Genetic factors affecting progression of nonalcoholic fatty liver disease

Ahmad Al-Serri

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

Institute of Cellular Medicine

Newcastle University

October 2011

Abstract

Non-alcoholic fatty liver disease (NAFLD) ranges from steatosis to the more progressive form non-alcoholic steatohepatitis. Genetic factors may be important risk determinants for disease progression. This study aimed to assess association of polymorphisms in candidate genes with NAFLD severity and to investigate functional significance of selected polymorphisms

Two approaches were used for association studies, case-control analysis on adults (n=351) with biopsy-proven NAFLD and transmission disequilibrium on family trios (n=71) with an index child with NAFLD. A total of 37 polymorphisms in 14 candidate genes selected on the basis of either biological relevance or previous data suggesting a role in NAFLD were genotyped. Significant differences were seen for polymorphisms in 4 genes between stages of NAFLD or in transmission within families. For *SOD2*, which encodes manganese-dependent superoxide dismutase, the homozygous T genotype for rs4880 was more common in severe fibrosis (OR 2.23 (95% CI 1.19-4.17); p=0.01) and the T-allele was preferentially transmitted in the family trios (p=0.038). For the adiponutrin gene (*PNPLA3)*, carriage of the variant G-allele (rs738409) was associated with severe steatosis (OR 1.87 (95% CI 1.15-3.04); p=0.01) and severe fibrosis (OR 2.02 (95% CI 1.29-3.1); p=0.002) in adults and preferentially transmitted in the family trios (p=0.001). For the claudin-10 gene (*CLDN10*), carriage of rs4143093 was associated with severe steatosis (OR 2.82 (95% CI 1.5-5.3); p=0.0009). To further assess the relevance of claudin-10 to NAFLD, immunohistochemistry was performed. Expression in liver sections was confirmed. The effect of rs4143093 genotype on expression was investigated but insufficient samples were available to reach a firm conclusion. For the gamma-glutamyl cysteine ligase catalytic subunit gene (*GCLC*), rs17883901 was preferentially transmitted in the family trios (p= 0.046), but did not affect disease severity in adults. Studies on functional significance of this polymorphism showed that no significant difference in promoter activity between allelic variants.

Acknowledgements

I would like to show my gratitude to my supervisor Professor Ann K. Daly for giving me the opportunity to work in her lab and for her full guidance and support throughout this journey. I am grateful for the support and help of my colleagues: Julian Leathart, Julia Patch, Elise Andrews, Pallav Bhatnagar, Ching Ng, Mohammad Alshabeeb and Tom Chamberlain. I would also like to thank Professor Christopher Day, Professor Heather Cordell and Professor Helen Reeves.

This thesis would not have been possible without the support of my family. I thank my wonderful parents, my brother and my sister for their love, encouragement and many years of support. I owe my deepest gratitude to my wife for her endless love, patience and understanding throughout this time.

Last but not least, I thank God who made all things possible and for giving me the will and patience to achieve my dreams.

The work described in this thesis, except where specifically stated otherwise, is entirely my own.

Ahmad Al-Serri

Table of Contents

Publications

Valenti L, **Al-Serri A**, Daly AK, Galmozzi E, Rametta R, Dongiovanni P, Nobili V, Mozzi E, Roviaro G, Vanni E, Bugianesi E, Maggioni M, Fracanzani AL, Fargion S, Day CP. Homozygosity for the Patatin-Like Phospholipase-3/Adiponutrin I148M Polymorphism Influences Liver Fibrosis in Patients with Nonalcoholic Fatty Liver Disease. *Hepatology* 2010, **51**(4), 1209-1217.

Valenti L, Nobili V, **Al-Serri A**, Rametta R, Leathart JBS, Zappa MA, Dongiovanni P, Fracanzani AL, Alterio A, Roviaro G, Daly AK, Fargion S, Day CP. 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. *Journal of Hepatology* 2011. In Press.

Al-Serri A, Anstee QM, Valenti L, Nobili V, Leathart JBS, Dongiovanni P, Patch J, Fracanzani A, Fargion S, Day CP, Daly AK. The SOD2 C47T Polymorphism Influences NAFLD Fibrosis Severity: Evidence From Case-Control and Intra-Familial Allele Association Studies. *Journal of Hepatology* 2011. In Press.

Abbreviations

µ	Micro
4-HNE	4-hydroxytransnonenal
ACC	acetyl-CoA carboxylase
ALD	Alcoholic liver disease
ALT	Alanine transaminase
APOC3	apolipoprotein C-3
ARE	antioxidant response element
ATP	Adenosine triphosphate
bp	Base pair
BSA	Bovine serum albumin
cAMP	cyclic adenosine monophosphate
CEU	Caucasians
CI	Confidence interval
DMSO	Dimethylsulphoxide
ECM	Extracellular matrix
EMEM	Eagle's minimal essential mediums
ENPP1	ectoenzyme nucleotide pyrophosphate phosphodiesterase 1
FA	Fatty acid
FAS	fatty acid synthetase
FFA	Free fatty acid
GCL	glutamate cysteine ligase
GCLC	glutamate cysteine ligase catalyze
GCLM	glutamate cysteine ligase modifier
GLUT	glucose transporters
GPX	glutathione peroxidise
GS	glutathione synthetase
GSH	Glutathione
GSR	glutathione reductase
GWAS	genome-wide association study
H_2O_2	hydrogen peroxide
HCC	Hepatocellular carcinoma
Het	Heterozygous
HGP	human genome project
HMOX1	haem oxygenase 1
IHC	Immunohistochemistry
IRS	insulin receptor substrates

Keap1	Kelch-like ECH associating protein 1
KLF	Kruppel-like factor
L	Litter
LD	linkage disequilibrium
leu	leucine
M	Moles
MAT	methionine adenosyl transferase
Met	Mutant
MnSOD	Mangenese-dependent superoxide dismutase
MTP	microsomal triglyceride transfer protein
MTS	mitochondrial targeting sequence
NAFLD	Non-alcoholic fatty liver disease
NASH	Non-alcoholic steatohepatitis
NEFA	non-esterified fatty acid
NHANES	National health and nutritional examination survey
Nrf1	Nuclear respiratory factor 1
Nrf2	NF-E2-related factor 2
O_2^-	superoxide anions
OH*	hydroxyl radicals
OR	Odds ratio
p	p-value
PBS	Phosphate buffer saline
PCR	Polymerase chain reaction
PEMT	Phosphatidylethanolamine methyltransferase
PNPLA3	patatin-like phospholipase 3
pro	proline
RFLP	Restriction fragment length polymorphism
ROS	Reactive oxygen species
SAMe	S-adenosylmethionine
SNP	Single nucleotide polymorphism
SOD	superoxide dismutase
SREBP-1c	sterol regulatory element-binding protein-1c
TBARS	Thiobarbituric acid reactive substances
TBE	Tris-borate/EDTA electrophoresis buffer
TBS	Tris buffer saline
TDT	Transmission disequilibrium test
TEMED	N, N, N', N-tetramethyl-ethylenediamine
TG	Triglyceride

TNR	trinucleotide repeat
Tris	Tris (hydroxymethyl) methylamine: 2 amino-2-hydroxymethyl) propane-1,3-idol
TSC	the SNP consortium
UGT1A1	UDP-glucuronosyltransferase 1A1
VLDL	very low density lipoproteins
WT	Wild type
X^2	Chi-square
ɣ-GC	ɣ-glutamylcysteine

List of figures

List of Tables

Chapter 1

Introduction

1 Introduction

1.1 NAFLD

Non-alcoholic fatty liver disease (NAFLD), the hepatic manifestation of the metabolic syndrome, is the most common cause of chronic liver disease in western countries. It is estimated that 1 in 3 Western adults now have NAFLD [Day, 2010]. The term NAFLD encompasses a spectrum of liver diseases, ranging from the initiation stage steatosis to its more progressive form non-alcoholic steatohepatitis (NASH). NASH can then lead to advanced fibrosis, cirrhosis, liver failure and hepatocellular carcinoma (HCC) (Figure 1.1) [Day, 2006c]. NASH, the progressive form of NAFLD was described more that 30 years ago when Adler and colleague in 1979 studied 29 cases and found a resemblance between patients with fatty liver that were non-alcoholics to alcoholic steatohepatitis. The term NASH was then coined by Ludwig in 1980 [Adler and Schaffner, 1979;Ludwig et al., 1980]. Because of the different diagnostic methods used for NAFLD the true population prevalence is unknown. However, prevalence of NAFLD and NASH are estimated to be 20-30% and 2-3% respectively in the general population [Matteoni et al., 1999;Bellentani et al., 2010;Moore, 2010]. In a meta-analysis study by Musso and colleagues, NAFLD was found associated with increased mortality compared with the general population. NASH also had a higher liver-related mortality compared to steatosis and increased with advance fibrosis [Musso et al., 2010]. The national health and nutritional examination survey (NHANES) conducted a study between 1988-2008 and results showed an increase in the prevalence of NAFLD compared to other chronic liver diseases which remained stable [Younossi et al., 2011]. NAFLD is now also considered to be the most common cause of liver disease in children due to the increase in childhood obesity. A study between 1974-2003 in the UK revealed an increase in prevalence of childhood obesity from 1.8% to 6% in boys and 1.3% to 6.6% in girls aged 5-10 years [Stamatakis et al., 2005]. NAFLD is found in all ethnic groups, however studies have shown Hispanics to have the highest frequency compared to Caucasians and Africans (45%, 33%, and 24% respectively) [Browning et al., 2004].

1.2 Diagnosis

The lack of consistency in diagnostic tests to identify NAFLD and its different stages has been a major issue in the world of NAFLD. Liver biopsy is considered to be the golden method for diagnosis of NAFLD; however its invasiveness has lead to cases of

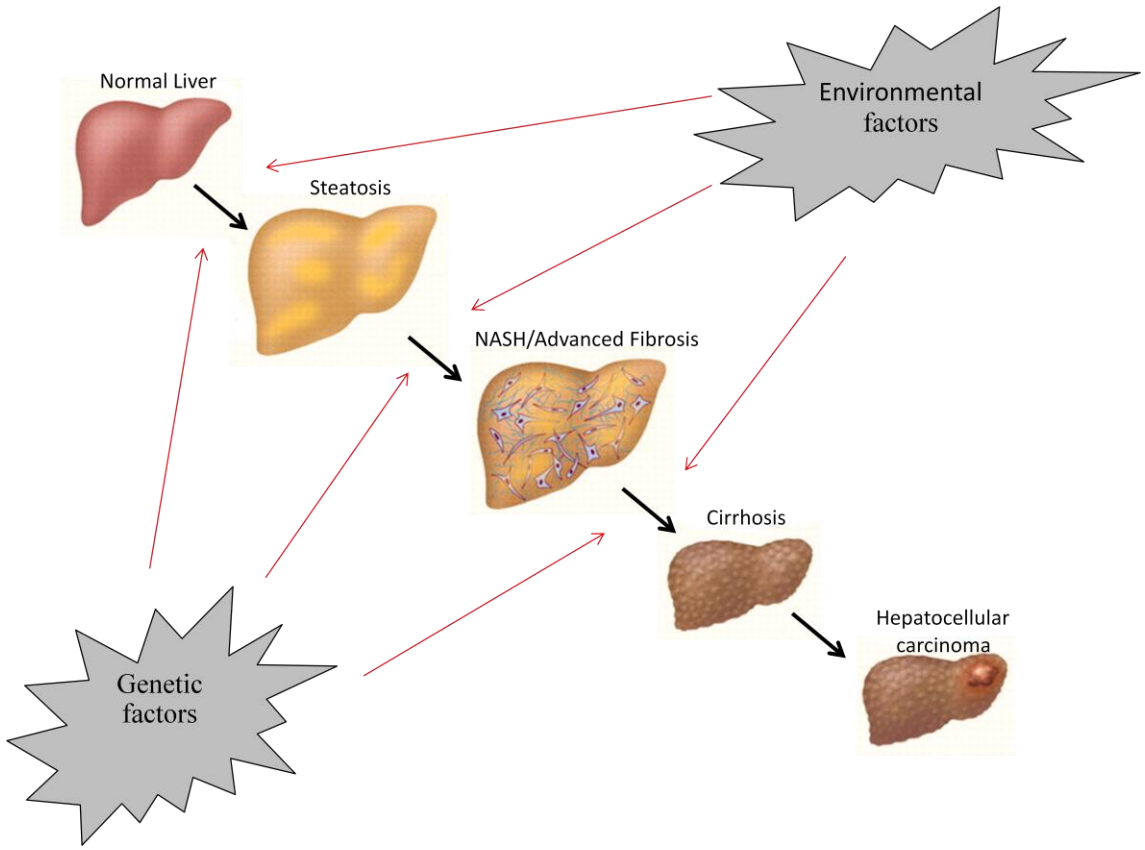

Figure 1.1: Diagram representing the stages of NAFLD and the factors affecting its progression. Adapted and modified from [Cohen et al., 2011]

morbidity and mortality [Cadranel et al., 2000;Gilmore et al., 1995]. Other methods such as proton magnetic resonance spectroscopy, ultrasound, MRI and computed tomography are non-invasive. However these non-invasive methods do have their advantages and disadvantages as reviewed by [Schwenzer et al., 2009]. Although methods are able to detect steatosis they lack the ability to differentiate between the different stages [Angulo, 2007]. Measurements of liver enzymes such as alanine transaminase (ALT) have been used quite often and high levels are a sign for further diagnostic tests, although sometimes relying on this can be misleading. A study by Browning and colleagues revealed that 79% of subjects with hepatic steatosis measured by proton magnetic resonance spectroscopy had normal levels of ALT [Browning et al., 2004]. The combined biochemical and clinical parameters have been used for fibrosis prediction [Hossain et al., 2009;Neuschwander-Tetri et al., 2010;Lebovics and Rubin, 2011]. Other non-invasive tests such as FibroTest and FibroScan have been used in NAFLD however this goes outside the scope of this thesis for review [Pais et al., 2009].

1.3 Stages of NAFLD

1.3.1 Steatosis

Steatosis, the hallmark feature and initiation stage of NAFLD, is suggested to occur from the excess accumulation (more than 5%) of lipids, mainly triglycerides (TG), although other lipids such as cholesterol esters and sphingolipids also exist within hepatocytes [Bircher, 1999b] [Kleiner et al., 2005]. Increased fat in the hepatocytes may be divided into macrovesicular (large lipid droplets) and microvesicular (small lipid droplets) steatosis. Macrovesicular results from imbalance of hepatic TG synthesis whereas microvesicular is a result of mitochondrial dysfunction [Bacon and O'Grady, 2006]. It is thought that fat initiates as small droplets (microvesicular) and eventually lead to large droplets (macrovesicular) [Bacon and O'Grady, 2006]. Excess accumulation of lipids is a sign of abnormal synthesis and disposal of TG (discussed 1.7.1). Having steatosis is considered to be benign, however a follow-up clinical study on patients with fatty liver by Matteoni and colleagues showed that 4% develop cirrhosis and 2% liver related death [Matteoni et al., 1999]. A scoring system has been developed by Kleiner and colleagues that defines the stages (0-3) of steatosis. Presence of fat between 5-33% is noted as stage one, >33-66% as stage two and above 66% as stage three [Kleiner et al., 2005].

1.3.2 NASH

NASH, the progressive form of NAFLD develops on the background of steatosis and features inflammation, ballooning, Mallory bodies, lobular necrosis, and pericellular and perivenular fibrosis which are changes that differentiate steatosis from NASH [Larter et al., 2010] [Sherlock and Dooley, 2002]. It is estimated that only a quarter of subjects progress from NAFLD to its more severe stages NASH/fibrosis, advanced fibrosis and cirrhosis [Wanless and Lentz, 1990;Dixon et al., 2001]. The progression from NASH/fibrosis to advanced fibrosis occurs over time in NASH patients which then leads to cirrhosis [Starley et al., 2010] [Angulo, 2007].

1.3.3 Fibrosis and cirrhosis

Fibrosis is described as increase in connective tissue in the liver with various patterns such as pericellular or perisinusoidal fibrosis and cirrhosis [Bircher, 1999b]. Fibrosis is a result of response to wound healing liver injury [Albanis and Friedman, 2001]. Excessive accumulation of extracellular matrix (ECM) proteins including collagen results in liver fibrosis [Bataller and Brenner, 2005]. The main ECM producing cells are hepatic satellite (HCS) [Gabele et al., 2003]. NASH, is a main cause of liver fibrosis [Brunt, 2004]. It is estimated that patients with NASH develop fibrosis in over 3.5-5 years and about 20% will develop advanced fibrosis [Starley et al., 2010;Ong and Younossi, 2003;Argo and Caldwell, 2009]. Kleiner and colleagues further scored fibrosis into different stages (0-4), zero having no fibrosis and four, development of cirrhosis [Kleiner et al., 2005]. Cirrhosis, the result of ongoing fibrosis is the end stage of many chronic liver diseases [Bircher, 1999a;Bataller and Brenner, 2005]. Studies show that 5 to 20% of NASH patients develop liver cirrhosis [Matteoni et al., 1999;Adams et al., 2005;Angulo, 2007].

1.3.4 Hepatocellular carcinoma

Hepatocellular carcinoma is the third most common cause of cancer death. The aetiology of 15-50% of HCC is unclear suggesting other risk factors such as NAFLD may be involved [Bugianesi, 2007]. A retrospective study suggested up to 27% of NASH cases may develop HCC [Ratziu et al., 2002;Siegel and Zhu, 2009] while other studies found that 40-62% of patients with NASH-related cirrhosis develop HCC after 5-7 years [Hui et al., 2003;Adams et al., 2005]. Up to 70% of cryptogenic cirrhosis is caused by NASH and this is also known to lead to HCC [Angulo, 2007;Yoneda, 2010].

Identifying NAFLD at an early stage before it progresses to end stage liver disease may allow lifestyle modification to prevent further disease progression. Liver transplantation is suggested to be the only effective therapy for cirrhosis, including that linked to NASH, [Davis et al., 2003] and it has been projected that the leading cause of liver transplantation would be NASH in 2020 [Baumeister et al., 2008;Charlton, 2004].

1.4 Risk factors in NAFLD

Obesity, type 2 diabetes and other components of the metabolic syndrome have all been found to be risk factors for the development of NAFLD and its progressive stages.

1.4.1 NAFLD and the metabolic syndrome

The metabolic syndrome, also known as syndrome X, is a cluster of risk factors that once accumulated increases the risk of developing stroke, coronary artery disease, type 2 diabetes, kidney disease, and NAFLD. The exact definition of the metabolic syndrome has been a controversial topic and definitions vary internationally, as reviewed by [Eckel et al., 2005]. Central obesity, dyslipidaemia, hypertension, hyperinsulinaemia, impaired glucose tolerance and microalbuminuria have all been proposed as risk factors relevant to the metabolic syndrome [Eckel et al., 2005]. Whether obesity and diabetes is a component of the metabolic syndrome or a consequence is also a continuing controversy [Eckel et al., 2005] [Einhorn et al., 2003]. The prevalence of the metabolic syndrome increased 12-41% worldwide between 1988 to 2004 [Deepa et al., 2007]. Abdominal obesity, hypertension and hyperglycemia are the most common risk factors frequently found in individuals with metabolic syndrome [Ervin, 2009]. More than 90% of patients with NAFLD have at least one component of the metabolic syndrome [Marchesini et al., 2003]. NAFLD prevalence increased in patients from 27% in normal fasting blood glucose to 43% in impaired fasting blood glucose to 62% among patients with type 2 diabetes [Jimba et al., 2005].

1.4.2 NAFLD and obesity

Obesity is a major risk factor in the development of NAFLD. In obese individuals the prevalence of NAFLD is increased to 74% and up to 90% in morbidly obese [Angulo, 2002b;Abrams et al., 2004]. It is estimated that 15-30% of morbidly obese individuals will develop NASH [Ratziu et al., 2000;Qureshi and Abrams, 2007]. Farrell and colleagues reviewed that central obesity is strongly associated with NASH and not overall obesity and that 20% of obese subjects develop NASH [Chitturi et al., 2002;Kral et al., 1993;Omagari et al., 2002;Thomas et al., 2005;Farrell and Larter, 2006].

Improvement of NAFLD is seen in obese patients that have undergone bariatric surgery [Angulo, 2006].

1.4.3 NAFLD and diabetes

Like obesity, type 2 diabetes is a risk factor in the development of NAFLD. NAFLD is seen in over 70% in patients with Type 2 diabetes [Cusi, 2009;Targher et al., 2007;Leite et al., 2009]. In a study on 458 Italians with NAFLD, diabetes was found to be an independent predictor factor for severe fibrosis [Fracanzani et al., 2008]. Obesity is also a well documented risk factor for type 2 diabetes [Goran et al., 2003]. NAFLD correlates with both obesity and type 2 diabetes in over 70% of patients [Bjornsson and Angulo, 2007;Angulo, 2002a].

This closely related presence of metabolic syndrome, diabetes and obesity with NAFLD suggests a shared pathogenesis in the development of these disorders. The combined effects of these risk factors increases the risk of developing NAFLD and its progressive forms and therefore avoiding these risk factors could help prevent the development of NAFLD and its progression.

1.5 NAFLD in Children

NAFLD is the leading cause of chronic liver disease in paediatric population as a result of the increase in prevalence of overweight and obesity in children in the past decades [Matthiessen et al., 2008;Ji, 2008]. It has been suggested by Schwimmer and colleagues that NAFLD in children could be represented in two different forms type 1 and type 2 NASH [Schwimmer et al., 2005]. Type 1 NASH is characterised by steatosis, ballooning and perisinusoidal fibrosis which resembles the same pattern seen in adults where as type 2 NASH is characterised by steatosis, portal inflammation, and portal fibrosis without ballooning [Schwimmer et al., 2005;Nobili et al., 2006]. Whether the pattern of initiation of NAFLD in children is distinct or develops into the same pattern seen in adults remains unclear [Loomba et al., 2009].

1.6 Treatments

Weight reduction in association with exercise and balanced nutrition is the gold method treatment for patients with NAFLD. A steady and modest weight lost results in lower ALT levels, reduces steatosis and resolves steatohepatitis and reverses hepatic fibrosis [Kral et al., 2004;Hickman et al., 2004;Ueno et al., 1997;Dixon et al., 2004] whereas a dramatic weight reduction can result in hepatic fibrosis and liver failure [Chitturi et al., 2002;Powell et al., 1990;Caldwell and Hespenheide, 2002;Capron et al., 1982]. A

considerable interest has been towards pharmacological therapies of NASH. Rosigalitazone and pioglitazone, PPAR-γ thiazolidinedione agonists, have been shown to improve ALT levels and liver histology. Pioglitazone improves insulin sensitivity and glucose and lipid metabolism in type 2 diabetics [Miyazaki et al., 2002]. Pioglatizone improves hepatic steatosis, ballooning, and inflammation in patients with NASH when compared with placebo [Belfort et al., 2006]. Rosiglitazone also improved steatosis and level of ALT when compared with placebo. However, weight gain was observed in NASH treatment trials for both drugs [Neuschwander-Tetri et al., 2003;Promrat et al., 2004]. Metformin is another drug involved in improving insulin sensitivity and is used for type 2 diabetes treatment [Knowler et al., 2002] [Yoneda, 2010]. Meta-analysis on metformin showed improvement of ALT and steatosis [Angelico et al., 2007]. Metformin also improved steatosis, liver inflammation and fibrosis in NAFLD patients more effectively than vitamin E which has an antioxidant effect [Bugianesi et al., 2005].

1.7 Pathogenic mechanisms involved in NAFLD

The mechanisms involved in the pathogenesis of NAFLD are not fully understood and are complex. However due to NAFLD being common and associated with other common disorders such as the metabolic syndrome, Type 2 diabetes and obesity, a considerable progress has been made in understanding the mechanisms involved. Alcoholic liver disease (ALD) and NAFLD share similar pathogenesis but are differentiated by the amount of alcohol consumed. Excessive alcohol intake should be excluded to diagnose NAFLD (> 40mg/d in males and >20mg/d in females) [Neuschwander-Tetri and Caldwell, 2003].

In the past decade considerable progress has been made in understanding the pathogenesis involved in the initiation and progression of NAFLD. A “two-hit” theory that spread in the NAFLD world briefly suggested that the initiation of steatosis known as the “first-hit” and the progression to its more severe stages (NASH) known as the “second hit” are caused by different sets of molecular mechanisms and environmental factors [Day and James, 1998;Day, 2002]. Day suggested that some mechanisms may be involved in the first hit as well as the second hit [Day, 2006a]. However with ongoing studies on NAFLD in recent years the initiation and progression of NAFLD is thought to be multi-factorial with combined mechanisms and environmental factors participating in disease initiation and progression. Obesity, type 2 diabetes and components of the metabolic syndrome are common risk factors causing NAFLD and therefore disease-related pathogenesis such as abnormal lipid metabolism, insulin

resistance and oxidative stress are thought to play an integral role in the pathogenesis of NAFLD. The main focus of this thesis is the involvement of oxidative stress however other mechanisms such as lipid metabolism and insulin resistance will be covered briefly.

1.7.1 Oxidative stress in NAFLD

Oxidative stress is a result of an imbalance of reactive oxygen species (ROS) and the production of antioxidants. There have been studies on human and animal models suggesting oxidative stress involvement in the pathogenesis of NAFLD. Sanyal and colleagues demonstrated the involvement of oxidative stress in NASH [Sanyal et al., 2001]. The study used 3-nitrotyrosine an immune-histochemical staining marker for oxidative stress and the results showed elevated levels of staining in NASH biopsies compared to controls and NAFLD. A recent study suggested oxidative stress was higher in NAFLD cases compared to controls after demonstrating elevated serum levels of the oxidative stress markers thiobarbituric acid reactive substances (TBARS) and protein carbonyl (PCC). This study also demonstrated lower levels of serum glutathione in NAFLD patients compared to controls [Narasimhan et al., 2010]. Also, increased oxidation of lipids due to accumulation of TG results in oxidative stress. A study on Sprague-Dawley rats fed with high-fat liquid diet or the standard Lieber-DeCarli diet found that the high-fat fed rats demonstrated oxidative stress and lipid peroxidation on the basis of an increase in 4-hydroxynonenal, a product of lipid peroxidation [Lieber et al., 2004].

1.7.1.1 Reactive oxygen species

Reactive oxygen species are produced from oxygen during normal cell function which include superoxide anions (O_2^-), hydrogen peroxide (H_2O_2), and hydroxyl radicals (OH*) [Fridovich, 1978]. ROS in the liver can be generated by the mitochondria, peroxisomes, microsomes (including cytochromes P450), reduced nicotinamide adenine dinucleotide oxidase, cyclooxygenase, lipoxygenase and xanthine oxidase [Malaguarnera et al., 2009;Koek et al., 2011]. Mitochondria are the main ROS-generating organelle [Poyton et al., 2009]. In NASH patients mitochondrial abnormalities are observed [Sanyal et al., 2001;Pessayre et al., 2001]. The majority of oxygen (about 95%) during mitochondrial oxygen metabolism is reduced to H_2O whereas the remaining 5% shifts away from the main pathway and is converted into O_2^-. O_2^- forms when some electrons undergoing transfer along the respiratory chain react with molecular oxygen leading to production of ROS. Excessive production of these

ROS can interact with and damage macromolecules (DNA and proteins) in the cells [Betteridge, 2000]. In animal models, increased production of ROS has been shown to be linked to liver diseases including NAFLD [Calvisi et al., 2004;Wei et al., 2008].

1.7.1.2 Antioxidants

Antioxidants are involved in protecting organisms from oxidative stress. The availability of antioxidants may arise endogenously or from the diet. Antioxidants can be divided into enzymatic and non-enzymatic antioxidants all of which take part in scavenging ROS [Koek et al., 2011]. The pro-oxidant O_2^- can be dismutated into H_2O_2 by the antioxidant superoxide dismutase (SOD) and then detoxified by glutathione peroxidase (GPx) and catalase into H_2O and O_2, a pathway when imbalanced leads to oxidative stress [Sasaki, 2006]. In this thesis we have been interested in the mechanisms involved in reducing excessive production of ROS specially superoxide anions (Figure 1.2).

Superoxide dismutase

Three forms of SOD exists, SOD1 (Cu/ZnSOD) which contains cooper and zinc and is located in the cytoplasm, SOD2 (MnSOD) which contains manganese and is located in the mitochondrial matrix, and SOD3 (EC-SOD) which is localized extracellularly [McCord and Fridovich, 1969;Weisiger and Fridovich, 1973;Marklund, 1982]. SOD2 is of particular interest in relation to NAFLD due to it being located in the mitochondria (main ROS generator). Previous reports have shown the lack of MnSOD is lethal in mice and mice heterozygous for knockout and showing decreased enzyme activity have an increased susceptibility to oxidative stress and mitochondrial dysfunction [Melov et al., 1998;Macmillan-Crow and Cruthirds, 2001;Li et al., 1995b;Lebovitz et al., 1996].

Glutathione peroxidase

Four forms of GPx exist (GPx1-4) as reviewed by [Arthur, 2000]. GPx1 is the main glutathione peroxidase in the liver [Arthur, 2000]. It plays a role in reducing the H_2O_2 produced by the MnSOD into H_2O leaving the iron stores less H_2O_2 to react with [Sutton et al., 2006]. GPx1 is a selenium dependent enzyme. Adequate selenium is essential for mRNA stability, translation and GPx1 activity and therefore insufficient levels could lead to termination of the protein at the UGA codon by incorporation of

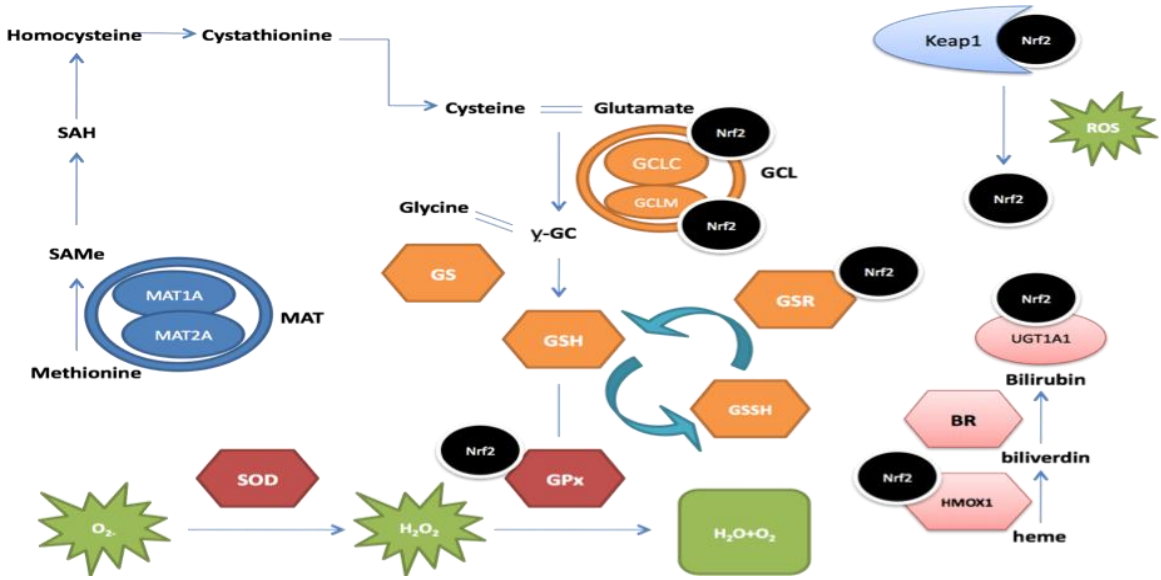

Figure 1.2: Mechanisms involved in balancing antioxidants and ROS. O_{2^-} is dismutatated by SOD into H_2O_2 which is then detoxified by GPX and its cofactor GSH into H_2O+O_2. Activity of GSR and GCL is essential for GSH. MAT plays an important role in the transmethylation and transsulfuration pathways. HMOX1, BR and UGT1A1 play a role in controlling Bilirubin. NEF2/KEAP1 complexes and the role of NRF2 in regulating antioxidants.

selenocysteine resulting in translation termination [Bermano et al., 1996;Papp et al., 2007]. In mouse studies over-expression of GPx1 protects from oxidative stress [Lei, 2001].

Glutathione

Glutathione (GSH) is the most abundant intracellular thiol and is the major antioxidant in the liver [Chen et al., 2007]. GSH is a tri-peptide, composed of glutamate, cysteine and glycine (Figure 1.3). It exists in two forms GSH (thiol-reduced) and GSSG (disulfide oxidized) [DeLeve and Kaplowitz, 1991]. GSH is found in the cell in three sites, cytosol, mitochondria and endoplasmic reticulum [Meredith and Reed, 1982;Hwang et al., 1992]. GSH functions include detoxification of electrophiles, antioxidant defence, cysteine storage and DNA synthesis [Kaplowitz et al., 1985;Meister and Anderson, 1983;Lu, 2009]. However detoxification is the major function of GSH [Lu, 2009]. Cysteine storage is an important function of GSH and availability of cysteine plays a role in controlling the rate of GSH synthesis [Meister, 1988]. Cysteine is derived from diet, protein breakdown and the conversion of methionine to cysteine in the transsulfuration pathway [Lu, 1999].

GSH is synthesised by two ATP-dependent reactions, with the first step involving glutamate cysteine ligase (GCL), the rate-limiting enzyme, which is a heterodimer composed of two subunits that dissociate into a heavy catalytic (GCLC) and a modifier (GCLM) subunits (Figure 1.4) [Meister and Anderson, 1983]. The catalytic activity is possessed by GCLC whereas GCLM possess the regulatory effect when it is associated with GCLC. Glutamate and cysteine are ligated and catalyzed by GCL (ATP dependent) to form y-glutamylcysteine (y-GC) (Figure 1.4). The second reaction is catalyzed by GSH synthetase (GS) which is a homodimeric protein [Dickinson et al., 2004]. GS ligates glycine to y-GC (Figure 1.4). The product of GCL (y-GC) is seen in low concentration levels in cells when GS is present. In humans, low levels of GS results in y-GC being converted to oxoproline which if it accumulates results in metabolic acidosis and GSH deficiency [Dalton et al., 2004]. In *Saccharomyces cerevisiae*, over expression of GCL resulted in an increased GSH level but this was not seen if GS was overexpressed [Grant et al., 1997]. These results are consistent with GCL being the rate-limiting enzyme.

GSH synthesis is determined by GCL activity which is affected by both the GCLC and GCLM subunits [Lu, 2009]. Chen and colleagues disrupted the GCLC gene in

hepatocytes which led to GSH depletion resulting in steatosis and mitochondrial injury [Chen et al., 2007]. GCLM knockout mice result in 10% to 40% of the normal GSH levels [Yang et al., 2002] because GCLM interacts with GCLC to increase its catalytic efficiency [Franklin et al., 2009]. GCLC and GCLM are suggested to being co-ordinately up-regulated however it is still a controversial area where some studies have found them independently regulated [Zheng et al., 2007;Gipp et al., 1995].

GSH content is usually decreased as part of a cellular response to oxidative stress. Such depletion in content must be replaced. Replacement can be achieved by *de novo* synthesis or by reducing GSSG (a product of GPx and GSH to remove peroxides from the cell) to generate GSH by the enzyme glutathione reductase (GSR) which is an essential enzyme maintaining the level of GSH [Kanzok et al., 2001;Satoh et al., 2010;Dickinson et al., 2004]. GSH functions in GPx-catalyzed reactions by acting as a cofactor to detoxify H_2O_2 into H_2O [Sasaki, 2006]. Depletion of GSH can therefore lead to mitochondrial dysfunction and cell death [Fernandez-Checa and Kaplowitz, 2005].

S-adenosylmethionine (SAMe)

SAMe is the precursor of GSH and deficiency in SAMe results in deficiency of GSH [Mato et al., 2002]. SAMe biosynthesis from methionine is catalyzed by methionine adenosyl transferase (MAT) [Mato et al., 2002]. The two genes *MAT1A* and *MAT2A* encode the two MAT catalytic subunits of MAT (alpha1 and alpha2). MAT1A knockout mice are more susceptible to oxidative stress, demonstrated by lower levels of hepatic GSH [Martinez-Chantar et al., 2002]. MAT1A is expressed in the liver and the gene product forms part of two isoforms MATI and MATIII which are the main enzymes for SAMe synthesis in the liver. MATI is a tetramer and MATIII a dimer of the catalytic subunit. MATII is encoded by MAT2A which is widely distributed and expressed in foetal liver as well as adult tissues [Kotb et al., 1997;Horikawa and Tsukada, 1992]. MAT1A knockout studies in mice were consistent with lack of MAT1A predisposing to liver injury due to depletion of hepatic SAMe, changes in expression of methionine metabolism enzymes and expression of growth-related genes resulting in the development of NASH and liver proliferation [Lu et al., 2001]. Depletion of SAMe occurred in wild-type mice fed with methionine choline-deficient diet but supplementation with SAMe resulted in reduced severity of NASH [Oz et al., 2008;Oz et al., 2006]. Depletion of SAMe occurred in db/db mice (a model of type 2 diabetes and the metabolic syndrome harbouring a mutation in the leptin receptor) progressing

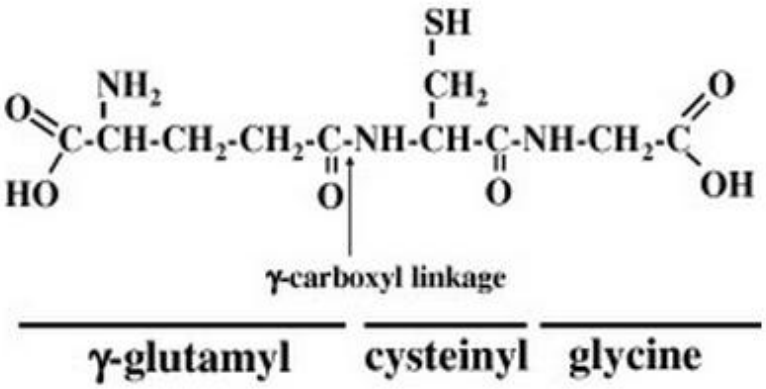

Figure 1.3: The structure of glutathione showing its tri-peptides. adapted from [Lu, 2009]

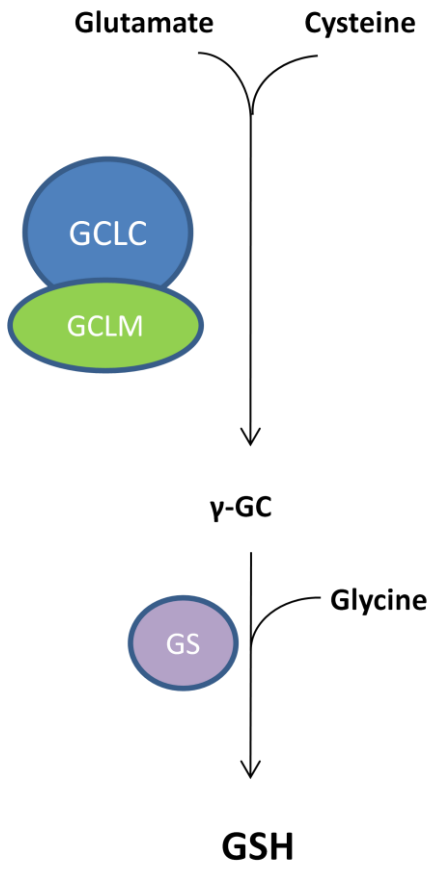

Figure 1.4: The first and second reaction of GSH synthesis demonstrating the role of the GCL subunits GCLC and GCLM. Adapted and modified from [Franklin et al., 2009]

from hepatic steatosis to NASH [Wortham et al., 2008]. Potential relevance of this work to human disease is shown by Mato and colleagues in a study on alcoholic liver cirrhosis patients where treatment with SAMe improved survival [Mato et al., 1999]. SAMe is required for the transmethylation and transsullfuration pathways leading to the formation of cysteine which is required for controlling GSH synthesis.

Other antioxidants

Vitamins (C and E) are free radical scavengers [Valko et al., 2004;van Acker et al., 2000], and bilirubin also function as an antioxidant [Stocker et al., 1987]. Three enzymes control bilirubin concentration, UDP-glucuronosyltransferase 1A1 (UGT1A1), haem oxygenase 1 (HMOX1) and biliverdin reductase [Schwertner and Vitek, 2008]. UGT1A1 modulates levels of serum bilirubin by conjugating glucuronic acid to bilirubin facilitating its removal from the human body [Jansen et al., 1992]. HMOX1 and biliverdin reductase are responsible for the formation of bilirubin [Schwertner and Vitek, 2008]. HMOX1 converts haem to biliverdin which is then converted to bilirubin via biliverdin reductase (Figure 1.1). HMOX1 is induced by oxidative stress and is the rate limiting enzyme for bilirubin synthesis [Vasavda et al., 2007].

Transcription factors regulating antioxidants

Many enzymes contribute to cellular defences against oxidative stress (mentioned above). Several of the key enzymes have an upstream sequence known as antioxidant response element (ARE) which regulates expression of these genes and can increase gene expression when ROS levels are high [Cho et al., 2005]. Two transcription factors known as Nrf1 (Nuclear respiratory factor 1) and Nrf2 (NF-E2-related factor 2) can interact with AREs. Both Nrf1 and Nrf2 are members of the basic leucine zipper CNC (cap 'n' collar) family of proteins. However, in knockout mouse experiments, complete loss of Nrf1 is lethal while animals lacking Nrf2 show increased susceptibility to liver damage when treated with paracetamol and to lung damage with a variety of toxins [Leung et al., 2003;Aleksunes and Manautou, 2007]. Conditional mutants for Nrf1 where Nrf1 is inactivated only in the liver have also been prepared. In these animals, a phenotype of NASH followed by hepatoma development is seen [Xu et al., 2005]. Both Nrf1 and Nrf2 are involved in activating genes regulated by ARE however studies suggest Nrf2 to be the dependent factor for expressing these genes [Biswas and Chan, 2010]. Nrf2 functions in protecting cells from oxidant injury by inducing antioxidant genes at the level of transcription by translocation from cytoplasm to the nucleus and

binding to the ARE. Briefly, in the cytoplasm Nrf2 is normally protected from activation by the bound Kelch-like ECH associating protein 1 (Keap1) which regulates its levels by promoting rapid turnover (Figure 1.5). Once oxidative stress occurs in unstressed cells Keap1 is disrupted allowing Nrf2 to translocate to the nucleus to dimerize with small Maf or AP-1 family of proteins resulting in it binding to the ARE [Itoh et al., 1999;Lee et al., 2005].

GPx2 but not GPx1 was found to be regulated by Nrf2 in mice. Expression of *GPx2* mRNA was increased when small interference RNA (siRNA) was used to knockout the *Nrf2* inhibitor *Keap1* and decreased when siRNA was used for Nrf2 [Singh et al., 2006]. Sequence analysis has shown GCLC promoter contains many cis-acting elements such as several AREs [Dickinson et al., 2004]. Evidence from knockout experiments have shown that Nrf2 regulates GCLC expression [Chan and Kan, 1999]. GCLM also has an ARE and studies have found both Nrf1 and Nrf2 to be involved in inducing and up-regulating GCLM [Zhang et al., 2007;Moinova and Mulcahy, 1999;Kwong et al., 1999]. Harvey and colleagues described the role of Nrf2 in regulating GSR during oxidative stress in Nrf2-defecient mice [Harvey et al., 2009]. UGT1A1 has been found to be regulated by Nrf2. A study by Yueh and Tukey demonstrated the over-expression of Nrf2 when UGT1A1 was induced in HepG2 cells treated with pro-oxidants. Over-expression of Keap1 blocked UGT1A1 expression [Yueh and Tukey, 2007]. Increase in Nrf2 levels results in up-regulation of HMOX1 in HepG2 cells and is prevented by siRNA-Nrf2 [Gong and Cederbaum, 2006]. These studies suggest an important role for Nrf1 and particularly Nrf2 in oxidative stress conditions.

1.7.2 Lipid metabolism involvement in NAFLD

Excess accumulation of TG in hepatocytes is a sign of abnormality in lipid metabolism. There are several pathways involved in lipid metabolism all of which in some way overlap and are reviewed by [Nguyen et al., 2008]. Briefly TG are formed by glycerol and three fatty acid molecules, a process that is influenced by insulin for both energy (ATP) and cell growth [Donnelly et al., 2005]. The liver plays an important role in controlling TG metabolism by maintaining the balance between delivery, synthesising and secreting of FA into or out of the liver. FA supply can be obtained from the diet via lipoproteins, hydrolysis of TG from adipose tissue which releases FFA, hydrolysis of circulating TG, *de novo* lipogenesis, and non-esterified fatty acid (NEFA) [Fabbrini et al., 2010;Moore, 2010].

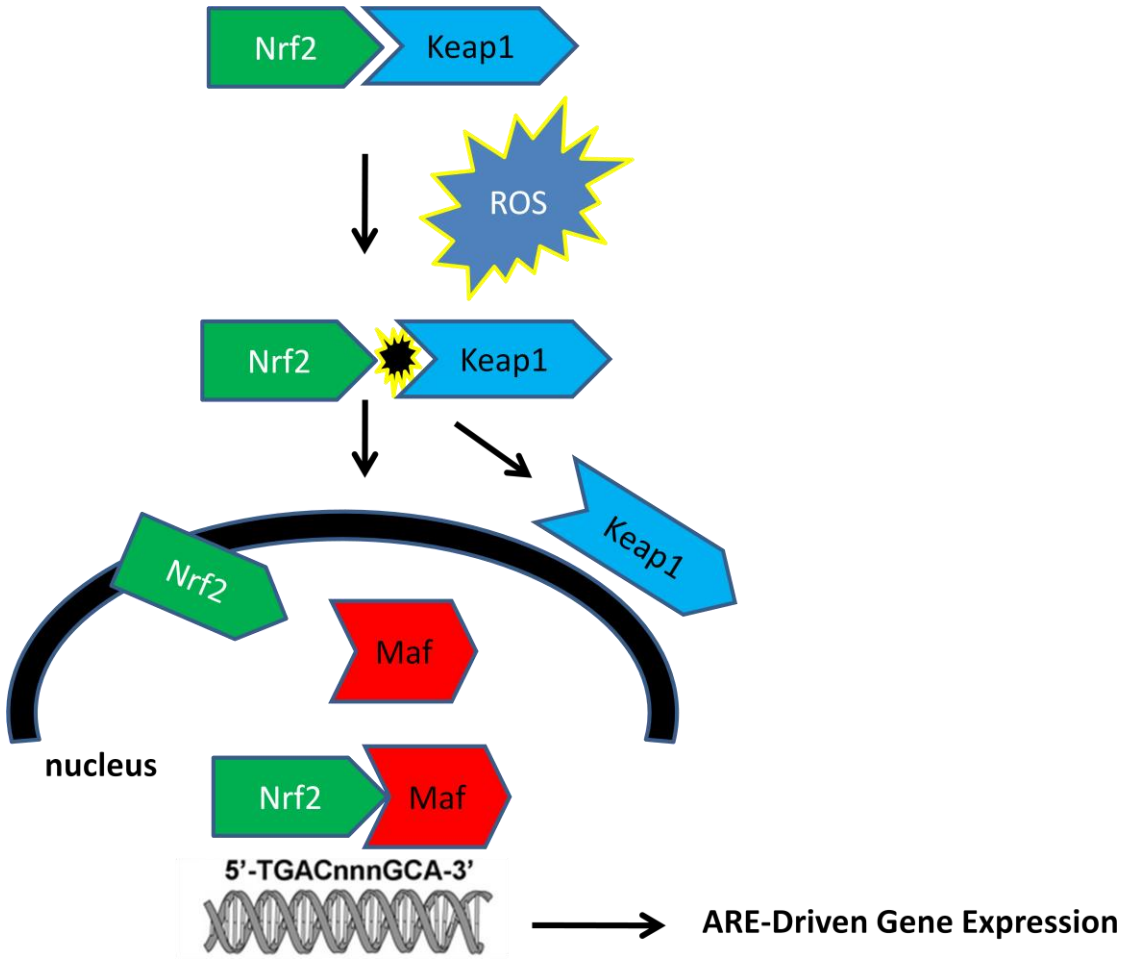

Figure 1.5: Activation of transcription factor Nrf2-Keap1 in the presence of ROS. Adapted and modified from [Lee et al., 2005]

On the other hand FFA can be removed via very low density lipoproteins (VLDL) and mitochondrial beta-oxidation [Hickman et al., 2004;Moore, 2010]. In individuals with NAFLD, there may be a disturbance of TG formation, supply, consumption, oxidation and disposal resulting in lipid accumulation [Schreuder et al., 2008]. Increased FFA is associated with inflammation via production of inflammatory cytokines that lead to Kupffer cell activation which is suggested to be involved in progression from steatosis to NASH [Day, 2006a].

1.7.2.1 FFA supply

TG stored in adipose tissue or present in the circulation can release FFA by hydrolysis via lipoprotein lipase. In obese subjects with NAFLD gene expression of lipoprotein lipase is higher compared in subjects without NAFLD [Pardina et al., 2009;Westerbacka et al., 2007]. Intracellular cyclic adenosine monophosphate (cAMP) levels control the hydrolysis of TG and in the fed state insulin depletes cAMP preventing TG lipolysis (hydrolysis of TG into FFA) [Bacon and O'Grady, 2006]. Donnelly and colleagues demonstrated that NEFAs are the main contributor to hepatic TG, and the elevation of *de novo* lipogenesis in NAFLD suggests its contribution in accumulation of hepatic TG [Adams et al., 2005]. The major regulatory enzymes for *de novo* lipogenesis are acetyl-CoA carboxylase (ACC) and fatty acid synthetase (FAS) [Mendez-Sanchez et al., 2007]. NAFLD has been found to be associated with up-regulation of ACC and FAS and other related genes involved in lipid metabolism [Mitsuyoshi et al., 2009;Kohjima et al., 2007]. Transcription factor sterol regulatory element-binding protein-1c (SREBP-1c) is activated by insulin to regulate *de novo* lipogenesis [Grefhorst et al., 2002;Shimomura et al., 1999]. SREBP-1c over-expression in transgenic mice results in hepatic steatosis [Horton et al., 2002].

1.7.2.2 FFA removal

VLDL is the only disposal mechanism of TG and the rate-limiting step in removal of TG from liver to peripheral stores [Bacon and O'Grady, 2006]. VLDL requires molecules and enzymes for its assembly. The apolipoprotein B-100 (apoB-100) is required for assembly and microsomal triglyceride transfer protein (MTP) is required to incorporate TG [Bacon and O'Grady, 2006;Fabbrini et al., 2010]. Peroxisome proliferators –activated receptors (PPAR-α and γ) are necessary for VLDL secretion [Yoneda, 2010]. Phosphatidylethanolamine methyltransferase (PEMT) involved in the synthesis of phosphatidylcholine which is also needed for VLDL synthesis [Yao and Vance, 1988].

FFA oxidation occurs mainly in mitochondria (β-oxidation), but also by peroxisomes and microsomes to prevent fat accumulation [Hickman et al.;Malaguarnera et al.]. Excessive oxidation can result in reduction in antioxidant defences and therefore result in oxidative stress [Sanyal et al., 2001]. Lipid peroxidation is correlated with FA, therefore an increase in FA results in oxidative stress [Benzie, 1996;Koek et al., 2011]. Malondialdehyde (MDA) and 4-hydroxytransnonenal (4-HNE), both products of lipid peroxidation are seen in 90% of NASH patients serum suggesting oxidative stress has occurred [Loguercio et al., 2001].

1.7.3 Insulin resistance in NAFLD

Insulin resistance is now considered a universal finding in NAFLD. There is accumulating evidence associating insulin resistance with NAFLD/NASH. Sanyal and colleagues demonstrated that non-diabetic patients with hepatic steatosis and NASH had decreased insulin sensitivity while Cortez-Pinto and colleagues showed association of insulin resistance and NASH [Sanyal et al., 2001;Cortez-Pinto et al., 1999]. Pagano and colleagues noted that insulin resistance was seen more in patients with NASH when compared to control subjects matched with age, gender and BMI [Pagano et al., 2002]. Day suggested that insulin resistance may be involved in both the initiation and progression of NAFLD [Day, 2006a]. Insulin resistance is suggested to be involved in hepatic lipid accumulation. However it is still not known whether NAFLD is caused by insulin resistance or whether it is a consequence of it.

Insulin resistance is the inability of a cell to perform its normal cellular function in response to insulin. Insulin plays a major role in glucose and lipid metabolism [Schreuder et al., 2008]. Briefly there are two important insulin receptor substrates (IRS) proteins (IRS-1 and IRS-2) which are tyrosine-phosphorylated when insulin bonds to its receptor [Schreuder et al., 2008]. IRS-1 initiates glucose metabolism and recruits glucose transporters (GLUT) whereas IRS-2 plays a role in lipid metabolism working as a regulator of the *de novo* lipogenesis via SREBP-1c [Schreuder et al., 2008]. SREBP-1c is over-expressed and *de novo* lipogenesis is up-regulated in hepatic insulin resistant states resulting in lipogenesis, caused by down-regulation of the IRS-2 receptor [Schultz et al., 2000]. Larter and colleagues demonstrated the development of steatosis was associated with increase insulin levels and FAS activity (involved in lipogenesis) [Larter et al., 2009]. FFA in hepatocytes may cause insulin resistance through defects in the insulin signalling pathway by activating serine kinases [Hotamisligil, 2005;Perseghin et al., 2003;Shoelson et al., 2006]. This will change the

tyrosine-phosphorylation in IRS-1 to serine-phosphorylation instead resulting in impairment of GLUT transporters [Liu et al., 2010;Schreuder et al., 2008]. Lack of ability to cease hepatic glucose production is a common feature of insulin resistance. Insulin works on inhibiting glycogenolysis and gluconeogenesis resulting in lower glucose production [Rector et al., 2008]. Both insulin and glucose are involved in lipogenesis [Browning and Horton, 2004]. Insulin plays a role in restraining the production of VLDL and in hepatic insulin resistance states in NAFLD patients. TG-rich VLDL are over produced suggesting a role of insulin resistance in lipid metabolism [Vanni et al., 2010]. Insulin resistance therefore results in FFA accumulation in the liver which requires the increase in oxidation that could result in oxidative stress [Cave et al., 2007].

Adiponectin an adipose derived-cytokine, considered as an anti-inflammatory adipokine is found to be in low levels in patients with NAFLD [Targher et al., 2004;Pagano et al., 2005;Pajvani et al., 2003]. Adeponectin is thought to play a role in regulating insulin resistance and lipid metabolism [Tsochatzis et al., 2006;Beltowski, 2003]. It plays a role in reducing TG accumulation in hepatocytes and increasing fatty acid oxidation by down-regulating SREBP-1c and stimulating mitochondrial beta-oxidation by the transport of fatty acids respectively [Rector et al., 2008].

Tumour necrosis factor alpha (TNF-α), an inflammatory mediator which is a candidate in transition from hepatic steatosis to NASH, has been shown to have increased serum levels in obese and insulin resistance subject and elevated levels in NASH when compared with hepatic steatosis [Palmentieri et al., 2006]. TNF-α plays a role in activating inflammatory cells and Kupffer cells causing hepatocyte injury and oxidative stress [Day, 2006b]. Adiponectin and TNF-α are antagonists. Adiponectin has anti-TNF affect which can be a possible therapeutic target for NASH [Cave et al., 2007;Farrell and Larter, 2006].

1.8 Genetics of NAFLD

Genetic factors involvement in NAFLD has recently been reviewed in depth by [Daly et al., 2011]. Identification of genetic factors involved in susceptibility to NAFLD and its progressive stages would allow a non-invasive strategy that could help prevent and control disease initiation and progression.

1.8.1 Genetic component in NAFLD

Family and interethnic studies have suggested a genetic component in NAFLD. Schwimmer and colleagues performed a familial aggregation study to demonstrate the heritability of NAFLD. Two groupshere studied; the first were family members with obese children that had biopsy-proven NAFLD and the second group were family members of obese children without NAFLD. Measures of fat fraction using magnetic resonance imaging were performed on family members. The study found a significant increase in fatty liver in the group with NAFLD children and their family members (59% and 78% respectively) compared to the group with children without NAFLD (17% and 37% respectively) suggesting a familial factor in determining risk of NAFLD [Schwimmer et al., 2009]. A recent study evaluated the correlation between liver fat and levels of ALT between monozygotic and dizygotic twins. The study found that correlation of liver fat and ALT levels were significantly higher in monozygotic twins ($p<0.001$) than dizygotic twins ($p=0.37$) suggesting the involvement of genetic factors [Makkonen et al., 2009]. A study on 8 kindreds (2 with cryptogenic cirrhosis, 4 NASH/cirrhosis and 2 NASH without cirrhosis) and 10 relatives showed the co-existence of NASH with or without cirrhosis and with or without cryptogenic cirrhosis in kindreds suggests a genetic risk [Struben et al., 2000]. Willner and colleagues studied 90 patients with NASH and found that 18% had a first degree relative that is also affected [Willner et al., 2001]. Prevalence of NAFLD is suggested to be different between different ethnic backgrounds. African Americans are suggested to be at lower risk compared with European Americans and Hispanic Americans which are regarded to be at highest risk [Romeo et al., 2008]. Difference between ethnicity is an important factor that could help understand genetic susceptibility to NAFLD.

1.8.2 Genomic approaches to determining novel candidate genes

Before the completion of the human genome project (HGP) in 2004 [HGP, 2004] and the establishment of organisations such as the HapMap project [Gibbs, 2003] and recently the 1000 genome project [Durbin et al., 2010], identifying susceptibility genes was achieved using linkage analysis also known as positional cloning. Linkage analysis used to be a major technique in identifying candidate genes however it did require large families to be studied [Kwon and Goate, 2000]. Linkage analysis was successful in identifying genes involved in single gene disorders but of more limited value for complex genetic disorders [Risch, 2000].

The candidate gene approach however does not require large families and can be used on unrelated cases and controls or small families [Kwon and Goate, 2000]. However it requires the biological knowledge of the complex disease being studied which in some cases is difficult when the aetiology of the disease is unknown. Once a candidate gene has been selected to study, selecting a polymorphism with functional variation that influences susceptibility to the complex disease can be challenging. In the past, this was frequently based on arbitrary knowledge concerning a very limited number of polymorphisms. However more recently, the SNP and HapMap projects have provided comprehensive information on common polymorphisms genome-wide, including the provision of detailed information on linkage disequilibrium within the human genome. This has facilitated more detailed candidate gene studies where, as discussed in more detail below, once a particular gene is chosen for study on the basis of biological relevance, a range of polymorphisms that tag the main haplotypes can be studied, together with possibly rarer SNPs of established functional significance.

The HapMap project, together with improvements in genotyping technology, has particularly facilitated the increasing use of genome-wide association studies (GWAS) to identify disease genes without prior knowledge of disease aetiology.

1.8.2.1 HapMap project and its relevance to genetic studies on disease susceptibility

Though the existence of many polymorphisms in human genes was known from the 1980s when cloning and sequencing of human genes was initiated, following the human genome project larger numbers of polymorphisms were identified systematically by the SNP consortium (TSC). The aim of TSC was to produce a public resource of SNPs in the human genome in 1999 when it was first established and the goal was to reach 300,000 SNPs. However results exceeded the goal and 1.4 million SNPs was recorded and released to the public in 2001 [Sachidanandam et al., 2001;Thorisson and Stein, 2003]. The HapMap project was then initiated to investigate the relationships between these SNPs, characterising linkage disequilibrium (LD) patterns and identifying haplotype-tagging SNPs, and to examine haplotypes in a range of different ethnic groups. [Gibbs, 2003;Seng and Seng, 2008]. HapMap started with 270 DNA samples (Phase I and Phase II) and now includes 1,301 samples (Phase III) from 11 different populations and covers around 10 million SNPs showing their frequencies and LD between nearby SNPs.

SNPs, haplotype blocks and tag SNPs

SNPs are the most common types of genetic variations found on average every 300 base pairs (Figure 1.6a) [Gibbs, 2003]. LD as defined by Cardon is “the non-random association of alleles at tightly linked markers” [Cardon and Abecasis, 2003]. The measure of LD is r^2 which ranges from zero (no LD) to one (complete LD) [Wall and Pritchard, 2003]. The stronger the LD between two alleles the more likely they will be co-inherited together and therefore studying one will result in knowing the other. Co-inherited alleles on the same chromosome where little recombination occurs between allele are called haplotype (Figure 1.6b) [Gibbs, 2003]. Gabriel and colleagues demonstrated that the human genome is divided into haplotype blocks where each block has little recombination [Gabriel et al., 2002]. The study suggested only three to five haplotypes within each block are needed to capture around 90% of all chromosomes in the population [Gibbs, 2003]. LD allows the selection of a few tag SNPs that are able to predict the remainder of the SNPs in the genomic region and determine the common haplotypes (Figure 1.6c) [Gibbs, 2003]. It is suggested that in order to construct a haplotype map for the human genome 300,000 to 1,000,000 chosen tag SNPs are required [Gibbs, 2003]. Gabriel also pointed out the LD between adjacent blocks could also lower the number of tag SNPs needed to be studied [Gabriel et al., 2002]. The LD between SNPs and the availability to select tag SNPs for identifying common haplotypes in selected genes for association studies is the hallmark of the HapMap project today.

