked in transfer buffer (1X CAPS (3-(Cyclohexylamino)-1-propanesulfonic acid) buffer: 0.01 M CAPS, 10% (v/v) methanol and 80% (v/v) H_2O). The membrane was laid on two sheets of thick filter paper with a foam sponge underneath with the sponge supported by a stiff plastic grid. The gel to be blotted was removed carefully from the gel electrophoresis tank and layered onto the membrane, avoiding air bubbles. A second set of filter papers, foam sponge and a stiff plastic grid in the same order as described above was applied to the top of the gel. All items were kept submerged with transfer buffer to avoid any dryness. The blotting sandwich was inserted into the support holder of the transfer tank with the orientation from gel side/cathode (-) to membrane

side/anode (+). Electrophoretic blotting was conducted in the same transfer buffer at 50 mA (overnight) or 250 mA (one hour) alternatively.

2.4.3.5 *Blocking, probing and immunodetection of protein blots*

The membranes were rinsed by PBST buffer (1X PBS and 0.1% Tween 20) twice. An incubation of membranes in blocking solution (5% dried skimmed milk in PBST buffer) was performed at room temperature for 1 hour to achieve blocking of non-specific binding by the primary antibody. After the primary antibody and anti-GAPDH (1:25000) were diluted in the blocking solution, the membranes were incubated with the blocking solution at 4^{o}C overnight on a shaker. On the following morning, membranes were washed four times for 5 min each in PBST buffer. The goat anti-rabbit IgG conjugated to horseradish peroxidase was incubated with membranes in a dilution of 1:5000 for 1 hour at room temperature with shaking. Membranes were subjected to four washes in PBST buffer. Immunoblots were visualised using Super Signal West Pico chemiluminescent substrate (Thermo scientific) and blue X-ray film (CL-Xposure film, Thermo scientific). To measure protein expression levels, intensities of specific bands, and corresponding to the proteins of interest, densitometric analyses of the blots were performed using a GS-800 calibrated densitometer (Bio-Rad), and signal intensities were analysed using Quantity One – 4.2.3 software (GraphPad Prism version 5.0a, San Diego, USA). Images of the blots were imported into the software, the area around each band was selected by drawing a tight boundary around them. The signal intensities of bands were then displayed in an excel format which can be exported for statistical analyses, graphic representation as mean ± S.E.M.

2.4.3.6 *Staining of the PVDF membrane*

Membranes were stained by 0.05% copper solution for 1 min on a shaker. Next, membranes were rinsed three times in 10 mM HCl to remove unbound dye. Membranes were left to dry at room temperature and stored as a loading reference.

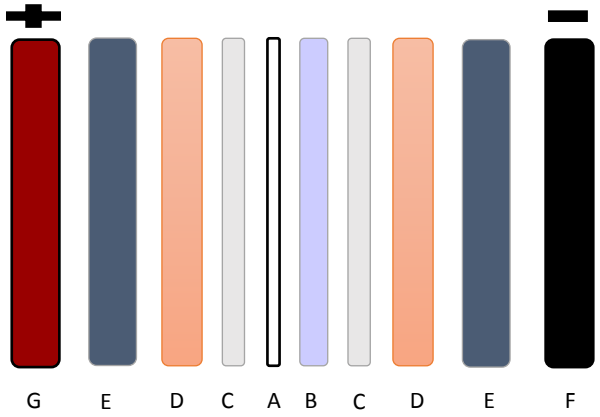


Figure 2.3: The assembly of electrophoretic blotting sandwich. A) PVDF membrane, B) Polyacrylamide gel, C) two sheets of thick filter paper, D) foam sponges, E) stiff plastic grids, F)cathode (-), G) anode (+).

Table 2.8: Materials used for Western blotting.

Protein extraction	**Supplier**	**Cat. No.**
CelLytic M buffer Lysis buffer	Sigma Aldrich, Gillingham, UK	C2978
Protease inhibitor	Roche, Welwyn Garden City, UK	11 836 170 001
Protein measurement	**Supplier**	**Cat. No.**
Pierce Coomassie Plus Assay and	Thermo Scientific, Loughborough, UK	23236
SDS-PAGE	**Supplier**	**Cat. No.**
30% Acrylamide	Sigma Aldrich, Gillingham, UK	A3699
Ammonium persulphate	Sigma Aldrich, Gillingham, UK	A3678
SDS	Sigma Aldrich, Gillingham, UK	L3771
Tris-base	Sigma Aldrich, Gillingham, UK	T1503
TEMED	Sigma Aldrich, Gillingham, UK	T9281
Butanol	BDH Laboratory Supplies, Leicestershire, UK	275006E
Electrophoretic transfer	**Supplier**	**Cat. No.**
TE 22 Mighty Small Transphor	Amersham Biosciences, Buckinghamshire, UK	80-6204-26
PVDF membrane	Merck Millipore, Hertfordshire, UK	IPVH00010
CAPS	Sigma Aldrich, Gillingham, UK	C2632
Glycine	Sigma Aldrich, Gillingham, UK	G8898
Blocking, probing and immunodetection	**Supplier**	**Cat. No.**
Methanol	VWR Chemicals, Leicestershire, UK	67-56-1
10X PBS	Lonza, Slough, UK	BE17-517Q
Tween 20	Sigma Aldrich, Gillingham, UK	P2287
Rabbit polyclonal anti-PNPLA3	Abcam, Cambridge, UK	ab81874
Rabbit polyclonal anti-GAPDH	Santa Cruz Biotechnology, Inc.	sc-25778
Horseradish peroxidase conjugated secondary antibody	Sigma Aldrich, Gillingham, UK	A6154
SuperSignal West Pico chemiluminescence	Thermo Scientific, Loughborough, UK	23236
CL-Xposure film	Thermo scientific, Loughborough, UK	34089

Chapter 3. Candidate Gene Association Study between *PNPLA3* and Severity of NAFLD in FLIP Cohort

3.1 Introduction

The basis for the effect of *PNPLA3* on genetic susceptibility to NAFLD up to the present has been described in detail in Chapter 1. Given the numerous gene association studies that had been performed for *PNPLA3*, and that its association has already been extensively replicated worldwide, there are limitations to be mentioned. The majority of genome-wide association studies (GWAS) reported to date are radiology-based but the diagnostic gold standard that can reliably assess NAFLD severity is liver biopsy. Although a few GWAS and candidate gene association studies reporting the *PNPLA3* association were histology-based, many have weaknesses in study design and have very limited cohort size. Lack of statistical power due to relatively small sample size could easily produce misinterpretation of both false positive and false negative results. To validate the genetic correlation between the *PNPLA3* gene and severity of NAFLD and to perform a better and more reliable genetic analysis, a large unrelated cohort of European-Caucasian descent (n=1,005) with histologically characterized NAFLD (representing the full spectrum of disease from simple steatosis through steatohepatitis to advanced fibrosis and cirrhosis) was assembled from centres across Europe by the FLIP consortium. This cohort was analysed by a GWAS using the Illumina OmniExpress chip platform however the *PNPLA3* rs738409 SNP, that is considered to be the causative variation, is not represented on this chip so direct genotyping was needed.

In addition to performing a candidate-gene association study upon the major SNP rs738409 in the *PNPLA3* gene using the FLIP cohort, another SNP rs139051 was also tested. This polymorphism has been reported to be associated with obesity and NAFLD (Johansson *et al.*, 2009; Peng *et al.*, 2012). A Swedish study with a total number of 466 obese children and adolescents compared to 491 non-obese reference group first demonstrated that carriage of the minor allele (A) of rs139051 was linked to obesity (P=0.014) but this effect was attenuated after adjustment for age and gender (P>0.05) (Johansson *et al.*, 2009). The other study was performed in a Han Chinese population with 553 ultrasonography diagnosed NAFLD patients (Peng *et al.*, 2012). When a dominant model was adapted, Peng *et al.* reported that the rs139051 homozygous mutant (AA) carriers were significantly associated with

increased risk of NAFLD in a multivariate analyses including gender, age, body mass index and other clinical characteristics (Peng *et al.*, 2012).

Little is known about this rs135091 polymorphism. This variant is an intronic polymorphism, located in intron 2 of *PNPLA3* causing a single nucleotide substitution from guanine (G) to adenine (A). The rs135091 SNP is 51bp upstream of rs738409 and 28bp away from the intron-exon boundary. A recent study using a locus-wide expression quantitative trait (eQTL) approach to assess SNPs that might be relevant to transcriptional regulation of *PNPLA3* conducted by a collaborator, Prof. Wanqing Liu (Purdue University, USA), has suggested that rs135091 influences hepatic *PNPLA3* expression with the A variant associated with increased gene expression (Liu *et al.*, 2016a). HapMap data (CEU, Northern Europeans in Utah) indicates that rs738409 and rs135091 are in relatively low linkage disequilibrium, the D' (coefficient of linkage disequilibrium) value is 0.61.

The aims of the studies described in this chapter were:

- Firstly, to validate the *PNPLA3* rs738409 association with histologically characterised NAFLD in the FLIP cohort.
- Secondly, to assess the relationship between the *PNPLA3* rs139051 polymorphism and severity of NAFLD in view of the recent eQTL result.
- Thirdly, to also investigate the relevance of combined genotypes for rs738409/rs139051 to severity.

3.2 Methods

3.2.1 Patients

A large cohort of unrelated patients of European-Caucasian descent with biopsy-proven NAFLD was recruited from centres across Europe. These patients were enrolled in a FP7 project (FLIP, Fatty liver: Inhibition of Progression) funded by European Commission. The centers that involved in FLIP cohort recruitment are described in section 2.1. Baseline characteristics of the cohort are shown in Table 3.1.

3.2.2 Liver biopsy

Liver biopsy performance is described in section 2.2. Liver biopsies for Newcastle samples were reviewed by two expert liver pathologists: Alastair Burt and Dina Tiniakos. Liver biopsies for other FLIP samples were performed by the collaborators at each participating centre. The criteria adapted for biopsy assessment includes NAFLD activity score (NAS) designed by Kleiner (Kleiner *et al.*, 2005), and semi-quantitative SAF score (Bedossa and Consortium, 2014; Dyson *et al.*, 2014). The histological phenotypes including degree of steatosis, stages of steatohepatitis and fibrosis were listed in Table 3.1.

3.2.3 DNA preparation from blood samples

DNA extraction from blood samples was described in section 2.3.2**.** DNA preparation from some Newcastle samples was performed by Julian Leathart and Julia Patch (Newcastle University, ICM). DNA samples from cases enrolled at other centres across Europe were prepared locally by the collaborators.

3.2.4 PNPLA3 SNPs genotyping

Assays for the two SNPs, *PNPLA3* rs738409 (reference number: C___7241_10, Cat. #4351379) and rs139051 (reference number: C_176091868_10, Cat. #4351379) were purchased from Applied Biosystems Inc., USA. Genotypes for these SNPs were determined by allelic discrimination using TaqMan reagents according to the manufacturer's protocol. Procedures were described in section 2.3.4.2**.** Control

samples of known genotype were also included in every 48-well plate (blank, homozygous wild-type, homozygous mutant and heterozygous).

3.2.5 Statistical analysis

Statistical analyses were performed using SPSS v19.0 (IBM, USA) and PLINK v1.07 (Purcell *et al.*, 2007) (via the gPLINK v2.050 GUI). Initially, univariate chi-squared analysis was performed followed by multivariate linear/logistic regression analysis incorporating biologically relevant covariates that were associated with risk of NAFLD (age, gender, BMI, and presence of T2DM) to test the genetic association. An additive genetic model best fitted the data and was used throughout this chapter. Results were expressed as beta (β ± SEM) or odds ratio (OR) with 95% confidence intervals (CI) as appropriate. $P<0.05$ was considered as the statistically significant level. Advanced haplotype analyses examining interactions between the SNPs (rs738409 and rs139051) and NAFLD-related phenotypes were performed in collaboration with Professor Heather Cordell (Institute of Genetic Medicine, Newcastle University) and Professor Wanqing Liu (Department of Medicinal Chemistry and Molecular Pharmacology, Purdue University) using the software programs R, PLINK (Purcell *et al.*, 2007) and UNPHASED (Dudbridge, 2008; Dudbridge et al., 2011).

Table 3.1: Clinical characteristics of FLIP cohort.

	Newcastle (UK)	Nottingham (UK)	Cambridge (UK)	Paris (France)	Antwerp (Belgium)	Bern (Switzerland)	OVERALL
Number	311	56	32	292	254	60	1005
Age, years	50.23 ± 12.7	51.52 ± 11.54	51.43 ± 12.32	47.71 ± 12.07	44.87 ± 12.31	52.7 ± 11.16	48.4 ± 12.48
Gender (% Female)	124 (39.9%)	24 (42.9%)	14 (43.8%)	174 (59.6%)	164 (64.6%)	20 (33.3%)	520 (51.7%)
BMI, kg/m^2	34.29 ± 5.18	30.41 ± 4.95	32.87 ± 4.79	41.77 ± 10.88	39.03 ± 6.56	31.99 ± 4.94	37.26 ± 8.44
DM (YES %)	129 (41.5%)	11 (19.6%)	14 (43.8%)	132 (45.2%)	30 (11.8%)	30 (50.0%)	346 (34.4%)
ALT	81.97 ± 59.23	90.27 ± 54.37	70.97 ± 39.07	54.28 ± 33.05	47.63 ± 23.62	86.10 ± 62.53	65.60 ± 47.34
AST$^{\#}$	53.57 ± 34.90	-	52.64 ± 36.50	38.27 ± 22.28	33.91 ± 17.01	66.37 ± 39.63	44.05 ± 29.23
Steatosis*							
S0	4 (1.3%)	1 (1.8%)	0 (0.0%)	8 (2.7%)	51 (20.1%)	0 (0.0%)	64 (6.4%)
S1	88 (28.3%)	13 (23.2%)	12 (37.5%)	63 (21.6%)	91 (35.8%)	20 (33.3%)	287 (28.6%)
S2	150 (48.2%)	20 (35.7%)	11 (34.4%)	114 (39.0%)	61 (24.0%)	27 (45.0%)	383 (38.1%)
S3	68 (21.9%)	19 (33.9%)	9 (28.1%)	107 (36.6%)	51 (20.1%)	13 (21.7%)	267 (26.6%)
Activity score (composite hepatocyte ballooning and necro-inflammation scores)*							
A0	90 (28.9%)	2 (3.6%)	5 (15.6%)	58 (19.9%)	41 (16.1%)	12 (20.0%)	208 (20.7%)
A1	57 (18.3%)	7 (12.5%)	13 (40.6%)	63 (21.6%)	42 (16.5%)	12 (20.0%)	194 (19.3%)
A2	88 (28.3%)	17 (30.4%)	6 (18.8%)	100 (34.2%)	53 (20.9%)	36 (60.0%)	300 (29.9%)
A3	48 (15.4%)	15 (26.8%)	6 (18.8%)	48 (16.4%)	80 (31.5%)	0 (0.0%)	197 (19.6%)

A4	20 (6.4%)	13 (23.2%)	2 (6.3%)	22 (7.5%)	38 (15.0%)	0 (0.0%)	95 (9.5%)
Fibrosis							
F0	109 (35.0%)	14 (25.0%)	8 (25%)	75 (25.7%)	157 (61.8%)	15 (25.0%)	378 (37.65)
F1	77 (24.8%)	9 (16.1%)	14 (43.8%)	91 (31.2%)	42 (16.5%)	9 (15.0%)	242 (24.1%)
F2	52 (16.7%)	22 (39.3%)	3 (9.4%)	77 (26.4%)	33 (13.0%)	13 (21.7%)	200 (19.9%)
F3	47 (15.1%)	5 (8.9%)	5 (15.6%)	34 (11.6%)	21 (8.3%)	5 (8.3%)	117 (11.6%)
F4 (Cirrhosis)	26 (8.4%)	6 (10.7%)	2 (6.3%)	15 (5.1%)	1 (0.4%)	18 (30.0%)	68 (6.8%)

**Steatosis and Activity score data incomplete in 4 (0.4%) and 11 (1.1%) of samples respectively. [#]Phenotypes of AST levels were not obtained by the collaborators.*

3.3 Results

3.3.1 PNPLA3 rs738409 genotype analyses

3.3.1.1 *Increased PNPLA3 rs738409 C>G minor allele carriage in NAFLD*

The *PNPLA3* rs738409 genotypes in the overall NAFLD cohort showed borderline agreement with the Hardy-Weinberg equilibrium (*P*=0.048). The minor allele frequency was 0.34 which is significantly higher than that observed in a reference Northern European population sample (MAF 0.22, http://browser.1000genomes.org). This result provides some initial evidence for an association between this variant and NAFLD generally. There was a gene-dosage effect for this variant in the cohort with the incidence of NAFLD increasing with the number of minor alleles possessed (X^2 for trend, *P*<0.0001), Table 3.2. The relationship between the rs738409 SNP and specific histological components of the NAFLD disease phenotype were assessed individually.

Table 3.2: Genotype frequency of *PNPLA3* rs738409 in FLIP NAFLD Cohort vs. the 1000 Genomes European Caucasian population.

***PNPLA3* rs738409 Genotypes**	**FLIP NAFLD n=1005 (%)**	**EUR Pop. n=379 (%)**	**OR (95% CI)**	***P*-value**
CC	455 (45.3%)	233 (61.5%)	-	-
CG	421 (41.8%)	128 (33.8%)	1.68 (1.38-2.17)	**<0.0001**
GG	129 (12.8%)	18 (4.7%)	3.67 (2.19-6.16)	**<0.0001**

ChiSq for trend p<0.0001 (X^2=36.04, df=1). 1000 Genomes MAF (EUR) = 0.22 (http://browser.1000genomes.org)

3.3.1.2 *PNPLA3 rs738409 C>G polymorphism is associated with steatosis*

The *PNPLA3* rs738409 association with hepatic steatosis diagnosed either by radiology or histology was reported by several studies (Romeo *et al.*, 2008; Valenti *et al.*, 2010a). In the FLIP cohort, carriage of the *PNPLA3* rs738409 minor allele was significantly associated with degree of steatosis in the univariate analysis (β 0.256±0.039, P=1.18x10^{-12}). This strong effect remained significant after adopting an additive model and adjusted for age at biopsy, gender, BMI and presence of T2D (β 0.308±0.038, 95%CI 0.23-0.38, P=5.4x10^{-15}). When the patients were classified into mild steatosis (S0-1) compared to advanced steatosis (S2-3), the *PNPLA3* rs738409 C>G minor allele was associated with increased risk of greater steatosis in an additive model (OR 1.87, 95%CI 1.52-2.29, P=2.98x10^{-9}), Table 3.3.

Table 3.3: Multivariate analysis of association between *PNPLA3* rs738409 genotype and histological steatosis.

Variables	Steatosis (S0-S3)		Steatosis (S0-1 vs. S2-3)	
	β (95% CI)	*P*-value	OR (95% CI)	*P*-value
rs738409 genotype	0.3 (0.23-0.38)	**5.4 x10^{-15}**	1.87 (1.52-2.29)	**2.98 x10^{-9}**
Age	-0.002 (-0.007-0.001)	0.22	0.99 (0.98-1.006)	0.29
Gender (Female)	-0.16 (-0.27- -0.05)	**0.0034**	0.51 (0.38-0.68)	**6.8 x10^{-6}**
BMI	0.018 (0.01-0.02)	**8.45 x10^{-8}**	1.05 (1.03-1.07)	**2.18 x10^{-6}**
T2DM	0.22 (0.1-0.33)	**0.0001**	1.59 (1.17-2.15)	**0.003**

3.3.1.3 *PNPLA3 rs738409 C>G polymorphism is associated with NASH*

The NAS scoring system covers the degree of steatosis (0-3), hepatocyte ballooning (0-2) and lobular inflammation (0-3), and a separate fibrosis stage (0-4) (Kleiner *et al.*, 2005). This scoring system aimed to provide a numerical order for assessing those patients who most likely to have NASH. The association with steatohepatitis activity was then tested using NAS score. The *PNPLA3* rs738409 was associated with severity of steatohepatitis in the FLIP cohort by both univariate (β 0.45±0.08, P=4.32x10-8) and multivariate analysis. The modifier effect remained strong after adjustment for gender, age at biopsy, BMI, and presence of T2DM (β 0.48±0.008, 95%CI 0.32-0.64, P=5.24x10-9). The other NASH scoring system, the validated semi-quantitative SAF score, is similar to the NAS score but separates the degree of steatosis from those of inflammation and ballooning to provide three separate domain measures of NAFLD severity (Bedossa and Consortium, 2014). The association with steatohepatitis activity was secondly tested using a composite score incorporating severity of necroinflammation and ballooning hepatocyte degeneration, generating the activity score (A0-A4). Similar results were found in the second analyses using SAF score. *PNPLA3* rs738409 was associated with severity of steatohepatitis in the FLIP cohort by both univariate (β 0.172±0.06, *P*=0.004) and multivariate analysis. The modifier effect remained strong after adjustment with gender, age at biopsy, BMI, and presence of T2DM (β 0.174±0.059, 95%CI 0.056-0.291, *P*=0.0037), Table 3.4. Beta coefficients instead of odds ratio were applied due to the nature of continuous variables.

Table 3.4: Multivariate analysis of association between *PNPLA3* rs738409 genotype and histological steatohepatitis.

Variables	SAF activity score (A0-A4)		NAS (0-8)	
	β (95% CI)	*P*-value	β (95% CI)	*P*-value
***PNPLA3* genotype**	0.17 (0.05-0.29)	**0.004**	0.48 (0.32-0.64)	**5.24×10^{-9}**
Age	0.13 (0.006-0.02)	**0.0001**	0.01(0.001-0.02)	**0.02**
Gender (Female)	-0.04 (-0.21-0.12)	0.58	-0.2 (-0.43-0.02)	0.07
BMI	0.01 (0.004-0.02)	**0.004**	0.03 (0.01-0.05)	**4.41×10^{-6}**
T2DM	0.1 (-0.06-0.28)	0.22	0.33 (0.09-0.56)	**0.006**

3.3.1.4 *PNPLA3 rs738409 C>G polymorphism is associated with fibrosis*

The association with NAFLD fibrosis stage were tested using linear and binary parameters. When a linear parameter of fibrosis score (F0-4) was applied, carriage of the minor allele of *PNPLA3* rs738409 was significantly associated with stage of fibrosis; univariate (β 0.22±0.06, P=9.87x10^{-5}), and linear regression analysis with major covariates included (gender, age at biopsy, BMI, and presence of T2DM) (β 0.22±0.05, 95%CI 0.11-0.32, P=4.67x10^{-5}). When patients were subdivided into mild fibrosis (F0-1) and advanced fibrosis (F2-4), individuals with the minor allele were at higher risk of developing NAFLD-related fibrosis in an additive model (OR 1.38, 95%CI 1.14-1.66, P=0.0008). A significant result was also found in the multivariate analysis adopting an additive model adjusted for gender, age at biopsy, BMI, and presence of T2DM (OR 1.4, 95%CI 1.15-1.7, P=0.0006), Table 3.5.

Table 3.5: Multivariate analysis of association between *PNPLA3* rs738409 genotype and histological fibrosis.

Variables	Fibrosis (F0-F4)		Fibrosis (F0-1 VS. F2-4)	
	β (95% CI)	*P*-value	OR (95% CI)	*P*-value
***PNPLA3* genotype**	0.3 (0.11-0.32)	**4.67 x10^{-5}**	1.41 (1.16-1.70)	**0.0005**
Age	-0.02 (0.01-0.02)	**1.99 x10^{-11}**	1.03 (1.01-1.03)	**2.01 x10^{-5}**
Gender (Female)	-0.26 (-0.41~-0.1)	**0.0006**	0.67 (0.51-0.89)	**0.007**
BMI	0.02 (0.01-0.03)	**1.06 x10^{-5}**	1.03 (1.02-1.05)	**1.42 x10^{-5}**
T2DM	0.58 (0.4-0.73)	**2.7 x10^{-12}**	2.27 (1.70-3.00)	**1.77 x10^{-8}**

3.3.2 PNPLA3 rs139051 genotype analyses

3.3.2.1 *PNPLA3 rs139051 genotype distributions between NAFLD patients and general population*

The *PNPLA3* rs139051 genotypes were confirmed to be in Hardy-Weinberg equilibrium (*P*=0.08) with a minor allele frequency of 0.40. This MAF is slightly higher than that observed in the same reference population as mentioned in section 3.3.1.1 (MAF 0.38, http://browser.1000genomes.org). However, the genotype distributions of *PNPLA3* rs139051 in the FLIP cohort and Northern European population were very similar, and did not show statistical significance (X^2 for trend, *P*=0.46), Table 3.6.

Table 3.6: Genotype frequency of *PNPLA3* rs139051 in FLIP NAFLD Cohort vs. the 1000 Genomes European Caucasian population.

PNPLA3 **rs139051 Genotypes**	**FLIP NAFLD n=961 (%)**	**EUR Pop. n=503 (%)**	**OR (95% CI)**	***P*-value**
GG	330 (34.3%)	184 (36.6%)	-	-
AG	489 (50.9%)	247 (49.1%)	1.10 (0.87-1.39)	**0.43**
AA	142 (14.7%)	72 (14.3%)	1.10 (0.78-1.54)	**0.31**

ChiSq for trend p=0.47 (X2=0.53, df=1). 1000 Genomes MAF (EUR) = 0.38 (http://browser.1000genomes.org)

3.3.2.2 *PNPLA3 rs139051 G>A polymorphism and histological NAFLD severity*

To investigate whether rs139051 contributes to NAFLD risk, the relationship between rs139051 and NAFLD phenotypes in FLIP cohort was then tested. A significant effect was observed between carriage of rs139051 G>A polymorphism and severity of NAFLD, including steatosis (β 0.13±0.04, *P*=0.003), and steatohepatitis (*P*=0.02 and *P*=0.002 for SAF score and NAS score, respectively) but not fibrosis (*P*=0.8) in the univariate analyses. These strong genetic modifier effects on hepatic steatosis and steatohepatitis remained significant in the multivariate analyses when adjusted for age at biopsy, gender, BMI and presence of T2D; $P=2.0\times10^{-4}$ and *P*=0.02 (for both SAF and NAS scoring systems), respectively. The association with stage of fibrosis remained non-significant after testing in a multivariate analysis (*P*=0.6).

Since the *PNPLA3* rs738409 SNP is a well-known and well-replicated genetic modifier for NAFLD, another sets of multivariate analyses that included gender, age at biopsy, BMI, T2DM, and *PNPLA3* rs738409 genotype were performed. Interestingly, the significant effects described above were completely abolished when the *PNPLA3* rs738409 genotype was taken into account (P>0.05 for all phenotypes), suggesting that rs139051 is not independent of rs738409 as a risk factor for NAFLD, Table 3.7.

3.3.2.3 *Haplotype analyses of PNPLA3 SNPs in FLIP cohort*

As previously identified by our collaborator, Prof. Wanqing Liu, during his eQTL studies, rs139051 (51bp upstream of rs738409 in intron 2, and 28bp away from the intron-exon boundary) was the most significant *cis*-acting eQTL for *PNPLA3* transcription [$P=6.6\times10^{-8}$, false discovery rate (FDR) < 0.05] (Liu *et al.*, 2016a). Detailed LD analysis of variants across the entire locus revealed that rs139051 was in low LD with all other significant eQTLs ($r^2 \leq 0.50$ for all tests). The data indicated that the minor (A) allele of rs139051 was associated with increased *PNPLA3* gene expression. In contrast, after correcting for multiple testing, the rs738409 polymorphism was not a significant eQTL for hepatic *PNPLA3* expression (FDR > 0.05) (Liu *et al.*, 2016a).

To investigate further whether rs139051 modifies the effect of rs738409 alleles in contributing to NAFLD risk, we tested the association between the rs139051-rs738409 haplotype and aforementioned phenotypes. However, analyses examining the relative contribution of the SNP to the haplotype effect did not identify a statistically significant independent effect of rs139051 in the FLIP population. Indeed, the carriage of the A-G haplotype which combines both the high-expression rs139051 A allele and the disease-risk rs738409 G (148M) allele, conferred similar risk for NAFLD phenotypes as compared to the G-G haplotype which contains the rs738409 risk allele, but possesses a relatively lower expression level (P=0.51), Table 3.8 (Liu *et al.*, 2016a).

Table 3.7: Multivariate analysis of association between *PNPLA3* rs139051 genotype and histological NAFLD.

Variables	Steatosis (S0-S3)		Steatosis (S0-1 vs. S2-3)	
Steatosis	**β (95% CI)**	***P*-value**	**OR (95% CI)**	***P*-value**
rs139051 genotype	0.02 (-0.06-0.21)	0.6	0.98 (0.78-1.23)	0.84
Age	-0.002 (-0.006-0.002)	0.35	0.99 (0.98-1.007)	0.39
Gender (Female)	-0.15 (-0.26~-0.04)	**0.009**	0.52 (0.39-0.70)	**1.64×10^{-5}**
BMI	0.018 (0.012-0.026)	**8.93×10^{-8}**	1.05 (1.03-1.07)	**2.23×10^{-6}**
T2DM	0.23 (0.11-0.35)	**1.4×10^{-4}**	1.62 (1.15-2.21)	**0.003**
rs738409 genotype	0.29 (0.21-0.38)	**4.18×10^{-11}**	1.87 (1.48-2.37)	**1.28×10^{-7}**
Variables	**SAF activity score (A0-A4)**		**NAS (0-8)**	
Steatohepatitis	**β (95% CI)**	***P*-value**	**β (95% CI)**	***P*-value**
rs139051 genotype	0.08 (-0.05-0.22)	0.22	0.1 (-0.08-0.28)	0.29
Age	0.014 (0.006-0.02)	**1.4×10^{-4}**	0.01 (0.002-0.02)	**0.021**
Gender (Female)	-0.06 (-0.23-0.11)	0.48	-0.21 (-0.44-0.02)	0.08
BMI	0.013 (0.002-0.02)	**0.013**	0.03 (0.02-0.05)	**1.52×10^{-5}**
T2DM	0.08 (-0.09-0.26)	0.36	0.03 (0.002-0.002)	**0.01**
rs738409 genotype	0.13 (0.004-0.27)	**0.042**	0.43 (0.25-0.61)	**2.93×10^{-6}**
Variables	**Fibrosis (F0-F4)**		**Fibrosis (F0-1 vs. F2-4)**	
Fibrosis	**β (95% CI)**	***P*-value**	**OR (95% CI)**	***P*-value**
rs139051 genotype	-0.08 (-0.2-0.04)	0.18	0.87 (0.69-1.01)	0.25
Age	0.02 (0.01-0.03)	**4.89×10^{-12}**	1.03 (1.02-1.04)	**7.64×10^{-6}**
Gender (Female)	-0.27 (-0.42~-0.12)	**5.56×10^{-4}**	0.66 (0.49-0.88)	**0.005**
BMI	0.02 (0.01-0.03)	**2.92×10^{-5}**	1.04 (1.02-1.06)	**2.38×10^{-6}**
T2DM	0.54 (0.38-0.71)	**1.10×10^{-10}**	2.09 (1.56-2.80)	**6.56×10^{-7}**
rs738409 genotype	-0.27 (-0.43~-0.12)	**9.78×10^{-5}**	1.49 (1.19-1.86)	**4.10×10^{-4}**

Table 3.8: Haplotype association with NAFLD histological features and effect of each haplotype on phenotype severity in the FLIP NAFLD Cohort (n=949). Adapted from (Liu *et al.*, 2016a).

Histological Phenotype*	Haplotype†	Added Value (95%CI)#	*P*#	*P**	Beta‡	*P*‡
Steatosis (0-3)	A-G	0	-	9.85×10^{-7}	0.33	3.36×10^{-11}
	G-G	-0.08 (-0.31, 0.16)	0.51	3.78×10^{-4}	0.23	0.0014
	A-C	-0.46 (-0.65, -0.28)	9.85×10^{-7}	-	-0.15	0.0108
	G-C	-0.46 (-0.60, -0.32)	1.33×10^{-11}	0.96	-0.23	3.15×10^{-8}
NAS (0-8)	A-G	0	-	4.81×10^{-4}	0.56	4.27×10^{-8}
	G-G	-0.06 (-0.17, 0.05)	0.30	0.05	0.27	0.06
	A-C	-0.16 (-0.25, -0.07)	4.81×10^{-4}	-	-0.15	0.2
	G-C	-0.18 (-0.25, -0.12)	2.66×10^{-8}	0.61	-0.4	2.45×10^{-6}
Fibrosis (0-4)	A-G	0	-	0.04	0.17	0.01
	G-G	0.11 (-0.04, 0.27)	0.15	9.64×10^{-4}	0.32	0.0013
	A-C	-0.14 (-0.28, -0.008)	0.04	-	-0.15	0.06
	G-C	-0.12 (-0.22, -0.03)	0.0089	0.74	-0.17	0.0092

†Haplotype frequencies across entire NAFLD cohort: A-G 0.2425; G-G 0.09797; A-C 0.1628; G-C 0.4968. Estimated r^2 between rs139051 and rs738409 = 0.20. #Tests for the effect of each haplotype compared to A-G. *Tests for the effect of each haplotype compared to A-C. ‡Tests for the effect of each haplotype compared to all other haplotypes combined. *NAFLD histologically characterized using the semi-quantitative NASH CRN Score (Kleiner *et al.*, 2005). The NAFLD Activity Score (NAS) equals the sum of the scores for steatosis, hepatocyte ballooning degeneration and lobular inflammation and reflects disease activity.

3.4 Discussion

Performing a case-control association study within a well-characterized population and with an appropriate sample size, we have validated the association between *PNPLA3* rs738409 and severity of NAFLD. Carriage of the minor allele (G) is profoundly associated with disease severity across the entire NAFLD disease spectrum (steatosis, steatohepatitis, and fibrosis). As described in Chapter 5, we also investigated its association with NAFLD-related HCC, the ultimate stage of NAFLD.

The other SNP rs139051 (G>A) which is 51bp upstream of rs738409 (Liu *et al.*, 2016a) has been suggested to be relevant to NAFLD susceptibility (Peng *et al.*, 2012). However, an earlier study demonstrated in obese children and adolescents provided opposite results (Johansson *et al.*, 2009). It would be inappropriate to compare these two studies directly as their study aims (obesity vs. NAFLD) and population design (children and adolescents vs. adults) were totally different, as is the variation in ethnicities. Since NAFLD is closely associated with obesity, these findings should not be ignored entirely and require further investigations to be ascertain the relevance of rs139051 as a genetic risk factor.

The rs139051 variant was identified as a strong eQTL, playing an independent role in regulating *PNPLA3* transcription in man (Liu *et al.*, 2016a). In our association analyses, a weaker but statistically significant association between rs139051 and steatosis, steatohepatitis but not fibrosis was observed. However, in the multivariate analysis with the rs738409 genotype included, the association was completely abolished ($P>0.05$ for all tests), suggesting that rs139051 is not independent of rs738409 as a risk factor for NAFLD. Furthermore, the apparent changes in hepatic *PNPLA3* expression due to rs139051 carriage do not modify the effects of rs738409 G-allele carriage in promoting NAFLD and related phenotypes. This was supported by the analysis of associations between rs139051-rs738409 haplotypes and histological features of NAFLD. While the A-G haplotype potentially represents an increased expression of the high-risk allele than the G-G haplotype, there was no significant difference between these two haplotypes in conferring risk to NAFLD or disease severity. A similar effect was observed between the A-C and G-C haplotype in conferring reduced risk for NAFLD.

As has already been discussed in section 1.7.1, a complex disease trait is a contribution of strong and weak genetic risk factors when combined together in the presence of a permissive environment (Hirschhorn and Gajdos, 2011). Taken Crohn's disease as one of the examples for polygenic disease, several strong (P<5X10^{-8} with OR ≥1.8) genetic risk factors that have been identified and replicated: particularly, *CARD15* (also known as *NOD2*, *nucleotide binding oligomerization domain containing 2*) on chromosome 16q12, *IL23R* (*interleukin 23 receptor*) on chromosome 1p31, and *ATG16L1* (*autophagy related 16 like 1*) on chromosome 2p37 are considered as the well-demonstrated SNPs associated with Crohn's disease. Unlike Crohn's disease, T2DM is best considered as a polygenic disease where multiple common variants with weak effects (P>5X10^{-4} with OR≤ 1.3) contribute to the outcomes and appear to combine in an additive manner to increase overall disease susceptibility, such as *PPARG* on chromosome 3q25 (*peroxisome proliferator activated receptor gamma*) and *KCNJ11* (*potassium voltage-gated channel subfamily J member 11*) on chromosome 11q15. It has been estimated that only 10% of the genetic risk were identified in T2DM despite there have been intensive effort to identify the genetics of T2DM in the past two to three decades (Wellcome Trust Case Control, 2007; Donaldson *et al.*, 2015).

The effect size of the *PNPLA3* rs738409 identified by the first GWAS-like study (P=5.9x 10^{-10}) (Romeo *et al.*, 2008) is even stronger than those identified in T2DM and Crohn's disease, and to date it remains the strongest and the most widely replicated genetic risk factor for NAFLD (P<5X10^{-8} with OR≥1.8). In addition, there are multiple SNPs in the chromosome 22 region (22q13.31) on the Illumina chip that were reported to be associated with NAFLD but with weaker effect sizes compared to rs738409 even in a strong LD, the rs3761472 located in the neighbouring gene *SAMM50* (*SAMM50 sorting and assembly machinery component*) would be an example, pointing that rs738409 is the true causative locus (Kawaguchi *et al.*, 2012; Kitamoto *et al.*, 2013).

In summary, the rs738409 *PNPLA3* polymorphism represents a biologically plausible candidate gene and remains the strongest genetic signal for NAFLD across the whole genome. However, the biological effect of rs738409 remains elusive as discussed in detail in section 1.7.3.6, investigation into the functional role of PNPLA3

on pathogenic mechanisms of NAFLD is therefore needed in order to possibly provide therapeutic approaches to NAFLD management in the future.

Chapter 4. *TM6SF2* rs58542926 Influences Hepatic Fibrosis Progression in Patients with Non-Alcoholic Fatty Liver Disease

4.1 Introduction

As described in section 1.7.3, genome-wide association studies (GWAS) (Romeo *et al.*, 2008; Chambers *et al.*, 2011; Speliotes *et al.*, 2011) and candidate-gene studies (Dongiovanni *et al.*, 2010b; Valenti *et al.*, 2010a; Al-Serri *et al.*, 2012; Aravinthan *et al.*, 2014; Tan *et al.*, 2014) have contributed greatly to the understanding of the genetic contribution to NAFLD pathogenesis and variability of prognosis (reviewed (Anstee and Day, 2013)). Amongst the loci identified, the non-synonymous SNP in *PNPLA3* (rs738409 c.444 C>G, p.Ile148Met), has been validated across multiple patient cohorts (Romeo *et al.*, 2008; Valenti *et al.*, 2010a; Speliotes *et al.*, 2011; Kawaguchi *et al.*, 2012). Importantly, carriage of this SNP has been robustly associated not only with steatosis, but also with clinically relevant factors including severity of hepatic fibrosis/cirrhosis (Valenti *et al.*, 2010a). However, its association with NAFLD-related HCC remains to be elucidated (Trepo *et al.*, 2014).

Recently, Kozlitina *et al*. showed that a non-synonymous SNP in *TM6SF2* (rs58542926 c.449 C>T, p.Glu167Lys (E167K)), *transmembrane 6 superfamily member 2*, a gene of unknown function on chromosome 19, was associated with ^{1}H-MRS quantified HTGC based on genotyping with a genome-wide exome chip (Kozlitina *et al.*, 2014). This variant has also been associated with dyslipidaemia and cardiovascular risk (Holmen *et al.*, 2014). The *TM6SF2* rs58542926 SNP lies within 50kb of an *NCAN* gene variant (rs2228603 c.274 C>T, p.Pro92Ser) that has previously been associated with HTGC in another GWAS (Speliotes *et al.*, 2011; Gorden *et al.*, 2013). Both SNPs are in strong linkage disequilibrium (D'=0.926, r^2=0.798). Conditioning on the *TM6SF2* variant abrogated the effect of the *NCAN* variant whilst the reverse did not occur, suggesting that *TM6SF2* rs58542926 is more strongly associated with the HTGC phenotype. Homozygote *TM6SF2* rs58542926 minor (T) allele carriage was shown to be associated with a modest but statistically significant increase in ^{1}H-MRS measured HTGC from 5.86±0.25% in CC homozygotes to 15.04±2.23% in TT homozygotes (Kozlitina *et al.*, 2014).

Limited knowledge has been established for the protein structure or the functional role of *TM6SF2*. The *TM6SF2* rs58542926 c.449 C>T variant is a non-synonymous change producing a glutamate to lysine amino acid substitution at residue 167 (Glu167Lys), which is highly conserved across mammals (Kozlitina *et al.*, 2014). First descried as a multi-pass membrane protein in 2000 (Carim-Todd *et al.*, 2000),

TM6SF2 was later discovered to be highly expressed in liver, kidney and intestines (Kozlitina *et al.*, 2014; Surakka *et al.*, 2015). *In vitro* study using confocal microscopy, GFP-tagged TM6SF2 was mainly localized to the endoplasmic reticulum (ER) and ER-Golgi intermediate compartments (ERGIC) (Mahdessian *et al.*, 2014). In the same study, *TM6SF2* siRNA inhibition led to a reduction of lipid secretion (triglycerides (TG) and Apolipoprotein B (ApoB)) and an increase in the number and overall size of lipid droplets, which represented as a manifestation of cellular triglyceride accumulation. On the contrary, overexpression of TM6SF2 caused a decrease in the lipid droplet contents (Mahdessian *et al.*, 2014). Prediction of mouse knockout phenotype by analysis of coexpressed gene profiles based on Mouse Genome Informatics (MGI) Database, TM6SF2 is linked to abnormal lipids levels (decreased Total cholesterol (TC), low-density lipoprotein-cholesterol (LDL-C) and very-low-density lipoprotein (VLDL)) and may act as lipid transporter and interact with proteins involved in intestinal absorption (Surakka *et al.*, 2015). *In vivo* studies may provide more clues for above statements, adenovirus-mediated short hairpin RNA knockdown of *Tm6sf2* in mice has been shown to increased hepatic triglyceride content and reduced plasma cholesterol and VLDL secretion, suggesting that *TM6SF2* activity is necessary for normal VLDL secretion and that impaired *TM6SF2* function causally contributes to NAFLD (Kozlitina *et al.*, 2014). On the other hand, transiently overexpressed of *Tm6sf2* in C57BL/6J mice resulted in a significant increase of serum TC, LDL-C and TG. These *in vivo* (Holmen et al., 2014; Kozlitina et al., 2014) and *in vitr*o (Holmen *et al.*, 2014; Kozlitina *et al.*, 2014; Mahdessian *et al.*, 2014) functional studies indicate that *TM6SF2* has gene effects on fat retention in the liver, lipid efflux and alteration in the number and average size of lipid droplets. However, the genetic and biological studies previously mentioned only focused on the scope of hepatic steatosis. Whether the effect of *TM6SF2* was limited to steatosis or had broader clinical relevance, possessing the susceptibility to steatohepatitis or fibrosis as has already been shown for *PNPLA3,* remains undetermined.

The aim of the current study was firstly to determine whether the association with NAFLD reported by Kozlitina *et al.* (Kozlitina *et al.*, 2014) could be independently validated; and secondly to establish whether the *TM6SF2* rs58542926 variant was

associated with clinically important disease endpoints that have prognostic relevance (in particular stage of hepatic fibrosis).

To address this, I performed a quantitative analysis within a well-characterised European Caucasian *'discovery' cohort* with histologically characterised NAFLD, controlling for relevant co-morbidities and factors that have previously been linked with disease progression (age, gender, BMI, presence of T2DM and *PNPLA3* rs738409 genotype), and replicated our findings in a separate histologically characterised European Caucasian *'validation' cohort*. The association between this SNP and development of NAFLD-related HCC was also examined (described in Chapter 5).

4.2 Methods

4.2.1 Patients

Patients were recruited from hepatology clinics at several European specialist centres, including UK, Switzerland, Belgium and France (details are listed in Table 4.1). The study had all the necessary ethical approvals (Table 4.1). All participants gave informed consent. Criteria for NAFLD diagnosis and patients' clinical and laboratory data are described in section 2.1.

- An initial *discovery cohort* of 349 consecutive European Caucasian patients from the UK with histologically characterised NAFLD of different stages of disease. These were unrelated patients with histologically characterised NAFLD, derived from a patient population originally identified as having ultrasonographically detected bright liver and abnormal biochemical tests (ALT and/or GGT).

- A *validation cohort* of 725 consecutive European Caucasian patients from centres in UK and mainland Europe with histologically characterised NAFLD of different stages of disease. Patients in this cohort were unrelated patients with histologically characterised NAFLD, derived from a patient population originally identified as having ultrasonographically detected bright liver and abnormal biochemical tests (ALT and/or GGT) or identified as having evidence of NAFLD at the time of bariatric surgery.

Together, these comprised the *combined cohort* of 1,074 patients with histologically characterised NAFLD. Demographic and histological details are shown in Table 4.2. A description of the 'healthy workers' cohort recruited in the North East of the UK has previously been published (Velaga *et al.*, 2004).

4.2.2 Liver biopsy

Liver biopsies for Newcastle samples were reviewed by two expert liver pathologists: Alastair Burt and Dina Tiniakos. Liver biopsies for FLIP samples were performed by the collaborators at each participating centre. Liver biopsy performance is written in section 2.2. The degree of steatosis (S0-3), activity of steatohepatitis (A0-4) and

stage of fibrosis (F0-4) were scored according to the validated semi-quantitative SAF score (Bedossa and Consortium, 2014; Dyson *et al.*, 2014), Table 1.1.

4.2.3 DNA preparation from blood samples

DNA preparation from blood samples were descried in section 2.3.2. Some DNA preparation on Newcastle samples was performed by Julian Leathart and Julia Patch (Newcastle University, ICM). DNA isolation from FLIP samples was performed locally by the collaborators.

4.2.4 TaqMan SNP genotyping assays

The TaqMan SNP genotyping assays for *TM6SF2* rs58542926 (reference number: C___8709053_10, Cat. # 4351379), *NCAN* rs2228603 (reference number: C___16171492_10, Cat. # 4351379) and *PNPLA3* rs738409 (reference number: C___7241_10, Cat. # 4351379) were purchased from Applied Biosystems Inc., USA. Genotypes for these SNPs were determined by allelic discrimination using TaqMan reagents according to the manufacturer's protocol. Procedures were described in section 2.3.4.2. Control samples of known genotype were also included in every 48-well plate (blank, homozygous wild-type, homozygous mutant and heterozygous).

4.2.5 Statistical analysis

Statistical analyses were performed using SPSS v19.0 (IBM, USA) to collate and to analyze cohort phenotype data. Continuous variables were analysed using Student's t-test/one-way ANOVA and categorical variables were analysed by Chi-squared test unless otherwise stated. PLINK v1.07 (Purcell *et al.*, 2007) (via the gPLINK v2.050 GUI) was used to conduct the genetic analysis. An initial univariate chi-squared analysis was performed. Subsequently, multivariate logistic regression analysis was conducted incorporating biologically relevant covariates that were associated with risk of NAFLD (age, gender, BMI, presence of T2DM and *PNPLA3* rs738409 genotype) to test the genetic association. An additive genetic model best fitted the data and was reported. Results were expressed as beta β ± SEM or odds ratio (OR)

with 95% confidence intervals (CI) as appropriate. *P*<0.05 was considered as statistically significant level.

Table 4.1: Patients were recruited from hepatology clinics across Europe.

Recruitment Centre	Country	Ethical Approval
Freeman Hospital	Newcastle-upon-Tyne, UK	Newcastle & North Tyneside 1 REC [10/H0906/41]
Addenbrooke's Hospital	Cambridge, UK	Norfolk REC [06/Q0106/70]
Nottingham University Hospitals NHS Trust	Nottingham, UK	Nottingham 2 REC [GM010201]
Inselspital Hospital	Bern, Switzerland	Inselspital Bern Local Ethics Committee
Antwerp University Hospital	Belgium	Antwerp University Hospital Ethics Committee
Pitié-Salpêtrière Hospital	Paris, France	CPP (Comité de Protection des Personnes) Paris VI IDF Pitié - Salpêtrière Hospital

Table 4.2: Demographic characteristics of patient cohorts.

	Discovery Cohort	Validation Cohort	Combined cohort
Number	349	725	1074
Ethnicity	European Caucasian	European Caucasian	European Caucasian
Gender (Female)	147 (42.1%)	407 (56.1%)	554 (51.6%)
Age, years	51 .5 ± 12.0	47.6 ± 12.4	48.9 ± 12.4
BMI, kg/m^2	34.5 ± 5.7	38.5 ± 9.1	37.2 ± 8.3
T2DM (Yes)	161 (46.1%)	235 (32.4%)	396 (36.9%)
Steatosis score*			
S0	5 (1.4%)	60 (8.3%)	65 (6.1%)
S1	99 (28.4%)	206 (28.4%)	305 (28.4%)
S2	166 (47.6%)	247 (34.1%)	413 (38.5%)
S3	79 (22.6%)	204 (28.1%)	283 (26.4%)
Activity score (composite hepatocyte ballooning and necro-inflammation scores)*			
A0	81 (23.2%)	132 (18.2%)	213 (19.8%)
A1	65 (18.6%)	133 (18.3%)	198 (18.4%)
A2	101 (28.9%)	214 (29.5%)	315 (29.3%)
A3	64 (18.3%)	149 (20.6%)	213 (19.8%)
A4	31 (8.9%)	89 (12.3%)	120 (11.2%)
Fibrosis score			
F0	108 (30.9%)	277 (38.2%)	385 (35.8%)
F1	90 (25.8%)	162 (22.3%)	252 (23.5%)
F2	55 (15.8%)	161 (22.2%)	216 (20.1%)
F3	66 (18.9%)	75 (10.3%)	141 (13.1%)
F4 (Cirrhosis)	30 (8.6%)	50 (6.9%)	80 (7.4%)

**Steatosis and Activity score data incomplete in 8 (0.7%) and 15 (1.3%) of samples respectively.*

4.3 Results

4.3.1 Increased TM6SF2 rs58542926 C>T minor allele carriage in NAFLD

In the NAFLD *discovery cohort*, the *TM6SF2* rs58542926 genotypes were confirmed to be in Hardy-Weinberg equilibrium with a minor allele frequency of 0.12, significantly higher than that observed in a reference Northern European population sample (MAF 0.07, http://browser.1000genomes.org) or a cohort of 265 Caucasian self-reported 'healthy workers' recruited from offices and factories locally in the North East of England (MAF 0.07) and so supportive of an association between these variants and NAFLD. Indeed, a gene-dosage effect was observed for both variants in the *discovery cohort* with the incidence of NAFLD increasing with the number of minor alleles possessed (X^2 for trend, P=0.0008), Table 4.3. A similar association was also confirmed for *PNPLA3* rs738409, P=0.0001 (Table 4.4). Specific histological components of the NAFLD disease phenotype were next assessed individually.

Table 4.3: Genotype frequency of *TM6SF2* rs58542926 in Discovery Cohort NAFLD patients vs. the 1000 Genomes European Caucasian population.

TM6SF2 Genotypes	NAFLD n=349 (%)	EUR Pop. n=379 (%)	OR (95% CI)	*P*-value
CC	271 (77.6%)	328 (86.5%)	-	-
CT	70 (20%)	49 (12.9%)	OR 1.72 (1.16-2.57)	**0.008**
TT	8 (2.3%)	2 (0.5%)	OR 4.84 (1.01-22.9)	**0.049**

ChiSq for trend p=0.0008 (X^2=11.34, df =1). 1000 Genomes MAF (EUR) = 0.07 (http://browser.1000genomes.org).

Table 4.4: Genotype frequency of *PNPLA3* rs738409 in Discovery Cohort NAFLD patients vs. the 1000 Genomes European Caucasian population.

PNPLA3 Genotypes	NAFLD N=349 (%)	EUR Pop. N=379 (%)	OR (95% CI)	*P*-value
CC	152 (43.5%)	233 (61.5%)	-	-
CG	148 (42.4%)	128 (33.8%)	OR 1.77 (1.29-2.42)	**0.0003**
GG	49 (14%)	18 (4.7%)	OR 4.17 (2.34-7.43)	**3.04×10^{-7}**

ChiSq for trend p<0.0001 (X^2=31.29, df =1). 1000 Genomes MAF (EUR) = 0.22 (http://browser.1000genomes.org).

4.3.2 TM6SF2 and degree of histological steatosis

As a positive control, and consistent with our previously reported analysis (Valenti *et al.*, 2010a), carriage of the *PNPLA3* rs738409 minor allele was significantly associated with degree of steatosis in multivariate analysis adopting an additive model adjusted for gender, age at biopsy, BMI and presence of T2DM (β 0.192±0.056, 95%CI 0.082-0.301, P=6.74x10^{-4}). However, in contrast to the report by Kozlitina *et al.*, neither *TM6SF2* rs58542926 (β 0.087±0.083, 95%CI -0.076-0.250, P=0.296) nor *NCAN* rs2228603 (β 0.050±0.085, 95%CI -0.116-0.216, P=0.554) were found to be significantly associated with degree of histologically determined steatosis in the 349-patient *discovery cohort*. This was also the case in the 725-patient *validation cohort* (P=0.17). However, a trend towards significance was observed when the two cohorts were combined (β 0.111±0.059, 95%CI -0.0041-0.2268, P=0.053), suggesting that an underlying effect on degree of steatosis may be present but of relatively small size. An effect became apparent when the multivariate analysis in the *combined cohort* was repeated after subdividing the cohort into those with mild steatosis (S0-1) and pronounced steatosis (S2-3). Here, carriage of each copy of the *TM6SF2* rs58542926 C>T minor allele was associated with increased risk of greater steatosis (OR 1.379, 95%CI 1.019-1.865, P=0.037), although with a marginal level of significance (Table 4.5).

4.3.3 TM6SF2 and severity of histological steatohepatitis

Next, the association with steatohepatitis activity was tested using a composite score incorporating severity of necroinflammation and ballooning hepatocyte degeneration. *TM6SF2* rs58542926, but not *NCAN* rs2228603, was associated with severity of steatohepatitis in the *discovery cohort* by multivariate analysis adopting an additive model adjusted for gender, age at biopsy, BMI, T2DM and *PNPLA3* rs738409 genotype (β 0.288±0.139, 95%CI 0.015-0.561, P=0.039). However, this effect was not replicated in the *validation* or *combined cohorts* (Table 4.6). Beta coefficients instead of odds ratio were applied due to the nature of continuous variables.

Table 4.5: Multivariate analysis of association between *TM6SF2* rs58542926 genotype and steatosis stage S0-1 (mild) vs. S2-3 (advanced).

Variables	Discovery cohort (n=349)		Validation cohort (n=725)		Combined cohort (n=1074)	
	OR (95% CI)	*P*-value	OR (95% CI)	*P*-value	OR (95% CI)	*P*-value
***TM6SF2* genotype**	1.39 (0.83-2.35)	0.202	1.32 (0.90-1.92)	0.15	**1.38 (1.01-1.86)**	**0.037**
***PNPLA3* genotype**	**1.51 (1.06-2.13)**	**0.021**	**2.02 (1.59-2.59)**	**1.66 x10^{-8}**	**1.84 (1.51-2.25)**	**1.57 x10^{-9}**
Age	**0.97 (0.95-0.99)**	**0.022**	0.99 (0.98-1.008)	0.41	0.99 (0.98-1.003)	0.137
Gender (Male)	1.02 (0.61-1.71)	0.94	**0.48 (0.34-0.68)**	**2.70 x10^{-5}**	**0.57 (0.43-0.75)**	**7.53 x10^{-5}**
BMI	1.03 (0.98-1.08)	0.174	**1.05 (1.03-1.07)**	**1.04 x10^{-6}**	**1.04 (1.02-1.06)**	**6.00 x10^{-6}**
T2DM	1.40 (0.86-2.28)	0.175	1.43 (0.99-2.09)	0.058	**1.52 (1.13-2.03)**	**5.01 x10^{-3}**

Additive model including age, gender, BMI, T2DM and *PNPLA3* rs738409 genotype as covariates.
Discovery/Validation/Combined cohorts: Stage S0-1 (mild) n=104/266/370, Stage S2-3 (advanced) n=245/451/696.

Table 4.6: Multivariate analysis of association between *TM6SF2* rs58542926 genotype and steatohepatitis stage (Activity Score A0-4).

Variables	Discovery cohort (n=349)		Validation cohort (n=725)		Combined cohort (n=1074)	
	β±SEM (95% CI)	*P*-value	β±SEM (95% CI)	*P*-value	β±SEM (95% CI)	*P*-value
***TM6SF2* genotype**	0.288±0.139 **(**0.015-0.561**)**	**0.039**	0.01±0.11 (-0.207-0.22)	0.921	0.1±0.087 (-0.067-0.276)	0.236
***PNPLA3* genotype**	0.19±0.096 (0.005-0.382)	**0.044**	0.17±0.07 (0.032-0.31)	**0.016**	0.17±0.057 (0.065-0.289)	**0.002**
Age	0.017±0.006 (0.006-0.03)	**0.0034**	0.014±0.004 (0.006-0.02)	**7.0×10^{-4}**	0.014±0.003 (0.007-0.02)	**2.87×10^{-5}**
Gender (Male)	-0.005±0.146 (-0.29-0.28)	0.972	-0.146±0.1(-0.34-0.051)	0.146	-0.049±0.08 (-0.21-0.11)	0.546
BMI	0.02±0.012 (-0.0028-0.04)	0.085	0.01±0.005 (-0.001-0.02)	0.108	0.013± 0.005(0.003-0.02)	**0.006**
T2DM	0.463±0.137 (0.194-0.733)	**0.0008**	0.079±0.11 (-0.13-0.29)	0.477	0.189± 0.086(0.01-0.358)	**0.029**

Additive model including age, gender, BMI, T2DM and *PNPLA3* rs738409 genotype as covariates.

4.3.4 TM6SF2 and stage of histological fibrosis

Finally, the association with NAFLD fibrosis stage was tested. In the *discovery cohort* multivariate analyses adopting an additive model adjusted for gender, age at biopsy, BMI, T2DM and *PNPLA3* rs738409 genotype found that *TM6SF2* rs58542926 (β 0.549±0.135, 95%CI 0.285-0.813, P=5.57x10^{-5}) and *NCAN* rs2228603 (β 0.419±0.138, 95%CI 0.148-0.689, P=0.0026) were both significantly associated with stage of fibrosis. The association between *TM6SF2* rs58542926 and fibrosis stage persisted when analysis included both the *NCAN* rs2228603 and the *PNPLA3* rs738409 SNPs as co-variates (β 0.552±0.205, 95%CI 0.151-0.953, P=0.0074). However, the association with *NCAN* rs2228603 was lost when the analysis was conditioned on rs58542926. Thus, the association is driven by the *TM6SF2* rs58542926 variant and carriage of its minor allele confers significantly greater NAFLD-related hepatic fibrosis independent of gender, age at biopsy, BMI, T2DM, and *PNPLA3* rs738409 genotype.

This strong association between *TM6SF2* rs58542926 and fibrosis stage was replicated independently in the *validation cohort* (β 0.238±0.097, 95%CI 0.047-0.428, P=0.014) and also clearly demonstrated in the *combined cohort* (β 0.357±0.079, 95%CI 0.203-0.511, P=6.36 x10^{-6}) using an additive model adjusted for gender, age at biopsy, BMI, T2DM and *PNPLA3* rs738409 genotype in both cases. To illustrate the potential clinical relevance of this finding, when the multivariate analysis was repeated subdividing the NAFLD cohort into those with mild fibrosis (F0-1) and advanced fibrosis (F2-4), carriage of each copy of the *TM6SF2* rs58542926 C>T minor allele was associated consistently with a significant increased risk of advanced fibrosis, independent of gender, age at biopsy, BMI, T2DM, and *PNPLA3* rs738409 genotype across each cohort studied (Table 4.7).

Table 4.7: Multivariate analysis of association between *TM6SF2* rs58542926 genotype and fibrosis stage F0-1 (mild) vs. F2-4 (advanced).

Variables	Discovery cohort (n=349)		Validation cohort (n=725)		Combined cohort (n=1074)	
	OR (95% CI)	*P*-value	OR (95% CI)	*P*-value	OR (95% CI)	*P*-value
***TM6SF2* genotype**	**2.94 (1.76-4.89)**	**3.44×10^{-5}**	**1.46 (1.03-2.09)**	**0.0362**	**1.88 (1.41-2.5)**	**1.63×10^{-5}**
***PNPLA3* genotype**	**1.57 (1.21-2.19)**	**0.0086**	**1.32 (1.05-1.66)**	**0.0183**	**1.40 (1.16-1.69)**	**4.84×10^{-4}**
Age	**1.03 (1.01-1.06)**	**0.0045**	**1.02 (1.01-1.04)**	**0.0041**	**1.03 (1.01-1.04)**	**1.57×10^{-5}**
Gender (Male)	1.05 (0.64-1.74)	0.8297	**0.55 (0.39-0.77)**	**4.50×10^{-4}**	**0.69 (0.53-0.91)**	**0.0096**
BMI	**1.05 (1.00-1.10)**	**0.0368**	**1.03 (1.01-1.05)**	**9.80×10^{-4}**	**1.04 (1.02-1.05)**	**3.78×10^{-5}**
T2DM	**2.39 (1.49-3.84)**	**0.0003**	**2.73 (1.93-3.88)**	**1.68×10^{-8}**	**2.57 (1.95-3.39)**	**1.78×10^{-11}**

Additive model including age, gender, BMI, T2DM and *PNPLA3* rs738409 genotype as covariates.
Discovery/Validation/Combined cohorts: Stage F0-1 (mild) n=198/439/637, Stage F2-4 (advanced) n=151/286/437.

4.4 Discussion

A region on chromosome 19 (19p13.11) that contains multiple genes previously known as *NCAN*/*CLIP2*/*PBX4*/*TM6SF2* has been reported to be associated with NAFLD (Speliotes *et al.*, 2011; Gorden *et al.*, 2013; Kozlitina *et al.*, 2014) as well as variations in plasma cholesterol, triglyceride and low-density lipoprotein levels in several earlier studies (Kathiresan *et al.*, 2008; Teslovich *et al.*, 2010; Holmen *et al.*, 2014). In particular, a variant within the *NCAN* gene (rs2228603 C>T) that is in strong linkage disequilibrium (D'=0.926, r^2=0.798) with *TM6SF2* rs58542926 was reported to be associated with radiologically and histologically characterized NAFLD in both GWAS and candidate-gene studies (Speliotes *et al.*, 2011; Gorden *et al.*, 2013). Prior to the recent publication by Kozlitina *et al.* (Kozlitina *et al.*, 2014), examination of linkage disequilibrium patterns across the region had already brought that association into question (Anstee and Day, 2013). It was however the use of a genome-wide exome-chip genotyping approach, combined with detailed association analysis conditioning on previously published variants across the 19p13.11 region, which determined that the causative variant affecting HTGC was *TM6SF2* rs58542926 (Kozlitina *et al.*, 2014). In clinical practice, simple steatosis is generally considered to have a benign course and so degree of HTGC is of limited prognostic relevance (Ekstedt *et al.*, 2006; Anstee *et al.*, 2011b; Anstee *et al.*, 2013b). In contrast, progressive hepatic fibrosis leading to cirrhosis is the principal common pathway to hepatic failure and a liver-related death (Ekstedt *et al.*, 2006; Musso *et al.*, 2011). Using two large, well-characterised European Caucasian cohorts with biopsy-proven NAFLD I demonstrate that carriage of the *TM6SF2* rs58542926 variant is strongly associated with the presence of NAFLD and in particular with a significantly greater risk of developing advanced hepatic fibrosis/cirrhosis.

Evidence to support a modifier effect of the *TM6SF2* rs58542926 variant on histologically determined HTGC (steatosis), seen only when the 1,074-patient strong *combined cohort* was studied, is arguably more modest than might be expected. Our findings do support the previously reported association (Speliotes *et al.*, 2011; Gorden *et al.*, 2013; Kozlitina *et al.*, 2014), although differences in sensitivity to subtle changes in HTGC between radiological and histological modalities may have reduced the power to detect this effect (McPherson *et al.*, 2009). Kozlitina *et al.* reported that the maximal effect of the *TM6SF2* variant in European Caucasians was

only a mean 9.2% increase in ^{1}H-MRS quantified HTGC in TT homozygotes above the ~5.9% observed in CC homozygotes (Kozlitina *et al.*, 2014). Histological assessment of hepatic steatosis uses broad microscopic categories reflecting the proportion of hepatocytes that are visibly steatotic (S0 <5%, S1 5-33%, S2 33-66%, S3 >66%) (Bedossa and Consortium, 2014). Based on data from previous comparative modality analysis (McPherson *et al.*, 2009), the modest gene effect size reported by Kozlitina (less-than a 3-fold increase in HTGC above normal) would likely be encompassed within the histological S1 bracket and therefore may not be apparent histologically. Combined with the relatively low minor allele frequency in the background population, smaller cohorts may therefore have insufficient statistical power for an association to become evident.

The modifier effect of the *TM6SF2* variant on grade of steatohepatitis (disease activity) was apparent in the initial *discovery cohort* analysis however statistical significance was not reached in the subsequent validation analysis. The *validation cohort* comprised a mixture of patients recruited from both hepatology and bariatric services and, although the cohorts appear well matched histologically, the *validation cohort* exhibited higher mean BMI levels (38.5±9.1 vs. 34.5±5.7 kg/m^2, *P*<0.0001), a younger mean age (47.6±12.4 vs. 51.5±12.0 years, p<0.0001), a greater female preponderance (56.1% vs. 32.1%, *P*<0.0001) and a lower prevalence of T2DM (32.4% vs. 46.1%, *P*<0.0001) than the *discovery cohort* (Table 4.2). These factors may have impacted on our ability to replicate the initial association with steatohepatitis in a multivariate analysis. Further study of the variant in other patient cohorts and exploration of the functional effects of *TM6SF2* on inflammatory response will be needed to address this point.

The key finding of the current study is that carriage of the *TM6SF2* rs58542926 C>T minor allele is unequivocally associated with an increased risk of advanced NAFLD-associated hepatic fibrosis. This highly significant effect was consistently demonstrated across all the cohorts studied and was independent of potentially confounding factors including gender, age at time of biopsy, BMI, T2DM, and *PNPLA3* rs738409 genotype. Conditional analysis undertaken as part of the present study adds further weight to the assertion that the 19p13 signal is causally related to *TM6SF2* and not *NCAN*, not only for HTGC as was previously reported (Kozlitina *et al.*, 2014) but now also for stage of hepatic fibrosis. These findings therefore

establish a new and important clinical relevance to the recently described association between *TM6SF2* and NAFLD and suggest that *TM6SF2* should be considered alongside *PNPLA3* (Romeo *et al.,* 2008*;* Valenti *et al.,* 2010a) and *GCKR* (Speliotes *et al.*, 2011; Anstee *et al.*, 2013a; Petta *et al.*, 2014), as one of a handful of genes so far identified that are associated not only with variations in hepatic triglyceride accumulation but also with fibrogenesis (Anstee and Day, 2013). It is noteworthy that across all the cohorts studied, the odds ratio for advanced fibrosis conferred by each copy of the *TM6SF2* variant carried was consistently of similar or up to 2-fold greater magnitude than that which was observed, or has previously been reported (Valenti *et al.*, 2010a; Valenti *et al.*, 2010b), for the widely-studied *PNPLA3* rs738409 variant (Table 4.7).

After our findings were published in mid-2014, the role of genetic modifier of *TM6SF2* rs58542926 variant on severity of NAFLD, particularly the stage of simple steatosis, was soon replicated in studies performed in adults (Dongiovanni *et al.*, 2015; Sookoian *et al.*, 2015; Wang *et al*., 2015; Zhou *et al.*, 2015) and in paediatric patients (Mancina *et al.*, 2015; Goffredo *et al.*, 2016; Grandone *et al.*, 2016). One study with large European cohort identified that NAFLD patients with *TM6SF2* rs58542926 E167K variant are not only more steatotic but also more prone to have steatohepatitis (necroinflammation and lobular ballooning) and fibrosis which validated our key finding (Dongiovanni *et al.*, 2015). Two studies from China (Wong *et al.*, 2014) and Argentina (Sookoian *et al.*, 2015) failed to replicate the effect of *TM6SF2* rs58542926 variant upon hepatic fibrosis. The former study used a community-based cohort rather than selecting individuals afflicted by NAFLD, suggesting the cohort is more likely to be a representative of ‘healthy population’. Together with the given generally low minor allele frequency of *TM6SF2* rs58542926 (MAF 0.07), they reported that this variant has limited impact on NAFLD. Whereas the other Chinese group with a more appropriate study design (a case-control association study between equivalent numbers of healthy controls and biopsy-proven NAFLD patients) demonstrated a positive association between NAFLD and rs58542926 (Wang *et al.*, 2015). The latter Argentinian study only contained a cohort of 226 NAFLD cases with histologically characterised disease: 96 of them had simple steatosis and the remaining 130 exhibited relatively mild fibrosis (mean fibrosis stage of 1.4 out of 4). The statistical power of this low quality study would

therefore be questionable, and Type II error may often occur. Interestingly, this significant genetic influence of *TM6SF2* E167K on fibrosis even remained in a GWAS recently published suggesting carrier of this variant increased susceptibility to alcohol-induced cirrhosis (Buch *et al.*, 2015).

When Kozlitina *et al.* first linked the *TM6SF2* rs58542926 minor allele (T) to NAFLD, a separate GWAS published around the same time assessing modifiers of serum lipid levels and cardiovascular disease (CVD) revealed that carriage of the common allele (C) of the same SNP was strongly associated with raised circulating TG/LDL-C/TC levels and increased risk of myocardial infraction/cardiovascular disease while carriage of the (T) minor allele was protective (Holmen *et al.*, 2014). One study applied two separate groups (a NAFLD patient cohort from Italy and Finland and a Swedish obese population) reinforced above findings (Dongiovanni *et al.*, 2015). In 427 NAFLD patients evaluated carotid atherosclerosis, the minor allele carriers of E167K had lower risk of developing carotid plaques (OR 0.49, 95%CI 0.25-0.94) while decreased serum lipid levels (TC, non HDL-C and ApoB) and lower prevalence of cardiovascular events (hazard ratio 0.61, 95%CI 0.39-0.95) were found in E167K carriers among 1,819 obese subjects (Dongiovanni *et al.*, 2015). Above findings suggest that *TM6SF2* rs58542926 C-allele carriage increases circulating lipid profiles whilst T-allele carriage promotes hepatic triglyceride/cholesterol retention. Although hepatic steatosis is often accompanied by insulin resistance, this dual role of *TM6SF2* rs58542926 on fatty liver and serum lipid profiles was not found to be correlated (Mancina *et al.*, 2015; Zhou *et al.*, 2015). In fact, TM6SF2 E167K carriers had a higher level of hepatic insulin sensitivity of glucose production and adipose tissue lipolysis than those wild-type subjects (Zhou *et al.*, 2015).

In vivo (Holmen et al., 2014; Kozlitina et al., 2014) and *in vitr*o (Holmen *et al.*, 2014; Kozlitina *et al.*, 2014; Mahdessian *et al.*, 2014) functional studies indicate that *TM6SF2* has gene effects on fat retention in the liver, lipid efflux and alteration in the number and average size of lipid droplets. However, these studies were of too short a duration to adequately address the effects on steatohepatitis or fibrogenesis. Furthermore, previous experimental evidence has shown that hepatic triglyceride accumulation may not itself be directly hepatotoxic. This was elegantly demonstrated in mice by silencing hepatic gene expression of *diacylglycerol O-acyltransferase 2* (*Dgat2*), a key enzyme mediating the conversion of free fatty acids to triglyceride

(Yamaguchi *et al.*, 2007). Rather than ameliorating steatohepatitis, the consequent reduction in hepatocyte triglyceride synthesis was associated with increased fatty acid oxidation, particularly through Cyp2e1, leading to greater oxidative stress, cellular damage and higher serum transaminase levels (Yamaguchi *et al.*, 2007). It is therefore tempting to speculate that the function of *TM6SF2* and the mechanism through which *TM6SF2* drives NAFLD-associated hepatic fibrosis may be other than through increased triglyceride accumulation.

In conclusion, the current study confirms that *TM6SF2* is associated with histologically defined NAFLD and is the first demonstration that this gene serves as a powerful modifier of hepatic fibrogenesis. That this gene is also associated with disturbed cholesterol metabolism and so may modify risk of cardiovascular events including myocardial infarction (Holmen *et al.*, 2014), suggesting that *TM6SF2* may be an important determinant of clinical outcome across several facets of metabolic syndrome related end-organ damage.

Based on the available data it appears that *TM6SF2* is an important determinant of clinical outcome across several facets of metabolic syndrome related end-organ damage. It is tempting to speculate that *TM6SF2* may act as a 'switch' with *TM6SF2* rs58542926 T-allele mediated hepatic retention of triglyceride and cholesterol predisposing to NAFLD-fibrosis whilst C-allele carriage promotes VLDL excretion, protecting the liver but at the expense of an increased risk of atherosclerosis and ultimately, cardiovascular disease (Figure 4.1). NAFLD is globally acknowledged to be associated with CVD which is the most common causes of mortality in NAFLD patients rather than malignancy and a liver-related death (Targher *et al.*, 2007a; Rafiq *et al.*, 2009; Angulo, 2013; Anstee *et al.*, 2013b), therefore carriage of different alleles of *TM6SF2* may dissociate this. The minor (T) allele of *TM6SF2* rs58542926 carriers may likely to experience liver-induced death while individuals carrying the major (C) allele are prone to afflicted by cardiovascular morbidity and mortality. These data surely mandate further mechanistic study to determine the physiological and pathophysiological role of this gene in various tissues and cell types as a modifier of fibrogenesis, and hopefully could develop a putative therapeutic target.

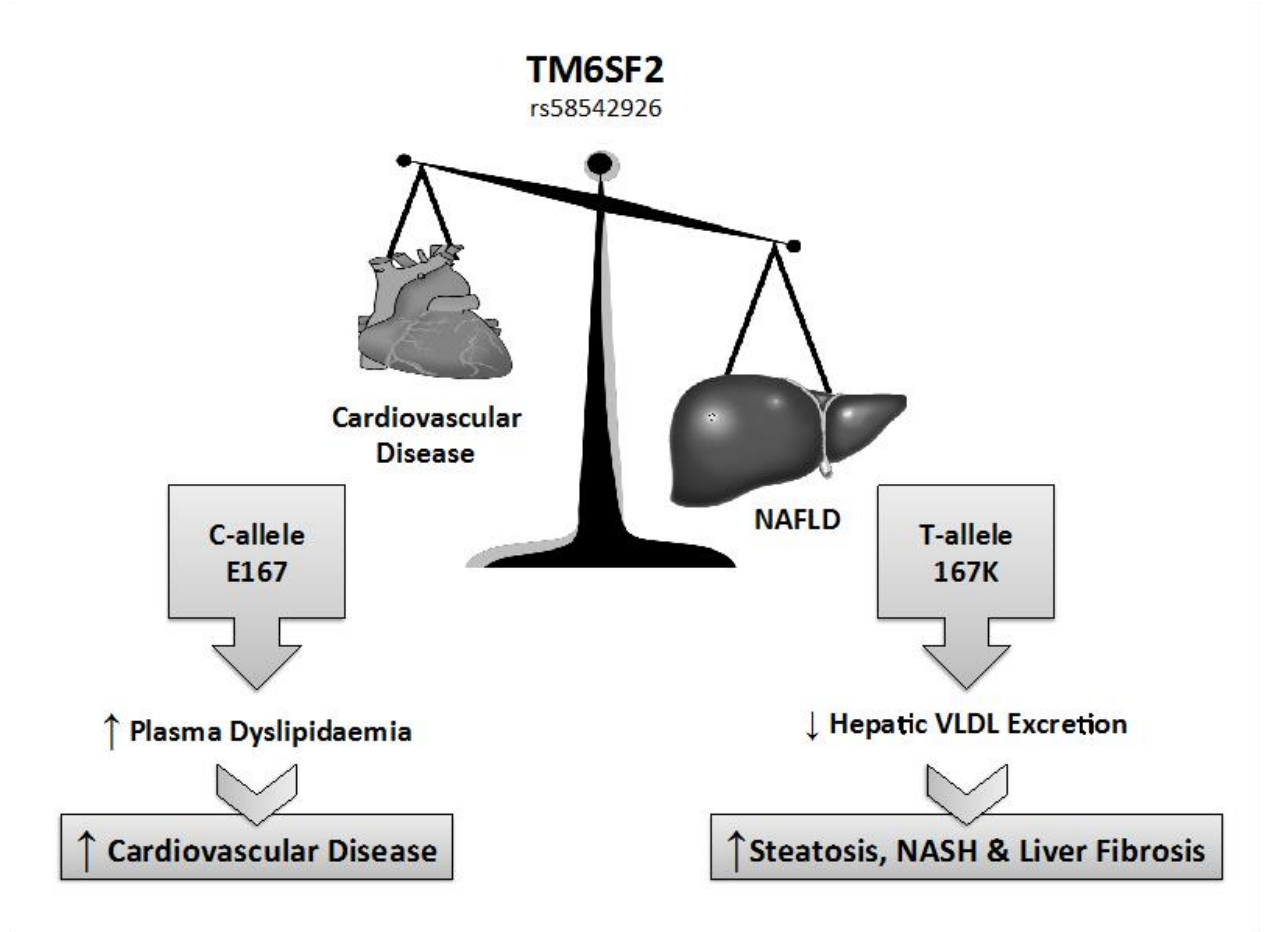


Figure 4.1: Outcomes of the metabolic syndrome: *TM6SF2* dissociates NAFLD from cardiovascular disease. Adapted from (Kahali *et al.*, 2015).

Chapter 5. Genetic Modifiers for NAFLD-Associated Hepatocellular Carcinoma

5.1 Introduction

The rise in the burden of NAFLD coincides with a marked increase in the incidence of HCC in many countries (El-Serag and Rudolph, 2007; European Association For The Study Of The Liver and Cancer, 2012; Dyson *et al.*, 2014). Worldwide, most HCC cases are related to chronic viral hepatitis; however, more than half of all HCC cases in developed countries occur in the viral hepatitis negative population (Parkin, 2006; Baffy *et al.*, 2012). Features of the metabolic syndrome, particularly obesity and T2DM (with gender, age and ethnicity), are well-known independent risk factors for HCC (Guzman *et al.*, 2008; Starley *et al.*, 2010) whilst the pathogenic processes that favour progression from steatosis to steatohepatitis are also pro-carcinogenic (El-Serag and Rudolph, 2007). Accumulating evidence suggests that a substantial proportion of HCC occurs in the absence of advanced fibrosis (Hashimoto *et al.*, 2009; Ascha *et al.*, 2010; Sanyal *et al.*, 2010; Ertle *et al.*, 2011).

Despite the high prevalence of NAFLD, only a minority of patients exhibit steatohepatitis, progress to significant fibrosis or experience associated morbidity and mortality (Anstee *et al.*, 2013b). The reasons for the apparent variation in individual susceptibility to progressive NAFLD in general and NAFLD-related HCC in particular are incompletely understood (Baffy *et al*., 2012; Anstee and Day, 2013). NAFLD-related HCC develops through the complex interplay of environmental and genetic factors that determine individual risk ('EASL-EORTC clinical practice guidelines: management of hepatocellular carcinoma,' 2012; Valenti *et al.*, 2013). The role of the *PNPLA3* c.444C>G SNP (rs738409) is well recognized as a modifier of hepatic triacylglycerol accumulation and NAFLD progression (Romeo *et al.*, 2008; Valenti *et al.*, 2010a; Anstee and Day, 2013). This variant has been associated *with* increased HCC risk in alcohol-related liver disease (Nischalke *et al.*, 2011; Trepo *et al.*, 2012; Guyot *et al.*, 2013; Trepo *et al.*, 2014) and, more variably, in chronic viral hepatitis (Falleti *et al.*, 2011; Nischalke *et al.*, 2011; Valenti *et al.*, 2011b; Guyot *et al.*, 2013; Trepo *et al.*, 2014). Data has also been presented showing an association with HCC in morbidly obese patients *(*Burza *et al.,* 2012*)* and a mixed-aetiology cohort (Hassan *et al.*, 2013). Although it may be hypothesised that these latter associations are related to underlying NAFLD, to date no studies have specifically addressed the effect of *PNPLA3* rs738409 C>G carriage on HCC risk in a NAFLD cohort.

As part of the current project, the *TM6SF2* association with NAFLD was expanded after Kozlitina *et al*. first linked the rs58542926 minor allele (T) to steatosis (Kozlitina, 2014); this gene was also found to serve as a powerful modifier of hepatic fibrogenesis (Liu *et al.*, 2014b). Whether the *TM6SF2* SNP is also relevant to NAFLD-associated HCC was therefore tested.

The aim of the current study was to determine whether carriage of the *PNPLA3* rs738409 C>G or *TM6SF2* rs58452926 C>T polymorphisms confers an increased risk of NAFLD-related HCC, and whether that effect is independent of the presence of cirrhosis. To address this, I performed a case-control association study in a well-characterised Northern European cohort with NAFLD-related HCC.

5.2 Methods

5.2.1 Patients

Patients were recruited from hepatology clinics at two European specialist centres, the Freeman Hospital, Newcastle-upon-Tyne, UK and Inselspital Hospital, Bern, Switzerland. The study had all the necessary ethical approvals in both the countries and all participants gave informed consent (Table 5.1). A cohort of 100 Northern European Caucasian patients, were recruited prospectively, with primary HCC arising on a background of NAFLD was identified (UK 82, Switzerland 18 patients). The diagnosis of HCC was established histologically or through non-invasive assessment according to the EASL-EORTC clinical practice guidelines ('EASL-EORTC clinical practice guidelines: management of hepatocellular carcinoma,' 2012). The presence of NAFLD was determined through histological assessment of non-tumour liver tissue or, when biopsy was not clinically appropriate, through radiological evidence of hepatic steatosis. As a comparator, a cohort of 275 UK patients recruited prospectively, with histologically characterised NAFLD of different stages of disease but without clinical evidence of HCC was assembled. These were unrelated patients with histologically characterised NAFLD, derived from a patient population originally identified as having ultrasonographically detected bright liver and abnormal biochemical tests (ALT and/or GGT). In all cases, alternative diagnoses were excluded, including excess alcohol intake (alcohol intake <20g/day for women; <30g/day for men), chronic viral hepatitis (hepatitis B and hepatitis C), autoimmune liver diseases, hereditary hemochromatosis, α1-antitrypsin deficiency, Wilson's disease and drug induced liver injury.