1.8.2.2 Genome wide association approach

Following the completion of the human genome project and advances in technology that allowed genotyping (Illumina and Affymetrix), HapMap project (LD), statistical analysis (PLINK) and collection of samples has lead to the feasibility of GWAS for review see [Seng and Seng, 2008]. After the success of GWAS of age-related macular degeneration (AMD) in 2005 [Klein et al., 2005], several GWAS were published in 2006 [Seng and Seng, 2008]. Among those were studied of type 2 diabetes [Altshuler et al., 2000;Grant et al., 2006]. The success of GWAS in type 2 diabetes came from the convincing evidence of TCF7L2 association found in more than 20 studies from different populations [Cauchi et al., 2007].Three loci were found associated with type 2 diabetes risk in

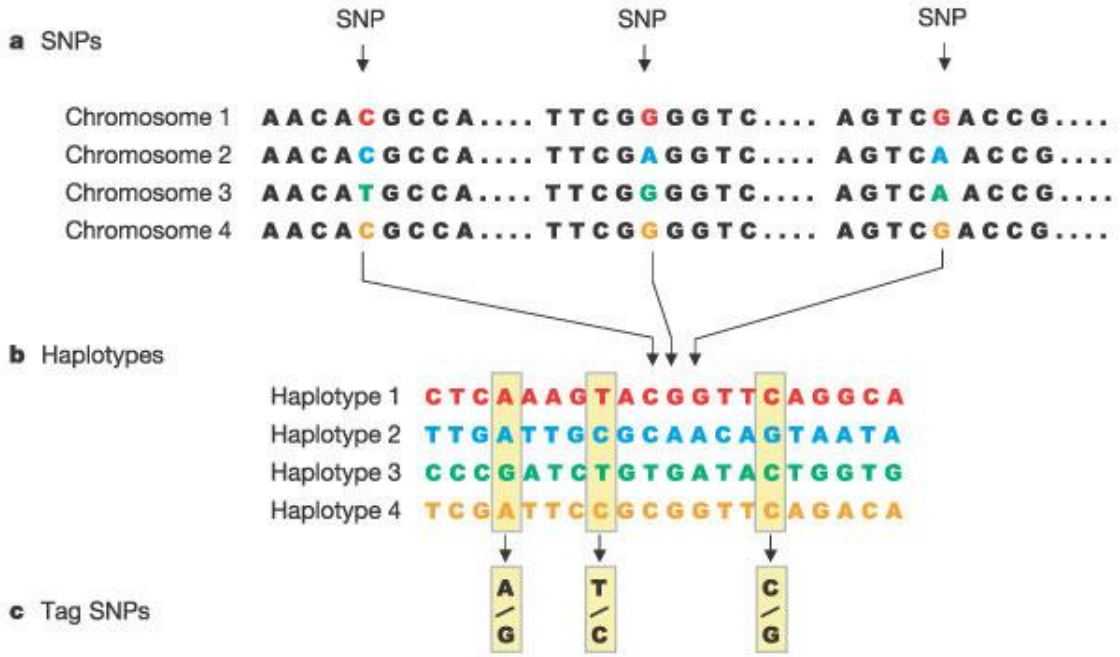

Figure 1.6: representation of SNPs, and assembly of haplotypes from tag SNP [Gibbs, 2003]

three GWASs [Zeggini et al., 2007;Saxena et al., 2007;Scott et al., 2007]. The sample size of these three studies reached 32 000 allowing the detection of modest size effects (OR 1.1) [Seng and Seng, 2008]. A recent GWAS on a multi-ethnic cohort from southeast Asia (3781 cases and 4354 controls) reported 17 genes to be associated with type 2 diabetes ($p<9x10^{-6}$) [Sim et al., 2011]. A GWAS on extremely obese children from France and German study groups identified FTO gene to be associated with obesity ($p=7x10^{-13}$) [Scherag et al., 2010]. The same gene was also found associated to obesity in a recent GWAS on non-Hispanics Caucasians ($p=2x10^{-12}$) [Wang et al., 2011].

1.9 Candidate gene association studies in NAFLD

A relatively large number of candidate gene association studies in relation to both NAFLD susceptibility and progression have now appeared. An overview of all such studies to date is provided in table 1.1. Many of these studies have been on single SNPs in candidate genes though a few recent studies have involved more comprehensive studies on a range of SNPs in particular candidates. There are a number of general issues that need to be considered when performing genetic studies on NAFLD. One important issue concerns controls. Since NAFLD is a very common disease, finding unaffected controls requires use of diagnostic tests with the most accurate being liver biopsy. This is not an appropriate investigation for controls but some studies have used alternatives such as ultrasound scanning. An alternative approach is to study genes concerned with disease progression rather than susceptibility and not include a healthy control group. Other important issues for studies on NAFLD genetics include the need for consistency in determining histological stages of NAFLD between studies using the same technique and for ethnicity to be taken into consideration.

In this section, candidate gene association studies in relation to NAFLD susceptibility and severity will be considered with particular focus on genes relevant to oxidative stress and lipid metabolism, the main subjects of this thesis.

Table 1.1: Genetic predictors of NAFLD susceptibly and severity

Gene	SNP	Suggested Effect	MVA performed	Replication and result	References
PNPLA3	rs738409; I148M	Both NAFLD only and NAFLD cases and healthy controls studied. NAFLD cases show carrying the variant allele show higher steatosis and fibrosis levels than homozygous wild-type cases.	Yes	Yes, Positive	[Romeo et al., 2008;Kotronen et al., 2009b;Speliotes et al., 2010;Sookoian et al., 2010;Valenti et al., 2010b;Romeo et al., 2010a;Rotman et al., 2010;Santoro et al., 2010]
FDFT1	rs2645424	Higher NAFLD activity score in NAFLD cohort for carriers of variant allele compared with wild-type cases.	Yes	Yes but negative	[Chalasani et al., 2010;Ballestri et al., 2011]
COL13A1	rs1227756	Higher lobular inflammation in NAFLD cohort for carriers of variant allele compared with wild-type cases.	Yes	No	[Chalasani et al., 2010]
TNF-α	Various	Increased frequency of carriage of variant alleles in NAFLD and NASH compared with healthy controls		Yes but negative	[Valenti et al., 2002;Tokushige et al., 2007;Hu et al., 2009]
IL-6	rs1800795 (G-174C)	Increased frequency of carriage of variant allele in NAFLD and NASH compared with healthy controls		No	[Carulli et al., 2009]

CD14	rs2569190 (C-159T)	Increased frequency of carriage of variant allele in NASH compared with simple steatosis	No	[Day CP, 2006]
GCLC	rs4140528 (C-129T)	Increased carriage of variant allele in NASH compared with simple steatosis	No	[Oliveira et al., 2010]
SOD2	rs4880 (A16V)	Carriage of variant (T) allele associated with higher fibrosis score in NAFLD cases and families	Yes, Positive	[Namikawa et al., 2004;Nobili V, 2007]
PEMT	rs7946 (V175M)	Frequency of carriage of variant allele more common in (i) NASH cases compared with healthy controls (ii) NAFLD cases compared with healthy controls	Yes but negative	[Dong et al., 2007;Song et al., 2005;Jun et al., 2009]
MTP	rs1800591	"Wild-type" GG (i) more common in NASH compared with healthy controls (ii) associated with higher ALT in type II diabetics	Yes but negative	[Namikawa et al., 2004;Bernard et al., 2000;Oliveira et al., 2010]
APOC	rs2854116/2854117	Carriage of variant alleles more common in subjects with NAFLD compared with those without NAFLD	Yes but negative	[Petersen et al., 2010;Kozlitina et al., 2011][Valenti et al., 2011]

HFE	rs1800562 (C282Y)	(i) Increased incidence of NASH compared with simple steatosis in those with the variant allele, (ii) increased incidence of fibrosis in NASH patients with the variant allele		Yes, negative and positive	[George et al., 1998;Nelson et al., 2007;Valenti et al., 2010c;Bugianesi et al., 2004]
ATGR1	rs3772622	Carriage of variant associated with NAFLD susceptibility compared with healthy controls. Fibrosis severity also higher		No	[Yoneda et al., 2009]
KLF6	rs3750861	Homozygous wild-type associated with increased fibrosis risk in NAFLD cohort.	Yes	No	[Miele et al., 2008]
ENPP1	rs1044498(L121Q)	Carriage of variant associated with increased fibrosis risk in NAFLD cohort	Yes	Yes but negative	[Dongiovanni et al., 2010;Carulli et al., 2009]
IRS1	rs1801278(G972R)	Carriage of variant associated with increased fibrosis risk in NAFLD cohort	Yes	No	[Dongiovanni et al., 2010]
ADIPOQ	rs2241766	Increased frequency of homozygous variant genotype in NAFLD cases compared with healthy controls. This genotype also associated with increased NASH severity		No	[Musso et al., 2008]

PXR	rs7643645	Carriage of variant shows "higher disease severity" in NAFLD cohort and more frequent in NAFLD cohort compared with healthy controls	No	[Sookoian et al., 2010]
UGT1A1	Upstream TA repeat	Extra TA repeat more common in paediatric NAFLD cases compared with obese controls without NAFLD	No	[Lin et al., 2009]

Table adapted from [Daly et al., 2011] and slightly modified.

1.9.1 Candidate gene association related to oxidative stress in NAFLD

Only a few candidate genes related to oxidative stress in NAFLD have been investigated to date.

1.9.1.1 SOD2

The gene *SOD2* encodes for the mitochondrial enzyme MnSOD (see also section 1.7.1.2). Following translation, an N-terminal mitochondrial targeting sequence (MTS) of 24 amino acids is cleaved in the mitochondrial matrix and the enzyme is then assembled into an active tetramer [Sutton et al., 2005]. A cytosine to thymine polymorphism at position 47 which is associated with an alanine to valine amino acid substitution occurs in the targeting sequence [Degoul et al., 2001]. The presence of ala-*SOD2* is predicted to result in an α-helix in this region of the protein and val-*SOD2* in a β-sheet structure [Shimoda-Matsubayashi et al., 1996]. The ala-*SOD2* variant appears to achieve high mitochondrial activity due to its structure favouring import to the mitochondria whereas the val-*SOD2* variant may be stuck in the narrow translocase of the inner mitochondria resulting in lower total activity (approximately 40% lower compared with the Ala-*SOD2* variant) [Sutton et al., 2005]. In one of the earliest studies on NAFLD genetic susceptibility, the Ala16Val *SOD2* polymorphism was found to be more common in 63 Japanese NASH patients compared with 150 controls [Namikawa et al., 2004]. However, paradoxically, the ala-*SOD2* variant has been suggested to be associated with severe alcoholic liver disease though not all studies confirm this [Degoul et al., 2001;Nahon et al., 2005;Stewart et al., 2002]. The val-*SOD2* was also found to be associated with other oxidative stress related conditions including type I diabetic nephropathy in a Swedish [Mollsten et al., 2007] and Denmark [Mollsten et al., 2009] populations. The val-*SOD2* was also found to be associated with pancreatic adenocarcinoma in a US-based population [Wheatley-Price et al., 2008].

1.9.1.2 GCLC

A promoter region polymorphism C-129T in GCLC suggested to influence promoter activity has been investigated in NAFLD subjects. The C-129T *GCLC* has recently been studied in 131 biopsy-proven NAFLD patients from Brazil with heterozygous being more common in NASH patients compared with those with steatosis only [Oliveira et al., 2010]. The association result is consistent with the T-allele being associated with lower promoter activity compared to the C-allele [Nakamura et al., 2002]. However no replication studies have been available to confirm these findings.

1.9.1.3 UGT1A1

A common TA repeat polymorphism located 41 nucleotides upstream of the TATA box reduces levels of transcription of UGT1A1 [Ritter et al., 1992]. Normally 6 TA repeats occur but some individuals may show either 7 repeats or occasionally 5or 8 repeats. The 7 repeat allelic variant is referred to as *UGT1A1*28.* Individuals with Gilbert's syndrome who show higher than normal plasma bilirubin levels are usually homozygous for this *28 variant allele. This is because the lower UGT1A1 expression associated with this genotype results in higher levels of serum bilirubin [Lampe et al., 1999]. Because bilirubin can act as an antioxidant, the possible relevance of UGT1A1 genotype to diseases concerned with oxidative stress has been investigated. For example, possession of UGT1A1*28 was associated with lower risk of cardiovascular disease [Lin et al., 2006]. *UGT1A*6* but not the UGT1A1*28 variants have been found to be protective against NAFLD in 234 Taiwanese children [Lin et al., 2009].

1.9.1.4 HFE

Increase in hepatic iron is associated to severity of fibrosis [George et al., 1998]. The common *HFE* gene mutation has been found to be associated with NAFLD and severity of fibrosis [Nelson et al., 2007] however the results have been conflicting [Valenti et al., 2010c].

1.9.2 Candidate gene association studies related to lipid metabolism in NAFLD

1.9.2.1 PEMT

A non-synonymous polymorphism Val175Met in the *PEMT* gene was found to be associated with biopsy-proven NAFLD in 28 cases (67.9% Mut/Mut) compared to 59 controls (40.7% Mut/Mut) ($p=0.03$) without fatty liver in a US-based study [Song et al., 2005]. The Val175Met was found to be associated with NASH (107 biopsy-proven Japanese NASH patients) compared to controls (150 healthy volunteers) ($p < 0.001$) [Dong et al., 2007]. A recent Korean study on 588 ALD and NAFLD patients did not find an association with fatty liver disease on *PEMT* Val175Met [Jun et al., 2009]. The inconsistency between the studies could be due to sample size differences.

1.9.2.2 MTP

A promoter region polymorphism G-493T in the *MTP* gene, previously described to be involved in TG-VLDL assembly was found to be associated to NASH cases in 63 Japanese biopsy-proven NASH patients compared to 150 healthy controls. The G-allele is associated to decreased transcription of *MTP* and therefore leads to less exports of TG

[Namikawa et al., 2004]. However, this was not confirmed in a recent study on 131 biopsy-proven NAFLD and 141 healthy volunteers from Brazil [Oliveira et al., 2010]. The inconsistency of association could be due to the different populations being studied. An association of *MTP* with ALT was found suggesting an indirect association with NASH [Bernard et al., 2000].

1.9.2.3 APOC3

Apolipoproteins are components of lipoproteins and are involved in regulating lipoproteins metabolism [Mahley et al., 1984]. Two promoter region polymorphisms C-482T and T-455C in the *APOC3* gene have been found to be associated with NAFLD in 95% healthy Asian Indian men [Petersen et al., 2010]. However, dissociation of the polymorphisms were found in a separate study on Americans [Kozlitina et al., 2011]. The inconsistency could be due to the differences in ethnicity.

1.9.2.4 PXR

PXR (pregnane X receptor) is suggested to be involved in lipid homeostasis. Activation of PXR in mice is suggested to cause hepatic steatosis by increased TG deposition in the liver [Zhou et al., 2006]. Recently polymorphisms in PXR have been found to be associated to progression of NAFLD and to its more severe stages [Sookoian et al., 2010].

1.9.3 Candidate gene association related to insulin resistance in NAFLD

1.9.3.1 ENPP1 and IRS-1

Polymorphisms in *ENPP1* (ectoenzyme nucleotide pyrophosphate phosphodiesterase 1) and *IRS-1* (insulin receptor substrate-1) genes affecting insulin receptor activity have been investigated in 702 biopsy-proven NAFLD patients from Europe (Italy and UK) [Dongiovanni et al., 2010]. The variant-alleles were found to be associated with increased prevalence of fibrosis scores above 1. However the *ENPP1* polymorphism was not found to be associated to fibrosis in a smaller (n=114) study based in Italy [Carulli et al., 2009]. The inconsistency could be due to the lack of availability of biopsies in the [Carulli et al., 2009] study (59 out 114 were biopsied) which might have limited the recognition of association.

1.9.3.2 TNF-α

TNF-α -238 SNP located in the promoter region was found to be significantly prevalent in NAFLD patients (n=99) compared to controls (n=172) (31% vs 15% $p<0.0001$) in an

Italian-based study [Valenti et al., 2002]. Recently the -238 polymorphism was found to be significant in NAFLD patients (n=189) compared to controls (n=138) in a Chinese population [Hu et al., 2009]. However a Japanese study on 102 NAFLD patients (36 with steatosis and 66 with NASH) compared against 100 controls showed -1031 and -863 but not the -238 of *TNF-α* was found to be associated with NASH [Tokushige et al., 2007]. The inconsistency between studies could be due to the differences in the stages of NAFLD being studied.

1.9.3.3 ADIPOQ

Musso and colleagues described two common SNPs in the adiponectin gene (*ADIPOQ*) exon/intron 2 (45GT and 276GT) which showed an association with NAFLD and severity of disease [Musso et al., 2008].

1.9.4 Other candidate gene association related to NAFLD

1.9.4.1 KLF6

The Kruppel-like factor (KLF) belongs to a family of transcription factors suggested to be involved in liver fibrogenesis. A functional polymorphism in KLF6 is associated to severity of fibrosis in a European (UK and Italy) cohort [Miele et al., 2008].

1.9.4.2 AGTRI

Angiotensin II type 1 receptor (AGTRI) is found expressed on HSC which are involved in fibrosis progression in NASH [Togashi et al., 2000]. Recently, *AGTRI* polymorphism was found to be associated with severity of fibrosis [Yoneda et al., 2009]. The same study also found five other SNPs to be associated with NAFLD [Yoneda et al., 2009].

1.9.4.3 IL-6

IL-6 acts as a pro-inflammatory cytokine, has been found to be associated to NAFLD in a relatively small study (n=114 of which 59 undergone liver biopsy) [Carulli et al., 2009]. The IL-6 C-174 was found to be prevalent in NAFLD subjects compared to controls, and higher in NASH compared to NAFLD [Carulli et al., 2009].

1.10 GWAS in NAFLD

Up to the present date, three GWAS on NAFLD patients have been reported. In a limited GWAS involving only nonsynonymous polymorphisms, 383 Hispanics, 1032 African-Americans and 696 European-Americans all were assessed for hepatic fat content by magnetic resonance imaging [Romeo et al., 2008]. The study found a non-synonymous variant in the patatin-like phospholipase domain-containing protein 3

(PNPLA3, rs738409 I148M) to be associated with hepatic fat accumulation. The variant allele (G-allele) was most commonly found in Hispanics which are known to be most susceptible to NAFLD compared to African-Americans and Europeans [Browning et al., 2004]. The I148M PNPLA3 variant was also associated with ALT levels, which is consistent with population-based GWAS carried by Yuan and colleagues on plasma liver enzyme where variants in PNPLA3 predicted plasma ALT levels [Yuan et al., 2008]. Recently a GWAS on 236 female patients with well-characterized NAFLD has been reported [Chalasani et al., 2010]. A SNP (rs343062) in the gene *FDFT1* encode for the enzyme farnesyl diphosphate farnesyl transferase 1 (FDFT1) is associated with NAFLD activity score. FDFT1 is involved in cholesterol biosynthesis [Schechter et al., 1994]. The collagen gene *COL3A1* SNP (rs1227756), *EFCAB4B* SNP (rs887304) and *LTBP3* SNP (rs6591182) was associated with lobular inflammation. However no association was found with PNPLA3 rs738409 which was not captured by GWAS. However the PNPLA3 variant was in complete LD with five other PNPLA3 genotyped SNPs. A reason for this dissociation is the modest sample size and that only females were selected. Recently a genome wide association study on 7176 individuals from four studies was carried out in a meta-analysis study [Speliotes et al., 2011]. Measures of hepatic steatosis were carried out using computed tomography. Five SNPs were found to be associated with increased hepatic steatosis, the common *PNPLA3* rs738409 and three novel SNPs rs4240624 near *PPP1R3B*, rs2228603 near *NCAN*, rs12137855 near *LYPLAL1* and rs780094 near *GCKR*. In a separate cohort of 592 subjects with biopsy proven NAFLD and 1405 healthy controls, *PNPLA3, NCAN, GCKR and LYPLAL1* variants were all found associated with NASH.

The associations found in GWAS are interesting but need to be replicated in independent cohorts. To date, the association with the PNPLA3 variant I148M has been replicated in a number of studies [Kotronen et al., 2009b;Sookoian et al., 2009;Valenti et al., 2010a;Romeo et al., 2010a;Speliotes et al., 2010;Valenti et al., 2010b;Rotman et al., 2010;Santoro et al., 2010], as discussed in more detail in Chapter 6. Only one replication study to date on the reported FDFT1 association has been performed. This was on 340 biopsy proven NAFLD cases from the UK but no association to severity of fibrosis could be demonstrated [Ballestri et al., 2011]. The strong association with *NCAN*, *GCKR* and *LYPLAL1* detected recently also require replication [Speliotes et al., 2011].

1.11 Aim

Although some progress has been achieved in identifying genetic risk factors for NAFLD, there are still considerable gaps to be filled. As a result the genes involved in the initiation and progression of NAFLD are still unclear. The aim of this study was to identify candidate genes that confer progression of NAFLD to its progressive stages. The SNPs to be studied in candidate genes were selected based on both the published literature and bioinformatic databases but were consisted of both tag SNPs and non-synonymous and promoter region SNPs. Their relevance was assessed by single association studies, haplotype analysis, and gene-gene interaction analysis. Two complementary approaches will be used to assess the relationship of the selected SNPs and severity of disease in association studies. Firstly, case-control studies assessing patients with mild NAFLD against severe NAFLD were carried out and secondly, intra-familial association will be studied in "trios" by assessing the transmission of alleles from parents to the affected child. As increased oxidative stress due to lipid accumulation appears to be an important contributor to NAFLD severity though information on genetic associations is limited, the study of candidate genes relevant to oxidative stress such as SOD2 were emphasised with follow up studies on assessment of functional significance being performed if appropriate. In addition, studies on the relevance to NAFLD severity of genes already reported as risk factors in association studies on other NAFLD cohorts were performed. Studies by others on differences in gene expression between normal liver biopsies and biopsies from NAFLD patients provided additional candidate genes for genotyping analyses and for functional studies. This combination of genetic approaches enabled multiple regression models to be applied to get a better understanding of overall risk of NAFLD development from several genetic risk factors.

Chapter 2

Materials and Methods

2 Materials and Methods

2.1 Patient samples

Samples from 351 unrelated patients from Newcastle with biopsy-proven NAFLD and 71 Italian family 'trios' (213 individuals) from Rome, Italy, each comprising two living parents and an index child with biopsy-proven fibrotic NAFLD were collected. The study had ethical approval from the Newcastle upon Tyne Local Research Ethics committee and from the relevant Italian committee and all participants (or their parents) gave informed consent. Other causes of liver disease were excluded such as chronic viral hepatitis, autoimmune hepatitis, hereditary hemochromatosis, Wilson's disease and drug induced liver disease, including increased alcoholic intake (>30/>20 g/day for males/females, respectively), as confirmed by at least one family member or friend.

Clinical and laboratory data were collected on the date a diagnostic liver biopsy was performed for the adult NAFLD Newcastle samples (Table 2.1). Liver histology scoring was performed according to [Kleiner et al., 2005]. Kleiner and colleagues assigned score 0 to livers without fibrosis, score 1 was divided further into scores 1A, 1B and 1C characterised with perisinusoidal or periportal fibrosis (score 1), mild zone 3, perisinusoidal (1A), moderate zone 3 perisinusoidal (1B) and portal/periportal (1C). Score 2 was assigned for livers with both perisinusoidal and periportal fibrosis, score 3 with those showing fibrosis bridges and score 4 with those showing cirrhosis. Due to the existence of sampling error in biopsy and the differences between 0 and 1 being subtle, stages 0 and all scores of 1 (mild disease) were grouped together and compared with stages 2-4 (severe disease). Body mass index (BMI was calculated using the formula: weight (kilograms)/height (m^2) and presence of diabetes mellitus (fasting glucose $\geq$7.1 mmol/L mg/dl or treatment with anti-diabetic drugs) was recorded. Laboratory evaluation included routine liver biochemistry such as alanine aminotransferase IU/L (ALT), total cholesterol mmol/L, triglycerides (TG) mmol/L and homeostasis model assessment of insulin resistance (HOMA-IR). For the Italian children, 61 of the 71 children had fibrosing steatohepatitis (59 patients with stage 1 fibrosis and 2 patients with stage 2 fibrosis). Detailed clinical data concerning these children has been presented previously [Nobili et al., 2006].

2.2 Laboratory methods

2.2.1 Materials

The name of suppliers of all chemical and enzymes used in this work is listed (Table 2.2).

2.2.1.1 Preparation of solutions and stock solutions

Reverse osmosis distilled water was used to prepare solutions. Sterilisation via autoclave for solutions was applied were appropriate at 120 ^{o}C, 15 pounds per inch (PSI) pressure for 20 min. Plastic ware was autoclaved unless purchased sterilised. Tissue culture solutions were filter-sterilised using 0.2mm filters (Millipore, MA, USA).

Table 2.1: Clinical characteristics of Newcastle NAFLD patients

	Newcastle samples
Number	351
Sex (Male)	218 (62.1%)
Age, years	49.7 ± 12.7
BMI, kg/m^2	34.05 ± 5.3
DM (Yes)	131 (37.3%)
HOMA-IR	6.3 ± 6.2
ALT, IU/L	80.6 ± 64.05
Total Cholesterol, mmol/L	5.6 ± 1.3
TG, mmol/L	2.81 ± 2.0
Steatosis Score	
1	99 (29%)
2	151 (44%)
3	93 (27%)
NASH (yes)	194 (56%)
Fibrosis Score	
0	134 (38%)
1	65 (19%)
2	57 (16.5%)
3	55 (15.5%)
4	40 (11%)

Table 2.2: List of suppliers and addresses

Supplier	Address
Bioline	London, UK
Bio-Rad	Hemel Hempstead, UK
Eurofins MWG Operon	London, UK
Fisher Scientific	Loughborough, UK
Invitrogen	Paisley, UK
New England Biolabs	Hitchin, UK
Promega	Southampton, UK
Qiagen	Crawley, UK
Sigma Aldrich	Gillingham, UK

Table 2.3: components of frequently used stock solutions

Solution	Constituents
10x TBE	0.9 M Tris-base
	0.9 M boric Acid
	20 mM EDTA PH 7.0
DNA gel loading buffer	0.25% (w/v) bromophenol blue
	0.25% (w/v) xylene cyanol
	30 % glycerol

Purchased sterile water was used for PCR (Fresenius Kabi Limited, Chesire, UK). Frequently used stock solution and their compositions are listed in Table 2.3

2.2.2 DNA extraction

DNA extraction from peripheral blood leukocytes was performed by Julian Leathart and Julia Patch using a perchlorate-chloroform extraction method as described previously [Daly et al., 1996].

2.2.3 Polymerase Chain Reaction (PCR)

Lyophilised primers were purchased. Primers were resuspended to 200 µM using sterile water (Fresenius Kabi Limited, Chesshire, UK). Working dilutions (25 µM) were made and stored at 4^{o}C with the main stock stored at -20^{o}C. Genomic DNA (0.2 ug) was amplified in a total reaction volume 20 µl, containing 0.1 mM dNTPs (Bioline, London, UK), 0.25 µM forward primer, 0.25 µM reverse primer (Eurofins MWG, London, UK), 1 X supplied buffer and 0.025 units *Taq* DNA polymerase (New England Biolabs (NEB), Hitchin, UK). Reactions were performed in 0.2 ml sterile tubes (Fisher Scientific, Loughborough, UK). An Applied Biosystems 2720 Thermal cycler was used for PCR. Cycling conditions were standard, with only the annealing temperature varying depending on primer set. Quality of PCR was checked on agarose gel (see section 2.1.5.1).

2.2.4 Restriction enzyme digestion

Restriction enzymes were purchased from Fermentas and NEB. Digestion was achieved in a reaction volume of 25 µl with 2 U of digestion enzyme, digestion buffer provided and 5 µl of PCR product. Samples were incubated overnight at the temperature specified by the manufacturer. Digestion products were then run on either 2% agarose or polyacrylamide gels depending on the product sizes.

2.2.5 Electrophoresis and visualisation of DNA

2.2.5.1 Agarose gel electrophoresis

DNase and RNase free agarose powder (Bioline, UK) was dissolved by heating in 1 X TBE buffer. Normally 2% gels containing 0.5 µg/ml ethidium bromide (Sigma-Aldrich, Gillingham, UK) were used. DNA gel loading buffer (1 µl) was added to 5 µl of PCR product. A 100 - 1000bp size marker ladder (NEB) was included on the gel. Electrophoresis was carried out for 15-30 minutes at ~80V in 1 x TBE buffer.

2.2.5.2 Polyacrylamide gel electrophoresis (PAGE) and ethidium bromide staining

Polyacrylamide gels (10%) were made up from 30% (29:1) acrylamide solution (Fisher Scientific) in 10 X TBE buffer, 0.1% (v/v) TEMED (Fisher Scientific) and 0.4 mg/ml ammonium persulphate (Bio-Rad, Hemel Hempstead, UK) in a final volume of 50 ml. The solution was poured between two sealed glasses plates and left for 20-30 minutes to set. Loading buffer (2 µl) was added to the 25 µl DNA digest products. Gels were run at 150V for 4-5 hours in 1 X TBE buffer, followed by gel staining in 1 X TBE and 0.5 µg/ml ethidium bromide for 30 minutes.

2.2.5.3 Gel visualisation

Agarose and polyacylamide gels were visualised and photographed initially on BioRad gel documentation system and photographed using Fluro S-multi imager Quantity One software and on an AlphaImager 2200 documentation and analysis system.

2.2.6 KBiosciences genotyping

Genotyping was also carried out by KBiosciences, Essex where appropriate. The KBiosciences competitive allele-specific PCR (KASP) genotyping system is a homogenous, fluorescent, endpoint genotyping system that uses three unlabelled primers and a universal reaction mix. KBiosciences achieves a >99.5% success rate and an error rate of <0.03% in genotyping. 96-well plates were set up with DNA according to the manufacturer's guidelines. Quality control was performed by including negative (water only) controls and genotyping replicate DNA samples (in every 96 well plate sent).

2.3 Statistical analysis

The chi-square test was used to ensure genotypes were in Hardy-Weinberg Equilibrium (HWE) using web-based calculator available at www.tufts.edu/. Pearson's chi-square test was performed with odds ratio (OR) and 95% confidence intervals (CI) and chi-square test for trend were calculated for the Newcastle NAFLD samples using PRISM version 5.0a software.

Where appropriate statistical analysis was performed using various packages of the R program including 'gtools', 'gdata', 'MASS', 'mvtnorm', 'genetics', 'combinat' and 'dgc.genetics' running in the R software environment version 2.7.1 [Team, 2008]. Multivariate analysis by logistic regression to evaluate the factors associated with progressive NAFLD was performed using R statistical package. The results of multivariate analysis were expressed as OR with 95% CI. The transmission

disequilibrium test (TDT) was performed on the Italian trio families using the R statistical package.

Haplotypes and their frequencies were determined using PHASE version 2.1 [Stephens et al., 2001]. Frequencies between Newcastle NAFLD samples and HapMap Caucasian samples were then compared by Pearson's chi-square analysis.

SPSS version 17.0 was used for calculations on clinical parameters such as means and standard deviations.

2.3.1 Power calculation

EpiInfo version 6 was used to calculate the number of samples required to show statistical significance. To detect a difference between mild and severe cases at an odds ratio of 2.0 with $p<0.05$ and a power of 80%, 165 controls and 165 cases would be required if the variant genotype frequency in controls is 25%.

Chapter 3

Superoxide dismutase 2 (SOD2) genotype and its association with NAFLD severity

3 Superoxide dismutase 2 (SOD2) genotype and its association with NAFLD severity

3.1 Introduction

Oxidative stress results from an imbalance between ROS and antioxidants as discussed in section 1.7.1.1. The gene *SOD2* encodes the mitochondrial antioxidant enzyme manganese-dependent superoxide dismutase (MnSOD). MnSOD dismutates superoxide anions (O_2^-) into hydrogen peroxide (H_2O_2). An N-terminal mitochondrial targeting sequence (MTS) of 24 amino acids is cleaved in the mitochondrial matrix and the enzyme is then assembled into an active tetramer [Sutton et al., 2005]. A common non-synonymous cytosine to thymine polymorphism at position 47 (C47T) in the *SOD2* gene is associated with alanine to valine amino acid substitution at codon 16 (Ala16Val) in the mitochondrial targeting sequence [Sutton et al., 2003]. The presence of the ala-*SOD2* (C-allele) is predicted to result in an α-helix structure favouring import to the mitochondria and therefore achieving high enzyme activity. However, the val-*SOD2* (T-allele) is predicted to result in a β-sheet structure which may get stuck in the narrow translocase of the inner mitochondrial membrane resulting in lower enzyme activity (approximately 40% lower compared with the ala-*SOD2* variant) [Shimoda-Matsubayashi et al., 1996;Sutton et al., 2005;Sutton et al., 2003]. In one of the earliest studies on NAFLD genetic susceptibility, homozygosity for the val-*SOD2* was found to be more common in 63 biopsy-proven Japanese NASH patients compared with 150 controls (p=0.016) [Namikawa et al., 2004]. In oxidative stress related diseases homozygotes for the val-*SOD2* were also found in two studies [Mollsten et al., 2007] [Mollsten et al., 2009] from Sweden and Denmark respectively on type 1 diabetes to be associated with increased risk of diabetic nephropathy (p=0.049 and p=0.008 respectively). The val-*SOD2* was also found to be associated with pancreatic adenocarcinoma in a US-based population (p=0.04) [Wheatley-Price et al., 2008].

The aim of the studies described in this chapter is to assess the relationship between the Ala16Val polymorphism and susceptibility to severity of NAFLD by two complementary approaches using Newcastle NAFLD patients with biopsy proven disease and secondly Italian family trios.

3.2 Methods

3.2.1 SOD2 rs4880 genotyping assay

PCR-RFLP was used to genotype for rs4880. Primers for PCR were previously published by [Shimoda-Matsubayashi et al., 1996] but slightly modified for both forward and reverse primers (5`-CAGCCCAGCCTGCGTAGACGG-3` and 5`-GCGTTGATGTGAGGTTCCAG-3`). PCR amplification (35 cycles of denaturation at 95^{o}C for 1 minute, annealing at 63^{o}C for 1 minute and extension at 72^{o}C for 1 minute) was then performed in a GeneAmp PCR system 9700 thermal cycler. PCR product was then digested with *BsaWI* restriction enzyme. The PCR digest products were viewed on a 10% acrylamide gel. PCR-RFLP analysis on the Newcastle NAFLD patients was performed previously by Julian Leathart who supplied the genotyping data but the genotyping of the Italian families was performed as part of the present study.

3.3 Results

3.3.1 SOD2 genotype analysis

The non-synonymous rs4880 polymorphism in *SOD2* at position 47 on exon 2, results in an alanine to valine substitution (C to T). The *SOD2* rs4880 polymorphism was genotyped by PCR-RFLP with a typical result shown in Figure 3.1. Genotypes were in Hardy-Weinberg equilibrium (X^2=0.74, p=0.38). The relationship between genotype and severity of fibrosis was analysed (Table 3.1). Incidence of advance fibrosis (stage 2-4) increased with the number of T (Val) alleles (p=0.01 chi-square for trend) (Table 3.1). 65% of CC individuals had fibrosis scores 0 or 1 compared with 35% of homozygous T patients. Homozygosity for the T-allele vs. C-allele was associated with an odds ratio of 2.23 (95% CI 1.19-4.17; p=0.01) for the development of fibrosis of grade 2 and above (Table 3.1). There was no significant difference between CT vs. CC genotype (odds ratio 1.38 (95% CI 0.8-2.3); p=0.23) or CC vs. CT+TT (odds ratio 1.6 (95% CI 0.96-2.7); p=0.06).

The TDT showed 55 of the 71 families to be informative in that one or both parents were heterozygous. In these families the T-allele was found to be transmitted on 47/76 (62%) possible occasions and the C-allele was transmitted only 29/76 (38%) occasions, p=0.038 (Table 3.2).

A multivariate analysis using logistic regression was performed on the Newcastle NAFLD samples to control for the effect of factors such as age, sex, BMI, diabetes, steatosis and NASH that are relevant to disease severity. The multivariate results demonstrate that *SOD2* genotype, age, BMI, diabetes and NASH were all important predictors for advance fibrosis (Table 3.3). Carrying the T-allele was a significant and independent predictor of fibrosis with odds ratio 1.68 (95% CI 1.09 – 2.64; p=0.019). NASH appeared to be the strongest predictor of fibrosis (odds ratio 29.1 (95% CI 13.8 – 67.9); $p=2x10^{-16}$).

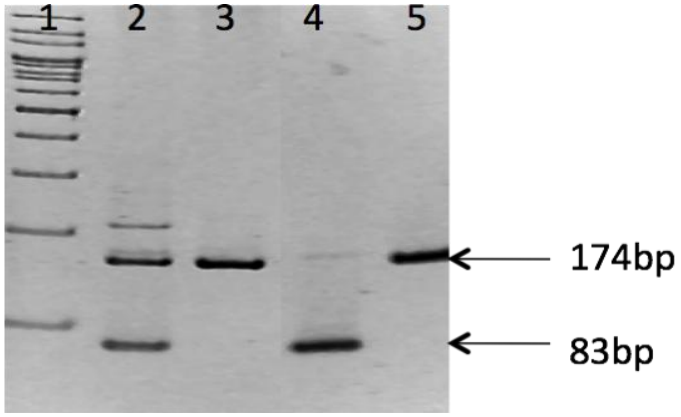

Figure 3.1: 10% acrylamide gel showing *SOD2* rs4880 digested with *BsaWI*. Lane 2 shows heterozygous (CT), lane 3 wild-type (CC) and lane 4 mutant (TT).

Table 3.1: *SOD2* genotyping distribution in Newcastle NAFLD patients – mild fibrosis vs. severe fibrosis

SOD2	Genotype	Fibrosis 0+1 (n=197)	Fibrosis 2+3+4 (n=151)
Rs4880	CC	53 (0.27)	28 (0.19)
	CT	105 (0.53)	77 (0.51)
	TT	39 (0.20)	46 (0.30)

Population Hardy-Weinberg calculation (X^2=0.74, p=0.38)

Chi-square for trend p=0.01

Odds ratio for TT against CC= 2.23 (1.19 – 4.17); p=0.01

Odds ratio for CC against CT+TT= 1.6 (95% CI 0.96-2.7); p=0.06

Table 3.2: *SOD2* transmission disequilibrium test, comparing the distribution of the transmitted and untransmitted allele from heterozygous parents to affected child

Gene and SNP number	Allele	Transmitted	Untransmitted	p-value
SOD2 rs4880	T	47 (62%)	29 (38%)	0.038
	C	29 (38%)	47 (62%)	

Table 3.3: Multivariate analysis using logistic regression for mild fibrosis vs. severe fibrosis to control for the effect of predictor factors which are relevant to fibrosis severity

Variables	OR (95% CI)	p-value
SOD2	1.68 (1.09 – 2.64)	0.019
Age	1.03 (1.00 – 1.06)	0.011
Sex	1.4 (0.72 – 2.74)	0.32
BMI	1.07 (1.01 – 1.14)	0.017
Diabetes	2.07 (1.13 – 3.86)	0.019

SOD2, Age, BMI and diabetes appear to be predictor factors

Calculations assume a co-dominant model for SOD2 genotype.

3.4 Discussion

We have shown a consistent association between a functional SNP in the mitochondrial targeting sequence of *SOD2* and severity of fibrosis in both case-control and intra-familial association methodologies. These results provide genetic evidence for the importance of mitochondria-derived oxidative stress and pathogenesis of NAFLD. The TDT analysis on the Italian family samples, 61 of the 71 children had fibrosing steatohepatitis (59 patients with stage 1 fibrosis and 2 patients with stage 2 fibrosis) is a novel feature which is significantly more powerful at detecting true associations compared to case-control approach and therefore inclusion of TDT analysis in an association study adds strength to any positive findings [Spielman and Ewens, 1996]. Nevertheless, our case-control approach on the adult NAFLD patients demonstrated an association with *SOD2* genotype and severity of fibrosis. The role of MnSOD in protection against oxidative stress is well demonstrated and discussed earlier in section 1.7.1.2. Previous reports have shown the lack of MnSOD is lethal in mice and mice heterozygous for knockout and showing decreased enzyme activity have an increased susceptibility to oxidative stress and mitochondrial dysfunction [Melov et al., 1998;Macmillan-Crow and Cruthirds, 2001;Li et al., 1995b;Lebovitz et al., 1996].

Our finding is consistent with the idea that the ala-allele (C-allele) resulting in an α-helix structure favours import to the mitochondria. This means this allele codes for a gene product which has greater enzyme activity compared with the val-allele (T-allele) where the β-sheet structure is suggested to be less efficiently imported resulting in lower enzyme activity (approximately 40% lower compared to C-allele) [Shimoda-Matsubayashi et al., 1996;Sutton et al., 2005]. Carrying the Val-allele is likely to result in more accumulation of the toxic radical (O_2^-) and less production of less toxic radical (H_2O_2) [Naganuma et al., 2008]. Our association of the T-allele with NAFLD was found consistent with the previous small study (n=63) on Japanese population with biopsy-proven NAFLD [Namikawa et al., 2004] and also with studies on other diseases where oxidative stress is likely to be relevant to disease pathogenesis such as diabetic nephropathy [Mollsten et al., 2007] [Mollsten et al., 2009] and pancreatic adenocarcinoma [Wheatley-Price et al., 2008]. SOD2 is a relatively small gene of 14 kb and consists of one haplotype block. From HapMap data on European subjects, four SNPs (rs2855116, rs8031, rs2758331 and rs2842964) are in strong LD with rs4880. The intronic polymorphism rs2855116 has been previously studied [Tomkins et al., 2001]

and found not to be a risk factor for amyotrophic lateral sclerosis, however rs4880 is the only common non-synonymous polymorphism described in *SOD2*.

FA oxidation occurs mainly in the mitochondria (β -oxidation) but also by peroxisome and microsomes. Excessive oxidation can result in reduction of antioxidant defences and therefore result in oxidative stress [Sanyal et al., 2001]. The synthesis of ROS by FA metabolism and the mitochondrial respiratory chain is increased in NAFLD patients [Pessayre and Fromenty, 2005]. There is evidence in both humans and rodents of an association between an increase in ROS and mitochondrial abnormalities in fibrotic NAFLD [Pessayre, 2007]. Therefore the high activity of the ala-allele (C-allele) is likely to protect the mitochondria from the high toxicity of superoxide anion (O_2^-) better than the val-allele (T-allele).

In conclusion the rs4880 *SOD2* polymorphism represents a biologically plausible candidate gene with a well established significant function. The study highlights the role of defence mechanisms involved in oxidative stress. However there still remains a need for validation in independent large cohorts by either candidate gene studies or genome wide association studies.

Chapter 4

Oxidative stress related genes and NAFLD severity

4 Oxidative stress related genes and NAFLD severity

4.1 Introduction

The association of a non-synonymous polymorphism in *SOD2* gene discussed earlier (see chapter 3) suggested that investigation of variants in additional genes encoding proteins involved in protection against oxidative stress would be appropriate (Figure 1.1). I have therefore assessed a number of common SNPs in selected genes to determine a possible relationship of these genes to NAFLD severity. A number of SNPs in these genes have already been studied by others either in NAFLD or in other diseases where oxidative stress may be important.

Following the dismutation of O_2^- into H_2O_2 by SOD2, GPX1 is involved in detoxifying H_2O_2 into H_2O. A cytosine to thymine nucleotide substitution at position 537 on exon 2, results in a proline (pro) to leucine (leu) change at codon 198 (Pro198Leu) of GPX1. This non-synonymous SNP is suggested to affect the levels of activity of GPX1, with the Leu-GPX1 being the less active form [Hu and Diamond, 2003].

As discussed in section 1.7.1.2, GSH is the major antioxidant molecule within the cell. The rate limiting step of GSH synthesis is catalyzed by GCL which has two subunits (GCLC and GCLM) encoded by different genes. A common promoter region SNP C-129T in *GCLC* has previously been suggested to be involved in *GCLC* promoter activity with the C-allele reported to be associated with higher promoter activity [Koide et al., 2003]. A common promoter region SNP C-588T has also been suggested to be involved in *GCLM* promoter activity with the C-allele reported to be associated with higher promoter activity [Nakamura et al., 2002]. Two additional SNPs in *GCLM* which "tag" common haplotypes have been investigated as susceptibility factors for methamphetamine-induced disorder and schizophrenia though with negative findings [Kishi et al., 2008]. GSR regenerates GSH from GSSG and is encoded by *GSR*. There have been few previous studies on disease susceptibility in relation to this gene but a promoter region SNP and six SNPs which tag common haplotypes have been investigated as risk factors for breast cancer survival though with negative results [Udler et al., 2007]. *MAT1A* encodes a subunit of MAT which catalyzes formation of SAMe, a precursor of GSH, making it an additional good candidate for relevance to oxidative stress. A recent study on hypertension and stroke investigated 8 polymorphisms in the *MAT1A* gene. No association was observed with 8 polymorphisms however haplotype

analysis using three tag SNPs was associated to hypertension, stroke and CVD [Lai et al., 2010]. Bilirubin is a metabolite of haem. Though this compound is toxic at high concentrations, it has been demonstrated to have antioxidant properties and to be of potential benefit in a range of diseases [Schwertner and Vitek, 2008]. Genes encoding two enzymes that regulate levels of bilirubin are of interest. First of all, haem oxygenase 1 is the main enzyme which converts haem to bilirubin encoded by *HMOX1*. Three SNPS which tag common haplotypes in *HMOX1* which encodes haem oxygenase 1 have been previously investigated in oxidative stress related disorders [Sheu et al., 2009]. The study found a common haplotype in *HMOX1* to be associated with increased risk of acute respiratory distress syndrome (ADRS). *UGT1A1* encodes the enzyme which conjugates bilirubin by glucuronidation increasing its solubility and therefore promoting excretion. A common TA repeat in *UGT1A1* (*UGT1A1*28*) is associated with reduced UGT1A1 activity and higher than normal levels of bilirubin in plasma (Gilbert's syndrome). Variants of *UGT1A1* are found associated with NAFLD. *UGT1A1*6* but not *UGT1A1*28* has been previously suggested to be a protective factor against NAFLD in children [Lin et al., 2009].

Finally, the NRF1 and NRF2-KEAP1 transcription factors are important in regulating enzymes with ARE such as SOD2 and GCLC and GCLM. Two *NRF1* SNPs have been previously investigated in diabetes patients, however the results did not reach its significance after correcting for multiple-testing [Gaulton et al., 2008]. Yamamoto and colleagues previously identified promoter region polymorphisms in *NRF2* [Yamamoto et al., 2004]. Fukushima-Uesaka and colleagues investigated variations in both *NRF1* and *KEAP1* and inferred haplotypes for both genes [Fukushima-Uesaka et al., 2007]. However recently there has been an explosion of interest mainly in *NRF2* polymorphisms and haplotypes [von Otter et al., 2010;Cordova et al., 2010]. Von Otter and colleagues identified a protective haplotype against Parkinson disease. Cordova and colleagues found a promoter region SNP in *NRF1* to be associated with nephritis.

Of the genes discussed above, only *UGT1A1* has been previously studied in relation to NAFLD susceptibility and that study was confined to children. The aim of the experimental work described in this chapter was to determine whether there was an association between selected polymorphisms in the genes described above (see summary list in Table 4.1) and severity of NAFLD in both our Newcastle NAFLD patients and our Italian family trios. An additional aim was to test for epistatic effects between genes which are biologically suggested to interact with one another.

Table 4.1: Candidate genes relating to oxidative stress chosen for study

Genes
GPX1
GCLC
GCLM
GSR
MAT1A
HMOX1
UGT1A1
NRF1
NRF2
KEAP1

4.2 Methods

4.2.1 SNP selection

Selection of SNPs for the candidate genes listed in Table 4.1 was based on (i) their importance from published literature, (ii) their position within the gene (promoter region, exon/intron boundary), (iii) whether synonymous or non-synonymous) and (iv) whether they were SNPs that tag common (>10% frequency) haplotypes based on Caucasian (CEU) genotypes from the HapMap. The precise approach to SNP selection was gene-dependent. A summary of all SNPs studied is shown in Table 4.1.

4.2.2 Genotyping assays

Most SNPs were genotyped by either PCR-RFLP or allelic discrimination described in detail in chapter 2. The primers, PCR conditions and restriction enzymes used for PCR-RFLP assays are shown in Table 4.2. Assays were all developed in the laboratory except for *GPX1* rs1050450 [Forsberg et al., 1999] and *UGT1A1*28* [Lin et al., 2009] which involved slight modification of the published methods. The PCR-RFLP assays were viewed on acrylamide and agarose gels where appropriate. Genotyping by fluorescent allelic discrimination was performed by K-Biosciences (described section 2.1.6).

*4.2.2.1 UGT1A1*28 assay*

The primers selected and PCR conditions were obtained from [Lin et al., 2009]. The standard PCR method described in section 2.1.2 was used but 5% DMSO was added to the master mix. Cycling conditions were 30 cycles of 1 min at 95 ^{o}C, 1 min 55 ^{o}C annealing and 72 ^{o}C extension for 1 min, terminated by 7 min at 72 ^{o}C resulting in 77 or 75 bp products. PCR products of 77 and 75 were viewed on 13% acrylamide gels to resolve products within 2 bp.

Table 4.2: Primers, PCR conditions and restriction enzymes of PCR-RFLP assays

Assay	Primers (Forward and Reverse)	PCR conditions	Restriction enzyme
GPX1 rs1050450	5'-GCCTGGTGGTGGGTTCGAGCC-3' 5'-GACAGCAGCACTGCAACTGCC-3'	Annealing 64 ^{o}C for 1min	*ApaI*
GCLC rs17883901	5' GGTGAGTTCGTCATTGATTC -3' 5'- ACTTGTGACCAAAACCTGCG -3'	Annealing 55 ^{o}C for 1min	*TaaI*
MAT1A rs17677908	5'- GAAAAGCGAAAGTGCCTGAA-3' 5'- CTCTGAGAAAGAGTTTCTTG-3'	Annealing 53 ^{o}C for 1min	*HpyCH-4V*
MAT1A rs4934027	5'- ACTGCCCCCTCTCTTGTACT -3' 5'- CACCAGTCAATAGTGATGTC -3'	Annealing 55 ^{o}C for 1min	*MspI*
MAT1A rs10788546	5' CAGGACATCCTTCCTGAGTT -3' 5'- TCCTCCACTGCAACAGGAC -3'	Annealing 57 ^{o}C for 1min	*AccI*
MAT1A rs2993763	5'- ACCAGGTGCCTCCAGGGTGAGA-3' 5'- GTGGGGAGGCGTAGTGCCATG-3'	Annealing 65 ^{o}C for 1min	*MscI*
NRF2 rs6721961	5'-GGCTAAAGATTTGGACCCAGAC-3' 5'-TTCCCGTTTTTCTCCCAGCT-3'	Annealing 57 ^{o}C for 1min	*HaeII*
NRF2 rs2364722	5'-GCAATTGAATAAATCTTGGCCTGT-3' 5'-GGAGGACAAGTTTACTAGATGTC-3'	Annealing 57 ^{o}C for 1min	*RsaI*
NRF2 rs1806649	5`- CTCCAGGATATCCTCTCTGA-3` 5`-AGGTATCTACACACCCTAGG-3`	Annealing 57 ^{o}C for 1min	*Bpu10I*
NRF2 rs4243387	5'-CTCAGACCTACACCTTGGCA-3` 5`-ATATCCTGGAGGAACTGCCA-3`	Annealing 57 ^{o}C for 1min	*BstEII*
NRF2 rs10183914	5'-GCTGGTCTGTTTCAGCTTAAGC-3' 5'-CTTGTCAAGGAAACCTTAGCC-3'	Annealing 57 ^{o}C for 1min	*HphI*
UGT1A1*28	5`-TAACTTGGTGTATCGATTGGT-3` 5`-CTTTGCTCCTGCCAGAGGT-3`	Annealing 55 ^{o}C for 1min	*

All PCRs were carried out using 35 cycles of 94^{o}C denaturation for 1 min and 72^{o}C extension for 1 min and all PCRs ended with 7 min at 72^{o}C to ensure elongation.

*PCR conditions 30 cycles assay obtained from [Lin et al., 2009] described 4.2.2.1

4.2.3 Statistical analysis

27 SNPs and 1 repeat from 12 genes (including *SOD2* and *PNPLA3*) were genotyped in both Newcastle NAFLD and the Italian family cohorts and the level of significance was therefore set at 0.002 to correct for the number of SNPs studied (Bonferroni correction). For the Newcastle adult NAFLD cases, the frequency of carriage of the minor allele between cases of mild fibrosis (score 0 or 1) and severe fibrosis (score 2 or higher) were compared. For the Italian families with children suffering from NAFLD, transmission rates between alleles were compared. Logistic regression analysis using the R statistical package was used for analysis of gene-gene interaction.

4.2.4 Haplotype assignments

SNP genotype data for Caucasians for each gene was downloaded from HapMap.org and uploaded onto Haploview version 4.1 [Barrett et al., 2005] to determine the tag SNPs that identify the common (>10%) haplotypes. An input file was created for the Newcastle genotyped data which was then run in PHASE to generate all the possible haplotypes and their frequencies. Frequencies between Newcastle and HapMap Caucasians were then compared as described in section 2.1.7.

4.3 Results:

4.3.1 Candidate gene SNP selection

For GPX1, UGT1A1 and GCLC single SNPs which had been reported to be of clear functional significance were investigated only on the grounds that these genes had already been well studied by others. For GCLM, GSR, MAT1A, HMOX1, NRF1, NRF2 and KEAP1 a tag SNP approach was adopted due, at the time of initiation of these studies, to the limited information on the existence of functionally significant polymorphisms in these genes. The studies on genes involving single SNPs are covered in the first section below with the studies involving tag SNPs described later. For all genes, a description of the assay and genotyping results obtained is provided. In addition, where tag SNPs were genotyped, the frequency of overall haplotypes compared with those reported previously [Gibbs, 2003] is described.

4.3.2 Candidate gene association studies

The overall results obtained for the Newcastle adult NAFLD patients are summarised in Table 4.3 with the results of the TDT analysis on the Italian families shown in Table 4.4. Results for each gene are discussed individually below.

Table 4.3: Candidate genes genotyping distribution in Newcastle NAFLD patients for severity of fibrosis

Gene and SNP	Genotype	Fibrosis 0+1	Fibrosis 2+3+4	Odd ratio for carriage of minor allele (95% CI)	P value for carriage of minor allele
GPX1 rs1050450	CC	91 (0.48)	75 (0.53)	0.85 (0.55 – 1.31)	0.88
	CT	88 (0.46)	61 (0.42)		
	TT	11 (0.06)	8 (0.05)		
GCLC rs17883901	CC	145 (0.77)	110 (0.78)	0.95 (0.56 – 1.6)	0.85
	CT	42 (0.22)	31 (0.22)		
	TT	1 (0.01)	0		
GCLM rs41303970	CC	136 (0.73)	104 (0.72)	1.03 (0.63 - 1.67)	0.92
	CT	40 (0.21)	39 (0.27)		
	TT	11 (0.06)	1 (0.01)		
GCLM rs12140446	AA	89 (0.48)	66 (0.46)	1.1 (0.71 – 1.7)	0.67
	AT	78 (0.42)	67 (0.46)		
	TT	19 (0.1)	12 (0.08)		
GCLM rs7549683	CC	80 (0.43)	57 (0.41)	1.1 (0.7 – 1.7)	0.67
	CA	83 (0.45)	66 (0.47)		
	AA	23 (0.12)	17 (0.12)		
	GG	129 (0.68)	112 (0.75)	0.69 (0.42 –	

GSR rs1002149	GT	52 (0.28)	33 (0.22)	1.12)	0.13
	TT	8 (0.04)	3 (0.03)		
GSR rs3779647	TT	54 (0.28)	50 (0.34)	0.77 (0.48 – 1.23)	0.28
	TC	95 (0.49)	74 (0.49)		
	CC	43 (0.23)	25 (0.17)		
GSR rs8190996	GG	61 (0.32)	50 (0.33)	0.93 (0.59 – 1.47)	0.77
	GA	99 (0.52)	68 (0.46)		
	AA	30 (0.16)	31 (0.21)		
GSR rs2911678	TT	102 (0.53)	94 (0.63)	0.68 (0.44 – 1.05)	0.09
	TA	80 (0.42)	50 (0.33)		
	AA	9 (0.05)	6 (0.04)		
MAT1A rs17677908	AA	147 (0.81)	115 (0.78)	1.2 (0.7 – 2.05)	0.49
	AG	31 (0.17)	31 (0.21)		
	GG	4 (0.02)	2 (0.01)		
MAT1A rs4934027	CC	109 (0.59)	71 (0.52)	1.3 (0.85 – 2.05)	0.22
	CT	68 (0.37)	57 (0.42)		
	TT	9 (0.04)	9 (0.06)		
MAT1A rs10788546	CC	86 (0.47)	63 (0.43)	1.2 (0.76 – 1.82)	0.46
	CT	76 (0.42)	72 (0.49)		
	TT	20 (0.11)	11 (0.08)		

MAT1A rs2993763	GG	54 (0.30)	40 (0.28)	1.1 (0.67-1.79)	0.69
	GA	90 (0.49)	68 (0.48)		
	AA	38 (0.21)	34 (0.24)		
HMOX1 rs2071748	AA	57 (0.3)	52 (0.36)	0.78 (0.49 - 1.24)	0.3
	AG	102 (0.55)	73 (0.5)		
	GG	28 (0.15)	20 (0.14)		
HMOX1 rs2071749	GG	58 (0.32)	55 (0.38)	0.75 (0.48 – 1.2)	0.24
	GA	93 (0.51)	61 (0.42)		
	AA	32 (0.17)	29 (0.2)		
UGT1A1*28	WT (6/6)	72 (0.39)	60 (0.42)	0.89 (0.57-1.39)	0.61
	HET (6/7)	98 (0.53)	69 (0.48)		
	MT (7/7)	15 (0.08)	15 (0.10)		
NRF1 rs1882095	CC	83 (0.45)	49 (0.35)	1.52 (0.96 – 2.4)	0.067
	CT	81 (0.44)	70 (0.5)		
	TT	19 (0.11)	20 (0.15)		
NRF1 rs4731608	GG	147 (0.77)	118 (0.81)	0.79 (0.46 – 1.35)	0.4
	GA	40 (0.21)	27 (0.18)		
	AA	4 (0.02)	1 (0.01)		
NRF1 rs7796553	TT	145 (0.76)	92 (0.63)	1.85 (1.15 – 2.96)	
	TC	42 (0.22)	46 (0.32)		

	CC	4 (0.02)	8 (0.05)		0.01
NRF1 rs10245818	AA	138 (0.75)	93 (0.64)	1.64 (1.02 – 2.68)	0.04
	AG	42 (0.23)	49 (0.34)		
	GG	5 (0.02)	3 (0.02)		
NRF2 rs6721961	CC	145 (0.78)	114 (0.77)	1.08 (0.64 – 1.87)	0.76
	CA	36 (0.2)	32 (0.22)		
	AA	4 (0.02)	2 (0.01)		
NRF2 rs2364722	AA	84 (0.45)	69 (0.49)	0.85 (0.55 – 1.32)	0.48
	AG	96 (0.51)	66 (0.46)		
	GG	8 (0.04)	7 (0.05)		
NRF2 rs1806649	CC	95 (0.51)	70 (0.47)	1.17 (0.76 – 1.81)	0.46
	CT	77 (0.42)	68 (0.46)		
	TT	13 (0.07)	10 (0.07)		
NRF2 rs4243387	TT	147 (0.78)	113 (0.76)	1.1 (0.67 – 1.85)	0.67
	TC	39 (0.21)	32 (0.21)		
	CC	3 (0.01)	4 (0.03)		
NRF2 rs10183914	CC	78 (0.4)	60 (0.4)	1 (0.65 – 1.55)	0.97
	CT	94 (0.49)	73 (0.49)		
	TT	21 (0.11)	16 (0.11)		
KEAP1	CC	69 (0.37)	53 (0.38)	0.96 (0.61 –	0.88

rs1048290	CG	88 (0.47)	69 (0.49)	1.53)	
	GG	29 (0.16)	18 (0.13)		

All variations were in Hardy-Weinberg equilibrium except for NRF2 rs2364722

Table 4.4: TDT in the trio Italian families comparing the distribution of the transmitted and untransmitted allele from heterozygous parentsto affected child

Gene and SNP number	Allele	Transmitted	Untransmitted	p-value
GPX1 rs1050450	C	31 (48.5%)	33 (51.5%)	0.9
	T	33 (51.5%)	31 (48.5%)	
GCLC rs17883901	C	3 (23%)	10 (77%)	0.046
	T	10 (77%)	3 (23%)	
GCLM rs41303970	C	12 (41%)	17 (59%)	0.45
	T	17 (59%)	12 (41%)	
GCLM rs12140446	A	37 (57%)	28 (43%)	0.32
	T	28 (43%)	37 (57%)	
GCLM rs7549683	C	26 (45%)	32 (55%)	0.51
	A	32 (55%)	26 (45%)	
GSR rs1002149	G	20 (55%)	16 (45%)	0.61
	T	16 (45%)	20 (55%)	
GSR rs3779647	T	36 (55%)	30 (45%)	0.54
	C	30 (45%)	36 (55%)	
GSR rs8190996	C	35 (48%)	38 (52%)	0.81
	T	38 (52%)	35 (48%)	
GSR rs2911678	A	27 (64%)	15 (36%)	0.08
	T	15 (36%)	27 (64%)	
MAT1A rs17677908	A	13 (46%)	15 (54%)	0.14

	G	15 (54%)	13 (46%)	
MAT1A rs4934027	C	30 (57%)	23 (43%)	0.41
	T	23 (43%)	30 (57%)	
MAT1A rs10788546	C	29 (54%)	25 (46%)	0.68
	T	25 (46%)	29 (54%)	
MAT1A rs2993763	G	30 (55%)	25 (45%)	0.66
	A	35 (45%)	30 (55%)	
HMOX1 rs2071748	G	33 (51%)	32 (49%)	1
	A	32 (49%)	33 (51%)	
HMOX1 rs2071749	G	29 (46%)	36 (54%)	0.45
	A	36 (54%)	29 (46%)	
UGT1A1*28	-TA	25 (44%)	32 (56%)	0.42
	+TA	32 (56%)	25 (44%)	
NRF1 rs1882095	C	28 (60%)	19 (40%)	0.24
	T	19 (40%)	28 (60%)	
NRF1 rs7796553	T	15 (65%)	8 (35%)	0.21
	C	8 (35%)	15 (65%)	
NRF1 rs4731608	G	11(44%)	14 (46%)	0.69
	A	14 (46%)	11 (44%)	
NRF1 rs10245818	A	15 (65%)	8 (35%)	0.21
	G	8 (35%)	15 (65%)	

NRF2 rs10183914	C	43 (59%)	30 (41%)	0.12
	T	30 (41%)	43 (59%)	
NRF2 rs2364722	A	26 (46%)	30 (54%)	0.59
	G	30 (54%)	26 (46%)	
NRF2 rs6721961	G	21 (60%)	14 (40%)	0.23
	T	14 (40%)	21 (60%)	
NRF2 rs1806649	C	38 (59%)	26 (41%)	0.16
	T	26 (41%)	38 (59%)	
NRF2 rs4243387	T	20 (51%)	19 (49%)	1
	C	19 (49%)	20 (51%)	
KEAP1 rs1048290	C	31 (56%)	24 (44%)	0.42
	G	24 (44%)	31 (56%)	

Percentage of the alleles being transmitted and untransmitted from parents to the affected child

4.3.3 Studies on SNPs of known functional significance

4.3.3.1 GPX1

The non-synonymous rs1050450 polymorphism in *GPX1* at position 537 on exon 2, results in a Proline to Leucine substitution at codon 198 (Pro198Leu). The GPX1 rs1050450 polymorphism was genotyped by PCR-RFLP with a typical result shown in Figure 4.1. The relationship between genotype and severity of fibrosis was analysed (Table 4.3). There was no difference in genotype frequency between mild fibrosis and severe fibrosis (p=0.85). TDT analysis on the Italian families did not show any preferential transmission of either allele to the affected child (p=0.9) (Table 4.4).

4.3.3.2 GCLC

The *GCLC* polymorphism rs17883901 is a C to T substitution located in the promoter region at position -129. The polymorphism was genotyped by PCR-RFLP (see Figure 4.2 for a typical result) and the relationship between genotype and severity of fibrosis was analysed (Table 4.3). There was no difference in carriage of the variant allele between mild fibrosis and severe fibrosis (p=0.85). The TDT on the Italian families showed a borderline significance prior to Bonferroni correction (Table 4.4). The T-allele was found to be transmitted on 10/13 (77%) occasions compared with the C-allele which was transmitted on 3/13 (23%) occasions from both or at least one heterozygous parent (p=0.046).

4.3.3.3 UGT1A1

The TA repeat polymorphism *UGT1A1*28* is located in the promoter region of *UGT1A1*. Repeat was differentiated by length of PCR product (see Figure 4.3). The relationship between genotype and severity of fibrosis was analysed (Table 4.3). There was no significant differences in carriage of the 7-TA repeat allele between the mild fibrosis and severe fibrosis groups (p=0.61). TDT was performed on the Italian families (Table 4.4). There was no preferential transmission of the 7-TA repeat allele to the affected child (p=0.42).

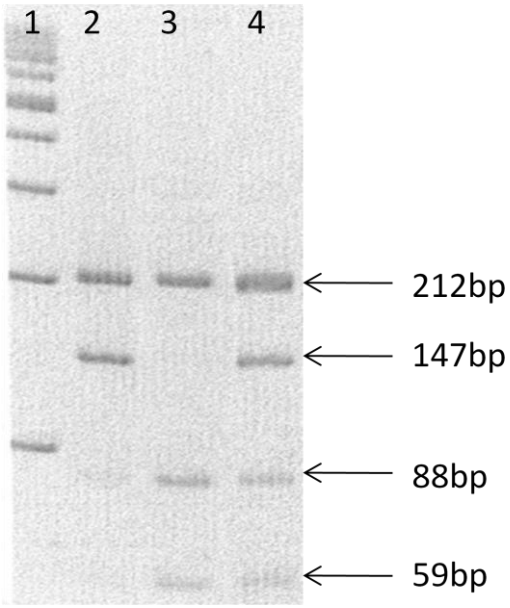

Figure 4.1: 10% acrylamide gel showing GPX1 rs1050450 digested with *ApaI*. Lane 2 shows mutant (TT), lane 3 shows wild type (CC), and lane 4 a heterozygote (CT)

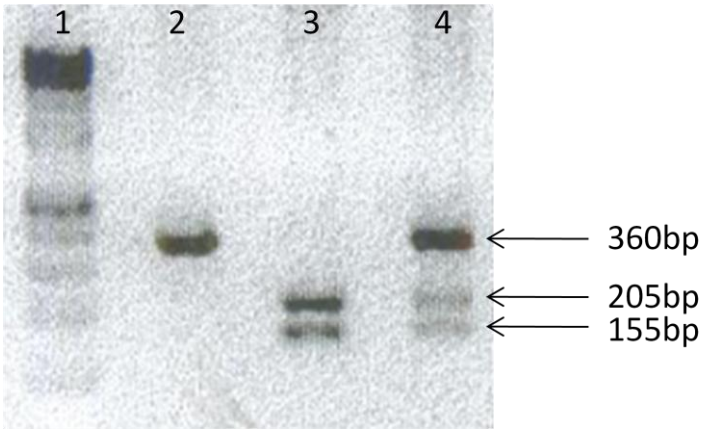

Figure 4.2: 2% agarose gel showing *GCLC* rs17883901 digested with *TaaI*. Lane 2 shows mutant (TT), lane 3 wild type (CC) and lane 4 heterozygous (CT).

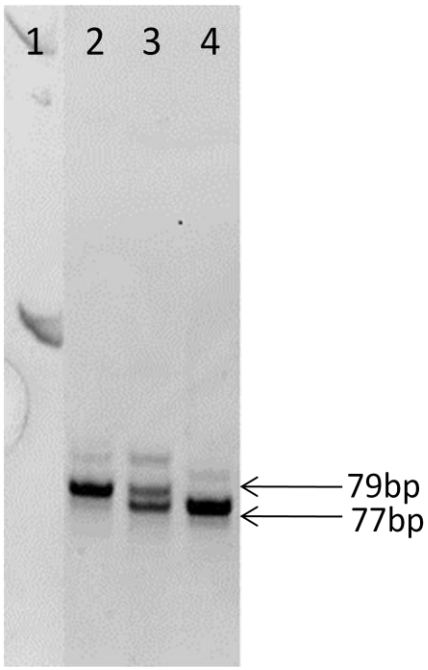

Figure 4.3: 13% acrylamide gel, UGT1A1*28 TA repeat PCR product. Lane 2, shows wild type TA 6/6, lane 3, heterozygous TA 6/7, and lane 4 TA 7/7.

4.3.3.4 GCLM

Three *GCLM* polymorphisms were selected for study. A promoter region SNP (rs41303970), an intronic SNP (rs12140446) and 3`-UTR SNP rs7549683 were selected. Using Haploview and genotype data on Europeans from the HapMap, *GCLM* showed only one haplotype block. The two tag SNPs (rs12140446 and rs7549683) selected were in LD (r^2>0.8) with four SNPs that were predicted to represent the common (>10%) haplotypes of *GCLM* (Figure 4.4). These SNPs were genotyped by allelic discrimination and the relationship between genotype and severity of fibrosis was analysed (Table 4.3). None of these polymorphisms showed any significant differences in carriage of the variant allele between the mild fibrosis and severe fibrosis groups (p=0.92, p=0.67, and p=0.67 respectively). TDT was performed on the Italian families (Table 4.4). There was no preferential transmission of any allele to the affected child for any of the 3 SNPs studied (p=0.45, p=0.32, and p=0.51 respectively). There was no HapMap data available on the upstream promoter region SNP and therefore LD between the upstream SNP and the tag SNPs selected was not available. Frequencies of common haplotypes were generated using the two tag SNPs (Table 4.5). There were no differences between frequencies of haplotypes generated compared against frequencies from HapMap CEU (p=0.24).

4.3.3.5 GSR

Four *GSR* polymorphisms were selected for genotyping. A promoter region SNP rs1002149, and three intronic region SNPs rs3779647, rs8190996 and rs2911678 were selected (Figure 4.5). Using Haploview, *GSR* consisted of three haplotype blocks. Each block represented a region of the GSR gene (Figure 4.5). The three tag SNPs (rs3779647, rs8190996 and rs2911678) selected were in LD (r^2>0.8) with six SNPs that were predicted to represent the common (>10%) haplotypes (Figure 4.4). These SNPs were genotyped by allelic discrimination and the relationship between genotype and severity of fibrosis was analysed (Table 4.3). None of these polymorphisms showed any significant differences in carriage of the variant allele between the mild fibrosis and severe fibrosis groups (p=0.13, p=0.28, p=0.77 and p=0.09 respectively). TDT was performed on the Italian families (Table 4.4). There was no preferential transmission of any allele to the affected child for any of the 4 SNPs studied (p=0.61, p=0.54, p=0.81 and p=0.08 respectively). Frequencies of common haplotypes were generated using the three tag SNPs (Table 4.6). There were no differences between frequencies of haplotypes generated compared against frequencies from HapMap CEU (p=0.95).

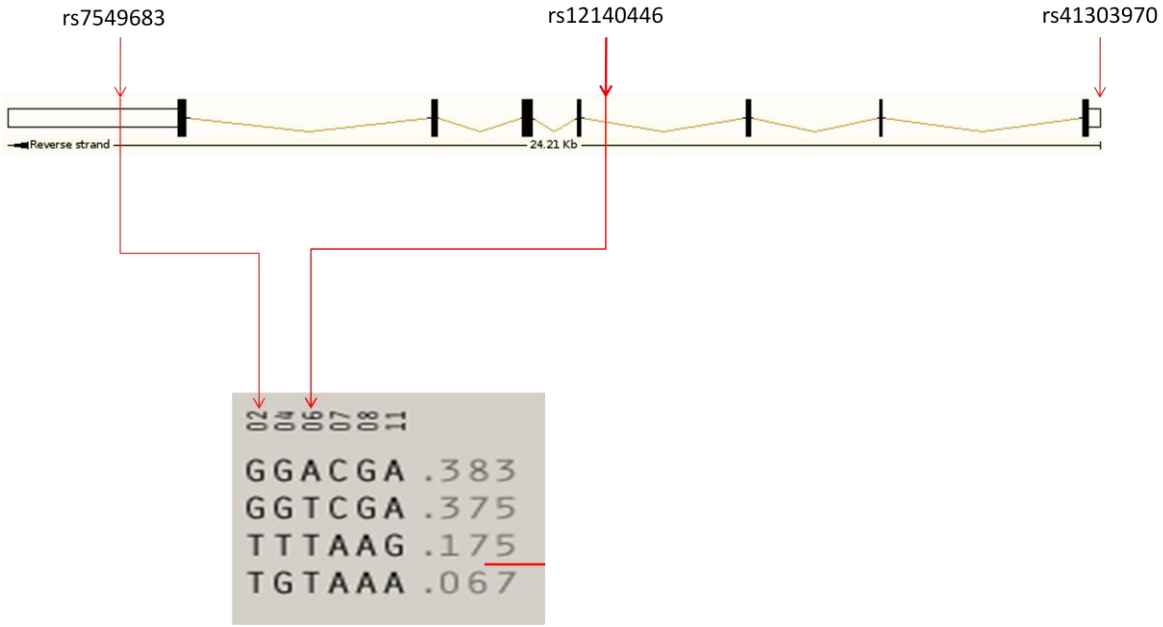

Figure 4.4: Location of rs41303970, rs12140446 (intron region) and rs7549683 (3' UTR) SNPs in the *GCLM* gene. Haplotypes were obtained from Haploview based on European data from the HapMap. Four haplotypes with their frequencies are shown. Three were common (>10%) haplotypes. Two SNPs indicated with arrows were sufficient to tag these common haplotypes.

Table 4.5: GCLM Haplotype frequencies compared in both Newcastle and HapMap

	Newcastle frequency	HapMap frequency	P-value
GCLM			
TT	0.349	0.242	0.24
GT	0.337	0.375	
GA	0.309	0.383	
TA	0.002	NA	

Chi-square test analysing the differences between frequencies of the Newcastle and HapMap populations

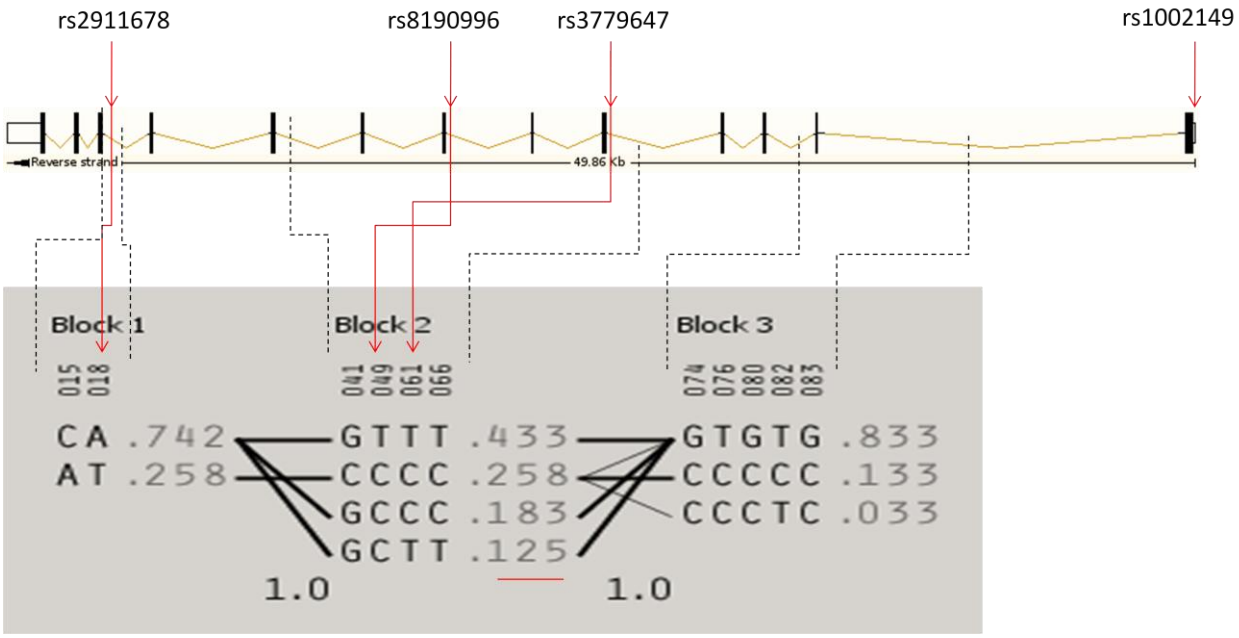

Figure 4.5: Four polymorphisms were selected and their position in the *GSR* gene is illustrated. Three haplotype blocks were identified using Haploview. Each haplotype block covers a region of the *GSR* gene represented by dotted lines. The dark lines between blocks indicate the strength of LD between haplotypes. The r^2 numbers (1.0) between blocks indicates LD between alleles from different blocks. Three tag SNPs illustrated with red arrows were selected to identify the common (>10%) haplotypes. Rs2911678 (018) was in complete LD (r^2=1) with rs2253409 (041).

Table 4.6: GSR Haplotype frequencies comparing both Newcastle and HapMap

	Newcastle frequency	HapMap frequency	P-value
GSR			
ATT	0.4	0.43	0.95
TCC	0.23	0.25	
ACC	0.19	0.183	
ACT	0.14	0.12	
ACT	0.02	NA	
TTT	0.003	NA	

Chi-square test analysing the differences between frequencies of the Newcastle and HapMap populations

4.3.3.6 MAT1A

Four polymorphisms were selected for *MAT1A*. A promoter region SNP rs17677908, an intronic SNP rs4934027, and two non-synonymous SNPs rs10788546 and rs2993763 were selected (Figure 4.6). Using Haploview, *MAT1A* consisted on three haplotype blocks, each block representing a region of the *MAT1A* gene (Figure 4.6). Three tag SNPs (rs4934027, rs10788549 and rs2993763) selected were in LD (r^2> 0.8) with 16 SNPs across the haplotype blocks that were predicted to represent the common (>10%) haplotypes (Figure 4.6). These SNPs were genotyped by PCR-RFLP (see Figures 4.7 for typical results). The relationship between genotype and severity of fibrosis was analysed (Table 4.3). None of these polymorphisms showed any significant differences in carriage of the variant allele between the mild fibrosis and severe fibrosis groups (p=0.49, p=0.22, p=0.46 and p=0.69 respectively). TDT was performed on the Italian families (Table 4.4). There was no preferential transmission of any allele to the affected child for any of the 4 SNPs studied (p=0.14, p=0.41, p=0.68 and p=0.69 respectively). The upstream promoter region SNP was not among the HapMap data SNPs available and therefore no LD was available with the selected tag SNPs. Frequencies of common haplotypes were generated using the three tag SNPs (Table 4.7). There were no differences between frequencies of haplotypes generated compared against frequencies from HapMap CEU (p=0.47).

4.3.3.7 HMOX1

Two SNPs rs2071748 and rs2071749 in the intron region were selected for *HMOX1* (Figure 4.8). Using Haploview, *HMOX1* consisted of one haplotype block. The two tag SNPs (rs2071748 and rs2071749) selected were in LD (r^2>0.8) with three SNPs across the haplotype block and were predicted to represent the common (>10%) haplotypes (Figure 4.8). These SNPs were genotyped by allelic discrimination and the relationship between genotype and severity of fibrosis was analysed (Table 4.3). None of these polymorphisms showed any significant differences in carriage of the variant allele between the mild fibrosis and severe fibrosis groups (p=0.3, and p=0.24 respectively). TDT was performed on the Italian families (Table 4.4). There was no preferential transmission of any allele to the affected child for any of the two SNPs studied (p=1 and p=0.45 respectively). Frequencies of common haplotypes were generated using the two tag SNPs (Table 4.8). There were no differences between frequencies of haplotypes generated compared against frequencies from HapMap CEU (p=0.27).

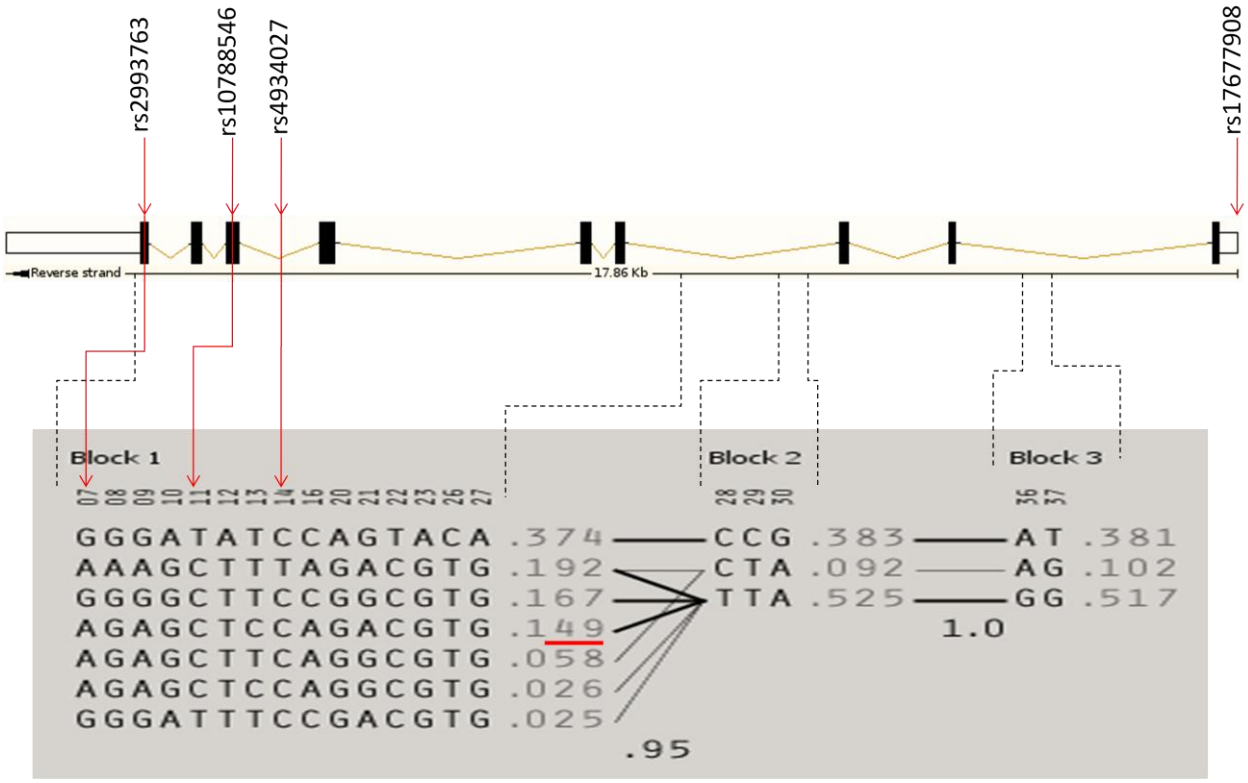

Figure 4.6: Four polymorphisms were selected and their position on the *MAT1A* gene is illustrated. Three *MAT1A* haplotype blocks identified, each block covers a region of the *MAT1A* gene represented with the dotted lines. The dark lines between blocks indicate the strength of LD between haplotypes. The r2 numbers (0.95 and 1.0) between blocks indicates LD between alleles from different blocks. Three tag SNPs illustrated with red arrows were selected to identify the common (>10%) haplotypes.

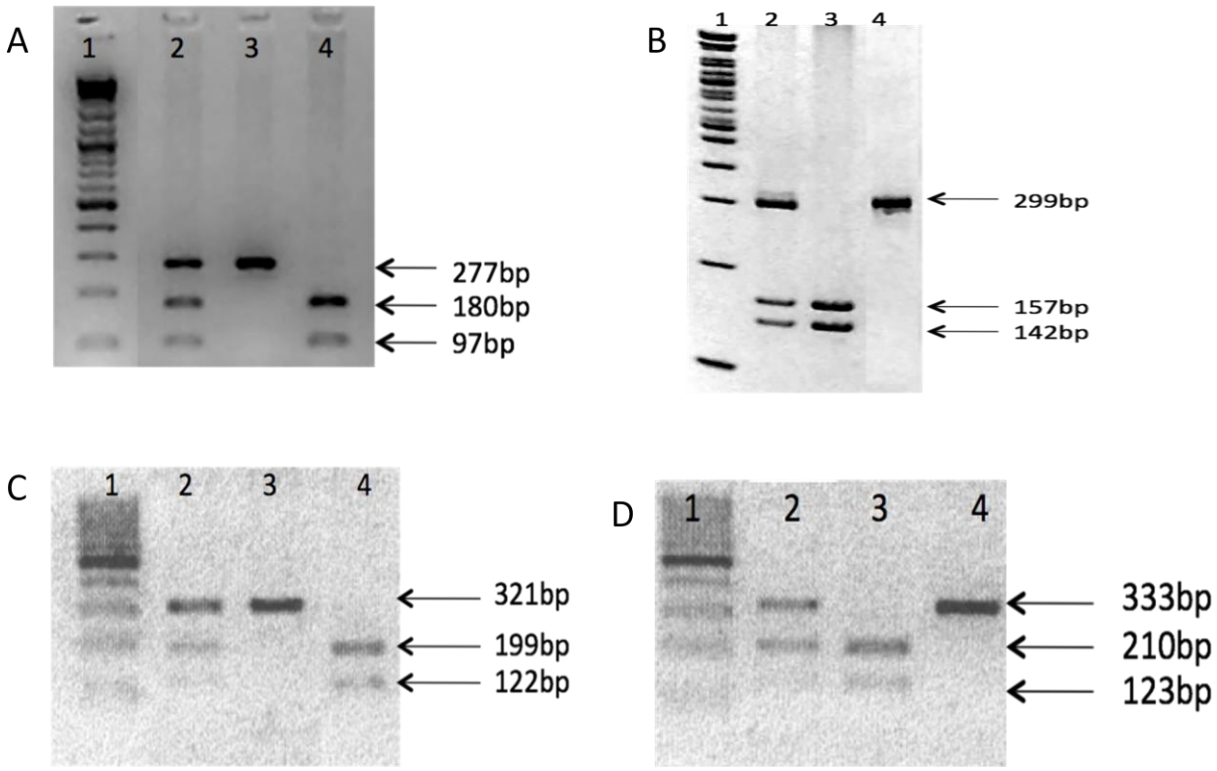

Figure 4.7: 2% agarose gel showing *MAT1A* selected SNPs A) rs2993763 digested with *MscI*. Lane 2 shows heterozygous (GA), lane 3 Wild type (GG), and lane 4 mutant (AA). B) 10% acrylamide rs17677908 digested with *HpyCH-4V*. Lane 2 shows heterozygous (AG), lane 3 wild type (AA), and lane 4 mutant (GG). C) 2% agarose gel showing *MAT1A* rs4934027 digested with *MspI*. Lane 2 shows, heterozygous (GA), lane 3 mutant (AA), and lane 4 wild type (GG). D) 2% agarose gel showing *MAT1A* rs10788546 digested with *AccI*. Lane 2 shows heterozygous (CT), lane 2 mutant (TT) and lane 4, wild type (CC).

Table 4.7: MAT1A Haplotype frequencies comparing both Newcastle and HapMap

	Newcastle frequency	HapMap frequency	P-value
MAT1A			
GTC	0.309	0.374	
ACT	0.24	0.192	
ACC	0.232	0.149	
GCC	0.206	0.167	0.47
GCT	0.008	NA	
GTT	0.003	NA	
ATT	0.002	NA	

Chi-square test analysing the differences between frequencies of the Newcastle and HapMap populations

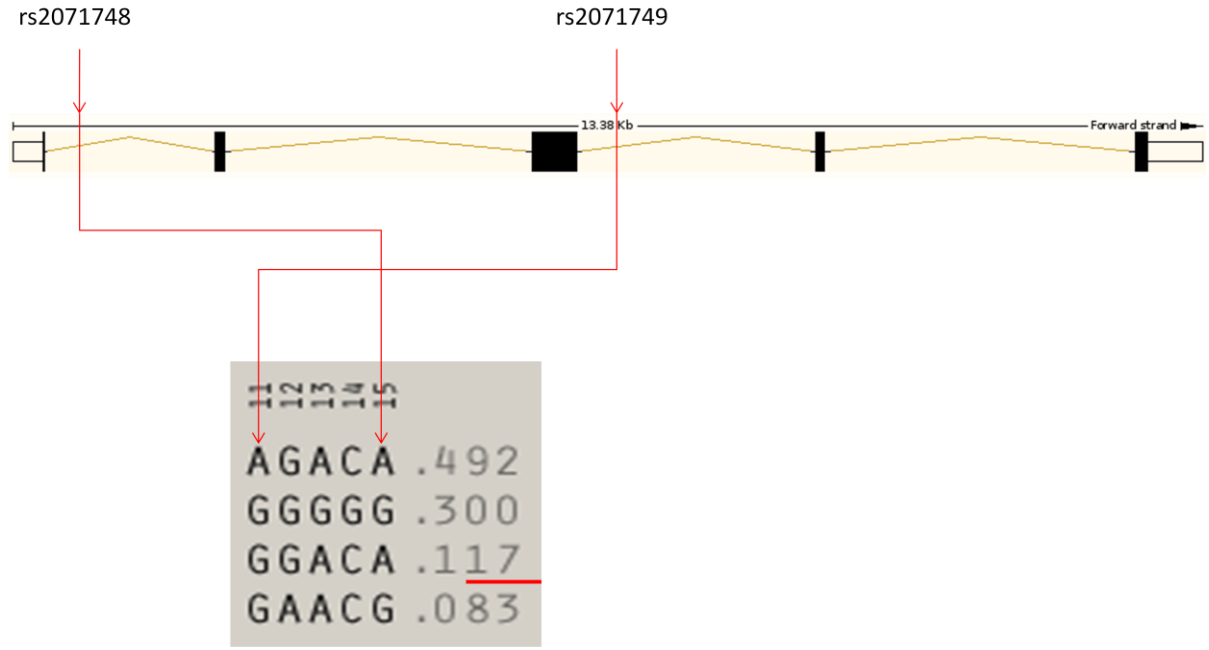

Figure 4.8: Location of the two tag SNPs rs2071749 (intronic) and rs2071748 (intronic) in the *HMOX1* gene structure. Haplotypes obtained from Haploview showing 4 haplotypes with their frequencies, only 3 were common (>10%) haplotypes. Two tag SNPs indicated with arrows were sufficient to determine these common haplotypes.

Table 4.8: HMOX1 Haplotype frequencies comparing both Newcastle and HapMap

	Newcastle frequency	HapMap frequency	P-value
HMOX1			
AA	0.418	0.492	0.27
GG	0.408	0.3	
GA	0.165	0.117	
GG	0.009	0.083	

Chi-square test analysing the differences between frequencies of the Newcastle and HapMap populations

4.3.3.8 NRF1

Four polymorphisms were selected for *NRF1*. One promoter region SNP rs1882095 and three intron region SNPs rs7796553, rs4731608 and rs10245818 were selected (Figure 4.9). Using Haploview, NRF1 consisted of 8 haplotype blocks. Each block representing a region of *NRF1* gene (Figure 4.9). The two tag SNPs (rs4731608 and rs10245818) selected were in LD (r^2>0.8) with 20 SNPs across the haplotype blocks which were predicted to determine the common (>10%) haplotypes (Figure 4.9). These SNPs were genotyped by allelic discrimination and the relationship between genotype and severity of fibrosis was analysed (Table 4.3). Two SNPs (rs7796553 and rs10245818) were associated with severity of fibrosis. Carriage of the C-allele of rs7796553 (OR 1.85 (1.15-2.96), p=0.01) was found associated with severity of fibrosis. Carriage of G-allele of rs10245818 (OR 1.64 (1.02-2.68), p=0.04) was found associated with severity of fibrosis. Both results were prior to Bonferroni correction. None of the remaining SNPs (rs4731608 and rs1882092) showed any significant differences in carriage of the variant allele between the mild fibrosis and severe fibrosis groups (p=0.4 and p=0.067 respectively). TDT was performed on the Italian families (Table 4.4). There was no preferential transmission of any allele to the affected child for any of the 4 SNPs studied (p=0.24, p=0.21, p=0.69 and p=0.21 respectively). The upstream promoter region SNP was not found in HapMap SNP data and therefore no LD with the other selected SNPs was available. However rs7796553 was selected from the literature and was found to be in LD (r^2= 0.86) with rs10245818. Frequencies of common haplotypes were generated using the two tag SNPs (Table 4.9). There were no differences between frequencies of haplotypes generated compared against frequencies from HapMap CEU (p=0.65).

4.3.3.9 NRF2

Five polymorphisms were selected for NRF2. A promoter region SNP rs6721961 and 4 intron region SNPs rs2364722, rs1806649, rs4243387 and rs10183914 were selected (Figure 4.10). Using Haploview, NRF2 consisted of one haplotype block. Four tag SNPs (rs2364722, rs1806649, rs4243387 and rs10183914) selected were in LD (r^2>0.8) with 10 SNPs found across the haplotype block that were able to determine the common (>10%) haplotypes (Figure 4.10). These SNPs were genotyped by PCR-RFLP (see Figure 4.11 for typical results) and rs4243387 by allelic discrimination. The relationship between genotype and severity of fibrosis was analysed (Table 4.3). None of these polymorphisms showed any significant differences in carriage of the variant allele between the mild fibrosis and severe fibrosis groups (p=0.76, p=0.48, p=0.46 p=0.67

and p=0.97 respectively). TDT was performed on the Italian families (Table 4.4). There was no preferential transmission of any allele to the affected child for any of the 4 SNPs studied (p=0.23, p=0.59, p=0.16, p=1 and p=0.12 respectively). The upstream promoter region SNP was not found in HapMap SNP data and therefore no LD with the other selected SNPs was available. Frequencies of common haplotypes were generated using the three tag SNPs (Table 4.10). There were no differences between frequencies of haplotypes generated compared against frequencies from HapMap CEU (p=0.32).

4.3.3.10 KEAP1

One polymorphism was selected for *KEAP1*. A synonymous SNP rs1048290 was selected (Figure 4.12). Using Haploview, KEAP1 consisted of two haplotype blocks. Each block represented a region of the *KEAP1* gene. One tag SNP (rs1048290) selected was in LD (r^2>0.9) with one SNP that was able to determine the common (>10%) haplotype (Figure 4.12). The polymorphism was genotyped by allelic discrimination. There was no differences in carriage of the variant allele between the mild fibrosis and severe fibrosis groups (p=0.88) (Figure 4.3). TDT was performed on the Italian families (Table 4.4). There was no preferential transmission of any allele to the affected child (p=0.42). There were no differences between frequencies of haplotypes generated compared against frequencies from HapMap CEU (p=0.8).

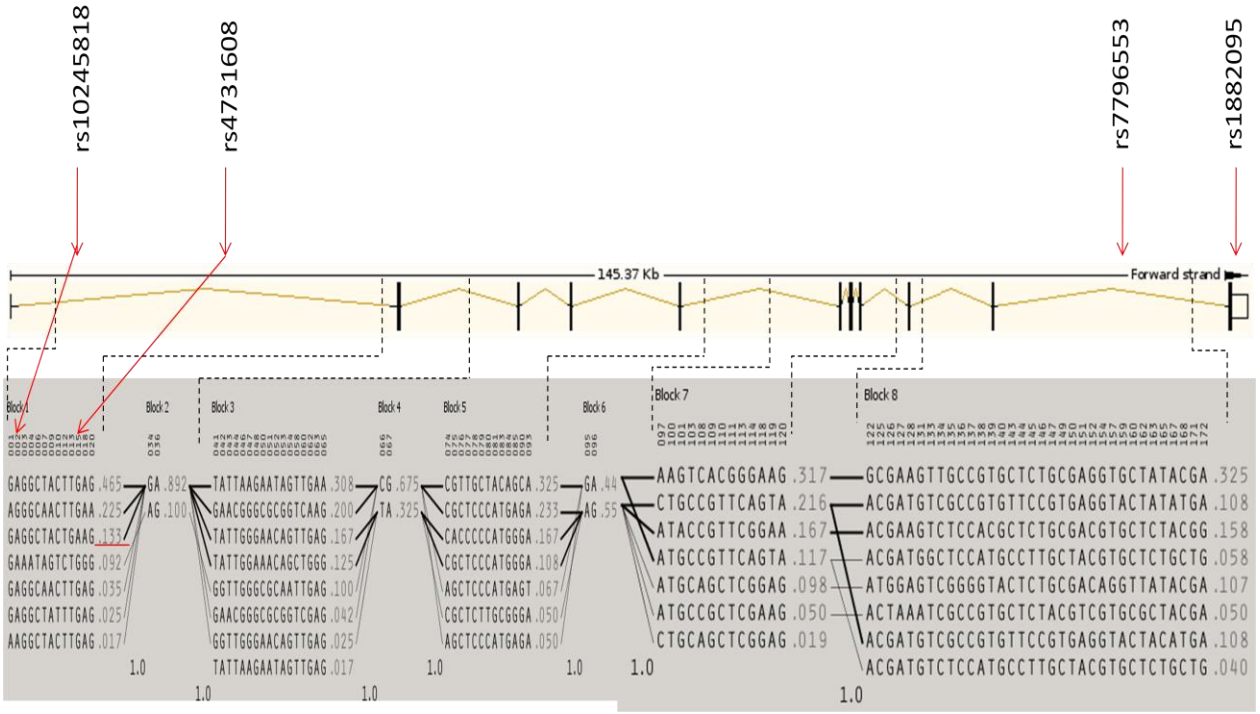

Figure 4.9: Four polymorphisms were selected and their positions are illustrated in the *NRF1* gene. Haploview identified 8 haplotype blocks in NRF1. Each block covers a region of the NRF1 gene represented with dotted lines. The dark lines between the blocks indicate strength of the LD between haplotypes from different blocks. The r^2 number (1.0) between blocks indicates complete LD between alleles across different blocks. Two tag SNPs illustrated with red arrows in block one were selected to identify the common (>10%) haplotypes.

Table 4.9: NRF1Haplotype frequencies comparing both Newcastle and HapMap

	Newcastle frequency	HapMap frequency	P-value
NRF1			
AG	0.73	0.465	0.65
GG	0.145	0.225	
AA	0.097	0.133	
GA	0.018	NA	

Chi-square test analysing the differences between frequencies of the Newcastle and HapMap populations

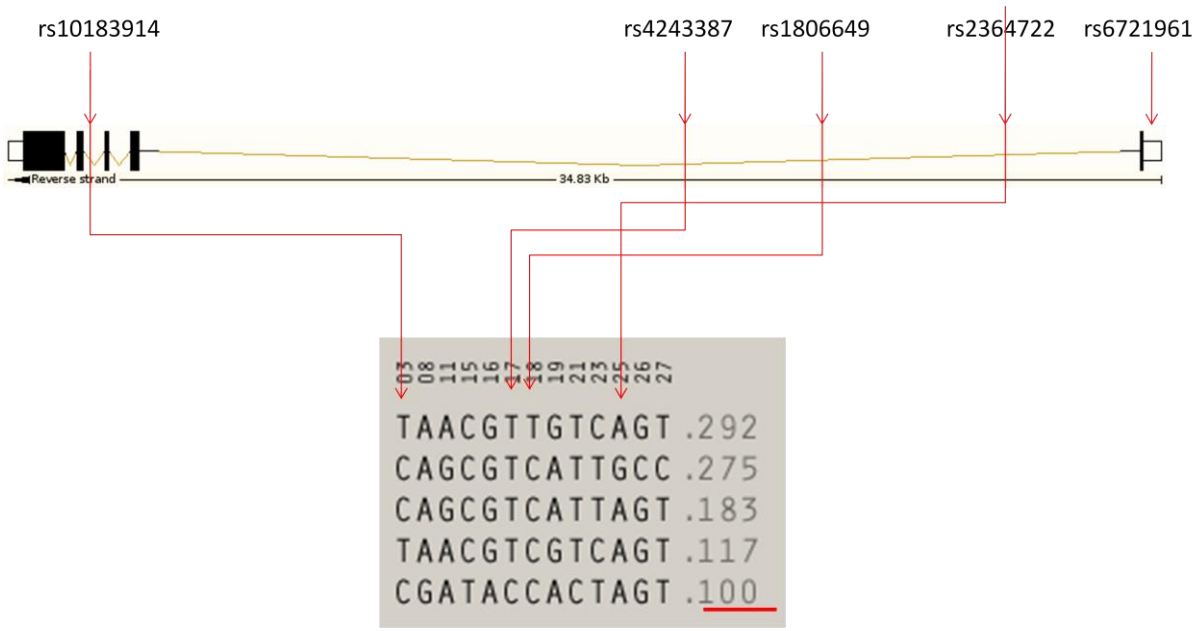

Figure 4.10: Five polymorphisms were selected for NRF2 and their positions in the *NRF2* gene are illustrated. Haplotypes obtained from Haploview showing 5 Haplotypes with their frequencies, all 5 were common (>10%) haplotypes. Four tag SNPs indicated with arrows were sufficient to determine these common haplotypes.

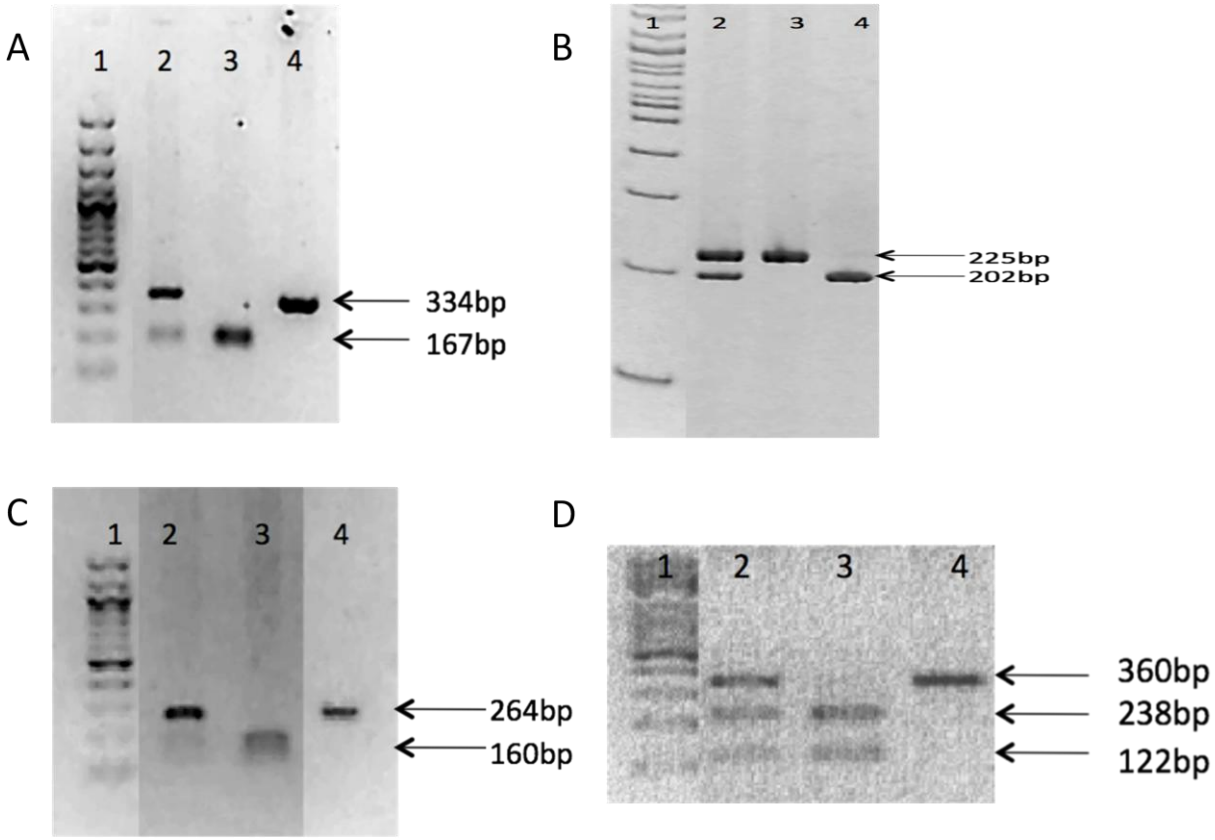

Figure 4.11: A) 2% agarose gel showing *NRF2* rs10183914 digested with *HphI*. Lane 2 showing heterozygous (CT), and lane 3 wild type (CC) and lane 4 mutant (TT). B) 10% acrylamide gel showing NRF2 rs2364722 digested with *RsaI*. Lane 2 shows heterozygous (AG), lane 3 mutant (GG) and lane 4 wild type (AA). C) 2% agarose gel showing *NRF2* rs6721961 digested with *HaeII*. Lane 2 shows heterozygous (CT), lane 3 wild type (CC) and lane 4 mutant (TT). D) 2% agarose gel showing *NRF2* rs1806649 digested with *Bpu10I*. Lane 2 shows heterozygous (CT), lane 3 wild type (CC) and lane 4 mutant (TT).

Table 4.10: NRF2 Haplotype frequencies comparing both Newcastle and HapMap

	Newcastle frequency	HapMap frequency	P-value
NRF2			
TTTA	0.235	0.292	
CTCG	0.265	0.275	
CTCA	0.227	0.183	0.32
TTCA	0.087	0.117	
CCCA	0.117	0.100	

Chi-square test analysing the differences between frequencies of the Newcastle and HapMap populations

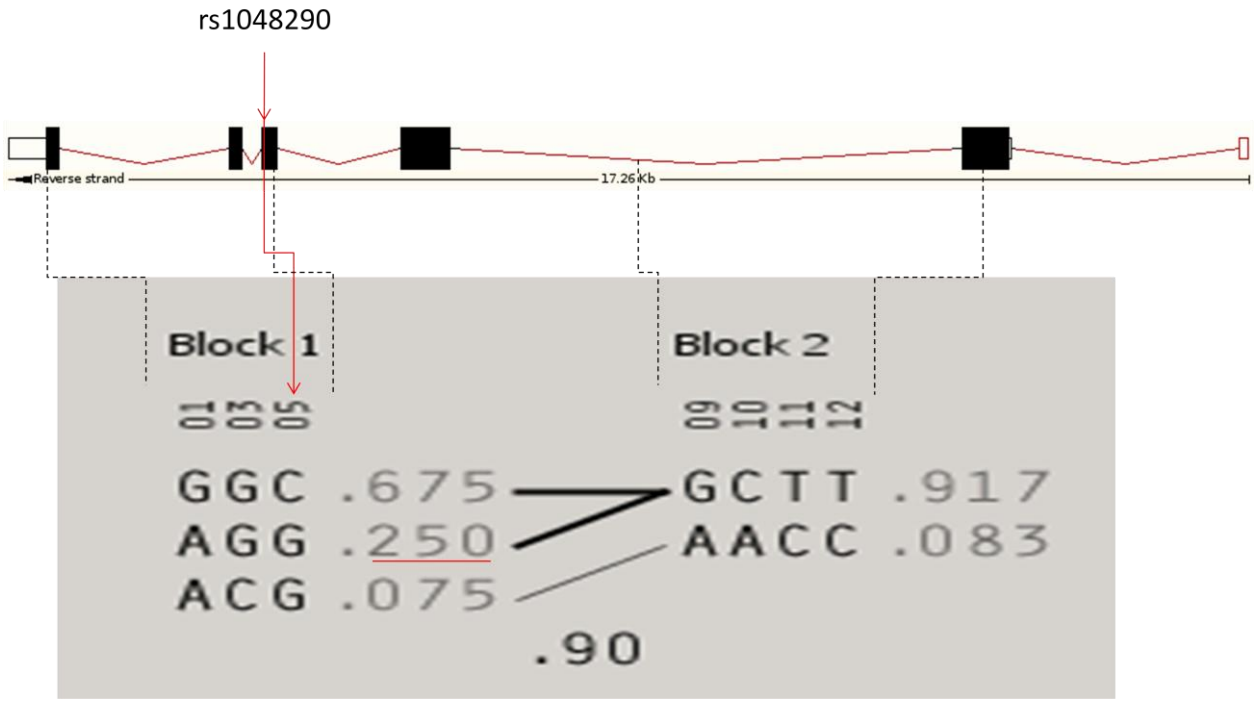

Figure 4.12: One SNP was selected for *KEAP1* and its position in *KEAP1* Gene is illustrated. Haplotypes obtained from Haploview showing 3 haplotypes with their frequencies, only 2 were common (>10%) haplotypes. One tag SNP indicated with arrow was sufficient to determine these common haplotypes.

4.3.4 Gene-Gene interaction

SNPs were selected based on their apparently significant functional effects in published studies and were analysed for gene-gene interaction. *SOD2*, *GPX1*, *GCLC* and *GCLM* SNPs were analysed. Interaction was examined between two SNPs at one time (*SOD2* and *GPX1*), (*SOD2* and *GCLC*), (*SOD2* and *GCLM*), (*GPX1* and *GCLC*), (*GPX1* and *GCLM*) and (*GCLC* and *GCLM*). Of the six interactions performed only *SOD2* with *GCLC* and *SOD2* with *GCLM* were significant (Table 4.11).

The results for *SOD2* and *GCLC* showed a significant interaction (P for interaction = **0.0008**) (Table 4.4). In table 4.11 the *GCLC* CC and the *SOD2* Ala/Ala were set as the base (OR = 1) for comparison with the other genotype combinations. The highest risk for having the disease is seen in individuals with the *SOD2* Ala/Ala and *GCLC* CT/TT with an OR of 5.5 (95% CI 1.62-18.5). Individuals with *SOD2* Ala/Val or Val/Val and *GCLC* CC also have a higher risk (OR = 2.565 (95% CI 1.36-4.9)). Individuals with low enzyme activity *SOD2* Ala/Val or Val/Val and low promoter activity *GCLC* CT or TT appeared to show a slightly lower risk but this may not be different to the overall risk for possession of the val-*SOD2* (OR = 1.5 (95% CI 0.7- 3.3)).

The results for *SOD2* and *GCLM* showed a significant interaction (p for interaction **0.003**) (Table 4.11). In table 4.11 *SOD2* Ala/Ala and *GCLM* CC were set as the base (OR=1) to compare with the other combined groups. The *SOD2* (Ala/Ala) *GCLM* (CT+TT) combination appeared to be protective against disease (OR=0.1(95% CI 0.03-0.78)) with the highest risk found in the *SOD2* (Ala/Val) and *GCLM* (CT+TT) group (OR=1.5 (95% CI 0.8-3.2)).

Table 4.11: Interaction between variants in *SOD2* gene and *GCLC* and *GCLM* genes in relation to severity of fibrosis (Mild vs. Severe)

Genotype	SOD2 rs4880 Ala/Ala (CC) OR 95% CI	SOD2 rs4880 Ala/Val + Val/Ala (CT+TT) OR 95% CI	P for interaction
GCLC rs17883901			
CC	1	5.5 (1.6-18.5)	**0.0008**
CT+TT	2.5 (1.3-4.9)	1.5 (0.7-3.3)	
GCLM rs41303970			
CC	1	0.1 (0.03-0.8)	**0.003**
CT+TT	1.1 (0.6-2.04)	1.5 (0.8-3.2)	

P-value for overall interaction of the different groups compared to the base group (OR =1)

4.4 Discussion

We examined the association of 25 SNPs and one TA repeat in 10 genes involved in defence against ROS. Between the 26 polymorphisms none appeared to be associated with severity of NAFLD in the Newcastle and Italian family samples after Bonferroni correction. Prior to the Bonferroni correction three SNPs, GCLC C-129T (rs17883901) and NRF1 rs7796553 and rs10245818 appeared to show a borderline significance.

A non-synonymous Pro198Leu polymorphism was investigated in the *GPX1* gene that encodes for the enzyme GPX1 involved in detoxifying the product (H_2O_2) of SOD2 into H_2O. Results showed no significance in both the Newcastle and Italian family samples. In contrast, Kuzuya and colleagues study on 2233 randomly selected, community-dwelling, middle-aged and older Japanese people found the CT and TT genotypes of the Pro198Leu polymorphism associated with anthropometric and metabolic variables such as the waist-hip ratio (WHR), triacylglycerol concentrations, immunoreactive insulin (IRA), homeostasis model assessment of insulin resistance (HOMA-IR) and body fat mass. The study also found men with high WHR and waist circumference without any difference in BMI or body fat mass suggests that CT and TT to be associated with central obesity [Kuzuya et al., 2008]. Such a suggestion agrees with the hypothesis from Farrell and Larter which suggests that many individuals that are not obese based on BMI have central obesity which had a stronger association with NASH than overall obesity [Farrell and Larter, 2006]. [Kuzuya et al., 2008] found an association with higher prevalence of metabolic syndrome (defined according to International Diabetes Federation and National Education Cholesterol Program criteria) in men. The study was done on Japanese population and on 2233 samples which may be the reason for such difference with our study [Kuzuya et al., 2008]. Sutton and colleagues used combinations of *SOD2* and *GPX1* genotypes and found that ALD patients with two val-*SOD2* alleles and two pro-*GPX1* alleles had a lower risk for developing HCC [Sutton et al., 2006]. However such a result can be a chance observation due to the small sample size (n=162) in this study [Sutton et al., 2006]. Our gene-gene study did not show an interaction between both the *SOD2* and *GPX1* variants which may be due to the sample size studied for gene-gene interaction was relatively small (n=330) for such type of analysis. However, though some patients with severe NAFLD may develop HCC in the future, it was not possible to do a comparable study in the present patient cohort.

We investigated two common variations in the genes encoding for the subunits (GCLC and GCLM) of GCL, the rate limiting enzyme of GSH. There was no association found

between the *GCLC* C-129T and severity of NAFLD in the Newcastle samples. However a borderline association was found in the Italian family samples. The T-allele was found to be preferentially transmitted from the parents to the effected children. Such conflicting results between the two cohorts could be a spurious result due to chance. Spurious results are commonly found in genetic association studies which require the need of replication studies [Hong et al., 2004]. A possible reason could also be due to the difference in the region between the two cohorts (UK and Italy) and therefore differences in environmental factors could mean that all factors affecting disease are not the same. In addition, NAFLD in children may have some differences to NAFLD in adults [Schwimmer et al., 2005]. Ideally, we would have performed the TDT studies in trios involving the UK-based adults NAFLD patients and their parents but it is very difficult to find such trios due to the typical NAFLD patients being 50 to 60 years old. The selection of extreme phenotypes has shown to be a powerful study design for stronger association studies [Zhang et al., 2006]. The children selected for our study with fibrotic NAFLD might exhibit such extreme phenotype leading to the explanation of this association. However, recently the T-allele was found to be associated with NASH compared to simple steatosis in 131 patients from Brazil [Oliveira et al., 2010]. This association found in Brazil and not in Newcastle cohorts could be due to difference in region and sample size between the two studies and that a larger cohort is required for replication. The association found in the Brazil study is consistent with the T-allele being associated with lower promoter activity [Koide et al., 2003], but this still remains debatable [Butticaz et al., 2011]. No association was found between the *GCLM* C-588T promoter polymorphism and severity of NAFLD. The C-588T polymorphism was reported to be associated with increased risk of myocardial infarction which was consistent with the T-allele showing 40% to 50% lower promoter activity compared with the C-allele [Nakamura et al., 2002].

Similarly to our other findings discussed above, we found no association between SNPs in the *GSR, MAT1A, HMOX1* and *UGT1A1* and severity of NAFLD. Our genotyping on the TA repeat variant of *UGT1A1*28* was not found to be associated to NAFLD severity which was consistent with the study on 234 Taiwanese children with NAFLD [Lin et al., 2009]. Likewise, we did not find any association with genes encoding for transcription factors NRF1 and NRF2-KEAP1, although prior to Bonferroni correction two SNPs rs10245818 and rs7796553 which were found to be in LD (r^2=0.86) in the *NRF1* gene showed a borderline association with NAFLD severity. This was consistent

with [Gaulton et al., 2008] study, where rs7796553 was found to be associated with susceptibility to type 2 diabetes prior to Bonferroni correction.

Though the various tag SNPs we selected did not show any association with severity of NAFLD, the common haplotypes they tag were the same common haplotypes identified by samples from HapMap CEU. There were no significant differences in frequencies of haplotypes between the Newcastle and HapMap CEU cohorts. Ideally, control cohorts from the UK should also have been used for comparison but it was not feasible to obtain a control population without NAFLD.

Complex diseases involve multiple genes and their interactions. Epistasis is defined as one gene (epistatic gene) masking the effect of another (hypostatic gene) which may result in determining a specific phenotype [Van Steen, 2011;de Visser et al., 2011;Cordell, 2002]. There are two types of epistasis; synergistic (increase activity) and antagonist (suppressing activity). There have been studies of epistasis in NAFLD related diseases such as diabetes and obesity [Neuman et al., 2010;Ochoa et al., 2004]. In type 2 diabetes patients gene-gene interaction resulted in higher risk of disease between four genes with prior evidence of association with diabetes [Neuman et al., 2010]. Gene-gene interaction was observed between two genes influencing fat accumulation and found in the aetiology of obesity [Ochoa et al., 2004]. The interest in gene-gene interaction has increased in the past decade with the availability of large data sets and statistical software available to analyse such interaction [Cordell, 2009]. We have therefore examined the interactions between four common SNPs suggested to have a functional role in *SOD2*, *GPX1*, *GCLC* and *GCLM*. Between the six interactions studied only the combined *SOD2* and *GCLC*, and *SOD2* and *GCLM* showed an interaction. Our gene-gene results supports and is consistent with the biological mechanisms involving the enzymes SOD2, GPX1 and GCLC which are involved in O_2^- being dismutated by SOD2 into H_2O_2 and then detoxified by GPX1 and its co-factor GSH into H_2O. Oestergaard and colleagues studied the interaction of 10 genes encoding proteins in the antioxidant defence system and breast cancer [Oestergaard et al., 2006]. *SOD2* and *GPX1* were among the genes studied for the gene-gene interaction on 2271 cases and 2280 controls. However there was no evidence of interaction among the 10 genes studied. The authors suggested that such lack of association may be due to the lack of statistical power. Since our study is considerably smaller than that from [Oestergaard et al., 2006], it is clear that our power to detect gene-gene interactions is also limited and our positive findings could be chance observations. Replication on a

larger sample size is required to confirm the interactions. However, our finding does appear biologically plausible and it seems more likely that oxidative stress genes could be important in overall susceptibility to advanced NAFLD compared with breast cancer.

In summary, we failed to see any strong associations between NAFLD severity and the genotypes described in this Chapter. Lack of association may be due to the limited number of genes that had functional significant polymorphisms available to assess. The information available on the majority of genes selected for this study was limited and therefore lead to a tag SNP haplotype approach. However, we must not ignore the involvement of oxidative stress and its related genes in susceptibility to NAFLD. There are more polymorphisms to be investigated in oxidative stress related genes however due to financial restrictions it was not possible to study all. Therefore a larger scale study such as GWAS on NAFLD that includes oxidative stress genes is needed.

Chapter 5

Investigation of the functional significance of the GCLC C-129T polymorphism

5 Investigation of the functional significance of the GCLC C-129T polymorphism

5.1 Introduction

GSH, as explained previously (section 1.7.1.2), is protective during oxidative stress and works as a cofactor with GPx1 to detoxify H_2O_2 into H_2O. GCL the rate limiting enzyme in synthesis of GSH is composed of two subunits encoded by *GCLC* and *GCLM* genes. Using deletion mutagenesis Mulcahy and colleagues demonstrated the promoter region sequences involved in the up and down regulation of *GCLC* gene activity [Mulcahy et al., 1997]. The 5'-flanking region of GCLC (-3802 to +465) was cloned into pGL3-basic vector and the construct transfected into HepG2 cells. A series of deletions generated various reporter constructs which were transfected into the HepG2 cells and tested for reporter activity. The results showed that the region between -202 and +22 influenced the levels of expression and that deletion of this region resulted in a significant reduction in reporter gene activity and that further deletion (to +358) resulted in complete elimination of promoter activity. Results also showed other regulating elements such as ARE and AP-1 sequences and transcription binding sequences were found present in the -3802 to +465 sequence of GCLC (Figure 5.1).

In 2003 Koide and colleagues identified two polymorphisms (A-3506G and C-129T) in the GCLC promoter region. The T allele of C-129T polymorphism was found to be associated with increased risk of myocardial infarction (p=0.03) and was also found to be associated with lower transcriptional activity using a reporter gene assay [Koide et al., 2003]. Another polymorphism, a GAG trinucleotide repeat (TNR) polymorphism located downstream of the transcription start site in the 5`-untranslated region (5`-UTR) has been found to be associated with age of onset of type I diabetes and in LD with the C-129T polymorphism [Bekris et al., 2007]. The number of repeats varies from 4 to 10 repeats and it is estimated that 90% of the general population have a 7 or 9 GAG repeat [Walsh et al., 2001]. Though not explicitly stated in any article, it seems likely that T-129 variant is in linkage disequilibrium with the 7 GAG repeat allele, based on the report that 42% of a small group with a 7/7 repeat genotype were heterozygous for C-129T [Bekris et al., 2007] [Butticaz et al., 2011]. The originally reported upstream sequence shows a C at position -129 combined with 7 TAG repeats [Mulcahy et al., 1997]. GSH content has been suggested to be influenced by the number of GAG repeats, however not all results are consistent [Walsh et al., 2001;Gysin et al.,

2007;Nichenametla et al., 2008]. Butticaz and colleagues recently examined the functional significance of the GAG TNR using a reporter gene assay by cloning the region including the repeats downstream of a SV40 promoter but upstream of a luciferase coding sequence. When a 7 repeat sequence was combined with T-129, lower luciferase activity was seen compared with all other combinations [Butticaz et al., 2011]. A very recent study which considered the TNRs only also found higher luciferase activity associated with 8 and 9 repeats compared with 7 [Nichenametla et al., 2011].

In view of the genotyping data presented in chapter 4 showing a significant association for the T-129 allele with NAFLD in the Italian families, though no significance in the adult cohort, we decided to further assess the functional significance of this SNP alone by studying the region upstream of the transcription start site by reporter gene assays.