Clinical and laboratory data were collected at the time of diagnosis. These included basic anthropometrics so that body mass index (BMI) could be calculated. Relevant co-morbidity including the presence of diabetes mellitus (fasting glucose ≥7.1 mmol/L [≥128 mg/dl] or treatment with anti-diabetic drugs) and evidence of underlying cirrhosis was recorded. Laboratory evaluation included routine liver biochemistry (alanine and aspartate aminotransferase, total bilirubin, albumin, alkaline phosphatase and gamma glutamyl transpeptidase); full blood count; total- and HDL-cholesterol and total triglycerides; viral serology for hepatitis B and C infection and autoantibodies. Demographic details of the cohorts are shown in Table 5.1.

5.2.2 Liver biopsy

Liver biopsy performance is described in section 2.2. The severity of steatosis, necroinflammatory grade and stage of fibrosis were scored according to the validated Kleiner criteria (Kleiner *et al.*, 2005). This assessment allowed confirmation of a NAFLD diagnosis and assessment of fibrosis so that patients were classified as cirrhotic or non-cirrhotic. In 25 HCC patients, the diagnosis of HCC was confirmed histologically and staged according to Edmondson (Edmondson and Steiner, 1954), adapted for needle biopsy specimens.

5.2.3 DNA preparation from blood samples

DNA extraction from blood samples was described in section 2.3.2. DNA preparation from some Newcastle samples was performed by Julian Leathart and Julia Patch (Newcastle University, ICM). DNA samples from Bern were isolated locally by the collaborators.

5.2.4 PNPLA3 rs738409 and TM6SF2 rs58452926 genotyping

Genotypes for the two SNPs were determined by allelic discrimination using TaqMan reagents. TaqMan SNP genotyping assays for *PNPLA3* rs738409 (reference number: C___7241_10, Cat. # 4351379) and *TM6SF2* rs58542926 (reference number: C___8709053_10, Cat. # 4351379) were purchased from Applied Biosystems Inc., USA. Genotypes for these SNPs were determined by allelic discrimination using TaqMan reagents according to the manufacturer's protocol. Procedures were described in section 2.3.4.2.

5.2.5 Statistical analysis

Statistical analyses was performed using SPSS v19.0 (IBM, USA) to collate and analyse cohort phenotype data. Continuous variables were tested using Student's t-test/one-way ANOVA and categorical variables were analysed by Chi-squared test unless otherwise stated. PLINK v1.07 (Purcell *et al.*, 2007) (via the gPLINK v2.050 GUI) was used to conduct the genetic analysis. An initial univariate chi-squared analysis was performed to determine whether *PNPLA3* rs738409 C>G carriage

differed between the NAFLD-HCC and control NAFLD-Cohort. Subsequently multivariate logistic regression analysis was conducted, incorporating rs738409 genotype and those biologically relevant covariates that were associated with risk of disease progression to HCC (age, gender, coexisting T2DM, BMI and presence of cirrhosis) to test the genetic association. Consistent with previous studies (Romeo *et al.*, 2008; Sookoian *et al.*, 2009b; Rotman *et al.*, 2010), an additive genetic model best fitted the data and was reported. Results were expressed as odds ratio (OR) with 95% confidence intervals (CI). Significance was taken as $p<0.05$ throughout.

5.3 Results

5.3.1 Cohort characteristics

One hundred NAFLD-HCC patients and 275 histologically characterised NAFLD control patients were recruited. Characteristics of the NAFLD-HCC and NAFLD control populations are provided in Table 5.1. Patients with NAFLD-HCC were significantly older (mean age 70.3±8.0 vs. 50.9±12.4 years, *P*<0.0001) than a general NAFLD cohort. NAFLD-HCC patients were also significantly more likely to be male (82% vs. 59%, *P*<0.0001) and to be diabetic (68% vs. 43%, *P*<0.0001). The presence of NAFLD-HCC was also significantly associated with underlying advanced fibrosis/cirrhosis (Kleiner F4 67% vs. 9%, *P*<0.0001). NAFLD-HCC patients exhibited a lower mean BMI than the NAFLD control population.

Table 5.1: Details of NAFLD-HCC and NAFLD Cohorts.

Phenotype	NAFLD-HCC Cohort n=100	NAFLD Cohort n=275	*P*-Value
***PNPLA3* rs738409 G-Allele Frequency**	0.505	0.333	**<0.0001**
Age (Mean±SD)	70.3±8.0	50.9±12.4	**<0.0001**
Male Gender (%)	82 (0.82)	161 (0.59)	**<0.0001**
BMI (Mean±SD)	32.0±6.6	34.4±5.2	**0.0003**
Diabetes (%)	68 (0.68)	117 (0.43)	**<0.0001**
Cirrhosis (%)	67 (0.67)	26 (0.09)*	**<0.0001**

Categorical values are shown as n (%). Continuous variables are shown as mean ± SD. *Fibrosis stage distribution in NAFLD Cohort: F0 89 (32.4%), F1 97 (35.3%), F2 37 (13.5%), F3 26 (9.5%), F4 26 (9.5%).

5.3.2 *PNPLA3* rs738409 C>G polymorphism carriage is associated with increased risk of HCC relative to a tertiary centre NAFLD cohort

To determine whether carriage of the *PNPLA3* rs738409 C>G allele influenced susceptibility to NAFLD-related HCC, we studied a cohort of Northern European adult patients with established NAFLD. The total study population of 375 individuals (100 NAFLD-HCC, 275 NAFLD only) was genotyped for *PNPLA3* rs738409. *PNPLA3* rs738409 genotypes were confirmed to be in Hardy-Weinberg equilibrium. Reflecting the known association with NAFLD, the minor allele (G) frequency observed in the present study (G617: 0.38) was slightly higher than that observed in a population of North-Western European descent by the International HapMap project (http://www.ncbi.nlm.nih.gov/projects/SNP/snp_ref.cgi?rs=738409) but similar to that reported in previous NAFLD studies (International HapMap Consortium, 2003; Valenti *et al.*, 2010a). Genotype frequencies are summarised in Table 5.2.

Compared to a cohort of NAFLD patients with varying severity, carriage of the *PNPLA3* rs738409 minor (G) allele (I148M variant) was strongly associated with the presence of NAFLD-related HCC (unadjusted OR 2.046, 95%CI 1.47-2.84, X^2 18.50, P<0.0001) and exhibited a gene-dosage effect with the incidence of HCC increasing with the number of G alleles present (Cochran-Armitage X^2 for trend 16.92, P<0.0001). Adopting an additive model relative to the NAFLD patient cohort, an approximate doubling of HCC risk was observed for each copy of the minor (G) allele carried (unadjusted OR 1.95, 95%CI 1.40-2.70, P<0.0001). The unadjusted odds ratio increased to 3.92 (95%CI 2.06-7.48, P=0.0001) when only homozygotes were considered (Table 5.2). The relationship between *PNPLA3* genotype and a number of relevant patient-specific and clinical parameters is summarised in Table 5.3. Amongst those samples where HCC was histologically characterised, carriage of the rs738409 C>G polymorphism was associated with the presence of more poorly differentiated tumour (Fisher's Exact Test P=0.018), (Table 5.3).

5.3.3 *PNPLA3* rs738409 C>G polymorphism carriage in NAFLD-HCC relative to an unselected population cohort

The published *PNPLA3* rs738409 C>G minor allele frequency in a UK general population sample (the MRC/Wellcome Trust UK 1958 Birth Cohort, http://www.b58cgene.sgul.ac.uk/) is 0.23 (95%CI 0.21-0.24) with genotype frequencies of CC 0.59, CG 0.36, GG 0.05. Carriage of the rs738409 C>G

polymorphism was increased in the NAFLD-HCC cohort relative to this unselected population (Table 5.2). As limited phenotype data is available on the population cohort only unadjusted odds ratios could be calculated however, adopting an additive model relative to the UK general population, a greater than 3-fold increased HCC risk was observed for each copy of the minor (G) allele carried (OR 3.43, 95%CI 2.54-4.62, $P<0.0001$). When only homozygotes were considered an odds ratio of 12.19 (95%CI 6.89-21.58, $P<0.0001$) for GG over CC was observed.

Table 5.2: *PNPLA3* rs738409 genotype frequencies and their relationship to risk of HCC development.

Genotype	NAFLD-HCC n(%)	NAFLD Cohort n(%)	Unadjusted OR (95%CI)	*P*-value	Adjusted OR (95%CI)♯	*P*-value	UK Pop[n.] n(%)†	Unadjusted OR (95%CI)	*P*-value
CC	28 (0.28)	125 (0.46)	-	-	-	-	871 (0.59)	-	-
GC	43 (0.43)	117 (0.42)	1.64 (0.96-2.81)	0.072	2.35 (0.90-6.13)	0.082	531 (0.36)	2.52 (1.55-4.10)	**0.0002**
GG	29 (0.29)	33 (0.12)	3.92 (2.06-7.48)	**<0.0001**	5.05 (1.47-17.29)	**0.01**	74 (0.05)	12.19 (6.89-21.58)	**<0.0001**

♯ Odds ratio for HCC relative to CC genotype adjusted for age, gender, diabetes, BMI and cirrhosis.

† The MRC/Wellcome Trust UK 1958 Birth Cohort.

Table 5.3: Comparison of selected characteristics according to *PNPLA3* rs738409 genotype within NAFLD-HCC and NAFLD Cohorts.

	NAFLD-HCC Cohort				NAFLD Cohort			
Genotype	**CC**	**CG**	**GG**	***P*-Value†**	**CC**	**CG**	**GG**	***P*-value†**
n(%)	**n = 28 (0.28)**	**n = 43 (0.43)**	**n = 29 (0.29)**		**n = 125 (0.46)**	**n = 117 (0.42)**	**n = 33 (0.12)**	
Age (Mean±SD)	73.6±8.1	68.9±7.9	69.2±7.4	**0.034**	50.9±12	51±12	50±15.2	0.906
Male Gender (%)	23 (0.82)	36 (0.84)	23 (0.79)	0.892	75 (0.60)	74 (0.63)	12 (0.36)	**0.020**
BMI (Mean±SD)	30.4±6.7	31.9±4.5	33.6±8.9	0.201	34.4±4.8	34.6±5.7	33.9±4.8	0.765
Diabetes (%)	18 (0.64)	31 (0.72)	19 (0.66)	0.744	51 (0.41)	56 (0.48)	10 (0.30)	0.171
Cirrhosis (%)	13 (0.46)	30 (0.70)	24 (0.83)	**0.012**	11 (0.01)	9 (0.01)	6 (0.18)	0.181
Histological Grade of HCC*	5 (0.45) / 6 (0.54)	8 (0.80) / 2 (0.20)	0 (0.00) / 4 (1.00)	**0.018**	-	-	-	-

Categorical values are shown as n (%). Continuous variables are shown as mean ± SD.

† Comparison between *PNPLA3* genotypes within each study cohort (chi-squared test used for categorical variables and one-way ANOVA for continuous variables unless otherwise stated).

* 25 cases had histologically characterised HCC. Results expressed as n (%) of cases with 'well differentiated'/'moderate-poorly differentiated' tumour. P-value calculated by Fisher's Exact Test.

5.3.4 Contribution of the PNPLA3 rs738409 C>G polymorphism to NAFLD-HCC risk is independent of presence of cirrhosis

Given the previous reports demonstrating a role for rs738409 C>G as a modifier of disease progression and fibrosis in NAFLD, and the association between presence of age, T2DM, obesity and cirrhosis with HCC, a multivariate logistic regression analysis using an additive model was performed to control for any potential confounding effects. Along with genotype, gender, age at diagnosis, presence of advanced fibrosis/cirrhosis, T2DM and BMI were included in the analysis. *PNPLA3* genotype (OR 2.26, 95%CI 1.23-4.14, *P*=0.0082), gender, age and presence of cirrhosis were independent predictors of NAFLD-HCC (Table 5.4). Of these factors, male gender (OR 11.11, 95%CI 4.17-33.33, *P*<0.0001) and the presence of underlying cirrhosis (OR 9.37, 95%CI 3.82-23.00, *P*<0.0001) conferred the greatest risk. The presence of T2DM or raised BMI were not significant predictors of HCC risk in this analysis, although a trend towards significance was observed for T2DM (OR 2.33, 95%CI 0.93-5.81, *P*=0.070). No significant interactions were identified within the regression model. The results of multivariate analysis correcting for age, gender, T2DM and BMI but not cirrhosis are shown in Table 5.5 and Table 5.6.

Table 5.4: Multivariate analysis of the effect of *PNPLA3* genotype on NAFLD-related HCC risk.

Variables	OR (95% CI)	*P*-value
***PNPLA3* rs738409 genotype**	**2.26 (1.23-4.14)**	**0.0082**
Age	**1.24 (1.17-1.32)**	**<0.0001**
Gender (Male)	**11.11 (4.17-33.33)**	**<0.0001**
BMI	0.94 (0.87-1.02)	0.148
Diabetes	2.33 (0.93-5.81)	0.070
Cirrhosis	**9.37 (3.82-23.00)**	**<0.0001**

Additive Model including age, gender, BMI, diabetes and cirrhosis as covariates.

Table 5.5: Multivariate analysis of the effect of *PNPLA3* genotype on NAFLD-related HCC risk adjusted for age, gender, BMI and diabetes but not cirrhosis.

Variables	OR (95% CI)	*P*-value
***PNPLA3* rs738409 genotype**	**3.06 (1.76-5.29)**	**<0.0001**
Age	**1.28 (1.21-1.36)**	**<0.0001**
Gender (Male)	**9.52 (3.85-23.26)**	**<0.0001**
BMI	0.98 (0.91-1.05)	0.593
Diabetes	**2.71 (1.19-6.17)**	**0.018**

Additive Model including age, gender, BMI and diabetes (but not cirrhosis) as covariates.

Table 5.6: *PNPLA3* rs738409 genotype frequencies and their relationship to risk of HCC development adjusted for age, gender, BMI and diabetes but not cirrhosis.

Genotype	NAFLD-HCC n(%)	NAFLD Cohort n(%)	Adjusted Odds Ratio (95%CI)♯	*P*-value
CC	28 (0.27)	125 (0.46)	-	-
CG	43 (0.43)	117 (0.42)	2.50 (1.03-6.08)	0.043
GG	29 (0.3)	33 (0.12)	9.61 (3.21-28.75)	**<0.0001**

♯ Adjusted odds ratio for HCC relative to CC genotype (covariates age, gender, T2DM and BMI).

NAFLD-HCC patients carrying the G allele on average developed HCC at a younger age than those that do not (Table 5.3). When NAFLD-HCC patients were stratified by median age (</>70 years) a significant enrichment in G allele carriage was present in the younger HCC patients over the older ones (MAF Young 0.59 vs. Old 0.42, X^2 5.76, P=0.016). Similarly, when compared to the NAFLD-Cohort control group, the effect of rs738409 carriage was statistically significant in those younger than the median age (OR 2.85, 95%CI 1.85-4.37, X^2 24.09, P<0.0001) but not in those older.

Consistent with previous reports (Valenti *et al.*, 2010a), carriage of the rs738409 C>G polymorphism was also significantly associated with the presence of NAFLD-related cirrhosis (OR 2.33, 95%CI 1.66-3.27, X^2 24.8, P<0.0001) across the entire 375 patient NAFLD cohort and exhibited a gene-dosage effect with the incidence of cirrhosis increasing with the number of G alleles present (Cochran-Armitage X^2 for trend 22.68, P<0.0001). To further examine the effect of *PNPLA3* on HCC risk independent of degree of fibrosis, only those patients with coexistent cirrhosis were studied (NAFLD-HCC n=67, NAFLD-Cohort n=26). Study-group sizes were relatively small in this sub-analysis however carriage of the rs739409 C>G polymorphism remained significantly associated with NAFLD-HCC (OR 2.06, 95%CI 1.07-3.94, X^2 4.78, P=0.029) in patients with cirrhosis. Adopting an additive model incorporating genotype, gender, age at diagnosis, T2DM and BMI, *PNPLA3* remained significantly associated with HCC (OR 3.41, 95%CI 1.39-8.37, P=0.0074) amongst patients with cirrhosis. The effect did not reach statistical significance in the non-cirrhotic cohort, but for HCC this group included only 23 cases.

5.3.5 TM6SF2 rs58542926 C>T polymorphism and risk of hepatocellular carcinoma

We also sought to determine whether *TM6SF2* rs58542926 had a similar effect to *PNPLA3* rs738409. *TM6SF2* rs58542926 allele and genotype frequencies in this cohort were compared to the NAFLD cohort described above. In the univariate analysis, homozygote carriage of the *TM6SF2* rs58542926 minor allele was associated with an increased risk of NAFLD-HCC with respect to CC (OR 1.922, 95%CI 1.31-2.81, P=6.81 x10^{-4}), however significance was lost in the multivariate analysis incorporating known risk factors including age, gender, BMI, T2DM and presence of cirrhosis (P=0.42).

5.4 Discussion

Using a large, well-characterised Northern European cohort with biopsy-proven NAFLD we show a strong association between *PNPLA3* rs738409 genotype and HCC risk, but not with *TM6SF2* rs58542926 genotype. This highly significant effect between the *PNPLA3* rs738409 SNP and HCC risk was independent of potentially confounding factors including age, gender, co-existent diabetes, obesity and the presence of cirrhosis. Although *PNPLA3* rs738409 is principally recognised as a disease risk factor in NAFLD (Anstee and Day, 2013), studies thus far reporting associations with HCC have been in individuals with viral hepatitis (Falleti *et al.*, 2011; Nischalke *et al.*, 2011; Valenti *et al.*, 2011b; Guyot *et al.*, 2013), alcoholic liver disease (Nischalke *et al.*, 2011; Trepo *et al*., 2012; Guyot *et al.*, 2013) or obesity (Burza *et al*., 2012) (reviewed (Valenti *et al.*, 2013; Trepo *et al.*, 2014). We report the first NAFLD study in which multivariate analysis relative to a well-characterised NAFLD population demonstrates that carriage of each G allele is associated with a doubling of HCC risk (adjusted OR 2.26, 95% CI 1.23-4.14, *P*=0.0082).

As they were recruited in a tertiary centre, our NAFLD cohort is likely to represent a more severe spectrum of disease than is present in an unselected population cohort, and so is likely to be enriched for rs738409 C>G minor allele carriage. We therefore also compared the NAFLD-HCC cohort with a UK general population sample (the MRC/Wellcome Trust funded UK 1958 Birth Cohort) to provide a measure of effect relative to an unselected background population. Although additional phenotype data for this secondary analysis was limited and so only univariate comparisons could be made, the results were striking. Carriage of the G allele was strongly associated with HCC with the homozygote GG genotype being associated with an unadjusted odds ratio for HCC of 12.19 (95%CI 6.89-21.58) over CC. By its nature, this secondary analysis cannot control for potential confounders including age, gender, co-existent diabetes, obesity and the presence of cirrhosis as these data are not available for the 1958 Birth Cohort and so care should be taken not to over-interpret the estimates of effect-size in this comparison. However, taken together with the results of our primary analysis comparing NAFLD-HCC with histologically characterised NAFLD, the results consistently demonstrate significantly increased HCC risk with *PNPLA3* rs738409 C>G minor allele carriage and suggest that the effect-size reported by our primary analysis using a tertiary centre NAFLD cohort may be a conservative

estimate. These results have implications for our understanding of HCC pathogenesis in NAFLD and, if supported by further validation, are also potentially of clinical relevance.

In keeping with established predictors of progression from steatosis to NASH and fibrosis, and the recognised independent associations of HCC with both cirrhosis and T2DM (El-Serag *et al.*, 2004; Anstee *et al.*, 2013b), we show that NAFLD-HCC patients are likely to be older males with T2DM and underlying cirrhosis. In contrast, although NAFLD-HCC patients were generally obese, at the time of HCC diagnosis the mean BMI was lower than that of the reference NAFLD population. The reason for this is not clear however this observation is possibly attributable to HCC induced cachexia.

A key finding in the current study is that the influence of rs738409 C>G on HCC risk was greater than can be accounted for by the associated increased risk of progression to cirrhosis. This was demonstrated both in the multivariate analysis, where the effect of cirrhosis was controlled by inclusion as a covariate, and also in a sub-analysis with the cohort stratified according to presence of cirrhosis. Amongst cirrhotics, rs738409 C>G remained a highly significant factor even after the other covariates were included in the model (OR 3.41, 95%CI 1.39-8.37, P=0.0074), indeed the adjusted odds ratio increased. Thus, supporting the conclusions of the multivariate regression analysis, HCC risk conferred by *PNPLA3* genotype is not mediated solely through progression to advanced fibrosis. The effect of *PNPLA3* did not reach significance in the non-cirrhotic sub-group and so it is tempting to speculate that the effect of rs738409 on HCC risk is largely confined to those with cirrhosis; however, only 23 HCC patents were included in that analysis and overall G allele carriage was low, severely limiting statistical power and so the negative result should be interpreted with caution. Furthermore, no significant interaction between genotype and cirrhosis was identified in the logistic regression model making this interpretation less likely. Validation in a larger NAFLD-HCC cohort will be required to clarify this point; irrespective of this our data supports the view that HCC promotion by *PNPLA3* in NAFLD is independent of its role in fibrosis progression.

It is noteworthy rs738409 C>G polymorphism carriage was associated with a mean 4-year younger age at tumour presentation than the CC genotype (Table 5.3).

Consistent with a recent report that *PNPLA3* genotype may have a greater impact on hepatic fibrosis progression at younger onset of alcohol consumption (Burza *et al.*, 2013), our observation that minor allele carriers were younger at HCC diagnosis possibly suggests the influence of genetic factors on HCC risk diminishes with age, whilst acquired exposure to environmental factors exerts a greater effect. Stratification by median age showed the greatest effect of *PNPLA3* in those aged less than seventy and so would support this view however, as discussed earlier, such an interaction is not supported by the multivariate analysis and so further studies in larger cohorts will be needed to address this.

The *PNPLA3* gene encodes a protein that is closely related to adipose triglyceride lipase (*ATGL/PNPLA2*), the major TAG hydrolase in adipose tissue (Zimmermann *et al.*, 2004; Romeo *et al.*, 2010). *PNPLA3* represents one of a small number of genes that has been consistently identified as a modifier of NAFLD severity (reviewed (Anstee and Day, 2013)). Notably, the variant has also been associated with inflammation and fibrosis independent of TAG accumulation (Sookoian *et al.*, 2009b; Rotman *et al.*, 2010; Valenti *et al.*, 2010a; Al-Serri *et al.*, 2012). The pathogenesis of HCC in NAFLD has been the subject of recent reviews, highlighting the likely contributions of lipotoxicity (Baffy *et al.*, 2012), metabolic or stress response pathways (Michelotti *et al.*, 2013), gut microbiota, bile acid receptors, vitamin D, senescence and autophagy in hepatic stellate cells as well as progenitor cell dysregulation (Lade *et al.*, 2014). However, our findings suggest that the contribution of *PNPLA3*, probably in the context of these proposed mechanisms, warrants further study.

In terms of lipotoxicity, the phenomenon called ‘adipose tissue expansion’ - is a compensatory response to an overwhelming uptake of fat/nutrition when obesity occurs. This process releases several cytokines, typically leptin, tumour necrosis factor alpha (TNF-α) and interleukin-6 (IL-6), promoting a pro-inflammatory and pro-fibrotic environment through activation of JAK/STAT, PI3K/Akt, mTOR and Nf-KB signalling pathways (Saxena *et al.*, 2007; Baffy *et al.*, 2012; Nakagawa *et al.*, 2014). Sustained lipid/nutrition overloading results in an increase in *de novo* lipogenesis coincident with a decrease in excretion of VLDL and a reduction in free fatty acids (FFA) clearance. Aberrant level of circulating FFA could induce ER stress and JNK-dependent hepatocyte lipoapoptosis (Malhi *et al.*, 2006; Wei *et al.*, 2006). In the

context of overwhelmed fatty acid oxidation, reactive oxygen species (ROS) can therefore be produced at high levels. These free radicals target DNA, RNA, proteins and lipids, causing inflammation, damage to genomic DNA and other macromolecules, mitochondrial dysfunction, ER stress and apoptosis, forming a fundamental process for oncogenesis (Hussain *et al.*, 2003; Malhi *et al.*, 2006; Wei *et al.*, 2006). Because PNPLA3 is known to be expressed in the liver and adipose tissue, suggested to be localized in the ER and in lipid droplets and possibly functioning as a lipid hydrolase (Pirazzi *et al.*, 2012), it is tempting to speculate that mutation of PNPLA3-I148M may stimulate or accelerate the conditions described above by means of loss of function. Namely, the mutated protein causes excess fat retention within lipid droplets via a reduced activity of lipid hydrolysis and an impaired secretion of VLDL in the ER. Hence, PNPLA3-I148M may profoundly promote the effects of adipose tissue remodelling and following cascades of lipotoxicity.

Moreover, insulin resistance (IR) and hyperinsulinemia are essential elements in obesity-associated conditions. The actions of insulin involve a complex signalling network, with insulin receptor-mediated tyrosine phosphorylation of insulin receptor substrates (IRS), activation of PI3K/Akt pathway, and enzymatic regulators for glucose uptake, glycogen and lipid synthesis (Saltiel and Kahn, 2001). Multiple insults driven by elevated circulating FFA and ROS may cause impaired insulin signalling, resulting in insulin resistance and hyperinsulinemia (Saltiel and Kahn, 2001; Jou *et al.*, 2008; Karagozian *et al.*, 2014). Evidence indicates that consistently high levels of insulin, insulin-like growth factors (IGF) and IRS-1 are involved in the development of angiogenesis and HCC, by activating oncogenic pathways of PI3K/Akt, vascular endothelial growth factor (VEGF) and mitogen-activated protein kinase (MAPK) (Tanaka *et al.*, 1997; Kim *et al.*, 1998). Insulin also activates the gene expression of transcription factor steroid regulatory element-binding protein (SREBP) -1c to promote hepatic *de novo* lipogenesis (Ferre and Foufelle, 2007). Increased activation of SREBP-1c due to insulin resistance/hyperinsulinemia would therefore worsen steatosis. Additionally, carbohydrate response element-binding protein (ChREBP) is another key transcription factor responsible for promotion of glycolysis and *de novo* lipogenesis (Iizuka and Horikawa, 2008). It has been demonstrated that the degree of steatosis was markedly improved by liver-specific inhibition of ChREBP in mouse, decreasing lipogenic rates and improving insulin

sensitivity in liver, skeletal muscles and adipose tissue (Dentin *et al.*, 2006). Hence, *PNPLA3* is very likely to share a part of the underpinning mechanisms of NAFLD-HCC as it is regulated by both SREBP-1c and ChREBP (Rae-Whitcombe *et al.*, 2010; Perttila *et al.*, 2012a). Besides, PNPLA3 also has been suggested to have retinyl-palmitate lipase activity in human hepatic stellate cells, which may explain its pathophysiological role in liver fibrosis (Pirazzi *et al.*, 2014).

Taken together with our data, the observation that carriage of the rs738409 GG genotype was associated with more poorly differentiated tumours amongst those with histologically characterised NAFLD-HCC, these findings suggest that *PNPLA3* genotype may directly influence tumour biology. Additional studies will be required to establish how these effects of *PNPLA3* are mediated. However, the general concept of *PNPLA3* in determining the susceptibility to NAFLD-related HCC could be that accumulating fat promoted by PNPLA3 I148M, oxidative stress and the low-grade inflammatory response present in NAFLD, favour a pro-carcinogenic *milieu* within the liver (Baffy *et al.*, 2012). This interpretation would be consistent with previous studies that suggest a greater effect of the *PNPLA3* variant on HCC risk in steatotic liver diseases (e.g. ALD, and here NAFLD) than that observed in non-steatotic conditions such as viral hepatitis (Sookoian and Pirola, 2013; Trepo *et al.*, 2014).

The high prevalence of NAFLD within the general population and the potential for disease progression both to cirrhosis and HCC poses a major challenge to existing healthcare infrastructure (Anstee *et al.*, 2011b). Understanding the contribution of *PNPLA3* to this process may theoretically have relevance to future preventive or therapeutic strategies targeting NAFLD-HCC. Ultrasonography and serological examinations (alpha-fetoprotein, AFP) ('EASL-EORTC clinical practice guidelines: management of hepatocellular carcinoma,' 2012) are presently used surveillance strategies, but even when targeting the cirrhotic population, their cost-effectiveness is debatable (Sarasin *et al.*, 1996; Trevisani *et al.*, 2007). Like most common diseases, NAFLD-related HCC is a complex disease trait with risk influenced by a combination of genetic and environmental factors. Whilst the odds ratios reported here are substantial and highly statistically significant, these do not mean that the *PNPLA3* rs738409 variant is the sole driver for HCC or could alone be used to stratify individual HCC risk (Manolio, 2013). However, if the results of the current study are corroborated, *PNPLA3* rs738409 genotype could potentially be included in

a broader multi-factorial risk assessment to help physicians to identify those amongst the expanding population of obese individuals with NAFLD at greatest risk of HCC, both in the presence and absence of cirrhosis (Baffy *et al.*, 2012; Dyson *et al.*, 2014). If the current findings are validated, a prospective economic evaluation of strategies incorporating assessment of *PNPLA3* rs738409 C>G status to target surveillance to those at greatest HCC risk may be warranted.

It should however be noted that the NAFLD-HCC cohort contained only 100 patients in the case-control analyses (NAFLD vs. NAFLD-HCC) for both *PNPLA3* and *TM6SF2*. However, a significant association was only observed between the *PNPLA3* rs738409 genotype and NAFLD-related HCC, not *TM6SF2* rs58542926. One possible explanation for the differences between the two SNPs in the association with NAFLD-HCC could be due to their minor allele frequencies. In the Northern European population, minor allele frequencies of 0.22 and 0.07 were reported for *PNPLA3* rs738409 and *TM6SF2* rs58542926, respectively. The numbers increased slightly and were found to be 0.34 (rs738409) and 0.12 (rs58542926) in our NAFLD cohort. Combined with a relatively modest *TM6SF2* rs58542926 minor allele frequency, the current study had approximately 70% power to detect an effect if an additive genetic model and risk similar to that seen for fibrosis (described in Chapter 4) is assumed (α=0.05) (Purcell *et al.*, 2003). Failure to detect an association for *TM6SF2* due to limited statistical power cannot therefore be completely excluded but would seem unlikely. Studies using larger cohorts of NAFLD-HCC patients than are presently available will be required to provide sufficient power to study this further. Alternatively, differences between the two proteins in overall function (e.g. the reported additional role for PNPLA3 in retinyl-palmitate lipase activity) may explain the association of *PNPLA3* with fibrosis and HCC but *TM6SF2* with fibrosis only.

In summary, we report a striking association between *PNPLA3* rs738409 and the development of NAFLD-HCC, which is independent of other known risk factors. These data highlight the importance of understanding the contribution of PNPLA3 I148M to NAFLD-HCC pathogenesis, and if validated, may contribute to a tailored approach to the cost effective surveillance and detection of NAFLD-HCC in those at greatest risk.

Chapter 6. Investigation of the Functional Significance of the *PNPLA3* rs738409 and *TM6SF2* rs58542926

6.1 Introduction

As described in Chapter 1**,** the two SNPs in the genes *PNPLA3* (rs738409 c.444 C>G, p.I148M) and *TM6SF2* (rs58542926 c.449 C>T, p.E167K) are significantly associated with severity of NAFLD. Our statistical data in the previous Chapters also suggested that the *PNPLA3* rs738409 SNP is a strong genetic risk factor associated with susceptibility to steatosis, inflammation and fibrosis using a well-characterised histologically proven NAFLD cohort. Moreover, we demonstrated the novel finding that this SNP was independently associated with the end-stage of NAFLD - hepatocellular carcinoma (data described in Chapter 5). Therefore, investigating the functional role of *PNPLA3* in NAFLD pathogenesis is important, including how this variant may biologically influence development of steatosis and fibrosis.

As described in detail in section 1.7.3.6, recent research suggests that PNPLA3 mediates hydrolysis of acylglycerols and the I148M, which results in reduced enzymatic activity, causes fat accumulation in the liver. (He *et al.*, 2010; Huang *et al.*, 2011). The PNPLA3 I148M variant has also been reported to alter hepatic lipid metabolism especially TG remodelling in lipid droplets within hepatocytes (Li *et al.*, 2012; Ruhanen *et al.*, 2014). Beyond steatosis, its role in underlying biological mechanisms through which *PNPLA3* influences progression to NASH and hepatic fibrosis still remain obscure. One study recently published by Pirazzi *et al.* has shed some light on the association between PNPLA3 I148M and hepatic fibrogenesis, demonstrating that PNPLA3 may possess a role of retinyl-palmitate hydrolase in human HSCs (Pirazzi *et al.*, 2014). *PNPLA3* was reported to be highly expressed in both the retina and the liver thus pointing to a key feature within the two organs, retinol metabolism. Retinol is converted into retinyl ester for storage in quiescent HSCs and is released during stellate cell activation (Lee and Jeong, 2012; Pirazzi *et al.*, 2014). Pirazzi *et al.* further demonstrated that *PNPLA3* is indeed highly expressed in HSCs, the dominant contributors to fibrogenesis. Incubation of primary human HSCs or immortalized human HSCs – LX-2 with retinol and palmitate up to 48 hours showed that a time-dependent increase in lipid droplet accumulation coexisted with a down-regulation of PNPLA3 protein expression, suggesting that PNPLA3 expression may also be regulated by retinol availability in human HSCs. Additionally, overexpressed wild-type PNPLA3 in HSCs resulted in a substantial reduction of lipid droplet content, an effect that was lost with I148M. These findings

reinforce that PNPLA3 is a regulator of lipid metabolism influenced by retinol levels and the I148M variant is a loss-of-function mutation (Pirazzi *et al.*, 2014). How activation of HSC, collagen deposition and fibrogenesis may be affected by this mutated protein remains unknown but however indicates that PNPLA3 may be expressed in multiple cell types. Thus, challenges remain in targeting relevant cell types and optimising experimental conditions to further investigate the key role of *PNPLA3* in NAFLD.

In respect of *TM6SF2* rs58542926 SNP, unlike studies for *PNPLA3*, little is known in terms of the pathogenic role of TM6SF2 in NAFLD. The current understandings of this protein are that: 1) TM6SF2 is a multi-pass membrane protein (Carim-Todd *et al.*, 2000), 2) is highly expressed in liver, kidney and intestines (Kozlitina *et al.*, 2014; Surakka *et al.*, 2015), 3) is mainly localized to the ER and ER-Golgi intermediate compartments (ERGIC) (Mahdessian *et al.*, 2014), and 4) both *in vivo* (Holmen et al., 2014; Kozlitina et al., 2014) and *in vitr*o (Holmen *et al.*, 2014; Kozlitina *et al.*, 2014; Mahdessian *et al.*, 2014) functional studies indicate that *TM6SF2* has effects on fat retention in the liver, lipid efflux and alteration in the number and average size of lipid droplets. Taking the current knowledge of *TMS6F2* together with our data in Chapter 4 that the *TM6SF2* rs58542926 polymorphism was strongly linked to hepatic fibrosis, it appears that this variant is an important determinant of clinical outcome across several facets of metabolic syndrome related end-organ damage.

To investigate of the functional significance of the *PNPLA3* rs738409 and *TM6SF2* rs58542926 in fatty liver disease, a series of studies were performed as described in this Chapter. In this preliminary investigation, I aimed to establish *in vitro* models to demonstrate the effect of fatty acid treatment on PNPLA3 and TM6SF2 expression. Commercially available human hepatoblastoma/hepatoma cell lines were characterised in detail and compared with primary hepatocytes. Additionally, the effect of retinol-palmitate treatment on PNPLA3 expression in the immortalized human HSCs – LX-2 was also investigated.

6.2 Methods

6.2.1 Cell culture

6.2.1.1 *Cell lines employed and cell culture conditions*

Three human HCC cell lines (Hep3B, HUH-7 and SNU182), and one human hepatoblastoma cell line (HepG2) were obtained from Gillian Patman, Northern Institute for Cancer Research, Newcastle University. The human hepatic stellate cell line LX-2 was kindly provided by Prof. Fiona Oakley, Institute for Cellular Medicine, Newcastle University. Cell culture conditions and DNA/RNA isolation are described in section 2.4.1 and section 2.4.2, respectively. Materials used are listed in Table 6.1 and Table 6.2.

Cultured primary human hepatocytes were purchased from Biopredic International (Saint-Grégoire, France, catalogue number: HEP220). These primary hepatocytes were isolated from a single donor, 54-year old male Caucasian, who had been diagnosed with liver metastases from carcinoma of the colon. Alternative diagnoses were excluded, including excess alcohol intake, chronic viral hepatitis (hepatitis B and hepatitis C), hypertension, diabetes, cardiovascular disease, and human immunodeficiency virus infection (HIV).

The isolation of hepatocytes was performed by the supplier using an undisclosed protocol (for commercial reasons). Primary hepatocytes were cultured by the supplier in a 6-well plate for one day after isolation with undisclosed culture conditions. Cell morphology was maintained and cell viability and culture confluence were reported to be 88% and 98%, respectively.

The 6-well plate was dispatched and delivered on the second day after isolation. After arrival, the cells were washed once with 1X PBS (Lonza, Slough, UK; catalogue number: BE17-516F) and the culture medium was replaced by the basal hepatic cell medium (Biopredic International, Saint-Grégoire, France, catalogue number: MIL600, 100 ml) supplemented with the 10X long term culture medium (Biopredic International, Saint-Grégoire, France, catalogue number: ADD271, 11 ml). All the solutions for cell culture were filter-sterilized by 0.22 μm filter (Merck Millipore, Hertfordshire, UK). The cells were incubated at 37°C in an atmosphere with 5% humidified CO_2. On the following morning, the hepatocytes were harvested for DNA (section 2.4.2.1), RNA (section 2.4.2.1) and protein extraction (section 2.4.2.3).

Table 6.1: Complete RPMI medium for growing HepG2, Huh-7, Hep3B and SNU182 cells.

Name	Supplier
RPMI 1640 500 ml	Lonza, Cat: BE12-167F
Sodium Pyruvate 100 mM (5 ml)	Lonza, Cat: BE13-115E
L-glutamine 200 mM (5 ml)	Lonza, Cat: BE17-605E
Non-essential amino acids 100X (5 ml)	Lonza, Cat: BE13-114E
10% Foetal bovine serum FBS (50 ml)	Lonza, Cat: DE14-801F
Penicillin/ Streptomycin (5 ml)	Lonza, Cat: DE17-603E

Table 6.2: Complete RPMI medium for growing LX-2 cells.

Name	Supplier
RPMI 1640 500 ml	Lonza, Cat: BE12-167F
Sodium Pyruvate 100 mM (5 ml)	Lonza, Cat: BE13-115E
L-glutamine 200 mM (5 ml)	Lonza, Cat: BE17-605E
Non-essential amino acids 100X (5 ml)	Lonza, Cat: BE13-114E
5% Foetal bovine serum FBS (50 ml)	Lonza, Cat: DE14-801F
Penicillin/ Streptomycin (5 ml)	Lonza, Cat: DE17-603E

6.2.1.2 *Genotyping and gene expression*

Genotyping was conducted using TaqMan SNP genotyping assays for *PNPLA3* rs738409 (reference number: C___7241_10, Cat. # 4351379) and *TM6SF2* rs58542926 (reference number: C___8709053_10, Cat. # 4351379) purchased from Applied Biosystems Inc., USA. The procedure of genotyping is described in section 2.3.4.2. TaqMan gene expression assays were purchased from Applied Biosystems Inc., USA for *PNPLA3* (Hs00228747_m1, Cat: #4331182), *TM6SF2* (Hs00403495_m1, Cat: #4331182) and *GAPDH* (Dm01843827_s1, Cat. # 4331182) as the reference gene. Materials used and performance of reverse transcription and gene expression are described in section 2.4.2.3. Unpaired student T-test was adapted for statistical analyses using 2-^delta CT value.

6.2.1.3 *Preparation of BSA-bound oleic acid*

Oleic acid (317 µl) was taken from the stock solution (SIGMA, Cat: O1008, 3.15M), and was added to 50 ml of a 0.15 M (8.76 mg/ml) NaCl solution at 70°C to reach a concentration of 20 mM. Another 25 ml of 3.4 mM (224 mg/ml) bovine serum albumin (BSA) already dissolved in Dulbecco's modified Eagle's medium-F12 (DMEM-F12) at 37°C was prepared, filtered by 0.22 µm filter (Millex-GP), and kept at 37°C. Once the 20 mM oleic acid solution was fully mixed, 25 ml of it was then added to 25 ml 3.4 mM BSA-DMEM solution at 70°C, under constant agitation to provide the final solution of BSA-bound oleic acid (10 mM). The final solution of BSA-bound oleic acid was adjusted to pH 7.4, sterilized by 0.22 µm filter (Millex-GP), and stored at -20°C for fatty acid treatments on cell lines. Materials used are listed in Table 6.3. Selected cells were grown under the same culture conditions and treated with a final concentration of 1 mM oleic acid, loading for 8, 24 and 48 hours.