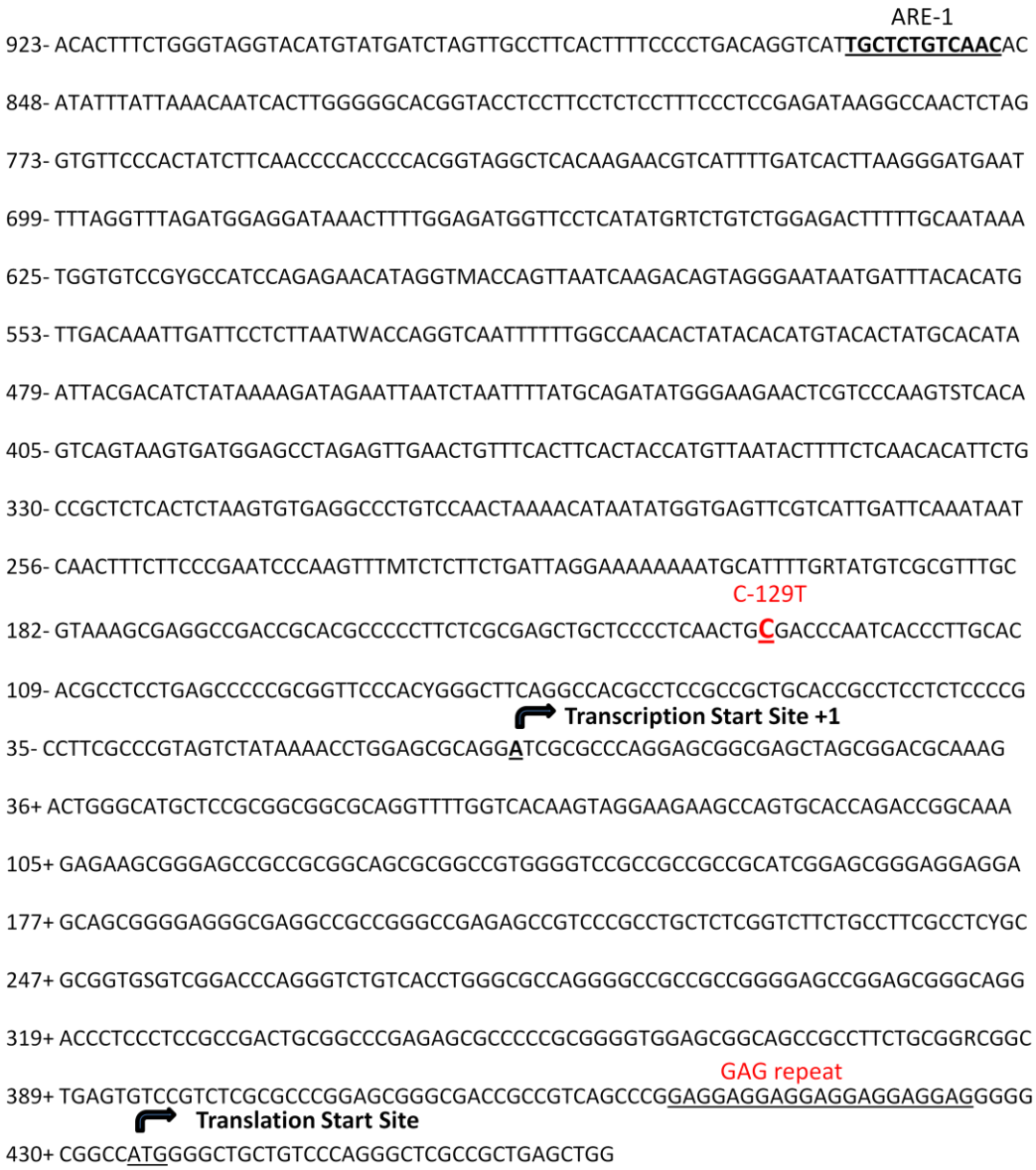

Figure 5.1: 5'-flanking region of GCLC showing the transcription start site (+1), the C-129T polymorphism and the GAG repeat.

5.2 Methods

5.2.1 Insert Design

792 base pairs upstream of the transcription start site (where +1 is the transcription start site) was introduced to the pGL3 basic vector. The area contained a TATA box and the C-129T *GCLC* polymorphism. *Kpn*I and *Sac*I restriction enzymes determined the orientation of the insert into pGL3 vector by introducing restriction sites into the primers. The 792 base pairs were amplified using the primers: 5'-TTTCGGTACC**AGATAAGGCCAACTCTAGGT**-3' and 5'CTGGGAGCTC**CCTGCGCTCCAGGTTTTATA**-3'. The bold type indicates the GCLC sequence. PCR was performed on two DNA samples, one with the -129TT and other -129CC genotypes which resulted in the two different constructs.

5.2.2 Purification of PCR product for cloning

DNA was amplified by PCR according to methods described in chapter 2. The annealing temperature used was 66^{o}C. PCR products were run on a 2% agarose gel. A scalpel was used to remove the bands from the agarose gel on a UV transilluminator. A QIAGEN clean up kit was then used to purify the PCR products. The PCR product was ligated into a pGEM T-easy vector according to the manufacturer guidelines (Promega).

5.2.3 Ligation and Transformation

Ligation reaction was set up using 5 µl of 10x T4 ligation buffer, 1 µl of 50 ng T-easy vector, 1 µl T4 DNA ligase and 10 µl of PCR product. Reactions were incubated overnight at 4^{o}C.

Component JM109 *E.coli* cells (100 µl) were mixed with 10 µl of ligation reaction in a 1.5 ml tube on ice for 20 minutes. The mixture was then heat shocked for 55 seconds in 42^{o}C in a water bath and placed on ice for 2 minutes. Sterile SOC medium (0.5% yeast extract, 2% tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgSO_4$, 10 mM $MgCl_2$, 20 mM glucose) was prepared and 900 µl was added to the tubes. Tubes were then incubated for 1.5 h with shaking at 37^{o}C at 150 rpm.

Transformed cells (100µl and 200µl) were then spread onto duplicate LB agar plates which had been overlaid previously with 100µl of 10µg/ml ampicillin and 20µl of 0.03mg/ml X-gal and 0.3mM IPTG and then incubated overnight at 37^{o}C.

5.2.4 Plasmid DNA extraction

White colonies were picked by sterile pipette tips into 10 ml of LB medium containing 100µg/ml ampicillin in universal tubes and incubated on an orbital shaker at 150 rpm at 37^{o}C overnight. The bacterial culture (1.5 ml) was then centrifuged at 13000 x rpm in 1.5 ml eppendorf tube to pellet cells which were then used to extract plasmid DNA using a QIAGEN miniprep plasmid extraction kit.

5.2.5 Preparation of GCLC promoter-pGL3 constructs

Both the pGEM T-easy vector containing the insert and pGL3-vector were digested with both *Kpn*I and *Sac*I (2U in 10x buffer and 1% BSA for 3h) in separate tubes. Digests were then run on 2% agarose gels. Both inserts and pGL3 digested vector were extracted and ligation and transformation was performed as before. The pGL3-insert were then extracted and purified again. To confirm the presence of the pGL3 and insert we digested 5 µl of plasmid DNA with *Kpn*I and *Sac*I as above and ran the samples on a 2% agarose gel (Figure 5.2). Plasmid DNA was then sent for sequencing (Eurofins MWG, UK) to confirm that the variation was only found at position -129 and not elsewhere. Primers for sequencing were supplied by Eurofins MWG and the plasmid DNA concentration was adjusted to 50-100 ng/µl.

5.2.6 HepG2 and LS180 tissue culture

HepG2 cells (hepatocellular carcinoma cell line) were supplied by Fiona Fenwick, Institute of Cellular Medicine, Newcastle University. LS180 (colon cancer cell line) were obtained from European collection of cell cultures (ECACC) (Porton Down, UK). Cells were cultured in 75cm^2 filtered flasks with 90 ml sterile distilled H_2O, 10 ml of 10x Minimum Essential Medium – Eagles (EMEM) (Bio Whittaker) supplemented with 10 ml of 10% heat-inactivated calf serum, 1 ml of 100U/mL penicillin and 100 µg/ml streptomycin, 1 ml of 200 mM L-Glutamine, 1 ml of non-essential amino acid and 5 ml of 4.4% gassed sodium bicarbonate containing 0.4% phenol red. Flasks were incubated at 37^{o}C with 5% CO_2. Confluent cells were split to 1:3 ratio every five days. Medium was removed and discarded. Cells were washed with PBS and 0.25% trypsin in Ca^{2+} and Mg^{2+} free PBS with 0.02% EDTA at 37^{o}C was used to detach cells from flasks.

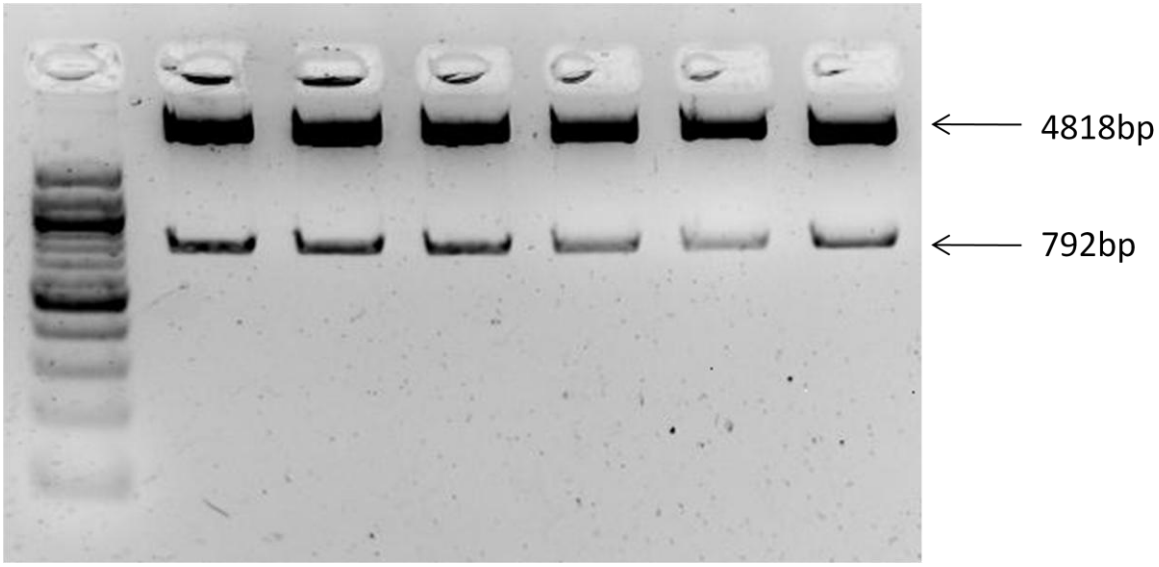

Figure 5.2: 2% agarose gel of pGL3 GCLC promoter plasmid DNA digested with *SacI* and *KpnI*. The upper band shows the linearised vector and the lower band shows the -129 insert.

5.2.7 Seeding cells

Cells were seeded into 24-well plates at a density between $4x10^{-4}$ cells/ml per well in complete medium to reach 50-80% confluence overnight. Medium was replaced with 200 µl of full fresh EMEM per well.

5.2.8 Cell Transfection and treatment

GeneJuice (Merck, Nottingham, UK) was used with serum free medium for transfection. Serum free medium (100 µl) and 1.5 µl of GeneJuice was vortexed and incubated at room temperature for 5 minutes before adding 0.5 µg of the pGL3-promoter vector or pGL3-empty vector and 0.05 µg of Renilla control vector per well. This master mix was then left for 20 minutes at room temperature before adding 100 µl of master mix to the existing 200 µl fresh EMEM in the wells. Cells were then left overnight before replacing the transfection reagent with 1ml of fresh complete medium. Once cells were transfected they were left for 72 h before setting up luciferase reporter assay. One hour prior to performing the luciferase assay, cells were treated with 100 µM H_2O_2 for 1h.

5.2.9 Luciferase reporter assay readings

Medium was removed and surface of the wells were washed with sterile PBS before adding 100 µl of passive lysis buffer (PLB) (Promega). Wells were left with PLB for 15-20 minutes at room temperature on a shaker to ensure even coverage of PLB and to disrupt the cells. Luciferase assay reagent II (LAR II) (100 µl) was added to 1.5 ml microcentrifuge tubes containing 20 µl of cell lysate. Cells were mixed thoroughly by pipetting before taking the initial reading (Firefly Luciferase). Stop and Glo reagent (100 µl) was then added and vortexed for the second reading (Renilla Luciferase). The luminometer (Turner Biosystems) was set up to perform a 2 second delay followed by 10 second measurement according to the Dual-Luciferase reporter assay protocol (Promega).

5.3 Results

5.3.1 Sequencing

The sequencing results showed that each construct only varied at the -129 position. All other SNPs in the region were found to be the same in each construct and all were wild-type (Figure 5.3A and 5.3B).

5.3.2 Luciferase reporter assay

The first measurement taken was the Firefly luciferase activity. The second reading was Renilla luciferase activity which were then normalised by Firefly/Renilla. Normalised readings for each sample were expressed as a fold increase and compared to that of the pGL3-empty vector. Results of the control vector (pGL3-empty) transfected into the cells showed small or no sign of reporter activity. Average normalised reporter gene activity for pGL3-T was 51.3 ± 3 fold greater than pGL3-empty (Figure 5.4). For C, the average normalised reporter activity was 59 ± 4.2 fold greater than pGL3-empty (Figure 5.4). To determine the difference between the pGL3-T and pGL3-C reporter gene activity a two-tailed t-test was performed. The slight difference was not found to be significant (p=0.3).

HepG2 cells were also treated with H_2O_2 for an hour before measuring the reporter activity. Measurements and analysis were similarly done to that of the untreated cells. The average normalised reporter activity for pGL3-T-H_2O_2 was 45.8 ± 2.3 fold greater than the pGL3-empty-H_2O_2 (Figure 5.5). The average normalised reporter activity for pGL3-C-H_2O_2 was 51.5 ± 1.5 fold greater than the pGL3-empty-H_2O_2 (Figure 5.5). There was no significant difference between the pGL3-T-H_2O_2 and pGL3-C-H_2O_2 reporter activity (p=0.3).

Reporter activity was also tested on a different cell line (LS180). The results appeared to be similar to that of HepG2 cell line. The average normalised activity for pGL3-T-LS180 was 16.5 ± 1.1 fold greater than pGL3-empty-LS180 (Figure 5.6). The average normalised reporter activity in pGL3-CC-LS180 was 16.1 ± 0.4 fold greater than pGL3-empty-LS180 (Figure 5.6). There was no significant difference between pGL3-TT-LS180 and pGL3-CC-LS180 reporter activity (p=0.7).

A

GCAGACTTTCTCTATCGATAGGTACCAGATAAGGCCAACTCTAGGTGTTCCCACTATCTT
CAACCCCACCCCACGGTAGGCTCACAAGAACGTCATTTTGATCACTTAAGGGATGAATTT
TAGGTTTAGATGGAGGATAAACTTTTTGAGATGGTTCCTCATATG**G**TCTGTCTGGAGACT
TTTTGCAATAAATGGTGTCCG**T**GCCATCCAGAGAACATAGGT**C**ACCAGTTAACCAAGACA
GTAGGGAATAATGATTTACACATGTTGACAAATTGATTCCTCTTAAT**A**ACCAGGTCAATT
TTTTGGCCAACACTATACACATGTACACTATGCACATAATTACGACATCTATAAAAGATA
GAATTAATCTAATTTTATGCAGATATGGGAAGAACTCGTCCCAAGT**C**TCACAGTCAGTAA
GTGATGGAGCCTAGAGTTGAACTGTTTCACTTCACTACCATGTTAATACTTTTCTCAACA
CATTCTGCCGCTCTCACTCTAAGTGTGAGGCCCTGTCCAACTAAAACATAATATGGTGAG
TTCGTCATTGATTCAAATAATCAACTTTCTTCCCGAATCCCAAGTTT**A**TCTCTTCTGATT
AGGAAAAAAAAATGCATTTTG**A**TATGTCGCGTTTGCGTAAAGCGAGGCCGACCGCACGCCC
CCTTCTCGCGAGCTGCTCCCCTCAACTG**T**GACCCAATCACCCTTGCACACGCCTCCTGAG
CCCCCCGCGGTTCCCAC**C**GGGCTTCAGGCCACGCCTCCGCCGCTGCACCGCCTCCTCTCCC
CGCCTTCGCCCGTAGTCTATAAAACCTGGAGCGCAGGGAGCTCTTACGCGTGCTAGCCCG
GGCTCGAGATCTGCGATCTAAGTAAGCTTGGCATTCCGGTACTGTTGGTAAAGCCACCAT
GGAAAGACGCCAAAAACATAAAG

B

GCAGACATTTCTCTATCGATAGGTACCAGATAAGGCCAACTCTAGGTGTTCCCACTATCT
TCAACCCCACCCCACGGTAGGCTCACAAGAACGTCATTTTGATCACTTAAGGGATGAATT
TTAGGTTTAGATGGAGGATAAACTTTTGGAGATGGTTCCTCATATG**G**TCTGTCTGGAGAC
TTTTTGCAATAAATGGTGTCCG**T**GCCATCCAGAGAACATAGGT**C**ACCAGTTAATCAAGAC
AGTAGGGAATAATGATTTACACATGTTGACAAATTGATTCCTCTTAAT**A**ACCAGGTCAAT
TTTTTGGCCAACACTATACACATGTACACTATGCACATAATTACGACATCTATAAAAGAT
AGAATTAATCTAATTTTATGCAGATATGGGAAGAACTCGTCCCAAGT**C**TCACAGTCAGTA
AGTGATGGAGCCTAGAGTTGAACTGTTTCACTTCACTACCATGTTAATACTTTTCTCAAC
ACATTCTGCCGCTCTCACTCTAAGTGTGAGGCCCTGTCCAACTAAAACATAATATGGTGA
GTTCGTCATTGATTCAAATAATCAACTTTCTTCCCGAATCCCAAGTTT**A**TCTCTTCTGAT
TAGGAAAAAAAATGCATTTTG**A**TATGTCGCGTTTGCGTAAAGCGAGGCCGACCGCACGCC
CCCTTCTCGCGAGCTGCTCCCCTCAACTG**C**GACCCAATCACCCTTGCACACGCCTCCTGA
GCCCCCGCGGTTCCCAC**C**GGGCTTCAGGCCACGCCTCCGCCGCTGCACCGCCTCCTCTCC
CCGCCTTCGCCCGTAGTCTATAAAACCTGGAGCGCAGGGAGCTCTTACGCGTGCTAGCCC
GGGCTCGAGATCTGCGATCTAAGTAAGCTTGGCATTCCGGTACTGTTGGTAAAGCCACCA
TGGAAGACGCCAAAAACATAAAGAAAG

Figure 5.3: Sequencing results indicating GCLC sequence (in Black) and pGL3-vector sequence (in Blue) with variation only in -129. A) T-allele highlighted in red and 8 other SNPs in green. B) C-allele highlighted in red and other SNPs in green. No differences found between the other 8 SNPs in the two sequences.

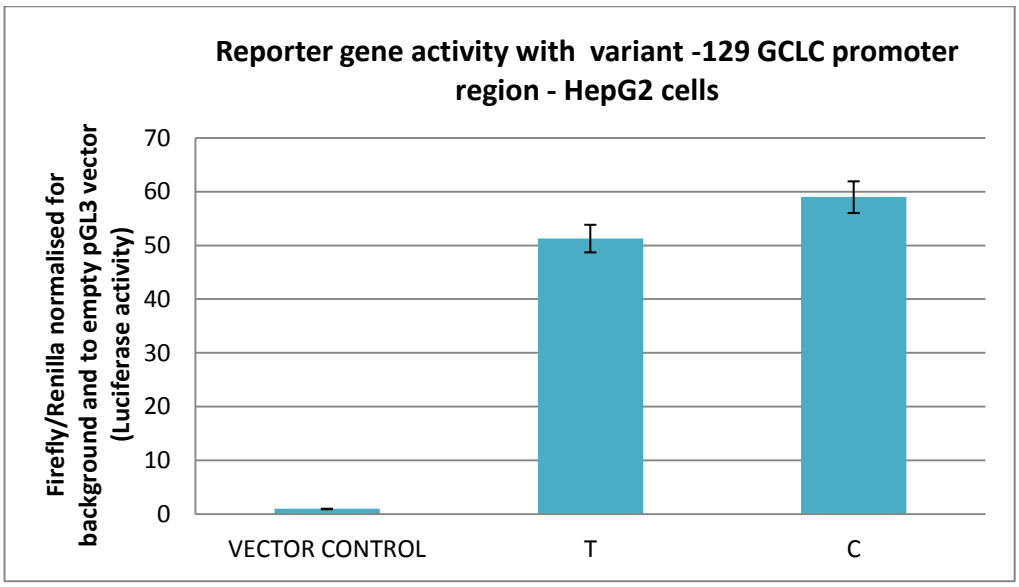

Figure 5.4: Luciferase reporter gene activity normalised to renilla activity, background and empty pGL3 vector, between a -129 T promoter and a -129 C promoter. T-test (2 tailed) p-value= 0.3 (n=3)

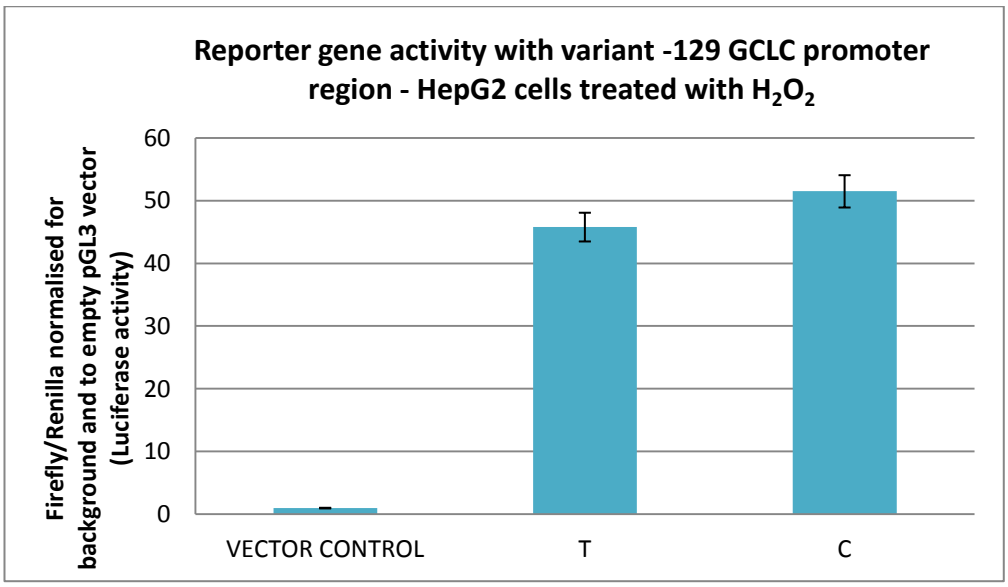

Figure 5.5: Luciferase reporter gene activity normalised to renilla activity, background and empty pGL3 vector, between a -129 T promoter and a -129 C promoter. T-test (2 tailed) p-value= 0.3 (n=3)

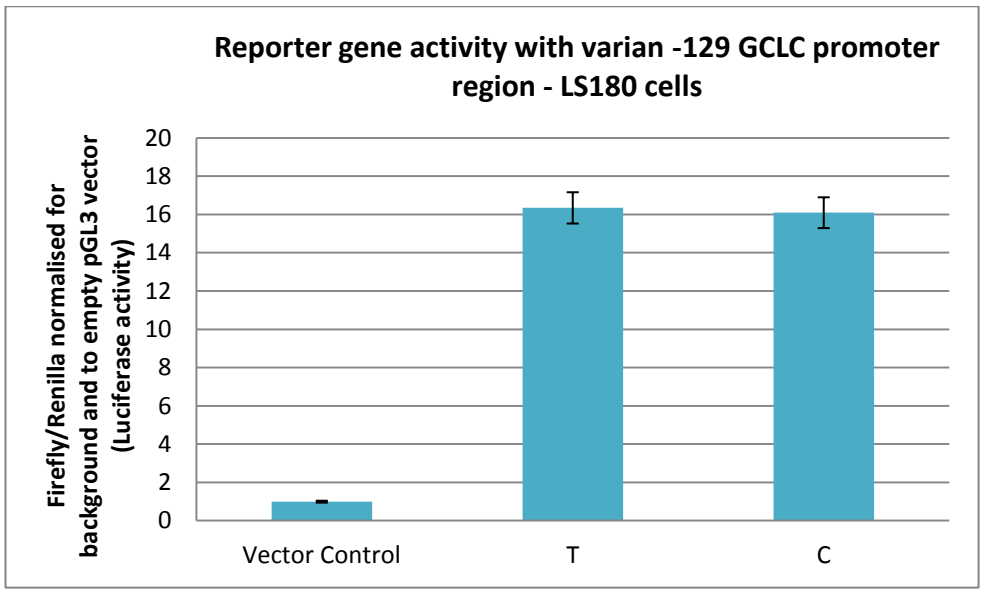

Figure 5.6: LS180 cells. Luciferase reporter gene activity normalised to renilla activity, background and empty pGL3 vector, between a -129 T promoter and a -129 C promoter. T-test (2 tailed) p-value = 0.7 (n=3)

5.4 Discussion

Because of the involvement of GCLC in GSH synthesis described earlier (section 1.7.1.2), it is important to understand the regulation of the *GCLC* gene. In this chapter we therefore described the reporter activity of the *GCLC* promoter region sequence that included the C-129T polymorphism. This appears to be the first study of the *GCLC* promoter region sequence that did not include transcription binding regulatory sequences further upstream and the 5`-UTR previously reported to be involved in gene expression as described in other studies [Koide et al., 2003;Butticaz et al., 2011]. The objective was to determine the effect of the C-129T SNP directly on gene transcription as effects from differences from the GAG repeats in the 5'-UTR may be at the level of mRNA stability or at the level of translation. Therefore for this purpose we designed a smaller construct (792bp) that only included the C-129T polymorphism compared with the larger construct (4267bp) previously studied which included upstream transcription binding regulatory sequences and the 5`-UTR that included a GAG TNR as shown in figure 5.7 [Koide et al., 2003].

Koide and colleagues suggested that the C-129T *GCLC* polymorphism showed functional significance with the luciferase activity for C significantly higher than for T [Koide et al., 2003]. However, our reporter gene assay results suggested the -129 polymorphism had no functional significance. Also, treatment with H_2O_2 did not influence reporter activity between the two constructs in our study conflicting with the earlier findings [Koide et al., 2003]. This could be due to our chosen promoter sequence construct not including any ARE or AP-1 sequences that could bind nuclear protein factors and therefore treatment with H_2O_2 did not influence reporter activity. Using a different cell line did not change the outcome of the reporter activity and the results showed no difference between the two constructs.

On the other hand our study was consistent with another study [Butticaz et al., 2011], where they found no differences between T-129 and C-129 constructs. However their constructs only had C-129 polymorphisms and therefore a point mutation by site-direct mutagenesis was introduced to construct T-129. The study demonstrated a decrease in activity only when both 7 GAG TNR and the T-129 were combined compared to the C-129 constructs. The GAG TNR is located in the 5`-UTR, a region which has been suggested by various studies to be involved in transcriptional mechanisms affecting gene expression [Mignone et al., 2002;Chen et al., 2008;Coppotelli et al., 2006]. The two very recent published studies involving reporter genes have investigated the activity

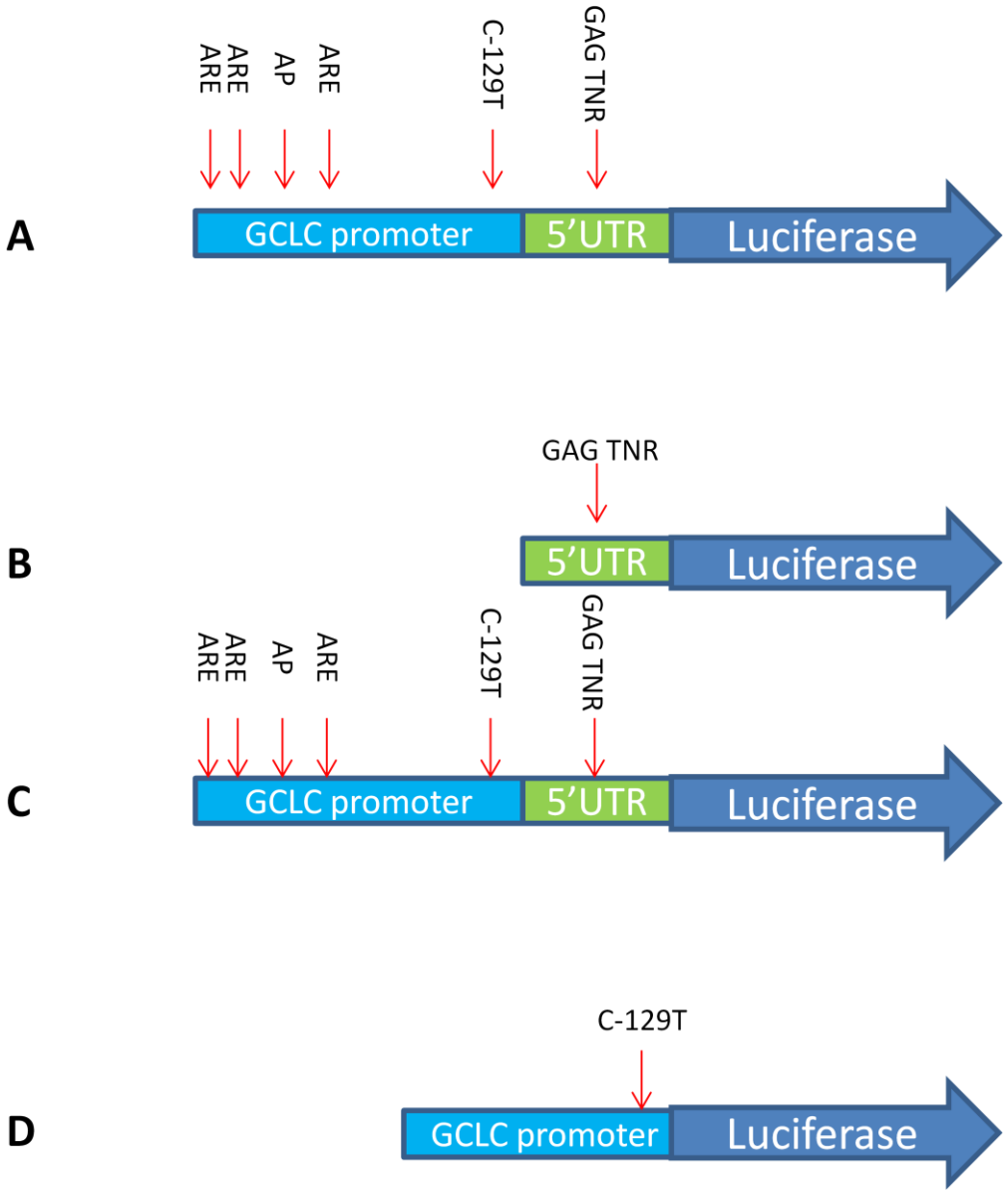

Figure 5.7: Schematic representation of the PGL3-basic vector constructs showing the location of the C-129T, GAG TNR and both ARE and AP sequences. A) Construct used in Koide *et al*. study included 5`-UTR and a large region of the promoter sequence. B&C) Constructs from Butticaz *et al*. study showing only the 5`-UTR in construct B and both 5`-UTR and the large region of the promoter sequence in C. D) Construct from our study showing only a smaller region of the promoter sequence.

of the GAG TNR without the *GCLC* promoter region sequence. There are discrepancies between them with one reporting no differences for luciferase with respect to the number of repeats alone but the other reporting higher luciferase for constructs with 8 or 9 repeats compared with 7 [Butticaz et al., 2011;Nichenametla et al., 2011]. In addition, Butticaz and colleagues measured GCLC protein levels in skin fibroblasts from individuals who were homozygous for 7 GAG repeats but had either CC or CT genotypes for C-129T. In this case, they found a significantly lower GCLC protein content (1.45 fold lower) in cells from individuals with the CT genotype compared with the CC genotype using Western blot analysis. A smaller not significant effect was obtained for total GSH content though the levels in the CT cells were lower than the CC.

The discrepancy between our findings and those of Koide et al. and Buttincaz et al. for C-129T might be due to factors involved in reporter gene assays. It has been previously reported that conflicting and inconsistent results in reporter gene activity in well-studied promoter SNPs may be due to the difference in reporter constructs, cell line used and transfection techniques reviewed by [Karimi et al., 2009].

We cannot rule out the interaction between the two polymorphisms (C-129T and GAG repeats) recently suggested by [Butticaz et al., 2011]. Ideally it would have been useful to have included as controls sequences where differences between constructs would be expected and are well established but suitable sequences were not available to me. In addition, studies involving a range of different reporter constructs with varying lengths of the promoter region would have been helpful.

My reporter gene assay results shows no evidence that the C-129T polymorphism directly influences reporter activity. This result is consistent with the case/control genotyping. The apparent preferential transmission of T over C in NAFLD families was only of borderline significance and may be a spurious finding. However, it is interesting that it is in line with suggestions in the literature that T-containing alleles are associated with lower GCLC levels than C-alleles.

Chapter 6

Patatin-like phospholipase-3 (PNPLA3) and Apolipoprotein-3 (APOC3) genotypes and their association with NAFLD severity

6 Patatin-like phospholipase-3 (PNPLA3) and Apolipoprotein-3 (APOC3) genotypes and their association with NAFLD severity

6.1 Introduction

In 2008, a genome-wide association study (GWAS) on the association of genotype with hepatic fat content which involved nonsynonymous polymorphisms only found an association with a SNP (rs738409) that results in a cytosine to guanine substitution encoding for isoleucine to methionine substitution at residue 148 (I148M) in *patatin-like phospholipase-3 (PNPLA3)* gene also known as adiponutrin [Romeo et al., 2008]. The limited GWAS involved 383 Hispanics, 1032 African-Americans and 696 European-Americans and hepatic fat content was assessed by magnetic resonance imaging. The *PNPLA3* G-allele was found to be associated with both increased hepatic fat ($p= 5.9 \times 10^{-10}$) and elevation of serum alanine aminotransferase (ALT) levels ($p= 1.3 \times 10^{-5}$) [Romeo et al., 2008]. This was consistent with a separate population-based GWAS carried by Yuan and colleagues on plasma liver enzyme levels where variants in *PNPLA3* predicted ALT levels [Yuan et al., 2008]. The association with liver fat was confirmed in two studies with smaller series of histological patients [Kotronen et al., 2009a] [Sookoian et al., 2009]. Kotronen and colleagues genotyped the I148M variant in 291 Finish individuals with hepatic fat measured by proton magnetic resonance. The G-allele was found associated with increased liver fat (p=0.01) [Kotronen et al., 2009a]. Sookoian and colleagues on the other hand studied 103 biopsy-proven NAFLD and found the G-allele to be associated with severe steatosis ($P<0.005$) [Sookoian et al., 2009]. The exact function of PNPLA3 is still controversial, however it is localised in adipose tissue and in liver and suggested to be functionally involved in both lipolysis and lipogenesis with the Met variant apparently associated with decreased catalytic activity [Romeo et al., 2010b] [Zimmermann et al., 2009] [He et al., 2010].

Lipoproteins are complexes of lipids (TG, cholesterol and phospholipids) and proteins known as apolipoproteins (APOC) [Mahley et al., 1984]. The different classes of lipoproteins (Chylomicrons, Chylomicron remnants, VLDL, and HDL) are involved in transport and distribution of lipids in cells and tissues [Mahley et al., 1984]. The APOC family play an important role in regulating lipoprotein metabolism [Mahley et al., 1984]. APOC3 is synthesised in the liver and the intestine and its main function is to regulate lipolysis through inhibition of hydrolysis enzymes (Lipoprotein Lipase and hepatic

lipase) [Bruns et al., 1984;Karathanasis, 1985;Ooi et al., 2008]. Studies have found that insulin is involved in regulating APOC3 [Li et al., 1995a;Chen et al., 1994] which suggests the dysregulation of APOC3 during diseases linked with insulin resistance. Two common *APOC3* polymorphisms T-455C and C-482T have been found to influence insulin resistance, lipid metabolism and susceptibility to NAFLD in a relatively small (n=95) study of Indian men [Petersen et al., 2010].

The aim of the work described in this chapter is to replicate the association of *PNPLA3* I148M variant with steatosis and to assess its relationship to severity of fibrosis. The effect of the *APOC3* promoter region polymorphisms (T-455C and C-482T) on the severity of NAFLD in UK patients is also assessed.

6.2 Methods

6.2.1 PNPLA3 and APOC3 SNP selection

The non-synonymous *PNPLA3* rs738409 polymorphism (I148M) was selected due to its association found in GWAS [Romeo et al., 2008]. Both *APOC3* rs2854116 (T-455C) and rs2854117 (C-482T) were selected due to the association with NAFLD reported previously [Petersen et al., 2010].

6.2.2 PNPLA3 rs738409 genotyping assay

PCR-RFLP was used to genotyped for rs738409. The forward and reverse primers were (5`-CTGCTCACTTGGAGAAAGCT-`3 and 5`-TGAAAGGCAGTGAGGCATGG-`3) . PCR amplification (35 cycles of denaturation at 95^{o}C for 1 minute, annealing at 57^{o}C for 1 minute and extension at 72^{o}C for 1 minute) was then performed in a GeneAmp PCR system 9700 thermal cycler. PCR products were digested with *Bts*CI. The PCR digestion products were viewed on a 2% agarose gel.

6.2.3 APOC3 rs2854116 and rs2854117

APOC3 genotyping was initially performed by Julian Leathart using the CVD53 multi-locus genotyping assay (Roche Molecular Diagnostic) published previously by [Cheng et al., 1998] and by PCR-RFLP previously published by [Hegele et al., 1997] which was performed as part of the present study. The forward and reverse primers (5'-TGTAACCAGGCCTTGTCGGA -3' and 5'-TTTCACACTGGAAATTTCAGG-3) were used to assess both SNPs. PCR amplification (35 cycles of denaturation at 95^{o}C for 1 minute, annealing at 57^{o}C for 1 minute and extension at 72^{o}C for 1 minute) was then performed in GeneAmp PCR system 9700 thermal cycler. PCR products were digested with *Msp*I for C-482T and *Fok*I for T-455C). PCR digestion products were viewed on a 2% agarose gel.

6.3 Results

6.3.1 PNPLA3

The *PNPLA3* rs738409 polymorphism was genotyped by PCR-RFLP with a typical result shown in Figure 6.1. Genotypes were in Hardy-Weinberg equilibrium (X^2=0.36, p=0.54). The relationship between genotype and severity of steatosis was analysed (Table 6.1). Incidence with severe steatosis (Grade 2+3) was 59% in patients carrying the G-allele compared to 41% homozygous wild type (CC) with an odds ratio 1.87 (95%CI 1.15 – 3.04; p=0.01). There was difference in genotype frequencies between mild fibrosis and severe fibrosis (Table 6.2). 28% of mild fibrosis patients carried the CC genotype in comparison with only 19% of severe fibrosis patients. However the GG genotype was found to be lower in mild fibrosis patients with only 20% compared with 30% of severe fibrosis patients (chi-square for trend p=0.0007). Carrying the G-allele was found to be associated with development of severe fibrosis with an OR of 2.02 (95% CI 1.29 – 3.1; p=0.002).

The TDT on the Italian families showed 56 out of 71 families showed at least one parent to be heterozygous. In these families the G-allele was transmitted on 52/76 (68%) occasions compared with the C-allele which was only transmitted on 24/76 (32%) occasions from both or at least one heterozygous parents to the affected child (p=0.001) (Table 6.3).

Multivariate analysis using logistic regression was performed on the Newcastle NAFLD samples to control for the effect of factors such as Genotype, age, sex, BMI, diabetes, steatosis and NASH that are relevant to disease severity (steatosis severity and fibrosis severity). The multivariate results demonstrated *PNPLA3* genotype was a predictor for severe steatosis with an odds ratio 1.8 (95% CI 1.09-2.99; p=0.021) (Table 6.4A). The multivariate results demonstrated *PNPLA3* genotype was a predictor for severe steatosis with an odds ratio 1.8 (95% CI 1.09-2.99; p=0.021) (Table 6.4A). *PNPLA3* genotype did not reach significance for predicting of NASH and severity of fibrosis (p=0.05 and p=0.06 respectively) (Table 6.4B and C).

6.3.2 Multivariate analysis combining PNPLA3, SOD2, KLF6 and ENPP1 genotypes

PNPLA3, *SOD2* previously described (Chapter 3) and previously published significant findings of *KLF6* [Miele et al., 2008] and ENPP1 [Dongiovanni et al., 2010] on the UK cohort were analysed to predict the factors involved in severe fibrosis (Table 6.5).

PNPLA3, *SOD2* and *KLF6* genotypes were found to be predictor factors for severity of fibrosis (p=0.03, p=0.001 and p=0.03 respectively) (Table 6.5). BMI, Diabetes and NASH were also found to be predictor factor for severity of fibrosis (Table 6.5).

6.3.3 APOC3

The promoter region polymorphisms of *APOC3* rs2854116 (T-455C) and rs2854117 (C-482T) were genotyped by CVD53 multi-locus and PCR-RFLP genotyping with a typical result for the latter shown in Figure 6.2A and B respectively. Genotypes were in Hardy-Weinberg equilibrium (X^2=0.61, p=0.43 for T-455C and X^2=0.008, p=0.92 for C-482T). The relation between genotype and severity of steatosis and fibrosis was analysed (Table 6.6 and Table 6.7 respectively). There was no difference in genotype frequency between mild steatosis and severe steatosis for both SNPs (p=0.9 and 0.6 respectively). There was also no difference in genotype frequency between mild fibrosis and severe fibrosis for both SNPs (p=0.4 and 0.9 respectively). LD between the two SNPs was moderate (r^2=0.5). The combined genotypes of both SNPs in relation to severity of steatosis and severity of fibrosis were analysed (Table 6.8 and 6.9 respectively). There was no significant difference between the combined wild type genotypes (TT and CC) against the remaining combined genotypes in relation to steatosis and fibrosis (p=0.46 and p=0.75 respectively) (Table 6.8 and 6.9).

The TDT showed 66 out 71 families to be at least heterozygous for one of the parents for T-455C polymorphism. In these families the T-allele was found to be transmitted on 43/66 (69%) occasion compared with the C-allele which was transmitted only 23/66 (31%) occasions, p=0.02 (Table 6.8). The TDT for C-482T polymorphism showed 61 out of 71 families to be informative. The T-allele was transmitted on 25/61 (41%) occasions compared with the C allele 36/61 (59%) occasions p=0.2 (Table 6.10).

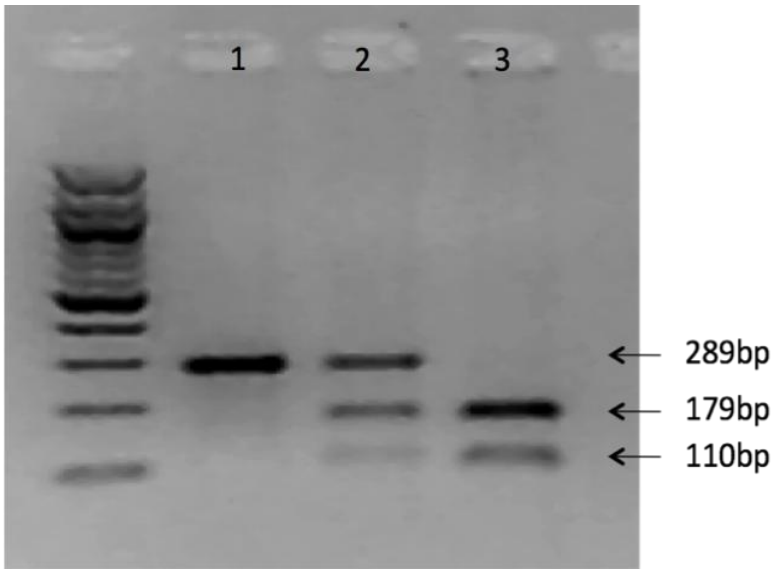

Figure 6.1: 2% agarose gel showing *PNPLA3* rs738409 digested with *BsaWI*. Lane 1 shows homozygous mutant (GG), Lane 2 shows heterozygous (CG) and Lane 3 shows homozygous wild type (CC).

Table 6.1: *PNPLA3* genotype distribution in Newcastle NAFLD patients in relation to steatosis severity

Gene and SNP	Genotype	Steatosis 0+1 (n=95)	Steatosis 2+3 (n=226)
PNPLA3 rs738409	CC	53 (0.55)	91 (0.41)
	CG	34 (0.36)	104 (0.45)
	GG	8 (0.09)	31 (0.14)

OR 1.87 (95% CI 1.15 – 3.04); p=0.01

Table 6.2: *PNPLA3* genotype distribution in Newcastle NAFLD patients in relation to fibrosis severity

Gene and SNP	Genotype	Fibrosis 0+1 (n=188)	Fibrosis 2-4 (n=143)
PNPLA3 rs738409	CC	98 (0.28)	50 (0.19)
	CG	74 (0.53)	69 (0.51)
	GG	16 (0.20)	24 (0.30)

OR 2.02 (95% CI 1.29 – 3.16); p=0.002

Chi-square for trend p=0.0007

Table 6.3: PNPLA3 transmission disequilibrium test, comparing the distribution on the transmitted and untransmitted allele from heterozygous parents to affected child

Gene and SNP number	Allele	Transmitted	Untransmitted	p-value
PNPLA3 rs738409	C	24 (32%)	52 (68%)	0.001
	G	52 (68%)	24 (32%)	

Table 6.4: Multivariate analysis of PNPLA3 using logistic regression for liver damage to control for predictor variables affecting severity of disease

A. Steatosis 0+1 vs. 2+3

Variables for	**OR (95% CI)**	**p-value**
PNPLA3	1.80 (1.09-2.99)	0.021
Age	0.97 (0.95-0.99)	0.025
Sex	0.68 (0.39-1.19)	0.182
BMI	1.04 (0.99-1.10)	0.080
Diabetes	1.19 (0.70-2.06)	0.513

PNPLA3 and Age found to be predictor factors

B. NASH (0 vs 1)

Variables for	**OR (95% CI)**	**p-value**
PNPLA3	1.62 (0.98-2.66)	0.05
Age	1.0 2 (1.00-1.04)	0.04
Sex	0.96 (0.55-1.69)	0.91
BMI	1.04 (0.99-1.09)	0.08
Diabetes	2.21 (1.30-3.81)	0.003

Age and diabetes found to be predictor factors

C. Fibrosis (0+1 vs 2-4)

Variables for	**OR (95% CI)**	**p-value**
PNPLA3	1.79 (0.96-3.35)	0.063
Age	1.03 (1.01-1.06)	0.004
Sex	1.37 (0.69-2.73)	0.358
BMI	1.06 (1.00-1.13)	0.031
Diabetes	2.13 (1.14-4.01)	0.017

Age, BMI and diabetes found to be predictor factors

Table 6.5: Multivariate analysis combining PNPLA3, SOD2, KLF6 and ENPP1 in relation to severe fibrosis (0+1 vs. 2-4) and to control for predictor variables

Variables for	OR (95% CI)	p-value
SOD2	2.3 (1.41 – 4.19)	0.001
PNPLA3	2.2 (1.09 – 4.66)	0.03
KLF6	0.28 (0.08 – 0.86)	0.03
ENPP1	1.8 (0.85 – 4.22)	0.11
Age	1.02 (099 – 1.05)	0.06
Sex	1.6 (0.62 – 3.02)	0.43
BMI	1.09 (1.02 – 1.18)	0.009
Diabetes	2.5 (1.24 – 5.38)	0.011

PNPLA3, SOD2 and KLF6 genotypes and both BMI and diabetes found to be predictor factors for severe fibrosis

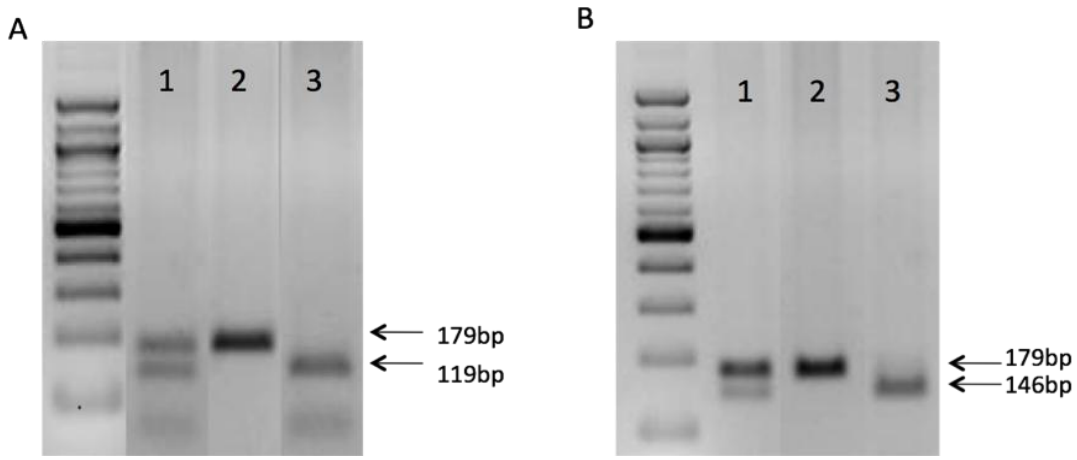

Figure 6.2: A) 2% agarose gel showing digest of APOC3 T-455C with *FokI*. Lane 1 shows heterozygous (TC), Lane 2 shows homozygous mutant (CC) and Lane 3 shows homozygous wild type (TT). B) 2% agarose gel showing APOC3 C-482T digested with *MspI*. Lane 1 shows the heterozygous (CT), lane 2 shows the homozygous mutant (TT) and lane 3 showing homozygous wild type (CC)

Table 6.6: *APOC3* genotype distribution in Newcastle NAFLD patients in relation to steatosis severity

Gene and SNP	Genotype	Steatosis 0+1	Steatosis 2+3	Odd ratio for carriage of minor allele (95% CI)	P value for carriage of minor allele
APOC3-455 rs2854116	TT	35 (0.41)	93 (0.41)	0.9 (0.6 - 1.6)	0.9
	TC	40 (0.45)	105 (0.46)		
	CC	12 (0.14)	37 (0.13)		
APOC3-482 rs2854117	CC	51 (0.58)	125 (0.55)	1.1 (0.7 – 1.9)	0.6
	CT	32 (0.36)	85 (0.38)		
	TT	5 (0.06)	15 (0.07)		

Table 6.7: *APOC3* genotype distribution in Newcastle NAFLD patients in relation to fibrosis severity

Gene and SNP	Genotype	Fibrosis 0+1	Fibrosis 2-4	Odd ratio for carriage of minor allele (95% CI)	P value for carriage of minor allele
APOC3-455 rs2854116	TT	78 (0.43)	54 (0.39)	1.2 (0.76 – 1.9)	0.4
	TC	79 (0.44)	71 (0.50)		
	CC	24 (0.13)	15 (0.11)		
APOC3-482 rs2854117	CC	103 (0.57)	80 (0.57)	0.9 (0.62 – 1.5)	0.9
	CT	67 (0.37)	52 (0.37)		
	TT	12 (0.07)	8 (0.06)		

Table 6.8: Combination of APOC3 rs2854116 and rs3854117 wild type genotypes compared with all other genotypes in relation to steatosis scores

APOC3 combined Genotypes	Steatosis 0+1 (n=78)	Steatosis 2+3 (n=197)
TT+CC	33 (42%)	93 (49%)
All other	45 (58%)	104 (51%)

OR 0.82 (95% CI 0.48-1.39); p=0.46

Table 6.9: Combination of APOC3 rs2854116 and rs3854117 wild type genotypes compared with all other genotypes in relation to fibrosis scores

APOC3 combined Genotypes	Fibrosis 0+1 (n=163)	Fibrosis 2-4 (n=112)
TT+CC	74 (45%)	53 (47%)
All other	89 (55%)	59 (53%)

OR 0.92 (95% CI 0.57-1.5); p=0.75

Table 6.10: *APOC3* transmission disequilibrium test comparing the transmitted and untransmitted allele from heterozygous parents to affect child

Gene and SNP number	Allele	Transmitted	Untransmitted	p-value
APOC3-455 rs2854116	T	43 (69%)	23 (31%)	0.02
	C	23 (31%)	43 (69%)	
APOC3-482 rs2854117	C	36 (59%)	25 (41%)	0.2
	T	25 (41%)	36 (59%)	

6.4 Discussion

We examined the association of the *PNPLA3* I148M polymorphisms that has been identified in a GWAS to be associated with increased levels of fat [Romeo et al., 2008]. Our finding that carriage of the G-allele of the *PNPLA3* polymorphism is associated with severity of steatosis (>1) is consistent with the GWAS finding [Romeo et al., 2008] and with a study on 103 biopsied NAFLD patients of European descent from Argentina [Sookoian et al., 2009] and a study on a Finnish population with 291 samples with liver fat measured by proton magnetic resonance [Kotronen et al., 2009a]. Our association on the association of the G allele with severe fibrosis in the adult Newcastle NAFLD samples is a novel finding. This was also replicated in a separate Italian cohort and both UK and Italian cohorts were combined and the results remained significant [Valenti et al., 2010a]. The PNPLA3 variant was found to be a predictor factor for both severe steatosis and severe fibrosis. The TDT on the Italian family trios with fibrotic NAFLD was consistent and confirmed the finding in the Newcastle samples. A study by Rotman and colleagues on Caucasians from the US reported an association of the G-allele with both steatosis and fibrosis in 894 biopsy-proven NAFLD samples with reported liver histological information [Rotman et al., 2010]. This was the first study to be consistent with our study on the association with fibrosis. A Japanese study (253 biopsy-proven NAFLD) reported the association of the G-allele with only fibrosis but not steatosis [Hotta et al., 2010] but this lack of association with steatosis may be due to differences in ethnicity. The I148M variant has also been found to be associated with severity of steatosis and fibrosis in children [Valenti et al., 2010b] which is consistent with our TDT results where 61 out the 71 children had fibrosis steatohepatitis (59 patients with stage 1 fibrosis and 2 patients with stage 2 fibrosis). To date there are 16 published studies on the influence of the I148M variant on the susceptibility and histological severity of NAFLD dating from 2008 to 2011, according to a meta-analysis study by [Sookoian and Pirola, 2011]. This combined study showed that the I148M variant influence liver fat accumulation with individuals with GG genotype having 73% higher fat compared to CC genotypes and also influences more aggressive disease where the GG genotype is 3.24 and 3.2-fold greater risk for developing NASH and fibrosis respectively compared with CC genotypes. The study also suggests the additive model (CG+GG vs. CC) best explains the effect of the I148M variant [Sookoian and Pirola, 2011]. In a GWAS on NAFLD performed on females, the 15 SNPs in *PNPLA3* that were studied did not show any significant association but rs734809 was not among the SNPs studied. A SNP in moderate LD (r^2=0.65) (rs2076211) was included but did not

show genome-wide significance [Chalasani et al., 2010]. A recent GWAS however found the PNPLA3 rs738409 variant to be associated with increasing computed tomography (CT) hepatic steatosis (n=7167) and associated to NASH in subjects with biopsy proven NAFLD (n=592) [Speliotes et al., 2011].

The exact function of PNPLA3 and the functional significance of its non-synonymous polymorphism (rs738409) that leads to a C to G substitution resulting in change of amino acid from isoleucine (ATC) to methionine (ATG) remain unclear. However it is suggested that substitution of methionine for isoleucine inhibits catalytic activity of the enzyme causing in appropriate access to substrate [He et al., 2010]. In vitro studies have shown that Met-PNPLA3 but not the wild-type Iso-PNPLA3 abolishes the activity of TG hydrolysis (lipolysis) resulting in accumulation of hepatic TG content and therefore suggesting a loss of PNPLA3 function [He et al., 2010]. However Chen and colleagues creation of a *Pnpla3* gene-deleted mouse produced by gene targeting resulted in *Pnpla3*$^{-/-}$ mouse failure to cause increase in hepatic TG content [Chen et al., 2010] adding further mystery to the function of Pnpla3. The human *PNPLA3* gene consists of four haplotype blocks and is 23kb long. It is possible that there are other genetic variants highly linked to rs738409.

Since *PNPLA3* genotype appeared to predict severity of fibrosis and SOD2 is also a predictor of fibrosis severity (Chapter 3) and genotype for *KLF6* has also been shown to predict severity in a separate study using these NAFLD DNA samples [Miele et al., 2008], the overall contribution of *PNPLA3* to severity alongside both *SOD2* and *KLF6* genotypes was also assessed. The overall effects of both SOD2 and PNPLA3 are broadly similar in terms of odds ratio (approx. 2) though the p value seen for SOD2 is lower, possibly because the variant genotype associated with increased risk is more common. The effect of the *KLF6* variant genotype is different since it decreases fibrosis risk but this had a slightly larger effect (approx 3 fold decrease in risk).

On the other hand, data on *APOC3* C-482T and T-455C did not show any significant association with severity of steatosis and fibrosis in the Newcastle cohort. However borderline significance for transmission of the -455 T-allele was found in the family study. Although Petersen and colleagues reported the T-allele to protect against steatosis in 95 male Asians from India with no liver histology available [Petersen et al., 2010], our result would indicate the opposite effect. The significance in our TDT may be a chance observation. Our finding for the adult Newcastle cohort was also replicated

in a separate Italian cohort. Both UK and Italian cohorts were also combined to increase the numbers and the results remained non-significant [Valenti et al 2011]. The data therefore suggests *APOC3* SNPs do not influence severity of NAFLD in Europeans. Our finding has been the first to report the lack of association of *APOC3* genotype with NAFLD severity. However, the same US cohort used to originally identify the *PNPLA3* variant association with hepatic fat was recently assessed for the *APOC3* variants and the results showed no significant association with hepatic fat content [Kozlitina et al., 2010].

In conclusion, the *PNPLA3* I148M is associated with risk of developing both steatosis and fibrosis severity in NAFLD patients suggests a link between lipid metabolism and fibrogenesis. The *PNPLA3* rs734809 is the strongest variant to be reported to date in susceptibility to NAFLD. Multivariate analysis showed that *PNPLA3*, *SOD2* and *KLF6* genotypes are independent risk factors for fibrosis. The *APOC3* promoter region variants do not affect severity of NAFLD in UK and Italian patients.

Chapter 7

Claudin-10 (CLDN10) – a candidate gene for NAFLD progression

7 Chapter 7 Claudin-10 (CLDN10) – a candidate gene for NAFLD progression

7.1 Introduction

There has been accumulating evidence supporting the liver-gut axis in the pathogenesis of NAFLD and therefore investigating the mechanisms involved in gut permeability is of interest [Dumas et al., 2006] [Solga and Diehl, 2003] [Farhadi et al., 2008]. In addition, there is evidence that intestinal tight junctions in NAFLD patients are disrupted, increasing gut permeability though it is possible that this could be due to small intestinal bacterial overgrowth rather than an abnormality within the tight junctions [Miele et al., 2009]. Increased gut permeability could result in increased fat deposition in the liver. A study by Beale and colleagues on liver tissue by expression microarray analysis suggested up-regulation of *CLDN10* gene (which encodes for CLDN-10 which is involved in paracellular permeability) is associated with advanced fibrosis [Beale et al., 2009].

Tight junctions are related to the cell-cell membrane contact of polarised cells (endothelia and epithelia). They form tissue barriers and pores that act as fences and gates. Tight junctions are involved in localizing membrane proteins (for example they occur at the boundary between the sinuosoidal and canalicular regions of the hepatocyte cell membrane) and regulating movement of solutes and electrolytes by the paracellular pathway [Berry and Boulpaep, 1975;Diamond, 1978] [Krause et al., 2008] [Amasheh et al., 2011]. Electrical resistance and preference for cations or anions are the characteristics of tight junctions permeability [Powell, 1981]. In transmission electron microscopy, tight junctions display a network of strands and groves [Staehelin, 1974]. Several transmembrane proteins such as occludin, claudins, tricellulin and marvelD3 have been identified in these tight junction strands suggesting their roles in barrier formation [Furuse et al., 1993] [Furuse et al., 1998] [Ikenouchi et al., 2005] [Steed et al., 2009].

Claudins are proteins with four transmembrane segments and two extracellular domains [Van Itallie et al., 2006]. The claudin family consists of 24 members, some of which are involved in forming paracellular channels and others play a role in tightening epithelia [Amasheh et al., 2011]. Claudins can therefore be divided into three functional groups,

barrier builders, pore formers and claudins of ambiguous function according to Amasheh and colleagues for review [Amasheh et al., 2011].

Claudin-10, a member of the claudin family is involved in mediating paracellular permeability and is encoded by *CLDN10* gene. It is mainly expressed in kidney and at low levels in the liver of mouse tissues [Van Itallie et al., 2006] [Gunzel et al., 2009]. Van Itallie and colleagues showed weak claudin-10 immunoreactivity in tight junctions of hepatocytes [Van Itallie et al., 2006]. The study also found two isoforms of claudin-10, claudin-10a and claudin-10b. The study reported that the two isoforms in both mouse and human differed in their first exon with claudin-10b having two extra amino acids (214bp in 10a and 220bp in 10b) and revealed that claudin-10a creates pore permeable for anions and claudin-10b for cations [Van Itallie et al., 2006]. On the other hand Gunzel and colleagues' study on mouse found six spliced isoforms in claudin-10, claudin-10a (expressed in kidney and uterus) and claudin-10b (expressed ubiquitously) being the main isoforms [Van Itallie et al., 2006] [Gunzel et al., 2009]. Gunzel and colleagues' study was consistent with claudin-10a and claudin-10b being the common isoforms, but also found another isoform of claudin-10a (claudin-10a_v1) that lacked 19 amino acids in the first exon. Claudin-10a_v1 variant has no direct effect on the ion selectivity as does the two common (10a and 10b) isoforms, however it does modulate claudin-10a's effect on organic anion permeability. The study also found three other variants of claudin-10 that lacked exon four which is predicted to contain the fourth transmembrane region. The remaining variants that lack exon four were not found in the tight junctions [Gunzel et al., 2009]. No studies to date have performed experiments on *cldn10*-deficient mice, however claudin-2 which is highly expressed with claudin-10 in the kidney has been studied in *cldn2*-deficient mice [Muto et al., 2010]. The *cldn2*-deficient mice show a loss of cation selectivity suggesting a role for the protein in cation permeability [Muto et al., 2010].

Down-regulation of CLDN-10 expression appears to be associated with prolonged disease free-survival in hepatocellular carcinoma (HCC) [Ip et al., 2007]. Though claudins are generally localized along the plasma membrane, immunostaining on HCC found CLDN-10 localised strongly around the nuclei [Inai et al., 2005] [Sanada et al., 2007]. Liver tissue microarray analysis showed the up-regulation of *CLDN10* is associated with advanced fibrosis [Beale et al., 2009]. Since hepatic claudin-10 expression in fibrosis cases is increased, it is possible that a similar effect could be

present in intestinal epithelial cells of NAFLD patients which could explain the increased gut permeability.

To date no polymorphisms have been investigated in *CLDN10* gene and therefore it would be of interest to study its variations, especially since studies in mice show the existence of six splice variant isoforms which could be due to splice site SNPs. A similar situation could occur in humans.

The aim of this chapter is to determine whether common haplotypes in *CLDN10* predict progression of NAFLD and also to perform immunohistochemistry (IHC) for claudin-10 in liver tissues of NAFLD patients to assess its expression in liver and possible functional significance of polymorphisms in *CLDN10*.

7.2 Materials and Methods

7.2.1 Immunhistochemistry (IHC)

Formalin-fixed and paraffin-embedded tissues were prepared by Helen Reeves' group, Newcastle University. Section (5-µm thick) slides were dewaxed in xylene for 5 minutes and rehydrated in alcohol gradient (100% to 75%). Slides were then incubated in 180ml of water with 3ml of 30% hydrogen peroxide for 10 minutes at room temperature to inactivate endogenous peroxidase. Slides were then incubated in 0.01M citrate buffer (pH 6.0) and microwaved for 10 minutes for antigen retrieval. Slides were then incubated with tris buffered saline (TBS, pH 7.6, 0.14 M NaCl, 5 mM Tris-HCl pH 7.6) for 5 minutes. Sections were then incubated overnight at 4^{o}C with appropriate primary antibodies with optimized dilutions (Table 7.1). Sections were then rinsed with TBS two times for 5 minutes. For the secondary antibody Menapath Kit was used (Table 7.1). The secondary antibody is a new horseradish-peroxidase (HRP) polymerization technology and was applied for 20 minutes at room temperature according to the manufacturer's guidelines, followed by a rinse in water for 20 minutes. Sections were then incubated with 3,3'-diaminobenzidine (DAB) substrate kit for 1 minute, producing a brown colour. Slides were then rinsed with water and then counterstained with Mayer's haematoxylin, followed by further rinse in water and brief dehydration in graded ethanol (75% to 100%) and xylene. Sections were mounted with DPX mounting medium.

7.2.2 Visualization

Slides were visualized using the ScanScope digital scanner. Slides were then viewed on Aperios Imagescope viewer software from www.aperio.com.

7.2.3 Evaluation of Immunohistochemical finding

The slides were analysed using a scoring system that incorporates both staining intensity (0 absent, 1 weak, 2 moderate or 3 strong) which was multiplied to overall percentage (from 0 to 100%) of staining resulting in score ranging from 0 to 300 (expression score) [Stefano et al., 2011].

7.2.4 CLDN10 genotyping

Tag SNP selection was carried out by Gary Beale, Newcastle University. All SNPs selected were genotyped via K-bioscience fluorescent allele-specific PCR system as mentioned previously in section 2.1.6 for the Newcastle NAFLD samples. For the Italian 'trio' families, only rs4143093 was genotyped due to financial restrictions.

7.2.5 Statistical analysis

Genotyping analysis were carried out as discussed in section 2.4

Table 7.1: Summary of antibodies used for immunohistochemistry

Primary Antibodies	**Source**	**Optimal dilution**
Rabbit anti-claudin-10 (polyclonal)	Abcam Ab66053	1:200
Mouse anti-claudin-10 (monoclonal)	Invitrogen 415100	1:100
Secondary Antibody		
X-cell plus Universal polymer HRP detection kit	MenaPath MP-XCP-U6	One drop/section

7.3 Results

7.3.1 CLDN10 candidate gene association study

A tag SNP approach was adopted for *CLDN10* due to the limited information on the existence of functionally significant polymorphisms in the *CLDN10* gene. Eight *CLDN10* polymorphisms were selected for genotyping. Five tag SNPs (rs4143093, rs1000600, rs2031540, rs12855856 and rs1000601) were originally selected by Gary Beale. In addition, I selected three promoter region SNPs (rs12870505, rs9584304 and rs3751334) which were not found to be in LD with the tag SNPs selected in HapMap CEU data (Figure 7.1). HapMap data on CLDN10 consisted of information on two isoforms (claudin-10a and claudin-10b). Using Haploview, Claudin-10a consisted of eight haplotype blocks whereas claudin-10b consisted of one haplotype block which was the only block shared between the two isoforms (Figure 7.1). Five tag SNPs (rs4143093, rs1000600, rs2031540, rs12855856 and rs1000601) from the shared haplotype were selected, these tag SNPs were in LD ($r^2 > 0.8$) with 12 other SNPs found across the haplotype blocks that were able to determine the common haplotypes (Figure 7.1). These selected SNPs were genotyped by allelic discrimination and the relationship between genotype and severity of steatosis (summarised in table 7.2), presence of NASH (summarised table 7.3) and severity of fibrosis (summarised in table 7.4) were analysed. For the eight SNPs selected all were in Hardy-Weinberg equilibrium (P>0.05). There was a significant difference between the mild steatosis (0+1) group compared to the severe steatosis (2+3) group in carriage of the A-allele of rs4143093 with OR 2.82 (95% CI (1.5-5.3), p=0.0009) (Table 7.2). There was no significant difference between mild steatosis and severe steatosis in the remaining SNPs (Table 7.2). There were no significant differences in carriage of the variant alleles between presence and absence of NASH (Table 7.3) or between mild fibrosis and severe fibrosis groups (Table 7.4). The polymorphism rs4143093 was genotyped in the Italian family trio samples but there was no preferential transmission of any allele to the affected children (p=0.88) (Table 7.5). Multivariate analysis using logistic regression analysis was performed to evaluate the factors associated with severe steatosis. *CLDN10* genotype for rs4143093 was the only significant predictor factors for severe steatosis with OR 2.66 (95% CI (1.42-5.25); p=0.003) for CLDN10 (Table 7.6A). The common *PNPLA3* rs738409 variant previously described (Chapter 6) was added to the multivariate analysis and both genotype factors were found to be predictor factors for steatosis (Table 7.6B). Frequencies of common haplotypes were generated using the five tag SNPs (Table 7.7).

There was no difference between frequencies of haplotypes generated compared against frequencies from HapMap CEU (p=0.73) (Table 7.7).

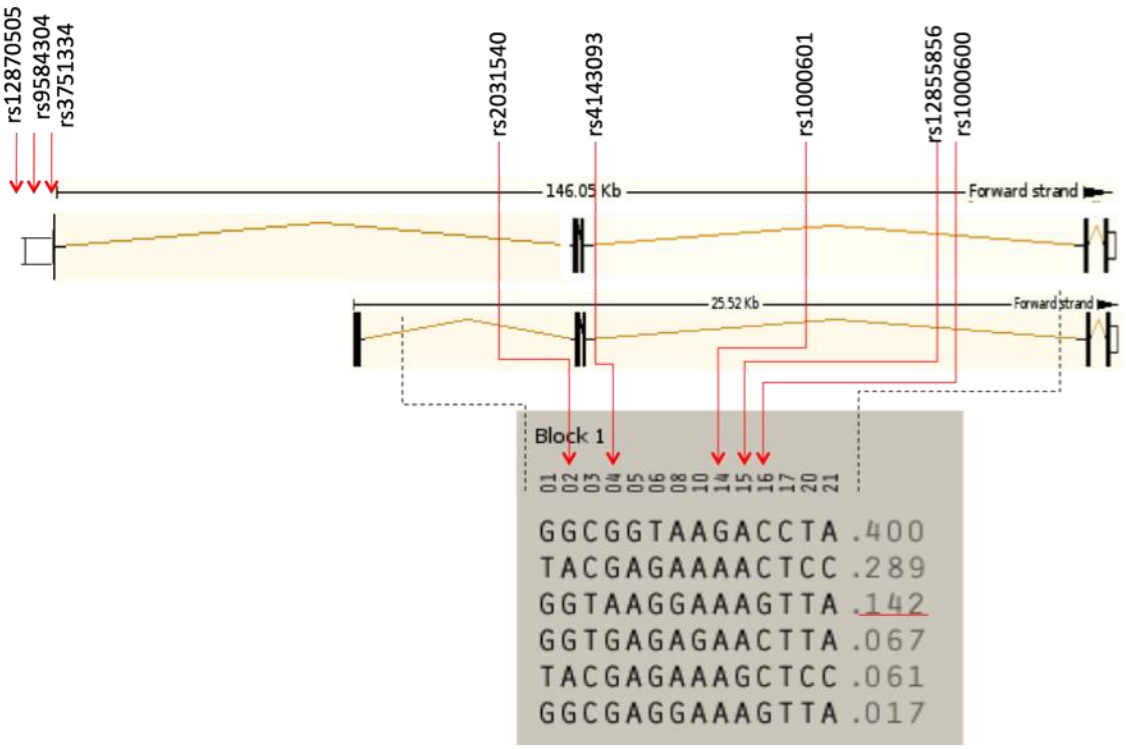

Figure 7.1: Five polymorphisms were selected and their position in the *CLDN10* gene is illustrated. One common haplotype block was identified between the two isoforms using Haploview. The haplotype block covers a region of the *CLDN10* gene represented by dotted lines. Five tag SNPs illustrated with red arrows were selected to identify all the haplotypes. Three promoter region polymorphisms from isoform one were selected.

Table 7.2: *CLDN10* genotype distribution in Newcastle NAFLD patients in relation to steatosis severity

CLDN10	Genotype	Steatosis 0+1	Steatosis 2+3	Odd ratio for carriage of minor allele (95% CI)	P value for carriage of minor allele
rs4143093	GG	59 (0.8)	121 (0.58)	2.82 (1.5-5.3)	0.0009
	GA	13 (0.18)	79 (0.38)		
	AA	2 (0.02)	8 (0.04)		
rs1000600	CC	54 (0.73)	122 (0.59)	0.94 (0.58 – 1.52)	0.08
	CG	19 (0.26)	79 (0.38)		
	GG	1 (0.01)	7 (0.03)		
rs2031540	GG	30 (0.41)	90 (0.44)	0.87 (0.51 - 1.5)	0.63
	GA	36 (0.49)	99 (0.48)		
	AA	8 (0.1)	17 (0.08)		
rs1000601	AA	27 (0.36)	85 (0.41)	0.81 (0.46 – 1.4)	0.45
	AG	34 (0.46)	94 (0.46)		
	GG	13 (0.18)	26 (0.13)		
rs12855856	AA	71 (0.97)	198 (0.95)	1.97 (0.42 – 9.1)	0.37
	AG	2 (0.03)	11 (0.05)		
	GG	0	0		
rs12870505	AA	59 (0.65)	153 (0.67)	0.91 (0.54 –	0.73

	AC	27 (0.3)	70 (0.31)	1.52)	
	CC	5 (0.05)	6 (0.02)		
rs3751334	AA	66 (0.72)	171 (0.75)		0.54
	AG	22 (0.24)	54 (0.24)	0.84 (0.49 – 1.45)	
	GG	4 (0.04)	3 (0.01)		
rs9584304	CC	38 (0.43)	106 (0.46)		0.56
	CT	38 (0.43)	102 (0.45	0.86 (0.52 – 1.41)	
	TT	13 (0.14)	21 (0.09)		

Table 7.3: *CLDN10* genotype distribution in Newcastle NAFLD patients in relation to presence of NASH

CLDN10	Genotype	NASH (No)	NASH (Yes)	Odd ratio for carriage of minor allele (95% CI)	P value for carriage of minor allele
rs4143093	GG	80 (0.68)	103 (0.61)	1.3 (0.8 - 2.1)	0.26
	GA	33 (0.28)	60 (0.36)		
	AA	5 (0.04)	5 (0.03)		
rs1000600	CC	77 (0.65)	102 (0.60)	1.2 (0.7 – 1.9	0.43
	CG	37 (0.31)	62 (0.37)		
	GG	4 (0.04)	4 (0.03)		
rs2031540	GG	45 (0.41)	79 (0.44)	0.69 (0.43 – 1.12)	0.13
	GA	60 (0.49)	75 (0.48)		
	AA	12 (0.1)	13 (0.08)		
rs1000601	AA	48 (0.36)	64 (0.41)	1.1 (0.68 – 1.7)	0.67
	AG	57 (0.46)	73 (0.46)		
	GG	12 (0.18)	29 (0.13)		
rs12855856	AA	111 (0.97)	162 (0.95)	0.79 (0.26 – 2.44)	0.69
	AG	6 (0.03)	7 (0.05)		
	GG	0	0		
rs12870505	AA	97 (0.65)	118 (0.67)	1.1 (0.74 –	0.45

	AC	36 (0.3)	61 (0.31)	1.9)	
	CC	8 (0.05)	3 (0.02)		
rs3751334	AA	99 (0.72)	140 (0.75)		
	AG	36 (0.24)	41 (0.24)	0.77 (0.47 – 1.28)	0.32
	GG	4 (0.04)	3 (0.01)		
rs9584304	CC	61 (0.43)	85 (0.46)		
	CT	64 (0.43)	77 (0.45	0.85 (0.54 – 1.32)	0.47
	TT	16 (0.14)	18 (0.09)		

Table 7.4: *CLDN10* genotype distribution in Newcastle NAFLD patients in relation to fibrosis severity

CLDN10	Genotype	Fibrosis 0+1	Fibrosis 2-4	Odd ratio for carriage of minor allele (95% CI)	P value for carriage of minor allele
rs4143093	GG	103 (0.65)	84 (0.64)	1.05 (0.64-1.7)	0.83
	GA	50 (0.31)	44 (0.33)		
	AA	6 (0.04)	4 (0.03)		
rs1000600	CC	99 (0.62)	84 (0.64)	0.94 (0.58 – 1.52)	0.8
	CG	55 (0.35)	45 (0.34)		
	GG	5 (0.03)	3 (0.02)		
rs2031540	GG	70 (0.44)	56 (0.43)	1.06 (0.66 - 1.7)	0.79
	GA	75 (0.47)	62 (0.47)		
	AA	13 (0.09)	13 (0.1)		
rs1000601	AA	60 (0.38)	53 (0.41)	0.88 (0.55 – 1.4)	0.62
	AG	76 (0.48)	57 (0.44)		
	GG	22 (0.14)	20 (0.15)		
rs12855856	AA	153 (0.96)	124 (0.94)	1.64 (0.55 – 4.86)	0.41
	AG	6 (0.04)	8 (0.06)		
	GG	0	0		
rs12870505	AA	129 (0.69)	89 (0.63)	1.29 (0.81 –	0.26

	AC	51 (0.27)	48 (0.34)	2.06)	
	CC	7 (0.04)	4 (0.03)		
rs3751334	AA	132 (0.71)	112 (0.78)		0.13
	AG	50 (0.27)	28 (0.19)	0.67 (0.4 – 1.12)	
	GG	4 (0.02)	3 (0.03)		
rs9584304	CC	84 (0.45)	65 (0.46)		0.78
	CT	83 (0.44)	61 (0.44)	0.94 (0.6 – 1.46)	
	TT	20 (0.1)	14 (0.1)		

Table 7.5: *CLDN10* Transmission disequilibrium test, comparing the transmitted and untransmitted allele from heterozygous parents to affected child

Gene and SNP number	Allele	Transmitted	Untransmitted	p-value
CLDN10 rs4143093	G	25 (52%)	23 (48%)	0.88
	A	23 (48%)	25 (52%)	

Table 7.6: Multivariate analysis of *CLDN10* rs4143093 for severity steatosis (0+1 vs. 2+3), controlling for predictor variables

A. Steatosis 0+1 vs. 2+3

Variables for	OR (95% CI)	p-value
CLDN10	2.66 (1.42-5.25)	0.003
Age	0.98 (0.96-1.00)	0.21
Sex	0.66 (0.35-1.23)	0.192
BMI	1.04 (0.98-1.10)	0.125
Diabetes	1.37 (0.76-2.51)	0.291

CLDN10 appeared to be the only predictor factor for severe steatosis

B. Steatosis 0+1 vs. 2+3

Variables for	OR (95% CI)	p-value
PNPLA3	1.8 (1.0 – 3.2)	0.04
CLDN	2.7 (1.4 – 5.5)	0.003
Age	0.9 (0.9 – 1.0)	0.23
Sex	0.6 (0.3 – 1.1)	0.15
BMI	1.03 (0.9 – 1.1)	0.18
Diabetes	1.1 (0.6 – 2.2)	0.57

MVA showing PNPLA3 and CLDN to be predictor variables for severe steatosis

Table 7.7: *CLDN10* Haplotype frequencies compression in both Newcastle and HapMap populations

	Newcastle frequency	CEU samples from HapMap frequency	P-value
CLDN10			
GGGAC	0.373	0.400	0.73
AGAAC	0.305	0.289	
GAAAG	0.184	0.142	
GGAAC	0.086	0.067	
AGAGC	0.024	0.061	
GGAAG	0.014	0.017	

Chi-square test analysing the differences between frequencies of the Newcastle and HapMap populations

7.3.2 Claudin-10 immunohistochemical analysis

Two claudin-10 antibodies, a rabbit polyclonal and a mouse monoclonal antibody were used for immunostaining. Claudin-10 expression was found in three cell types, hepatocytes, bile ducts and inflammatory cells but mainly in bile ducts (Figure 7.2). Intensity and localisation of expression correlated with severity of disease (Figure 7.2 A-F). Expression of claudin-10 in bile ducts and hepatocytes was seen in both polyclonal and monoclonal antibodies (Figure 7.2 C-D). Expression of claudin-10 in inflammatory cells was only positive in the polyclonal antibodies compared to the monoclonal antibody (Figure 7.2 C and D).

The relationship between IHC expression and rs4143093 genotype was analysed to assess the possible functional significance of this SNP (Table 7.8). Only 7 NAFLD patients were analysed by IHC and all had rs4143093 genotypes available. One subject was heterozygous (GA) and the remaining six were homozygous wild-types (GG). The patient with the GA genotype had both severe fibrosis and severe steatosis (Table 7.8). The GA patient appeared to have the highest expression score of claudin-10 compared to the GG patients. However due to the small number of samples genotyped this needs to be further assessed on a larger number of samples.

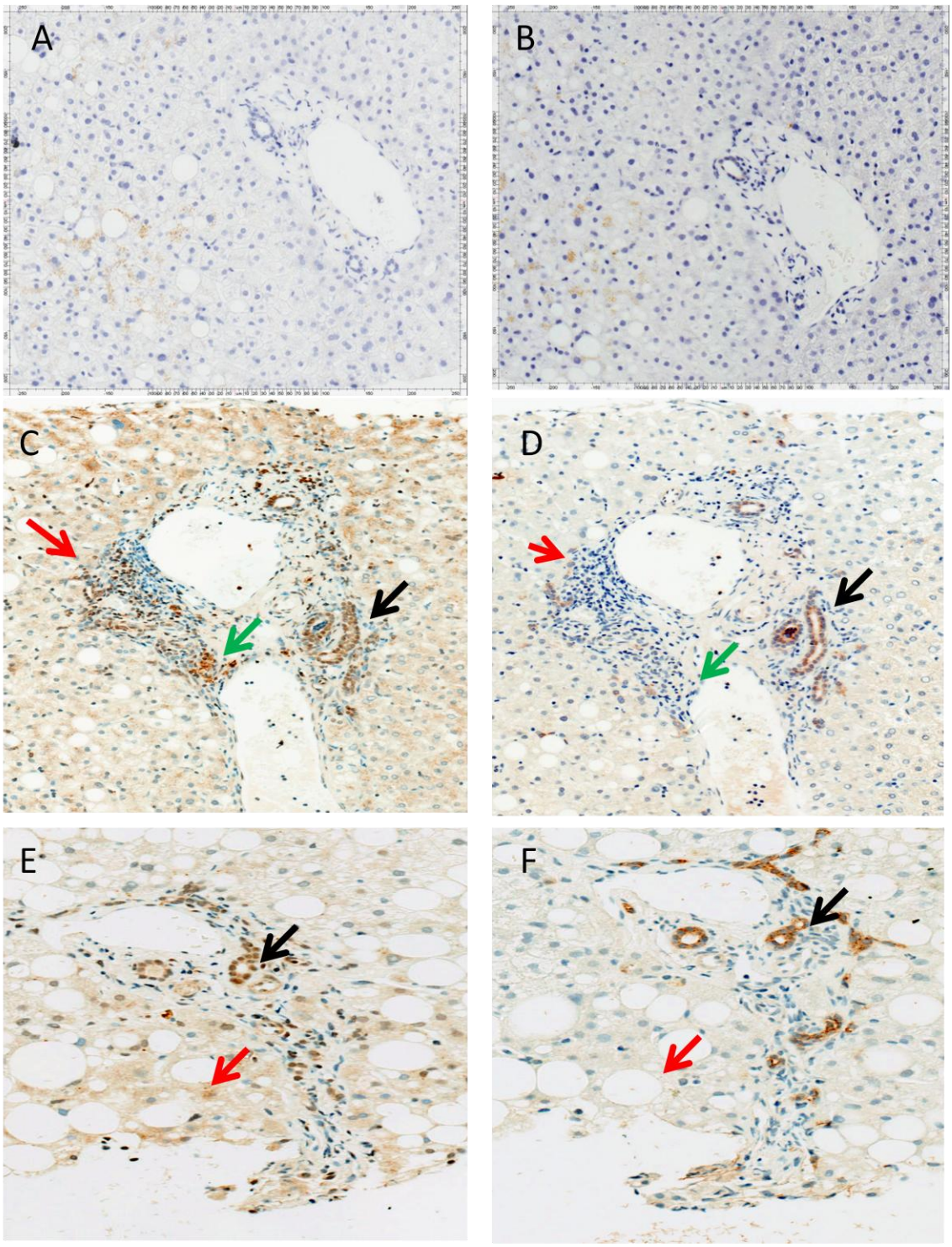

Figure 7.2: Immunohistochemistry using polyclonal antibodies (A, C and E x20) and monoclonal antibodies (B, D and F x20) A&B) represents mild steatosis, absence of NASH and fibrosis, no expression is seen in both antibodies C&D) represents severe steatosis, presence of NASH, and absence of fibrosis, expression seen in bile ducts (black arrow) for both antibodies, expression in inflammatory cells (green arrow) only positive in polyclonal antibody and in hepatocytes (red arrow). E&F) represents severe steatosis and presence of fibrosis, expression found in both bile ducts (black arrow) and hepatocytes (red arrow).

Table 7.8: Histology, claudin-10 expression and rs4143093 genotype

Case No.	Steatosis	Nicro-inflamation	Fibrosis	Abcam (polyclonal) Cell type (Intensity, over all %) (expression score)	Invitrogen (monoclonal) Cell type (Intensity, over all %) (expression score)	Genotype
1	1	0	0	HEP (1, 10%) (10) BD (1, 20%) (20)	HEP (1, 5%) (5) BD (2, 25%) (50)	GG
2	1	0	0	HEP (1, 5%) (5) BD (1, 5%) (5)	HEP (1, 5%) (5) BD (0, 100%) (0)	GG
3	3	0	1	HEP (1, 90%) (90) BD-nucleus (1, 20%) (20)	HEP (0, 100%) (0) BD-membrane (2, 25%) (50)	GG
4	2	0	0	HEP (2, 45%) (90) BD (3, 90%) (270) INF (3, 90%) (270)	HEP (0, 100%) (0) BD (2, 80%) (160)	GG
5	2	1	0	HEP (1, 90%) (90) BD (3, 90%) (270)	HEP (1, 100%) (100) BD (1, 35%) (35)	GG
6	3	0	0	HEP (1, 5%) (5) BD (1, 10%) (10) Other	HEP (0, 100%) (0) BD (1, 25%) (25)	GG
7	3	1	2	HEP (3, 90%) (270) BD (3, 90%) (270)	HEP (1, 90%) (90) BD (3, 90%) (270)	GA

HEP: Hepatocytes; BD: bile ducts; INF: inflammatory cells

Group1 Mild Steatosis (0+1) and Mild Fibrosis (0+1)

Group2 Severe Steatosis (2+3) and Mild Fibrosis (0+1)

Group3 Severe Steatosis (2+3) and Severe Fibrosis (2-4)

7.4 Discussion

The finding of increased gut permeability in NAFLD subjects and its correlation with severity of steatosis suggests the disruption of tight junctions [Miele et al., 2009]. The possibility of up-regulation of *CLDN10* in advanced fibrosis [Beale et al., 2009] suggests the involvement of claudin-10 in NAFLD. We have therefore investigated variants in the *CLDN10* gene that encodes for claudin-10 which is involved in mediating paracellular permeability. We have shown an association between an intronic polymorphism (rs4143093) and severity of steatosis which is consistent with a role for tight junctions in NAFLD [Miele et al., 2009]. The polymorphism rs4143093 is found 31 nucleotides downstream of exon 3 and might be involved in splicing due to its location. Using exon sequence enhancer (ESE) finder [Cartegni et al., 2003], the A-allele is suggested to destroy an SC35 splicing factor recognition site and create novel binding sites for both SRp40 and SRp55 (Helen Reeves, Newcastle University, unpublished observations) but this requires further assessment. The two common isoforms of claudin-10 (claudin-10a and claudin-10b) show a difference in exon one. It is therefore less likely that the sequence in the area of rs4143093 is involved in creating these two variants. The possibility that other SNPs could affect the relative levels of the two isoforms requires further investigation. Gunzel and colleagues work on mice found three isoforms of claudin-10 that lacked exon 4 [Gunzel et al., 2009]. However, it is not clear that this can be applied to humans and that rs4143093 could create a novel isoform that lacks an exon, though addressing this hypothesis requires further work.

The current study demonstrates the immunoexpression of claudin-10 in the spectrum of NAFLD and confirms the expression of claudin-10 in human liver tissues as was previously shown in mouse liver [Gunzel et al., 2009]. The increase in expression of claudin-10 in severe disease is consistent with the earlier finding that *CLDN10* gene expression was increased in severe NAFLD (severe fibrosis) [Beale et al., 2009]. The discrepancy between the expression of claudin-10 in inflammatory cells between the polyclonal and monoclonal antibodies is consistent with the differences in immunogen of the antibodies used where the monoclonal antibody is a mixture of two peptides from the middle and C-terminal regions of claudin-10 compared with only the C-terminal of the polyclonal antibody which makes the monoclonal more specific and may suggest the existence of two different isoforms.

Although results from IHC found expression of claudin-10 in liver tissues of NAFLD subjects with varying severity of disease, the study still has its limitations. The scoring

method used for IHC relies on a trained pathologist, which involves two characteristics; stain intensity and overall percentage. The errors in assessing for positive staining could lead to false results specially by not having control samples to eliminate any negative staining. Setting for a positive threshold should give better measurements. Computer imaging systems such as Aperio supplemented by a pathologist is able to set a positive threshold for staining that should result in accurate measurements [Nassar et al., 2011].

The relationship between IHC expression scores and genotype show increased expression scores of claudin-10 in GA subjects with severe NAFLD compared with GG genotype with mild NAFLD which may suggest the polymorphism to be associated with higher expression of claudin-10. However due to the small number of subject available for both IHC and genotyping these results need to be validated on a larger scale. Claudin-10 expression in hepatocytes, bile ducts and inflammatory cells suggests it could be of direct relevance to NAFLD. Tight junction function within hepatocytes could be a contributory factor to disease in addition to the previous suggestion that tight junction function in the intestine is important [Miele et al., 2009].

In conclusion, the finding from the association study suggests claudin-10 genotype is a contributor to severe steatosis in NAFLD. The findings from IHC confirm that the gene is expressed in the liver, and the overall expression of claudin-10 is increased in severe NAFLD. It remains unclear whether there is any relationship between this finding and the increased frequency of rs4143093 in severe steatosis with further IHC or RNA expression studies needed on more diseased samples of known genotype.

Chapter 8

General discussion

8 General discussion

8.1 General discussion

Non-alcoholic fatty liver disease is a complex disorder influenced by both genetic and environmental factors. NAFLD encompasses a spectrum of liver diseases ranging from steatosis to its more severe forms NASH/advanced fibrosis and cirrhosis. However not all individuals with steatosis will progress to the more severe forms. Although it is yet unclear what triggers such progression, it is suggested that knowledge of genetic variation in genes related to the mechanisms in the pathogenesis of NAFLD could enable us to determine those that are at high risk in developing severe NAFLD. Therefore those at high risk could be considered for preventive treatments.

This study was designed to compare unrelated patients of adults with NAFLD of varying severity assessed using the gold standard method liver biopsy. Because of the invasive nature of liver-biopsy, normal control samples were not studied. This therefore lead to the comparison of mild NAFLD patients against severe NAFLD patients. A second group of children with fibrotic NAFLD and their two parents in whom we have performed the transmission disequilibrium test was used to test for association and to confirm effects seen for the adult NAFLD group. It would have been ideal to use a normal control group free of NAFLD but the lack of non-invasive methods to establish absence of NAFLD was not feasible. There has been progress in assessing NAFLD severity using non-invasive methods that are based on a number of laboratory measurements in combination with clinical parameters such as BARD score, NAFLD fibrosis score, FIB-4 test, AST/ALT ratio and FibroMeter reviewed by [Dowman et al., 2011]. Other tests use the combination of fibrosis biomarkers such as the FibroTest [Poynard et al., 2007]. Radiological assessments such as ultrasound, computed tomography (CT) and magnetic resonance imaging (MRI) have been used to distinguish between NAFLD stage however not all studies have been reliable, as reviewed recently [Dowman et al., 2011]. This recent article suggests that SteatoTest and NASHTest can be used for steatosis and NASH assessments, combining 10 and 13 available blood tests respectively with age, gender and BMI [Dowman et al., 2011]. The ELF test also involves enzyme measurements and has recently been used as a measure of fibrosis in patients with hepatitis C [Trepo et al., 2011]. Despite the availability currently of these improved tests, they were not available when the current study was initiated. In addition, the cost of performing these on healthy controls is likely to be quite high and there are

potential ethical problems if normal controls were to be diagnosed with liver fibrosis. In view of all these limitations, the design used in the current study seemed the best approach when the study began and continues to have value except that genetic factors involved in susceptibility to simple steatosis would not be detected unless a control group without steatosis was included.

Using candidate gene association studies, our work has found significant associations in *SOD2* with severity of fibrosis, *PNPLA3* with severity of both steatosis and fibrosis and *CLDN10* with severity of steatosis (Table 8.1). Due to the correlation between NASH and fibrosis, assessing severity of fibrosis was sufficient to predict severe NAFLD. Assessing these SNPs alone would not be sufficient to predict severity of NAFLD but they are factors contributing to disease initiation and progression that could be targeted in the design of novel treatments. When we extend the findings described in the current thesis to other genetic studies performed on the same sample cohort where significant associations with disease severity have been detected, we should consider studies on KLF6 [Miele et al., 2008], ENPP1 and IRS-1 [Dongiovanni et al., 2010]. All three have been demonstrated to affect risk of fibrosis though not steatosis. However, in multivariate analysis on fibrosis severity, it was decided only to add both KLF6 and ENPP1 as additional genetic risk factors because IRS1 was found to be a significant risk factor for fibrosis severity only when additional cases from Italy were included as well as the Newcastle cohort [Dongiovanni et al., 2010]. Although the four genes were included in the multivariate analysis to predict severity of fibrosis, only SOD2, PNPLA3 and KLF6 were found to be significant predictor factors. It appears that genetic risk varies over an almost 15 fold range with those positive for the KLF6 variant but wild-type for SOD2 and PNPLA3 showing a 3 fold lower risk of severe fibrosis compared with those wild-type for all 3 but an individual positive for the "at risk" SOD2 and PNPLA3 genotypes and wild-type for KLF6 would show a 4 fold increase in risk compared with homozygous wild-types for the 3 genes. The risk of severe steatosis in individuals positive for the risk alleles of *CLDN10* and *PNPLA3* are approximately 4-fold higher compared to those negative for the risk alleles.

Table 8.1: Summary of significant findings on NAFLD severity genes on the Newcastle NAFLD cohort results

Gene and rs. Number	Steatosis	Fibrosis
CDN10 rs4143093	2.66 (1.42 - 5.25); p=0.003	NS
PNPLA3 rs738409	1.8 (1.09 - 2.99); p=0.02	2.2 (1.09 – 4.66); p=0.03
SOD2 rs4880	NA	2.3 (1.41 – 4.19); p=0.001
KLF6	NA	0.28 (0.08 – 0.86); p=0.03

Unlike the other genes studied, the selection of claudin-10 as a candidate gene for this study was due to the microarray expression analysis resulting in up-regulation of *CLDN10* in association with advanced fibrosis [Beale et al., 2009]. The positive association of *CLDN10* polymorphism with severity of steatosis has added mechanisms involving tight junctions to possible mechanisms involved in NAFLD. This finding supports use of expression microarray studies to identify candidate genes. Therefore in the future the choices of candidate genes can be selected from genome and proteome expression studies in liver tissues with different stages of disease, quantitative trait locus mapping in animal models and mouse mutagenesis studies [Day, 2010].

Our association findings are consistent with the involvement of lipid metabolism and oxidative stress in the pathogenesis of NAFLD as previously described in the literature. However the association with claudin-10 has opened a new venue towards identifying tight junctions as a new mechanism involved in the pathogenesis of NAFLD and therefore more consideration should be given to this newly involved mechanism. The *SOD2* supports the current use of antioxidants in patients with NAFLD [Foster et al., 2011] which may open up new venues that lead to designing new drugs or using current drugs for directing lifestyle advice.

A mechanistic explanation for the current findings can be proposed. Subjects having the risk allele of *PNPLA3* will have excess fat in the liver due to either the variant allele abolishing hydrolysis (breakdown of TG) and therefore causing accumulation of hepatic TG or increasing lipogenesis (increase in FA synthesis and subsequently TG synthesis). Carrying the *CLDN10* risk allele may also result in increased liver fat deposition from dietary fat due to increased gut-liver permeability. This excess fat results in lipid peroxidation (due to an increase in free radical production). Having the *SOD2* risk allele (lower enzyme activity) will result in less antioxidant activity. An imbalance between formation of free radicals and antioxidant activity will result in oxidative stress leading to the increased level of liver fibrosis seen in fibrosis stages 2 to 4. The reason for PNPLA3 contributing to risk of fibrosis in addition to steatosis (unlike CLDN10) may be due to its lipogenesis role whereas CLDN10 will simply allow more dietary fat to accumulate in the liver. Clearly both this study and the recent findings from others [Speliotes et al., 2010] suggest that PNPLA3 genotype is a important risk factor for NAFLD progression and measures to correct this deficiency (e.g. by drug treatment) might be of value in limiting disease severity.

The work done by our group has shed further light on the aetiology of NAFLD. The previous 'two-hit' hypothesis by [Day and James, 1998] suggested a set of factors to be involved in the initiation of NAFLD (first hit) where as a different set is involved in the progression of NAFLD (second hit). Although Day suggested the 'two-hit' theory, Day also suggested the involvement of insulin resistance in both initiation and progression of NAFLD. Other authors on the other hand suggested the interaction of factors [Malaguarnera et al., 2009]. Our findings of *CLDN10* and *SOD2* being associated with only severity of steatosis and severity of fibrosis respectively and *PNPLA3* being associated with both severe steatosis and severe fibrosis supports the observations that a group of factors may be only involved in initiation or progression whereas other factors may be involved in both initiation and progression together. This observation supports the complex nature of NAFLD.

Although the studies conducted in our lab did not find any association with the other oxidative stress related genes besides *SOD2* despite studying a number of SNPs in a range of candidate genes, these findings cannot be used to eliminate the role of oxidative stress in disease progression. This observation highlights the limited number of genes that had functionally significant polymorphisms available to assess and that information available on the majority of genes selected for this study was limited. It is therefore clear from our study that future studies examining susceptibility of NAFLD need to have a considerably larger number of samples than those performed so far and to be able to increase the coverage of functional and rare variants (rare variants < 5%) being studied. This can be achieved by collaborations and the establishment of consortia as well as use of the more contemporary approaches such as GWAS and whole-genome sequencing discussed below.

Today it is clear the main issue underlining NAFLD genetic association studies is the lack of consistency in selection of methods for liver assessment and the number of sample size between different studies. This inconsistency starts with the different non-invasive and invasive techniques being used for histological determination, a well-phenotyped cases and controls would be hard to achieve and also with the lack of collaboration the low number of samples used will result in non-robust findings that usually lead to most studies not being replicated by others. Another challenge today is identifying the actual functional disease variant when the surrounding variants are in strong LD which will result in similar association strength. Re-sequencing the genomic region is usually necessary to identify novel and uncommon variants. Functional studies

are necessary however they are only feasible for well-characterized genes and therefore challenging for SNPs located in unknown genes or non-coding regions and gene deserts.

GWAS studies have been a major tool in identifying novel variants that lead to increase understanding of disease biology. Their advantage is that no prior knowledge of disease aetiology is needed. In GWAS, 500,000 to 1,000,000 SNPs are analysed at once, to avoid detecting incorrect associations due to type 1 error, a stringent p-value for significance is needed. Because of this requirement, some true associations may not reach significance unless a very large sample size is used and such a large sample size may not be available in many cases. GWAS will not be able to identify all disease genes due to rare variants not being covered, resulting in missed variants [Goldstein, 2009]. However, work on improving the genotyping platforms (Illumina and Affymetrix) that have poor coverage for rare variants which are currently biased towards only common alleles is in progress. The study design of GWAS is hypothesis-generating rather than hypothesis testing, and therefore independent replication studies are required are also needed further increasing the sample size required [Seng and Seng, 2008]. However, for NAFLD, the *PNPLA3* association which was originally detected by a limited genome-wide association study has already been widely replicated in independent studies and this should also be possible in other cases.

To date there have been three GWAS on NAFLD which have been limited in the number of SNPs being assessed (only non-synonymous) [Romeo et al., 2008] or on small cohorts of 236 female NAFLD patients [Chalasani et al., 2010]. However, recently a GWAS (n=7176) by Speliotes and colleagues combined four cohorts from four different studies where NAFLD was not the primary diagnosed disease, the Old Order Amish, Age Gene/Environemt Susceptibility-Reykjavik study (AGES-Reykjavik), Family Heart, and Framingham Heart Studies [Speliotes et al., 2011]. Computed tomography (CT) was used for liver attenuation measuring hepatic steatosis. Although it is a non-invasive method, it is limited in assessing NASH/Fibrosis histology. However the study did replicate its findings in an independent cohort of 592 samples with biopsy-proven NAFLD from the NASH CRN samples from the US compared with 1405 control samples from iCONT. Our group are currently in process of a GWAS on biopsy-proven NAFLD samples from Northern Europe which is the first study to be conducted on only NAFLD samples assessed using the golden standard method. The study aims to reach 1000 samples for GWAS and an independent cohort will be available for replication to confirm the findings from the GWAS.

The value of the GWAS approach is well illustrated by recent progress on type 2 diabetes genetics. To date 44 different polymorphisms mostly in separate genes in type 2 diabetes have been discovered [Wheeler and Barroso, 2011]. Most of the associations have been well replicated. Although only 10% of the risk for type 2 diabetes is explained when combining all these loci [Wheeler and Barroso, 2011], knowledge of the risk factors can be used as targets for potential therapeutic drugs. There is strong evidence to date of the use of these discovered genes as drug targets, for example *PPARG* and thiazolidinediones and *KCNJ11* and sulfonylureas [Hakonarson and Grant, 2011]. Type 2 diabetes clearly overlaps with NAFLD and there is no reason why further GWAS studies on NAFLD should not quickly advance knowledge of genetics of this disease to the level now seen for type 2 diabetes. Nevertheless, it is desirable that if GWAS on NAFLD leads to the discovery of a relatively large number of genes affecting severity that more than 10% of risk will be explained.

SNPs alone cannot explain the genetic susceptibility of complex disease and therefore copy number variations (CNV) which are currently provided by GWAS platforms can be utilized to increase our understanding of the genetics of complex diseases. Other factors should be taken into consideration such as the involvement of microRNAs in posttranscriptional regulation of gene expression [Kerr et al., 2011] and DNA methylation and histone modification (the area of epigenetics) [Cooney, 2007] which deserve recognition as the understanding of these areas increases.

The newest approach for detecting disease-associated variants involves either whole genome sequencing covering the entire genome which is still costly or exome sequencing which is an approach for analysing polymorphisms in exons only. Although exome sequencing targeting only the coding regions is relatively simple using modern next generation sequencing methods, this does not detect other variations in the promoter and intronic regions that may also be functionally significant. The ongoing 1000 genome project is a large scale human genome sequencing programme divided into three projects; low-coverage genome sequencing, high coverage-genome sequencing and exon-targeted sequencing [Durbin et al., 2010]. Up to the present, the number of disease associated variants detected by these approaches is limited but very recently a number of new variants associated with rare diseases have been detected by exome sequencing. For example exome sequencing in Parkinson's disease (PD) has identified a missense mutation in the *VPS35* gene (component of the retromer cargo-recognition complex) suggested to cause late-onset disease [Zimprich et al., 2011].

Other diseases such as gray platelet syndrome has identified *NBEAL2* (Neurobeachin-like 2) as a causative gene [Albers et al., 2011]. Although there are an increasing number of published studies using exome sequencing, the majority of these are on Mendelian disorders rather than complex genetic disorders. It is possible that whole genome sequencing not exome sequencing may be needed to progress on identifying rarer variants important in complex diseases such as NAFLD.

Finally, though they suffer from a number of limitations, candidate gene association studies have resulted in improved understanding of NAFLD risk though currently the genes detected account for only a small proportion of the overall disease risk. With improved technology such as GWAS and genome sequencing and also the increased incidence of NAFLD due to the rise in obesity which makes finding cases relatively easy, it seems likely that additional genes will be identified in the near future. Knowledge of genetic factors will be useful for improving the understanding of disease biology which may lead to the development of new strategies such as directing lifestyle advice and development of new drugs for NAFLD.

References

References

ABRAMS, G. A.,KUNDE, S. S.,LAZENBY, A. J. & CLEMENTS, R. H. 2004. Portal fibrosis and hepatic steatosis in morbidly obese subjects: A spectrum of nonalcoholic fatty liver disease. *Hepatology,* 40**,** 475-83.

ADAMS, L. A.,LYMP, J. F.,ST SAUVER, J.,SANDERSON, S. O.,LINDOR, K. D.,FELDSTEIN, A. & ANGULO, P. 2005. The natural history of nonalcoholic fatty liver disease: a population-based cohort study. *Gastroenterology,* 129**,** 113-21.

ADLER, M. & SCHAFFNER, F. 1979. Fatty liver hepatitis and cirrhosis in obese patients. *Am J Med,* 67**,** 811-6.

ALBANIS, E. & FRIEDMAN, S. L. 2001. Hepatic fibrosis. Pathogenesis and principles of therapy. *Clin Liver Dis,* 5**,** 315-34, v-vi.

ALBERS, C. A.,CVEJIC, A.,FAVIER, R.,BOUWMANS, E. E.,ALESSI, M. C.,BERTONE, P.,JORDAN, G.,KETTLEBOROUGH, R. N.,KIDDLE, G.,KOSTADIMA, M.,READ, R. J.,SIPOS, B.,SIVAPALARATNAM, S.,SMETHURST, P. A.,STEPHENS, J.,VOSS, K.,NURDEN, A.,RENDON, A.,NURDEN, P. & OUWEHAND, W. H. 2011. Exome sequencing identifies NBEAL2 as the causative gene for gray platelet syndrome. *Nat Genet*.

ALEKSUNES, L. M. & MANAUTOU, J. E. 2007. Emerging role of Nrf2 in protecting against hepatic and gastrointestinal disease. *Toxicol Pathol,* 35**,** 459-73.

ALTSHULER, D.,HIRSCHHORN, J. N.,KLANNEMARK, M.,LINDGREN, C. M.,VOHL, M. C.,NEMESH, J.,LANE, C. R.,SCHAFFNER, S. F.,BOLK, S.,BREWER, C.,TUOMI, T.,GAUDET, D.,HUDSON, T. J.,DALY, M.,GROOP, L. & LANDER, E. S. 2000. The common PPARgamma Pro12Ala polymorphism is associated with decreased risk of type 2 diabetes. *Nat Genet,* 26**,** 76-80.

AMASHEH, S.,FROMM, M. & GUNZEL, D. 2011. Claudins of intestine and nephron - a correlation of molecular tight junction structure and barrier function. *Acta Physiol (Oxf),* 201**,** 133-40.

ANGELICO, F.,BURATTIN, M.,ALESSANDRI, C.,DEL BEN, M. & LIRUSSI, F. 2007. Drugs improving insulin resistance for non-alcoholic fatty liver disease and/or non-alcoholic steatohepatitis. *Cochrane Database Syst Rev***,** CD005166.

ANGULO, P. 2002a. Nonalcoholic fatty liver disease. *N Engl J Med,* 346**,** 1221-31.

ANGULO, P. 2002b. Treatment of nonalcoholic fatty liver disease. *Ann Hepatol,* 1**,** 12-9.

ANGULO, P. 2006. NAFLD, obesity, and bariatric surgery. *Gastroenterology,* 130**,** 1848-52.

ANGULO, P. 2007. GI epidemiology: nonalcoholic fatty liver disease. *Aliment Pharmacol Ther,* 25**,** 883-9.

ARGO, C. K. & CALDWELL, S. H. 2009. Epidemiology and natural history of non-alcoholic steatohepatitis. *Clin Liver Dis,* 13**,** 511-31.

ARTHUR, J. R. 2000. The glutathione peroxidases. *Cell Mol Life Sci,* 57**,** 1825-35.

BACON, B. R. & O'GRADY, J. G. 2006. *Comprehensive clinical hepatology,* St. Louis, Mo.; London, Elsevier Mosby.

BALLESTRI, S.,DAY, C. P. & DALY, A. K. 2011. Polymorphism in the farnesyl diphosphate farnesyl transferase 1 gene and nonalcoholic Fatty liver disease severity. *Gastroenterology,* 140**,** 1694-5.

BARRETT, J. C.,FRY, B.,MALLER, J. & DALY, M. J. 2005. Haploview: analysis and visualization of LD and haplotype maps. *Bioinformatics,* 21**,** 263-5.

BATALLER, R. & BRENNER, D. A. 2005. Liver fibrosis. *J Clin Invest,* 115**,** 209-18.

BAUMEISTER, S. E.,VOLZKE, H.,MARSCHALL, P.,JOHN, U.,SCHMIDT, C. O.,FLESSA, S. & ALTE, D. 2008. Impact of fatty liver disease on health care utilization and costs in a general population: a 5-year observation. *Gastroenterology,* 134**,** 85-94.

BEALE, G.,AL-SERRI, A.,MIELE, L.,PATMAN, G.,ABATE, M. L.,DALY, A. K.,BUGIANESI, E.,DAY, C. P. & REEVES, H. 2009. CLAUDIN 10 (CLDN10)- A CANDIDATE SUSCEPTIBILITY GENE FOR NAFLD PROGRESSION. *Hepatology,* 50**,** 1899.

BEKRIS, L. M.,SHEPHARD, C.,JANER, M.,GRAHAM, J.,MCNENEY, B.,SHIN, J.,ZARGHAMI, M.,GRIFFITH, W.,FARIN, F.,KAVANAGH, T. J. & LERNMARK, A. 2007. Glutamate cysteine ligase catalytic subunit promoter polymorphisms and associations with type 1 diabetes age-at-onset and GAD65 autoantibody levels. *Exp Clin Endocrinol Diabetes,* 115**,** 221-8.

BELFORT, R.,HARRISON, S. A.,BROWN, K.,DARLAND, C.,FINCH, J.,HARDIES, J.,BALAS, B.,GASTALDELLI, A.,TIO, F.,PULCINI, J.,BERRIA, R.,MA, J. Z.,DWIVEDI, S.,HAVRANEK, R.,FINCKE, C.,DEFRONZO, R.,BANNAYAN, G. A.,SCHENKER, S. & CUSI, K. 2006. A placebo-controlled trial of pioglitazone in subjects with nonalcoholic steatohepatitis. *N Engl J Med,* 355**,** 2297-307.

BELLENTANI, S.,SCAGLIONI, F.,MARINO, M. & BEDOGNI, G. 2010. Epidemiology of non-alcoholic fatty liver disease. *Dig Dis,* 28**,** 155-61.

BELTOWSKI, J. 2003. Adiponectin and resistin--new hormones of white adipose tissue. *Med Sci Monit,* 9**,** RA55-61.

BENZIE, I. F. 1996. Lipid peroxidation: a review of causes, consequences, measurement and dietary influences. *Int J Food Sci Nutr,* 47**,** 233-61.

BERMANO, G.,NICOL, F.,DYER, J. A.,SUNDE, R. A.,BECKETT, G. J.,ARTHUR, J. R. & HESKETH, J. E. 1996. Selenoprotein gene expression during selenium-repletion of selenium-deficient rats. *Biol Trace Elem Res,* 51**,** 211-23.

BERNARD, S.,TOUZET, S.,PERSONNE, I.,LAPRAS, V.,BONDON, P. J.,BERTHEZENE, F. & 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**,** 995-9.

BERRY, C. A. & BOULPAEP, E. L. 1975. Nonelectrolyte permeability of the paracellular pathway in Necturus proximal tubule. *Am J Physiol,* 228**,** 581-95.

BETTERIDGE, D. J. 2000. What is oxidative stress? *Metabolism,* 49**,** 3-8.

BIRCHER, J. 1999a. *Oxford textbook of clinical hepatology. Volume 1, Sections 1-13 and index,* Oxford; New York, Oxford University Press.

BIRCHER, J. 1999b. *Oxford textbook of clinical hepatology. Volume 2, Sections 14-32 and index,* Oxford; New York, Oxford University Press.

BISWAS, M. & CHAN, J. Y. 2010. Role of Nrf1 in antioxidant response element-mediated gene expression and beyond. *Toxicol Appl Pharmacol,* 244**,** 16-20.

BJORNSSON, E. & ANGULO, P. 2007. Non-alcoholic fatty liver disease. *Scand J Gastroenterol,* 42**,** 1023-30.

BROWNING, J. D. & HORTON, J. D. 2004. Molecular mediators of hepatic steatosis and liver injury. *J Clin Invest,* 114**,** 147-52.

BROWNING, J. D.,SZCZEPANIAK, L. S.,DOBBINS, R.,NUREMBERG, P.,HORTON, J. D.,COHEN, J. C.,GRUNDY, S. M. & HOBBS, H. H. 2004. Prevalence of hepatic steatosis in an urban population in the United States: impact of ethnicity. *Hepatology,* 40**,** 1387-95.

BRUNS, G. A.,KARATHANASIS, S. K. & BRESLOW, J. L. 1984. Human apolipoprotein A-I--C-III gene complex is located on chromosome 11. *Arteriosclerosis,* 4**,** 97-102.

BRUNT, E. M. 2004. Nonalcoholic steatohepatitis. *Semin Liver Dis,* 24**,** 3-20.
BUGIANESI, E. 2007. Non-alcoholic steatohepatitis and cancer. *Clin Liver Dis,* 11**,** 191-207, x-xi.
BUGIANESI, E.,GENTILCORE, E.,MANINI, R.,NATALE, S.,VANNI, E.,VILLANOVA, N.,DAVID, E.,RIZZETTO, M. & MARCHESINI, G. 2005. A randomized controlled trial of metformin versus vitamin E or prescriptive diet in nonalcoholic fatty liver disease. *Am J Gastroenterol,* 100**,** 1082-90.
BUGIANESI, E.,MANZINI, P.,D'ANTICO, S.,VANNI, E.,LONGO, F.,LEONE, N.,MASSARENTI, P.,PIGA, A.,MARCHESINI, G. & RIZZETTO, M. 2004. Relative contribution of iron burden, HFE mutations, and insulin resistance to fibrosis in nonalcoholic fatty liver. *Hepatology,* 39**,** 179-87.
BUTTICAZ, C.,GYSIN, R.,CUENOD, M. & DO, K. Q. 2011. Interaction of GAG trinucleotide repeat and C-129T polymorphisms impairs expression of the glutamate-cysteine ligase catalytic subunit gene. *Free Radic Biol Med,* 50**,** 617-23.
CADRANEL, J. F.,RUFAT, P. & DEGOS, F. 2000. Practices of liver biopsy in France: results of a prospective nationwide survey. For the Group of Epidemiology of the French Association for the Study of the Liver (AFEF). *Hepatology,* 32**,** 477-81.
CALDWELL, S. H. & HESPENHEIDE, E. E. 2002. Subacute liver failure in obese women. *Am J Gastroenterol,* 97**,** 2058-62.
CALVISI, D. F.,LADU, S.,HIRONAKA, K.,FACTOR, V. M. & THORGEIRSSON, S. S. 2004. Vitamin E down-modulates iNOS and NADPH oxidase in c-Myc/TGF-alpha transgenic mouse model of liver cancer. *J Hepatol,* 41**,** 815-22.
CAPRON, J. P.,DELAMARRE, J.,DUPAS, J. L.,BRAILLON, A.,DEGOTT, C. & QUENUM, C. 1982. Fasting in obesity: another cause of liver injury with alcoholic hyaline? *Dig Dis Sci,* 27**,** 265-8.
CARDON, L. R. & ABECASIS, G. R. 2003. Using haplotype blocks to map human complex trait loci. *Trends Genet,* 19**,** 135-40.
CARTEGNI, L.,WANG, J.,ZHU, Z.,ZHANG, M. Q. & KRAINER, A. R. 2003. ESEfinder: A web resource to identify exonic splicing enhancers. *Nucleic Acids Res,* 31**,** 3568-71.
CARULLI, L.,CANEDI, I.,RONDINELLA, S.,LOMBARDINI, S.,GANAZZI, D.,FARGION, S.,DE PALMA, M.,LONARDO, A.,RICCHI, M.,BERTOLOTTI, M.,CARULLI, N. & 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**,** 823-8.
CAUCHI, S.,EL ACHHAB, Y.,CHOQUET, H.,DINA, C.,KREMPLER, F.,WEITGASSER, R.,NEJJARI, C.,PATSCH, W.,CHIKRI, M.,MEYRE, D. & FROGUEL, P. 2007. TCF7L2 is reproducibly associated with type 2 diabetes in various ethnic groups: a global meta-analysis. *J Mol Med,* 85**,** 777-82.
CAVE, M.,DEACIUC, I.,MENDEZ, C.,SONG, Z.,JOSHI-BARVE, S.,BARVE, S. & MCCLAIN, C. 2007. Nonalcoholic fatty liver disease: predisposing factors and the role of nutrition. *J Nutr Biochem,* 18**,** 184-95.
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. & ROTTER, J. I. 2010. Genome-wide association study identifies variants associated with histologic features of nonalcoholic Fatty liver disease. *Gastroenterology,* 139**,** 1567-76, 1576 e1-6.
CHAN, K. & KAN, Y. W. 1999. Nrf2 is essential for protection against acute pulmonary injury in mice. *Proc Natl Acad Sci U S A,* 96**,** 12731-6.

CHARLTON, M. 2004. Nonalcoholic fatty liver disease: a review of current understanding and future impact. *Clin Gastroenterol Hepatol,* 2**,** 1048-58.
CHEN, G. L.,VALLENDER, E. J. & MILLER, G. M. 2008. Functional characterization of the human TPH2 5' regulatory region: untranslated region and polymorphisms modulate gene expression in vitro. *Hum Genet,* 122**,** 645-57.
CHEN, M.,BRESLOW, J. L.,LI, W. & LEFF, T. 1994. Transcriptional regulation of the apoC-III gene by insulin in diabetic mice: correlation with changes in plasma triglyceride levels. *J Lipid Res,* 35**,** 1918-24.
CHEN, W.,CHANG, B.,LI, L. & CHAN, L. 2010. Patatin-like phospholipase domain-containing 3/adiponutrin deficiency in mice is not associated with fatty liver disease. *Hepatology,* 52**,** 1134-42.
CHEN, Y.,YANG, Y.,MILLER, M. L.,SHEN, D.,SHERTZER, H. G.,STRINGER, K. F.,WANG, B.,SCHNEIDER, S. N.,NEBERT, D. W. & DALTON, T. P. 2007. Hepatocyte-specific Gclc deletion leads to rapid onset of steatosis with mitochondrial injury and liver failure. *Hepatology,* 45**,** 1118-28.
CHENG, S.,PALLAUD, C.,GROW, M. A.,SCHARF, S. J.,ERLICH, H. A.,KLITZ, W.,PULLINGER, C. R.,MALLOY, M. J.,KANE, J. P.,SIEST, G. & VISVIKIS, S. 1998. A multilocus genotyping assay for cardiovascular disease. *Clin Chem Lab Med,* 36**,** 561-6.
CHITTURI, S.,ABEYGUNASEKERA, S.,FARRELL, G. C.,HOLMES-WALKER, J.,HUI, J. M.,FUNG, C.,KARIM, R.,LIN, R.,SAMARASINGHE, D.,LIDDLE, C.,WELTMAN, M. & GEORGE, J. 2002. NASH and insulin resistance: Insulin hypersecretion and specific association with the insulin resistance syndrome. *Hepatology,* 35**,** 373-9.
CHO, H. Y.,REDDY, S. P.,DEBIASE, A.,YAMAMOTO, M. & KLEEBERGER, S. R. 2005. Gene expression profiling of NRF2-mediated protection against oxidative injury. *Free Radic Biol Med,* 38**,** 325-43.
COHEN, J. C.,HORTON, J. D. & HOBBS, H. H. 2011. Human fatty liver disease: old questions and new insights. *Science,* 332**,** 1519-23.
COONEY, C. A. 2007. Epigenetics--DNA-based mirror of our environment? *Dis Markers,* 23**,** 121-37.
COPPOTELLI, G.,SUMMERS, A.,CHIDAKEL, A.,ROSS, J. M. & CELI, F. S. 2006. Functional characterization of the 258 A/G (D2-ORFa-Gly3Asp) human type-2 deiodinase polymorphism: a naturally occurring variant increases the enzymatic activity by removing a putative repressor site in the 5' UTR of the gene. *Thyroid,* 16**,** 625-32.
CORDELL, H. J. 2002. Epistasis: what it means, what it doesn't mean, and statistical methods to detect it in humans. *Hum Mol Genet,* 11**,** 2463-8.
CORDELL, H. J. 2009. Detecting gene-gene interactions that underlie human diseases. *Nat Rev Genet,* 10**,** 392-404.
CORDOVA, E. J.,VELAZQUEZ-CRUZ, R.,CENTENO, F.,BACA, V. & OROZCO, L. 2010. The NRF2 gene variant, -653G/A, is associated with nephritis in childhood-onset systemic lupus erythematosus. *Lupus,* 19**,** 1237-42.
CORTEZ-PINTO, H.,CAMILO, M. E.,BAPTISTA, A.,DE OLIVEIRA, A. G. & DE MOURA, M. C. 1999. Non-alcoholic fatty liver: another feature of the metabolic syndrome? *Clin Nutr,* 18**,** 353-8.
CUSI, K. 2009. Nonalcoholic fatty liver disease in type 2 diabetes mellitus. *Curr Opin Endocrinol Diabetes Obes,* 16**,** 141-9.
DALTON, T. P.,CHEN, Y.,SCHNEIDER, S. N.,NEBERT, D. W. & SHERTZER, H. G. 2004. Genetically altered mice to evaluate glutathione homeostasis in health and disease. *Free Radic Biol Med,* 37**,** 1511-26.

DALY, A. K.,BALLESTRI, S.,CARULLI, L.,LORIA, P. & DAY, C. P. 2011. Genetic determinants of susceptibility and severity in nonalcoholic fatty liver disease. *Expert Rev Gastroenterol Hepatol,* 5**,** 253-63.
DALY, A. K.,FAIRBROTHER, K. S.,ANDREASSEN, O. A.,LONDON, S. J.,IDLE, J. R. & STEEN, V. M. 1996. Characterization and PCR-based detection of two different hybrid CYP2D7P/CYP2D6 alleles associated with the poor metabolizer phenotype. *Pharmacogenetics,* 6**,** 319-28.
DAVIS, G. L.,ALBRIGHT, J. E.,COOK, S. F. & ROSENBERG, D. M. 2003. Projecting future complications of chronic hepatitis C in the United States. *Liver Transpl,* 9**,** 331-8.
DAY, C. P. 2002. Non-alcoholic steatohepatitis (NASH): where are we now and where are we going? *Gut,* 50**,** 585-8.
DAY, C. P. 2006a. From fat to inflammation. *Gastroenterology,* 130**,** 207-10.
DAY, C. P. 2006b. Genes or environment to determine alcoholic liver disease and non-alcoholic fatty liver disease. *Liver Int,* 26**,** 1021-8.
DAY, C. P. 2006c. Non-alcoholic fatty liver disease: current concepts and management strategies. *Clin Med,* 6**,** 19-25.
DAY, C. P. 2010. Genetic and environmental susceptibility to non-alcoholic fatty liver disease. *Dig Dis,* 28**,** 255-60.
DAY, C. P. & JAMES, O. F. 1998. Steatohepatitis: a tale of two "hits"? *Gastroenterology,* 114**,** 842-5.
DAY CP, L. J., MCTERNAN P, MATHEW C, DALY A 2006. Genetic evidence for a role of gut flora in the pathogenesis of NASH in humans. *Hepatology,* 44**,** 261A-261A.
DE VISSER, J. A.,COOPER, T. F. & ELENA, S. F. 2011. The causes of epistasis. *Proc Biol Sci*.
DEEPA, M.,FAROOQ, S.,DATTA, M.,DEEPA, R. & MOHAN, V. 2007. Prevalence of metabolic syndrome using WHO, ATPIII and IDF definitions in Asian Indians: the Chennai Urban Rural Epidemiology Study (CURES-34). *Diabetes Metab Res Rev,* 23**,** 127-34.
DEGOUL, F.,SUTTON, A.,MANSOURI, A.,CEPANEC, C.,DEGOTT, C.,FROMENTY, B.,BEAUGRAND, M.,VALLA, D. & PESSAYRE, D. 2001. Homozygosity for alanine in the mitochondrial targeting sequence of superoxide dismutase and risk for severe alcoholic liver disease. *Gastroenterology,* 120**,** 1468-74.
DELEVE, L. D. & KAPLOWITZ, N. 1991. Glutathione metabolism and its role in hepatotoxicity. *Pharmacol Ther,* 52**,** 287-305.
DIAMOND, J. M. 1978. Channels in epithelial cell membranes and junctions. *Fed Proc,* 37**,** 2639-43.
DICKINSON, D. A.,LEVONEN, A. L.,MOELLERING, D. R.,ARNOLD, E. K.,ZHANG, H.,DARLEY-USMAR, V. M. & FORMAN, H. J. 2004. Human glutamate cysteine ligase gene regulation through the electrophile response element. *Free Radic Biol Med,* 37**,** 1152-9.
DIXON, J. B.,BHATHAL, P. S.,HUGHES, N. R. & O'BRIEN, P. E. 2004. Nonalcoholic fatty liver disease: Improvement in liver histological analysis with weight loss. *Hepatology,* 39**,** 1647-54.
DIXON, J. B.,BHATHAL, P. S. & O'BRIEN, P. E. 2001. Nonalcoholic fatty liver disease: predictors of nonalcoholic steatohepatitis and liver fibrosis in the severely obese. *Gastroenterology,* 121**,** 91-100.
DONG, H.,WANG, J.,LI, C.,HIROSE, A.,NOZAKI, Y.,TAKAHASHI, M.,ONO, M.,AKISAWA, N.,IWASAKI, S.,SAIBARA, T. & ONISHI, S. 2007. The phosphatidylethanolamine N-methyltransferase gene V175M single nucleotide

polymorphism confers the susceptibility to NASH in Japanese population. *J Hepatol,* 46**,** 915-20.