Table 6.3: Materials used for BSA-bound oleic/palmitic acids.

Name	Supplier
Oleic acid	SIGMA, Cat: O1008 Mwt 282.46; Density 0.891/ml
Palmitate (palmitic acid)	SIGMA, Cat: P9767
BSA	SIGMA, Cat: A7030
DMEM F-12 **(Dulbecco's modified Eagle's medium-F12) 500ml**	Lonza, Cat: BE12-719F

6.2.1.4 *Retinol/palmitate treatment on LX-2*

25 mg of retinol powder purchased from Sigma (Cat: R7632) was first dissolved in 1 ml of ethanol to make the original stock at 87.2 mM, 25.8 µl of it was then diluted in a 45 ml of 5% BSA-PBS solution to make the 50 µM BSA-bound retinol stock solution. The stock solution was adjusted to pH 7.4 and was sterilized by a 0.22 µm filter.

Sodium palmitate (palmitic acid) was purchased from SIGMA (Cat: P9767). 278 mg palmitate was first dissolved in a 50 ml of 0.15 M NaCl solution (8.76mg/ml) at 70°C for 1h to reach a concentration of 20 mM. 25 ml of the palmitate solution was then added to another 25 ml solution of 3.4 mM BSA (224 mg/ml) already dissolved in DMEM-F12 and stored at 37°C. The 50 ml mixture was under constant agitation at 70°C to make the final BSA-conjugated palmitate solution (10 mM). Afterwards, the final solution was adjusted to pH 7.4, sterilized using 0.22 µm filter (Millex-GP), and stored at -20°C (materials are listed in Table 6.3).

The study design and the procedure for retinol/palmitate treatment of LX-2 are shown in Figure 6.1- 6.4. Cells were divided into 4 groups depending on the culture conditions: untreated (group A), treated with retinol 10 µM alone (group B), with palmitate 300 µM alone (group C) and group D treated with both retinol (10 µM) and palmitate (300 µM). Four different time points were selected as read-out references (0 hour, 12 hours, 24 hours, and 48 hours). It has been suggested that an increased retinol uptake within HSCs occurred in the presence of palmitate pre-incubation prior to retinol loading (Randolph and Ross, 1991; Vogel *et al.*, 2000). Except for the group A and B, group C and D were therefore pre-incubated with palmitate (300 µM) for 12 hours prior to further incubation of palmitate or a mixture of retinol and palmitate, thus the actual durations of palmitate loading for these two groups were 0 hour, 24 hours, 36 hours, and 60 hours. All the standard media or conditioned media (supplemented with either retinol, palmitate or both) were changed at the same time to minimise variations. Once the experiment was complete, cells were washed twice by 1X PBS and harvested for Western blotting, gene expression analysis and Oil red O staining.

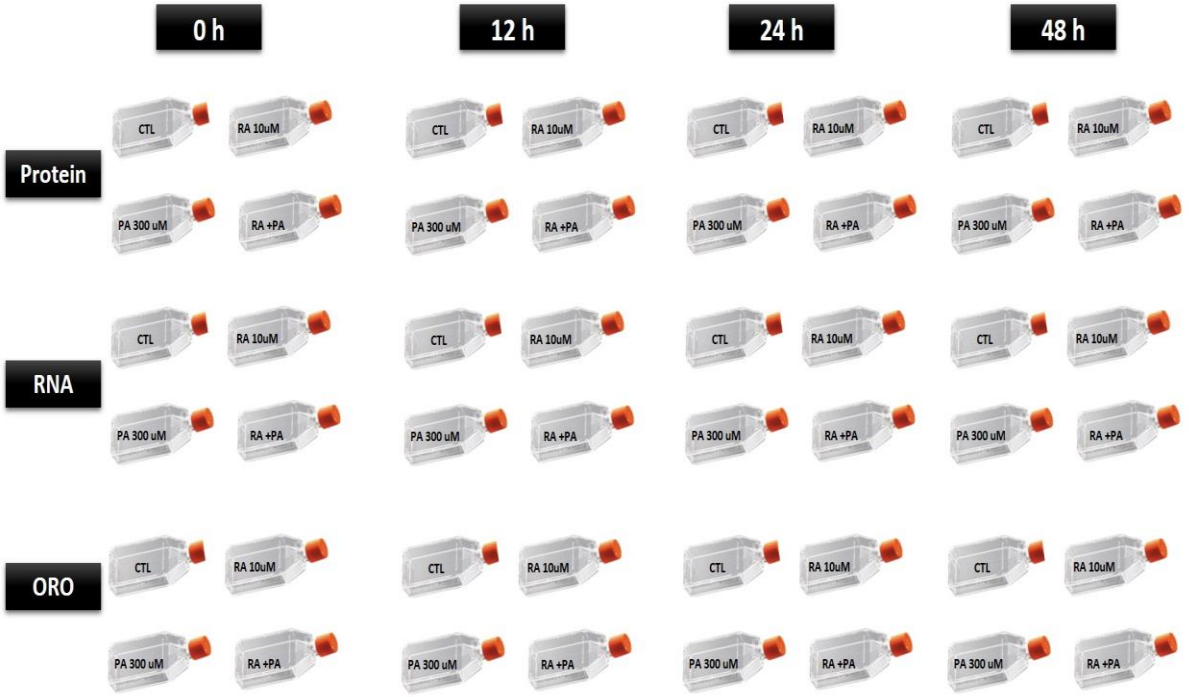


Figure 6.1: Layout of study design and purposes for retinol or/and palmitate treatment on LX-2. PA, palmitate; RA, retinol.

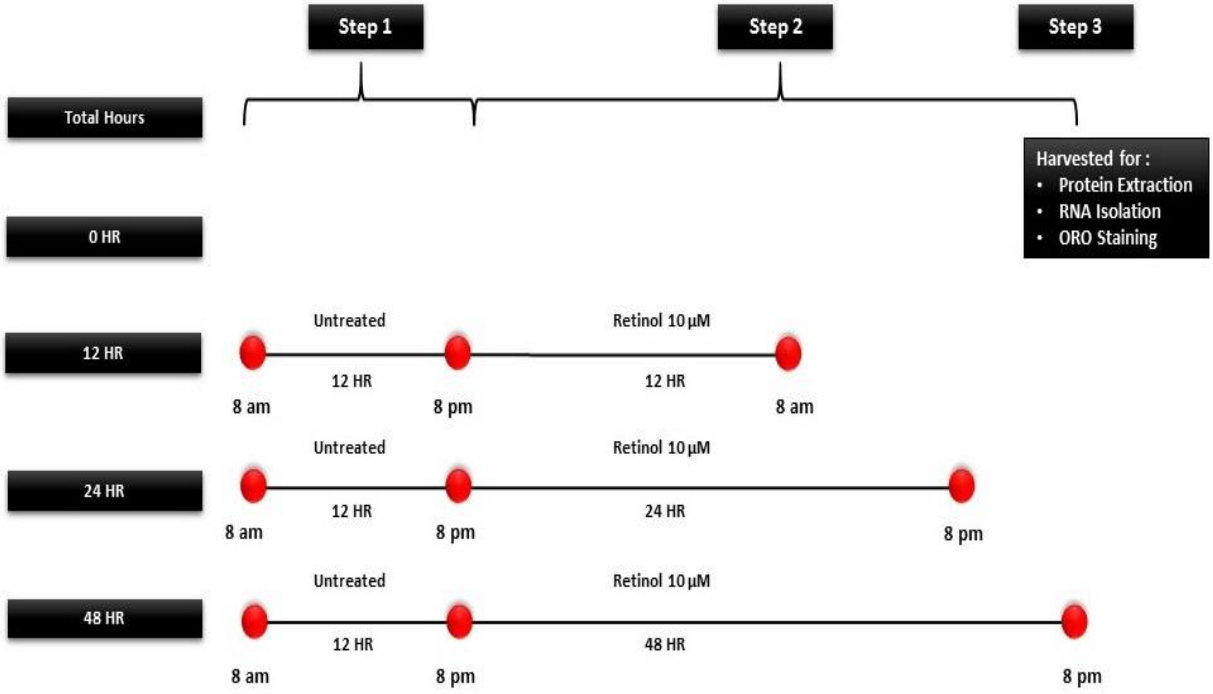


Figure 6.2: Study procedure for group B, LX-2 treated with retinol (10 µM) alone.

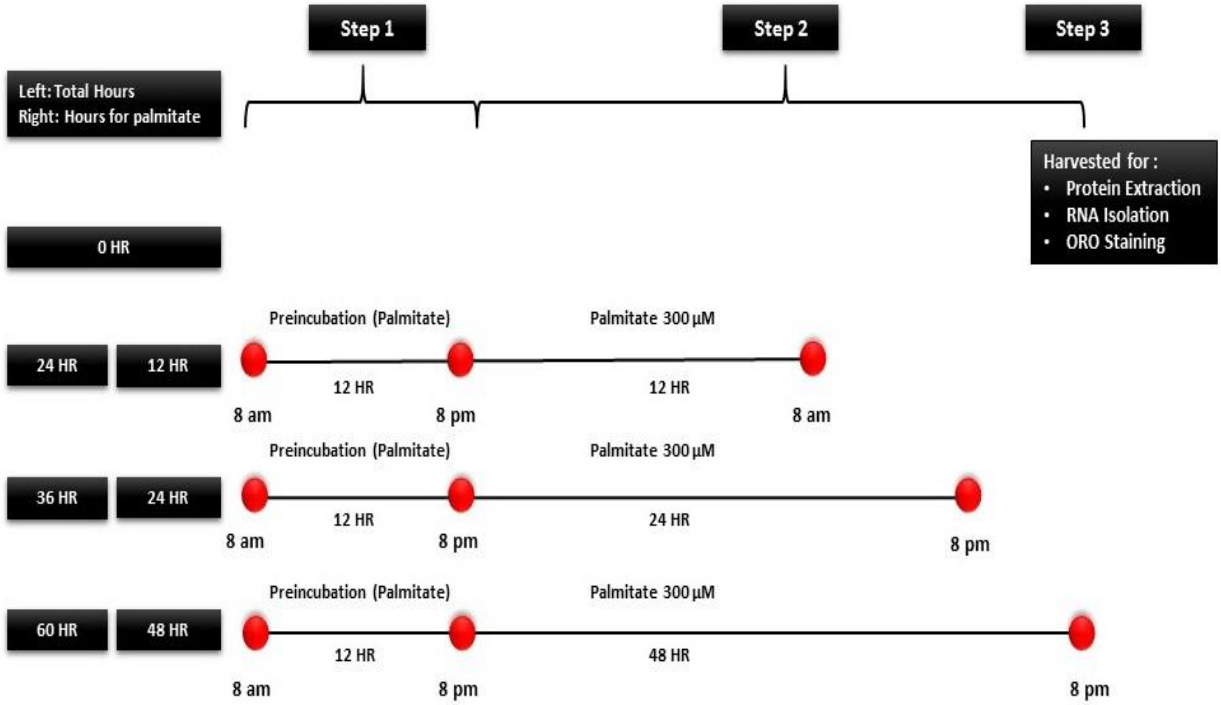


Figure 6.3: Study procedure for group C, LX-2 treated with palmitate (300 µM) alone.

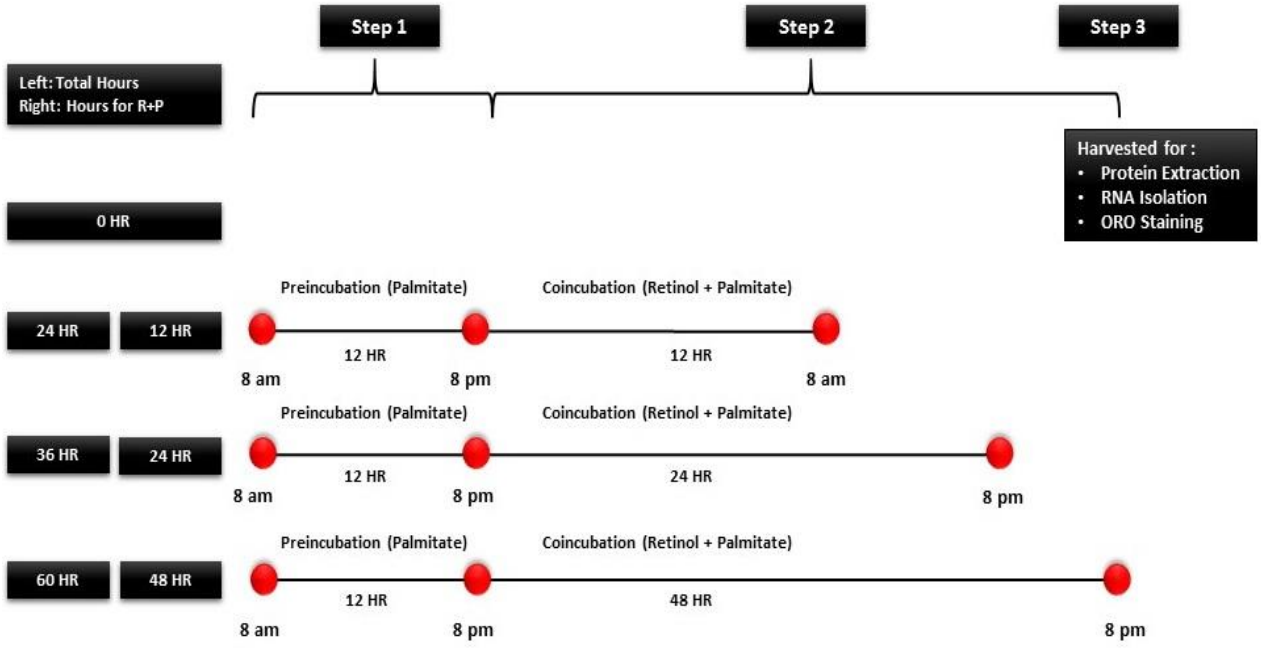


Figure 6.4: Study procedure for group D, LX-2 treated with both retinol (10 µM) and palmitate (300 µM).

6.2.1.5 *Western-blotting*

A series of Western-blotting procedures including protein extraction, protein concentration determination, polyacrylamide gel electrophoresis of protein samples, electrophoretic blotting procedure of proteins separated by SDS-PAGE, blocking, probing and immunodetection of protein blots, and staining of the PVDF membrane are described thoroughly in section 2.4.3. The rabbit polyclonal anti-PNPLA3 was purchased from Abcam, Cambridge, UK (ab81874, with 1:900 dilution) and the rabbit polyclonal anti-GAPDH was purchased from Santa Cruz Biotechnology, Inc. (sc-25778, with 1:25000 dilution). Two commercial anti-TM6SF2 were employed, the mouse polyclonal anti-TM6SF2 from Abcam (ab169629) and the rabbit polyclonal anti-TM6SF2 from ANTYBODY VERIFY (AAS00444C).

6.2.1.6 *Oil Red O staining*

Oil Red O is a fat-soluble dye used for staining of neutral triglycerides and lipids on frozen or paraffin tissues. Cells were fixed overnight using a 10% formalin solution (VWR, prolabo, Cat: 361387P), and were then incubated with 60% isopropanol (VWR, prolabo, Cat: 20842.330) for 5 min at room temperature after washing with 1 ml of sterile water twice. Having been dried completely, cells were loaded with 1 ml of 60% Oil Red O working solution (6 parts of Oil Red O purchased from SIGMA (Cat: O-0625) and 4 parts of 100% isopropanol) and left on the bench for 10 min incubation at room temperature. At the final stage, cells were washed 4 times by sterile water, and all images were acquired under the microscope.

6.3 Results

6.3.1 The role of PNPLA3 polymorphism rs738409 in vitro

6.3.1.1 *PNPLA3 rs738409 genotypes of selected cell lines*

To investigate the effect of PNPLA3-I148M *in vitro* three human hepatoma cell lines (Hep3B, HUH-7 and SNU182), one hepatoblastoma cell line (HepG2), primary human hepatocytes and one human hepatic stellate cell line (LX-2) were employed. Each cell line was genotyped for *PNPLA3* rs738409. Interestingly, 50% of cell lines targeted are homozygous mutant (GG). Subsequently, the basal expression of *PNPLA3* (both RNA and protein levels) among these cell lines were investigated and compared to primary human hepatocytes. Genotypes of *PNPLA3* (rs738409) in these cell lines are listed in Table 6.4.

6.3.1.2 *Basal mRNA levels and protein expression of PNPLA3*

Primary hepatocytes and the five cell lines were characterized for *PNPLA3* mRNA expression in triplicate. It should be noted that the primary hepatocytes used in this gene expression analysis were technical replicates (n=3) as they were isolated from a single donor whilst the other commercial cell lines were biological replicates (n=3). However, primary human hepatocytes showed the highest expression of *PNPLA3*, whereas the lowest level was found in LX-2. Using primary hepatocytes as the reference, *PNPLA3* is highly expressed in primary hepatocytes compared to HepG2 (P=0.001), HUH-7 (P=0.038), SNU182 (P=0.001), and LX-2 (P=0.02). No significant difference was found between Hep3B and primary hepatocytes (P=0.43). With regard to expression of PNPLA3, it is highly expressed in primary human hepatocytes compared to the others which shared a similar level, Figure 6.5 (this preliminary data contains a low n-value, n=1).

Table 6.4: *PNPLA3* rs738409 genotypes of the selected cell lines.

***PNPLA3* rs738409 Genotypes**	**HepG2**	**Hep3B**	**HUH7**	**SNU182**	**LX-2**	**Human Hepatocytes**
	GG	**CC**	**GG**	**CG**	**GG**	**CC**

CC for homozygous wild-type and GG for homozygous mutant.

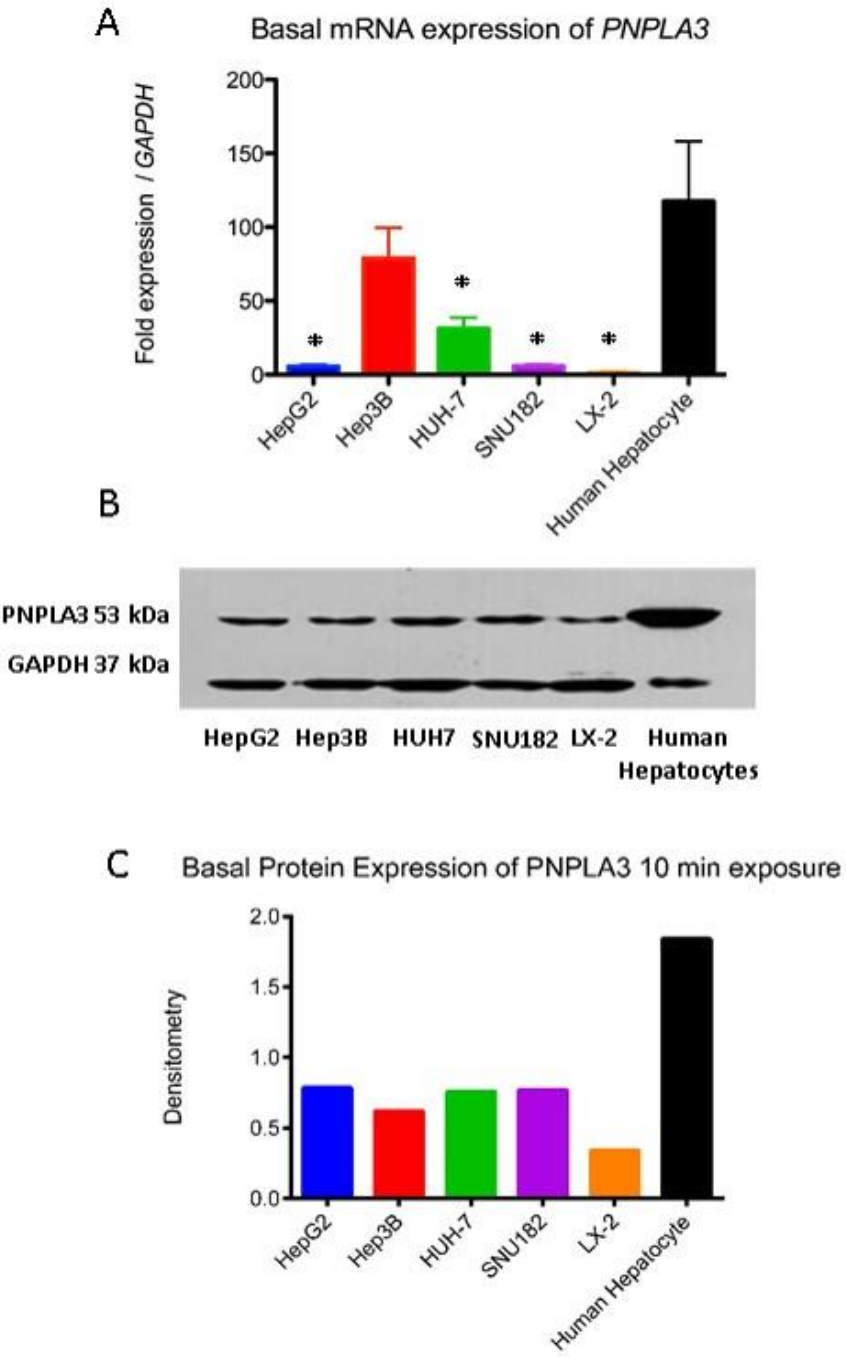


Figure 6.5: Basal messenger RNA levels and protein expression of PNPLA3 in HepG2, Hep3B, HUH-7, SNU182, LX-2 and primary human hepatocytes. (A) Basal mRNA levels of *PNPLA3* (n=3) and (B) Western blot analysis of PNPLA3 (n=1) in the selected cell lines. (C) Densitometry based on Western blot results. Unpaired student T-test was used as the test for statistical significance test using 2-^delta CT values, with human hepatocytes as the reference. * represents *p* value <0.05.

6.3.2 The effect of oleic acid treatment on PNPLA3 expression in vitro

To investigate differences in expression of PNPLA3 following fat loading in cells of different genotypes, two cell lines which possess two copies of common (C) or minor (G) allele respectively were selected: Hep3B homozygous wild-type CC and HUH-7 homozygous mutant GG. Hep3B and HUH-7 were treated with the same culture condition, the medium supplemented with 1 mM BSA-bound oleic acid, and were fat-loaded for the same time points of 8, 24, and 48 hours, Figure 6.6.

For the experiments performed by using Hep3B (homozygous wild-type, CC), protein intensities were quantified by densitometry at three exposure times, 2, 5, and 17 minutes (Figure 6.6. A). An increased expression of PNPLA3 was observed in Hep3B compared to the control 8 hours after start of treatment. However, this result could be affected by a bubble between the SDS gel and the blotting membrane, therefore no firm conclusion could be made upon this. Distinct alterations in expression of PNPLA3 were not detectable when comparisons were made between the untreated and the treated groups at the 24- and 48-hour time points.

For the results of fat-loading on HUH-7 displayed in Figure 6.6 (B), the densitometric analyses were performed by three different exposure times, 15, 30 seconds, and 2 minutes. Once again, no major differences could be identified in the protein level of PNPLA3 between each groups at the three time points.

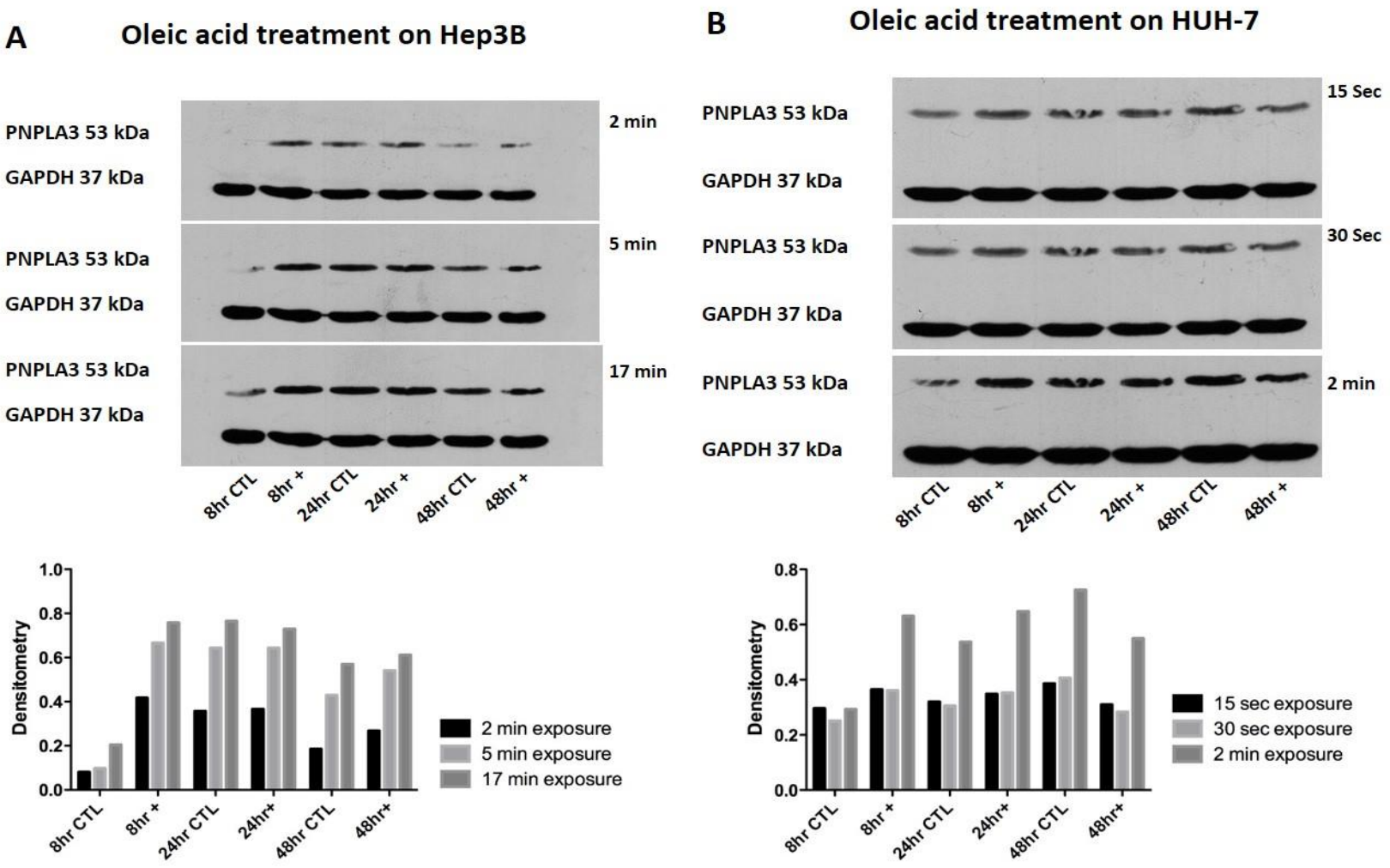


Figure 6.6: Effect of oleic acid on expression of PNPLA3 in the selected cell lines. (A) Hep3B and (B) HUH-7 cells were treated for 8 hours, 24 hours, and 48 hours with 1 mM Oleic acid (n=1). The expression of PNPLA3 were visualised by Western blot analysis and quantified using densitometry at different time exposures (+ represents treated condition, CTL, control).

6.3.3 The effect of retinol/palmitate treatment on PNPLA3 expression in LX-2

A recent study reported that PNPLA3 may be involved in retinyl-palmitate hydrolysis in hepatic stellate cells (HSC/LX-2), by demonstrating that a decreased expression of PNPLA3 coexisted with an increased lipid droplet accumulation in the presence of retinol and palmitate incubation for up to 48 hours (Pirazzi *et al.*, 2014). To investigate the role of PNPLA3 in a fibrogenic cell type, I therefore performed a similar study to that from Pirazzi *et al*. but with LX-2 employed only to test the effect of retinol/palmitate treatment upon expression of PNPLA3.

As described in section 6.2.1.4 (methods), this study is composed of 4 groups (untreated, treated with retinol or palmitate alone, and with both) for 4 time points (0, 12, 24, 48 hours). The study design and procedure is shown in Figure 6.1.-6.4. At each time point, cells were harvested for RNA/protein extraction and Oil red O staining.

Before reporting our results, the limitations of these preliminary data should be noted. A key limitation is that the cells were overly stressed by the incubations with palmitate (alone or in combination with retinol) due to the nature of saturated fat toxicity. In fact, the viability of LX-2 could only be maintained for 24 hours of incubation. Beyond this time point, the LX-2 were severely stressed and cell death/cell detachment from the flasks could be easily observed. The palmitate treatment of the LX-2 cells resulted in a strongly reduced RNA 280/260 ratio (data not shown) due to a large loss of cell numbers. Therefore, I focussed on analysing protein expression rather than gene expression as a more acceptable amount of protein extracted from the cells allowed Western blotting to be performed (results are displayed in Figure 6.7).

Taken the results from those blots and densitometric data, an increased expression of PNPLA3 was found in the untreated controls and those treated with palmitate alone. The expression of PNPLA3 in those treated with retinol alone increased slightly at the beginning followed by a trend towards a decrease. Treatment with combined retinol and palmitate showed a reduced level of PNPLA3 in a time-dependent manner, similar to the findings published by (Pirazzi *et al.*, 2014).

Lastly, Oil Red O staining was used to visualise intracellular neutral lipid content in LX-2 incubated with palmitate, retinol or both combined (Figure 6.8). No major

differences were detectable for the group incubated with retinol alone (Figure 6.8 - B). In regard to the groups treated with palmitate only (Figure 6.8 - C) or in combination with retinol (Figure 6.8 - D), it is clear that both palmitate and combined treatments induced intracellular lipid uptake in a time-dependent manner. However, reduced cell viability and cell loss after 24 hours incubation were observed with treatment by palmitate or both compounds during the experiment thus quantification of intracellular lipid accumulation was not comparable. Also noteworthy is that the study described in this section has only been conducted once, hence all data should be interpreted with caution.

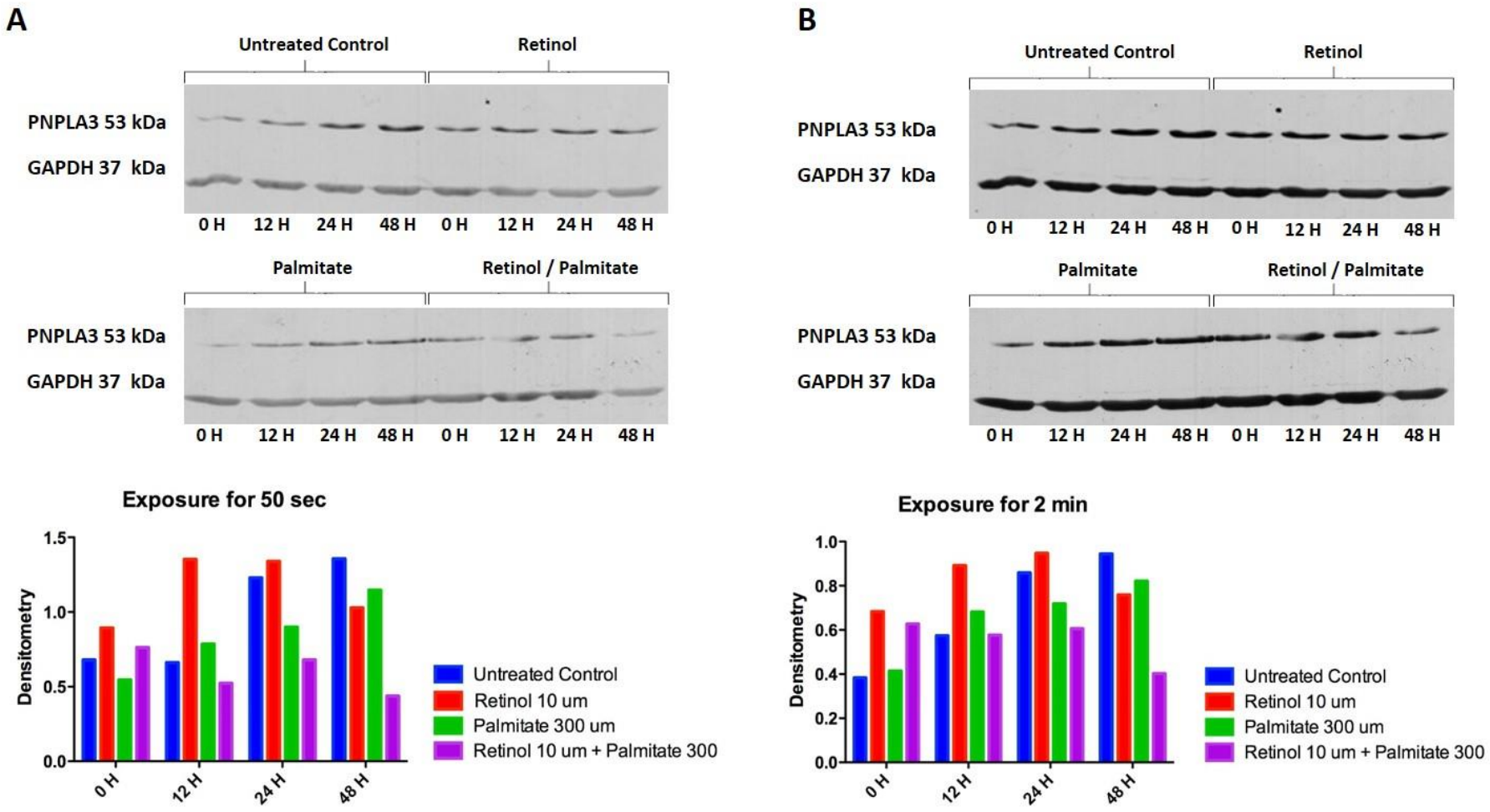


Figure 6.7: Effect of retinol/palmitate treatment on expression of PNPLA3 in LX-2 for 0, 12, 24, and 48 hours. (A) Exposure time for 50 seconds, and (B) exposure time for 2 minutes. The signal intensities presented by densitometric analyses with different time exposures are shown underneath.

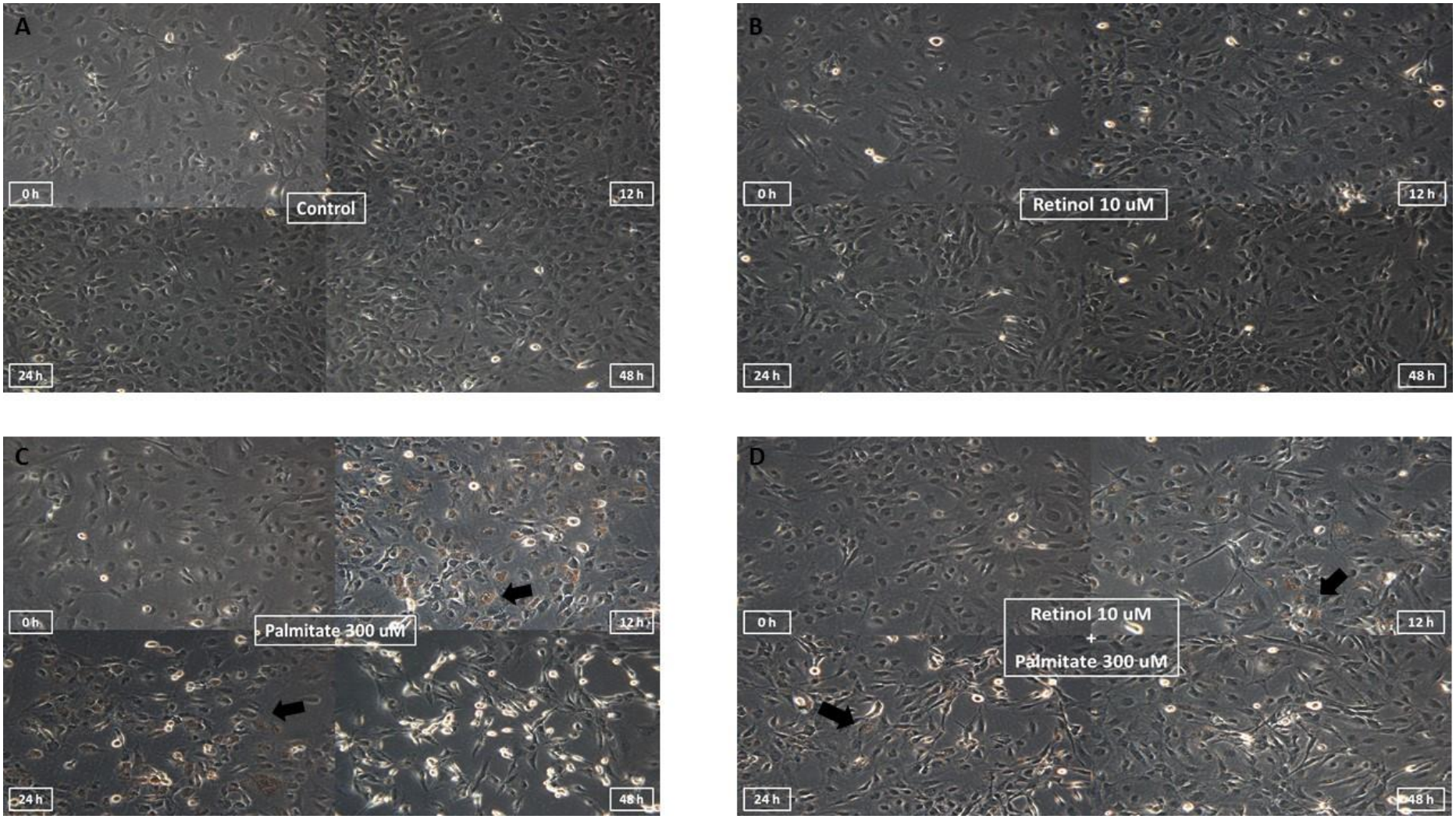


Figure 6.8: Oil red O staining of LX-2 at 4 time points. (A) Untreated control, (B) LX-2 treated with 10 µM retinol alone, (C) LX-2 treated with 300 µM palmitate alone, (D) LX-2 treated with 10 µM retinol and 300 µM palmitate. Arrows showed above are the examples of stained lipids within the cells.

6.3.4 The role of TM6SF2 rs58542926 polymorphism in vitro

6.3.4.1 *TM6SF2 rs58542926 genotypes of each cell line*

The functional study of TM6SF2-E167K is covered briefly in this thesis as it is an ongoing study. Primary human hepatocytes, human hepatoma and hepatic stellate cell lines (HepG2, Hep3B, HUH-7, SNU182 and LX-2) were genotyped for *TM6SF2* rs58542926 (results displayed in Table 6.5). All of them are homozygous wild-type (CC) except the SNU182 cells are heterozygotes (CT). A high genotype frequency of homozygous wild-type identified in the cell lines is conceivable as the minor allele frequency of this SNP is only 0.07, observed in a Northern European population (http://browser.1000genomes.org). Interestingly, the low frequency of minor allele (T) carriage among these cell line is also in line with the non-statistically significant association between rs58542926 C>T and NAFLD-related HCC (already described in Chapter 5).

6.3.4.2 *Basal mRNA levels and protein expression of TM6SF2*

To get a general idea of *TM6SF2* expression, quantitative gene expression (n=1) and Western blot analyses were used to investigate mRNA levels and protein expression of this gene in the aforementioned cell lines. All the cell lines used in this gene expression analysis were technical replicates. Primary human hepatocytes showed the highest expression level of *TM6SF2*, whereas HUH-7 are the second highest and the others shared a similar level of *TM6SF2* mRNA (Hep3B, LX-2 and SNU182), Figure 6.9. The expression of TM6SF2 of these cell lines was tested by Western blot analysis. However, the two commercial antibodies employed in this study failed to detect the protein (data not shown).

Table 6.5: *TM6SF2* rs58542926 genotypes of the selected cell lines.

TM6SF2 rs58542926 Genotypes	HepG2	Hep3B	HUH7	SNU182	LX-2	Human Hepatocytes
	CC	CC	CC	CT	CC	CC

CC for homozygous wild-type and CT for heterozygous.

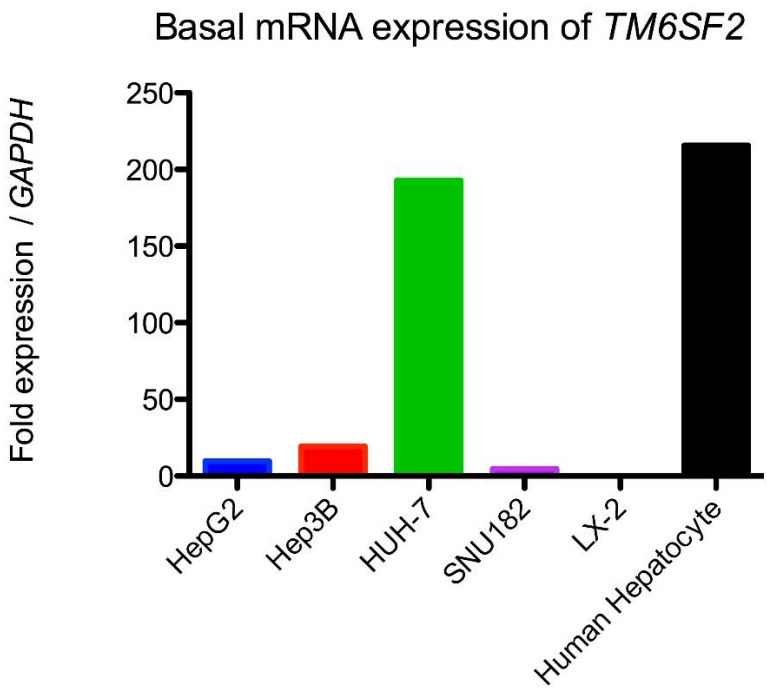


Figure 6.9: Basal expression levels of *TM6SF2* mRNA in the selected cell lines, n=1.