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. & FARGION, S. 2010. Genetic variants regulating insulin receptor signalling are associated with the severity of liver damage in patients with non-alcoholic fatty liver disease. *Gut,* 59**,** 267-73.

DONNELLY, K. L.,SMITH, C. I.,SCHWARZENBERG, S. J.,JESSURUN, J.,BOLDT, M. D. & 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**,** 1343-51.

DOWMAN, J. K.,TOMLINSON, J. W. & NEWSOME, P. N. 2011. Systematic review: the diagnosis and staging of non-alcoholic fatty liver disease and non-alcoholic steatohepatitis. *Aliment Pharmacol Ther,* 33**,** 525-40.

DUMAS, M. E.,BARTON, R. H.,TOYE, A.,CLOAREC, O.,BLANCHER, C.,ROTHWELL, A.,FEARNSIDE, J.,TATOUD, R.,BLANC, V.,LINDON, J. C.,MITCHELL, S. C.,HOLMES, E.,MCCARTHY, M. I.,SCOTT, J.,GAUGUIER, D. & NICHOLSON, J. K. 2006. Metabolic profiling reveals a contribution of gut microbiota to fatty liver phenotype in insulin-resistant mice. *Proc Natl Acad Sci U S A,* 103**,** 12511-6.

DURBIN, R. M.,ABECASIS, G. R.,ALTSHULER, D. L.,AUTON, A.,BROOKS, L. D.,GIBBS, R. A.,HURLES, M. E. & MCVEAN, G. A. 2010. A map of human genome variation from population-scale sequencing. *Nature,* 467**,** 1061-73.

ECKEL, R. H.,GRUNDY, S. M. & ZIMMET, P. Z. 2005. The metabolic syndrome. *Lancet,* 365**,** 1415-28.

EINHORN, D.,REAVEN, G. M.,COBIN, R. H.,FORD, E.,GANDA, O. P.,HANDELSMAN, Y.,HELLMAN, R.,JELLINGER, P. S.,KENDALL, D.,KRAUSS, R. M.,NEUFELD, N. D.,PETAK, S. M.,RODBARD, H. W.,SEIBEL, J. A.,SMITH, D. A. & WILSON, P. W. 2003. American College of Endocrinology position statement on the insulin resistance syndrome. *Endocr Pract,* 9**,** 237-52.

ERVIN, R. B. 2009. Prevalence of metabolic syndrome among adults 20 years of age and over, by sex, age, race and ethnicity, and body mass index: United States, 2003-2006. *Natl Health Stat Report***,** 1-7.

FABBRINI, E.,SULLIVAN, S. & KLEIN, S. 2010. Obesity and nonalcoholic fatty liver disease: biochemical, metabolic, and clinical implications. *Hepatology,* 51**,** 679-89.

FARHADI, A.,GUNDLAPALLI, S.,SHAIKH, M.,FRANTZIDES, C.,HARRELL, L.,KWASNY, M. M. & KESHAVARZIAN, A. 2008. Susceptibility to gut leakiness: a possible mechanism for endotoxaemia in non-alcoholic steatohepatitis. *Liver Int,* 28**,** 1026-33.

FARRELL, G. C. & LARTER, C. Z. 2006. Nonalcoholic fatty liver disease: from steatosis to cirrhosis. *Hepatology,* 43**,** S99-S112.

FERNANDEZ-CHECA, J. C. & KAPLOWITZ, N. 2005. Hepatic mitochondrial glutathione: transport and role in disease and toxicity. *Toxicol Appl Pharmacol,* 204**,** 263-73.

FORSBERG, L.,DE FAIRE, U. & MORGENSTERN, R. 1999. Low yield of polymorphisms from EST blast searching: analysis of genes related to oxidative stress and verification of the P197L polymorphism in GPX1. *Hum Mutat,* 13**,** 294-300.

FOSTER, T.,BUDOFF, M. J.,SAAB, S.,AHMADI, N.,GORDON, C. & GUERCI, A. D. 2011. Atorvastatin and antioxidants for the treatment of nonalcoholic fatty liver disease: the St Francis Heart Study randomized clinical trial. *Am J Gastroenterol,* 106**,** 71-7.

FRACANZANI, A. L.,VALENTI, L.,BUGIANESI, E.,ANDREOLETTI, M.,COLLI, A.,VANNI, E.,BERTELLI, C.,FATTA, E.,BIGNAMINI, D.,MARCHESINI, G. & FARGION, S. 2008. Risk of severe liver disease in nonalcoholic fatty liver disease with normal aminotransferase levels: a role for insulin resistance and diabetes. *Hepatology,* 48**,** 792-8.

FRANKLIN, C. C.,BACKOS, D. S.,MOHAR, I.,WHITE, C. C.,FORMAN, H. J. & KAVANAGH, T. J. 2009. Structure, function, and post-translational regulation of the catalytic and modifier subunits of glutamate cysteine ligase. *Mol Aspects Med,* 30**,** 86-98.

FRIDOVICH, I. 1978. The biology of oxygen radicals. *Science,* 201**,** 875-80.

FUKUSHIMA-UESAKA, H.,SAITO, Y.,MAEKAWA, K.,KAMATANI, N.,KAJIO, H.,KUZUYA, N.,NODA, M.,YASUDA, K. & SAWADA, J. 2007. Genetic variations and haplotype structures of transcriptional factor Nrf2 and its cytosolic reservoir protein Keap1 in Japanese. *Drug Metab Pharmacokinet,* 22**,** 212-9.

FURUSE, M.,FUJITA, K.,HIIRAGI, T.,FUJIMOTO, K. & TSUKITA, S. 1998. Claudin-1 and -2: novel integral membrane proteins localizing at tight junctions with no sequence similarity to occludin. *J Cell Biol,* 141**,** 1539-50.

FURUSE, M.,HIRASE, T.,ITOH, M.,NAGAFUCHI, A.,YONEMURA, S. & TSUKITA, S. 1993. Occludin: a novel integral membrane protein localizing at tight junctions. *J Cell Biol,* 123**,** 1777-88.

GABELE, E.,BRENNER, D. A. & RIPPE, R. A. 2003. Liver fibrosis: signals leading to the amplification of the fibrogenic hepatic stellate cell. *Front Biosci,* 8**,** d69-77.

GABRIEL, S. B.,SCHAFFNER, S. F.,NGUYEN, H.,MOORE, J. M.,ROY, J.,BLUMENSTIEL, B.,HIGGINS, J.,DEFELICE, M.,LOCHNER, A.,FAGGART, M.,LIU-CORDERO, S. N.,ROTIMI, C.,ADEYEMO, A.,COOPER, R.,WARD, R.,LANDER, E. S.,DALY, M. J. & ALTSHULER, D. 2002. The structure of haplotype blocks in the human genome. *Science,* 296**,** 2225-9.

GAULTON, K. J.,WILLER, C. J.,LI, Y.,SCOTT, L. J.,CONNEELY, K. N.,JACKSON, A. U.,DUREN, W. L.,CHINES, P. S.,NARISU, N.,BONNYCASTLE, L. L.,LUO, J.,TONG, M.,SPRAU, A. G.,PUGH, E. W.,DOHENY, K. F.,VALLE, T. T.,ABECASIS, G. R.,TUOMILEHTO, J.,BERGMAN, R. N.,COLLINS, F. S.,BOEHNKE, M. & MOHLKE, K. L. 2008. Comprehensive association study of type 2 diabetes and related quantitative traits with 222 candidate genes. *Diabetes,* 57**,** 3136-44.

GEORGE, D. K.,GOLDWURM, S.,MACDONALD, G. A.,COWLEY, L. L.,WALKER, N. I.,WARD, P. J.,JAZWINSKA, E. C. & POWELL, L. W. 1998. Increased hepatic iron concentration in nonalcoholic steatohepatitis is associated with increased fibrosis. *Gastroenterology,* 114**,** 311-8.

GIBBS, R. A. 2003. The International HapMap Project. *Nature,* 426**,** 789-96.

GILMORE, I. T.,BURROUGHS, A.,MURRAY-LYON, I. M.,WILLIAMS, R.,JENKINS, D. & HOPKINS, A. 1995. Indications, methods, and outcomes of percutaneous liver biopsy in England and Wales: an audit by the British Society of Gastroenterology and the Royal College of Physicians of London. *Gut,* 36**,** 437-41.

GIPP, J. J.,BAILEY, H. H. & MULCAHY, R. T. 1995. Cloning and sequencing of the cDNA for the light subunit of human liver gamma-glutamylcysteine synthetase

and relative mRNA levels for heavy and light subunits in human normal tissues. *Biochem Biophys Res Commun,* 206**,** 584-9.

GOLDSTEIN, D. B. 2009. Common genetic variation and human traits. *N Engl J Med,* 360**,** 1696-8.

GONG, P. & CEDERBAUM, A. I. 2006. Nrf2 is increased by CYP2E1 in rodent liver and HepG2 cells and protects against oxidative stress caused by CYP2E1. *Hepatology,* 43**,** 144-53.

GORAN, M. I.,BALL, G. D. & CRUZ, M. L. 2003. Obesity and risk of type 2 diabetes and cardiovascular disease in children and adolescents. *J Clin Endocrinol Metab,* 88**,** 1417-27.

GRANT, C. M.,MACIVER, F. H. & DAWES, I. W. 1997. Glutathione synthetase is dispensable for growth under both normal and oxidative stress conditions in the yeast Saccharomyces cerevisiae due to an accumulation of the dipeptide gamma-glutamylcysteine. *Mol Biol Cell,* 8**,** 1699-707.

GRANT, S. F.,THORLEIFSSON, G.,REYNISDOTTIR, I.,BENEDIKTSSON, R.,MANOLESCU, A.,SAINZ, J.,HELGASON, A.,STEFANSSON, H.,EMILSSON, V.,HELGADOTTIR, A.,STYRKARSDOTTIR, U.,MAGNUSSON, K. P.,WALTERS, G. B.,PALSDOTTIR, E.,JONSDOTTIR, T.,GUDMUNDSDOTTIR, T.,GYLFASON, A.,SAEMUNDSDOTTIR, J.,WILENSKY, R. L.,REILLY, M. P.,RADER, D. J.,BAGGER, Y.,CHRISTIANSEN, C.,GUDNASON, V.,SIGURDSSON, G.,THORSTEINSDOTTIR, U.,GULCHER, J. R.,KONG, A. & STEFANSSON, K. 2006. Variant of transcription factor 7-like 2 (TCF7L2) gene confers risk of type 2 diabetes. *Nat Genet,* 38**,** 320-3.

GREFHORST, A.,ELZINGA, B. M.,VOSHOL, P. J.,PLOSCH, T.,KOK, T.,BLOKS, V. W.,VAN DER SLUIJS, F. H.,HAVEKES, L. M.,ROMIJN, J. A.,VERKADE, H. J. & KUIPERS, F. 2002. Stimulation of lipogenesis by pharmacological activation of the liver X receptor leads to production of large, triglyceride-rich very low density lipoprotein particles. *J Biol Chem,* 277**,** 34182-90.

GUNZEL, D.,STUIVER, M.,KAUSALYA, P. J.,HAISCH, L.,KRUG, S. M.,ROSENTHAL, R.,MEIJ, I. C.,HUNZIKER, W.,FROMM, M. & MULLER, D. 2009. Claudin-10 exists in six alternatively spliced isoforms that exhibit distinct localization and function. *J Cell Sci,* 122**,** 1507-17.

GYSIN, R.,KRAFTSIK, R.,SANDELL, J.,BOVET, P.,CHAPPUIS, C.,CONUS, P.,DEPPEN, P.,PREISIG, M.,RUIZ, V.,STEULLET, P.,TOSIC, M.,WERGE, T.,CUENOD, M. & DO, K. Q. 2007. Impaired glutathione synthesis in schizophrenia: convergent genetic and functional evidence. *Proc Natl Acad Sci U S A,* 104**,** 16621-6.

HAKONARSON, H. & GRANT, S. F. 2011. GWAS and its impact on elucidating the etiology of diabetes. *Diabetes Metab Res Rev.*

HARVEY, C. J.,THIMMULAPPA, R. K.,SINGH, A.,BLAKE, D. J.,LING, G.,WAKABAYASHI, N.,FUJII, J.,MYERS, A. & BISWAL, S. 2009. Nrf2-regulated glutathione recycling independent of biosynthesis is critical for cell survival during oxidative stress. *Free Radic Biol Med,* 46**,** 443-53.

HE, S.,MCPHAUL, C.,LI, J. Z.,GARUTI, R.,KINCH, L.,GRISHIN, N. V.,COHEN, J. C. & HOBBS, H. H. 2010. A sequence variation (I148M) in PNPLA3 associated with nonalcoholic fatty liver disease disrupts triglyceride hydrolysis. *J Biol Chem,* 285**,** 6706-15.

HEGELE, R. A.,CONNELLY, P. W.,HANLEY, A. J.,SUN, F.,HARRIS, S. B. & ZINMAN, B. 1997. Common genomic variation in the APOC3 promoter associated with variation in plasma lipoproteins. *Arterioscler Thromb Vasc Biol,* 17**,** 2753-8.

HGP 2004. Finishing the euchromatic sequence of the human genome. *Nature,* 431**,** 931-45.

HICKMAN, I. J.,JONSSON, J. R.,PRINS, J. B.,ASH, S.,PURDIE, D. M.,CLOUSTON, A. D. & POWELL, E. E. 2004. Modest weight loss and physical activity in overweight patients with chronic liver disease results in sustained improvements in alanine aminotransferase, fasting insulin, and quality of life. *Gut,* 53**,** 413-9.

HONG, C. J.,HUO, S. J.,LIAO, D. L.,LEE, K.,WU, J. Y. & TSAI, S. J. 2004. Case-control and family-based association studies between the neuregulin 1 (Arg38Gln) polymorphism and schizophrenia. *Neurosci Lett,* 366**,** 158-61.

HORIKAWA, S. & TSUKADA, K. 1992. Molecular cloning and developmental expression of a human kidney S-adenosylmethionine synthetase. *FEBS Lett,* 312**,** 37-41.

HORTON, J. D.,GOLDSTEIN, J. L. & BROWN, M. S. 2002. SREBPs: activators of the complete program of cholesterol and fatty acid synthesis in the liver. *J Clin Invest,* 109**,** 1125-31.

HOSSAIN, N.,AFENDY, A.,STEPANOVA, M.,NADER, F.,SRISHORD, M.,RAFIQ, N.,GOODMAN, Z. & YOUNOSSI, Z. 2009. Independent predictors of fibrosis in patients with nonalcoholic fatty liver disease. *Clin Gastroenterol Hepatol,* 7**,** 1224-9, 1229 e1-2.

HOTAMISLIGIL, G. S. 2005. Role of endoplasmic reticulum stress and c-Jun NH2-terminal kinase pathways in inflammation and origin of obesity and diabetes. *Diabetes,* 54 Suppl 2**,** S73-8.

HOTTA, K.,YONEDA, M.,HYOGO, H.,OCHI, H.,MIZUSAWA, S.,UENO, T.,CHAYAMA, K.,NAKAJIMA, A.,NAKAO, K. & SEKINE, A. 2010. Association of the rs738409 polymorphism in PNPLA3 with liver damage and the development of nonalcoholic fatty liver disease. *BMC Med Genet,* 11**,** 172.

HU, Y. J. & DIAMOND, A. M. 2003. Role of glutathione peroxidase 1 in breast cancer: loss of heterozygosity and allelic differences in the response to selenium. *Cancer Res,* 63**,** 3347-51.

HU, Z. W.,LUO, H. B.,XU, Y. M.,GUO, J. W.,DENG, X. L.,TONG, Y. W. & TANG, X. 2009. Tumor necrosis factor--alpha gene promoter polymorphisms in Chinese patients with nonalcoholic fatty liver diseases. *Acta Gastroenterol Belg,* 72**,** 215-21.

HUI, J. M.,KENCH, J. G.,CHITTURI, S.,SUD, A.,FARRELL, G. C.,BYTH, K.,HALL, P.,KHAN, M. & GEORGE, J. 2003. Long-term outcomes of cirrhosis in nonalcoholic steatohepatitis compared with hepatitis C. *Hepatology,* 38**,** 420-7.

HWANG, C.,SINSKEY, A. J. & LODISH, H. F. 1992. Oxidized redox state of glutathione in the endoplasmic reticulum. *Science,* 257**,** 1496-502.

IKENOUCHI, J.,FURUSE, M.,FURUSE, K.,SASAKI, H. & TSUKITA, S. 2005. Tricellulin constitutes a novel barrier at tricellular contacts of epithelial cells. *J Cell Biol,* 171**,** 939-45.

INAI, T.,SENGOKU, A.,GUAN, X.,HIROSE, E.,IIDA, H. & SHIBATA, Y. 2005. Heterogeneity in expression and subcellular localization of tight junction proteins, claudin-10 and -15, examined by RT-PCR and immunofluorescence microscopy. *Arch Histol Cytol,* 68**,** 349-60.

IP, Y. C.,CHEUNG, S. T.,LEE, Y. T.,HO, J. C. & FAN, S. T. 2007. Inhibition of hepatocellular carcinoma invasion by suppression of claudin-10 in HLE cells. *Mol Cancer Ther,* 6**,** 2858-67.

ITOH, K.,WAKABAYASHI, N.,KATOH, Y.,ISHII, T.,IGARASHI, K.,ENGEL, J. D. & YAMAMOTO, M. 1999. Keap1 represses nuclear activation of antioxidant responsive elements by Nrf2 through binding to the amino-terminal Neh2 domain. *Genes Dev,* 13**,** 76-86.

JANSEN, P. L.,MULDER, G. J.,BURCHELL, B. & BOCK, K. W. 1992. New developments in glucuronidation research: report of a workshop on "glucuronidation, its role in health and disease". *Hepatology,* 15**,** 532-44.

JI, C. Y. 2008. The prevalence of childhood overweight/obesity and the epidemic changes in 1985-2000 for Chinese school-age children and adolescents. *Obes Rev,* 9 Suppl 1**,** 78-81.

JIMBA, S.,NAKAGAMI, T.,TAKAHASHI, M.,WAKAMATSU, T.,HIROTA, Y.,IWAMOTO, Y. & WASADA, T. 2005. Prevalence of non-alcoholic fatty liver disease and its association with impaired glucose metabolism in Japanese adults. *Diabet Med,* 22**,** 1141-5.

JUN, D. W.,HAN, J. H.,JANG, E. C.,KIM, S. H.,JO, Y. J.,PARK, Y. S. & CHAE, J. D. 2009. Polymorphisms of microsomal triglyceride transfer protein gene and phosphatidylethanolamine N-methyltransferase gene in alcoholic and nonalcoholic fatty liver disease in Koreans. *Eur J Gastroenterol Hepatol,* 21**,** 667-72.

KANZOK, S. M.,FECHNER, A.,BAUER, H.,ULSCHMID, J. K.,MULLER, H. M.,BOTELLA-MUNOZ, J.,SCHNEUWLY, S.,SCHIRMER, R. & BECKER, K. 2001. Substitution of the thioredoxin system for glutathione reductase in Drosophila melanogaster. *Science,* 291**,** 643-6.

KAPLOWITZ, N.,AW, T. Y. & OOKHTENS, M. 1985. The regulation of hepatic glutathione. *Annu Rev Pharmacol Toxicol,* 25**,** 715-44.

KARATHANASIS, S. K. 1985. Apolipoprotein multigene family: tandem organization of human apolipoprotein AI, CIII, and AIV genes. *Proc Natl Acad Sci U S A,* 82**,** 6374-8.

KARIMI, M.,GOLDIE, L. C.,CRUICKSHANK, M. N.,MOSES, E. K. & ABRAHAM, L. J. 2009. A critical assessment of the factors affecting reporter gene assays for promoter SNP function: a reassessment of -308 TNF polymorphism function using a novel integrated reporter system. *Eur J Hum Genet,* 17**,** 1454-62.

KERR, T. A.,KORENBLAT, K. M. & DAVIDSON, N. O. 2011. MicroRNAs and liver disease. *Transl Res,* 157**,** 241-52.

KISHI, T.,IKEDA, M.,KITAJIMA, T.,YAMANOUCHI, Y.,KINOSHITA, Y.,KAWASHIMA, K.,INADA, T.,HARANO, M.,KOMIYAMA, T.,HORI, T.,YAMADA, M.,IYO, M.,SORA, I.,SEKINE, Y.,OZAKI, N.,UJIKE, H. & IWATA, N. 2008. Glutamate cysteine ligase modifier (GCLM) subunit gene is not associated with methamphetamine-use disorder or schizophrenia in the Japanese population. *Ann N Y Acad Sci,* 1139**,** 63-9.

KLEIN, R. J.,ZEISS, C.,CHEW, E. Y.,TSAI, J. Y.,SACKLER, R. S.,HAYNES, C.,HENNING, A. K.,SANGIOVANNI, J. P.,MANE, S. M.,MAYNE, S. T.,BRACKEN, M. B.,FERRIS, F. L.,OTT, J.,BARNSTABLE, C. & HOH, J. 2005. Complement factor H polymorphism in age-related macular degeneration. *Science,* 308**,** 385-9.

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. & SANYAL, A. J. 2005. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology,* 41**,** 1313-21.

KNOWLER, W. C.,BARRETT-CONNOR, E.,FOWLER, S. E.,HAMMAN, R. F.,LACHIN, J. M.,WALKER, E. A. & NATHAN, D. M. 2002. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *N Engl J Med,* 346**,** 393-403.

KOEK, G. H.,LIEDORP, P. R. & BAST, A. 2011. The role of oxidative stress in non-alcoholic steatohepatitis. *Clin Chim Acta*.

KOHJIMA, M.,ENJOJI, M.,HIGUCHI, N.,KATO, M.,KOTOH, K.,YOSHIMOTO, T.,FUJINO, T.,YADA, M.,YADA, R.,HARADA, N.,TAKAYANAGI, R. & NAKAMUTA, M. 2007. Re-evaluation of fatty acid metabolism-related gene expression in nonalcoholic fatty liver disease. *Int J Mol Med,* 20**,** 351-8.

KOIDE, S.,KUGIYAMA, K.,SUGIYAMA, S.,NAKAMURA, S.,FUKUSHIMA, H.,HONDA, O.,YOSHIMURA, M. & OGAWA, H. 2003. Association of polymorphism in glutamate-cysteine ligase catalytic subunit gene with coronary vasomotor dysfunction and myocardial infarction. *J Am Coll Cardiol,* 41**,** 539-45.

KOTB, M.,MUDD, S. H.,MATO, J. M.,GELLER, A. M.,KREDICH, N. M.,CHOU, J. Y. & CANTONI, G. L. 1997. Consensus nomenclature for the mammalian methionine adenosyltransferase genes and gene products. *Trends Genet,* 13**,** 51-2.

KOTRONEN, A.,JOHANSSON, L. E.,JOHANSSON, L. M.,ROOS, C.,WESTERBACKA, J.,HAMSTEN, A.,BERGHOLM, R.,ARKKILA, P.,AROLA, J.,KIVILUOTO, T.,FISHER, R. M.,EHRENBORG, E.,ORHO-MELANDER, M.,RIDDERSTRALE, M.,GROOP, L. & YKI-JARVINEN, H. 2009a. A common variant in PNPLA3, which encodes adiponutrin, is associated with liver fat content in humans. *Diabetologia,* 52**,** 1056-60.

KOTRONEN, A.,PELTONEN, M.,HAKKARAINEN, A.,SEVASTIANOVA, K.,BERGHOLM, R.,JOHANSSON, L. M.,LUNDBOM, N.,RISSANEN, A.,RIDDERSTRALE, M.,GROOP, L.,ORHO-MELANDER, M. & YKI-JARVINEN, H. 2009b. Prediction of non-alcoholic fatty liver disease and liver fat using metabolic and genetic factors. *Gastroenterology,* 137**,** 865-72.

KOZLITINA, J.,BOERWINKLE, E.,COHEN, J. C. & HOBBS, H. H. 2010. Dissociation between APOC3 variants, hepatic triglyceride content and insulin resistance. *Hepatology*.

KOZLITINA, J.,BOERWINKLE, E.,COHEN, J. C. & HOBBS, H. H. 2011. Dissociation between APOC3 variants, hepatic triglyceride content and insulin resistance. *Hepatology,* 53**,** 467-74.

KRAL, J. G.,SCHAFFNER, F.,PIERSON, R. N., JR. & WANG, J. 1993. Body fat topography as an independent predictor of fatty liver. *Metabolism,* 42**,** 548-51.

KRAL, J. G.,THUNG, S. N.,BIRON, S.,HOULD, F. S.,LEBEL, S.,MARCEAU, S.,SIMARD, S. & MARCEAU, P. 2004. Effects of surgical treatment of the metabolic syndrome on liver fibrosis and cirrhosis. *Surgery,* 135**,** 48-58.

KRAUSE, G.,WINKLER, L.,MUELLER, S. L.,HASELOFF, R. F.,PIONTEK, J. & BLASIG, I. E. 2008. Structure and function of claudins. *Biochim Biophys Acta,* 1778**,** 631-45.

KUZUYA, M.,ANDO, F.,IGUCHI, A. & SHIMOKATA, H. 2008. Glutathione peroxidase 1 Pro198Leu variant contributes to the metabolic syndrome in men in a large Japanese cohort. *Am J Clin Nutr,* 87**,** 1939-44.

KWON, J. M. & GOATE, A. M. 2000. The candidate gene approach. *Alcohol Res Health,* 24**,** 164-8.

KWONG, M.,KAN, Y. W. & CHAN, J. Y. 1999. The CNC basic leucine zipper factor, Nrf1, is essential for cell survival in response to oxidative stress-inducing agents. Role for Nrf1 in gamma-gcs(l) and gss expression in mouse fibroblasts. *J Biol Chem,* 274**,** 37491-8.

LAI, C. Q.,PARNELL, L. D.,TROEN, A. M.,SHEN, J.,CAOUETTE, H.,WARODOMWICHIT, D.,LEE, Y. C.,CROTT, J. W.,QIU, W. Q.,ROSENBERG, I. H.,TUCKER, K. L. & ORDOVAS, J. M. 2010. MAT1A variants are associated with hypertension, stroke, and markers of DNA damage

and are modulated by plasma vitamin B-6 and folate. *Am J Clin Nutr,* 91**,** 1377-86.

LAMPE, J. W.,BIGLER, J.,HORNER, N. K. & POTTER, J. D. 1999. UDP-glucuronosyltransferase (UGT1A1*28 and UGT1A6*2) polymorphisms in Caucasians and Asians: relationships to serum bilirubin concentrations. *Pharmacogenetics,* 9**,** 341-9.

LARTER, C. Z.,CHITTURI, S.,HEYDET, D. & FARRELL, G. C. 2010. A fresh look at NASH pathogenesis. Part 1: the metabolic movers. *J Gastroenterol Hepatol,* 25**,** 672-90.

LARTER, C. Z.,YEH, M. M.,VAN ROOYEN, D. M.,TEOH, N. C.,BROOLING, J.,HOU, J. Y.,WILLIAMS, J.,CLYNE, M.,NOLAN, C. J. & FARRELL, G. C. 2009. Roles of adipose restriction and metabolic factors in progression of steatosis to steatohepatitis in obese, diabetic mice. *J Gastroenterol Hepatol,* 24**,** 1658-68.

LEBOVICS, E. & RUBIN, J. 2011. Baffled by NAFLD? non-alcoholic fatty liver disease: why you should care, when you should worry, what you should do. *Diabetes Metab Res Rev*.

LEBOVITZ, R. M.,ZHANG, H.,VOGEL, H.,CARTWRIGHT, J., JR.,DIONNE, L.,LU, N.,HUANG, S. & MATZUK, M. M. 1996. Neurodegeneration, myocardial injury, and perinatal death in mitochondrial superoxide dismutase-deficient mice. *Proc Natl Acad Sci U S A,* 93**,** 9782-7.

LEE, J. M.,LI, J.,JOHNSON, D. A.,STEIN, T. D.,KRAFT, A. D.,CALKINS, M. J.,JAKEL, R. J. & JOHNSON, J. A. 2005. Nrf2, a multi-organ protector? *Faseb J,* 19**,** 1061-6.

LEI, X. G. 2001. Glutathione peroxidase-1 gene knockout on body antioxidant defense in mice. *Biofactors,* 14**,** 93-9.

LEITE, N. C.,SALLES, G. F.,ARAUJO, A. L.,VILLELA-NOGUEIRA, C. A. & CARDOSO, C. R. 2009. Prevalence and associated factors of non-alcoholic fatty liver disease in patients with type-2 diabetes mellitus. *Liver Int,* 29**,** 113-9.

LEUNG, L.,KWONG, M.,HOU, S.,LEE, C. & CHAN, J. Y. 2003. Deficiency of the Nrf1 and Nrf2 transcription factors results in early embryonic lethality and severe oxidative stress. *J Biol Chem,* 278**,** 48021-9.

LI, W. W.,DAMMERMAN, M. M.,SMITH, J. D.,METZGER, S.,BRESLOW, J. L. & LEFF, T. 1995a. Common genetic variation in the promoter of the human apo CIII gene abolishes regulation by insulin and may contribute to hypertriglyceridemia. *J Clin Invest,* 96**,** 2601-5.

LI, Y.,HUANG, T. T.,CARLSON, E. J.,MELOV, S.,URSELL, P. C.,OLSON, J. L.,NOBLE, L. J.,YOSHIMURA, M. P.,BERGER, C.,CHAN, P. H.,WALLACE, D. C. & EPSTEIN, C. J. 1995b. Dilated cardiomyopathy and neonatal lethality in mutant mice lacking manganese superoxide dismutase. *Nat Genet,* 11**,** 376-81.

LIEBER, C. S.,LEO, M. A.,MAK, K. M.,XU, Y.,CAO, Q.,REN, C.,PONOMARENKO, A. & DECARLI, L. M. 2004. Model of nonalcoholic steatohepatitis. *Am J Clin Nutr,* 79**,** 502-9.

LIN, J. P.,O'DONNELL, C. J.,SCHWAIGER, J. P.,CUPPLES, L. A.,LINGENHEL, A.,HUNT, S. C.,YANG, S. & KRONENBERG, F. 2006. Association between the UGT1A1*28 allele, bilirubin levels, and coronary heart disease in the Framingham Heart Study. *Circulation,* 114**,** 1476-81.

LIN, Y. C.,CHANG, P. F.,HU, F. C.,CHANG, M. H. & NI, Y. H. 2009. Variants in the UGT1A1 gene and the risk of pediatric nonalcoholic fatty liver disease. *Pediatrics,* 124**,** e1221-7.

LIU, Q.,BENGMARK, S. & QU, S. 2010. The role of hepatic fat accumulation in pathogenesis of non-alcoholic fatty liver disease (NAFLD). *Lipids Health Dis,* 9**,** 42.

LOGUERCIO, C.,DE GIROLAMO, V.,DE SIO, I.,TUCCILLO, C.,ASCIONE, A.,BALDI, F.,BUDILLON, G.,CIMINO, L.,DI CARLO, A.,DI MARINO, M. P.,MORISCO, F.,PICCIOTTO, F.,TERRACCIANO, L.,VECCHIONE, R.,VERDE, V. & DEL VECCHIO BLANCO, C. 2001. Non-alcoholic fatty liver disease in an area of southern Italy: main clinical, histological, and pathophysiological aspects. *J Hepatol,* 35**,** 568-74.

LOOMBA, R.,SIRLIN, C. B.,SCHWIMMER, J. B. & LAVINE, J. E. 2009. Advances in pediatric nonalcoholic fatty liver disease. *Hepatology,* 50**,** 1282-93.

LU, S. C. 1999. Regulation of hepatic glutathione synthesis: current concepts and controversies. *FASEB J,* 13**,** 1169-83.

LU, S. C. 2009. Regulation of glutathione synthesis. *Mol Aspects Med,* 30**,** 42-59.

LU, S. C.,ALVAREZ, L.,HUANG, Z. Z.,CHEN, L.,AN, W.,CORRALES, F. J.,AVILA, M. A.,KANEL, G. & 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**,** 5560-5.

LUDWIG, J.,VIGGIANO, T. R.,MCGILL, D. B. & OH, B. J. 1980. Nonalcoholic steatohepatitis: Mayo Clinic experiences with a hitherto unnamed disease. *Mayo Clin Proc,* 55**,** 434-8.

MACMILLAN-CROW, L. A. & CRUTHIRDS, D. L. 2001. Invited review: manganese superoxide dismutase in disease. *Free Radic Res,* 34**,** 325-36.

MAHLEY, R. W.,INNERARITY, T. L.,RALL, S. C., JR. & WEISGRABER, K. H. 1984. Plasma lipoproteins: apolipoprotein structure and function. *J Lipid Res,* 25**,** 1277-94.

MAKKONEN, J.,PIETILAINEN, K. H.,RISSANEN, A.,KAPRIO, J. & 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**,** 1035-42.

MALAGUARNERA, M.,DI ROSA, M.,NICOLETTI, F. & MALAGUARNERA, L. 2009. Molecular mechanisms involved in NAFLD progression. *J Mol Med,* 87**,** 679-95.

MARCHESINI, G.,BUGIANESI, E.,FORLANI, G.,CERRELLI, F.,LENZI, M.,MANINI, R.,NATALE, S.,VANNI, E.,VILLANOVA, N.,MELCHIONDA, N. & RIZZETTO, M. 2003. Nonalcoholic fatty liver, steatohepatitis, and the metabolic syndrome. *Hepatology,* 37**,** 917-23.

MARKLUND, S. L. 1982. Human copper-containing superoxide dismutase of high molecular weight. *Proc Natl Acad Sci U S A,* 79**,** 7634-8.

MARTINEZ-CHANTAR, M. L.,CORRALES, F. J.,MARTINEZ-CRUZ, L. A.,GARCIA-TREVIJANO, E. R.,HUANG, Z. Z.,CHEN, L.,KANEL, G.,AVILA, M. A.,MATO, J. M. & LU, S. C. 2002. Spontaneous oxidative stress and liver tumors in mice lacking methionine adenosyltransferase 1A. *Faseb J,* 16**,** 1292-4.

MATO, J. M.,CAMARA, J.,FERNANDEZ DE PAZ, J.,CABALLERIA, L.,COLL, S.,CABALLERO, A.,GARCIA-BUEY, L.,BELTRAN, J.,BENITA, V.,CABALLERIA, J.,SOLA, R.,MORENO-OTERO, R.,BARRAO, F.,MARTIN-DUCE, A.,CORREA, J. A.,PARES, A.,BARRAO, E.,GARCIA-MAGAZ, I.,PUERTA, J. L.,MORENO, J.,BOISSARD, G.,ORTIZ, P. & RODES, J. 1999. S-adenosylmethionine in alcoholic liver cirrhosis: a randomized, placebo-controlled, double-blind, multicenter clinical trial. *J Hepatol,* 30**,** 1081-9.

MATO, J. M.,CORRALES, F. J.,LU, S. C. & AVILA, M. A. 2002. S-Adenosylmethionine: a control switch that regulates liver function. *Faseb J,* 16**,** 15-26.

MATTEONI, C. A.,YOUNOSSI, Z. M.,GRAMLICH, T.,BOPARAI, N.,LIU, Y. C. & MCCULLOUGH, A. J. 1999. Nonalcoholic fatty liver disease: a spectrum of clinical and pathological severity. *Gastroenterology,* 116**,** 1413-9.

MATTHIESSEN, J.,VELSING GROTH, M.,FAGT, S.,BILTOFT-JENSEN, A.,STOCKMARR, A.,ANDERSEN, J. S. & TROLLE, E. 2008. Prevalence and trends in overweight and obesity among children and adolescents in Denmark. *Scand J Public Health,* 36**,** 153-60.

MCCORD, J. M. & FRIDOVICH, I. 1969. The utility of superoxide dismutase in studying free radical reactions. I. Radicals generated by the interaction of sulfite, dimethyl sulfoxide, and oxygen. *J Biol Chem,* 244**,** 6056-63.

MEISTER, A. 1988. Glutathione metabolism and its selective modification. *J Biol Chem,* 263**,** 17205-8.

MEISTER, A. & ANDERSON, M. E. 1983. Glutathione. *Annu Rev Biochem,* 52**,** 711-60.

MELOV, S.,SCHNEIDER, J. A.,DAY, B. J.,HINERFELD, D.,COSKUN, P.,MIRRA, S. S.,CRAPO, J. D. & WALLACE, D. C. 1998. A novel neurological phenotype in mice lacking mitochondrial manganese superoxide dismutase. *Nat Genet,* 18**,** 159-63.

MENDEZ-SANCHEZ, N.,ARRESE, M.,ZAMORA-VALDES, D. & URIBE, M. 2007. Current concepts in the pathogenesis of nonalcoholic fatty liver disease. *Liver Int,* 27**,** 423-33.

MEREDITH, M. J. & REED, D. J. 1982. Status of the mitochondrial pool of glutathione in the isolated hepatocyte. *J Biol Chem,* 257**,** 3747-53.

MIELE, L.,BEALE, G.,PATMAN, G.,NOBILI, V.,LEATHART, J.,GRIECO, A.,ABATE, M.,FRIEDMAN, S. L.,NARLA, G.,BUGIANESI, E.,DAY, C. P. & REEVES, H. L. 2008. The Kruppel-like factor 6 genotype is associated with fibrosis in nonalcoholic fatty liver disease. *Gastroenterology,* 135**,** 282-291 e1.

MIELE, L.,VALENZA, V.,LA TORRE, G.,MONTALTO, M.,CAMMAROTA, G.,RICCI, R.,MASCIANA, R.,FORGIONE, A.,GABRIELI, M. L.,PEROTTI, G.,VECCHIO, F. M.,RAPACCINI, G.,GASBARRINI, G.,DAY, C. P. & GRIECO, A. 2009. Increased intestinal permeability and tight junction alterations in nonalcoholic fatty liver disease. *Hepatology,* 49**,** 1877-87.

MIGNONE, F.,GISSI, C.,LIUNI, S. & PESOLE, G. 2002. Untranslated regions of mRNAs. *Genome Biol,* 3**,** REVIEWS0004.

MITSUYOSHI, H.,YASUI, K.,HARANO, Y.,ENDO, M.,TSUJI, K.,MINAMI, M.,ITOH, Y.,OKANOUE, T. & YOSHIKAWA, T. 2009. Analysis of hepatic genes involved in the metabolism of fatty acids and iron in nonalcoholic fatty liver disease. *Hepatol Res,* 39**,** 366-73.

MIYAZAKI, Y.,MAHANKALI, A.,MATSUDA, M.,MAHANKALI, S.,HARDIES, J.,CUSI, K.,MANDARINO, L. J. & DEFRONZO, R. A. 2002. Effect of pioglitazone on abdominal fat distribution and insulin sensitivity in type 2 diabetic patients. *J Clin Endocrinol Metab,* 87**,** 2784-91.

MOINOVA, H. R. & MULCAHY, R. T. 1999. Up-regulation of the human gamma-glutamylcysteine synthetase regulatory subunit gene involves binding of Nrf-2 to an electrophile responsive element. *Biochem Biophys Res Commun,* 261**,** 661-8.

MOLLSTEN, A.,JORSAL, A.,LAJER, M.,VIONNET, N. & TARNOW, L. 2009. The V16A polymorphism in SOD2 is associated with increased risk of diabetic

nephropathy and cardiovascular disease in type 1 diabetes. *Diabetologia,* 52**,** 2590-3.

MOLLSTEN, A.,MARKLUND, S. L.,WESSMAN, M.,SVENSSON, M.,FORSBLOM, C.,PARKKONEN, M.,BRISMAR, K.,GROOP, P. H. & DAHLQUIST, G. 2007. A functional polymorphism in the manganese superoxide dismutase gene and diabetic nephropathy. *Diabetes,* 56**,** 265-9.

MOORE, J. B. 2010. Non-alcoholic fatty liver disease: the hepatic consequence of obesity and the metabolic syndrome. *Proc Nutr Soc,* 69**,** 211-20.

MULCAHY, R. T.,WARTMAN, M. A.,BAILEY, H. H. & GIPP, J. J. 1997. Constitutive and beta-naphthoflavone-induced expression of the human gamma-glutamylcysteine synthetase heavy subunit gene is regulated by a distal antioxidant response element/TRE sequence. *J Biol Chem,* 272**,** 7445-54.

MUSSO, G.,GAMBINO, R.,CASSADER, M. & PAGANO, G. 2010. Meta-analysis: Natural history of non-alcoholic fatty liver disease (NAFLD) and diagnostic accuracy of non-invasive tests for liver disease severity. *Ann Med*.

MUSSO, G.,GAMBINO, R.,DE MICHIELI, F.,DURAZZO, M.,PAGANO, G. & CASSADER, M. 2008. Adiponectin gene polymorphisms modulate acute adiponectin response to dietary fat: Possible pathogenetic role in NASH. *Hepatology,* 47**,** 1167-77.

MUTO, S.,HATA, M.,TANIGUCHI, J.,TSURUOKA, S.,MORIWAKI, K.,SAITOU, M.,FURUSE, K.,SASAKI, H.,FUJIMURA, A.,IMAI, M.,KUSANO, E.,TSUKITA, S. & FURUSE, M. 2010. Claudin-2-deficient mice are defective in the leaky and cation-selective paracellular permeability properties of renal proximal tubules. *Proc Natl Acad Sci U S A,* 107**,** 8011-6.

NAGANUMA, T.,NAKAYAMA, T.,SATO, N.,FU, Z.,SOMA, M.,AOI, N. & USAMI, R. 2008. A haplotype-based case-control study examining human extracellular superoxide dismutase gene and essential hypertension. *Hypertens Res,* 31**,** 1533-40.

NAHON, P.,SUTTON, A.,PESSAYRE, D.,RUFAT, P.,DEGOUL, F.,GANNE-CARRIE, N.,ZIOL, M.,CHARNAUX, N.,N'KONTCHOU, G.,TRINCHET, J. C.,GATTEGNO, L. & BEAUGRAND, M. 2005. Genetic dimorphism in superoxide dismutase and susceptibility to alcoholic cirrhosis, hepatocellular carcinoma, and death. *Clin Gastroenterol Hepatol,* 3**,** 292-8.

NAKAMURA, S.,KUGIYAMA, K.,SUGIYAMA, S.,MIYAMOTO, S.,KOIDE, S.,FUKUSHIMA, H.,HONDA, O.,YOSHIMURA, M. & OGAWA, H. 2002. Polymorphism in the 5'-flanking region of human glutamate-cysteine ligase modifier subunit gene is associated with myocardial infarction. *Circulation,* 105**,** 2968-73.

NAMIKAWA, C.,SHU-PING, Z.,VYSELAAR, J. R.,NOZAKI, Y.,NEMOTO, Y.,ONO, M.,AKISAWA, N.,SAIBARA, T.,HIROI, M.,ENZAN, H. & ONISHI, S. 2004. Polymorphisms of microsomal triglyceride transfer protein gene and manganese superoxide dismutase gene in non-alcoholic steatohepatitis. *J Hepatol,* 40**,** 781-6.

NARASIMHAN, S.,GOKULAKRISHNAN, K.,SAMPATHKUMAR, R.,FAROOQ, S.,RAVIKUMAR, R.,MOHAN, V. & BALASUBRAMANYAM, M. 2010. Oxidative stress is independently associated with non-alcoholic fatty liver disease (NAFLD) in subjects with and without type 2 diabetes. *Clin Biochem,* 43**,** 815-21.

NASSAR, A.,COHEN, C.,AGERSBORG, S. S.,ZHOU, W.,LYNCH, K. A.,HEYMAN, E. R.,OLSON, A.,LANGE, H. & SIDDIQUI, M. T. 2011. A new immunohistochemical ER/PR image analysis system: a multisite performance study. *Appl Immunohistochem Mol Morphol,* 19**,** 195-202.

NELSON, J. E.,BHATTACHARYA, R.,LINDOR, K. D.,CHALASANI, N.,RAAKA, S.,HEATHCOTE, E. J.,MISKOVSKY, E.,SHAFFER, E.,RULYAK, S. J. & KOWDLEY, K. V. 2007. HFE C282Y mutations are associated with advanced hepatic fibrosis in Caucasians with nonalcoholic steatohepatitis. *Hepatology,* 46**,** 723-9.

NEUMAN, R. J.,WASSON, J.,ATZMON, G.,WAINSTEIN, J.,YERUSHALMI, Y.,COHEN, J.,BARZILAI, N.,BLECH, I.,GLASER, B. & PERMUTT, M. A. 2010. Gene-gene interactions lead to higher risk for development of type 2 diabetes in an Ashkenazi Jewish population. *PLoS One,* 5**,** e9903.

NEUSCHWANDER-TETRI, B. A.,BRUNT, E. M.,WEHMEIER, K. R.,OLIVER, D. & BACON, B. R. 2003. Improved nonalcoholic steatohepatitis after 48 weeks of treatment with the PPAR-gamma ligand rosiglitazone. *Hepatology,* 38**,** 1008-17.

NEUSCHWANDER-TETRI, B. A. & CALDWELL, S. H. 2003. Nonalcoholic steatohepatitis: summary of an AASLD Single Topic Conference. *Hepatology,* 37**,** 1202-19.

NEUSCHWANDER-TETRI, B. A.,CLARK, J. M.,BASS, N. M.,VAN NATTA, M. L.,UNALP-ARIDA, A.,TONASCIA, J.,ZEIN, C. O.,BRUNT, E. M.,KLEINER, D. E.,MCCULLOUGH, A. J.,SANYAL, A. J.,DIEHL, A. M.,LAVINE, J. E.,CHALASANI, N. & KOWDLEY, K. V. 2010. Clinical, laboratory and histological associations in adults with nonalcoholic fatty liver disease. *Hepatology,* 52**,** 913-24.

NGUYEN, P.,LERAY, V.,DIEZ, M.,SERISIER, S.,LE BLOC'H, J.,SILIART, B. & DUMON, H. 2008. Liver lipid metabolism. *J Anim Physiol Anim Nutr (Berl),* 92**,** 272-83.

NICHENAMETLA, S. N.,ELLISON, I.,CALCAGNOTTO, A.,LAZARUS, P.,MUSCAT, J. E. & RICHIE, J. P., JR. 2008. Functional significance of the GAG trinucleotide-repeat polymorphism in the gene for the catalytic subunit of gamma-glutamylcysteine ligase. *Free Radic Biol Med,* 45**,** 645-50.

NICHENAMETLA, S. N.,LAZARUS, P. & RICHIE, J. P., JR. 2011. A GAG trinucleotide-repeat polymorphism in the gene for glutathione biosynthetic enzyme, GCLC, affects gene expression through translation. *FASEB J*.

NOBILI V, D. A., AL-SERRI A, LEATHART JB, DAY CP 2007. The mitochondrial superoxide dismutase 2 (SOD2) targeting sequence polymorphism is associated with fibrotic NAFLD: consistent evidence from case-control and intra-familial allelic association studies. *Hepatology,* 46.

NOBILI, V.,MARCELLINI, M.,DEVITO, R.,CIAMPALINI, P.,PIEMONTE, F.,COMPARCOLA, D.,SARTORELLI, M. R. & ANGULO, P. 2006. NAFLD in children: a prospective clinical-pathological study and effect of lifestyle advice. *Hepatology,* 44**,** 458-65.

OCHOA, M. C.,MARTI, A.,AZCONA, C.,CHUECA, M.,OYARZABAL, M.,PELACH, R.,PATINO, A.,MORENO-ALIAGA, M. J.,MARTINEZ-GONZALEZ, M. A. & MARTINEZ, J. A. 2004. Gene-gene interaction between PPAR gamma 2 and ADR beta 3 increases obesity risk in children and adolescents. *Int J Obes Relat Metab Disord,* 28 Suppl 3**,** S37-41.

OESTERGAARD, M. Z.,TYRER, J.,CEBRIAN, A.,SHAH, M.,DUNNING, A. M.,PONDER, B. A.,EASTON, D. F. & PHAROAH, P. D. 2006. Interactions between genes involved in the antioxidant defence system and breast cancer risk. *Br J Cancer,* 95**,** 525-31.

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. & 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**,** 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. & KOHNO, S. 2002. Fatty liver in non-alcoholic non-overweight Japanese adults: incidence and clinical characteristics. *J Gastroenterol Hepatol,* 17**,** 1098-105.

ONG, J. P. & YOUNOSSI, Z. M. 2003. Nonalcoholic fatty liver disease (NAFLD)--two decades later: are we smarter about its natural history? *Am J Gastroenterol,* 98**,** 1915-7.

OOI, E. M.,BARRETT, P. H.,CHAN, D. C. & WATTS, G. F. 2008. Apolipoprotein C-III: understanding an emerging cardiovascular risk factor. *Clin Sci (Lond),* 114**,** 611-24.

OZ, H. S.,CHEN, T. S. & NEUMAN, M. 2008. Methionine deficiency and hepatic injury in a dietary steatohepatitis model. *Dig Dis Sci,* 53**,** 767-76.

OZ, H. S.,IM, H. J.,CHEN, T. S.,DE VILLIERS, W. J. & MCCLAIN, C. J. 2006. Glutathione-enhancing agents protect against steatohepatitis in a dietary model. *J Biochem Mol Toxicol,* 20**,** 39-47.

PAGANO, C.,SOARDO, G.,ESPOSITO, W.,FALLO, F.,BASAN, L.,DONNINI, D.,FEDERSPIL, G.,SECHI, L. A. & VETTOR, R. 2005. Plasma adiponectin is decreased in nonalcoholic fatty liver disease. *Eur J Endocrinol,* 152**,** 113-8.

PAGANO, G.,PACINI, G.,MUSSO, G.,GAMBINO, R.,MECCA, F.,DEPETRIS, N.,CASSADER, M.,DAVID, E.,CAVALLO-PERIN, P. & RIZZETTO, M. 2002. Nonalcoholic steatohepatitis, insulin resistance, and metabolic syndrome: further evidence for an etiologic association. *Hepatology,* 35**,** 367-72.

PAIS, R.,LUPSOR, M.,POANTA, L.,SILAGHI, A.,RUSU, M. L.,BADEA, R. & DUMITRASCU, D. L. 2009. Liver biopsy versus noninvasive methods--fibroscan and fibrotest in the diagnosis of non-alcoholic fatty liver disease: a review of the literature. *Rom J Intern Med,* 47**,** 331-40.

PAJVANI, U. B.,DU, X.,COMBS, T. P.,BERG, A. H.,RAJALA, M. W.,SCHULTHESS, T.,ENGEL, J.,BROWNLEE, M. & SCHERER, P. E. 2003. Structure-function studies of the adipocyte-secreted hormone Acrp30/adiponectin. Implications fpr metabolic regulation and bioactivity. *J Biol Chem,* 278**,** 9073-85.

PALMENTIERI, B.,DE SIO, I.,LA MURA, V.,MASARONE, M.,VECCHIONE, R.,BRUNO, S.,TORELLA, R. & PERSICO, M. 2006. The role of bright liver echo pattern on ultrasound B-mode examination in the diagnosis of liver steatosis. *Dig Liver Dis,* 38**,** 485-9.

PAPP, L. V.,LU, J.,HOLMGREN, A. & KHANNA, K. K. 2007. From selenium to selenoproteins: synthesis, identity, and their role in human health. *Antioxid Redox Signal,* 9**,** 775-806.

PARDINA, E.,BAENA-FUSTEGUERAS, J. A.,CATALAN, R.,GALARD, R.,LECUBE, A.,FORT, J. M.,ALLENDE, H.,VARGAS, V. & PEINADO-ONSURBE, J. 2009. Increased expression and activity of hepatic lipase in the liver of morbidly obese adult patients in relation to lipid content. *Obes Surg,* 19**,** 894-904.

PERSEGHIN, G.,PETERSEN, K. & SHULMAN, G. I. 2003. Cellular mechanism of insulin resistance: potential links with inflammation. *Int J Obes Relat Metab Disord,* 27 Suppl 3**,** S6-11.

PESSAYRE, D. 2007. Role of mitochondria in non-alcoholic fatty liver disease. *J Gastroenterol Hepatol,* 22 Suppl 1**,** S20-7.

PESSAYRE, D.,BERSON, A.,FROMENTY, B. & MANSOURI, A. 2001. Mitochondria in steatohepatitis. *Semin Liver Dis,* 21**,** 57-69.
PESSAYRE, D. & FROMENTY, B. 2005. NASH: a mitochondrial disease. *J Hepatol,* 42**,** 928-40.
PETERSEN, K. F.,DUFOUR, S.,HARIRI, A.,NELSON-WILLIAMS, C.,FOO, J. N.,ZHANG, X. M.,DZIURA, J.,LIFTON, R. P. & SHULMAN, G. I. 2010. Apolipoprotein C3 gene variants in nonalcoholic fatty liver disease. *N Engl J Med,* 362**,** 1082-9.
POWELL, D. W. 1981. Barrier function of epithelia. *Am J Physiol,* 241**,** G275-88.
POWELL, E. E.,COOKSLEY, W. G.,HANSON, R.,SEARLE, J.,HALLIDAY, J. W. & 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**,** 74-80.
POYNARD, T.,MORRA, R.,HALFON, P.,CASTERA, L.,RATZIU, V.,IMBERT-BISMUT, F.,NAVEAU, S.,THABUT, D.,LEBREC, D.,ZOULIM, F.,BOURLIERE, M.,CACOUB, P.,MESSOUS, D.,MUNTEANU, M. & DE LEDINGHEN, V. 2007. Meta-analyses of FibroTest diagnostic value in chronic liver disease. *BMC Gastroenterol,* 7**,** 40.
POYTON, R. O.,BALL, K. A. & CASTELLO, P. R. 2009. Mitochondrial generation of free radicals and hypoxic signaling. *Trends Endocrinol Metab,* 20**,** 332-40.
PROMRAT, K.,LUTCHMAN, G.,UWAIFO, G. I.,FREEDMAN, R. J.,SOZA, A.,HELLER, T.,DOO, E.,GHANY, M.,PREMKUMAR, A.,PARK, Y.,LIANG, T. J.,YANOVSKI, J. A.,KLEINER, D. E. & HOOFNAGLE, J. H. 2004. A pilot study of pioglitazone treatment for nonalcoholic steatohepatitis. *Hepatology,* 39**,** 188-96.
QURESHI, K. & ABRAMS, G. A. 2007. Metabolic liver disease of obesity and role of adipose tissue in the pathogenesis of nonalcoholic fatty liver disease. *World J Gastroenterol,* 13**,** 3540-53.
RATZIU, V.,BONYHAY, L.,DI MARTINO, V.,CHARLOTTE, F.,CAVALLARO, L.,SAYEGH-TAINTURIER, M. H.,GIRAL, P.,GRIMALDI, A.,OPOLON, P. & POYNARD, T. 2002. Survival, liver failure, and hepatocellular carcinoma in obesity-related cryptogenic cirrhosis. *Hepatology,* 35**,** 1485-93.
RATZIU, V.,GIRAL, P.,CHARLOTTE, F.,BRUCKERT, E.,THIBAULT, V.,THEODOROU, I.,KHALIL, L.,TURPIN, G.,OPOLON, P. & POYNARD, T. 2000. Liver fibrosis in overweight patients. *Gastroenterology,* 118**,** 1117-23.
RECTOR, R. S.,THYFAULT, J. P.,WEI, Y. & IBDAH, J. A. 2008. Non-alcoholic fatty liver disease and the metabolic syndrome: an update. *World J Gastroenterol,* 14**,** 185-92.
RISCH, N. J. 2000. Searching for genetic determinants in the new millennium. *Nature,* 405**,** 847-56.
RITTER, J. K.,CHEN, F.,SHEEN, Y. Y.,TRAN, H. M.,KIMURA, S.,YEATMAN, M. T. & OWENS, I. S. 1992. A novel complex locus UGT1 encodes human bilirubin, phenol, and other UDP-glucuronosyltransferase isozymes with identical carboxyl termini. *J Biol Chem,* 267**,** 3257-61.
ROMEO, S.,KOZLITINA, J.,XING, C.,PERTSEMLIDIS, A.,COX, D.,PENNACCHIO, L. A.,BOERWINKLE, E.,COHEN, J. C. & HOBBS, H. H. 2008. Genetic variation in PNPLA3 confers susceptibility to nonalcoholic fatty liver disease. *Nat Genet,* 40**,** 1461-5.
ROMEO, S.,SENTINELLI, F.,CAMBULI, V. M.,INCANI, M.,CONGIU, T.,MATTA, V.,PILIA, S.,HUANG-DORAN, I.,COSSU, E.,LOCHE, S. & BARONI, M. G. 2010a. The 148M allele of the PNPLA3 gene is associated with indices of liver damage early in life. *J Hepatol,* 53**,** 335-8.

ROMEO, S.,SENTINELLI, F.,DASH, S.,YEO, G. S.,SAVAGE, D. B.,LEONETTI, F.,CAPOCCIA, D.,INCANI, M.,MAGLIO, C.,IACOVINO, M.,O'RAHILLY, S. & BARONI, M. G. 2010b. Morbid obesity exposes the association between PNPLA3 I148M (rs738409) and indices of hepatic injury in individuals of European descent. *Int J Obes (Lond),* 34**,** 190-4.

ROTMAN, Y.,KOH, C.,ZMUDA, J. M.,KLEINER, D. E. & 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**,** 894-903.

SACHIDANANDAM, R.,WEISSMAN, D.,SCHMIDT, S. C.,KAKOL, J. M.,STEIN, L. D.,MARTH, G.,SHERRY, S.,MULLIKIN, J. C.,MORTIMORE, B. J.,WILLEY, D. L.,HUNT, S. E.,COLE, C. G.,COGGILL, P. C.,RICE, C. M.,NING, Z.,ROGERS, J.,BENTLEY, D. R.,KWOK, P. Y.,MARDIS, E. R.,YEH, R. T.,SCHULTZ, B.,COOK, L.,DAVENPORT, R.,DANTE, M.,FULTON, L.,HILLIER, L.,WATERSTON, R. H.,MCPHERSON, J. D.,GILMAN, B.,SCHAFFNER, S.,VAN ETTEN, W. J.,REICH, D.,HIGGINS, J.,DALY, M. J.,BLUMENSTIEL, B.,BALDWIN, J.,STANGE-THOMANN, N.,ZODY, M. C.,LINTON, L.,LANDER, E. S. & ALTSHULER, D. 2001. A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature,* 409**,** 928-33.

SANADA, Y.,YOSHIDA, K. & ITOH, H. 2007. Comparison of CT enhancement patterns and histologic features in hepatocellular carcinoma up to 2 cm: assessment of malignant potential with claudin-10 immunohistochemistry. *Oncol Rep,* 17**,** 1177-82.

SANTORO, N.,KURSAWE, R.,D'ADAMO, E.,DYKAS, D. J.,ZHANG, C. K.,BALE, A. E.,CALI, A. M.,NARAYAN, D.,SHAW, M. M.,PIERPONT, B.,SAVOYE, M.,LARTAUD, D.,ELDRICH, S.,CUSHMAN, S. W.,ZHAO, H.,SHULMAN, G. I. & CAPRIO, S. 2010. A common variant in the patatin-like phospholipase 3 gene (PNPLA3) is associated with fatty liver disease in obese children and adolescents. *Hepatology,* 52**,** 1281-90.

SANYAL, A. J.,CAMPBELL-SARGENT, C.,MIRSHAHI, F.,RIZZO, W. B.,CONTOS, M. J.,STERLING, R. K.,LUKETIC, V. A.,SHIFFMAN, M. L. & CLORE, J. N. 2001. Nonalcoholic steatohepatitis: association of insulin resistance and mitochondrial abnormalities. *Gastroenterology,* 120**,** 1183-92.

SASAKI, Y. 2006. Does oxidative stress participate in the development of hepatocellular carcinoma? *J Gastroenterol,* 41**,** 1135-48.