6.4 Discussion

To investigate the functional significance of PNPLA3 (mainly) and TM6SF2, a series of preliminary *in vitro* studies were described in this chapter, including genotyping for both *PNPLA3* rs738409 and *TM6SF2* rs58542926 SNPs, the quantitative gene expression of endogenous mRNA levels and immunoblotting for protein expressions of the two genes, and fatty acids treatments with oleic acid/retinol/palmitate using human hepatoma, hepatoblastoma, and hepatic stellate cell lines.

Regarding the study for *PNPLA3* rs738409, 50% of the cell lines were identified as homozygous mutant (GG) whilst only 30% in homozygous wild-type (CC). Although the sample number of the cell lines employed is quite small, only 3 hepatoma cell lines and the primary hepatocytes were obtained from only one individual, enriched G-allele in these cell lines roughly mirrors the statistical association between the rs738409 variant and liver cancer. Investigations on mRNA/protein expression of *PNPLA3* were subsequent performed to test any variations by different genotypes. Relatively high levels of *PNPLA3* mRNA were observed in primary hepatocytes and Hep3B (homozygous wild-type for both) compared to the others either are heterozygous CG (SNU182) or homozygous mutant (HepG2, HUH-7, and LX-2). For the protein levels, PNPLA3 is once again highly expressed in the primary hepatocytes while the other cell lines possess a similar but lower level. Notably, the levels of *PNPLA3* mRNA varied among the commercial cell lines however no major differences in expression of PNPLA3 were detected. This discrepancy of the expression patterns between mRNA and protein of PNPLA3 in those cell lines may be explained by the following points: 1) Cancer cells or artificial immortalized cells (LX-2) may have alterations in transcription/translation/metabolism process compared to primary cells. Several biological factors have been suggested may influence translational efficiency, including RNA secondary structure, weak Shine-Dalgarmo sequence (imperfect complementarity to ribosomal RNA), blocking translation by regulatory proteins (i.e. R-protein) and small RNA, low ribosome occupancy, and codon bias. Additionally, protein turnover may also influence the correlation between mRNA and protein abundances as individual protein half-lives vary from seconds to hours (Maier *et al.*, 2009). 2) Technical factors: varied sensitivity in methodology used for quantitative analysis of mRNA and protein. Real time PCR provides fast, precise and accurate results, by collecting data as the reaction is proceeding not measured at End-Point (plateau) as traditional PCR does. An increase in reporter fluorescent signal is directly proportional to the number of

amplicons generated thus it is more sensitive for DNA and RNA quantitation. On the contrary, Western blot is a semiquantitative technique with several processing-steps for protein detection. Various parameters during the processing steps may influence sensitivity of the results: blot transfer conditions, blot irregularities, concentration of antibodies and ECL substrate, exposure time, camera sensitivity, colour format and adjustments, band definition and background correction (Heidebrecht *et al.*, 2009). 3) Ethnicity, genetic background may be one of the determinants which exert different levels of *PNPLA3* expression, as HepG2 cells are derived from a Caucasian patient while Hep3B, HUH-7, SNU182 are from African American, Japanese, and Korean, respectively. Notably, LX-2 is a transfected cell line which was made immortal for research purposes (Xu *et al.*, 2005), and the origin of this cell line remains elusive. 4) Complex disease trait is a phenotype contributed from more than one SNP, therefore other SNPs near rs738409 may also have effects on the expression of *PNPLA3*. 5) A coding SNP is generally considered to slightly alter function of a protein rather than changing expression. Though there is growing evidence for the significant role of *PNPLA3* in severity of NAFLD, it is uncertain that whether under this amino acid substitution of PNPLA3-I148M, the other 'rescue' pathways in human body are activated to cover this subtle defect so as to maintain the original function or metabolic role that *PNPLA3* is involved in.

The next step after characterization of the cell lines for *PNPLA3* rs738409 was to investigate whether any alteration/activation of PNPLA3 expression could be observed in response to fat loading by different genotypes as FFA metabolism is one of the key elements for NAFLD development (Ricchi *et al.*, 2009), and PNPLA3 was suggested to mediate hydrolysis of acylglycerols with a strong preference for oleic acid as the acyl moiety (He *et al.*, 2010; Huang *et al.*, 2011). Additionally, studies also revealed that a decreased unsaturated/saturated fatty acid ratio in serum, fat and liver tissue was found in obese NAFLD patients (de Almeida *et al.*, 2002; Videla *et al.*, 2004). An *in vitro* study also reported that oleic acid (unsaturated FFA) was more steatogenic but less toxic than palmitate (saturated FFA) (Ricchi *et al.*, 2009). Therefore, a preliminary study of oleic acid treatments on Hep3B (wild-type, PNPLA3-148I) and HUH-7 (mutant isoform, PNPLA3-148M) was performed with time points of 8, 24 and 48 hours selected.

Although the previous *in vitro* studies on PNPLA3 showing an inactivated TG hydrolysis coexisted with intracellular fat deposition resulted from overexpression of

PNPLA3-148M under fat loading (He *et al.*, 2010; Perttila *et al.*, 2012a), no apparent effect on expression of PNPLA3 was in either the homozygous wild-type or homozygous mutant cells after treatment with oleic acid in the present study. Once again, this preliminary study has low n-value (n=1) thus no firm conclusion could be made, and the present data is not comparable to those studies mentioned above as the study designs such as cell type employed and procedure of overexpression are totally different. Replicate studies are definitely required, including further optimization of experimental conditions. These include the concentration and components of fatty acids applied and the duration of treatment. Other co-culture variables such as glucose or insulin could also be considered as *PNPLA3* is regulated by *ChREBP* and *SREBP-1c* under hyperglycaemia and hyperinsulinemia, respectively, to promote *de novo* lipogenesis (Shimomura *et al.*, 1999b; Rae-Whitcombe *et al.*, 2010; Perttila *et al.*, 2012b).

As *PNPLA3* is highly expressed in HSCs, which have a key role in hepatic fibrogenesis and are the main reservoir of retinoids stored in lipid droplets (Senoo *et al.*, 2010), and suggested to be regulated by retinol availability in human HSCs (Pirazzi *et al.*, 2014), other fat loading experiments using retinol/palmitate/combined treatment on immortalized LX-2 cells were also performed in the present work to demonstrate the role of PNPLA3 in a fibrogenic cell type. In particular this was to further investigate the underlying mechanisms of the association between the *PNPLA3* rs738409 variant and fibrosis. The study design is similar to the work by Pirazzi *et al*. with the difference that LX-2 cells were not only treated with retinol and palmitate but also with the palmitate separately. Our results showed that treatment with palmitate alone induced increased expression of PNPLA3 whilst a reduced level was found in the combined treatment with no effect was detectable under incubation with retinol alone. Although our data shares a similar trend with the work demonstrated by Pirazzi *et al.,* the limitations of this present study should be considered; reduced cell viability due to palmitic toxicity together with low n-value of 1 and it is therefore inappropriate to make a solid conclusion. Replicate experiments with optimised procedures are required to clarify whether PNPLA3 expression can be induced by retinol, palmitate or a combined effect as Pirazzi *et al.* only reported the combined effect of retinyl-palmitate on expression of PNPLA3 in primary human HSCs. Nonetheless, they stated that the same results were found in LX-2 with these data not included in the publication, though the absence of these data makes a direct comparison difficult. However, studies recently published provided more information

for the functional role of PNPLA3 in fibrogenesis; overexpression of wild-type PNPLA3 induced a reduction in secretion of proteins involved in extracellular matrix formation and remodelling, secretion of matrix metallopeptidase 2 and tissue inhibitor of metalloproteinase 1 and 2 (Pingitore *et al.*, 2016). Conversly, PNPLA3-I148M potentiates the fibrogenic features of HSCs (Bruschi *et al.*, 2017).

To date, the majority of studies examining the functional role of PNPLA3 or TM6SF2 in NAFLD involve either overexpression or knock-down of the protein, resulting in alterations in the entire protein level, which cannot completely reflect functional effects due to a SNP mutation. CRISPR/Cas9 may be a possible approach to address this issue, it is a nuclease guided by small RNAs through Watson-Crick base pairing with target DNA, representing a system that is markedly easier to design, highly specific, efficient and well-suited for high-throughput and multiplexed gene editing for a variety of cell types and organisms (Ran *et al.*, 2013). By adopting CRISPR/Cas9, the ease of customization, higher targeting efficiency and the ability to facilitate multiplex genome editing (simply knock-in the SNP of interest), a greater precision and prediction in functional effect of one SNP could be easily achieved. Other challenges remain in targeting different organs or cell types to establish *in vitro* models for NAFLD. Lipotoxicity-mediated inflammation and prolonged liver injury by oxidative stress triggers proliferation/activation of hepatic progenitor cells which can differentiate into hepatocytes and cholangiocytes. Activated progenitor cells are correlated to ductular reaction initiating fibrosis via epithelial-mesenchymal transition contributing to the portal myofibroblast pool or production of chemotactic agents for inflammatory cells activating progenitor cells (Wruck *et al.*, 2017). Macrophage-derived foam cells due to excessive lipid deposition are also presented in progression to NASH (Walenbergh *et al.*, 2013). By future work targeting cell types not only involving hepatocytes and HSCs but also hepatic progenitor cells or macrophages (Kupffer cells in the liver), and optimized experimental variables (fatty acid compositions and their saturation degree for instance) in combination with CRISPR/Cas9, the functional roles of PNPLA3 and TM6SF2 in NAFLD progression may be better assessed.

Chapter 7. General Discussion

7.1 General discussion

Non-alcoholic fatty liver disease represents a spectrum of progressive liver disease characterised by increased HTGC in the absence of excess alcohol consumption. It incorporates NASH, fibrosis and ultimately cirrhosis and is strongly associated with features of the metabolic syndrome (obesity, insulin resistance/T2DM and dyslipidaemia) (Anstee *et al.*, 2013b). NAFLD is best considered as a complex disease trait resulting from a combination of environmental exposures acting on a susceptible polygenic background (Altshuler *et al.*, 2008; Hirschhorn and Gajdos, 2011; Anstee and Day, 2013). Although NAFLD is estimated to affect approximately one-third of the population in many developed countries, only a minority of patients exhibit progressive steatohepatitis leading to cirrhosis and/or hepatocellular carcinoma (HCC), suggesting that inter-patient genetic variations may exert a greater effect upon the pathogenesis of NAFLD (Anstee *et al.*, 2013b; Loomba and Sanyal, 2013). Therefore, investigation into the genetic contribution to NAFLD may be useful to determine those are at high risk in developing severe disease and possibly provide tailored surveillance.

This thesis reports a series of studies examining the association of genetic variations in two genes *patatin-like phospholipase domain-containing 3* (*PNPLA3,* rs738409 c.444 C>G, p.I148M) and *transmembrane 6 superfamily member 2*, (*TM6SF2*, rs58542926 c.449 C>T, p.E167K) with severity of NAFLD and risk of NAFLD-associated hepatocellular carcinoma (HCC). The patient cohort used is the largest histologically characterised NAFLD cohort assembled up to date and represents the full spectrum of disease from simple steatosis through steatohepatitis to advanced fibrosis and cirrhosis. Patients are of European-Caucasian descent (n=1,005) and were assembled during the FLIP (Fatty liver: Inhibition of Progression) project funded by the European Commission. To test the association between SNPs of interest and the end stage of NAFLD, another cohort of 100 Northern European Caucasian patients with primary HCC arising on a background of NAFLD was identified and were recruited from UK and Switzerland. As a comparator with NAFLD-related HCC population, a cohort of 275 biopsy-proven NAFLD patients of different stages of disease but without clinical evidence of HCC was also assembled. Healthy controls were not studied due to the invasive nature of liver-biopsy, hence the analyses were conducted comparing severity of disease within the NAFLD spectrum. This approach differs from most case-control studies where comparison is made to a normal healthy population. However, in the absence of non-invasive methods to establish absence

of NAFLD this would have required the conduct of liver biopsy in heathy people and so was not ethically appropriate or feasible.

Using candidate gene association studies in the FLIP cohort to address first the role of *PNPLA3*, the rs738409 variant was significantly associated with steatosis (OR 1.87, 95%CI 1.52-2.29, P=2.98x10^{-9}), steatohepatitis (β 0.174±0.059, 95%CI 0.056-0.291, P=0.0037) and fibrosis (OR 1.4, 95%CI 1.15-1.7, P=0.0006), independent of age at biopsy, gender, BMI and presence of T2DM. Our data supports the well-established assertion that the *PNPLA3* rs738409 is a major genetic determinant for susceptibility to NAFLD. Subsequently, adopting a case-control analysis in a cohort of 100 NAFLD-associated HCC and a cohort of 275 NAFLD patients without HCC, carriage of the rs738409 minor (G) allele was shown to significantly associate with an increased risk of developing NAFLD-associated HCC, independent of potential confounding factors including gender, age at diagnosis, T2DM, BMI, and presence of advanced fibrosis/cirrhosis (OR 2.26, 95%CI 1.23-4.14, P=0.0082). As 22-54% of NAFLD-HCC occur in the absence of cirrhosis (Sanyal *et al.*, 2010; Ertle *et al.*, 2011), the effect of *PNPLA3* on HCC risk independent of degree of fibrosis was further examined selecting those patients with coexistent cirrhosis (NAFLD-HCC n=67, NAFLD-Cohort n=26). Here, although the study-group sizes were relatively small in this sub-analysis, the effect of rs739409 variant remained significantly associated with NAFLD-HCC (OR 2.06, 95%CI 1.07-3.94, X^2 4.78, P=0.029). When the non-cirrhotic group were studied however, the effect did not attain statistical significance, although this may be due to insufficient statistical power given that only 23 cases were available in non-cirrhotic HCC group.

The work in this thesis not only confirms the association between *PNPLA3* with NAFLD severity but also provides the most robust evidence to date of its association with HCC risk which was published in 2014 (Liu *et al.*, 2014a). Several studies have validated the *PNPLA3* effect on NAFLD-related HCC (Casper *et al.*, 2016; Vespasiani-Gentilucci *et al.*, 2016). It is well-acknowledged that NAFLD as a complex disease trait, unlike Mendelian disorders, is often considered an outcome attributed from multiple genetic modifiers, where each genetic variant may alone be insufficient to cause disease (Hirschhorn and Gajdos, 2011; Anstee and Day, 2013). Strikingly, a 12-fold increased risk of HCC was demonstrated in our work for rs738409 minor (G) allele homozygotes relative to C-allele homozygotes (OR 12.19, 95%CI 6.89-21.58, P<0.0001), when comparing the genotype frequencies between the NAFLD-HCC

patients and an unselected UK general population sample (the MRC/Wellcome Trust UK 1958 Birth Cohort). In response to our data, Krawczyk *et al.* compared the effect of *PNPLA3* to the monogenic condition hereditary haemochromatosis due to *HFE* rs1800562 (A>G, p.C282Y) as the rs738409 SNP carries such strong genetic effect. Their detailed comparison demonstrated that carriage of either variant confers a similar risk of HCC (Krawczyk *et al.*, 2015). Although statistical techniques to assess proportion of variability explained by genetic variations are imperfect (Witte *et al.*, 2014), our data indicates a Population Attributable Risk (PAR) of *PNPLA3* rs738409 for HCC of 55% and an area under the receiver operator characteristics curve (AUROC) of 0.68 attributable to this variant (Liu *et al.*, 2014a). Further supporting the independent role of *PNPLA3* rs738409 as a risk factor for HCC, it is telling that other genetic variants that are strong modifiers of NAFLD severity and fibrosis progression such as *TM6SF2* (Liu *et al.*, 2014b), do not appear to confer an additional, independent increased risk of HCC (Liu *et al.*, 2014b; Dongiovanni *et al.*, 2015). We agree with Krawczyk *et al.* that *PNPLA3* is an important genetic risk factor for liver disease. For the moment, we would however caution against a move towards considering "*PNPLA3*-associated NAFLD", or by extension "*PNPLA3*-associated HCC", as distinct, monogenic conditions analogous to *HFE* in haemochromatosis but rather suggest that *PNPLA3* should be considered a strong modifier within a complex, polygenic disease trait that is subject both to genetic and substantial environmental influence (e.g. due to dietary factors and the intestinal microbiota) (Anstee *et al.*, 2013b).

Given the mounting evidence of an association between *PNPLA3* and HCC (Burza *et al.*, 2012; Trepo *et al.*, 2014), it is timely to consider the clinical utility of *PNPLA3* genotyping to assist in patient risk stratification. It has been suggested that, assuming a sensitivity of 80%, an odds ratio of >50 is required to control false-positive rates to an acceptable level of <10% (Manolio, 2013). A re-analysis of the data from our previously published article (Liu *et al.*, 2014a) to assess sensitivity/specificity is shown in Table 7.1. Based on these data, the use of *PNPLA3* genotyping alone to *positively* predict risk of HCC is unlikely to be tenable. However, even in this HCC-enriched dataset, the *negative predictive value* was substantially greater, both within the NAFLD vs. NAFLD-HCC comparison and the NAFLD-HCC vs. unselected background population comparison, and so it may be that knowledge of *PNPLA3* rs738409 genotype has utility to select out those individuals least likely to develop HCC and therefore least likely to benefit from surveillance. As Krawczyk *et al*.

suggest, studies addressing the interaction between *PNPLA3* and *HFE* in determining HCC risk would be of great interest. Given the increasing prevalence of NAFLD-HCC and consequent clinical need for improved risk-stratification care pathways (Dyson *et al.*, 2014), we would also suggest that large, prospective studies are needed urgently to validate our findings and to determine the utility and health-economic merits of a multi-factorial risk stratification that incorporates *PNPLA3* rs738409 genotype along with other recognised risk factors for HCC.

Recently, a novel non-synonymous SNP in *TM6SF2* (rs58542926 c.449 C>T, p.Glu167Lys (E167K) was identified to be associated with ^{1}H-MRS quantified HTGC by an exome-wide association study (Kozlitina *et al.*, 2014). This variant has also been associated with dyslipidaemia and cardiovascular risk (Holmen *et al.*, 2014). It was therefore pertinent to determine whether this variant also affected risk of steatohepatitis or fibrosis in NAFLD. Using the aforementioned cohorts (FLIP and NAFLD-HCC ones), our data demonstrated the first time that, in addition to its association with steatosis, the rs58542926 SNP is significantly associated with stage of fibrosis in NAFLD. *P*-values are $P=5.57 \times 10^{-5}$, $P=0.014$, and $P=6.36 \times 10^{-6}$ for our discovery (n=349), validation (n=725), and combined cohort (n=1,074), respectively (Liu *et al.*, 2014b). Our data once again confirmed that the causal SNP showing a strong genetic signal on 19p13.11 region for steatotic association would actually be the *TM6SF2* rs58542926 variant rather than the *NCAN* gene (rs2228603 C>T) that is in strong LD (D'=0.926, r^2 =0.798) with *TM6SF2* rs58542926.

After our findings were published, this genetic effect of *TM6SF2* rs58542926 has been well replicated, including a recent study carriage of this variant increased susceptibility to alcohol-induced cirrhosis (Buch *et al.*, 2015). Relevance to simple steatosis has been confirmed in studies performed in adults (Dongiovanni *et al.*, 2015; Sookoian *et al.*, 2015; Wang *et al*., 2015; Zhou *et al.*, 2015) and in paediatric patients (Mancina *et al.*, 2015; Goffredo *et al.*, 2016; Grandone *et al.*, 2016). This variant not only influences initiation of steatosis but also steatohepatitis (necroinflammation and lobular ballooning) and our key finding - its effect upon fibrosis was also validated (Dongiovanni *et al.*, 2015). Two studies from China (Wong *et al.*, 2014) and Argentina (Sookoian *et al.*, 2015) disagreed with our data by reporting no associations between rs58542926 and stage of fibrosis. The former study only used a community-based cohort rather than selecting individuals afflicted by NAFLD, given generally low MAF of *TM6SF2* rs58542926 (MAF 0.07), they reported that this variant

has limited impact on NAFLD. Whereas the other Chinese group with a more appropriate study design demonstrated a positive association between NAFLD and rs58542926 (Wang *et al.*, 2015).

The current thesis confirms the association of *PNPLA3* with NAFLD severity and provides novel evidence of its association with HCC risk. In addition, it demonstrates for the first time that *TM6SF2* is associated with NAFLD-fibrosis severity. Certainly, our data needs further validation such as GWAS or candidate gene association study with a considerably larger sample size of NAFLD patients and healthy populations of different ethnicities to test whether these effects apply to all ethnic groups. Moreover, it would be worthwhile to investigate whether there is a joint effect of these two SNPs as one study performed in Chinese Han population reported the strong additive effects of the risk alleles of *PNPLA3* and *TM6SF2* with an overall significance between the number of risk alleles and NAFLD (OR 1.64, 95%CI 1.34–2.01, P=1.4×10^{-6}) (Wang *et al.*, 2016). Furthermore, it would be interesting to test whether carriers of the *PNPLA3* rs738409 variant show a different response compared to those with wild-type genotypes to drugs currently under investigation for treatment of NAFLD such as Elafibranor (Ratziu *et al.*, 2016).

The association of *PNPLA3* and NAFLD-HCC demonstrated in this thesis may provide a useful tool that the *PNPLA3* rs738409 genotyping, acting upon a broader multi-factorial risk assessment, to identify those NAFLD patients that are at higher risk of advanced NAFLD and possibly HCC development (particularly homozygous mutant, GG) or those are at low risk of less progressive NAFLD (homozygous wild-type, CC). It seems very unlikely that this approach alone could predict or exclude the occurrence of HCC, but might lower the possibility of HCC by early life style interventions. Ideally, if genotyping data was available at the early stage of NAFLD, patients could be encouraged to modify their life style including diet and physical activity and the frequency of targeted surveillance could be individualised. Given the current knowledge that CVD is closely associated with NAFLD and is a very common cause of death in these patients (Adams *et al.*, 2005; Ekstedt *et al.*, 2006; Targher *et al.*, 2007b; Rafiq *et al.*, 2009; Soderberg *et al.*, 2010; Treeprasertsuk *et al.*, 2012; Zhou *et al.*, 2012), the dual role of the *TM6SF2* rs58452926 SNP cannot be ignored; the minor T-allele is associated with increased risk of NAFLD while the common C-allele is associated with dyslipidaemia and cardiovascular risk, making targeted advice on the basis of this genotype more difficult. However, genotyping for these

two genetic risk factors should be considered further as a means of stratifying clinical care.

The functions of both PNPLA3 and TM6SF2 remain elusive despite numerous studies already preformed especially for PNPLA3. In spite of this, positive associations have been detected and well replicated in genetic studies. Further studies on the roles of both proteins are needed with better cell culture models and well-designed experimental conditions. Our group is currently conducting a GWAS on additional biopsy-proven NAFLD samples from Northern Europe (EPoS: Elucidating Pathways of Steatohepatitis) to identify more SNPs associated with NAFLD. Use of this larger (approx. 2000 case) cohort may lead to the discovery of novel genetic risk factors. In the meantime, use of CRISPR/Cas9 to knock-in the SNPs of interest could provide higher targeting efficiency and the ability to facilitate multiplex genome editing to further investigate functional roles of PNPLA3 and TM6SF2. A better understanding of these roles may lead to novel therapeutic approaches to NAFLD management in the future.

Table 7.1: Reanalysis focusing on clinical utility of *PNPLA3* genotype testing in HCC risk prediction.

Cohorts (n)	Genotype Comparison	Sensitivity	Specificity	PPV*	NPV*
NAFLD-HCC (100) vs. NAFLD Cohort (275)	**GG vs CC:**	51%	79%	47%	82%
	CG/GG vs CC:	72%	45%	32%	82%
NAFLD-HCC (100) vs. Background UK Pop^{n}(1476)	**GG vs CC:**	51%	92%	28%	97%
	CG/GG vs CC:	72%	59%	10%	97%

*PPV, positive predictive value; NPV, negative predictive value. PPV/NPV will alter according to the underlying prevalence of HCC in population assessed.

References

ABI-Biosystems (1997) *User Bulletin #2 ABI PRISM 7700 Sequence Detection System.* ABI Biosystems Inc.

Abrams, G.A., Kunde, S.S., Lazenby, A.J. and Clements, R.H. (2004) 'Portal fibrosis and hepatic steatosis in morbidly obese subjects: A spectrum of nonalcoholic fatty liver disease', *Hepatology*, 40(2), pp. 475-83.

Adams, L.A., Lymp, J.F., St Sauver, J., Sanderson, S.O., Lindor, K.D., Feldstein, A. and Angulo, P. (2005) 'The natural history of nonalcoholic fatty liver disease: a population-based cohort study', *Gastroenterology*, 129(1), pp. 113-21.

Addison, T. (1836) 'Observations on fatty degeneration of the liver.', *Guy's Hospital Reports*, 1:476-485.

Adler, M. and Schaffner, F. (1979) 'Fatty liver hepatitis and cirrhosis in obese patients', *Am J Med*, 67(5), pp. 811-6.

Al-Serri, A., Anstee, Q.M., Valenti, L., Nobili, V., Leathart, J.B., Dongiovanni, P., Patch, J., Fracanzani, A., Fargion, S., Day, C.P. and Daly, A.K. (2012) 'The SOD2 C47T polymorphism influences NAFLD fibrosis severity: evidence from case-control and intra-familial allele association studies', *J Hepatol*, 56(2), pp. 448-54.

Albanis, E. and Friedman, S.L. (2001) 'Hepatic fibrosis. Pathogenesis and principles of therapy', *Clin Liver Dis*, 5(2), pp. 315-34, v-vi.

Alberti, K.G. and Zimmet, P.Z. (1998) 'Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus provisional report of a WHO consultation', *Diabet Med*, 15(7), pp. 539-53.

Altshuler, D., Daly, M.J. and Lander, E.S. (2008) 'Genetic mapping in human disease', *Science*, 322(5903), pp. 881-8.

Angulo, P. (2002) 'Treatment of nonalcoholic fatty liver disease', *Ann Hepatol*, 1(1), pp. 12-9.

Angulo, P. (2006) 'NAFLD, obesity, and bariatric surgery', *Gastroenterology*, 130(6), pp. 1848-52.

Angulo, P. (2007) 'GI epidemiology: nonalcoholic fatty liver disease', *Aliment Pharmacol Ther*, 25(8), pp. 883-9.

Angulo, P. (2013) 'The Natural History of NAFLD, in Non-Alcoholic Fatty Liver Disease: A Practical Guide (eds G. C. Farrell, A. J. McCullough and C. P. Day)', *Wiley-Blackwell, Oxford, UK.* , (doi: 10.1002/9781118556153.ch4).

Angulo, P., Keach, J.C., Batts, K.P. and Lindor, K.D. (1999) 'Independent predictors of liver fibrosis in patients with nonalcoholic steatohepatitis', *Hepatology*, 30(6), pp. 1356-62.

Angulo, P., Kleiner, D.E., Dam-Larsen, S., Adams, L.A., Bjornsson, E.S., Charatcharoenwitthaya, P., Mills, P.R., Keach, J.C., Lafferty, H.D., Stahler, A., Haflidadottir, S. and Bendtsen, F. (2015) 'Liver Fibrosis, but No Other Histologic Features, Is Associated With Long-term Outcomes of Patients With Nonalcoholic Fatty Liver Disease', *Gastroenterology*, 149(2), pp. 389-97 e10.

Anstee, Q.M., Daly, A.K. and Day, C.P. (2011a) 'Genetics of alcoholic and nonalcoholic fatty liver disease', *Semin Liver Dis*, 31(2), pp. 128-46.

Anstee, Q.M., Darlay, R., Leathart, J., Liu, Y.L., Tordjman, J., Clement, K., Aithal, G., Valenti, L., Van Gaal, L., Stickel, F., Alison, M., Romero-Gomez, M., Andrade, R., Reeves, H., Bedossa, P., Pihlajamaki, J., Kotronen, A., Yki-Jarvinen, H., Chalasani, N., Kleiner, D.E., Rotter, J., Burt, A., Ratziu, V., Cordell, H.J., Daly, A.K., Day, C.P., Inhibition, F.F.L. and Investigators, N.C. (2013a) 'A Candidate-Gene Approach to Validation of Genetic Modifier Associations Using a Large Cohort with Histologically Characterised Non-Alcoholic Fatty Liver Disease', *Journal of Hepatology*, 58, pp. S46-S46.

Anstee, Q.M. and Day, C.P. (2013) 'The genetics of NAFLD', *Nat Rev Gastroenterol Hepatol*, 10(11), pp. 645-55.

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

Anstee, Q.M., McPherson, S. and Day, C.P. (2011b) 'How big a problem is non-alcoholic fatty liver disease?', *BMJ*, 343, p. d3897.

Anstee, Q.M., Targher, G. and Day, C.P. (2013b) 'Progression of NAFLD to diabetes mellitus, cardiovascular disease or cirrhosis', *Nat Rev Gastroenterol Hepatol*, 10(6), pp. 330-44.

Arase, Y., Suzuki, F., Kobayashi, M., Suzuki, Y., Kawamura, Y., Matsumoto, N., Akuta, N., Kobayashi, M., Sezaki, H., Saito, S., Hosaka, T., Ikeda, K., Kumada, H., Ohmoto, Y., Amakawa, K., Tsuji, H., Hsieh, S.D., Kato, K., Tanabe, M., Ogawa, K., Hara, S. and Kobayashi, T. (2011) 'The development of chronic kidney disease in Japanese patients with non-alcoholic fatty liver disease', *Intern Med*, 50(10), pp. 1081-7.

Aravinthan, A., Mells, G., Allison, M., Leathart, J., Kotronen, A., Yki-Jarvinen, H., Daly, A.K., Day, C.P., Anstee, Q.M. and Alexander, G. (2014) 'Gene polymorphisms of cellular senescence marker p21 and disease progression in non-alcohol-related fatty liver disease', *Cell Cycle*, 13(9), pp. 1489-94.

Argo, C.K. and Caldwell, S.H. (2009) 'Epidemiology and natural history of non-alcoholic steatohepatitis', *Clin Liver Dis*, 13(4), pp. 511-31.

Ascha, M.S., Hanouneh, I.A., Lopez, R., Tamimi, T.A., Feldstein, A.F. and Zein, N.N. (2010) 'The incidence and risk factors of hepatocellular carcinoma in patients with nonalcoholic steatohepatitis', *Hepatology*, 51(6), pp. 1972-8.

Atherton, H.J., Gulston, M.K., Bailey, N.J., Cheng, K.K., Zhang, W., Clarke, K. and Griffin, J.L. (2009) 'Metabolomics of the interaction between PPAR-alpha and age in the PPAR-alpha-null mouse', *Mol Syst Biol*, 5, p. 259.

Auinger, A., Valenti, L., Pfeuffer, M., Helwig, U., Herrmann, J., Fracanzani, A.L., Dongiovanni, P., Fargion, S., Schrezenmeir, J. and Rubin, D. (2010) 'A promoter polymorphism in the liver-specific fatty acid transport protein 5 is associated with features of the metabolic syndrome and steatosis', *Horm Metab Res*, 42(12), pp. 854-9.

Bacon, B.R., Farahvash, M.J., Janney, C.G. and Neuschwander-Tetri, B.A. (1994) 'Nonalcoholic steatohepatitis: an expanded clinical entity', *Gastroenterology*, 107(4), pp. 1103-9.

Baffy, G., Brunt, E.M. and Caldwell, S.H. (2012) 'Hepatocellular carcinoma in non-alcoholic fatty liver disease: an emerging menace', *J Hepatol*, 56(6), pp. 1384-91.

Baldini, M., Lohman, I.C., Halonen, M., Erickson, R.P., Holt, P.G. and Martinez, F.D. (1999) 'A Polymorphism* in the 5' flanking region of the CD14 gene is associated with circulating soluble CD14 levels and with total serum immunoglobulin E', *Am J Respir Cell Mol Biol*, 20(5), pp. 976-83.

Balkau, B. and Charles, M.A. (1999) 'Comment on the provisional report from the WHO consultation. European Group for the Study of Insulin Resistance (EGIR)', *Diabet Med*, 16(5), pp. 442-3.

Ballestri, S., Day, C.P. and Daly, A.K. (2011) 'Polymorphism in the farnesyl diphosphate farnesyl transferase 1 gene and nonalcoholic fatty liver disease severity', *Gastroenterology*, 140(5), pp. 1694-5.

Bambha, K., Belt, P., Abraham, M., Wilson, L.A., Pabst, M., Ferrell, L., Unalp-Arida, A. and Bass, N. (2012) 'Ethnicity and nonalcoholic fatty liver disease', *Hepatology*, 55(3), pp. 769-80.

Basantani, M.K., Sitnick, M.T., Cai, L., Brenner, D.S., Gardner, N.P., Li, J.Z., Schoiswohl, G., Yang, K., Kumari, M., Gross, R.W., Zechner, R. and Kershaw, E.E. (2011) 'Pnpla3/Adiponutrin deficiency in mice does not contribute to fatty liver disease or metabolic syndrome', *J Lipid Res*, 52, pp. 318-329.

Bataller, R. and Brenner, D.A. (2005) 'Liver fibrosis', *J Clin Invest*, 115(2), pp. 209-18.

Bechmann, L.P., Gastaldelli, A., Vetter, D., Patman, G.L., Pascoe, L., Hannivoort, R.A., Lee, U.E., Fiel, I., Munoz, U., Ciociaro, D., Lee, Y.M., Buzzigoli, E., Miele, L., Hui, K.Y., Bugianesi, E., Burt, A.D., Day, C.P., Mari, A., Agius, L., Walker, M., Friedman, S.L. and Reeves, H.L. (2012) 'Glucokinase links Kruppel-like factor 6 to

the regulation of hepatic insulin sensitivity in nonalcoholic fatty liver disease', *Hepatology*, 55(4), pp. 1083-93.

Bedossa, P. and Consortium, F.P. (2014) 'Utility and appropriateness of the fatty liver inhibition of progression (FLIP) algorithm and steatosis, activity, and fibrosis (SAF) score in the evaluation of biopsies of nonalcoholic fatty liver disease', *Hepatology*, 60(2), pp. 565-75.

Beer, N.L., Tribble, N.D., McCulloch, L.J., Roos, C., Johnson, P.R., Orho-Melander, M. and Gloyn, A.L. (2009) 'The P446L variant in GCKR associated with fasting plasma glucose and triglyceride levels exerts its effect through increased glucokinase activity in liver', *Hum Mol Genet*, 18(21), pp. 4081-8.

Bellentani, S., Bedogni, G., Miglioli, L. and Tiribelli, C. (2004) 'The epidemiology of fatty liver', *Eur J Gastroenterol Hepatol*, 16(11), pp. 1087-93.

Bellentani, S., Saccoccio, G., Masutti, F., Croce, L.S., Brandi, G., Sasso, F., Cristanini, G. and Tiribelli, C. (2000) 'Prevalence of and risk factors for hepatic steatosis in Northern Italy', *Ann Intern Med*, 132(2), pp. 112-7.

Benzie, I.F. (1996) 'Lipid peroxidation: a review of causes, consequences, measurement and dietary influences', *Int J Food Sci Nutr*, 47(3), pp. 233-61.

Bernard, S., Touzet, S., Personne, I., Lapras, V., Bondon, P.J., Berthezene, F. and Moulin, P. (2000) 'Association between microsomal triglyceride transfer protein gene polymorphism and the biological features of liver steatosis in patients with type II diabetes', *Diabetologia*, 43(8), pp. 995-9.

Bhala, N., Angulo, P., van der Poorten, D., Lee, E., Hui, J.M., Saracco, G., Adams, L.A., Charatcharoenwitthaya, P., Topping, J.H., Bugianesi, E., Day, C.P. and George, J. (2011) 'The natural history of nonalcoholic fatty liver disease with advanced fibrosis or cirrhosis: an international collaborative study', *Hepatology*, 54(4), pp. 1208-16.

Bian, Z. and Ma, X. (2012) 'Liver fibrogenesis in non-alcoholic steatohepatitis', *Front Physiol*, 3(doi:10.3389), p. 248.

Bircher, J. (1999) 'Oxford textbook of Hepatology', *Oxford University Pre*, 1, pp. Section1-13.

Bosch, F.X., Ribes, J., Cleries, R. and Diaz, M. (2005) 'Epidemiology of hepatocellular carcinoma', *Clin Liver Dis*, 9(2), pp. 191-211, v.

Bradford, M.M. (1976) 'A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding', *Anal Biochem*, 72, pp. 248-54.

Braillon, A., Capron, J.P., Herve, M.A., Degott, C. and Quenum, C. (1985) 'Liver in obesity', *Gut*, 26(2), pp. 133-9.

Brand, M.D. (2010) 'The sites and topology of mitochondrial superoxide production', *Exp Gerontol*, 45(7-8), pp. 466-72.

Browning, J.D. and Horton, J.D. (2004) 'Molecular mediators of hepatic steatosis and liver injury', *J Clin Invest*, 114(2), pp. 147-52.

Browning, J.D., Kumar, K.S., Saboorian, M.H. and Thiele, D.L. (2004a) 'Ethnic differences in the prevalence of cryptogenic cirrhosis', *Am J Gastroenterol*, 99(2), pp. 292-8.

Browning, J.D., Szczepaniak, L.S., Dobbins, R., Nuremberg, P., Horton, J.D., Cohen, J.C., Grundy, S.M. and Hobbs, H.H. (2004b) 'Prevalence of hepatic steatosis in an urban population in the United States: impact of ethnicity', *Hepatology*, 40(6), pp. 1387-95.

Brunt, E.M., Janney, C.G., Di Bisceglie, A.M., Neuschwander-Tetri, B.A. and Bacon, B.R. (1999) 'Nonalcoholic steatohepatitis: a proposal for grading and staging the histological lesions', *Am J Gastroenterol*, 94(9), pp. 2467-74.

Brunt, E.M., Kleiner, D.E., Wilson, L.A., Belt, P., Neuschwander-Tetri, B.A. and Network, N.C.R. (2011) 'Nonalcoholic fatty liver disease (NAFLD) activity score and the histopathologic diagnosis in NAFLD: distinct clinicopathologic meanings', *Hepatology*, 53(3), pp. 810-20.

Brunt, E.M., Kleiner, D.E., Wilson, L.A., Unalp, A., Behling, C.E., Lavine, J.E., Neuschwander-Tetri, B.A. and Appendix, N.C.R.N.l.o.m.o.t.N.S.C.R.N.c.b.f.i.t. (2009) 'Portal chronic inflammation in nonalcoholic fatty liver disease (NAFLD): a histologic marker of advanced NAFLD-Clinicopathologic correlations from the nonalcoholic steatohepatitis clinical research network', *Hepatology*, 49(3), pp. 809-20.

Bruschi, F.V., Claudel, T., Tardelli, M., Caligiuri, A., Stulnig, T.M., Marra, F. and Trauner, M. (2017) 'The PNPLA3 I148M variant modulates the fibrogenic phenotype of human hepatic stellate cells', *Hepatology*.

Buch, S., Schafmayer, C., Volzke, H., Becker, C., Franke, A., von Eller-Eberstein, H., Kluck, C., Bassmann, I., Brosch, M., Lammert, F., Miquel, J.F., Nervi, F., Wittig, M., Rosskopf, D., Timm, B., Holl, C., Seeger, M., ElSharawy, A., Lu, T., Egberts, J., Fandrich, F., Folsch, U.R., Krawczak, M., Schreiber, S., Nurnberg, P., Tepel, J. and Hampe, J. (2007) 'A genome-wide association scan identifies the hepatic cholesterol transporter ABCG8 as a susceptibility factor for human gallstone disease', *Nat Genet*, 39(8), pp. 995-9.

Buch, S., Stickel, F., Trepo, E., Way, M., Herrmann, A., Nischalke, H.D., Brosch, M., Rosendahl, J., Berg, T., Ridinger, M., Rietschel, M., McQuillin, A., Frank, J., Kiefer, F., Schreiber, S., Lieb, W., Soyka, M., Semmo, N., Aigner, E., Datz, C., Schmelz, R., Bruckner, S., Zeissig, S., Stephan, A.M., Wodarz, N., Deviere, J., Clumeck, N., Sarrazin, C., Lammert, F., Gustot, T., Deltenre, P., Volzke, H., Lerch, M.M., Mayerle, J., Eyer, F., Schafmayer, C., Cichon, S., Nothen, M.M., Nothnagel, M., Ellinghaus, D., Huse, K., Franke, A., Zopf, S., Hellerbrand, C., Moreno, C., Franchimont, D., Morgan, M.Y. and Hampe, J. (2015) 'A genome-wide association study confirms PNPLA3 and

identifies TM6SF2 and MBOAT7 as risk loci for alcohol-related cirrhosis', *Nat Genet*, 47(12), pp. 1443-8.

Bugianesi, E. (2007) 'Non-alcoholic steatohepatitis and cancer', *Clin Liver Dis*, 11(1), pp. 191-207, x-xi.

Bugianesi, E., Gastaldelli, A., Vanni, E., Gambino, R., Cassader, M., Baldi, S., Ponti, V., Pagano, G., Ferrannini, E. and Rizzetto, M. (2005) 'Insulin resistance in non-diabetic patients with non-alcoholic fatty liver disease: sites and mechanisms', *Diabetologia*, 48(4), pp. 634-42.

Bugianesi, E., Leone, N., Vanni, E., Marchesini, G., Brunello, F., Carucci, P., Musso, A., De Paolis, P., Capussotti, L., Salizzoni, M. and Rizzetto, M. (2002) 'Expanding the natural history of nonalcoholic steatohepatitis: from cryptogenic cirrhosis to hepatocellular carcinoma', *Gastroenterology*, 123(1), pp. 134-40.

Bugianesi, E., Manzini, P., D'Antico, S., Vanni, E., Longo, F., Leone, N., Massarenti, P., Piga, A., Marchesini, G. and Rizzetto, M. (2004) 'Relative contribution of iron burden, HFE mutations, and insulin resistance to fibrosis in nonalcoholic fatty liver', *Hepatology*, 39(1), pp. 179-87.

Burgdorf, K.S., Sandholt, C.H., Sparso, T., Andersen, G., Witte, D.R., Jorgensen, T., Sandbaek, A., Lauritzen, T., Sorensen, T.I., Madsbad, S., Hansen, T. and Pedersen, O. (2010) 'Studies of association between LPIN1 variants and common metabolic phenotypes among 17,538 Danes', *Eur J Endocrinol*, 163(1), pp. 81-7.

Burt, A.D., Lackner, C. and Tiniakos, D.G. (2015) 'Diagnosis and Assessment of NAFLD: Definitions and Histopathological Classification', *Semin Liver Dis*, 35(3), pp. 207-20.

Burza, M.A., Molinaro, A., Attilia, M.L., Rotondo, C., Attilia, F., Ceccanti, M., Ferri, F., Maldarelli, F., Maffongelli, A., De Santis, A., Attili, A.F., Romeo, S. and Ginanni Corradini, S. (2013) 'PNPLA3 I148M (rs738409) genetic variant and age at onset of at-risk alcohol consumption are independent risk factors for alcoholic cirrhosis', *Liver Int*.

Burza, M.A., Pirazzi, C., Maglio, C., Sjoholm, K., Mancina, R.M., Svensson, P.A., Jacobson, P., Adiels, M., Baroni, M.G., Boren, J., Ginanni Corradini, S., Montalcini, T., Sjostrom, L., Carlsson, L.M. and Romeo, S. (2012) 'PNPLA3 I148M (rs738409) genetic variant is associated with hepatocellular carcinoma in obese individuals', *Dig Liver Dis*, 44(12), pp. 1037-41.

Caldwell, S.H., Oelsner, D.H., Iezzoni, J.C., Hespenheide, E.E., Battle, E.H. and Driscoll, C.J. (1999) 'Cryptogenic cirrhosis: clinical characterization and risk factors for underlying disease', *Hepatology*, 29(3), pp. 664-9.

Calle, E.E., Rodriguez, C., Walker-Thurmond, K. and Thun, M.J. (2003) 'Overweight, obesity, and mortality from cancer in a prospectively studied cohort of U.S. adults', *N Engl J Med*, 348(17), pp. 1625-38.

Cameron, A.J., Shaw, J.E. and Zimmet, P.Z. (2004) 'The metabolic syndrome: prevalence in worldwide populations', *Endocrinol Metab Clin North Am*, 33(2), pp. 351-75, table of contents.

Carim-Todd, L., Escarceller, M., Estivill, X. and Sumoy, L. (2000) 'Cloning of the novel gene TM6SF1 reveals conservation of clusters of paralogous genes between human chromosomes 15q24-->q26 and 19p13.3-->p12', *Cytogenet Cell Genet*, 90(3-4), pp. 255-60.

Carulli, L., Canedi, I., Rondinella, S., Lombardini, S., Ganazzi, D., Fargion, S., De Palma, M., Lonardo, A., Ricchi, M., Bertolotti, M., Carulli, N. and Loria, P. (2009) 'Genetic polymorphisms in non-alcoholic fatty liver disease: interleukin-6-174G/C polymorphism is associated with non-alcoholic steatohepatitis', *Dig Liver Dis*, 41(11), pp. 823-8.

Casper, M., Krawczyk, M., Behrmann, I., Glanemann, M. and Lammert, F. (2016) 'Variant PNPLA3 increases the HCC risk: prospective study in patients treated at the Saarland University Medical Center', *Z Gastroenterol*, 54(6), pp. 585-6.

Catlin, R. (1976) 'Liver dysfunction after intestinal bypass', *JAMA*, 236(15), pp. 1693-4.

Chalasani, N., Guo, X., Loomba, R., Goodarzi, M.O., Haritunians, T., Kwon, S., Cui, J., Taylor, K.D., Wilson, L., Cummings, O.W., Chen, Y.D. and Rotter, J.I. (2010) 'Genome-wide association study identifies variants associated with histologic features of nonalcoholic Fatty liver disease', *Gastroenterology*, 139(5), pp. 1567-76, 1576 e1-6.

Chambers, J.C., Zhang, W., Sehmi, J., Li, X., Wass, M.N., Van der Harst, P., Holm, H., Sanna, S., Kavousi, M., Baumeister, S.E., Coin, L.J., Deng, G., Gieger, C., Heard-Costa, N.L., Hottenga, J.J., Kuhnel, B., Kumar, V., Lagou, V., Liang, L., Luan, J., Vidal, P.M., Mateo Leach, I., O'Reilly, P.F., Peden, J.F., Rahmioglu, N., Soininen, P., Speliotes, E.K., Yuan, X., Thorleifsson, G., Alizadeh, B.Z., Atwood, L.D., Borecki, I.B., Brown, M.J., Charoen, P., Cucca, F., Das, D., de Geus, E.J., Dixon, A.L., Doring, A., Ehret, G., Eyjolfsson, G.I., Farrall, M., Forouhi, N.G., Friedrich, N., Goessling, W., Gudbjartsson, D.F., Harris, T.B., Hartikainen, A.L., Heath, S., Hirschfield, G.M., Hofman, A., Homuth, G., Hypponen, E., Janssen, H.L., Johnson, T., Kangas, A.J., Kema, I.P., Kuhn, J.P., Lai, S., Lathrop, M., Lerch, M.M., Li, Y., Liang, T.J., Lin, J.P., Loos, R.J., Martin, N.G., Moffatt, M.F., Montgomery, G.W., Munroe, P.B., Musunuru, K., Nakamura, Y., O'Donnell, C.J., Olafsson, I., Penninx, B.W., Pouta, A., Prins, B.P., Prokopenko, I., Puls, R., Ruokonen, A., Savolainen, M.J., Schlessinger, D., Schouten, J.N., Seedorf, U., Sen-Chowdhry, S., Siminovitch, K.A., Smit, J.H., Spector, T.D., Tan, W., Teslovich, T.M., Tukiainen, T., Uitterlinden, A.G., Van der Klauw, M.M., Vasan, R.S., Wallace, C., Wallaschofski, H., Wichmann, H.E., Willemsen, G., Wurtz, P., Xu, C., Yerges-Armstrong, L.M., et al. (2011) 'Genome-wide association study identifies loci influencing concentrations of liver enzymes in plasma', *Nat Genet*, 43(11), pp. 1131-8.

Chan, W.K., Ida, N.H., Cheah, P.L. and Goh, K.L. (2014) 'Progression of liver disease in non-alcoholic fatty liver disease: a prospective clinicopathological follow-up study', *J Dig Dis*, 15(10), pp. 545-52.

Chen, C.H., Huang, M.H., Yang, J.C., Nien, C.K., Yang, C.C., Yeh, Y.H. and Yueh, S.K. (2007a) 'Prevalence and etiology of elevated serum alanine aminotransferase level in an adult population in Taiwan', *J Gastroenterol Hepatol*, 22(9), pp. 1482-9.

Chen, S., Li, Y., Li, S. and Yu, C. (2008) 'A Val227Ala substitution in the peroxisome proliferator activated receptor alpha (PPAR alpha) gene associated with non-alcoholic fatty liver disease and decreased waist circumference and waist-to-hip ratio', *J Gastroenterol Hepatol*, 23(9), pp. 1415-8.

Chen, Y., Yang, Y., Miller, M.L., Shen, D., Shertzer, H.G., Stringer, K.F., Wang, B., Schneider, S.N., Nebert, D.W. and Dalton, T.P. (2007b) 'Hepatocyte-specific Gclc deletion leads to rapid onset of steatosis with mitochondrial injury and liver failure', *Hepatology*, 45(5), pp. 1118-28.

Cohen, J.C., Horton, J.D. and Hobbs, H.H. (2011) 'Human fatty liver disease: old questions and new insights', *Science*, 332(6037), pp. 1519-23.

Cortez-Pinto, H., Chatham, J., Chacko, V.P., Arnold, C., Rashid, A. and Diehl, A.M. (1999) 'Alterations in liver ATP homeostasis in human nonalcoholic steatohepatitis: a pilot study', *JAMA*, 282(17), pp. 1659-64.

Crespo, J., Cayon, A., Fernandez-Gil, P., Hernandez-Guerra, M., Mayorga, M., Dominguez-Diez, A., Fernandez-Escalante, J.C. and Pons-Romero, F. (2001) 'Gene expression of tumor necrosis factor alpha and TNF-receptors, p55 and p75, in nonalcoholic steatohepatitis patients', *Hepatology*, 34(6), pp. 1158-63.

D'Amico, G., Garcia-Tsao, G. and Pagliaro, L. (2006) 'Natural history and prognostic indicators of survival in cirrhosis: a systematic review of 118 studies', *J Hepatol*, 44(1), pp. 217-31.

Dalamaga, M., Diakopoulos, K.N. and Mantzoros, C.S. (2012) 'The role of adiponectin in cancer: a review of current evidence', *Endocr Rev*, 33(4), pp. 547-94.

Daly, A.K., Steen, V.M., Fairbrother, K.S. and Idle, J.R. (1996) 'CYP2D6 multiallelism', *Methods Enzymol*, 272, pp. 199-210.

Davila, J.A., Morgan, R.O., Shaib, Y., McGlynn, K.A. and El-Serag, H.B. (2005) 'Diabetes increases the risk of hepatocellular carcinoma in the United States: a population based case control study', *Gut*, 54(4), pp. 533-9.

Day, C.P. (2006) 'Genes or environment to determine alcoholic liver disease and non-alcoholic fatty liver disease', *Liver Int*, 26(9), pp. 1021-8.

Day, C.P. (2010) 'Genetic and environmental susceptibility to non-alcoholic fatty liver disease', *Dig Dis*, 28(1), pp. 255-60.

Day, C.P. and James, O.F. (1998) 'Steatohepatitis: a tale of two "hits"?', *Gastroenterology*, 114(4), pp. 842-5.

Day, C.P., Leathart, J.B., McTernan, P. G. (Philip G.), Mathew, Christopher and Daly, A.K. (2006) 'Genetic evidence for a role for gut flora in the pathogenesis of NASH in humans.', *HEPATOLOGY*, 44 (4 Suppl. 1), pp. 261A-261A.

de Almeida, I.T., Cortez-Pinto, H., Fidalgo, G., Rodrigues, D. and Camilo, M.E. (2002) 'Plasma total and free fatty acids composition in human non-alcoholic steatohepatitis', *Clin Nutr*, 21(3), pp. 219-23.

de Marco, R., Locatelli, F., Zoppini, G., Verlato, G., Bonora, E. and Muggeo, M. (1999) 'Cause-specific mortality in type 2 diabetes. The Verona Diabetes Study', *Diabetes Care*, 22(5), pp. 756-61.

Demirag, M.D., Onen, H.I., Karaoguz, M.Y., Dogan, I., Karakan, T., Ekmekci, A. and Guz, G. (2007) 'Apolipoprotein E gene polymorphism in nonalcoholic fatty liver disease', *Dig Dis Sci*, 52(12), pp. 3399-403.

Dentin, R., Benhamed, F., Hainault, I., Fauveau, V., Foufelle, F., Dyck, J.R., Girard, J. and Postic, C. (2006) 'Liver-specific inhibition of ChREBP improves hepatic steatosis and insulin resistance in ob/ob mice', *Diabetes*, 55(8), pp. 2159-70.

Dietrich, P. and Hellerbrand, C. (2014) 'Non-alcoholic fatty liver disease, obesity and the metabolic syndrome', *Best Pract Res Clin Gastroenterol*, 28(4), pp. 637-53.

Ding, X., Saxena, N.K., Lin, S., Xu, A., Srinivasan, S. and Anania, F.A. (2005) 'The roles of leptin and adiponectin: a novel paradigm in adipocytokine regulation of liver fibrosis and stellate cell biology', *Am J Pathol*, 166(6), pp. 1655-69.

DiStefano, J.K., Kingsley, C., Craig Wood, G., Chu, X., Argyropoulos, G., Still, C.D., Done, S.C., Legendre, C., Tembe, W. and Gerhard, G.S. (2015) 'Genome-wide analysis of hepatic lipid content in extreme obesity', *Acta Diabetol*, 52(2), pp. 373-82.

Doege, H., Grimm, D., Falcon, A., Tsang, B., Storm, T.A., Xu, H., Ortegon, A.M., Kazantzis, M., Kay, M.A. and Stahl, A. (2008) 'Silencing of hepatic fatty acid transporter protein 5 in vivo reverses diet-induced non-alcoholic fatty liver disease and improves hyperglycemia', *J Biol Chem*, 283(32), pp. 22186-92.

Donaldson, P., Daly, A., Ermini Luca and Debra, B. (2015) 'Genetics of Complex Disease', *Garland Science*.

Dong, H., Wang, J., Li, C., Hirose, A., Nozaki, Y., Takahashi, M., Ono, M., Akisawa, N., Iwasaki, S., Saibara, T. and Onishi, S. (2007) 'The phosphatidylethanolamine N-methyltransferase gene V175M single nucleotide polymorphism confers the susceptibility to NASH in Japanese population', *J Hepatol*, 46(5), pp. 915-20.

Dongiovanni, P., Petta, S., Maglio, C., Fracanzani, A.L., Pipitone, R., Mozzi, E., Motta, B.M., Kaminska, D., Rametta, R., Grimaudo, S., Pelusi, S., Montalcini, T.,

Alisi, A., Maggioni, M., Karja, V., Boren, J., Kakela, P., Di Marco, V., Xing, C., Nobili, V., Dallapiccola, B., Craxi, A., Pihlajamaki, J., Fargion, S., Sjostrom, L., Carlsson, L.M., Romeo, S. and Valenti, L. (2015) 'Transmembrane 6 superfamily member 2 gene variant disentangles nonalcoholic steatohepatitis from cardiovascular disease', *Hepatology*, 61(2), pp. 506-14.

Dongiovanni, P., Rametta, R., Fracanzani, A.L., Benedan, L., Borroni, V., Maggioni, P., Maggioni, M., Fargion, S. and Valenti, L. (2010a) 'Lack of association between peroxisome proliferator-activated receptors alpha and gamma2 polymorphisms and progressive liver damage in patients with non-alcoholic fatty liver disease: a case control study', *BMC Gastroenterol*, 10, p. 102.

Dongiovanni, P., Valenti, L., Rametta, R., Daly, A.K., Nobili, V., Mozzi, E., Leathart, J.B., Pietrobattista, A., Burt, A.D., Maggioni, M., Fracanzani, A.L., Lattuada, E., Zappa, M.A., Roviaro, G., Marchesini, G., Day, C.P. and Fargion, S. (2010b) 'Genetic variants regulating insulin receptor signalling are associated with the severity of liver damage in patients with non-alcoholic fatty liver disease', *Gut*, 59(2), pp. 267-73.

Donnelly, K.L., Smith, C.I., Schwarzenberg, S.J., Jessurun, J., Boldt, M.D. and Parks, E.J. (2005) 'Sources of fatty acids stored in liver and secreted via lipoproteins in patients with nonalcoholic fatty liver disease', *J Clin Invest*, 115(5), pp. 1343-51.

Dubuquoy, C., Robichon, C., Lasnier, F., Langlois, C., Dugail, I., Foufelle, F., Girard, J., Burnol, A.F., Postic, C. and Moldes, M. (2011) 'Distinct regulation of adiponutrin/PNPLA3 gene expression by the transcription factors ChREBP and SREBP1c in mouse and human hepatocytes', *Journal of hepatology*, 55(1), pp. 145-53.

Dudbridge, F. (2008) 'Likelihood-based association analysis for nuclear families and unrelated subjects with missing genotype data', *Hum Hered*, 66(2), pp. 87-98.

Dudbridge, F., Holmans, P.A. and Wilson, S.G. (2011) 'A flexible model for association analysis in sibships with missing genotype data', *Ann Hum Genet*, 75(3), pp. 428-38.

Dyson, J., Jaques, B., Chattopadyhay, D., Lochan, R., Graham, J., Das, D., Aslam, T., Patanwala, I., Gaggar, S., Cole, M., Sumpter, K., Stewart, S., Rose, J., Hudson, M., Manas, D. and Reeves, H.L. (2014) 'Hepatocellular cancer: the impact of obesity, type 2 diabetes and a multidisciplinary team', *J Hepatol*, 60(1), pp. 110-7.

'EASL-EORTC clinical practice guidelines: management of hepatocellular carcinoma', (2012) *Journal of hepatology*, 56(4), pp. 908-43.

Eckel, R.H., Grundy, S.M. and Zimmet, P.Z. (2005) 'The metabolic syndrome', *Lancet*, 365(9468), pp. 1415-28.

Edmondson, H.A. and Steiner, P.E. (1954) 'Primary carcinoma of the liver: a study of 100 cases among 48,900 necropsies', *Cancer*, 7(3), pp. 462-503.

Ekstedt, M., Franzen, L.E., Mathiesen, U.L., Thorelius, L., Holmqvist, M., Bodemar, G. and Kechagias, S. (2006) 'Long-term follow-up of patients with NAFLD and elevated liver enzymes', *Hepatology*, 44(4), pp. 865-73.

El-Serag, H.B., Hampel, H. and Javadi, F. (2006) 'The association between diabetes and hepatocellular carcinoma: a systematic review of epidemiologic evidence', *Clin Gastroenterol Hepatol*, 4(3), pp. 369-80.

El-Serag, H.B. and Rudolph, K.L. (2007) 'Hepatocellular carcinoma: epidemiology and molecular carcinogenesis', *Gastroenterology*, 132(7), pp. 2557-76.

El-Serag, H.B., Tran, T. and Everhart, J.E. (2004) 'Diabetes increases the risk of chronic liver disease and hepatocellular carcinoma', *Gastroenterology*, 126(2), pp. 460-8.

Ellervik, C., Birgens, H., Tybjaerg-Hansen, A. and Nordestgaard, B.G. (2007) 'Hemochromatosis genotypes and risk of 31 disease endpoints: meta-analyses including 66,000 cases and 226,000 controls', *Hepatology*, 46(4), pp. 1071-80.

Ertle, J., Dechene, A., Sowa, J.P., Penndorf, V., Herzer, K., Kaiser, G., Schlaak, J.F., Gerken, G., Syn, W.K. and Canbay, A. (2011) 'Non-alcoholic fatty liver disease progresses to hepatocellular carcinoma in the absence of apparent cirrhosis', *Int J Cancer*, 128(10), pp. 2436-43.

European Association For The Study Of The Liver and Cancer, E.O.F.R.A.T.O. (2012) 'EASL-EORTC clinical practice guidelines: management of hepatocellular carcinoma', *J Hepatol*, 56(4), pp. 908-43.

Expert Panel on Detection, E. and Treatment of High Blood Cholesterol in, A. (2001) 'Executive Summary of The Third Report of The National Cholesterol Education Program (NCEP) Expert Panel on Detection, Evaluation, And Treatment of High Blood Cholesterol In Adults (Adult Treatment Panel III)', *JAMA*, 285(19), pp. 2486-97.

Falleti, E., Fabris, C., Cmet, S., Cussigh, A., Bitetto, D., Fontanini, E., Fornasiere, E., Bignulin, S., Fumolo, E., Bignulin, E., Pirisi, M. and Toniutto, P. (2011) 'PNPLA3 rs738409C/G polymorphism in cirrhosis: relationship with the aetiology of liver disease and hepatocellular carcinoma occurrence', *Liver Int*, 31(8), pp. 1137-43.

Fan, J.G., Li, F., Cai, X.B., Peng, Y.D., Ao, Q.H. and Gao, Y. (2007) 'Effects of nonalcoholic fatty liver disease on the development of metabolic disorders', *J Gastroenterol Hepatol*, 22(7), pp. 1086-91.

Farrell, G.C. and Larter, C.Z. (2006) 'Nonalcoholic fatty liver disease: from steatosis to cirrhosis', *Hepatology*, 43(2 Suppl 1), pp. S99-S112.

Fassio, E., Alvarez, E., Dominguez, N., Landeira, G. and Longo, C. (2004) 'Natural history of nonalcoholic steatohepatitis: a longitudinal study of repeat liver biopsies', *Hepatology*, 40(4), pp. 820-6.

Fawcett, K.A., Grimsey, N., Loos, R.J., Wheeler, E., Daly, A., Soos, M., Semple, R., Syddall, H., Cooper, C., Siniossoglou, S., O'Rahilly, S., Wareham, N.J. and Barroso, I. (2008) 'Evaluating the role of LPIN1 variation in insulin resistance, body weight, and human lipodystrophy in U.K. Populations', *Diabetes*, 57(9), pp. 2527-33.

Feitosa, M.F., Wojczynski, M.K., North, K.E., Zhang, Q., Province, M.A., Carr, J.J. and Borecki, I.B. (2013) 'The ERLIN1-CHUK-CWF19L1 gene cluster influences liver fat deposition and hepatic inflammation in the NHLBI Family Heart Study', *Atherosclerosis*, 228(1), pp. 175-80.

Feldstein, A.E., Werneburg, N.W., Canbay, A., Guicciardi, M.E., Bronk, S.F., Rydzewski, R., Burgart, L.J. and Gores, G.J. (2004) 'Free fatty acids promote hepatic lipotoxicity by stimulating TNF-alpha expression via a lysosomal pathway', *Hepatology*, 40(1), pp. 185-94.

Ferre, P. and Foufelle, F. (2007) 'SREBP-1c transcription factor and lipid homeostasis: clinical perspective', *Horm Res*, 68(2), pp. 72-82.

Fraser, A., Harris, R., Sattar, N., Ebrahim, S., Davey Smith, G. and Lawlor, D.A. (2009) 'Alanine aminotransferase, gamma-glutamyltransferase, and incident diabetes: the British Women's Heart and Health Study and meta-analysis', *Diabetes Care*, 32(4), pp. 741-50.

Friedman, S.L., Rockey, D.C., McGuire, R.F., Maher, J.J., Boyles, J.K. and Yamasaki, G. (1992) 'Isolated hepatic lipocytes and Kupffer cells from normal human liver: morphological and functional characteristics in primary culture', *Hepatology*, 15(2), pp. 234-43.

Frith, J., Day, C.P., Henderson, E., Burt, A.D. and Newton, J.L. (2009) 'Non-alcoholic fatty liver disease in older people', *Gerontology*, 55(6), pp. 607-13.

Gabele, E., Brenner, D.A. and Rippe, R.A. (2003) 'Liver fibrosis: signals leading to the amplification of the fibrogenic hepatic stellate cell', *Front Biosci*, 8, pp. d69-77.

Gawrieh, S., Marion, M.C., Komorowski, R., Wallace, J., Charlton, M., Kissebah, A., Langefeld, C.D. and Olivier, M. (2012) 'Genetic variation in the peroxisome proliferator activated receptor-gamma gene is associated with histologically advanced NAFLD', *Dig Dis Sci*, 57(4), pp. 952-7.

George, D.K., Goldwurm, S., MacDonald, G.A., Cowley, L.L., Walker, N.I., Ward, P.J., Jazwinska, E.C. and Powell, L.W. (1998) 'Increased hepatic iron concentration in nonalcoholic steatohepatitis is associated with increased fibrosis', *Gastroenterology*, 114(2), pp. 311-8.

Goffredo, M., Caprio, S., Feldstein, A.E., D'Adamo, E., Shaw, M.M., Pierpont, B., Savoye, M., Zhao, H., Bale, A.E. and Santoro, N. (2016) 'Role of TM6SF2 rs58542926 in the pathogenesis of nonalcoholic pediatric fatty liver disease: A multiethnic study', *Hepatology*, 63(1), pp. 117-25.

Goh, G.B. and McCullough, A.J. (2016) 'Natural History of Nonalcoholic Fatty Liver Disease', *Dig Dis Sci*, 61(5), pp. 1226-33.

Gorden, A., Yang, R., Yerges-Armstrong, L.M., Ryan, K.A., Speliotes, E., Borecki, I.B., Harris, T.B., Chu, X., Wood, G.C., Still, C.D., Shuldiner, A.R., Gerhard, G.S. and Consortium, G. (2013) 'Genetic variation at NCAN locus is associated with inflammation and fibrosis in non-alcoholic fatty liver disease in morbid obesity', *Hum Hered*, 75(1), pp. 34-43.

Grandone, A., Cozzolino, D., Marzuillo, P., Cirillo, G., Di Sessa, A., Ruggiero, L., Di Palma, M.R., Perrone, L. and Miraglia Del Giudice, E. (2016) 'TM6SF2 Glu167Lys polymorphism is associated with low levels of LDL-cholesterol and increased liver injury in obese children', *Pediatr Obes*, 11(2), pp. 115-9.

Grarup, N., Urhammer, S.A., Ek, J., Albrechtsen, A., Glumer, C., Borch-Johnsen, K., Jorgensen, T., Hansen, T. and Pedersen, O. (2006) 'Studies of the relationship between the ENPP1 K121Q polymorphism and type 2 diabetes, insulin resistance and obesity in 7,333 Danish white subjects', *Diabetologia*, 49(9), pp. 2097-104.

Guo, J., Loke, J., Zheng, F., Hong, F., Yea, S., Fukata, M., Tarocchi, M., Abar, O.T., Huang, H., Sninsky, J.J. and Friedman, S.L. (2009) 'Functional linkage of cirrhosis-predictive single nucleotide polymorphisms of Toll-like receptor 4 to hepatic stellate cell responses', *Hepatology*, 49(3), pp. 960-8.

Guy, C.D., Suzuki, A., Burchette, J.L., Brunt, E.M., Abdelmalek, M.F., Cardona, D., McCall, S.J., Unalp, A., Belt, P., Ferrell, L.D., Diehl, A.M. and Nonalcoholic Steatohepatitis Clinical Research, N. (2012) 'Costaining for keratins 8/18 plus ubiquitin improves detection of hepatocyte injury in nonalcoholic fatty liver disease', *Hum Pathol*, 43(6), pp. 790-800.

Guyot, E., Sutton, A., Rufat, P., Laguillier, C., Mansouri, A., Moreau, R., Ganne-Carrie, N., Beaugrand, M., Charnaux, N., Trinchet, J.C. and Nahon, P. (2013) 'PNPLA3 rs738409, hepatocellular carcinoma occurrence and risk model prediction in patients with cirrhosis', *J Hepatol*, 58(2), pp. 312-8.

Guzman, G., Brunt, E.M., Petrovic, L.M., Chejfec, G., Layden, T.J. and Cotler, S.J. (2008) 'Does nonalcoholic fatty liver disease predispose patients to hepatocellular carcinoma in the absence of cirrhosis?', *Arch Pathol Lab Med*, 132(11), pp. 1761-6.

Ha, Y., Seo, N., Shim, J.H., Kim, S.Y., Park, J.A., Han, S., Kim, K.W., Yu, E., Kim, K.M., Lim, Y.S., Lee, H.C., Chung, Y.H. and Lee, Y.S. (2015) 'Intimate association of visceral obesity with non-alcoholic fatty liver disease in healthy Asians: A case-control study', *J Gastroenterol Hepatol*, 30(11), pp. 1666-72.

Halliwell, B., Murcia, M.A., Chirico, S. and Aruoma, O.I. (1995) 'Free radicals and antioxidants in food and in vivo: what they do and how they work', *Crit Rev Food Sci Nutr*, 35(1-2), pp. 7-20.

Hardy, T., Oakley, F., Anstee, Q.M. and Day, C.P. (2016) 'Nonalcoholic Fatty Liver Disease: Pathogenesis and Disease Spectrum', *Annu Rev Pathol*, 11, pp. 451-96.

Harmon, R.C., Tiniakos, D.G. and Argo, C.K. (2011) 'Inflammation in nonalcoholic steatohepatitis', *Expert Rev Gastroenterol Hepatol*, 5(2), pp. 189-200.

Hashimoto, E., Yatsuji, S., Tobari, M., Taniai, M., Torii, N., Tokushige, K. and Shiratori, K. (2009) 'Hepatocellular carcinoma in patients with nonalcoholic steatohepatitis', *J Gastroenterol*, 44 Suppl 19, pp. 89-95.

Hassan, M.M., Kaseb, A., Etzel, C.J., El-Serag, H., Spitz, M.R., Chang, P., Hale, K.S., Liu, M., Rashid, A., Shama, M., Abbruzzese, J.L., Loyer, E.M., Kaur, H., Hassabo, H.M., Vauthey, J.N., Wray, C.J., Hassan, B.S., Patt, Y.Z., Hawk, E., Soliman, K.M. and Li, D. (2013) 'Genetic variation in the PNPLA3 gene and hepatocellular carcinoma in USA: risk and prognosis prediction', *Mol Carcinog*, 52 Suppl 1, pp. E139-47.

He, S., McPhaul, C., Li, J.Z., Garuti, R., Kinch, L., Grishin, N.V., Cohen, J.C. and Hobbs, H.H. (2010) 'A sequence variation (I148M) in PNPLA3 associated with nonalcoholic fatty liver disease disrupts triglyceride hydrolysis', *J Biol Chem*, 285(9), pp. 6706-15.

Heidebrecht, F., Heidebrecht, A., Schulz, I., Behrens, S.E. and Bader, A. (2009) 'Improved semiquantitative Western blot technique with increased quantification range', *J Immunol Methods*, 345(1-2), pp. 40-8.

Hernaez, R., Yeung, E., Clark, J.M., Kowdley, K.V., Brancati, F.L. and Kao, W.H. (2011) 'Hemochromatosis gene and nonalcoholic fatty liver disease: a systematic review and meta-analysis', *J Hepatol*, 55(5), pp. 1079-85.

Hirsch, D., Stahl, A. and Lodish, H.F. (1998) 'A family of fatty acid transporters conserved from mycobacterium to man', *Proc Natl Acad Sci U S A*, 95(15), pp. 8625-9.

Hirschhorn, J.N. (2009) 'Genomewide association studies--illuminating biologic pathways', *N Engl J Med*, 360(17), pp. 1699-701.

Hirschhorn, J.N. and Daly, M.J. (2005) 'Genome-wide association studies for common diseases and complex traits', *Nat Rev Genet*, 6(2), pp. 95-108.

Hirschhorn, J.N. and Gajdos, Z.K. (2011) 'Genome-wide association studies: results from the first few years and potential implications for clinical medicine', *Annu Rev Med*, 62, pp. 11-24.

Hjelkrem, M., Stauch, C., Shaw, J. and Harrison, S.A. (2011) 'Validation of the non-alcoholic fatty liver disease activity score', *Aliment Pharmacol Ther*, 34(2), pp. 214-8.

Holmen, O.L., Zhang, H., Fan, Y., Hovelson, D.H., Schmidt, E.M., Zhou, W., Guo, Y., Zhang, J., Langhammer, A., Lochen, M.L., Ganesh, S.K., Vatten, L., Skorpen, F.,

Dalen, H., Zhang, J., Pennathur, S., Chen, J., Platou, C., Mathiesen, E.B., Wilsgaard, T., Njolstad, I., Boehnke, M., Chen, Y.E., Abecasis, G.R., Hveem, K. and Willer, C.J. (2014) 'Systematic evaluation of coding variation identifies a candidate causal variant in TM6SF2 influencing total cholesterol and myocardial infarction risk', *Nat Genet*, 46(4), pp. 345-51.

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

Huang, H., Shiffman, M.L., Friedman, S., Venkatesh, R., Bzowej, N., Abar, O.T., Rowland, C.M., Catanese, J.J., Leong, D.U., Sninsky, J.J., Layden, T.J., Wright, T.L., White, T. and Cheung, R.C. (2007) 'A 7 gene signature identifies the risk of developing cirrhosis in patients with chronic hepatitis C', *Hepatology*, 46(2), pp. 297-306.

Huang, Y., Cohen, J.C. and Hobbs, H.H. (2011) 'Expression and characterization of a PNPLA3 protein isoform (I148M) associated with nonalcoholic fatty liver disease', *J Biol Chem*, 286(43), pp. 37085-93.

Huang, Y., He, S., Li, J.Z., Seo, Y.K., Osborne, T.F., Cohen, J.C. and Hobbs, H.H. (2010) 'A feed-forward loop amplifies nutritional regulation of PNPLA3', *Proc Natl Acad Sci U S A*, 107(17), pp. 7892-7.

Hui, A.Y., Wong, V.W., Chan, H.L., Liew, C.T., Chan, J.L., Chan, F.K. and Sung, J.J. (2005) 'Histological progression of non-alcoholic fatty liver disease in Chinese patients', *Aliment Pharmacol Ther*, 21(4), pp. 407-13.

Hui, J.M., Kench, J.G., Chitturi, S., Sud, A., Farrell, G.C., Byth, K., Hall, P., Khan, M. and George, J. (2003) 'Long-term outcomes of cirrhosis in nonalcoholic steatohepatitis compared with hepatitis C', *Hepatology*, 38(2), pp. 420-7.

Hussain, S.P., Hofseth, L.J. and Harris, C.C. (2003) 'Radical causes of cancer', *Nat Rev Cancer*, 3(4), pp. 276-85.

Hwang, S.T., Cho, Y.K., Yun, J.W., Park, J.H., Kim, H.J., Park, D.I., Sohn, C.I., Jeon, W.K., Kim, B.I., Rhee, E.J., Oh, K.W., Lee, W.Y. and Jin, W. (2010) 'Impact of non-alcoholic fatty liver disease on microalbuminuria in patients with prediabetes and diabetes', *Intern Med J*, 40(6), pp. 437-42.

Iizuka, K., Bruick, R.K., Liang, G., Horton, J.D. and Uyeda, K. (2004) 'Deficiency of carbohydrate response element-binding protein (ChREBP) reduces lipogenesis as well as glycolysis', *Proc Natl Acad Sci U S A*, 101(19), pp. 7281-6.

Iizuka, K. and Horikawa, Y. (2008) 'ChREBP: a glucose-activated transcription factor involved in the development of metabolic syndrome', *Endocr J*, 55(4), pp. 617-24.

International HapMap, C. (2003) 'The International HapMap Project', *Nature*, 426(6968), pp. 789-96.

International HapMap, C. (2005) 'A haplotype map of the human genome', *Nature*, 437(7063), pp. 1299-320.

International HapMap Consortium (2003) 'The International HapMap Project', *Nature*, 426(6968), pp. 789-96.

Jimba, S., Nakagami, T., Takahashi, M., Wakamatsu, T., Hirota, Y., Iwamoto, Y. and Wasada, T. (2005) 'Prevalence of non-alcoholic fatty liver disease and its association with impaired glucose metabolism in Japanese adults', *Diabet Med*, 22(9), pp. 1141-5.

Johansson, L.E., Johansson, L.M., Danielsson, P., Norgren, S., Johansson, S., Marcus, C. and Ridderstrale, M. (2009) 'Genetic variance in the adiponutrin gene family and childhood obesity', *PLoS One*, 4(4), p. e5327.

Jou, J., Choi, S.S. and Diehl, A.M. (2008) 'Mechanisms of disease progression in nonalcoholic fatty liver disease', *Semin Liver Dis*, 28(4), pp. 370-9.

Kahali, B., Liu, Y.L., Daly, A.K., Day, C.P., Anstee, Q.M. and Speliotes, E.K. (2015) 'TM6SF2: catch-22 in the fight against nonalcoholic fatty liver disease and cardiovascular disease?', *Gastroenterology*, 148(4), pp. 679-84.

Karagozian, R., Derdak, Z. and Baffy, G. (2014) 'Obesity-associated mechanisms of hepatocarcinogenesis', *Metabolism*, 63(5), pp. 607-17.

Kathiresan, S., Melander, O., Guiducci, C., Surti, A., Burtt, N.P., Rieder, M.J., Cooper, G.M., Roos, C., Voight, B.F., Havulinna, A.S., Wahlstrand, B., Hedner, T., Corella, D., Tai, E.S., Ordovas, J.M., Berglund, G., Vartiainen, E., Jousilahti, P., Hedblad, B., Taskinen, M.R., Newton-Cheh, C., Salomaa, V., Peltonen, L., Groop, L., Altshuler, D.M. and Orho-Melander, M. (2008) 'Six new loci associated with blood low-density lipoprotein cholesterol, high-density lipoprotein cholesterol or triglycerides in humans', *Nat Genet*, 40(2), pp. 189-97.

Kawaguchi, T., Sumida, Y., Umemura, A., Matsuo, K., Takahashi, M., Takamura, T., Yasui, K., Saibara, T., Hashimoto, E., Kawanaka, M., Watanabe, S., Kawata, S., Imai, Y., Kokubo, M., Shima, T., Park, H., Tanaka, H., Tajima, K., Yamada, R., Matsuda, F. and Japan Study Group of Nonalcoholic Fatty Liver, D. (2012) 'Genetic polymorphisms of the human PNPLA3 gene are strongly associated with severity of non-alcoholic fatty liver disease in Japanese', *PLoS One*, 7(6), p. e38322.

Kern, P.A., Saghizadeh, M., Ong, J.M., Bosch, R.J., Deem, R. and Simsolo, R.B. (1995) 'The expression of tumor necrosis factor in human adipose tissue. Regulation by obesity, weight loss, and relationship to lipoprotein lipase', *J Clin Invest*, 95(5), pp. 2111-9.

Kern, W.H., Heger, A.H., Payne, J.H. and DeWind, L.T. (1973) 'Fatty metamorphosis of the liver in morbid obesity', *Arch Pathol*, 96(5), pp. 342-6.

Kienesberger, P.C., Lee, D., Pulinilkunnil, T., Brenner, D.S., Cai, L., Magnes, C., Koefeler, H.C., Streith, I.E., Rechberger, G.N., Haemmerle, G., Flier, J.S., Zechner, R., Kim, Y.B. and Kershaw, E.E. (2009) 'Adipose triglyceride lipase deficiency causes tissue-specific changes in insulin signaling', *J Biol Chem*, 284(44), pp. 30218-29.

Kim, H., Haluzik, M., Asghar, Z., Yau, D., Joseph, J.W., Fernandez, A.M., Reitman, M.L., Yakar, S., Stannard, B., Heron-Milhavet, L., Wheeler, M.B. and LeRoith, D. (2003) 'Peroxisome proliferator-activated receptor-alpha agonist treatment in a transgenic model of type 2 diabetes reverses the lipotoxic state and improves glucose homeostasis', *Diabetes*, 52(7), pp. 1770-8.

Kim, K.W., Bae, S.K., Lee, O.H., Bae, M.H., Lee, M.J. and Park, B.C. (1998) 'Insulin-like growth factor II induced by hypoxia may contribute to angiogenesis of human hepatocellular carcinoma', *Cancer Res*, 58(2), pp. 348-51.

Kitamoto, T., Kitamoto, A., Yoneda, M., Hyogo, H., Ochi, H., Nakamura, T., Teranishi, H., Mizusawa, S., Ueno, T., Chayama, K., Nakajima, A., Nakao, K., Sekine, A. and Hotta, K. (2013) 'Genome-wide scan revealed that polymorphisms in the PNPLA3, SAMM50, and PARVB genes are associated with development and progression of nonalcoholic fatty liver disease in Japan', *Hum Genet*, 132(7), pp. 783-92.

Kleiner, D.E., Brunt, E.M., Van Natta, M., Behling, C., Contos, M.J., Cummings, O.W., Ferrell, L.D., Liu, Y.C., Torbenson, M.S., Unalp-Arida, A., Yeh, M., McCullough, A.J. and Sanyal, A.J. (2005) 'Design and validation of a histological scoring system for nonalcoholic fatty liver disease', *Hepatology*, 41(6), pp. 1313-21.

Koek, G.H., Liedorp, P.R. and Bast, A. (2011) 'The role of oxidative stress in non-alcoholic steatohepatitis', *Clin Chim Acta*, 412(15-16), pp. 1297-305.

Kozlitina, J. (2014) 'Exome-wide association study identifies a TM6SF2 variant that confers susceptibility to nonalcoholic fatty liver disease', *Nat. Genet.*, 46, pp. 352-356.