SATOH, N.,WATANABE, N.,KANDA, A.,SUGAYA-FUKASAWA, M. & HISATOMI, H. 2010. Expression of glutathione reductase splice variants in human tissues. *Biochem Genet,* 48**,** 816-21.

SAXENA, R.,VOIGHT, B. F.,LYSSENKO, V.,BURTT, N. P.,DE BAKKER, P. I.,CHEN, H.,ROIX, J. J.,KATHIRESAN, S.,HIRSCHHORN, J. N.,DALY, M. J.,HUGHES, T. E.,GROOP, L.,ALTSHULER, D.,ALMGREN, P.,FLOREZ, J. C.,MEYER, J.,ARDLIE, K.,BENGTSSON BOSTROM, K.,ISOMAA, B.,LETTRE, G.,LINDBLAD, U.,LYON, H. N.,MELANDER, O.,NEWTON-CHEH, C.,NILSSON, P.,ORHO-MELANDER, M.,RASTAM, L.,SPELIOTES, E. K.,TASKINEN, M. R.,TUOMI, T.,GUIDUCCI, C.,BERGLUND, A.,CARLSON, J.,GIANNINY, L.,HACKETT, R.,HALL, L.,HOLMKVIST, J.,LAURILA, E.,SJOGREN, M.,STERNER, M.,SURTI, A.,SVENSSON, M.,TEWHEY, R.,BLUMENSTIEL, B.,PARKIN, M.,DEFELICE, M.,BARRY, R.,BRODEUR, W.,CAMARATA, J.,CHIA, N.,FAVA, M.,GIBBONS, J.,HANDSAKER, B.,HEALY, C.,NGUYEN, K.,GATES, C.,SOUGNEZ, C.,GAGE, D.,NIZZARI, M.,GABRIEL, S. B.,CHIRN, G. W.,MA, Q.,PARIKH,

H.,RICHARDSON, D.,RICKE, D. & PURCELL, S. 2007. Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science,* 316**,** 1331-6.

SCHECHTER, I.,CONRAD, D. G.,HART, I.,BERGER, R. C.,MCKENZIE, T. L.,BLESKAN, J. & PATTERSON, D. 1994. Localization of the squalene synthase gene (FDFT1) to human chromosome 8p22-p23.1. *Genomics,* 20**,** 116-8.

SCHERAG, A.,DINA, C.,HINNEY, A.,VATIN, V.,SCHERAG, S.,VOGEL, C. I.,MULLER, T. D.,GRALLERT, H.,WICHMANN, H. E.,BALKAU, B.,HEUDE, B.,JARVELIN, M. R.,HARTIKAINEN, A. L.,LEVY-MARCHAL, C.,WEILL, J.,DELPLANQUE, J.,KORNER, A.,KIESS, W.,KOVACS, P.,RAYNER, N. W.,PROKOPENKO, I.,MCCARTHY, M. I.,SCHAFER, H.,JARICK, I.,BOEING, H.,FISHER, E.,REINEHR, T.,HEINRICH, J.,RZEHAK, P.,BERDEL, D.,BORTE, M.,BIEBERMANN, H.,KRUDE, H.,ROSSKOPF, D.,RIMMBACH, C.,RIEF, W.,FROMME, T.,KLINGENSPOR, M.,SCHURMANN, A.,SCHULZ, N.,NOTHEN, M. M.,MUHLEISEN, T. W.,ERBEL, R.,JOCKEL, K. H.,MOEBUS, S.,BOES, T.,ILLIG, T.,FROGUEL, P.,HEBEBRAND, J. & MEYRE, D. 2010. Two new Loci for body-weight regulation identified in a joint analysis of genome-wide association studies for early-onset extreme obesity in French and german study groups. *PLoS Genet,* 6**,** e1000916.

SCHREUDER, T. C.,VERWER, B. J.,VAN NIEUWKERK, C. M. & MULDER, C. J. 2008. Nonalcoholic fatty liver disease: an overview of current insights in pathogenesis, diagnosis and treatment. *World J Gastroenterol,* 14**,** 2474-86.

SCHULTZ, J. R.,TU, H.,LUK, A.,REPA, J. J.,MEDINA, J. C.,LI, L.,SCHWENDNER, S.,WANG, S.,THOOLEN, M.,MANGELSDORF, D. J.,LUSTIG, K. D. & SHAN, B. 2000. Role of LXRs in control of lipogenesis. *Genes Dev,* 14**,** 2831-8.

SCHWENZER, N. F.,SPRINGER, F.,SCHRAML, C.,STEFAN, N.,MACHANN, J. & SCHICK, F. 2009. Non-invasive assessment and quantification of liver steatosis by ultrasound, computed tomography and magnetic resonance. *J Hepatol,* 51**,** 433-45.

SCHWERTNER, H. A. & VITEK, L. 2008. Gilbert syndrome, UGT1A1*28 allele, and cardiovascular disease risk: possible protective effects and therapeutic applications of bilirubin. *Atherosclerosis,* 198**,** 1-11.

SCHWIMMER, J. B.,BEHLING, C.,NEWBURY, R.,DEUTSCH, R.,NIEVERGELT, C.,SCHORK, N. J. & LAVINE, J. E. 2005. Histopathology of pediatric nonalcoholic fatty liver disease. *Hepatology,* 42**,** 641-9.

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. & SIRLIN, C. B. 2009. Heritability of nonalcoholic fatty liver disease. *Gastroenterology,* 136**,** 1585-92.

SCOTT, L. J.,MOHLKE, K. L.,BONNYCASTLE, L. L.,WILLER, C. J.,LI, Y.,DUREN, W. L.,ERDOS, M. R.,STRINGHAM, H. M.,CHINES, P. S.,JACKSON, A. U.,PROKUNINA-OLSSON, L.,DING, C. J.,SWIFT, A. J.,NARISU, N.,HU, T.,PRUIM, R.,XIAO, R.,LI, X. Y.,CONNEELY, K. N.,RIEBOW, N. L.,SPRAU, A. G.,TONG, M.,WHITE, P. P.,HETRICK, K. N.,BARNHART, M. W.,BARK, C. W.,GOLDSTEIN, J. L.,WATKINS, L.,XIANG, F.,SARAMIES, J.,BUCHANAN, T. A.,WATANABE, R. M.,VALLE, T. T.,KINNUNEN, L.,ABECASIS, G. R.,PUGH, E. W.,DOHENY, K. F.,BERGMAN, R. N.,TUOMILEHTO, J.,COLLINS, F. S. & BOEHNKE, M. 2007. A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science,* 316**,** 1341-5.

SENG, K. C. & SENG, C. K. 2008. The success of the genome-wide association approach: a brief story of a long struggle. *Eur J Hum Genet,* 16**,** 554-64.
SHERLOCK, S. & DOOLEY, J. 2002. *Diseases of the liver and biliary system,* Malden, MA, Blackwell Science.
SHEU, C. C.,ZHAI, R.,WANG, Z.,GONG, M. N.,TEJERA, P.,CHEN, F.,SU, L.,THOMPSON, B. T. & CHRISTIANI, D. C. 2009. Heme oxygenase-1 microsatellite polymorphism and haplotypes are associated with the development of acute respiratory distress syndrome. *Intensive Care Med,* 35**,** 1343-51.
SHIMODA-MATSUBAYASHI, S.,MATSUMINE, H.,KOBAYASHI, T.,NAKAGAWA-HATTORI, Y.,SHIMIZU, Y. & MIZUNO, Y. 1996. Structural dimorphism in the mitochondrial targeting sequence in the human manganese superoxide dismutase gene. A predictive evidence for conformational change to influence mitochondrial transport and a study of allelic association in Parkinson's disease. *Biochem Biophys Res Commun,* 226**,** 561-5.
SHIMOMURA, I.,BASHMAKOV, Y. & HORTON, J. D. 1999. Increased levels of nuclear SREBP-1c associated with fatty livers in two mouse models of diabetes mellitus. *J Biol Chem,* 274**,** 30028-32.
SHOELSON, S. E.,LEE, J. & GOLDFINE, A. B. 2006. Inflammation and insulin resistance. *J Clin Invest,* 116**,** 1793-801.
SIEGEL, A. B. & ZHU, A. X. 2009. Metabolic syndrome and hepatocellular carcinoma: two growing epidemics with a potential link. *Cancer,* 115**,** 5651-61.
SIM, X.,ONG, R. T.,SUO, C.,TAY, W. T.,LIU, J.,NG, D. P.,BOEHNKE, M.,CHIA, K. S.,WONG, T. Y.,SEIELSTAD, M.,TEO, Y. Y. & TAI, E. S. 2011. Transferability of type 2 diabetes implicated Loci in multi-ethnic cohorts from southeast Asia. *PLoS Genet,* 7**,** e1001363.
SINGH, A.,RANGASAMY, T.,THIMMULAPPA, R. K.,LEE, H.,OSBURN, W. O.,BRIGELIUS-FLOHE, R.,KENSLER, T. W.,YAMAMOTO, M. & BISWAL, S. 2006. Glutathione peroxidase 2, the major cigarette smoke-inducible isoform of GPX in lungs, is regulated by Nrf2. *Am J Respir Cell Mol Biol,* 35**,** 639-50.
SOLGA, S. F. & DIEHL, A. M. 2003. Non-alcoholic fatty liver disease: lumen-liver interactions and possible role for probiotics. *J Hepatol,* 38**,** 681-7.
SONG, J.,DA COSTA, K. A.,FISCHER, L. M.,KOHLMEIER, M.,KWOCK, L.,WANG, S. & ZEISEL, S. H. 2005. Polymorphism of the PEMT gene and susceptibility to nonalcoholic fatty liver disease (NAFLD). *FASEB J,* 19**,** 1266-71.
SOOKOIAN, S.,CASTANO, G. O.,BURGUENO, A. L.,GIANOTTI, T. F.,ROSSELLI, M. S. & PIROLA, C. J. 2009. A nonsynonymous gene variant in the adiponutrin gene is associated with nonalcoholic fatty liver disease severity. *J Lipid Res,* 50**,** 2111-6.
SOOKOIAN, S.,CASTANO, G. O.,BURGUENO, A. L.,GIANOTTI, T. F.,ROSSELLI, M. S. & PIROLA, C. J. 2010. The nuclear receptor PXR gene variants are associated with liver injury in nonalcoholic fatty liver disease. *Pharmacogenet Genomics,* 20**,** 1-8.
SOOKOIAN, S. & 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**,** 1883-94.
SPELIOTES, E. K.,BUTLER, J. L.,PALMER, C. D.,VOIGHT, B. F. & HIRSCHHORN, J. N. 2010. PNPLA3 variants specifically confer increased risk for histologic nonalcoholic fatty liver disease but not metabolic disease. *Hepatology,* 52**,** 904-12.

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.,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. 2011. Genome-wide association analysis identifies variants associated with nonalcoholic fatty liver disease that have distinct effects on metabolic traits. *PLoS Genet,* 7**,** e1001324.

SPIELMAN, R. S. & EWENS, W. J. 1996. The TDT and other family-based tests for linkage disequilibrium and association. *Am J Hum Genet,* 59**,** 983-9.

STAEHELIN, L. A. 1974. Structure and function of intercellular junctions. *Int Rev Cytol,* 39**,** 191-283.

STAMATAKIS, E.,PRIMATESTA, P.,CHINN, S.,RONA, R. & FALASCHETI, E. 2005. Overweight and obesity trends from 1974 to 2003 in English children: what is the role of socioeconomic factors? *Arch Dis Child,* 90**,** 999-1004.

STARLEY, B. Q.,CALCAGNO, C. J. & HARRISON, S. A. 2010. Nonalcoholic fatty liver disease and hepatocellular carcinoma: a weighty connection. *Hepatology,* 51**,** 1820-32.

STEED, E.,RODRIGUES, N. T.,BALDA, M. S. & MATTER, K. 2009. Identification of MarvelD3 as a tight junction-associated transmembrane protein of the occludin family. *BMC Cell Biol,* 10**,** 95.

STEFANO, J. T.,DE OLIVEIRA, C. P.,CORREA-GIANNELLA, M. L.,SOARES, I. C.,KUBRUSLY, M. S.,BELLODI-PRIVATO, M.,DE MELLO, E. S.,DE LIMA, V. M.,CARRILHO, F. J. & ALVES, V. A. 2011. Decreased immunoexpression of survivin could be a potential marker in human non-alcoholic fatty liver disease progression? *Liver Int,* 31**,** 377-85.

STEPHENS, M.,SMITH, N. J. & DONNELLY, P. 2001. A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet,* 68**,** 978-89.

STEWART, S. F.,LEATHART, J. B.,CHEN, Y.,DALY, A. K.,ROLLA, R.,VAY, D.,MOTTARAN, E.,VIDALI, M.,ALBANO, E. & DAY, C. P. 2002. Valine-alanine manganese superoxide dismutase polymorphism is not associated with alcohol-induced oxidative stress or liver fibrosis. *Hepatology,* 36**,** 1355-60.

STOCKER, R.,YAMAMOTO, Y.,MCDONAGH, A. F.,GLAZER, A. N. & AMES, B. N. 1987. Bilirubin is an antioxidant of possible physiological importance. *Science,* 235**,** 1043-6.

STRUBEN, V. M.,HESPENHEIDE, E. E. & CALDWELL, S. H. 2000. Nonalcoholic steatohepatitis and cryptogenic cirrhosis within kindreds. *Am J Med,* 108**,** 9-13.

SUTTON, A.,IMBERT, A.,IGOUDJIL, A.,DESCATOIRE, V.,CAZANAVE, S.,PESSAYRE, D. & DEGOUL, F. 2005. The manganese superoxide dismutase Ala16Val dimorphism modulates both mitochondrial import and mRNA stability. *Pharmacogenet Genomics,* 15**,** 311-9.

SUTTON, A.,KHOURY, H.,PRIP-BUUS, C.,CEPANEC, C.,PESSAYRE, D. & DEGOUL, F. 2003. The Ala16Val genetic dimorphism modulates the import of human manganese superoxide dismutase into rat liver mitochondria. *Pharmacogenetics,* 13**,** 145-57.

SUTTON, A.,NAHON, P.,PESSAYRE, D.,RUFAT, P.,POIRE, A.,ZIOL, M.,VIDAUD, D.,BARGET, N.,GANNE-CARRIE, N.,CHARNAUX, N.,TRINCHET, J. C.,GATTEGNO, L. & BEAUGRAND, M. 2006. Genetic polymorphisms in antioxidant enzymes modulate hepatic iron accumulation and hepatocellular

carcinoma development in patients with alcohol-induced cirrhosis. *Cancer Res,* 66**,** 2844-52.
TARGHER, G.,BERTOLINI, L.,PADOVANI, R.,RODELLA, S.,TESSARI, R.,ZENARI, L.,DAY, C. & ARCARO, G. 2007. Prevalence of nonalcoholic fatty liver disease and its association with cardiovascular disease among type 2 diabetic patients. *Diabetes Care,* 30**,** 1212-8.
TARGHER, G.,BERTOLINI, L.,SCALA, L.,POLI, F.,ZENARI, L. & FALEZZA, G. 2004. Decreased plasma adiponectin concentrations are closely associated with nonalcoholic hepatic steatosis in obese individuals. *Clin Endocrinol (Oxf),* 61**,** 700-3.
TEAM, R. C. 2008. R: A Language and Environment for Statistical Computing *eds. Vienna Austria R Foundation for Statistical Computing*.
THOMAS, E. L.,HAMILTON, G.,PATEL, N.,O'DWYER, R.,DORE, C. J.,GOLDIN, R. D.,BELL, J. D. & 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**,** 122-7.
THORISSON, G. A. & STEIN, L. D. 2003. The SNP Consortium website: past, present and future. *Nucleic Acids Res,* 31**,** 124-7.
TOGASHI, N.,URA, N.,HIGASHIURA, K.,MURAKAMI, H. & SHIMAMOTO, K. 2000. The contribution of skeletal muscle tumor necrosis factor-alpha to insulin resistance and hypertension in fructose-fed rats. *J Hypertens,* 18**,** 1605-10.
TOKUSHIGE, K.,TAKAKURA, M.,TSUCHIYA-MATSUSHITA, N.,TANIAI, M.,HASHIMOTO, E. & SHIRATORI, K. 2007. Influence of TNF gene polymorphisms in Japanese patients with NASH and simple steatosis. *J Hepatol,* 46**,** 1104-10.
TOMKINS, J.,BANNER, S. J.,MCDERMOTT, C. J. & SHAW, P. J. 2001. Mutation screening of manganese superoxide dismutase in amyotrophic lateral sclerosis. *Neuroreport,* 12**,** 2319-22.
TREPO, E.,POTTHOFF, A.,PRADAT, P.,BAKSHI, R.,YOUNG, B.,LAGIER, R.,MORENO, C.,VERSET, L.,CROSS, R.,DEGRE, D.,LEMMERS, A.,GUSTOT, T.,BERTHILLON, P.,ROSENBERG, W.,TREPO, C.,SNINSKY, J.,ADLER, M. & WEDEMEYER, H. 2011. Role of a cirrhosis risk score for the early prediction of fibrosis progression in hepatitis C patients with minimal liver disease. *J Hepatol,* 55**,** 38-44.
TSOCHATZIS, E.,PAPATHEODORIDIS, G. V. & ARCHIMANDRITIS, A. J. 2006. The evolving role of leptin and adiponectin in chronic liver diseases. *Am J Gastroenterol,* 101**,** 2629-40.
UDLER, M.,MAIA, A. T.,CEBRIAN, A.,BROWN, C.,GREENBERG, D.,SHAH, M.,CALDAS, C.,DUNNING, A.,EASTON, D.,PONDER, B. & PHAROAH, P. 2007. Common germline genetic variation in antioxidant defense genes and survival after diagnosis of breast cancer. *J Clin Oncol,* 25**,** 3015-23.
UENO, T.,SUGAWARA, H.,SUJAKU, K.,HASHIMOTO, O.,TSUJI, R.,TAMAKI, S.,TORIMURA, T.,INUZUKA, S.,SATA, M. & TANIKAWA, K. 1997. Therapeutic effects of restricted diet and exercise in obese patients with fatty liver. *J Hepatol,* 27**,** 103-7.
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. & 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**,** 1209-17.

VALENTI, L.,ALISI, A.,GALMOZZI, E.,BARTULI, A.,DEL MENICO, B.,ALTERIO, A.,DONGIOVANNI, P.,FARGION, S. & NOBILI, V. 2010b. I148M patatin-like phospholipase domain-containing 3 gene variant and severity of pediatric nonalcoholic fatty liver disease. *Hepatology*, 52**,** 1274-80.

VALENTI, L.,FRACANZANI, A. L.,BUGIANESI, E.,DONGIOVANNI, P.,GALMOZZI, E.,VANNI, E.,CANAVESI, E.,LATTUADA, E.,ROVIARO, G.,MARCHESINI, G. & FARGION, S. 2010c. HFE genotype, parenchymal iron accumulation, and liver fibrosis in patients with nonalcoholic fatty liver disease. *Gastroenterology,* 138**,** 905-12.

VALENTI, L.,FRACANZANI, A. L.,DONGIOVANNI, P.,SANTORELLI, G.,BRANCHI, A.,TAIOLI, E.,FIORELLI, G. & FARGION, S. 2002. Tumor necrosis factor alpha promoter polymorphisms and insulin resistance in nonalcoholic fatty liver disease. *Gastroenterology,* 122**,** 274-80.

VALKO, M.,IZAKOVIC, M.,MAZUR, M.,RHODES, C. J. & TELSER, J. 2004. Role of oxygen radicals in DNA damage and cancer incidence. *Mol Cell Biochem,* 266**,** 37-56.

VAN ACKER, F. A.,SCHOUTEN, O.,HAENEN, G. R.,VAN DER VIJGH, W. J. & BAST, A. 2000. Flavonoids can replace alpha-tocopherol as an antioxidant. *FEBS Lett,* 473**,** 145-8.

VAN ITALLIE, C. M.,ROGAN, S.,YU, A.,VIDAL, L. S.,HOLMES, J. & ANDERSON, J. M. 2006. Two splice variants of claudin-10 in the kidney create paracellular pores with different ion selectivities. *Am J Physiol Renal Physiol,* 291**,** F1288-99.

VAN STEEN, K. 2011. Travelling the world of gene-gene interactions. *Brief Bioinform*.

VANNI, E.,BUGIANESI, E.,KOTRONEN, A.,DE MINICIS, S.,YKI-JARVINEN, H. & SVEGLIATI-BARONI, G. 2010. From the metabolic syndrome to NAFLD or vice versa? *Dig Liver Dis,* 42**,** 320-30.

VASAVDA, N.,MENZEL, S.,KONDAVEETI, S.,MAYTHAM, E.,AWOGBADE, M.,BANNISTER, S.,CUNNINGHAM, J.,EICHHOLZ, A.,DANIEL, Y.,OKPALA, I.,FULFORD, T. & THEIN, S. L. 2007. The linear effects of alpha-thalassaemia, the UGT1A1 and HMOX1 polymorphisms on cholelithiasis in sickle cell disease. *Br J Haematol,* 138**,** 263-70.

VON OTTER, M.,LANDGREN, S.,NILSSON, S.,CELOJEVIC, D.,BERGSTROM, P.,HAKANSSON, A.,NISSBRANDT, H.,DROZDZIK, M.,BIALECKA, M.,KURZAWSKI, M.,BLENNOW, K.,NILSSON, M.,HAMMARSTEN, O. & ZETTERBERG, H. 2010. Association of Nrf2-encoding NFE2L2 haplotypes with Parkinson's disease. *BMC Med Genet,* 11**,** 36.

WALL, J. D. & PRITCHARD, J. K. 2003. Haplotype blocks and linkage disequilibrium in the human genome. *Nat Rev Genet,* 4**,** 587-97.

WALSH, A. C.,FEULNER, J. A. & REILLY, A. 2001. Evidence for functionally significant polymorphism of human glutamate cysteine ligase catalytic subunit: association with glutathione levels and drug resistance in the National Cancer Institute tumor cell line panel. *Toxicol Sci,* 61**,** 218-23.

WANG, K.,LI, W. D.,ZHANG, C. K.,WANG, Z.,GLESSNER, J. T.,GRANT, S. F.,ZHAO, H.,HAKONARSON, H. & PRICE, R. A. 2011. A genome-wide association study on obesity and obesity-related traits. *PLoS One,* 6**,** e18939.

WANLESS, I. R. & LENTZ, J. S. 1990. Fatty liver hepatitis (steatohepatitis) and obesity: an autopsy study with analysis of risk factors. *Hepatology,* 12**,** 1106-10.

WEI, Y.,CLARK, S. E.,MORRIS, E. M.,THYFAULT, J. P.,UPTERGROVE, G. M.,WHALEY-CONNELL, A. T.,FERRARIO, C. M.,SOWERS, J. R. & IBDAH, J. A. 2008. Angiotensin II-induced non-alcoholic fatty liver disease is mediated by oxidative stress in transgenic TG(mRen2)27(Ren2) rats. *J Hepatol,* 49**,** 417-28.

WEISIGER, R. A. & FRIDOVICH, I. 1973. Mitochondrial superoxide simutase. Site of synthesis and intramitochondrial localization. *J Biol Chem,* 248**,** 4793-6.
WESTERBACKA, J.,KOLAK, M.,KIVILUOTO, T.,ARKKILA, P.,SIREN, J.,HAMSTEN, A.,FISHER, R. M. & YKI-JARVINEN, H. 2007. Genes involved in fatty acid partitioning and binding, lipolysis, monocyte/macrophage recruitment, and inflammation are overexpressed in the human fatty liver of insulin-resistant subjects. *Diabetes,* 56**,** 2759-65.
WHEATLEY-PRICE, P.,ASOMANING, K.,REID, A.,ZHAI, R.,SU, L.,ZHOU, W.,ZHU, A.,RYAN, D. P.,CHRISTIANI, D. C. & LIU, G. 2008. Myeloperoxidase and superoxide dismutase polymorphisms are associated with an increased risk of developing pancreatic adenocarcinoma. *Cancer,* 112**,** 1037-42.
WHEELER, E. & BARROSO, I. 2011. Genome-wide association studies and type 2 diabetes. *Brief Funct Genomics,* 10**,** 52-60.
WILLNER, I. R.,WATERS, B.,PATIL, S. R.,REUBEN, A.,MORELLI, J. & RIELY, C. A. 2001. Ninety patients with nonalcoholic steatohepatitis: insulin resistance, familial tendency, and severity of disease. *Am J Gastroenterol,* 96**,** 2957-61.
WORTHAM, M.,HE, L.,GYAMFI, M.,COPPLE, B. L. & 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*.
XU, Z.,CHEN, L.,LEUNG, L.,YEN, T. S.,LEE, C. & CHAN, J. Y. 2005. Liver-specific inactivation of the Nrf1 gene in adult mouse leads to nonalcoholic steatohepatitis and hepatic neoplasia. *Proc Natl Acad Sci U S A,* 102**,** 4120-5.
YAMAMOTO, T.,YOH, K.,KOBAYASHI, A.,ISHII, Y.,KURE, S.,KOYAMA, A.,SAKAMOTO, T.,SEKIZAWA, K.,MOTOHASHI, H. & YAMAMOTO, M. 2004. Identification of polymorphisms in the promoter region of the human NRF2 gene. *Biochem Biophys Res Commun,* 321**,** 72-9.
YANG, Y.,DIETER, M. Z.,CHEN, Y.,SHERTZER, H. G.,NEBERT, D. W. & DALTON, T. P. 2002. Initial characterization of the glutamate-cysteine ligase modifier subunit Gclm(-/-) knockout mouse. Novel model system for a severely compromised oxidative stress response. *J Biol Chem,* 277**,** 49446-52.
YAO, Z. M. & VANCE, D. E. 1988. The active synthesis of phosphatidylcholine is required for very low density lipoprotein secretion from rat hepatocytes. *J Biol Chem,* 263**,** 2998-3004.
YONEDA, M. 2010. *Insulin resistance and nonalcoholic fatty liver disease (NAFLD)/nonalcoholic steatohepatitis (NASH),* New York, Nova Biomedial Books.
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. & NAKAJIMA, A. 2009. Association between angiotensin II type 1 receptor polymorphisms and the occurrence of nonalcoholic fatty liver disease. *Liver Int,* 29**,** 1078-85.
YOUNOSSI, Z. M.,STEPANOVA, M.,AFENDY, M.,FANG, Y.,YOUNOSSI, Y.,MIR, H. & SRISHORD, M. 2011. Changes in the Prevalence of the Most Common Causes of Chronic Liver Diseases in the United States from 1988-2008. *Clin Gastroenterol Hepatol*.
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. & MOOSER, V. 2008.

Population-based genome-wide association studies reveal six loci influencing plasma levels of liver enzymes. *Am J Hum Genet,* 83**,** 520-8.

YUEH, M. F. & TUKEY, R. H. 2007. Nrf2-Keap1 signaling pathway regulates human UGT1A1 expression in vitro and in transgenic UGT1 mice. *J Biol Chem,* 282**,** 8749-58.

ZEGGINI, E.,WEEDON, M. N.,LINDGREN, C. M.,FRAYLING, T. M.,ELLIOTT, K. S.,LANGO, H.,TIMPSON, N. J.,PERRY, J. R.,RAYNER, N. W.,FREATHY, R. M.,BARRETT, J. C.,SHIELDS, B.,MORRIS, A. P.,ELLARD, S.,GROVES, C. J.,HARRIES, L. W.,MARCHINI, J. L.,OWEN, K. R.,KNIGHT, B.,CARDON, L. R.,WALKER, M.,HITMAN, G. A.,MORRIS, A. D.,DONEY, A. S.,MCCARTHY, M. I. & HATTERSLEY, A. T. 2007. Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science,* 316**,** 1336-41.

ZHANG, G.,NEBERT, D. W.,CHAKRABORTY, R. & JIN, L. 2006. Statistical power of association using the extreme discordant phenotype design. *Pharmacogenet Genomics,* 16**,** 401-13.

ZHANG, H.,COURT, N. & FORMAN, H. J. 2007. Submicromolar concentrations of 4-hydroxynonenal induce glutamate cysteine ligase expression in HBE1 cells. *Redox Rep,* 12**,** 101-6.

ZHENG, S.,YUMEI, F. & CHEN, A. 2007. De novo synthesis of glutathione is a prerequisite for curcumin to inhibit hepatic stellate cell (HSC) activation. *Free Radic Biol Med,* 43**,** 444-53.

ZHOU, J.,ZHAI, Y.,MU, Y.,GONG, H.,UPPAL, H.,TOMA, D.,REN, S.,EVANS, R. M. & XIE, W. 2006. A novel pregnane X receptor-mediated and sterol regulatory element-binding protein-independent lipogenic pathway. *J Biol Chem,* 281**,** 15013-20.

ZIMMERMANN, R.,LASS, A.,HAEMMERLE, G. & ZECHNER, R. 2009. Fate of fat: the role of adipose triglyceride lipase in lipolysis. *Biochim Biophys Acta,* 1791**,** 494-500.

ZIMPRICH, A.,BENET-PAGES, A.,STRUHAL, W.,GRAF, E.,ECK, S. H.,OFFMAN, M. N.,HAUBENBERGER, D.,SPIELBERGER, S.,SCHULTE, E. C.,LICHTNER, P.,ROSSLE, S. C.,KLOPP, N.,WOLF, E.,SEPPI, K.,PIRKER, W.,PRESSLAUER, S.,MOLLENHAUER, B.,KATZENSCHLAGER, R.,FOKI, T.,HOTZY, C.,REINTHALER, E.,HARUTYUNYAN, A.,KRALOVICS, R.,PETERS, A.,ZIMPRICH, F.,BRUCKE, T.,POEWE, W.,AUFF, E.,TRENKWALDER, C.,ROST, B.,RANSMAYR, G.,WINKELMANN, J.,MEITINGER, T. & STROM, T. M. 2011. A Mutation in VPS35, Encoding a Subunit of the Retromer Complex, Causes Late-Onset Parkinson Disease. *Am J Hum Genet,* 89**,** 168-75.

Appendix

Homozygosity for the Patatin-Like Phospholipase-3/ Adiponutrin I148M Polymorphism Influences Liver Fibrosis in Patients with Nonalcoholic Fatty Liver Disease

Luca Valenti,[1] Ahmad Al-Serri,[3] Ann K. Daly,[3] Enrico Galmozzi,[1] Raffaela Rametta,[1] Paola Dongiovanni,[1] Valerio Nobili,[4] Enrico Mozzi,[2] Giancarlo Roviaro,[2] Ester Vanni,[5] Elisabetta Bugianesi,[5] Marco Maggioni,[6] Anna Ludovica Fracanzani,[1] Silvia Fargion,[1] and Christopher P. Day[3]

Inherited factors play a major role in the predisposition to nonalcoholic fatty liver disease (NAFLD), and the rs738409 C→G polymorphism of *PNPLA3/adiponutrin*, encoding for the isoleucine-to-methionine substitution at residue 148 (I148M) protein variant, has recently been recognized as a major determinant of liver fat content. However, the effect of the rs738409 polymorphism on the severity of liver fibrosis in patients with NAFLD is still unknown. In this study, we considered 253 Italian patients, 179 healthy controls, and 71 family trios with an affected child with NAFLD. Analyses were replicated in 321 patients from the United Kingdom. The rs738409 polymorphism was determined by TaqMan assays. Liver histology was scored according to Kleiner et al. Hepatic expression of genes regulating liver damage was assessed by real-time polymerase chain reaction in 52 patients. The rs738409 GG genotype was more prevalent in patients than in controls (14% versus 3%, adjusted odds ratio [OR] = 3.29, 95% confidence interval [CI] = 1.8-6.9), and in the family study, the G allele was overtransmitted to affected children ($P = 0.001$). In Italian and United Kingdom patients, *adiponutrin* genotype influenced alanine aminotransferase levels and the severity of steatosis. *Adiponutrin* genotype was associated with the expression of genes involved in the steatosis-related liver damage, including the proapoptotic molecule Fas ligand. In the whole series combined, *adiponutrin* genotype was associated with steatosis grade >1 (OR = 1.35, 95% CI = 1.04-1.76), nonalcoholic steatohepatitis (OR = 1.5, 95% CI = 1.12-2.04), and fibrosis stage >1 (OR = 1.5, 95% CI = 1.09-2.12), independent of age, body mass index, and diabetes. Adiponutrin genotype demonstrated a dose effect with heterozygote risk intermediate between CC and GG homozygotes. *Conclusion:* In patients with NAFLD, *adiponutrin* rs738409 C→G genotype, encoding for I148M, is associated with the severity of steatosis and fibrosis and the presence of nonalcoholic steatohepatitis. (HEPATOLOGY 2010;51:1209-1217)

Abbreviations: ALT, alanine aminotransferase; FASL, Fas ligand; FFA, free fatty acid; HDL, high-density lipoprotein; INSR, insulin receptor; LDL, low-density lipoprotein; mRNA, messenger RNA; NAFLD, nonalcoholic fatty liver disease; NASH, nonalcoholic steatohepatitis; PNPLA3, patatin-like phospholipase-3; PPAR-α, peroxisome proliferator-activated receptor-α; qRT-PCR, quantitative real-time polymerase chain reaction; α-SMA, α-smooth muscle actin; SNP, single-nucleotide polymorphism; SREBP1c, steroid regulatory element binding protein 1c.

From the [1]Department of Internal Medicine; and [2]Department of Surgery, Università degli Studi Milano, UO Medicina Interna 1B, Fondazione Ospedale Policlinico MaRE Istituto di Ricovero e Cura a Carattere Scientifico (IRCCS), Milan, Italy; [3]Institute of Cellular Medicine, Newcastle University, Newcastle upon Tyne, UK; [4]Liver Unit, Bambin Gesù Hospital, Rome, Italy; [5]Division of Gastroenterology, Università di Torino, Turin, Italy; and [6]UOC Anatomia Patologica, Ospedale San Paolo, Fondazione Ospedale Policlinico MaRE IRCCS, Milan, Italy.

Received January 4, 2010; accepted February 16, 2010.

The work was supported by the following grants: FIRST Università di Milano 2007, 2008 (to L.V., S.F.); Ricerca corrente Ospedale Maggiore Policlinico 2006 and 2008 (to L.V., S.F.); and Centro per lo Studio delle Malattie del Fegato e del Metabolismo..

Address reprint requests to: Silvia Fargion, Department of Internal Medicine, Università degli Studi Milano, UO Medicina Interna 1B, Fondazione Ospedale Policlinico MaRE IRCCS, Milano, Italy. E-mail: silvia.fargion@unimi.it; fax: +39 02 50320296.

Nonalcoholic fatty liver disease (NAFLD), affecting 20%-34% of the U.S. and European populations,[1,2] is currently considered a hepatic manifestation of the metabolic syndrome[3] and is associated with increased mortality due to cardiovascular and liver diseases.[4,5]

Hepatocellular fat accumulation and insulin resistance are the key factors in NAFLD pathophysiology and are deeply entangled with the progression of liver disease. Insulin resistance in the adipose tissue results in an excess of free fatty acid (FFA) supply to the liver, which may induce lipotoxicity, oxidative stress, and apoptosis, thus leading to nonalcoholic steatohepatitis (NASH).[6] NASH is considered the progressive form of NAFLD and is characterized by necroinflammatory changes, ballooning degeneration, and/or fibrosis. Insulin resistance per se may be involved in the onset of liver damage, and the presence of the metabolic syndrome predicts the severity of liver disease.[4,7]

Inherited factors play a major role in susceptibility to the metabolic syndrome and NASH. Predisposition to NASH appears to cluster with the heritability of metabolic risk factors within families, and ethnic differences have been reported.[8] Single-nucleotide polymorphisms (SNPs) in genes involved in inflammation, oxidative stress, and fibrogenesis have been associated with the severity of liver damage in NAFLD,[9-11] but data are contradicting and explain only a minority of fibrosis variations.

Recently, genomewide association studies identified an SNP in *adiponutrin/patatin-like phospholipase-3* (*PNPLA3*), rs738409 C→G, encoding the I148M (isoleucine-to-methionine substitution at residue 148) protein variant, as the strongest genetic determinant of liver fat and alanine aminotransferase (ALT) levels.[12-14] The 148M variant was more common in Hispanics, who are known to be at increased risk for NASH and cryptogenic cirrhosis.[12] Importantly, the I148M SNP influenced liver fat independently of body mass, dyslipidemia, and insulin resistance. In addition, *PNPLA3* variants have also been associated with alcoholic liver disease and alcohol-related cirrhosis.[15]

Adiponutrin/PNPLA3, which belongs to the patatin-like phospholipase family, is expressed in the liver and adipose tissue and possesses acyl hydrolase activity.[16] Adiponutrin expression is increased by carbohydrate feeding and a Western-type diet,[17,18] and has lipase activity against triglycerides and acylglycerol transacylase activity, thus being likely involved in energy mobilization and storage in lipid droplets.[19] Recently, it was reported that the 148M *adiponutrin* allele is a loss-of-function variant that predisposes an individual to steatosis by decreasing triglyceride hydrolysis in hepatocytes.[20] However, although the *adiponutrin* genotype has previously been reported to be associated with steatosis severity and NASH in a very small series of biopsied patients,[21,22] the relationship between adiponutrin expression and liver damage, as well as the effect of the I148M SNP on liver fibrosis, have not been evaluated in adequately powered samples.

Because the degree of steatosis has previously been reported to influence the severity of liver damage in NAFLD,[23] the aim of this study was to evaluate the relationship between the *adiponutrin* I148M polymorphism and the severity of liver fibrosis.

Patients and Methods

Patients. We analyzed DNA from 574 unrelated patients with biopsy-proven NAFLD diagnosed between January 1999 and January 2008. These included 190 Italian patients who underwent liver biopsy because of persistently abnormal liver enzymes/serum ferritin or a long-lasting history of steatosis associated with severe metabolic abnormalities, 63 Italian severely obese patients who were found to be affected by NAFLD at routine percutaneous liver biopsy performed during bariatric surgery, and 321 adult patients from Newcastle, UK. Other causes of liver disease were excluded, including increased alcohol intake (>30/>20 g/day for males/females, respectively), as confirmed by at least one family member or friend and carboxydesialylated transferrin determination, viral and autoimmune hepatitis, hereditary hemochromatosis, and alpha1-antitrypsin deficiency. Part of this group had previously been described.[9] The available demographic and clinical features are shown in Table 1. All patients included in the study had a fatty liver index >50.[24]

In addition, we tested 71 Italian family trios with both parents alive and an index child with biopsy-proven fibrotic NAFLD. We used transmission disequilibrium testing to look for preferential transmission of rs738409 C/G alleles to the affected children. This approach is not subject to the potential confounding effects inherent in case-control studies and is significantly more powerful at detecting true associations.[10,25]

Published online in Wiley InterScience (www.interscience.wiley.com).
DOI 10.1002/hep.23622
Potential conflict of interest: Nothing to report.
Additional Supporting Information may be found in the online version of this article.

Table 1. Demographic and Clinical Features of Patients with 179 Italian Healthy Control Subjects with Normal Liver Enzymes and Metabolic Parameters and 253 Patients with NAFLD, Including 190 Patients Who Underwent Liver Biopsy Because of Persistently Abnormal Liver Enzymes and/or Metabolic Abnormalities, and 63 Patients Found to Have NAFLD at Routine Liver Biopsy During Bariatric Surgery

Feature	Italian Controls	Italian NAFLD Patients	UK NAFLD Patients
Number	179	253	321
Female sex	38 (21)	76 (30)	123 (38)
Age (years)	48.4 ± 13	46.4 ± 11	49.5 ± 12
BMI (kg/m^2)	25.1 ± 2.7	30.5 ± 7.8	34.0 ± 5.3
LDL cholesterol (mg/dL)	118.7 ± 29	132.5 ± 43	127.0 ± 43
HDL cholesterol (mg/dL)	55.2 ± 13	44.9 ± 12	47.0 ± 13
Triglycerides (mg/dL)	90.1 ± 44	155.2 ± 84	248.3 ± 178
Glucose (mg/dL)	89.0 ± 10	98.3 ± 27	121.4 ± 54
HOMA-IR	2.7 ± 1.6	4.3 ± 3.4	6.4 ± 6.2
IGT-IFG or diabetes	0	63 (25)	120 (37)
ALT (IU/mL)	21.8 ± 7	55.5 ± 41	80.8 ± 64
GGT (IU/mL)	23.7 ± 16	86.2 ± 108	130.0 ± 184
NASH (%)	0	124 (49)	184 (57)
Fibrosis stage F0/F1/F2/F3/F4 n (%)	–	123/85/25/10/10 (49/33/10/4/4)	124/58/53/51/35 (39/18/16/16/11)

Values in parentheses are % values. IFG, impaired fasting glucose; IGT, impaired glucose tolerance; NASH, nonalcoholic steatohepatitis.

Controls. The Italian control group included 179 geographically-matched, age-matched, and sex-matched Italian subjects out of a larger series of 482 blood donors from Northern Italy who were selected because of the lack of clinical and biochemical evidence of liver and metabolic disease and no alcohol abuse (<30/20 g/day in males/females). We excluded subjects with ALT >35/30 IU/mL in males/females, gamma-glutamyl transferase >35 IU/mL, body mass index (BMI) >28, abdominal circumference >100 cm, glucose levels ≥100 mg/dL, triglycerides ≥150 mg/dL, high-density lipoprotein (HDL) ≤45/55 in males/females or a fatty liver index >35, a value with high specificity to rule out NAFLD in the general population.[24]

The study protocol was approved by the Institutional Review Board of the Ospedale Policlinico MaRE IRCCS, Milan, Italy, and by Local Research Ethics Committee in Newcastle, UK. Informed written consent was obtained from each patient and control subject, and the study conforms to the ethical guidelines of the 1975 Declaration of Helsinki.

Histological Assessment. Tissue sections were stained with hematoxylin and eosin, impregnated with silver for visualizing reticulin framework, and stained with trichrome for visualizing collagen. One expert pathologist unaware of clinical and genetic data reviewed all biopsies for fibrosis stage at each clinical center. The severity of steatosis, features of NASH, and fibrosis was assessed according to Kleiner et al.[26] NASH was considered to be present when steatosis, ballooning, and lobular inflammation were present. The minimum biopsy size was 1.7 cm and the number of portal areas was 10.

Genetic Analysis. DNA was extracted from peripheral blood collected at the time of liver biopsy by the phenol-chloroform method. Success rate in extracting DNA was 100% for each study group. The *adiponutrin/PNPLA3* rs738409 C→G SNP, encoding I148M, was genotyped by a TaqMan assay (assay on demand for rs738409; Applied Biosystems, Foster City, CA) by personnel unaware of clinical status of patients and controls. Post–polymerase chain reaction (PCR) allelic discrimination was carried out measuring allele-specific fluorescence on the Opticon2 detection system (MJ Research, Waltham, MA). Random samples were confirmed by direct genotyping that provided concordant results in all cases. Samples from both patients with NAFLD and controls were included in all batches analyzed, and quality controls were performed to verify the reproducibility of the results. Valid genotypic data were obtained for more than 99% of subjects analyzed.

Determination of Messenger RNA Levels of Adiponutrin and Genes Regulating Liver Metabolism and Damage. In 52 patients who underwent bariatric surgery for whom adequate liver biopsy samples collected at the time of surgery were available (comprising 42 subjects with NAFLD included in the genotype frequency study, and 10 subjects with steatosis <5% and no histological abnormalities), RNA was isolated by the Trizol reagent (Invitrogen, Carlsbad, CA), digested with deoxyribonuclease I, and quality evaluated by measuring the absorbance ratio at 260/280 nm (≥1.8) and by electrophoresis. First-strand complementary DNA was synthesized using equal amounts (0.5 μg) of total RNA, with the SuperScript VILO complementary DNA synthesis kit (Invitrogen, Carlsbad, CA). The

Table 2. Primers Used for Evaluation of Hepatic Gene Expression Analysis

Gene	Primers
Adiponutrin (PNPLA3)	Fw 5′-TTTACAGTGGCCTTATCCCT-3′
	Rv 5′-GAAAGTTCGTGGACTTGACT-3′
β-actin	Fw 5′-GGCATCCTCACCCTGAAGTA-3′
	Rv 5′-GGGGTGTTGAAGGTCTCAAA-3′
FASL	Fw 5′-GCACTTTGGGATTCTTTCCA-3′
	Rv 5′-CCTCCATTTGTCTGGCTCAT-3′
INSR	Fw 5′-TCAACGGGCAGTTTGTCGAA-3′
	Rv 5′-GCTGAAGTTCACACAGCGCC-3′
PPAR-α	Fw 5′-ATGGCATCCAGAACAAGGAG-3′
	Rv 5′-TCCCGTCTTTGTTCATCACA-3′

FASL, Fas ligand; Fw, forward primer; INSR, insulin receptor; PPAR-α, peroxisome proliferator-activated receptor-α; Rv, reverse primer; SREBP1c, steroid regulatory element binding protein 1c.

messenger RNA (mRNA) levels were analyzed by quantitative real-time PCR (qRT-PCR) with SYBR Green chemistry (Fluocycle II SYBR green mix; Euroclone, Pero, Italy). All the reactions were performed in triplicate with the Opticon2 qRT-PCR system (MJ Research, Waltham, MA). In patients with and without NAFLD, we evaluated the PNPLA3/adiponutrin mRNA levels, and we correlated the *PNPLA3* I148M genotype with the expression of insulin receptor (INSR), which regulates insulin signaling and is downregulated in the metabolic syndrome,[27] steroid regulatory element binding protein 1c (SREBP1c), which regulates lipogenesis,[28] peroxisome proliferator-activated receptor-α (PPAR-α), which regulates lipolysis and is decreased in NASH,[29] and Fas ligand (FASL), which is a proapoptotic molecule induced by insulin resistance.[30,31] Primers are shown in Table 2. Results were normalized for β-actin and 18S RNA, which were chosen as a control because of stable expression among different samples.

Statistical Analysis. Results are expressed as means $\pm$ standard deviation. Mean values were compared by analysis of variance (ANOVA) and post-hoc analysis or Wilcoxon test, when appropriate, and frequencies by Fisher's exact test, and chi-squared test for trend, when appropriate. Variables were correlated by the Spearman's rho test.

The association between the I148M variant and the presence of metabolic abnormalities, NASH, and fibrosis was evaluated by multivariate logistic regression analysis. In analyzing the association between I148M and steatosis and fibrosis, we compared allele and genotype frequencies in those with none or mild steatosis/fibrosis (0 and 1/3 for steatosis; 0 and 1/4 for fibrosis) versus those with moderate/severe steatosis/fibrosis (2 and 3/3 for steatosis; 2, 3, and 4/4 for fibrosis). Analyses were carried out with JMP 6.0 statistical analysis software (SAS Institute Inc, Cary, NC).

Results

Adiponutrin Genotype Influences the Risk of NAFLD. The frequency distribution of the rs738409 C→G *adiponutrin* genotype, which was in Hardy-Weinberg equilibrium in Italian and UK patients and in controls, is shown in Table 3. The frequency distribution of the mutant G allele was significantly higher in Italian patients with NAFLD than in geographically-matched, age-matched, and sex-matched healthy controls with normal liver enzymes and metabolic parameters and a normal fatty liver index ($P < 0.0001$; Table 2). In Italian adult subjects, the *adiponutrin* GG genotype was associated with an increased risk of NAFLD compared to the CC genotype independently of age, sex, and BMI (odds ratio = 3.29, 95% confidence interval = 1.8-6.9; Table 3).

The frequency distribution of the rs738409 C→G *adiponutrin* genotype, as well as the prevalence of the GG genotype, was very similar between Italian and UK patients with NAFLD (P = nonsignificant; Table 3).

In the family study conducted in 71 Italian trios, 56 families were "informative" in that one or both parents were heterozygous for the SNP (a total of 76 parents were heterozygous for the rs738409 C→G SNP). In these families, the variant allele was transmitted on 52 of 76 (68%) possible occasions, whereas the wild-type C allele was transmitted on only 24 of 76 (32%) occasions ($P = 0.001$).

Table 3. Frequency Distribution of the rs738409 C→G adiponutrin/PNPLA3 Genotype, Encoding for the I148M Protein Variant, in 253 Italian Patients with NAFLD and 197 Italian Healthy Subjects with Normal Liver Enzymes and Metabolic Parameters ($P < 0.0001$ for the Frequency Distribution of the G Allele Between Italian Patients and Controls), and in 321 NAFLD Patients from the United Kingdom

	PNPLA3 rs738409 C→G genotype (encoding for I148M)		
	CC (148I/I)	CG (148I/M)	GG (148M/M)
Italian patients	103 (40.7)	114 (45.1)	36 (14.2)
Italian controls	118 (65.9)	56 (31.3)	5 (2.8)
Adjusted OR*	Reference	0.80	3.29
95% CI	–	0.5-1.2	1.8-6.9
UK patients	142 (44.2)	140 (43.6)	39 (12.1)

*For Italian patients versus Italian controls. Adjusted for age, sex, and BMI. CI, confidence interval; OR, odds ratio.

Table 4. Association of the rs738409 C→G adiponutrin/PNPLA3 SNP, Encoding for the I148M Protein Variant, with Metabolic Features and Liver Damage in 253 Italian Patients with NAFLD

Characteristic	I148M PNPLA3 genotype			P†
	CC (148I/I)	CG (148I/M)	GG (148M/M)	
Number	103	114	36	
Age (years)	47 ± 11	46 ± 12	46 ± 10	0.59
Sex (F)	32 (31)	33 (28.9)	11 (30.6)	0.93
BMI (kg/m^2)	30.7 ± 8	30.7 ± 8	29.8 ± 7	0.10
LDL (mg/dL)	129 ± 41	129 ± 46	152 ± 35*,‡	0.02
HDL (mg/dL)	48 ± 13	42 ± 12*	43 ± 8*	0.004
Triglycerides (mg/dL)	145.6 ± 74	163.5 ± 94	158.5 ± 78	0.30
Glucose (mg/dL)	96 ± 19	99.9 ± 35	98 ± 15	0.71
Fasting insulin (IU/mL)	15.7 ± 9	16.7 ± 10	23.1 ± 18*,‡	0.01
HOMA-IR	3.9 ± 3.5	4 ± 2.7	5.2 ± 4.7*,‡	0.07
IFG-IGT / diabetes n (%)	16 (15)/9(8)	11 (10)/17(15)	7 (19)/3(8)	0.31
Hypertension	43 (41.7)	49 (43.0)	17 (47.2)	0.81
ALT (IU/mL)	47 ± 30	56 ± 44*	77 ± 56*,‡	0.0007
GGT (IU/mL)	91 ± 107	83 ± 119	73 ± 76	0.68

Values in parentheses are % values. M, male; F, female; NAS, nonalcoholic steatohepatitis activity score.
†P at ANOVA,
*$P < 0.05$ versus the I/I PNAPLA3 genotype,
‡$P < 0.05$ versus the I/M genotype.

Influence of Adiponutrin Genotype on Metabolic and Biochemical Parameters. The effect of the rs738409 C→G *adiponutrin* genotype, encoding for the I148M protein variant, on clinical features and metabolic and biochemical parameters in Italian patients is shown in Table 4. In patients with NAFLD, the GG genotype was associated with higher LDL, and both the CG and GG genotypes were associated with lower HDL cholesterol, although it was not associated with age, sex distribution, and BMI. The GG genotype was associated with significantly higher fasting insulin levels and homeostasis model assessment of insulin resistance (HOMA-IR) score, but not with hyperglycemia or type 2 diabetes. In addition, the G allele (148M variant) was associated with higher ALT in a dose-dependent manner ($P = 0.0007$).

Broadly similar data were obtained in the UK cohort (Supporting Information Table 1). Patients carrying the GG genotype had significantly higher ALT levels compared to those carrying the CC or GG genotypes (76.0 ± 54 IU/L CC, 74.5 ± 56 CG, 123.0 ± 103 GG; $P < 0.0001$ at ANOVA, $P < 0.05$ for GG versus either CC or CG genotypes). In contrast to the data in Italian patients, the rs738409 genotype was not associated with abnormal HDL or LDL levels.

Association Between Adiponutrin Genotype and Liver Damage. At histology, and as expected from previous studies, using an additive model for *adiponutrin* genotype, carriage of the G allele was an independent predictor for the presence of grade 2/3 steatosis together with BMI, sex, and diabetes in the combined cohort. In the UK patients alone, both genotype and age were predictors (Table 5A). More importantly, the *adiponutrin* genotype was strongly associated with the presence of NASH (Table 5B; $P = 0.007$ in the combined cohort, $P = 0.03$ in the UK cohort, and $P = 0.07$ in the Italian cohort) and stage >1 fibrosis (Table 5C; $P = 0.01$ in the combined cohort, $P = 0.06$ in the UK cohort, and $P = 0.04$ in the Italian cohort) independent of BMI, diabetes and steatosis (for NASH), and age, BMI, diabetes, steatosis, and NASH (for fibrosis). Furthermore, there was a dose effect of *adiponutrin* genotype, in that CG heterozygotes had a prevalence of NASH and fibrosis intermediate between CC and GG homozygotes (Fig. 1A,B).

Correlation Between Adiponutrin Expression, Metabolic Features, and Liver Damage. Hepatic adiponutrin mRNA levels, measured in 52 obese patients, were not significantly affected by the presence of the I148M variant (Fig. 2). Adiponutrin mRNA levels were significantly higher in females than in males (0.95 ± 0.39 versus 0.66 ± 0.24 arbitrary units, $P = 0.002$), and lower in patients with fibrosis >1 (0.64 ± 0.2 versus 0.94 ± 0.4 arbitrary units, $P = 0.01$). Adiponutrin mRNA levels correlated with serum triglycerides (rho = 0.63, $P < 0.0001$), but were not significantly associated with age, BMI, diabetes and insulin resistance, steatosis percentage, and the severity of necroinflammation.

Association Between Adiponutrin Genotype and Expression of Genes Associated with Steatosis-Related Liver Damage. Expression of INSR, SREBP1c, PPAR-

Table 5. Association of the adiponutrin/PNPLA3 rs738409 Genotype (Encoding for the 148 I/M Protein Variants) and Liver Damage, as Evaluated by Multivariate Logistic Regression Analysis, in 574 Adult Italian and UK Patients with NAFLD

	UK (n = 321)		Italy (n = 253)		Combined (n = 574)	
Variables	OR (95% CI)	*P* Value	OR (95% CI)	*P* Value	OR (95% CI)	*P* Value
A. Steatosis (0+1 versus 2+3)						
Genotype*	1.6 (1.10-2.37)	0.01	1.3 (0.88-1.9)	0.1	1.35 (1.04-1.76)	0.02
Age (years)	0.97 (0.95-0.99)	0.02	0.99 (0.97-1.02)	0.9	0.99 (0.97-1.1)	0.3
Sex	0.65 (0.37-1.14)	0.1	0.51 (0.24-2.02)	0.06	0.59 (0.38-90.89)	0.01
BMI (kg/m^2)	1.04 (0.99-1.1)	0.07	1.1 (1.05-1.15)	0.00001	1.1 (1.07-1.14)	5.03×10^{-10}
Diabetes	0.82 (0.47-1.39)	0.47	2.79 (1.45-5.55)	0.002	1.82 (1.2-2.78)	0.005
B. NASH (presence versus absence)						
Genotype*	1.5 (1.04-2.2)	0.03	1.7 (0.95-3.14)	0.07	1.5 (1.12-2.04)	0.007
Age (years)	1.0 (1.0-1.02)	0.02	1.0 (0.96-1.05)	0.2	1.01 (1.00-1.02)	0.006
Sex	0.76 (0.37-1.5)	0.4	0.44 (0.13-1.33)	0.09	0.68 (0.4-1.17)	0.1
BMI (kg/m^2)	1.04 (0.99-1.09)	0.08	1.04 (0.98-1.12)	0.2	1.03 (1.00-1.07)	0.02
Diabetes	2.38 (1.43-4.17)	0.001	4.54 (1.82-12.5)	0.001	2.78 (1.79-4.55)	0.000009
Steatosis	2.56 (1.8-3.72)	0.0000003	32.1 (13.2-95.2)	1.4×10^{-12}	4.01 (3.02-5.42)	2×10^{-16}
C. Fibrosis (0-1 versus 2-4)						
Genotype*	1.54 (0.98-2.45)	0.06	1.7 (1.02-3.1)	0.04	1.5 (1.09-2.12)	0.01
Age (years)	1.04 (1.01-1.06)	0.003	1.02 (0.98-1.06)	0.1	1.04 (1.02-1.06)	0.0001
Sex	1.28 (0.6-2.4)	0.48	0.57 (0.2-1.54)	0.2	1.1 (0.65-1.8)	0.7
BMI (kg/m^2)	1.07 (1.00-1.14)	0.02	0.99 (0.93-1.04)	0.8	1.03 (0.99-1.07)	0.08
Diabetes	2.22 (1.19-4.35)	0.01	4.2 (1.85-9.9)	0.0007	2.56 (1.59-4.17)	0.0001
Steatosis	1.13 (0.71-1.8)	0.59	0.68 (0.39-1.17)	0.16	0.99 (0.71-1.37)	0.9
NASH	22.7 (11.0-51.5)	1.03×10^{-15}	15.4 (4.8-60.4)	0.00001	16.4 (8.96-32.1)	2×10^{-16}

CI, confidence interval; OR, odds ratio; NS, not significant.
*Based on an additive model for rs738409 genotype.

α, and FASL was significantly different among *adiponutrin* genotypes ($P < 0.05$; Fig. 2). Homozygosity for the 148M variant was associated with higher expression of the proapoptotic molecule FASL, and lower mRNA levels of INSR, SREBP1c, and PPAR-α compared to the I/I 148 variant ($P < 0.05$; Fig. 2). The expression of INSR and PPAR-α were negatively correlated with percentage of steatosis and NAFLD activity score (NAS) (rho = −0.38, $P = 0.02$ and rho = −0.33, $P = 0.03$, respectively for the correlation of INSR with steatosis percentage and NAS; rho = −0.61, $P = 0.005$ and rho = −0.41, $P = 0.01$, respectively, for the correlation of PPAR-α with steatosis and NAS). SREBP1c was positively associated with steatosis percentage (rho = 0.38, $P = 0.02$). FASL expression correlated with fibrosis stage (rho = 0.36, $P = 0.02$).

Discussion

The *Adiponutrin/PNPLA3* genotype has recently been identified by genomewide association studies as a major inherited determinant of liver triglyceride accumulation and ALT levels at the population level, and the I148M variant as the most prevalent SNP predisposing to steatosis in Caucasians.[12,13] These results confirm that the frequency of the 148M variant is significantly higher in patients with NAFLD compared to controls with normal liver enzymes and metabolic parameters.[32] The risk of NAFLD was about 3.3-fold higher in subjects carrying the rs738409 GG genotype, and was detected in about 3% of healthy subjects but in 14% of patients, compared to the CC genotype. Because blood donors are often a healthy subset of the population, and we carefully further excluded subjects with any metabolic and liver enzyme abnormality, we cannot exclude that the prevalence of the G allele is higher in the general population, 20%-30% of whom have NAFLD, than in our controls.

In line with previous data,[12] the rs738409 SNP was not associated with BMI and serum triglycerides, but the 148M allele was associated with ALT levels in a dose-dependent manner. The altered HDL cholesterol levels observed in patients carrying the minor allele are consistent with recent data indicating that adiponutrin genotype influences lipoprotein metabolism.[33] However, because this association has not been observed in other populations[12] and was not confirmed in patients from the UK, additional studies are required to define whether an interaction exists between liver steatosis and the rs738409 SNP in determining lipoprotein metabolism.

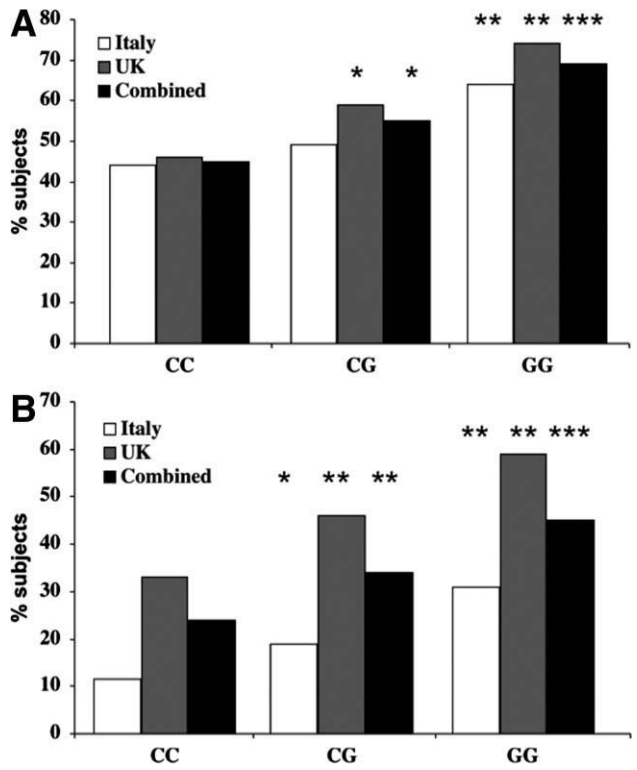

Fig. 1. Prevalence of (A) NASH and (B) fibrosis in 591 biopsied patients with NAFLD subdivided according to the adiponutrin rs738409 C→G genotype (encoding for I148M) and the referral center. (A) $P = 0.0002$ at chi-squared for trend for the prevalence of NASH across the adiponutrin genotypes in the combined series. (B) $P < 0.0001$ at chi-squared test for trend for the prevalence of fibrosis >1 across the adiponutrin genotypes in the combined series. *$P <$ 0.05, **$P < 0.005$, ***$P < 0.0001$ versus CC genotype.