Kozlitina, J., Boerwinkle, E., Cohen, J.C. and Hobbs, H.H. (2011) 'Dissociation between APOC3 variants, hepatic triglyceride content and insulin resistance', *Hepatology*, 53(2), pp. 467-74.

Kozlitina, J., Smagris, E., Stender, S., Nordestgaard, B.G., Zhou, H.H., Tybjaerg-Hansen, A., Vogt, T.F., Hobbs, H.H. and Cohen, J.C. (2014) 'Exome-wide association study identifies a TM6SF2 variant that confers susceptibility to nonalcoholic fatty liver disease', *Nat Genet*, 46(4), pp. 352-6.

Kral, J.G., Schaffner, F., Pierson, R.N., Jr. and Wang, J. (1993) 'Body fat topography as an independent predictor of fatty liver', *Metabolism*, 42(5), pp. 548-51.

Krawczyk, M., Stokes, C.S., Romeo, S. and Lammert, F. (2015) 'HCC and liver disease risks in homozygous PNPLA3 p.I148M carriers approach monogenic inheritance', *J Hepatol*, 62(4), pp. 980-1.

Kubik, J.F., Gill, R.S., Laffin, M. and Karmali, S. (2013) 'The impact of bariatric surgery on psychological health', *J Obes*, 2013, p. 837989.

Kumari, M., Schoiswohl, G., Chitraju, C., Paar, M., Cornaciu, I., Rangrez, A.Y., Wongsiriroj, N., Nagy, H.M., Ivanova, P.T., Scott, S.A., Knittelfelder, O., Rechberger, G.N., Birner-Gruenberger, R., Eder, S., Brown, H.A., Haemmerle, G., Oberer, M., Lass, A., Kershaw, E.E., Zimmermann, R. and Zechner, R. (2012) 'Adiponutrin functions as a nutritionally regulated lysophosphatidic acid acyltransferase', *Cell metabolism*, 15(5), pp. 691-702.

Lade, A., Noon, L.A. and Friedman, S.L. (2014) 'Contributions of metabolic dysregulation and inflammation to nonalcoholic steatohepatitis, hepatic fibrosis, and cancer', *Curr Opin Oncol*, 26(1), pp. 100-7.

Lake, A.C., Sun, Y., Li, J.L., Kim, J.E., Johnson, J.W., Li, D., Revett, T., Shih, H.H., Liu, W., Paulsen, J.E. and Gimeno, R.E. (2005) 'Expression, regulation, and triglyceride hydrolase activity of Adiponutrin family members', *J Lipid Res*, 46(11), pp. 2477-87.

Larsson, S.C. and Wolk, A. (2007) 'Overweight, obesity and risk of liver cancer: a meta-analysis of cohort studies', *Br J Cancer*, 97(7), pp. 1005-8.

Lassailly, G., Caiazzo, R., Buob, D., Pigeyre, M., Verkindt, H., Labreuche, J., Raverdy, V., Leteurtre, E., Dharancy, S., Louvet, A., Romon, M., Duhamel, A., Pattou, F. and Mathurin, P. (2015) 'Bariatric Surgery Reduces Features of Nonalcoholic Steatohepatitis in Morbidly Obese Patients', *Gastroenterology*, 149(2), pp. 379-88; quiz e15-6.

Latruffe, N., Cherkaoui Malki, M., Nicolas-Frances, V., Clemencet, M.C., Jannin, B. and Berlot, J.P. (2000) 'Regulation of the peroxisomal beta-oxidation-dependent pathway by peroxisome proliferator-activated receptor alpha and kinases', *Biochem Pharmacol*, 60(8), pp. 1027-32.

Le May, C., Pineau, T., Bigot, K., Kohl, C., Girard, J. and Pegorier, J.P. (2000) 'Reduced hepatic fatty acid oxidation in fasting PPARalpha null mice is due to impaired mitochondrial hydroxymethylglutaryl-CoA synthase gene expression', *FEBS Lett*, 475(3), pp. 163-6.

Leclercq, I.A., Farrell, G.C., Schriemer, R. and Robertson, G.R. (2002) 'Leptin is essential for the hepatic fibrogenic response to chronic liver injury', *J Hepatol*, 37(2), pp. 206-13.

Lee, Y.S. and Jeong, W.I. (2012) 'Retinoic acids and hepatic stellate cells in liver disease', *J Gastroenterol Hepatol*, 27 Suppl 2, pp. 75-9.

Lehninger, A.L., Nelson, D.L. and Cox, M.M. (2000) 'Lehninger principles of biochemistry.'.

Leite, N.C., Salles, G.F., Araujo, A.L., Villela-Nogueira, C.A. and Cardoso, C.R. (2009) 'Prevalence and associated factors of non-alcoholic fatty liver disease in patients with type-2 diabetes mellitus', *Liver Int*, 29(1), pp. 113-9.

Leung, C., Rivera, L., Furness, J.B. and Angus, P.W. (2016) 'The role of the gut microbiota in NAFLD', *Nat Rev Gastroenterol Hepatol*, 13(7), pp. 412-25.

Levene, A.P. and Goldin, R.D. (2012) 'The epidemiology, pathogenesis and histopathology of fatty liver disease', *Histopathology*, 61(2), pp. 141-52.

Li, J.Z., Huang, Y., Karaman, R., Ivanova, P.T., Brown, H.A., Roddy, T., Castro-Perez, J., Cohen, J.C. and Hobbs, H.H. (2012) 'Chronic overexpression of PNPLA3I148M in mouse liver causes hepatic steatosis', *J Clin Invest*, 122(11), pp. 4130-44.

Liu, W., Anstee, Q.M., Wang, X., Gawrieh, S., Gamazon, E.R., Athinarayanan, S., Liu, Y.L., Darlay, R., Cordell, H.J., Daly, A.K., Day, C.P. and Chalasani, N. (2016a) 'Transcriptional regulation of PNPLA3 and its impact on susceptibility to nonalcoholic fatty liver Disease (NAFLD) in humans', *Aging (Albany NY)*, 9(1), pp. 26-40.

Liu, Y., Lu, J., Xu, M., Xu, Y., Li, M., Wang, T., Zhang, J., Xu, B., Sun, J., Dai, M., Bi, Y., Wang, W. and Ning, G. (2013) 'Association between history of abortion and nonalcoholic fatty liver disease in middle-aged and elderly Chinese women', *Ann Epidemiol*, 23(3), pp. 119-23.

Liu, Y.L., Day, C.P. and Anstee, Q.M. (2016b) 'Can genetic influence in non-alcoholic fatty liver disease be ignored?', *Clinical Dilemmas in Non-Alcoholic Fatty Liver Disease (eds R. Williams and S. D. Taylor-Robinson)*.

Liu, Y.L., Patman, G.L., Leathart, J.B., Piguet, A.C., Burt, A.D., Dufour, J.F., Day, C.P., Daly, A.K., Reeves, H.L. and Anstee, Q.M. (2014a) 'Carriage of the PNPLA3 rs738409 C >G polymorphism confers an increased risk of non-alcoholic fatty liver disease associated hepatocellular carcinoma', *J Hepatol*, 61(1), pp. 75-81.

Liu, Y.L., Reeves, H.L., Burt, A.D., Tiniakos, D., McPherson, S., Leathart, J.B., Allison, M.E., Alexander, G.J., Piguet, A.C., Anty, R., Donaldson, P., Aithal, G.P., Francque, S., Van Gaal, L., Clement, K., Ratziu, V., Dufour, J.F., Day, C.P., Daly, A.K. and Anstee, Q.M. (2014b) 'TM6SF2 rs58542926 influences hepatic fibrosis progression in patients with non-alcoholic fatty liver disease', *Nat Commun*, 5, p. 4309.

Lonardo, A., Lombardini, S., Scaglioni, F., Carulli, L., Ricchi, M., Ganazzi, D., Adinolfi, L.E., Ruggiero, G., Carulli, N. and Loria, P. (2006) 'Hepatic steatosis and insulin resistance: does etiology make a difference?', *J Hepatol*, 44(1), pp. 190-6.

Loomba, R., Abraham, M., Unalp, A., Wilson, L., Lavine, J., Doo, E., Bass, N.M. and Nonalcoholic Steatohepatitis Clinical Research, N. (2012) 'Association between diabetes, family history of diabetes, and risk of nonalcoholic steatohepatitis and fibrosis', *Hepatology*, 56(3), pp. 943-51.

Loomba, R. and Sanyal, A.J. (2013) 'The global NAFLD epidemic', *Nat Rev Gastroenterol Hepatol*, 10(11), pp. 686-90.

Loomba, R., Schork, N., Chen, C.H., Bettencourt, R., Bhatt, A., Ang, B., Nguyen, P., Hernandez, C., Richards, L., Salotti, J., Lin, S., Seki, E., Nelson, K.E., Sirlin, C.B., Brenner, D. and Genetics of, N.i.T.C. (2015) 'Heritability of Hepatic Fibrosis and Steatosis Based on a Prospective Twin Study', *Gastroenterology*, 149(7), pp. 1784-93.

Lu, S.C., Alvarez, L., Huang, Z.Z., Chen, L., An, W., Corrales, F.J., Avila, M.A., Kanel, G. and Mato, J.M. (2001) 'Methionine adenosyltransferase 1A knockout mice are predisposed to liver injury and exhibit increased expression of genes involved in proliferation', *Proc Natl Acad Sci U S A*, 98(10), pp. 5560-5.

Ludwig, J., Viggiano, T.R., McGill, D.B. and Oh, B.J. (1980) 'Nonalcoholic steatohepatitis: Mayo Clinic experiences with a hitherto unnamed disease', *Mayo Clin Proc*, 55(7), pp. 434-8.

Luo, Z., Zang, M. and Guo, W. (2010) 'AMPK as a metabolic tumor suppressor: control of metabolism and cell growth', *Future Oncol*, 6(3), pp. 457-70.

Ma, L., Robinson, L.N. and Towle, H.C. (2006) 'ChREBP*Mlx is the principal mediator of glucose-induced gene expression in the liver', *J Biol Chem*, 281(39), pp. 28721-30.

Machado, M., Marques-Vidal, P. and Cortez-Pinto, H. (2006) 'Hepatic histology in obese patients undergoing bariatric surgery', *J Hepatol*, 45(4), pp. 600-6.

Mahdessian, H., Taxiarchis, A., Popov, S., Silveira, A., Franco-Cereceda, A., Hamsten, A., Eriksson, P. and van't Hooft, F. (2014) 'TM6SF2 is a regulator of liver fat metabolism influencing triglyceride secretion and hepatic lipid droplet content', *Proc Natl Acad Sci U S A*, 111(24), pp. 8913-8.

Maier, T., Guell, M. and Serrano, L. (2009) 'Correlation of mRNA and protein in complex biological samples', *FEBS Lett*, 583(24), pp. 3966-73.

Makkonen, J., Pietilainen, K.H., Rissanen, A., Kaprio, J. and Yki-Jarvinen, H. (2009) 'Genetic factors contribute to variation in serum alanine aminotransferase activity independent of obesity and alcohol: a study in monozygotic and dizygotic twins', *J Hepatol*, 50(5), pp. 1035-42.

Malhi, H., Bronk, S.F., Werneburg, N.W. and Gores, G.J. (2006) 'Free fatty acids induce JNK-dependent hepatocyte lipoapoptosis', *J Biol Chem*, 281(17), pp. 12093-101.

Mancina, R.M., Dongiovanni, P., Petta, S., Pingitore, P., Meroni, M., Rametta, R., Boren, J., Montalcini, T., Pujia, A., Wiklund, O., Hindy, G., Spagnuolo, R., Motta, B.M., Pipitone, R.M., Craxi, A., Fargion, S., Nobili, V., Kakela, P., Karja, V., Mannisto, V., Pihlajamaki, J., Reilly, D.F., Castro-Perez, J., Kozlitina, J., Valenti, L. and Romeo,

S. (2016) 'The MBOAT7-TMC4 Variant rs641738 Increases Risk of Nonalcoholic Fatty Liver Disease in Individuals of European Descent', *Gastroenterology*, 150(5), pp. 1219-1230 e6.

Mancina, R.M., Matikainen, N., Maglio, C., Soderlund, S., Lundbom, N., Hakkarainen, A., Rametta, R., Mozzi, E., Fargion, S., Valenti, L., Romeo, S., Taskinen, M.R. and Boren, J. (2015) 'Paradoxical dissociation between hepatic fat content and de novo lipogenesis due to PNPLA3 sequence variant', *J Clin Endocrinol Metab*, 100(5), pp. E821-5.

Manolio, T.A. (2013) 'Bringing genome-wide association findings into clinical use', *Nat Rev Genet*, 14(8), pp. 549-58.

Manolio, T.A., Brooks, L.D. and Collins, F.S. (2008) 'A HapMap harvest of insights into the genetics of common disease', *J Clin Invest*, 118(5), pp. 1590-605.

Marchesini, G., Brizi, M., Bianchi, G., Tomassetti, S., Bugianesi, E., Lenzi, M., McCullough, A.J., Natale, S., Forlani, G. and Melchionda, N. (2001) 'Nonalcoholic fatty liver disease: a feature of the metabolic syndrome', *Diabetes*, 50(8), pp. 1844-50.

Marchesini, G., Bugianesi, E., Forlani, G., Cerrelli, F., Lenzi, M., Manini, R., Natale, S., Vanni, E., Villanova, N., Melchionda, N. and Rizzetto, M. (2003) 'Nonalcoholic fatty liver, steatohepatitis, and the metabolic syndrome', *Hepatology*, 37(4), pp. 917-23.

Marra, F. (2007) 'Leptin and liver tissue repair: do rodent models provide the answers?', *J Hepatol*, 46(1), pp. 12-8.

Marrero, J.A., Fontana, R.J., Su, G.L., Conjeevaram, H.S., Emick, D.M. and Lok, A.S. (2002) 'NAFLD may be a common underlying liver disease in patients with hepatocellular carcinoma in the United States', *Hepatology*, 36(6), pp. 1349-54.

Masarone, M., Federico, A., Abenavoli, L., Loguercio, C. and Persico, M. (2014) 'Non alcoholic fatty liver: epidemiology and natural history', *Rev Recent Clin Trials*, 9(3), pp. 126-33.

Mason, T.M. (1998) 'The role of factors that regulate the synthesis and secretion of very-low-density lipoprotein by hepatocytes', *Crit Rev Clin Lab Sci*, 35(6), pp. 461-87.

Mato, J.M., Corrales, F.J., Lu, S.C. and Avila, M.A. (2002) 'S-Adenosylmethionine: a control switch that regulates liver function', *FASEB J*, 16(1), pp. 15-26.

Matteoni, C.A., Younossi, Z.M., Gramlich, T., Boparai, N., Liu, Y.C. and McCullough, A.J. (1999) 'Nonalcoholic fatty liver disease: a spectrum of clinical and pathological severity', *Gastroenterology*, 116(6), pp. 1413-9.

McCullough, A.J. (2004) 'The clinical features, diagnosis and natural history of nonalcoholic fatty liver disease', *Clin Liver Dis*, 8(3), pp. 521-33, viii.

McGarry, J.D. and Brown, N.F. (1997) 'The mitochondrial carnitine palmitoyltransferase system. From concept to molecular analysis', *Eur J Biochem*, 244(1), pp. 1-14.

McGettrick, A.J., Feener, E.P. and Kahn, C.R. (2005) 'Human insulin receptor substrate-1 (IRS-1) polymorphism G972R causes IRS-1 to associate with the insulin receptor and inhibit receptor autophosphorylation', *J Biol Chem*, 280(8), pp. 6441-6.

McPherson, S., Hardy, T., Henderson, E., Burt, A.D., Day, C.P. and Anstee, Q.M. (2015) 'Evidence of NAFLD progression from steatosis to fibrosing-steatohepatitis using paired biopsies: implications for prognosis and clinical management', *J Hepatol*, 62(5), pp. 1148-55.

McPherson, S., Jonsson, J.R., Cowin, G.J., O'Rourke, P., Clouston, A.D., Volp, A., Horsfall, L., Jothimani, D., Fawcett, J., Galloway, G.J., Benson, M. and Powell, E.E. (2009) 'Magnetic resonance imaging and spectroscopy accurately estimate the severity of steatosis provided the stage of fibrosis is considered', *J Hepatol*, 51(2), pp. 389-97.

Meister, A. and Anderson, M.E. (1983) 'Glutathione', *Annu Rev Biochem*, 52, pp. 711-60.

Mendez-Sanchez, N., Arrese, M., Zamora-Valdes, D. and Uribe, M. (2007) 'Current concepts in the pathogenesis of nonalcoholic fatty liver disease', *Liver Int*, 27(4), pp. 423-33.

Michelotti, G.A., Machado, M.V. and Diehl, A.M. (2013) 'NAFLD, NASH and liver cancer', *Nat Rev Gastroenterol Hepatol*, 10(11), pp. 656-65.

Miele, L., Beale, G., Patman, G., Nobili, V., Leathart, J., Grieco, A., Abate, M., Friedman, S.L., Narla, G., Bugianesi, E., Day, C.P. and Reeves, H.L. (2008) 'The Kruppel-like factor 6 genotype is associated with fibrosis in nonalcoholic fatty liver disease', *Gastroenterology*, 135(1), pp. 282-291 e1.

Mittal, S., El-Serag, H.B., Sada, Y.H., Kanwal, F., Duan, Z., Temple, S., May, S.B., Kramer, J.R., Richardson, P.A. and Davila, J.A. (2016) 'Hepatocellular Carcinoma in the Absence of Cirrhosis in United States Veterans is Associated With Nonalcoholic Fatty Liver Disease', *Clin Gastroenterol Hepatol*, 14(1), pp. 124-31 e1.

Mummadi, R.R., Kasturi, K.S., Chennareddygari, S. and Sood, G.K. (2008) 'Effect of bariatric surgery on nonalcoholic fatty liver disease: systematic review and meta-analysis', *Clin Gastroenterol Hepatol*, 6(12), pp. 1396-402.

Musso, G., Gambino, R., Cassader, M. and Pagano, G. (2011) 'Meta-analysis: natural history of non-alcoholic fatty liver disease (NAFLD) and diagnostic accuracy of non-invasive tests for liver disease severity', *Ann Med*, 43(8), pp. 617-49.

Musso, G., Gambino, R., Pacini, G., Pagano, G., Durazzo, M. and Cassader, M. (2009) 'Transcription factor 7-like 2 polymorphism modulates glucose and lipid

homeostasis, adipokine profile, and hepatocyte apoptosis in NASH', *Hepatology*, 49(2), pp. 426-35.

Nadalin, S., Malago, M., Valentin-Gamazo, C., Testa, G., Baba, H.A., Liu, C., Fruhauf, N.R., Schaffer, R., Gerken, G., Frilling, A. and Broelsch, C.E. (2005) 'Preoperative donor liver biopsy for adult living donor liver transplantation: risks and benefits', *Liver Transpl*, 11(8), pp. 980-6.

Naik, A., Belic, A., Zanger, U.M. and Rozman, D. (2013) 'Molecular Interactions between NAFLD and Xenobiotic Metabolism', *Front Genet*, 4, p. 2.

Nakagawa, H., Umemura, A., Taniguchi, K., Font-Burgada, J., Dhar, D., Ogata, H., Zhong, Z., Valasek, M.A., Seki, E., Hidalgo, J., Koike, K., Kaufman, R.J. and Karin, M. (2014) 'ER stress cooperates with hypernutrition to trigger TNF-dependent spontaneous HCC development', *Cancer Cell*, 26(3), pp. 331-43.

Namikawa, C., Shu-Ping, Z., Vyselaar, J.R., Nozaki, Y., Nemoto, Y., Ono, M., Akisawa, N., Saibara, T., Hiroi, M., Enzan, H. and Onishi, S. (2004) 'Polymorphisms of microsomal triglyceride transfer protein gene and manganese superoxide dismutase gene in non-alcoholic steatohepatitis', *J Hepatol*, 40(5), pp. 781-6.

Nasrallah, S.M., Wills, C.E., Jr. and Galambos, J.T. (1981) 'Hepatic morphology in obesity', *Dig Dis Sci*, 26(4), pp. 325-7.

Nelson, J.E., Bhattacharya, R., Lindor, K.D., Chalasani, N., Raaka, S., Heathcote, E.J., Miskovsky, E., Shaffer, E., Rulyak, S.J. and Kowdley, K.V. (2007) 'HFE C282Y mutations are associated with advanced hepatic fibrosis in Caucasians with nonalcoholic steatohepatitis', *Hepatology*, 46(3), pp. 723-9.

Neuschwander-Tetri, B.A. and Caldwell, S.H. (2003) 'Nonalcoholic steatohepatitis: summary of an AASLD Single Topic Conference', *Hepatology*, 37(5), pp. 1202-19.

Nischalke, H.D., Berger, C., Luda, C., Berg, T., Muller, T., Grunhage, F., Lammert, F., Coenen, M., Kramer, B., Korner, C., Vidovic, N., Oldenburg, J., Nattermann, J., Sauerbruch, T. and Spengler, U. (2011) 'The PNPLA3 rs738409 148M/M genotype is a risk factor for liver cancer in alcoholic cirrhosis but shows no or weak association in hepatitis C cirrhosis', *PLoS One*, 6(11), p. e27087.

Oakley, F., Teoh, V., Ching, A.S.G., Bataller, R., Colmenero, J., Jonsson, J.R., Eliopoulos, A.G., Watson, M.R., Manas, D. and Mann, D.A. (2009) 'Angiotensin II activates I kappaB kinase phosphorylation of RelA at Ser 536 to promote myofibroblast survival and liver fibrosis', *Gastroenterology*, 136(7), pp. 2334-2344 e1.

Oliveira, C.P., Stefano, J.T., Cavaleiro, A.M., Zanella Fortes, M.A., Vieira, S.M., Rodrigues Lima, V.M., Santos, T.E., Santos, V.N., de Azevedo Salgado, A.L., Parise, E.R., Ferreira Alves, V.A., Carrilho, F.J. and Correa-Giannella, M.L. (2010) 'Association of polymorphisms of glutamate-cystein ligase and microsomal triglyceride transfer protein genes in non-alcoholic fatty liver disease', *J Gastroenterol Hepatol*, 25(2), pp. 357-61.

Omagari, K., Kadokawa, Y., Masuda, J., Egawa, I., Sawa, T., Hazama, H., Ohba, K., Isomoto, H., Mizuta, Y., Hayashida, K., Murase, K., Kadota, T., Murata, I. and Kohno, S. (2002) 'Fatty liver in non-alcoholic non-overweight Japanese adults: incidence and clinical characteristics', *J Gastroenterol Hepatol*, 17(10), pp. 1098-105.

Oni, E.T., Agatston, A.S., Blaha, M.J., Fialkow, J., Cury, R., Sposito, A., Erbel, R., Blankstein, R., Feldman, T., Al-Mallah, M.H., Santos, R.D., Budoff, M.J. and Nasir, K. (2013) 'A systematic review: burden and severity of subclinical cardiovascular disease among those with nonalcoholic fatty liver; should we care?', *Atherosclerosis*, 230(2), pp. 258-67.

Onnerhag, K., Nilsson, P.M. and Lindgren, S. (2014) 'Increased risk of cirrhosis and hepatocellular cancer during long-term follow-up of patients with biopsy-proven NAFLD', *Scand J Gastroenterol*, 49(9), pp. 1111-8.

Ortiz-Lopez, C., Lomonaco, R., Orsak, B., Finch, J., Chang, Z., Kochunov, V.G., Hardies, J. and Cusi, K. (2012) 'Prevalence of prediabetes and diabetes and metabolic profile of patients with nonalcoholic fatty liver disease (NAFLD)', *Diabetes Care*, 35(4), pp. 873-8.

Otto, D.A. and Veech, R.L. (1980) 'Isolation of a lipocyte-rich fraction from rat liver nonparenchymal cells', *Adv Exp Med Biol*, 132, pp. 509-17.

Pagano, C., Soardo, G., Esposito, W., Fallo, F., Basan, L., Donnini, D., Federspil, G., Sechi, L.A. and Vettor, R. (2005) 'Plasma adiponectin is decreased in nonalcoholic fatty liver disease', *Eur J Endocrinol*, 152(1), pp. 113-8.

Pais, R., Charlotte, F., Fedchuk, L., Bedossa, P., Lebray, P., Poynard, T., Ratziu, V. and Group, L.S. (2013) 'A systematic review of follow-up biopsies reveals disease progression in patients with non-alcoholic fatty liver', *J Hepatol*, 59(3), pp. 550-6.

Pajvani, U.B., Du, X., Combs, T.P., Berg, A.H., Rajala, M.W., Schulthess, T., Engel, J., Brownlee, M. and Scherer, P.E. (2003) 'Structure-function studies of the adipocyte-secreted hormone Acrp30/adiponectin. Implications fpr metabolic regulation and bioactivity', *J Biol Chem*, 278(11), pp. 9073-85.

Palmer, N.D., Musani, S.K., Yerges-Armstrong, L.M., Feitosa, M.F., Bielak, L.F., Hernaez, R., Kahali, B., Carr, J.J., Harris, T.B., Jhun, M.A., Kardia, S.L., Langefeld, C.D., Mosley, T.H., Jr., Norris, J.M., Smith, A.V., Taylor, H.A., Wagenknecht, L.E., Liu, J., Borecki, I.B., Peyser, P.A. and Speliotes, E.K. (2013) 'Characterization of European ancestry nonalcoholic fatty liver disease-associated variants in individuals of African and Hispanic descent', *Hepatology*, 58(3), pp. 966-75.

Paradis, V. (2013) 'Histopathology of hepatocellular carcinoma', *Recent Results Cancer Res*, 190, pp. 21-32.

Park, E.J., Lee, J.H., Yu, G.Y., He, G., Ali, S.R., Holzer, R.G., Osterreicher, C.H., Takahashi, H. and Karin, M. (2010) 'Dietary and genetic obesity promote liver

inflammation and tumorigenesis by enhancing IL-6 and TNF expression', *Cell*, 140(2), pp. 197-208.

Park, W.D., Blackwood, C., Mignery, G.A., Hermodson, M.A. and Lister, R.M. (1983) 'Analysis of the Heterogeneity of the 40,000 Molecular Weight Tuber Glycoprotein of Potatoes by Immunological Methods and by NH(2)-Terminal Sequence Analysis', *Plant Physiol*, 71(1), pp. 156-60.

Parkin, D.M. (2006) 'The global health burden of infection-associated cancers in the year 2002', *Int J Cancer*, 118(12), pp. 3030-44.

Payne, J.H., Dewind, L.T. and Commons, R.R. (1963) 'Metabolic Observations in Patients with Jejunocolic Shunts', *Am J Surg*, 106, pp. 273-89.

Peng, X.E., Wu, Y.L., Lin, S.W., Lu, Q.Q., Hu, Z.J. and Lin, X. (2012) 'Genetic variants in PNPLA3 and risk of non-alcoholic fatty liver disease in a Han Chinese population', *PLoS One*, 7(11), p. e50256.

Perry, R.J., Samuel, V.T., Petersen, K.F. and Shulman, G.I. (2014) 'The role of hepatic lipids in hepatic insulin resistance and type 2 diabetes', *Nature*, 510(7503), pp. 84-91.

Perttila, J., Huaman-Samanez, C., Caron, S., Tanhuanpaa, K., Staels, B., Yki-Jarvinen, H. and Olkkonen, V.M. (2012a) 'PNPLA3 is regulated by glucose in human hepatocytes, and its I148M mutant slows down triglyceride hydrolysis', *Am J Physiol Endocrinol Metab*, 302(9), pp. E1063-9.

Perttila, J., Huaman-Samanez, C., Caron, S., Tanhuanpaa, K., Staels, B., Yki-Jarvinen, H. and Olkkonen, V.M. (2012b) 'PNPLA3 is regulated by glucose in human hepatocytes, and its I148M mutant slows down triglyceride hydrolysis', *American journal of physiology. Endocrinology and metabolism*, 302(9), pp. E1063-9.

Pessayre, D. (2007) 'Role of mitochondria in non-alcoholic fatty liver disease', *J Gastroenterol Hepatol*, 22 Suppl 1, pp. S20-7.

Pessayre, D., Berson, A., Fromenty, B. and Mansouri, A. (2001) 'Mitochondria in steatohepatitis', *Semin Liver Dis*, 21(1), pp. 57-69.

Pessayre, D. and Fromenty, B. (2005) 'NASH: a mitochondrial disease', *J Hepatol*, 42(6), pp. 928-40.

Petersen, K.F., Dufour, S., Hariri, A., Nelson-Williams, C., Foo, J.N., Zhang, X.M., Dziura, J., Lifton, R.P. and Shulman, G.I. (2010) 'Apolipoprotein C3 gene variants in nonalcoholic fatty liver disease', *N Engl J Med*, 362(12), pp. 1082-9.

Petta, S., Miele, L., Bugianesi, E., Camma, C., Rosso, C., Boccia, S., Cabibi, D., Di Marco, V., Grimaudo, S., Grieco, A., Pipitone, R.M., Marchesini, G. and Craxi, A. (2014) 'Glucokinase regulatory protein gene polymorphism affects liver fibrosis in non-alcoholic fatty liver disease', *PLoS One*, 9(2), p. e87523.

Pingitore, P., Dongiovanni, P., Motta, B.M., Meroni, M., Lepore, S.M., Mancina, R.M., Pelusi, S., Russo, C., Caddeo, A., Rossi, G., Montalcini, T., Pujia, A., Wiklund, O., Valenti, L. and Romeo, S. (2016) 'PNPLA3 overexpression results in reduction of proteins predisposing to fibrosis', *Hum Mol Genet*, 25(23), pp. 5212-5222.

Pirazzi, C., Adiels, M., Burza, M.A., Mancina, R.M., Levin, M., Stahlman, M., Taskinen, M.R., Orho-Melander, M., Perman, J., Pujia, A., Andersson, L., Maglio, C., Montalcini, T., Wiklund, O., Boren, J. and Romeo, S. (2012) 'Patatin-like phospholipase domain-containing 3 (PNPLA3) I148M (rs738409) affects hepatic VLDL secretion in humans and in vitro', *J Hepatol*, 57(6), pp. 1276-82.

Pirazzi, C., Valenti, L., Motta, B.M., Pingitore, P., Hedfalk, K., Mancina, R.M., Burza, M.A., Indiveri, C., Ferro, Y., Montalcini, T., Maglio, C., Dongiovanni, P., Fargion, S., Rametta, R., Pujia, A., Andersson, L., Ghosal, S., Levin, M., Wiklund, O., Iacovino, M., Boren, J. and Romeo, S. (2014) 'PNPLA3 has retinyl-palmitate lipase activity in human hepatic stellate cells', *Hum Mol Genet*, 23(15), pp. 4077-85.

Powell, E.E., Cooksley, W.G., Hanson, R., Searle, J., Halliday, J.W. and Powell, L.W. (1990) 'The natural history of nonalcoholic steatohepatitis: a follow-up study of forty-two patients for up to 21 years', *Hepatology*, 11(1), pp. 74-80.

Purcell, S., Cherny, S.S. and Sham, P.C. (2003) 'Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits', *Bioinformatics*, 19(1), pp. 149-50.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J. and Sham, P.C. (2007) 'PLINK: a tool set for whole-genome association and population-based linkage analyses', *Am J Hum Genet*, 81(3), pp. 559-75.

Puri, P., Mirshahi, F., Cheung, O., Natarajan, R., Maher, J.W., Kellum, J.M. and Sanyal, A.J. (2008) 'Activation and dysregulation of the unfolded protein response in nonalcoholic fatty liver disease', *Gastroenterology*, 134(2), pp. 568-76.

Rae-Whitcombe, S.M., Kennedy, D., Voyles, M. and Thompson, M.P. (2010) 'Regulation of the promoter region of the human adiponutrin/PNPLA3 gene by glucose and insulin', *Biochem Biophys Res Commun*, 402(4), pp. 767-72.

Rafiq, N., Bai, C., Fang, Y., Srishord, M., McCullough, A., Gramlich, T. and Younossi, Z.M. (2009) 'Long-term follow-up of patients with nonalcoholic fatty liver', *Clin Gastroenterol Hepatol*, 7(2), pp. 234-8.

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A. and Zhang, F. (2013) 'Genome engineering using the CRISPR-Cas9 system', *Nat Protoc*, 8(11), pp. 2281-308.

Randolph, R.K. and Ross, A.C. (1991) 'Regulation of retinol uptake and esterification in MCF-7 and HepG2 cells by exogenous fatty acids', *J Lipid Res*, 32(5), pp. 809-20.

Raszeja-Wyszomirska, J., Kurzawski, G., Lawniczak, M., Miezynska-Kurtycz, J. and Lubinski, J. (2010) 'Nonalcoholic fatty liver disease and HFE gene mutations: a Polish study', *World J Gastroenterol*, 16(20), pp. 2531-6.

Ratziu, V., Bellentani, S., Cortez-Pinto, H., Day, C. and Marchesini, G. (2010) 'A position statement on NAFLD/NASH based on the EASL 2009 special conference', *J Hepatol*, 53(2), pp. 372-84.

Ratziu, V., Harrison, S.A., Francque, S., Bedossa, P., Lehert, P., Serfaty, L., Romero-Gomez, M., Boursier, J., Abdelmalek, M., Caldwell, S., Drenth, J., Anstee, Q.M., Hum, D., Hanf, R., Roudot, A., Megnien, S., Staels, B., Sanyal, A. and Group, G.-I.S. (2016) 'Elafibranor, an Agonist of the Peroxisome Proliferator-Activated Receptor-alpha and -delta, Induces Resolution of Nonalcoholic Steatohepatitis Without Fibrosis Worsening', *Gastroenterology*, 150(5), pp. 1147-1159 e5.

Ratziu, V., Lalazar, A., Wong, L., Dang, Q., Collins, C., Shaulian, E., Jensen, S. and Friedman, S.L. (1998) 'Zf9, a Kruppel-like transcription factor up-regulated in vivo during early hepatic fibrosis', *Proc Natl Acad Sci U S A*, 95(16), pp. 9500-5.

Reddy, J.K. (2001) 'Nonalcoholic steatosis and steatohepatitis. III. Peroxisomal beta-oxidation, PPAR alpha, and steatohepatitis', *Am J Physiol Gastrointest Liver Physiol*, 281(6), pp. G1333-9.

Repa, J.J., Liang, G., Ou, J., Bashmakov, Y., Lobaccaro, J.M., Shimomura, I., Shan, B., Brown, M.S., Goldstein, J.L. and Mangelsdorf, D.J. (2000) 'Regulation of mouse sterol regulatory element-binding protein-1c gene (SREBP-1c) by oxysterol receptors, LXRalpha and LXRbeta', *Genes Dev*, 14(22), pp. 2819-30.

Reue, K. (2009) 'The lipin family: mutations and metabolism', *Curr Opin Lipidol*, 20(3), pp. 165-70.

Reue, K. and Zhang, P. (2008) 'The lipin protein family: dual roles in lipid biosynthesis and gene expression', *FEBS Lett*, 582(1), pp. 90-6.

Rey, J.W., Noetel, A., Hardt, A., Canbay, A., Alakus, H., Zur Hausen, A., Dienes, H.P., Drebber, U. and Odenthal, M. (2010) 'Pro12Ala polymorphism of the peroxisome proliferator-activated receptor gamma2 in patients with fatty liver diseases', *World J Gastroenterol*, 16(46), pp. 5830-7.

Ricchi, M., Odoardi, M.R., Carulli, L., Anzivino, C., Ballestri, S., Pinetti, A., Fantoni, L.I., Marra, F., Bertolotti, M., Banni, S., Lonardo, A., Carulli, N. and Loria, P. (2009) 'Differential effect of oleic and palmitic acid on lipid accumulation and apoptosis in cultured hepatocytes', *J Gastroenterol Hepatol*, 24(5), pp. 830-40.

Risch, N. and Merikangas, K. (1996) 'The future of genetic studies of complex human diseases', *Science*, 273(5281), pp. 1516-7.

Robertson, G., Leclercq, I. and Farrell, G.C. (2001) 'Nonalcoholic steatosis and steatohepatitis. II. Cytochrome P-450 enzymes and oxidative stress', *Am J Physiol Gastrointest Liver Physiol*, 281(5), pp. G1135-9.

Romeo, S., Huang-Doran, I., Baroni, M.G. and Kotronen, A. (2010) 'Unravelling the pathogenesis of fatty liver disease: patatin-like phospholipase domain-containing 3 protein', *Curr Opin Lipidol*, 21(3), pp. 247-52.

Romeo, S., Kozlitina, J., Xing, C., Pertsemlidis, A., Cox, D., Pennacchio, L.A., Boerwinkle, E., Cohen, J.C. and Hobbs, H.H. (2008) 'Genetic variation in PNPLA3 confers susceptibility to nonalcoholic fatty liver disease', *Nat Genet*, 40(12), pp. 1461-5.

Rotman, Y., Koh, C., Zmuda, J.M., Kleiner, D.E. and Liang, T.J. (2010) 'The association of genetic variability in patatin-like phospholipase domain-containing protein 3 (PNPLA3) with histological severity of nonalcoholic fatty liver disease', *Hepatology*, 52(3), pp. 894-903.

Ruhanen, H., Perttila, J., Holtta-Vuori, M., Zhou, Y., Yki-Jarvinen, H., Ikonen, E., Kakela, R. and Olkkonen, V.M. (2014) 'PNPLA3 mediates hepatocyte triacylglycerol remodeling', *J Lipid Res*, 55(4), pp. 739-46.

Ryan, C.K., Johnson, L.A., Germin, B.I. and Marcos, A. (2002) 'One hundred consecutive hepatic biopsies in the workup of living donors for right lobe liver transplantation', *Liver Transpl*, 8(12), pp. 1114-22.

Ryu, S., Chang, Y., Choi, Y., Kwon, M.J., Kim, C.W., Yun, K.E., Jung, H.S., Kim, B.K., Kim, Y.J., Ahn, J., Cho, Y.K., Kim, K.H., Chung, E.C., Shin, H. and Cho, J. (2015a) 'Age at menarche and non-alcoholic fatty liver disease', *J Hepatol*, 62(5), pp. 1164-70.

Ryu, S., Suh, B.S., Chang, Y., Kwon, M.J., Yun, K.E., Jung, H.S., Kim, C.W., Kim, B.K., Kim, Y.J., Choi, Y., Ahn, J., Cho, Y.K., Kim, K.H., Ahn, Y., Park, H.Y., Chung, E.C., Shin, H. and Cho, J. (2015b) 'Menopausal stages and non-alcoholic fatty liver disease in middle-aged women', *Eur J Obstet Gynecol Reprod Biol*, 190, pp. 65-70.

Saltiel, A.R. and Kahn, C.R. (2001) 'Insulin signalling and the regulation of glucose and lipid metabolism', *Nature*, 414(6865), pp. 799-806.

Sanyal, A., Poklepovic, A., Moyneur, E. and Barghout, V. (2010) 'Population-based risk factors and resource utilization for HCC: US perspective', *Curr Med Res Opin*, 26(9), pp. 2183-91.

Sanyal, A.J. and American Gastroenterological, A. (2002) 'AGA technical review on nonalcoholic fatty liver disease', *Gastroenterology*, 123(5), pp. 1705-25.

Sanyal, A.J., Banas, C., Sargeant, C., Luketic, V.A., Sterling, R.K., Stravitz, R.T., Shiffman, M.L., Heuman, D., Coterrell, A., Fisher, R.A., Contos, M.J. and Mills, A.S.

(2006) 'Similarities and differences in outcomes of cirrhosis due to nonalcoholic steatohepatitis and hepatitis C', *Hepatology*, 43(4), pp. 682-9.

Sanyal, A.J., Brunt, E.M., Kleiner, D.E., Kowdley, K.V., Chalasani, N., Lavine, J.E., Ratziu, V. and McCullough, A. (2011) 'Endpoints and clinical trial design for nonalcoholic steatohepatitis', *Hepatology*, 54(1), pp. 344-53.

Sarasin, F.P., Giostra, E. and Hadengue, A. (1996) 'Cost-effectiveness of screening for detection of small hepatocellular carcinoma in western patients with Child-Pugh class A cirrhosis', *Am J Med*, 101(4), pp. 422-34.

Satapathy, S.K. and Sanyal, A.J. (2015) 'Epidemiology and Natural History of Nonalcoholic Fatty Liver Disease', *Semin Liver Dis*, 35(3), pp. 221-35.

Sato, R., Miyamoto, W., Inoue, J., Terada, T., Imanaka, T. and Maeda, M. (1999) 'Sterol regulatory element-binding protein negatively regulates microsomal triglyceride transfer protein gene transcription', *J Biol Chem*, 274(35), pp. 24714-20.

Saxena, N.K., Sharma, D., Ding, X., Lin, S., Marra, F., Merlin, D. and Anania, F.A. (2007) 'Concomitant activation of the JAK/STAT, PI3K/AKT, and ERK signaling is involved in leptin-mediated promotion of invasion and migration of hepatocellular carcinoma cells', *Cancer Res*, 67(6), pp. 2497-507.

Sazci, A., Akpinar, G., Aygun, C., Ergul, E., Senturk, O. and Hulagu, S. (2008) 'Association of apolipoprotein E polymorphisms in patients with non-alcoholic steatohepatitis', *Dig Dis Sci*, 53(12), pp. 3218-24.

Schaffner, F. and Thaler, H. (1986) 'Nonalcoholic fatty liver disease', *Prog Liver Dis*, 8, pp. 283-98.

Schreuder, T.C., Verwer, B.J., van Nieuwkerk, C.M. and Mulder, C.J. (2008) 'Nonalcoholic fatty liver disease: an overview of current insights in pathogenesis, diagnosis and treatment', *World J Gastroenterol*, 14(16), pp. 2474-86.

Schwimmer, J.B., Celedon, M.A., Lavine, J.E., Salem, R., Campbell, N., Schork, N.J., Shiehmorteza, M., Yokoo, T., Chavez, A., Middleton, M.S. and Sirlin, C.B. (2009) 'Heritability of nonalcoholic fatty liver disease', *Gastroenterology*, 136(5), pp. 1585-92.

Senoo, H., Yoshikawa, K., Morii, M., Miura, M., Imai, K. and Mezaki, Y. (2010) 'Hepatic stellate cell (vitamin A-storing cell) and its relative--past, present and future', *Cell Biol Int*, 34(12), pp. 1247-72.

Sentinelli, F., Romeo, S., Maglio, C., Incani, M., Burza, M.A., Scano, F., Coccia, F., Cossu, E., Leonetti, F. and Baroni, M.G. (2011) 'Lack of effect of apolipoprotein C3 polymorphisms on indices of liver steatosis, lipid profile and insulin resistance in obese Southern Europeans', *Lipids Health Dis*, 10, p. 93.

Seppala-Lindroos, A., Vehkavaara, S., Hakkinen, A.M., Goto, T., Westerbacka, J., Sovijarvi, A., Halavaara, J. and Yki-Jarvinen, H. (2002) 'Fat accumulation in the liver is associated with defects in insulin suppression of glucose production and serum free fatty acids independent of obesity in normal men', *J Clin Endocrinol Metab*, 87(7), pp. 3023-8.

Serviddio, G., Bellanti, F., Tamborra, R., Rollo, T., Romano, A.D., Giudetti, A.M., Capitanio, N., Petrella, A., Vendemiale, G. and Altomare, E. (2008) 'Alterations of hepatic ATP homeostasis and respiratory chain during development of non-alcoholic steatohepatitis in a rodent model', *Eur J Clin Invest*, 38(4), pp. 245-52.

Serviddio, G., Bellanti, F., Vendemiale, G. and Altomare, E. (2011) 'Mitochondrial dysfunction in nonalcoholic steatohepatitis', *Expert Rev Gastroenterol Hepatol*, 5(2), pp. 233-44.

Sheth, S.G., Gordon, F.D. and Chopra, S. (1997) 'Nonalcoholic steatohepatitis', *Ann Intern Med*, 126(2), pp. 137-45.

Shibata, M., Kihara, Y., Taguchi, M., Tashiro, M. and Otsuki, M. (2007) 'Nonalcoholic fatty liver disease is a risk factor for type 2 diabetes in middle-aged Japanese men', *Diabetes Care*, 30(11), pp. 2940-4.

Shimano, H., Yahagi, N., Amemiya-Kudo, M., Hasty, A.H., Osuga, J., Tamura, Y., Shionoiri, F., Iizuka, Y., Ohashi, K., Harada, K., Gotoda, T., Ishibashi, S. and Yamada, N. (1999) 'Sterol regulatory element-binding protein-1 as a key transcription factor for nutritional induction of lipogenic enzyme genes', *J Biol Chem*, 274(50), pp. 35832-9.

Shimomura, I., Bashmakov, Y. and Horton, J.D. (1999a) 'Increased levels of nuclear SREBP-1c associated with fatty livers in two mouse models of diabetes mellitus', *J Biol Chem*, 274(42), pp. 30028-32.

Shimomura, I., Bashmakov, Y., Ikemoto, S., Horton, J.D., Brown, M.S. and Goldstein, J.L. (1999b) 'Insulin selectively increases SREBP-1c mRNA in the livers of rats with streptozotocin-induced diabetes', *Proc Natl Acad Sci U S A*, 96(24), pp. 13656-61.

Silverman, J.F., Pories, W.J. and Caro, J.F. (1989) 'Liver pathology in diabetes mellitus and morbid obesity. Clinical, pathological, and biochemical considerations', *Pathol Annu*, 24 Pt 1, pp. 275-302.

Simon, H.U., Haj-Yehia, A. and Levi-Schaffer, F. (2000) 'Role of reactive oxygen species (ROS) in apoptosis induction', *Apoptosis*, 5(5), pp. 415-8.

Singh, S., Allen, A.M., Wang, Z., Prokop, L.J., Murad, M.H. and Loomba, R. (2015) 'Fibrosis progression in nonalcoholic fatty liver vs nonalcoholic steatohepatitis: a systematic review and meta-analysis of paired-biopsy studies', *Clin Gastroenterol Hepatol*, 13(4), pp. 643-54 e1-9; quiz e39-40.

Smagris, E., BasuRay, S., Li, J., Huang, Y., Lai, K.M., Gromada, J., Cohen, J.C. and Hobbs, H.H. (2015) 'Pnpla3I148M knockin mice accumulate PNPLA3 on lipid droplets and develop hepatic steatosis', *Hepatology*, 61(1), pp. 108-18.

Smith, B.W. and Adams, L.A. (2011) 'Nonalcoholic fatty liver disease and diabetes mellitus: pathogenesis and treatment', *Nat Rev Endocrinol*, 7(8), pp. 456-65.

Smith, K. (2013) 'Liver disease: Kupffer cells regulate the progression of ALD and NAFLD', *Nat Rev Gastroenterol Hepatol*, 10(9), p. 503.

Soderberg, C., Stal, P., Askling, J., Glaumann, H., Lindberg, G., Marmur, J. and Hultcrantz, R. (2010) 'Decreased survival of subjects with elevated liver function tests during a 28-year follow-up', *Hepatology*, 51(2), pp. 595-602.

Song, J., da Costa, K.A., Fischer, L.M., Kohlmeier, M., Kwock, L., Wang, S. and Zeisel, S.H. (2005) 'Polymorphism of the PEMT gene and susceptibility to nonalcoholic fatty liver disease (NAFLD)', *FASEB J*, 19(10), pp. 1266-71.

Sookoian, S., Castano, G., Gianotti, T.F., Gemma, C. and Pirola, C.J. (2009a) 'Polymorphisms of MRP2 (ABCC2) are associated with susceptibility to nonalcoholic fatty liver disease', *J Nutr Biochem*, 20(10), pp. 765-70.

Sookoian, S., Castano, G.O., Burgueno, A.L., Gianotti, T.F., Rosselli, M.S. and Pirola, C.J. (2009b) 'A nonsynonymous gene variant in the adiponutrin gene is associated with nonalcoholic fatty liver disease severity', *J Lipid Res*, 50(10), pp. 2111-6.

Sookoian, S., Castano, G.O., Burgueno, A.L., Gianotti, T.F., Rosselli, M.S. and Pirola, C.J. (2010) 'The nuclear receptor PXR gene variants are associated with liver injury in nonalcoholic fatty liver disease', *Pharmacogenet Genomics*, 20(1), pp. 1-8.

Sookoian, S., Castano, G.O., Scian, R., Mallardi, P., Fernandez Gianotti, T., Burgueno, A.L., San Martino, J. and Pirola, C.J. (2015) 'Genetic variation in transmembrane 6 superfamily member 2 and the risk of nonalcoholic fatty liver disease and histological disease severity', *Hepatology*, 61(2), pp. 515-25.

Sookoian, S. and Pirola, C.J. (2011) 'Meta-analysis of the influence of I148M variant of patatin-like phospholipase domain containing 3 gene (PNPLA3) on the susceptibility and histological severity of nonalcoholic fatty liver disease', *Hepatology*, 53(6), pp. 1883-94.

Sookoian, S. and Pirola, C.J. (2013) 'PNPLA3, the history of an orphan gene of the potate tuber PROTEIN family that found an organ: the Liver', *Hepatology*.

Speliotes, E.K., Yerges-Armstrong, L.M., Wu, J., Hernaez, R., Kim, L.J., Palmer, C.D., Gudnason, V., Eiriksdottir, G., Garcia, M.E., Launer, L.J., Nalls, M.A., Clark, J.M., Mitchell, B.D., Shuldiner, A.R., Butler, J.L., Tomas, M., Hoffmann, U., Hwang, S.J., Massaro, J.M., O'Donnell, C.J., Sahani, D.V., Salomaa, V., Schadt, E.E., Schwartz, S.M., Siscovick, D.S., Nash, C.R.N., Consortium, G., Investigators, M.,

Voight, B.F., Carr, J.J., Feitosa, M.F., Harris, T.B., Fox, C.S., Smith, A.V., Kao, W.H., Hirschhorn, J.N., Borecki, I.B. and Consortium, G. (2011) 'Genome-wide association analysis identifies variants associated with nonalcoholic fatty liver disease that have distinct effects on metabolic traits', *PLoS Genet*, 7(3), p. e1001324.

Spruss, A., Kanuri, G., Wagnerberger, S., Haub, S., Bischoff, S.C. and Bergheim, I. (2009) 'Toll-like receptor 4 is involved in the development of fructose-induced hepatic steatosis in mice', *Hepatology*, 50(4), pp. 1094-104.

Starley, B.Q., Calcagno, C.J. and Harrison, S.A. (2010) 'Nonalcoholic fatty liver disease and hepatocellular carcinoma: a weighty connection', *Hepatology*, 51(5), pp. 1820-32.

Struben, V.M., Hespenheide, E.E. and Caldwell, S.H. (2000) 'Nonalcoholic steatohepatitis and cryptogenic cirrhosis within kindreds', *Am J Med*, 108(1), pp. 9-13.

Sugimoto, K. and Takei, Y. (2011) 'Clinicopathological features of non-alcoholic fatty liver disease', *Hepatol Res*, 41(10), pp. 911-20.

Sung, K.C., Jeong, W.S., Wild, S.H. and Byrne, C.D. (2012) 'Combined influence of insulin resistance, overweight/obesity, and fatty liver as risk factors for type 2 diabetes', *Diabetes Care*, 35(4), pp. 717-22.

Surakka, I., Horikoshi, M., Magi, R., Sarin, A.P., Mahajan, A., Lagou, V., Marullo, L., Ferreira, T., Miraglio, B., Timonen, S., Kettunen, J., Pirinen, M., Karjalainen, J., Thorleifsson, G., Hagg, S., Hottenga, J.J., Isaacs, A., Ladenvall, C., Beekman, M., Esko, T., Ried, J.S., Nelson, C.P., Willenborg, C., Gustafsson, S., Westra, H.J., Blades, M., de Craen, A.J., de Geus, E.J., Deelen, J., Grallert, H., Hamsten, A., Havulinna, A.S., Hengstenberg, C., Houwing-Duistermaat, J.J., Hypponen, E., Karssen, L.C., Lehtimaki, T., Lyssenko, V., Magnusson, P.K., Mihailov, E., Muller-Nurasyid, M., Mpindi, J.P., Pedersen, N.L., Penninx, B.W., Perola, M., Pers, T.H., Peters, A., Rung, J., Smit, J.H., Steinthorsdottir, V., Tobin, M.D., Tsernikova, N., van Leeuwen, E.M., Viikari, J.S., Willems, S.M., Willemsen, G., Schunkert, H., Erdmann, J., Samani, N.J., Kaprio, J., Lind, L., Gieger, C., Metspalu, A., Slagboom, P.E., Groop, L., van Duijn, C.M., Eriksson, J.G., Jula, A., Salomaa, V., Boomsma, D.I., Power, C., Raitakari, O.T., Ingelsson, E., Jarvelin, M.R., Thorsteinsdottir, U., Franke, L., Ikonen, E., Kallioniemi, O., Pietiainen, V., Lindgren, C.M., Stefansson, K., Palotie, A., McCarthy, M.I., Morris, A.P., Prokopenko, I., Ripatti, S. and Consortium, E. (2015) 'The impact of low-frequency and rare variants on lipid levels', *Nat Genet*, 47(6), pp. 589-97.

Szczepaniak, L.S., Nurenberg, P., Leonard, D., Browning, J.D., Reingold, J.S., Grundy, S., Hobbs, H.H. and Dobbins, R.L. (2005) 'Magnetic resonance spectroscopy to measure hepatic triglyceride content: prevalence of hepatic steatosis in the general population', *Am J Physiol Endocrinol Metab*, 288(2), pp. E462-8.

Tan, H.L., Zain, S.M., Mohamed, R., Rampal, S., Chin, K.F., Basu, R.C., Cheah, P.L., Mahadeva, S. and Mohamed, Z. (2014) 'Association of glucokinase regulatory gene

polymorphisms with risk and severity of non-alcoholic fatty liver disease: an interaction study with adiponutrin gene', *J Gastroenterol*, 49(6), pp. 1056-64.

Tanaka, S., Mohr, L., Schmidt, E.V., Sugimachi, K. and Wands, J.R. (1997) 'Biological effects of human insulin receptor substrate-1 overexpression in hepatocytes', *Hepatology*, 26(3), pp. 598-604.

Targher, G., Bertolini, L., Padovani, R., Rodella, S., Tessari, R., Zenari, L., Day, C. and Arcaro, G. (2007a) 'Prevalence of nonalcoholic fatty liver disease and its association with cardiovascular disease among type 2 diabetic patients', *Diabetes Care*, 30(5), pp. 1212-8.

Targher, G., Bertolini, L., Rodella, S., Tessari, R., Zenari, L., Lippi, G. and Arcaro, G. (2007b) 'Nonalcoholic fatty liver disease is independently associated with an increased incidence of cardiovascular events in type 2 diabetic patients', *Diabetes Care*, 30(8), pp. 2119-21.

Targher, G., Bertolini, L., Rodella, S., Zoppini, G., Lippi, G., Day, C. and Muggeo, M. (2008) 'Non-alcoholic fatty liver disease is independently associated with an increased prevalence of chronic kidney disease and proliferative/laser-treated retinopathy in type 2 diabetic patients', *Diabetologia*, 51(3), pp. 444-50.

Targher, G., Bertolini, L., Scala, L., Poli, F., Zenari, L. and Falezza, G. (2004) 'Decreased plasma adiponectin concentrations are closely associated with nonalcoholic hepatic steatosis in obese individuals', *Clin Endocrinol (Oxf)*, 61(6), pp. 700-3.

Targher, G., Chonchol, M., Zoppini, G., Abaterusso, C. and Bonora, E. (2011) 'Risk of chronic kidney disease in patients with non-alcoholic fatty liver disease: is there a link?', *J Hepatol*, 54(5), pp. 1020-9.

Teslovich, T.M., Musunuru, K., Smith, A.V., Edmondson, A.C., Stylianou, I.M., Koseki, M., Pirruccello, J.P., Ripatti, S., Chasman, D.I., Willer, C.J., Johansen, C.T., Fouchier, S.W., Isaacs, A., Peloso, G.M., Barbalic, M., Ricketts, S.L., Bis, J.C., Aulchenko, Y.S., Thorleifsson, G., Feitosa, M.F., Chambers, J., Orho-Melander, M., Melander, O., Johnson, T., Li, X., Guo, X., Li, M., Shin Cho, Y., Jin Go, M., Jin Kim, Y., Lee, J.Y., Park, T., Kim, K., Sim, X., Twee-Hee Ong, R., Croteau-Chonka, D.C., Lange, L.A., Smith, J.D., Song, K., Hua Zhao, J., Yuan, X., Luan, J., Lamina, C., Ziegler, A., Zhang, W., Zee, R.Y., Wright, A.F., Witteman, J.C., Wilson, J.F., Willemsen, G., Wichmann, H.E., Whitfield, J.B., Waterworth, D.M., Wareham, N.J., Waeber, G., Vollenweider, P., Voight, B.F., Vitart, V., Uitterlinden, A.G., Uda, M., Tuomilehto, J., Thompson, J.R., Tanaka, T., Surakka, I., Stringham, H.M., Spector, T.D., Soranzo, N., Smit, J.H., Sinisalo, J., Silander, K., Sijbrands, E.J., Scuteri, A., Scott, J., Schlessinger, D., Sanna, S., Salomaa, V., Saharinen, J., Sabatti, C., Ruokonen, A., Rudan, I., Rose, L.M., Roberts, R., Rieder, M., Psaty, B.M., Pramstaller, P.P., Pichler, I., Perola, M., Penninx, B.W., Pedersen, N.L., Pattaro, C., Parker, A.N., Pare, G., Oostra, B.A., O'Donnell, C.J., Nieminen, M.S., Nickerson, D.A., Montgomery, G.W., Meitinger, T., McPherson, R., McCarthy, M.I., et al. (2010) 'Biological, clinical and population relevance of 95 loci for blood lipids', *Nature*, 466(7307), pp. 707-13.

Thomas, E.L., Hamilton, G., Patel, N., O'Dwyer, R., Dore, C.J., Goldin, R.D., Bell, J.D. and Taylor-Robinson, S.D. (2005) 'Hepatic triglyceride content and its relation to body adiposity: a magnetic resonance imaging and proton magnetic resonance spectroscopy study', *Gut*, 54(1), pp. 122-7.

Tilg, H. and Moschen, A.R. (2010) 'Evolution of inflammation in nonalcoholic fatty liver disease: the multiple parallel hits hypothesis', *Hepatology*, 52(5), pp. 1836-46.

Tokushige, K., Takakura, M., Tsuchiya-Matsushita, N., Taniai, M., Hashimoto, E. and Shiratori, K. (2007) 'Influence of TNF gene polymorphisms in Japanese patients with NASH and simple steatosis', *J Hepatol*, 46(6), pp. 1104-10.

Tonjes, A., Scholz, M., Loeffler, M. and Stumvoll, M. (2006) 'Association of Pro12Ala polymorphism in peroxisome proliferator-activated receptor gamma with Pre-diabetic phenotypes: meta-analysis of 57 studies on nondiabetic individuals', *Diabetes Care*, 29(11), pp. 2489-97.

Towbin, H., Staehelin, T. and Gordon, J. (1979) 'Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications', *Proc Natl Acad Sci U S A*, 76(9), pp. 4350-4.

Towle, H.C., Kaytor, E.N. and Shih, H.M. (1997) 'Regulation of the expression of lipogenic enzyme genes by carbohydrate', *Annu Rev Nutr*, 17, pp. 405-33.

Tran, T.T., Changsri, C., Shackleton, C.R., Poordad, F.F., Nissen, N.N., Colquhoun, S., Geller, S.A., Vierling, J.M. and Martin, P. (2006) 'Living donor liver transplantation: histological abnormalities found on liver biopsies of apparently healthy potential donors', *J Gastroenterol Hepatol*, 21(2), pp. 381-3.

Treeprasertsuk, S., Leverage, S., Adams, L.A., Lindor, K.D., St Sauver, J. and Angulo, P. (2012) 'The Framingham risk score and heart disease in nonalcoholic fatty liver disease', *Liver Int*, 32(6), pp. 945-50.

Trepo, E., Guyot, E., Ganne-Carrie, N., Degre, D., Gustot, T., Franchimont, D., Sutton, A., Nahon, P. and Moreno, C. (2012) 'PNPLA3 (rs738409 C>G) is a common risk variant associated with hepatocellular carcinoma in alcoholic cirrhosis', *Hepatology*, 55(4), pp. 1307-8.

Trepo, E., Nahon, P., Bontempi, G., Valenti, L., Falleti, E., Nischalke, H.D., Hamza, S., Corradini, S.G., Burza, M.A., Guyot, E., Donati, B., Spengler, U., Hillon, P., Toniutto, P., Henrion, J., Franchimont, D., Deviere, J., Mathurin, P., Moreno, C., Romeo, S. and Deltenre, P. (2014) 'Association between the PNPLA3 (rs738409 C>G) variant and hepatocellular carcinoma: Evidence from a meta-analysis of individual participant data', *Hepatology*, 59(6), pp. 2170-7.

Trevisani, F., Santi, V., Gramenzi, A., Di Nolfo, M.A., Del Poggio, P., Benvegnu, L., Rapaccini, G., Farinati, F., Zoli, M., Borzio, F., Giannini, E.G., Caturelli, E. and Bernardi, M. (2007) 'Surveillance for early diagnosis of hepatocellular carcinoma: is it

effective in intermediate/advanced cirrhosis?', *Am J Gastroenterol*, 102(11), pp. 2448-57; quiz 2458.

Utermann, G., Hees, M. and Steinmetz, A. (1977) 'Polymorphism of apolipoprotein E and occurrence of dysbetalipoproteinaemia in man', *Nature*, 269(5629), pp. 604-7.

Vacca, M., Allison, M., Griffin, J.L. and Vidal-Puig, A. (2015) 'Fatty Acid and Glucose Sensors in Hepatic Lipid Metabolism: Implications in NAFLD', *Semin Liver Dis*, 35(3), pp. 250-61.

Valenti, L., Al-Serri, A., Daly, A.K., Galmozzi, E., Rametta, R., Dongiovanni, P., Nobili, V., Mozzi, E., Roviaro, G., Vanni, E., Bugianesi, E., Maggioni, M., Fracanzani, A.L., Fargion, S. and Day, C.P. (2010a) 'Homozygosity for the patatin-like phospholipase-3/adiponutrin I148M polymorphism influences liver fibrosis in patients with nonalcoholic fatty liver disease', *Hepatology*, 51(4), pp. 1209-17.

Valenti, L., Alisi, A., Galmozzi, E., Bartuli, A., Del Menico, B., Alterio, A., Dongiovanni, P., Fargion, S. and Nobili, V. (2010b) 'I148M patatin-like phospholipase domain-containing 3 gene variant and severity of pediatric nonalcoholic fatty liver disease', *Hepatology*, 52(4), pp. 1274-80.

Valenti, L., Dongiovanni, P., Ginanni Corradini, S., Burza, M.A. and Romeo, S. (2013) 'PNPLA3 I148M variant and hepatocellular carcinoma: A common genetic variant for a rare disease', *Dig Liver Dis*, 45(8), pp. 619-24.

Valenti, L., Fracanzani, A.L., Bugianesi, E., Dongiovanni, P., Galmozzi, E., Vanni, E., Canavesi, E., Lattuada, E., Roviaro, G., Marchesini, G. and Fargion, S. (2010c) 'HFE genotype, parenchymal iron accumulation, and liver fibrosis in patients with nonalcoholic fatty liver disease', *Gastroenterology*, 138(3), pp. 905-12.

Valenti, L., Fracanzani, A.L., Dongiovanni, P., Santorelli, G., Branchi, A., Taioli, E., Fiorelli, G. and Fargion, S. (2002) 'Tumor necrosis factor alpha promoter polymorphisms and insulin resistance in nonalcoholic fatty liver disease', *Gastroenterology*, 122(2), pp. 274-80.

Valenti, L., Motta, B.M., Alisi, A., Sartorelli, R., Buonaiuto, G., Dongiovanni, P., Rametta, R., Pelusi, S., Fargion, S. and Nobili, V. (2012) 'LPIN1 rs13412852 polymorphism in pediatric nonalcoholic fatty liver disease', *J Pediatr Gastroenterol Nutr*, 54(5), pp. 588-93.

Valenti, L., Nobili, V., Al-Serri, A., Rametta, R., Leathart, J.B., Zappa, M.A., Dongiovanni, P., Fracanzani, A.L., Alterio, A., Roviaro, G., Daly, A.K., Fargion, S. and Day, C.P. (2011a) 'The APOC3 T-455C and C-482T promoter region polymorphisms are not associated with the severity of liver damage independently of PNPLA3 I148M genotype in patients with nonalcoholic fatty liver', *J Hepatol*, 55(6), pp. 1409-14.

Valenti, L., Rumi, M., Galmozzi, E., Aghemo, A., Del Menico, B., De Nicola, S., Dongiovanni, P., Maggioni, M., Fracanzani, A.L., Rametta, R., Colombo, M. and Fargion, S. (2011b) 'Patatin-like phospholipase domain-containing 3 I148M

polymorphism, steatosis, and liver damage in chronic hepatitis C', *Hepatology*, 53(3), pp. 791-9.

Valko, M., Izakovic, M., Mazur, M., Rhodes, C.J. and Telser, J. (2004) 'Role of oxygen radicals in DNA damage and cancer incidence', *Mol Cell Biochem*, 266(1-2), pp. 37-56.

Vanni, E., Marengo, A., Mezzabotta, L. and Bugianesi, E. (2015) 'Systemic Complications of Nonalcoholic Fatty Liver Disease: When the Liver Is Not an Innocent Bystander', *Semin Liver Dis*, 35(3), pp. 236-49.

Vazquez-Chantada, M., Gonzalez-Lahera, A., Martinez-Arranz, I., Garcia-Monzon, C., Regueiro, M.M., Garcia-Rodriguez, J.L., Schlangen, K.A., Mendibil, I., Rodriguez-Ezpeleta, N., Lozano, J.J., Banasik, K., Justesen, J.M., Joergensen, T., Witte, D.R., Lauritzen, T., Hansen, T., Pedersen, O., Veyrie, N., Clement, K., Tordjman, J., Tran, A., Le Marchand-Brustel, Y., Buque, X., Aspichueta, P., Echevarria-Uraga, J.J., Martin-Duce, A., Caballeria, J., Gual, P., Castro, A., Mato, J.M., Martinez-Chantar, M.L. and Aransay, A.M. (2013) 'Solute carrier family 2 member 1 is involved in the development of nonalcoholic fatty liver disease', *Hepatology*, 57(2), pp. 505-14.

Velaga, M.R., Wilson, V., Jennings, C.E., Owen, C.J., Herington, S., Donaldson, P.T., Ball, S.G., James, R.A., Quinton, R., Perros, P. and Pearce, S.H. (2004) 'The codon 620 tryptophan allele of the lymphoid tyrosine phosphatase (LYP) gene is a major determinant of Graves' disease', *J Clin Endocrinol Metab*, 89(11), pp. 5862-5.

Verrijken, A., Beckers, S., Francque, S., Hilden, H., Caron, S., Zegers, D., Ruppert, M., Hubens, G., Van Marck, E., Michielsen, P., Staels, B., Taskinen, M.R., Van Hul, W. and Van Gaal, L. (2013) 'A gene variant of PNPLA3, but not of APOC3, is associated with histological parameters of NAFLD in an obese population', *Obesity (Silver Spring)*, 21(10), pp. 2138-45.

Vespasiani-Gentilucci, U., Gallo, P., Porcari, A., Carotti, S., Galati, G., Piccioni, L., De Vincentis, A., Dell'Unto, C., Vorini, F., Morini, S., Riva, E. and Picardi, A. (2016) 'The PNPLA3 rs738409 C > G polymorphism is associated with the risk of progression to cirrhosis in NAFLD patients', *Scand J Gastroenterol*, 51(8), pp. 967-73.

Victor, R.G., Haley, R.W., Willett, D.L., Peshock, R.M., Vaeth, P.C., Leonard, D., Basit, M., Cooper, R.S., Iannacchione, V.G., Visscher, W.A., Staab, J.M. and Hobbs, H.H. (2004) 'The Dallas Heart Study: a population-based probability sample for the multidisciplinary study of ethnic differences in cardiovascular health', *Am J Cardiol*, 93(12), pp. 1473-80.

Videla, L.A., Rodrigo, R., Araya, J. and Poniachik, J. (2004) 'Oxidative stress and depletion of hepatic long-chain polyunsaturated fatty acids may contribute to nonalcoholic fatty liver disease', *Free Radic Biol Med*, 37(9), pp. 1499-507.

Vogel, S., Piantedosi, R., Frank, J., Lalazar, A., Rockey, D.C., Friedman, S.L. and Blaner, W.S. (2000) 'An immortalized rat liver stellate cell line (HSC-T6): a new cell model for the study of retinoid metabolism in vitro', *J Lipid Res*, 41(6), pp. 882-93.

Walenbergh, S.M., Koek, G.H., Bieghs, V. and Shiri-Sverdlov, R. (2013) 'Non-alcoholic steatohepatitis: the role of oxidized low-density lipoproteins', *J Hepatol*, 58(4), pp. 801-10.

Wang, J., Leclercq, I., Brymora, J.M., Xu, N., Ramezani-Moghadam, M., London, R.M., Brigstock, D. and George, J. (2009) 'Kupffer cells mediate leptin-induced liver fibrosis', *Gastroenterology*, 137(2), pp. 713-23.

Wang, X., Liu, Z., Peng, Z. and Liu, W. (2015) 'The TM6SF2 rs58542926 T allele is significantly associated with non-alcoholic fatty liver disease in Chinese', *J Hepatol*, 62(6), pp. 1438-9.

Wang, X., Liu, Z., Wang, K., Wang, Z., Sun, X., Zhong, L., Deng, G., Song, G., Sun, B., Peng, Z. and Liu, W. (2016) 'Additive Effects of the Risk Alleles of PNPLA3 and TM6SF2 on Non-alcoholic Fatty Liver Disease (NAFLD) in a Chinese Population', *Front Genet*, 7, p. 140.

Wang, Z., Xu, M., Peng, J., Jiang, L., Hu, Z., Wang, H., Zhou, S., Zhou, R., Hultstrom, M. and Lai, E.Y. (2013) 'Prevalence and associated metabolic factors of fatty liver disease in the elderly', *Exp Gerontol*, 48(8), pp. 705-9.

Wei, Y., Wang, D., Topczewski, F. and Pagliassotti, M.J. (2006) 'Saturated fatty acids induce endoplasmic reticulum stress and apoptosis independently of ceramide in liver cells', *Am J Physiol Endocrinol Metab*, 291(2), pp. E275-81.

Wellcome Trust Case Control, C. (2007) 'Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls', *Nature*, 447(7145), pp. 661-78.

Welzel, T.M., Graubard, B.I., Zeuzem, S., El-Serag, H.B., Davila, J.A. and McGlynn, K.A. (2011) 'Metabolic syndrome increases the risk of primary liver cancer in the United States: a study in the SEER-Medicare database', *Hepatology*, 54(2), pp. 463-71.

White, D.L., Kanwal, F. and El-Serag, H.B. (2012) 'Association between nonalcoholic fatty liver disease and risk for hepatocellular cancer, based on systematic review', *Clin Gastroenterol Hepatol*, 10(12), pp. 1342-1359 e2.

Wieckowska, A. and Feldstein, A.E. (2008) 'Diagnosis of nonalcoholic fatty liver disease: invasive versus noninvasive', *Semin Liver Dis*, 28(4), pp. 386-95.

Wiedmann, S., Fischer, M., Koehler, M., Neureuther, K., Riegger, G., Doering, A., Schunkert, H., Hengstenberg, C. and Baessler, A. (2008) 'Genetic variants within the LPIN1 gene, encoding lipin, are influencing phenotypes of the metabolic syndrome in humans', *Diabetes*, 57(1), pp. 209-17.

Williams, C.D., Stengel, J., Asike, M.I., Torres, D.M., Shaw, J., Contreras, M., Landt, C.L. and Harrison, S.A. (2011) 'Prevalence of nonalcoholic fatty liver disease and nonalcoholic steatohepatitis among a largely middle-aged population utilizing

ultrasound and liver biopsy: a prospective study', *Gastroenterology*, 140(1), pp. 124-31.

Williamson, R.M., Price, J.F., Glancy, S., Perry, E., Nee, L.D., Hayes, P.C., Frier, B.M., Van Look, L.A., Johnston, G.I., Reynolds, R.M., Strachan, M.W. and Edinburgh Type 2 Diabetes Study, I. (2011) 'Prevalence of and risk factors for hepatic steatosis and nonalcoholic Fatty liver disease in people with type 2 diabetes: the Edinburgh Type 2 Diabetes Study', *Diabetes Care*, 34(5), pp. 1139-44.

Willner, I.R., Waters, B., Patil, S.R., Reuben, A., Morelli, J. and Riely, C.A. (2001) 'Ninety patients with nonalcoholic steatohepatitis: insulin resistance, familial tendency, and severity of disease', *Am J Gastroenterol*, 96(10), pp. 2957-61.

Wilson, P.A., Gardner, S.D., Lambie, N.M., Commans, S.A. and Crowther, D.J. (2006) 'Characterization of the human patatin-like phospholipase family', *J Lipid Res*, 47(9), pp. 1940-9.

Witte, J.S. (2010) 'Genome-wide association studies and beyond', *Annu Rev Public Health*, 31, pp. 9-20 4 p following 20.

Witte, J.S., Visscher, P.M. and Wray, N.R. (2014) 'The contribution of genetic variants to disease depends on the ruler', *Nat Rev Genet*, 15(11), pp. 765-76.

Wong, V.W., Wong, G.L., Choi, P.C., Chan, A.W., Li, M.K., Chan, H.Y., Chim, A.M., Yu, J., Sung, J.J. and Chan, H.L. (2010) 'Disease progression of non-alcoholic fatty liver disease: a prospective study with paired liver biopsies at 3 years', *Gut*, 59(7), pp. 969-74.

Wong, V.W., Wong, G.L., Tse, C.H. and Chan, H.L. (2014) 'Prevalence of the TM6SF2 variant and non-alcoholic fatty liver disease in Chinese', *J Hepatol*, 61(3), pp. 708-9.

Wong, V.W., Wong, G.L., Yip, G.W., Lo, A.O., Limquiaco, J., Chu, W.C., Chim, A.M., Yu, C.M., Yu, J., Chan, F.K., Sung, J.J. and Chan, H.L. (2011) 'Coronary artery disease and cardiovascular outcomes in patients with non-alcoholic fatty liver disease', *Gut*, 60(12), pp. 1721-7.

Wortham, M., He, L., Gyamfi, M., Copple, B.L. and Wan, Y.J. (2008) 'The transition from fatty liver to NASH associates with SAMe depletion in db/db mice fed a methionine choline-deficient diet', *Dig Dis Sci*, 53(10), pp. 2761-74.

Wruck, W., Graffmann, N., Kawala, M.A. and Adjaye, J. (2017) 'Concise Review: Current Status and Future Directions on Research Related to Nonalcoholic Fatty Liver Disease', *Stem Cells*, 35(1), pp. 89-96.

Xu, L., Hui, A.Y., Albanis, E., Arthur, M.J., O'Byrne, S.M., Blaner, W.S., Mukherjee, P., Friedman, S.L. and Eng, F.J. (2005) 'Human hepatic stellate cell lines, LX-1 and LX-2: new tools for analysis of hepatic fibrosis', *Gut*, 54(1), pp. 142-51.

Yamada, T., Fukatsu, M., Suzuki, S., Wada, T., Yoshida, T. and Joh, T. (2010) 'Fatty liver predicts impaired fasting glucose and type 2 diabetes mellitus in Japanese undergoing a health checkup', *J Gastroenterol Hepatol*, 25(2), pp. 352-6.

Yamaguchi, K., Yang, L., McCall, S., Huang, J., Yu, X.X., Pandey, S.K., Bhanot, S., Monia, B.P., Li, Y.X. and Diehl, A.M. (2007) 'Inhibiting triglyceride synthesis improves hepatic steatosis but exacerbates liver damage and fibrosis in obese mice with nonalcoholic steatohepatitis', *Hepatology*, 45(6), pp. 1366-74.

Yamakawa-Kobayashi, K., Ishiguro, H., Arinami, T., Miyazaki, R. and Hamaguchi, H. (2002) 'A Val227Ala polymorphism in the peroxisome proliferator activated receptor alpha (PPARalpha) gene is associated with variations in serum lipid levels', *J Med Genet*, 39(3), pp. 189-91.

Yamashita, H., Takenoshita, M., Sakurai, M., Bruick, R.K., Henzel, W.J., Shillinglaw, W., Arnot, D. and Uyeda, K. (2001) 'A glucose-responsive transcription factor that regulates carbohydrate metabolism in the liver', *Proc Natl Acad Sci U S A*, 98(16), pp. 9116-21.

Yang, L., Li, P., Fu, S., Calay, E.S. and Hotamisligil, G.S. (2010) 'Defective hepatic autophagy in obesity promotes ER stress and causes insulin resistance', *Cell Metab*, 11(6), pp. 467-78.

Yasui, K., Sumida, Y., Mori, Y., Mitsuyoshi, H., Minami, M., Itoh, Y., Kanemasa, K., Matsubara, H., Okanoue, T. and Yoshikawa, T. (2011) 'Nonalcoholic steatohepatitis and increased risk of chronic kidney disease', *Metabolism*, 60(5), pp. 735-9.

Yatsuji, S., Hashimoto, E., Tobari, M., Taniai, M., Tokushige, K. and Shiratori, K. (2009) 'Clinical features and outcomes of cirrhosis due to non-alcoholic steatohepatitis compared with cirrhosis caused by chronic hepatitis C', *J Gastroenterol Hepatol*, 24(2), pp. 248-54.

Yeh, M.M. and Brunt, E.M. (2014) 'Pathological features of fatty liver disease', *Gastroenterology*, 147(4), pp. 754-64.

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

Yoneda, M., Hotta, K., Nozaki, Y., Endo, H., Uchiyama, T., Mawatari, H., Iida, H., Kato, S., Fujita, K., Takahashi, H., Kirikoshi, H., Kobayashi, N., Inamori, M., Abe, Y., Kubota, K., Saito, S., Maeyama, S., Wada, K. and Nakajima, A. (2009) 'Association between angiotensin II type 1 receptor polymorphisms and the occurrence of nonalcoholic fatty liver disease', *Liver Int*, 29(7), pp. 1078-85.

Younossi, Z.M., Otgonsuren, M., Henry, L., Venkatesan, C., Mishra, A., Erario, M. and Hunt, S. (2015) 'Association of nonalcoholic fatty liver disease (NAFLD) with hepatocellular carcinoma (HCC) in the United States from 2004 to 2009', *Hepatology*, 62(6), pp. 1723-30.

Younossi, Z.M., Stepanova, M., Rafiq, N., Makhlouf, H., Younoszai, Z., Agrawal, R. and Goodman, Z. (2011) 'Pathologic criteria for nonalcoholic steatohepatitis: interprotocol agreement and ability to predict liver-related mortality', *Hepatology*, 53(6), pp. 1874-82.

Yuan, X., Waterworth, D., Perry, J.R., Lim, N., Song, K., Chambers, J.C., Zhang, W., Vollenweider, P., Stirnadel, H., Johnson, T., Bergmann, S., Beckmann, N.D., Li, Y., Ferrucci, L., Melzer, D., Hernandez, D., Singleton, A., Scott, J., Elliott, P., Waeber, G., Cardon, L., Frayling, T.M., Kooner, J.S. and Mooser, V. (2008) 'Population-based genome-wide association studies reveal six loci influencing plasma levels of liver enzymes', *Am J Hum Genet*, 83(4), pp. 520-8.

Zain, S.M., Mohamed, Z., Mahadeva, S., Rampal, S., Basu, R.C., Cheah, P.L., Salim, A. and Mohamed, R. (2013) 'Susceptibility and gene interaction study of the angiotensin II type 1 receptor (AGTR1) gene polymorphisms with non-alcoholic fatty liver disease in a multi-ethnic population', *PLoS One*, 8(3), p. e58538.

Zhang, B., Xie, W. and Krasowski, M.D. (2008) 'PXR: a xenobiotic receptor of diverse function implicated in pharmacogenetics', *Pharmacogenomics*, 9(11), pp. 1695-709.

Zhou, G., Myers, R., Li, Y., Chen, Y., Shen, X., Fenyk-Melody, J., Wu, M., Ventre, J., Doebber, T., Fujii, N., Musi, N., Hirshman, M.F., Goodyear, L.J. and Moller, D.E. (2001) 'Role of AMP-activated protein kinase in mechanism of metformin action', *J Clin Invest*, 108(8), pp. 1167-74.

Zhou, J., Zhai, Y., Mu, Y., Gong, H., Uppal, H., Toma, D., Ren, S., Evans, R.M. and Xie, W. (2006) 'A novel pregnane X receptor-mediated and sterol regulatory element-binding protein-independent lipogenic pathway', *J Biol Chem*, 281(21), pp. 15013-20.

Zhou, Y., Llaurado, G., Oresic, M., Hyotylainen, T., Orho-Melander, M. and Yki-Jarvinen, H. (2015) 'Circulating triacylglycerol signatures and insulin sensitivity in NAFLD associated with the E167K variant in TM6SF2', *J Hepatol*, 62(3), pp. 657-63.

Zhou, Y.J., Li, Y.Y., Nie, Y.Q., Huang, C.M. and Cao, C.Y. (2012) 'Natural course of nonalcoholic fatty liver disease in southern China: a prospective cohort study', *J Dig Dis*, 13(3), pp. 153-60.

Zhu, J.Z., Dai, Y.N., Wang, Y.M., Zhou, Q.Y., Yu, C.H. and Li, Y.M. (2015) 'Prevalence of Nonalcoholic Fatty Liver Disease and Economy', *Dig Dis Sci*, 60(11), pp. 3194-202.

Zimmermann, R., Strauss, J.G., Haemmerle, G., Schoiswohl, G., Birner-Gruenberger, R., Riederer, M., Lass, A., Neuberger, G., Eisenhaber, F., Hermetter, A. and Zechner, R. (2004) 'Fat mobilization in adipose tissue is promoted by adipose triglyceride lipase', *Science*, 306(5700), pp. 1383-6.