In this study, we have confirmed in a large series of biopsied patients that the *adiponutrin* rs738409 genotype influences both the presence of NASH and the severity of fibrosis in patients with NAFLD independent of the degree of obesity, the presence of diabetes, and the previously demonstrated effect of genotype on steatosis. This effect was observed in two independent cohorts and confirmed in transmission disequilibrium testing of family trios with fibrotic NAFLD. The association between *adiponutrin* genotype, steatosis, ALT levels, and NASH is not a novel finding, although histological damage has been evaluated so far only in small series of patients.[21,22,34] In contrast, the association with fibrosis severity is completely novel. The observation that the association between genotype and NASH/fibrosis is independent of steatosis provides further evidence that steatosis may not be an *a priori* "first hit" in the development of progressive damage[35] and suggests that adiponutrin, possibly by regulating FFA fluxes, may be involved in pathogenic mechanisms of NASH and fibrosis that are independent of its effects on the amount of hepatic triglyceride accumulation. An effect of *adiponutrin* genotype on FFA flux, endoplasmic reticulum, and oxidative- or cytokine-mediated stress would independently influence both steatosis, necroinflammation, and fibrosis, and would reconcile the clinical observation that steatohepatitis and fibrosis severity correlate with the degree of steatosis.[19] However, we cannot exclude that the loss of liver fat, the presence of which is known to be associated with the development of severe fibrosis, may have negatively affected the association among *adiponutrin* genotype, steatosis, and fibrosis severity in this cross-sectional study.

Whether patients with the CG/GG genotype have different response to therapy and outcome, or may benefit from a tailored follow-up, needs to be evaluated in further studies.

In line with a functional effect of the I148M polymorphism on adiponutrin activity leading to steatosis,[20] the rs738409 GG genotype, coding for the 148M/M variant, did not affect hepatic adiponutrin mRNA, but influenced the expression of molecules associated with steatosis-related liver damage, such as INSR, PPAR-α, and the membrane receptor FASL. In particular, FASL sensitizes hepatocytes to apoptosis and has been implicated in the pathogenesis of NASH,[36] and indeed FASL expression was correlated with fibrosis stage in the patients evaluated. Decreased adiponutrin mRNA expression in patients with more severe liver fibrosis suggests also that decreased activity of the enzyme may predispose individuals to progressive liver disease. However, because reduced mRNA levels do not always translate to reduced protein levels,

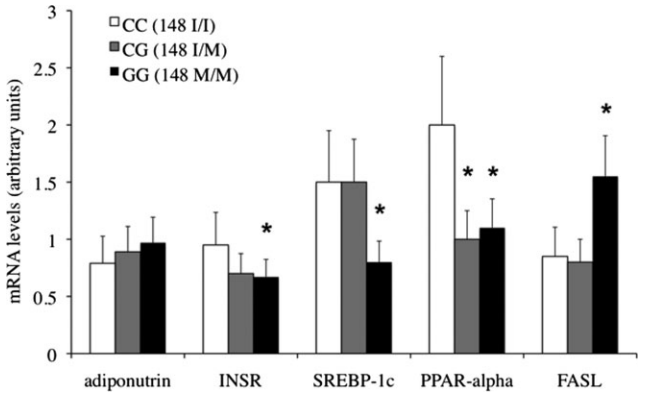

Fig. 2. Expression of adiponutrin and genes involved in insulin signaling, lipogenesis, apoptosis, and fibrogenesis according to the adiponutrin I148M genotype in 52 obese patients with NAFLD. Data are shown as mean ± standard error. FASL, Fas ligand; INSR, insulin receptor; PPAR-alpha, peroxisome proliferator-activated receptor-α; SREBP1c, steroid regulatory element binding protein 1c. $P < 0.05$ for the expression of INSR, SREBP1c, PPAR-alpha and FASL across adiponutrin genotypes. *$P <$ 0.05 versus I/I genotype.

and gene expression analysis was restricted to a subset of patients subjected to bariatric surgery, these data should be interpreted with caution. Additional mechanistic studies are required to clarify these issues.

The rationale for evaluating gene expression in bariatric patients was related to the availability at our center of larger samples of hepatic tissue from needle biopsies performed at the time of surgery, and of the full spectrum of liver pathology samples ranging from normal liver without steatosis to fibrotic steatohepatitis. Moreover, even if the high BMI of the evaluated subject may limit generalizability of the results, this population has the advantage of eliminating the confounding effect of peripheral adiposity (which was present in all subjects) on liver fat and gene expression.

A limitation of this study is the lack of ultrasonographic exclusion of liver steatosis in control subjects. However, these subjects were carefully selected among a much larger pool of healthy blood donors because of normal metabolic parameters, liver enzymes, and a very low fatty liver index.[24] Another limitation is related to the relatively small number of patients, all affected by severe obesity, who were subjected to gene expression evaluation. However, the strengths of this study are: (1) it is by some considerable margin the largest association study examining the influence of the adiponutrin I148M genotype on NAFLD histological severity; (2) the associations were present in two independent cohorts of patients and confirmed in a family study; and (3) this is the largest association study for which hepatic gene expression data in relation to the *adiponutrin* genotype are available.

In conclusion, the *adiponutrin* I148M genotype is associated with the risk of developing NAFLD and its progression to necroinflammation and fibrosis. Further studies on the function of adiponutrin are awaited with interest, but these data suggest that this is unlikely to be restricted to a role in the accumulation of hepatic triglyceride.

Acknowledgment: We thank Erika Fatta, Cristina Bertelli, Daniela Bignamini, Andrea Pietrobattista, Marco Antonio Zappa, and Ezio Lattuada for clinical assistance.

References

1. Browning JD, Szczepaniak LS, Dobbins R, Nuremberg P, Horton JD, Cohen JC, et al. Prevalence of hepatic steatosis in an urban population in the United States: impact of ethnicity. HEPATOLOGY 2004;40:1387-1395.
2. Bellentani S, Saccoccio G, Masutti F, Croce LS, Brandi G, Sasso F, et al. Prevalence of and risk factors for hepatic steatosis in Northern Italy. Ann Intern Med 2000;132:112-117.
3. Marchesini G, Brizi M, Bianchi G, Tomassetti S, Bugianesi E, Lenzi M, et al. Nonalcoholic fatty liver disease: a feature of the metabolic syndrome. Diabetes 2001;50:1844-1850.
4. Ekstedt M, Franzen LE, Mathiesen UL, Thorelius L, Holmqvist M, Bodemar G, et al. Long-term follow-up of patients with NAFLD and elevated liver enzymes. HEPATOLOGY 2006;44:865-873.
5. Bugianesi E, Leone N, Vanni E, Marchesini G, Brunello F, Carucci P, et al. Expanding the natural history of nonalcoholic steatohepatitis: from cryptogenic cirrhosis to hepatocellular carcinoma. Gastroenterology 2002;123:134-140.
6. Day CP. From fat to inflammation. Gastroenterology 2006;130:207-210.
7. Marchesini G, Bugianesi E, Forlani G, Cerrelli F, Lenzi M, Manini R, et al. Nonalcoholic fatty liver, steatohepatitis, and the metabolic syndrome. HEPATOLOGY 2003;37:917-923.
8. Wilfred de Alwis NM, Day CP. Genetics of alcoholic liver disease and nonalcoholic fatty liver disease. Semin Liver Dis 2007;27:44-54.
9. Valenti L, Fracanzani AL, Dongiovanni P, Santorelli G, Branchi A, Taioli E, et al. Tumor necrosis factor alpha promoter polymorphisms and insulin resistance in nonalcoholic fatty liver disease. Gastroenterology 2002;122:274-280.
10. Miele L, Beale G, Patman G, Nobili V, Leathart J, Grieco A, et al. The Kruppel-like factor 6 genotype is associated with fibrosis in nonalcoholic fatty liver disease. Gastroenterology 2008;135:282-291;e281.
11. Valenti L, Dongiovanni P, Piperno A, Fracanzani AL, Maggioni M, Rametta R, et al. Alpha1-antitrypsin mutations in NAFLD: high prevalence and association with altered iron metabolism but not with liver damage. HEPATOLOGY 2006;44:857-864.
12. Romeo S, Kozlitina J, Xing C, Pertsemlidis A, Cox D, Pennacchio LA, et al. Genetic variation in PNPLA3 confers susceptibility to nonalcoholic fatty liver disease. Nat Genet 2008;40:1461-1465.
13. Yuan X, Waterworth D, Perry JR, Lim N, Song K, Chambers JC, et al. Population-based genome-wide association studies reveal six loci influencing plasma levels of liver enzymes. Am J Hum Genet 2008;83:520-528.
14. Romeo S, Sentinelli F, Dash S, Yeo GS, Savage DB, Leonetti F, et al. Morbid obesity exposes the association between PNPLA3 I148M (rs738409) and indices of hepatic injury in individuals of European descent. Int J Obes (Lond) 2010;34:190-194.
15. Tian C, Stokowski RP, Kershenobich D, Ballinger DG, Hinds DA. Variant in PNPLA3 is associated with alcoholic liver disease. Nat Genet 2010;42:21-23.
16. Wilson PA, Gardner SD, Lambie NM, Commans SA, Crowther DJ. Characterization of the human patatin-like phospholipase family. J Lipid Res 2006;47:1940-1949.
17. Baulande S, Lasnier F, Lucas M, Pairault J. Adiponutrin, a transmembrane protein corresponding to a novel dietary- and obesity-linked mRNA specifically expressed in the adipose lineage. J Biol Chem 2001;276:33336-33344.
18. Hoekstra M, Li Z, Kruijt JK, Eck MV, Berkel TJ, Kuiper J. The expression level of non-alcoholic fatty liver disease-related gene PNPLA3 in hepatocytes is highly influenced by hepatic lipid status. J Hepatol 2010;52:244-251.
19. Jenkins CM, Mancuso DJ, Yan W, Sims HF, Gibson B, Gross RW. Identification, cloning, expression, and purification of three novel human calcium-independent phospholipase A2 family members possessing triacylglycerol lipase and acylglycerol transacylase activities. J Biol Chem 2004;279:48968-48975.
20. He S, McPhaul C, Li JZ, Garuti R, Kinch LN, Grishin NV, et al. A sequence variation (I148M) in PNPlA3 associated with nonalcoholic fatty liver disease disrupts triglyceride hydrolysis. J Biol Chem 2010;285:6706-6715.
21. Sookoian S, Castaño GO, Burgueno AL, Fernandez Gianotti T, Rosselli MS, Pirola CJ. A nonsynonymous gene variant in adiponutrin gene is associated with nonalcoholic fatty liver disease severity. J Lipid Res 2009;50:2111-2116.

22. Kotronen A, Johansson LE, Johansson LM, Roos C, Westerbacka J, Hamsten A, et al. A common variant in PNPLA3, which encodes adiponutrin, is associated with liver fat content in humans. Diabetologia 2009;52:1056-1060.
23. Chalasani N, Wilson L, Kleiner DE, Cummings OW, Brunt EM, Unalp A. Relationship of steatosis grade and zonal location to histological features of steatohepatitis in adult patients with non-alcoholic fatty liver disease. J Hepatol 2008;48:829-834.
24. Bedogni G, Bellentani S, Miglioli L, Masutti F, Passalacqua M, Castiglione A, et al. The Fatty Liver Index: a simple and accurate predictor of hepatic steatosis in the general population. BMC Gastroenterol 2006;6:33.
25. Spielman RS, McGinnis RE, Ewens WJ. Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). Am J Hum Genet 1993;52:506-516.
26. Kleiner DE, Brunt EM, Van Natta M, Behling C, Contos MJ, Cummings OW, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. HEPATOLOGY 2005;41: 1313-1321.
27. Accili D. Lilly lecture 2003: the struggle for mastery in insulin action: from triumvirate to republic. Diabetes 2004;53:1633-1642.
28. Qu S, Su D, Altomonte J, Kamagate A, He J, Perdomo G, et al. PPAR-{alpha} mediates the hypolipidemic action of fibrates by antagonizing FoxO1. Am J Physiol Endocrinol Metab 2007;292:E421-E434.
29. Nakamuta M, Kohjima M, Morizono S, Kotoh K, Yoshimoto T, Miyagi I, et al. Evaluation of fatty acid metabolism-related gene expression in nonalcoholic fatty liver disease. Int J Mol Med 2005;16: 631-635.
30. Zou C, Ma J, Wang X, Guo L, Zhu Z, Stoops J, et al. Lack of Fas antagonism by Met in human fatty liver disease. Nat Med 2007;13: 1078-1085.
31. Accili D, Arden KC. FoxOs at the crossroads of cellular metabolism, differentiation, and transformation. Cell 2004;117:421-426.
32. Kotronen A, Peltonen M, Hakkarainen A, Sevastianova K, Bergholm R, Johansson LM, et al. Prediction of non-alcoholic fatty liver disease and liver fat using metabolic and genetic factors. Gastroenterology 2009;137:865-872.
33. Kollerits B, Coassin S, Beckmann ND, Teumer A, Kiechl S, Doring A, et al. Genetic evidence for a role of adiponutrin in the metabolism of apolipoprotein B-containing lipoproteins. Hum Mol Genet 2009;18: 4669-4676.
34. Johansson LE, Lindblad U, Larsson CA, Rastam L, Ridderstrale M. Polymorphisms in the adiponutrin gene are associated with increased insulin secretion and obesity. Eur J Endocrinol 2008;159:577-583.
35. Yamaguchi K, Yang L, McCall S, Huang J, Yu XX, Pandey SK, et al. Inhibiting triglyceride synthesis improves hepatic steatosis but exacerbates liver damage and fibrosis in obese mice with nonalcoholic steatohepatitis. HEPATOLOGY 2007;45:1366-1374.
36. Wieckowska A, Zein NN, Yerian LM, Lopez AR, McCullough AJ, Feldstein AE. In vivo assessment of liver cell apoptosis as a novel biomarker of disease severity in nonalcoholic fatty liver disease. HEPATOLOGY 2006;44:27-33.

Research Article

EASL EUROPEAN ASSOCIATION FOR THE STUDY OF THE LIVER | JOURNAL OF HEPATOLOGY

The *SOD2* C47T polymorphism influences NAFLD fibrosis severity: Evidence from case-control and intra-familial allele association studies

Ahmad Al-Serri[1], Quentin M. Anstee[1,*], Luca Valenti[3], Valerio Nobili[2], Julian B.S. Leathart[1], Paola Dongiovanni[3], Julia Patch[1], Anna Fracanzani[3], Silvia Fargion[3], Christopher P. Day[1], Ann K. Daly[1]

[1]*Institute of Cellular Medicine, Newcastle University Medical School, Newcastle upon Tyne, UK;* [2]*Liver Unit, Bambino Gesù Children's Hospital and Research Institute, Rome, Italy;* [3]*Department of Internal Medicine, UO Medicina Interna IB, University of Milano, Ospedale Policlinico Mangiagalli e Regina Elena Fondazione IRCCS, Milano, Italy*

Background & Aims: Non-alcoholic fatty liver disease (NAFLD) is a complex disease trait where genetic variations and environment interact to determine disease progression. The association of *PNPLA3* with advanced disease has been consistently demonstrated but many other modifier genes remain unidentified. In NAFLD, increased fatty acid oxidation produces high levels of reactive oxygen species. Manganese-dependent superoxide dismutase (MnSOD), encoded by the *SOD2* gene, plays an important role in protecting cells from oxidative stress. A common non-synonymous polymorphism in *SOD2* (C47T; rs4880) is associated with decreased MnSOD mitochondrial targeting and activity making it a good candidate modifier of NAFLD severity.

Methods: The relevance of the *SOD2* C47T polymorphism to fibrotic NAFLD was assessed by two complementary approaches: we sought preferential transmission of alleles from parents to affected children in 71 family trios and adopted a case-control approach to compare genotype frequencies in a cohort of 502 European NAFLD patients.

Results: In the family study, 55 families were informative. The T allele was transmitted on 47/76 (62%) possible occasions whereas the C allele was transmitted on only 29/76 (38%) occasions, $p = 0.038$. In the case control study, the presence of advanced fibrosis (stage >1) increased with the number of T alleles, $p = 0.008$ for trend. Multivariate analysis showed susceptibility to advanced fibrotic disease was determined by *SOD2* genotype (OR 1.56 (95% CI 1.09–2.25), $p = 0.014$), *PNPLA3* genotype ($p = 0.041$), type 2 diabetes mellitus ($p = 0.009$) and histological severity of NASH ($p = 2.0 \times 10^{-16}$).

Conclusions: Carriage of the *SOD2* C47T polymorphism is associated with more advanced fibrosis in NASH.

Keywords: NAFLD; NASH; Steatohepatitis; Gene; Polymorphism; Oxidative stress.
Received 26 April 2011; received in revised form 19 May 2011; accepted 23 May 2011
* Corresponding author. Address: Institute of Cellular Medicine, The Medical School, Newcastle University, 3rd Floor, William Leech Building, Framlington Place, Newcastle-upon-Tyne NE2 4HH, UK. Tel.: +44 (0) 191 222 7012; fax: +44 (0) 191 222 6621.
E-mail address: quentin.anstee@newcastle.ac.uk (Q.M. Anstee).

Introduction

Non-alcoholic fatty liver disease (NAFLD), widely considered to be the hepatic manifestation of the metabolic syndrome, encompasses a spectrum of liver disease from simple steatosis (fatty liver) through non-alcoholic steatohepatitis (NASH) to fibrosis and ultimately cirrhosis in the absence of alcohol abuse [1]. The prevalence of NAFLD is rapidly increasing and it is now the most common cause of chronic liver disease in Western countries [2]. Despite its high prevalence, only a minority of patients with steatosis progress to develop inflammation and less than a quarter of subjects with NAFLD ever progress beyond steatosis to significant fibrosis and cirrhosis [3,4].

NAFLD is best considered a complex disease trait where subtle inter-patient variations including host genetic factors and environment interact to produce disease phenotype and determine disease progression [5–7]. Basic science is helping to elucidate the mechanisms perpetuating liver cell injury and fibrosis in NAFLD through *in vitro* and *in vivo* studies. Attention is currently focused on the role of increased free fatty acid (FFA) flux on a background of insulin resistance as key drivers of pathogenesis through hepatocellular oxidative stress secondary to reactive oxygen species (ROS) production during β- and ω-FFA oxidation, direct lipotoxicity, cytokine release and endoplasmic reticulum stress. Consequent cellular damage triggers a mixture of immune mediated hepatocellular injury and both necrotic and apoptotic cell death pathways [8–11]. Persistence of these processes culminates in hepatic fibrosis [12].

Whilst the reasons for the apparent variation in individual susceptibility to progressive disease are incompletely understood, family/ethnic studies suggest that genetic factors play a significant role [13,14]. Accordingly, genes encoding proteins influencing the magnitude of these different forms of cellular stress are obvious candidates as genetic factors contributing to susceptibility to progressive NAFLD [5]. Single nucleotide polymorphisms (SNPs) in several candidate genes involved in inflammation, oxidative stress and fibrogenesis have been associated with the severity of liver damage in NAFLD (reviewed [5]). As yet, only the association of *PNPLA3* with advanced NASH has been consistently replicated in large studies [5].

Please cite this article in press as: Al-Serri A et al. The *SOD2* C47T polymorphism influences NAFLD fibrosis severity: Evidence from case-control and intra-familial allele association studies. J Hepatol (2011), doi:10.1016/j.jhep.2011.05.029

Research Article

The mitochondrial enzyme manganese-dependent superoxide dismutase (MnSOD), encoded by the nuclear *SOD2* gene, plays an important role in protecting cells from superoxide radicals [15]. *SOD2* is subject to a common polymorphism (C47T, rs4880) which results in an amino acid substitution in the signal sequence targeting the enzyme to the mitochondrion (Ala16Val) and may induce a conformational change in the protein tertiary structure [16]. The presence of alanine at position −9 in this sequence (C47 allele) has been demonstrated to be associated with more efficient protein import than valine (T47 allele) by *in vitro* expression studies and is predicted to result in higher enzyme activity [17,18]. This *SOD2* polymorphism (rs4880) has been investigated as a possible susceptibility factor in NASH and several other diseases where oxidative stress is considered to play a role in pathogenesis including hereditary hemochromatosis [19] and drug induced liver injury [20]. A small study in 63 Japanese patients found a significantly increased prevalence of the lower activity homozygous T genotype among cases [21]. More recently, the homozygous T genotype was reported to be associated with an increased incidence of diabetic nephropathy in large studies of type 1 diabetics from Sweden [22] and Denmark [23] and with susceptibility to pancreatic cancer in a US-based study [24]. *SOD2* has also been investigated in relation to susceptibility to alcoholic liver disease but the results have been inconsistent and inconclusive [25,26].

To further examine this biologically plausible association between *SOD2* genotype and susceptibility to fibrosing steatohepatitis, we have used two complementary approaches. Firstly, we have carried out a family study analyzing trios consisting of children with fibrotic NAFLD and their two parents in which we have performed transmission disequilibrium test (TDT) analysis to determine whether there is preferential transmission of a particular parental allele to the affected children [27]. Secondly, we performed a classical case-control allelic association study in unrelated patients with NAFLD of varying severity including five times more patients than that reported previously [21].

Materials and methods

Patients

A large patient cohort was recruited across centers in UK and Italy. The study had all necessary ethical approvals in both countries and all participants (or their parents) gave informed consent.

The Italian family study, collected DNA from 71 Italian family 'trios' in Rome (210 individuals) each comprising two living parents and an index child with biopsy-proven fibrotic NAFLD. Other causes of liver disease were excluded, including increased alcohol intake, chronic viral hepatitis, autoimmune hepatitis, hereditary hemochromatosis, α1-antitrypsin deficiency, Wilson's disease, and drug induced liver disease. Detailed clinical data concerning these children has been presented previously [28].

For the case-control study, 502 European Caucasian patients with biopsy-proven NAFLD of different stages of disease were enrolled (338 patients from Newcastle upon Tyne, UK and a further 164 patients from Milan, Italy). Baseline characteristics of the cohort are shown in Table 1. These were unrelated patients with NAFLD, derived from a patient population originally identified as having ultrasonographically detected bright liver and abnormal biochemical tests (ALT and GGT) between January 1999 and 2007 [29]. Alternate diagnoses were excluded, including increased alcohol intake (males and females consuming greater than 21/14 U of alcohol per week [>30/20 g/day ethanol] respectively were excluded), as were any individuals with chronic viral hepatitis (hepatitis B and hepatitis C), autoimmune liver diseases, hereditary hemochromatosis, α1-antitrypsin deficiency, Wilson's disease, and drug induced liver disease.

Clinical and laboratory data were collected on the date a diagnostic liver biopsy was performed. Body mass index (BMI) was calculated using the formula: weight (kilograms)/height (m^2). The presence of diabetes mellitus (fasting glucose ⩾7.1 mmol/L mg/dl or treatment with anti-diabetic drugs) and hypertension (blood pressure ⩾130/85 or on treatment for previously diagnosed hypertension) was recorded. Laboratory evaluation included routine liver biochemistry (alanine and aspartate aminotransferase, total bilirubin, albumin, alkaline phosphatase and gamma glutamyl transpeptidase); complete blood count; total- and HDL-cholesterol and total triglycerides; fasting glucose; fasting insulin; viral serology for hepatitis B and C infection, and autoantibodies.

Liver biopsy

Ultrasound guided liver biopsy was performed in all patients. Specimens (at least 1.6 cm length and 5 μm thick) were fixed in formalin for evaluation. Tissue sections were stained with hematoxylin and eosin, impregnated with silver for visualizing reticulin framework, and stained with trichrome for visualizing collagen. Liver biopsies were reviewed by a single expert liver pathologist at each participating center, unaware of clinical or genetic data. The severity of steatosis, necroinflammatory grade and stage of fibrosis were scored according to modified Brunt criteria [30]. For fibrosis, stage 0 = no fibrosis; stage 1 = isolated perisinusoidal or portal fibrosis; stage 2 = perisinusoidal and portal/periportal fibrosis; stage 3 = septal or bridging fibrosis; and stage 4 = cirrhosis. As the main aim of this study was to evaluate the possible role of the polymorphism of *SOD2* in predicting patients with progressive NAFLD, *a priori* we considered the presence of stage 2 fibrosis in adults to be indicative of a more progressive disease phenotype. Given the younger age and therefore limited duration of disease exposure of pediatric patients, the presence of fibrosis stage 1 or greater was considered indicative of a progressive disease phenotype in the family study.

DNA preparation

Venous blood was collected at the time of liver biopsy and DNA was prepared from peripheral blood lymphocytes as described previously [31]. Genotyping was performed by personnel unaware of clinical status or histology of patients.

SOD2 rs4880 genotyping

SOD2 genotype was determined by PCR-RFLP analysis using a minor modification of our previously described method [26]. The final volume (20 μl) of the PCR reaction contained 0.5–1 μg genomic DNA, 0.625 U Taq DNA polymerase, 1× Taq DNA polymerase reaction buffer (50 mM potassium chloride, 10 mM Tris–HCl pH 9.0, 0.1% (v/v) Triton X-100, 1.5 mM $MgCl_2$), 0.1 mM dNTPs and 0.25 μM both forward and reverse primers (5′-CAGCCCAGCCTGCGTAGACGG-3′ and 5′-GCGCGTTGATG TGAGGTTCCAG-3′). Amplification (35 cycles of denaturation at 95 °C for 1 min, annealing at 63 °C for 1 min and extension at 72 °C for 1 min) was then performed in a GeneAmp PCR system 9700 thermal cycler. After successful amplification, PCR products (6 μl) were diluted with the appropriate restriction enzyme buffer to a final volume of 20 μl and 5 U *Bsa*WI added. The digestion was then incubated at 60 °C for 3 h. The presence of a T at the polymorphic site results in the creation of a restriction site for *Bsa*WI. Digests were analyzed by polyacrylamide gel electrophoresis on 10% polyacrylamide gels in 1× TBE buffer and visualized by staining with ethidium bromide.

Statistical analysis

Statistical analyses were performed using various packages including 'genetics', 'combinat' and 'dgc.genetics' running in the R software environment (R version 2.7.1) [32]. Transmission disequilibrium tests (TDT) were performed to determine preferential transmission of alleles and statistical significance in the family study. This approach is not subject to many of the potential confounding effects inherent in case control studies and is significantly more powerful at detecting true associations [27]. In the cohort study, we evaluated by univariate and multivariate analysis the capability of *SOD2* to predict progressive NAFLD (fibrosis 2–4). Parametric and non-parametric data were presented as means ± standard deviation and percentage when appropriate. The statistical analysis was performed using ANOVA, Student *t* test, Pearson Chi-square test and Chi-square test for trend, when appropriate. Significance was taken as $p < 0.05$. Hardy–Weinberg equilibrium was determined for each study population using the web-based calculator available at www.tufts.edu/ which confirmed that UK and Italian study populations were in equilibrium. The multivariate analysis was performed by logistic

Please cite this article in press as: Al-Serri A et al. The *SOD2* C47T polymorphism influences NAFLD fibrosis severity: Evidence from case-control and intra-familial allele association studies. J Hepatol (2011), doi:10.1016/j.jhep.2011.05.029

JOURNAL OF HEPATOLOGY

Table 1. Cohort characteristics.

	All	UK	Italy	*p* value*
Number	502	338	164	-
Country/Ethnicity	Caucasian	UK Caucasian	Italy Caucasian	-
Sex (Male)	339 (67.5%)	210 (62.1%)	129 (78.6%)	
SOD2 rs4880 C-Allele frequency	0.491	0.494	0.484	
Age, years	49.1 ± 12.3	49.7 ± 12.7	47.9 ± 11.4	0.1
BMI, kg/m^2	31.8 ± 5.7	34.05 ± 5.3	27.4 ± 3.8	$4.5x10^{-45}$
DM	166 (33.5%)	127 (38.3%)	39 (23.7%)	0.001
HOMA-IR	5.5 ± 5.3	6.3 ± 6.2	4.3 ± 3.5	0.0003
ALT, IU/L	75.1 ± 59.2	80.6 ± 64.05	63.7± 46.1	0.001
Total cholesterol	5.44 ± 1.5	5.6 ± 1.3	5.09 ± 1.7	0.0007
TG	2.42 ± 1.8	2.81 ± 2.0	1.64 ± 1.0	$2x10^{-16}$
Steatosis score				
1	198 (40%)	96 (29%)	102 (62%)	<0.0001
2	178 (36%)	145 (44%)	33 (20%)	<0.0001
3	115 (24%)	86 (27%)	29 (18%)	0.03
NASH (yes)	263 (53%)	184 (55%)	79 (48%)	0.18
Fibrosis score				
0	215 (43%)	129 (38%)	86 (52.4%)	0.002
1	108 (22%)	62 (18%)	46 (28.1%)	0.01
2	76 (15%)	56 (17%)	20 (12.1%)	0.23
3	62 (12%)	55 (16%)	7 (4.3%)	0.00008
4	41 (8%)	36 (11%)	5 (3.1%)	0.004

*Statistical analysis of UK *vs.* Italian recruitment centers.

regression analysis to evaluate the factors associated with progressive NAFLD (fibrosis stage $\geqslant 2$). The results of the multivariate analysis are expressed as odds ratio (OR) with 95% confidence intervals (CI).

Results

The SOD2 C47T allele is preferentially transmitted to children with fibrosing steatohepatitis

In the Italian family study, 61 out of the 71 children had fibrosing steatohepatitis (59 patients exhibited stage 1 fibrosis and 2 patients had stage 2 disease). Transmission disequilibrium testing (TDT) [27] was used to seek preferential transmission of either *SOD2* allele to affected children in the family study. For *SOD2* rs4880, 55 out of the 71 families were informative in that one or both parents were heterozygous for this SNP. In these families, the T allele was transmitted on 47/76 (62%) possible occasions whereas the C allele was transmitted on only 29/76 (38%) occasions, $p = 0.038$.

Carriage of the C47T SOD2 genotype is associated with advanced fibrosis

To determine whether carriage of the *SOD2* rs4880 SNP influenced susceptibility to fibrosing steatohepatitis in an adult population, we examined whether it was associated with histological disease progression in a large cohort of NAFLD patients. The total study population of 502 patients with biopsy-proven NAFLD was genotyped for *SOD2* rs4880. Clinical details are reported in Table 1. Broadly, the UK study population exhibited a more severe metabolic syndrome and NASH phenotype than the Italian population. *SOD2* genotypes were in Hardy–Weinberg equilibrium. Similar *SOD2* allele frequencies were observed at both the UK (C47: 49.4%) and Italian (C47: 48.4%) centers which were consistent with those observed in a cohort of North-Western European descent by the International HapMap project (www.ncbi.nlm.nih.gov/projects/SNP; rs4880).

Table 2 summarizes the relationship between *SOD2* genotype and a number of patient-specific and clinical parameters relevant to the disease. Both age and a fibrosis score >1 were significantly associated with *SOD2* genotype. In particular, a gene-dosage effect was observed with the incidence of advanced fibrosis (stage >1) increasing with the number of T (Val) alleles ($p = 0.008$, χ for trend) (Fig. 1). 72.2% of CC individuals had a fibrosis score of 0 or 1 compared with 56.2% of homozygous T patients. Homozygosity for the T allele *vs.* C was associated with an odds ratio of 2.02 (95% CI 1.19–3.45; $p = 0.008$) for the development of fibrosis of grade 2 or above. There was no significant difference for steatosis score or other clinical parameters, including the presence of diabetes, between *SOD2* genotype groups.

Multivariate analysis and relevance of PNPLA3 genotype

Given the previous reports demonstrating a role for *PNPLA3* I148M [33,34] as a modifier of disease progression in NAFLD, the cohort was also genotyped for this polymorphism. A multivariate logistic analysis was performed to control for the effect of *PNPLA3* together with other factors relevant to disease severity. In line with our previous findings [34], a dominant effect for the

Please cite this article in press as: Al-Serri A et al. The *SOD2* C47T polymorphism influences NAFLD fibrosis severity: Evidence from case-control and intra-familial allele association studies. J Hepatol (2011), doi:10.1016/j.jhep.2011.05.029

Research Article

Table 2. Comparison of selected patient characteristics according to *SOD2* genotype.

Phenotype	Combined Cohort			
	CC (n = 119)	CT (n = 255)	TT (n = 128)	*p* value
Age (yrs)	49.6 ± 12.6	47.9 ± 12.1	51.3 ± 12.1	0.036
BMI (kg/m^2)	31.9 ± 5.9	32.1 ± 6.1	30.9 ± 5.1	0.1
Gender				
Male	76 (0.64)	173 (0.68)	90 (0.70)	0.2
Female	43 (0.36)	82 (0.32)	38 (0.30)	
Diabetes				
Yes	32 (0.27)	87 (0.35)	47 (0.38)	0.07
No	87 (0.73)	164 (0.65)	78 (0.62)	
Fibrosis >1				
Yes	33 (0.19)	90 (0.50)	56 (0.31)	0.008
No	86 (0.27)	165 (0.51)	72 (0.22)	
NASH				
Yes	59 (0.22)	131 (0.5)	73 (0.27)	0.1
No	60 (0.26)	121 (0.52)	51 (0.22)	
Steatosis >1				
Yes	66 (0.23)	146 (0.5)	81 (0.27)	0.11
No	52 (0.26)	104 (0.53)	42 (0.21)	
Hypertension				
Yes	58 (0.23)	118 (0.47)	74 (0.3)	0.1
No	53 (0.24)	124 (0.56)	44 (0.2)	
ALT (IU/L)	70.4 ± 49.7	77.1 ± 62.2	75.5 ± 61.4	0.6
Glucose	35.4 ± 41.7	37.8 ± 49.2	40 ± 47	0.7
Insulin	21 ± 18.8	20.3 ± 17.8	19.2 ± 14.6	0.8
HOMA-IR	5.6 ± 6.2	5.6 ± 5.2	5.1 ± 4.5	0.8
Total cholesterol (mmol/L)	5.3 ± 1.4	5.5 ± 1.4	5.2 ± 1.6	0.08
TG (mmol/L)	2.3 ± 1.5	2.5 ± 2.1	2.2 ± 1.3	0.1

Population Hardy–Weinberg calculation (UK: $\chi^2 = 0.58$, $p = 0.44$; Italy: $\chi^2 = 0.21$, $p = 0.65$; combined: $\chi^2 = 0.13$, $p = 0.72$) i.e. in equilibrium. Advanced fibrosis with TT genotype *vs.* CC 2.02 (95% CI 1.19–3.45), $p = 0.008$; CC *vs.* CT 1.4 (0.88–2.28), $p = 0.14$.

Table 3. Multivariate analysis by disease element. (A) Steatosis 0 + 1 *vs.* 2 + 3. (B) NASH 0 *vs.* 1. (C) Fibrosis 0 + 1 *vs.* 2 + 3 + 4.

A

Variables	OR (95% CI)	*p* value
SOD2	1.35 (1.01-1.80)	0.039
PNPLA3	1.64 (1.09-2.46)	0.016
Age	0.98 (0.96-0.99)	0.040
Sex	0.74 (0.46-1.19)	0.222
BMI	1.06 (1.02-1.11)	0.003
Diabetes	1.46 (0.94-2.27)	0.09
Recruitment Centre	0.33 (0.20-0.54)	1.2×10^{-5}

B

Variables	OR (95% CI)	*p* value
SOD2	1.09 (0.81-1.48)	0.538
PNPLA3	1.44 (0.94-2.21)	0.091
Age	1.02 (1.00-1.04)	0.024
Sex	0.74 (0.44-1.22)	0.244
BMI	1.06 (1.01-1.11)	0.008
Diabetes	2.97 (1.86-4.78)	5.3×10^{-6}
Steatosis	3.91 (2.86-5.45)	2×10^{-16}
Recruitment Centre	2.21 (1.27-3.91)	0.005

C

Variables	OR (95% CI)	*p* value
SOD2	1.56 (1.09-2.25)	0.014
PNPLA3	1.69 (1.02-2.84)	0.041
Age	1.02 (0.99-1.04)	0.056
Sex	1.41 (0.81-2.5)	0.231
BMI	1.03 (0.97-1.08)	0.261
Diabetes	1.94 (1.17-3.21)	0.009
Steatosis	0.81 (0.56-1.16)	0.273
NASH	21.9 (1.16-44.3)	2×10^{-16}
Recruitment Centre	0.33 (0.17-0.62)	0.0006

Calculations assume a co-dominant model for *SOD2* and a dominant model for *PNPLA3* variants.

PNPLA3 variant allele was assumed. Recruitment center (UK *vs.* Italy) was included within the analysis to ensure that variations in environmental influences were controlled. In multivariate analysis, degree of steatosis was found to be significantly associated with BMI and *SOD2* genotype (odds ratio 1.35 (95% CI 1.01–1.80), $p = 0.039$) as well as, in agreement with our previous findings, *PNPLA3* genotype (Table 3A). Steatohepatitis was independently associated with a greater degree of steatosis, age, BMI and the presence of diabetes (Table 3B). Consistent with its metabolic function and supporting the view that *SOD2* genotype exerts its modifier effect through direct modulation of fibrogenic response to oxidative stress rather than initiation of steatohepatitic injury, carriage of neither *SOD2* allele predicted severity of steatohepatitis (Table 3B). The analysis confirmed that carriage of the *SOD2* rs4880 polymorphism was an independent risk factor for advanced fibrosis (OR 1.56 (95% CI 1.09–2.25), $p = 0.014$) (Table 3C). The effect of *SOD2* on fibrosis was comparable to that of *PNPLA3* (rs738409; I148M) (OR 1.69, $p = 0.041$). Histological steatohepatitis grade also remained a strong predictor of fibrosis stage (OR 21.9 (95% CI 1.16–44.3), $p = 2.0 \times 10^{-16}$).

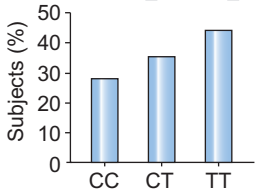

Fig. 1. Gene dosage effect of *SOD2* C47T SNP on NASH associated fibrosis. Prevalence of advanced fibrosis in 502 patients with biopsy proven NASH subdivided according to SOD2 C47T genotype. Chi-squared for trend $p = 0.008$. CC *vs.* CT + TT: OR 1.6 (1.02–2.5) $p = 0.038$.

Discussion

Using both case-control and intra-familial association methodologies, we have shown a consistent association between a

Please cite this article in press as: Al-Serri A et al. The *SOD2* C47T polymorphism influences NAFLD fibrosis severity: Evidence from case-control and intra-familial allele association studies. J Hepatol (2011), doi:10.1016/j.jhep.2011.05.029

functional SNP in the mitochondrial targeting sequence of *SOD2* and fibrosis severity in NAFLD. These results provide persuasive genetic evidence that mitochondria-derived oxidative stress is important in the pathogenesis of advanced NAFLD. In particular, use of TDT analysis in a young NAFLD population is a novel feature. In addition, we have also demonstrated a strong association with *SOD2* genotype in a European case-control study population ($p = 0.008$, χ for trend). The central role of MnSOD in cellular protection against oxidative stress has been well demonstrated by the finding that deficiency of this enzyme is lethal in mice [35]. Homozygous knockout mice typically die within the first 10 days post-natally, with impairment of function of several organs, particularly the heart [36]. Hepatic steatosis is also observed in this model, reflecting the interaction between ROS production and exacerbation of hepatic insulin resistance and lipid accumulation. An effect that was also observed in the current study where increased steatosis with *SOD2* genotype was found ($p = 0.039$), possibly reflecting oxidative stress mediated Apolipoprotein B degradation and impaired VLDL excretion [37]. Heterozygous animals have a normal lifespan but evidence for increased hydroperoxide accumulation in hepatocyte mitochondria followed by apoptosis has been reported [36]. More recently, conditional knockouts with deficiency of MnSOD only in hepatocytes have been generated. These animals show a number of signs of liver failure as well as loss of metabolic zonation [38].

The presence of alanine at position-9 in the *SOD2* sequence (C47 allele) has been demonstrated to be associated with more efficient protein import than valine (T47 allele) by *in vitro* expression studies and is predicted to result in greater *SOD2* enzyme activity[17,18]. Expression of the *SOD2* gene is inducible by a variety of modulators ranging from cytokines to UV irradiation [15] but it is likely that even when induced, the difference between the two allelic variants will be maintained. Our association of T allele carriage with greater NAFLD severity is consistent with the previous small study in a Japanese population [21] and also with studies on other diseases where oxidative stress is likely to be relevant to disease pathogenesis including diabetic nephropathy [22], pancreatic adenocarcinoma [24], exocrine pancreatic insufficiency in chronic alcoholic pancreatitis [39] and diabetic retinopathy [40]. *SOD2* is a relatively small gene of 14 kb and consists of a single haplotype block. SNPs other than rs4880 have been described but most of those studied previously, for example rs2855116 [41], are in strong linkage disequilibrium with the rs4880 SNP, which is the only common non-synonymous polymorphism described in *SOD2*.

In the mammalian liver, fatty acid oxidation occurs in three organelles; β-oxidation takes place in the mitochondria and peroxisomes and cytochrome P4504A mediated ω-oxidation takes place in the microsomes [42,43]. The synthesis of reactive oxygen species by fatty acid metabolism and the mitochondrial respiratory chain is increased in patients with NAFLD [44]. ROS are able to induce lipid peroxidation of the expanded lipid stores, compromise mitochondrial integrity and promote apoptotic cell death [44]. There is increasing evidence from studies in both humans and rodents that fibrotic NAFLD is associated with increased levels of ROS and mitochondrial abnormalities [45]. Superoxide-generated lipid hydroperoxides degrade to several hydroxy-alkenals including 4-hydroxynonenal. Studies in both humans and animals suggest that 4-HNE induces and/or activates uncoupling protein 2 (*UCP2*) [46]. Though this process results in uncoupling of substrate oxidation from ATP synthesis and should avoid a further increase in mitochondrial hydrogen peroxide production, ATP depletion makes cells more vulnerable to damage if exposed to further insults. The high activity (C47) form of MnSOD is likely to protect mitochondria from superoxide exposure better than the T47 form, thus avoiding the need to uncouple the electron transfer process.

As with many other complex genetic diseases, susceptibility to fibrotic NAFLD is likely to be determined by epistatic interaction of a number of different genes and environmental influences, indeed several genetic modifiers have been reported across a range of progressive liver diseases [5,34,47–53]. In clinical studies, it is difficult or impossible to control for all subtle inter-patient variations, whether genetic or environmental, that interact to determine disease phenotype and progression [7] and so there remains a need for validation in independent patient cohorts, either through further candidate gene studies or large-scale genome-wide association studies. However, *SOD2* represents a good example of a biologically plausible candidate gene with a common, well-established, functionally significant polymorphism. This study clearly demonstrates the predicted association with disease susceptibility for fibrosing steatohepatitis and has several key methodological strengths: (i) it is the largest association study examining the influence of *SOD2* on NASH associated fibrosis; (ii) the cohort studied comprised patients with biopsy proven NASH; and (iii) the association was further demonstrated in a family study where preferential transmission of the low activity (T47) allele was seen in offspring with biopsy proven NASH. Much focus is given to changes in gene expression that increase ROS in NASH pathogenesis whilst the role of modifiers of host defense is too often ignored. The current study highlights the importance of this aspect of disease pathogenesis. Although results of some therapeutic trials of anti-oxidants in NASH have been inconclusive [54–59], other large studies have been supportive [60]. Taken together with the data presented here, these provide a strong rationale for further investigation of the utility of anti-oxidants to ameliorate fibrosis progression in steatohepatitis.

Conflict of interest

The authors who have taken part in this study declared that they do not have anything to disclose regarding funding or conflict of interest with respect to this manuscript.

Acknowledgements

Q.M.A. is the recipient of a Clinical Senior Lectureship Award from the Higher Education Funding Council for England (HEFCE). Q.M.A., A.K.D. and C.P.D. are recipients of an MRC Programme Grant to identify and study novel genetic modifiers of liver disease and are members of the European Union FP7 funded 'Fatty Liver Inhibition of Progression (FLIP)' consortium.

References

[1] Cortez-Pinto L, De Moura MC, Day CP. Non-alcoholic steatohepatitis: from cell biology to clinical practice. J Hepatol 2006;44:197–208.
[2] Neuschwander-Tetri BA, Caldwell SH. Nonalcoholic steatohepatitis: summary of an AASLD Single Topic Conference. Hepatology 2003;37:1202–1219.
[3] Wanless IR, Lentz JS. Fatty liver hepatitis (steatohepatitis) and obesity: an autopsy study with analysis of risk factors. Hepatology 1990;12:1106–1110.

Please cite this article in press as: Al-Serri A et al. The *SOD2* C47T polymorphism influences NAFLD fibrosis severity: Evidence from case-control and intra-familial allele association studies. J Hepatol (2011), doi:10.1016/j.jhep.2011.05.029

Research Article

[4] Dixon J, Bhathal P, O'Brian P. Non-alcoholic fatty liver disease: predictors of non-alcoholic steatohepatitis and liver fibrosis in the severely obese. Gastroenterology 2001;121:91–100.
[5] Anstee QM, Daly A, Day CP. Genetics of alcoholic and non-alcoholic fatty liver disease. Semin Liver Dis 2011;31:128–146.
[6] Day CP. Pathogenesis of steatohepatitis. Best Pract Res Clin Gastroenterol 2002;16:663–678.
[7] Anstee QM, Goldin RD. Mouse models in non-alcoholic fatty liver disease and steatohepatitis research. Int J Exp Pathol 2006;87:1–16.
[8] Malhi H, Gores GJ, Lemasters JJ. Apoptosis and necrosis in the liver: a tale of two deaths? Hepatology 2006;43:S31–S44.
[9] Anstee QM, Concas D, Kudo H, Levene A, Pollard J, Charlton P, et al. Impact of pan-caspase inhibition in animal models of established steatosis and non-alcoholic steatohepatitis. J Hepatol 2010;53:542–550.
[10] Farrell GC, Larter CZ, Hou JY, Zhang RH, Yeh MM, Williams J, et al. Apoptosis in experimental NASH is associated with p53 activation and TRAIL receptor expression. J Gastroenterol Hepatol 2009;24:443–452.
[11] Day CP. From fat to inflammation. Gastroenterology 2006;130:207–210.
[12] Iredale JP. Models of liver fibrosis: exploring the dynamic nature of inflammation and repair in a solid organ. J Clin Invest 2007;117:539–548.
[13] Struben VM, Hespenheide EE, Caldwell SH. Nonalcoholic steatohepatitis and cryptogenic cirrhosis within kindreds. Am J Med 2000;108:9–13.
[14] Willner IR, Waters B, Patil SR, Reuben A, Morelli J, Riely CA. Ninety patients with nonalcoholic steatohepatitis: insulin resistance, familial tendency, and severity of disease. Am J Gastroenterol 2001;96:2957–2961.
[15] MacMillan-Crow LA, Cruthirds DL. Invited review – manganese superoxide dismutase in disease. Free Radic Res 2001;34:325–336.
[16] Shimoda-Matsubayashi S, Matsumine H, Kobayashi T, Nakagawa-Hattori Y, Shimizu Y, Mizuno Y. Structural dimorphism in the mitochondrial targeting sequence in the human manganese superoxide dismutase gene. A predictive evidence for conformational change to influence mitochondrial transport and a study of allelic association in Parkinson's disease. Biochem Biophys Res Commun 1996;226:561–565.
[17] Sutton A, Imbert A, Igoudjil A, Descatoire V, Cazanave S, Pessayre D, et al. The manganese superoxide dismutase Ala16Val dimorphism modulates both mitochondrial import and mRNA stability. Pharmacogenet Genomics 2005;15:311–319.
[18] Sutton A, Khoury H, Prip-Buus C, Cepanec C, Pessayre D, Degoul F. The Ala(16)Val genetic dimorphism modulates the import of human manganese superoxide dismutase into rat liver mitochondria. Pharmacogenetics 2003;13:145–157.
[19] Valenti L, Conte D, Piperno A, Dongiovanni P, Fracanzani AL, Fraquelli M, et al. The mitochondrial superoxide dismutase A16V polymorphism in the cardiomyopathy associated with hereditary haemochromatosis. J Med Genet 2004;41:946–950.
[20] Lucena MI, Garcia-Martin E, Andrade RJ, Martinez C, Stephens C, Ruiz JD, et al. Mitochondrial superoxide dismutase and glutathione peroxidase in idiosyncratic drug-induced liver injury. Hepatology 2010;52:303–312.
[21] Namikawa C, Shu-Ping Z, Vyselaar JR, Nozaki Y, Nemoto Y, Ono M, et al. Polymorphisms of microsomal triglyceride transfer protein gene and manganese superoxide dismutase gene in non-alcoholic steatohepatitis. J Hepatol 2004;40:781–786.
[22] Mollsten A, Marklund SL, Wessman M, Svensson M, Forsblom C, Parkkonen M, et al. A functional polymorphism in the manganese superoxide dismutase gene and diabetic nephropathy. Diabetes 2007;56:265–269.
[23] Mollsten A, Jorsal A, Lajer M, Vionnet N, Tarnow L. The V16A polymorphism in SOD2 is associated with increased risk of diabetic nephropathy and cardiovascular disease in type 1 diabetes. Diabetologia 2009;52:2590–2593.
[24] Wheatley-Price P, Asomaning K, Reid A, Zhai R, Su L, Zhou W, et al. Myeloperoxidase and superoxide dismutase polymorphisms are associated with an increased risk of developing pancreatic adenocarcinoma. Cancer 2008;112:1037–1042.
[25] Nahon P, Sutton A, Pessayre D, Rufat P, Degoul F, Ganne-Carrie N, et al. Genetic dimorphism in superoxide dismutase and susceptibility to alcoholic cirrhosis, hepatocellular carcinoma, and death. Clin Gastroenterol Hepatol 2005;3:292–298.
[26] Stewart SF, Leathart JB, Chen Y, Daly AK, Rolla R, Vay D, et al. Valine–alanine manganese superoxide dismutase polymorphism is not associated with alcohol-induced oxidative stress or liver fibrosis. Hepatology 2002;36:1355–1360.
[27] Spielman RS, Ewens WJ. The TDT and other family-based tests for linkage disequilibrium and association. Am J Hum Genet 1996;59:983–989.
[28] Nobili V, Marcellini M, Devito R, Ciampalini P, Piemonte F, Comparcola D, et al. NAFLD in children: a prospective clinical-pathological study and effect of lifestyle advice. Hepatology 2006;44:458–465.
[29] Valenti L, Fracanzani AL, Dongiovanni P, Santorelli G, Branchi A, Taioli E, et al. Tumor necrosis factor alpha promoter polymorphisms and insulin resistance in nonalcoholic fatty liver disease. Gastroenterology 2002;122:274–280.
[30] Brunt EM, Janney CG, Di Bisceglie AM, Neuschwander-Tetri BA, Bacon BR. Nonalcoholic steatohepatitis: a proposal for grading and staging the histological lesions. Am J Gastroenterol 1999;94:2467–2474.
[31] Daly AK, Fairbrother KS, Andreassen OA, London SJ, Idle JR, Steen VM. Characterization and PCR-based detection of two different hybrid CYP2D7P/CYP2D6 alleles associated with the poor metabolizer phenotype. Pharmacogenetics 1996;6:319–328.
[32] R-Development-Core-Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2010.
[33] Romeo S, Kozlitina J, Xing C, Pertsemlidis A, Cox D, Pennacchio LA, et al. Genetic variation in PNPLA3 confers susceptibility to nonalcoholic fatty liver disease. Nat Genet 2008;40:1461–1465.
[34] Valenti L, Al-Serri A, Daly AK, Galmozzi E, Rametta R, Dongiovanni P, et al. Homozygosity for the patatin-like phospholipase-3/adiponutrin I148M polymorphism influences liver fibrosis in patients with nonalcoholic fatty liver disease. Hepatology 2010;51:1209–1217.
[35] Li YB, Huang TT, Carlson EJ, Melov S, Ursell PC, Olson TL, et al. Dilated cardiomyopathy and neonatal lethality in mutant mice lacking manganese superoxide-dismutase. Nat Genet 1995;11:376–381.
[36] Kokoszka JE, Coskun P, Esposito LA, Wallace DC. Increased mitochondrial oxidative stress in the Sod2 (+/−) mouse results in the age-related decline of mitochondrial function culminating in increased apoptosis. Proc Natl Acad Sci USA 2001;98:2278–2283.
[37] Pan M, Cederbaum AI, Zhang YL, Ginsberg HN, Williams KJ, Fisher EA. Lipid peroxidation and oxidant stress regulate hepatic apolipoprotein B degradation and VLDL production. J Clin Invest 2004;113:1277–1287.
[38] Lenart J, Dombrowski F, Gorlach A, Kietzmann T. Deficiency of manganese superoxide dismutase in hepatocytes disrupts zonated gene expression in mouse liver. Arch Biochem Biophys 2007;462:238–244.
[39] Osterreicher CH, Schultheiss J, Wehler M, Homann N, Hellerbrand C, Kunzli B, et al. Genetic polymorphisms of manganese-superoxide dismutase and glutathione-S-transferase in chronic alcoholic pancreatitis. Mutagenesis 2007;22:305–310.
[40] Petrovic MG, Cilensek I, Petrovic D. Manganese superoxide dismutase gene polymorphism (V16A) is associated with diabetic retinopathy in Slovene (Caucasians) type 2 diabetes patients. Dis Markers 2008;24:59–64.
[41] Tomkins J, Banner SJ, McDermott CJ, Shaw PJ. Mutation screening of manganese superoxide dismutase in amyotrophic lateral sclerosis. Neuro-Report 2001;12:2319–2322.
[42] Reddy JK, Rao MS. Lipid metabolism and liver inflammation. II. Fatty liver disease and fatty acid oxidation. Am J Physiol Gastrointest Liver Physiol 2006;290:G852–858.
[43] Rao MS, Reddy JK. Peroxisomal beta-oxidation and steatohepatitis. Semin Liver Dis 2001;21:43–55.
[44] Pessayre D, Fromenty B. NASH: a mitochondrial disease. J Hepatol 2005;42:928–940.
[45] Pessayre D. Role of mitochondria in non-alcoholic fatty liver disease. J Gastroenterol Hepatol 2007;22:S20–27.
[46] Serviddio G, Bellanti F, Tamborra R, Rollo T, Capitanio N, Romano AD, et al. Uncoupling protein-2 (UCP2) induces mitochondrial proton leak and increases susceptibility of non-alcoholic steatohepatitis (NASH) liver to ischaemia-reperfusion injury. Gut 2008;57:957–965.
[47] Martinelli A, Knapp S, Anstee Q, Worku M, Tommasi A, Zucoloto S, et al. Effect of a thrombin receptor (protease-activated receptor 1, PAR-1) gene polymorphism in chronic hepatitis C liver fibrosis. J Gastroenterol Hepatol 2008;23:1403–1409.
[48] Wright M, Goldin R, Hellier S, Knapp S, Frodsham A, Hennig B, et al. Factor V Leiden polymorphism and the rate of fibrosis development in chronic hepatitis C virus infection. Gut 2003;52:1206–1210.
[49] Miele L, Beale G, Patman G, Nobili V, Leathart J, Grieco A, et al. The Kruppel-like factor 6 genotype is associated with fibrosis in nonalcoholic fatty liver disease. Gastroenterology 2008;135:282–291, e281.
[50] Bataller R, North KE, Brenner DA. Genetic polymorphisms and the progression of liver fibrosis: a critical appraisal. Hepatology 2003;37:493–503.
[51] Nelson JE, Bhattacharya R, Lindor KD, Chalasani N, Raaka S, Heathcote EJ, et al. HFE C282Y mutations are associated with advanced hepatic fibrosis in Caucasians with nonalcoholic steatohepatitis. Hepatology 2007;46:723–729.
[52] Anstee QM, Wright M, Goldin R, Thomas HC, Thursz M. The Factor V Leiden mutation accelerates disease progression in a mouse model of hepatic fibrosis induced by chronic carbon tetrachloride exposure. Hepatology 2004;40:854.

Please cite this article in press as: Al-Serri A et al. The *SOD2* C47T polymorphism influences NAFLD fibrosis severity: Evidence from case-control and intra-familial allele association studies. J Hepatol (2011), doi:10.1016/j.jhep.2011.05.029

JOURNAL OF **HEPATOLOGY**

[53] Yoneda M, Hotta K, Nozaki Y, Endo H, Uchiyama T, Mawatari H, et al. Association between angiotensin II type 1 receptor polymorphisms and the occurrence of nonalcoholic fatty liver disease. Liver Int 2009;29: 1078–1085.
[54] Merat S, Malekzadeh R, Sohrabi MR, Sotoudeh M, Rakhshani N, Sohrabpour AA, et al. Probucol in the treatment of non-alcoholic steatohepatitis: a double-blind randomized controlled study. J Hepatol 2003;38:414–418.
[55] Abdelmalek MF, Angulo P, Jorgensen RA, Sylvestre PB, Lindor KD. Betaine, a promising new agent for patients with nonalcoholic steatohepatitis: results of a pilot study. Am J Gastroenterol 2001;96:2711–2717.
[56] Lavine JE. Vitamin E treatment of nonalcoholic steatohepatitis in children: a pilot study. J Pediatr 2000;136:734–738.
[57] Facchini FS, Hua NW, Stoohs RA. Effect of iron depletion in carbohydrate-intolerant patients with clinical evidence of nonalcoholic fatty liver disease. Gastroenterology 2002;122:931–939.
[58] Harrison SA, Torgerson S, Hayashi P, Ward J, Schenker S. Vitamin E and vitamin C treatment improves fibrosis in patients with nonalcoholic steatohepatitis. Am J Gastroenterol 2003;98:2485–2490.
[59] Dufour JF, Oneta CM, Gonvers JJ, Bihl F, Cerny A, Cereda JM, et al. Randomized placebo-controlled trial of ursodeoxycholic acid with vitamin e in nonalcoholic steatohepatitis. Clin Gastroenterol Hepatol 2006;4:1537–1543.
[60] Sanyal AJ, Chalasani N, Kowdley KV, McCullough A, Diehl AM, Bass NM, et al. Pioglitazone, vitamin E, or placebo for nonalcoholic steatohepatitis. N Engl J Med 2010;362:1675–1685.

Please cite this article in press as: Al-Serri A et al. The *SOD2* C47T polymorphism influences NAFLD fibrosis severity: Evidence from case-control and intra-familial allele association studies. J Hepatol (2011), doi:10.1016/j.jhep.2011.05.029

Research Article

EASL EUROPEAN ASSOCIATION FOR THE STUDY OF THE LIVER | JOURNAL OF HEPATOLOGY

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

Luca Valenti[1,*], Valerio Nobili[2], Ahmad Al-Serri[3], Raffaela Rametta[1], Julian B.S. Leathart[3], Marco A. Zappa[4], Paola Dongiovanni[1], Anna L. Fracanzani[1], Arianna Alterio[2], Giancarlo Roviaro[4], Ann K. Daly[3], Silvia Fargion[1,*], Christopher P. Day[3]

[1]*Department of Internal Medicine, Università degli Studi Milano, UO Medicina Interna 1B, Ospedale Maggiore "Ca' Granda" Fondazione Policlinico IRCCS, Milan, Italy;* [2]*Liver Unit, Bambin Gesù Hospital, Rome, Italy;* [3]*Institute of Cellular Medicine, Newcastle University, Newcastle upon Tyne, UK;* [4]*Department of Surgery, Università degli Studi Milano, Ospedale Maggiore "Ca' Granda" Fondazione Policlinico IRCCS, Milan, Italy*

Background & Aims: The T-455C and C-482T *APOC3* promoter region polymorphisms (SNPs) have recently been reported to predispose to dyslipidemia, insulin resistance, and nonalcoholic fatty liver disease (NAFLD) in Indian subjects, but the association with liver damage has not been evaluated so far. The aim was to assess the association between *APOC3* SNPs and liver damage in Caucasian patients.

Methods: We considered 437 Italian patients with histological diagnosis of NAFLD (including 137 children, 120 morbid obese) and 316 healthy controls, 71 Italian family trios, and 321 patients from the UK. *APOC3* SNPs were determined by sequencing, allele-specific oligonucleotide probes and PCR–restriction fragment length polymorphism analysis, hepatic *APOC3* mRNA levels by real-time PCR.

Results: *APOC3* SNPs were not associated with NAFLD in Italian subjects, although a borderline significance for the transmission of the −455T allele was observed in the family study. Homozygosity for the *APOC3* wild-type genotype (APOC3 WT) was associated with a more favorable lipid profile in control subjects, and consistently with lower hepatic *APOC3* mRNA levels in obese patients without diabetes. However, *APOC3* SNPs, alone or in combination, were not associated with insulin resistance, altered lipid levels, liver enzymes, and with liver damage (severity of steatosis, nonalcoholic steatohepatitis, and moderate/severe fibrosis) in Italian as well as in UK patients, and in the whole cohort. Stratification for the I148M *PNPLA3* mutation, associated with the susceptibility to NASH, did not alter the results.

Conclusions: *APOC3* genotype is not associated with progressive liver damage in Caucasian patients with NAFLD.

Keywords: Apolipoprotein C3; Dyslipidemia; Genetics; Insulin resistance; Nonalcoholic fatty liver disease; Promoter polymorphism; Steatohepatitis; Liver fibrosis.

Received 1 January 2011; received in revised form 4 March 2011; accepted 7 March 2011

* Corresponding authors. Address: Department of Internal Medicine, Università degli Studi Milano, UO Medicina Interna 1B, Fondazione Ospedale Policlinico MaRE IRCCS, Milano, Italy. Tel.: +39 02 55033301; fax: +39 02 50320296.

E-mail addresses: luca.valenti@unimi.it (L. Valenti), silvia.fargion@unimi.it (S. Fargion).

Abbreviations: NAFLD, nonalcoholic fatty liver disease; IR, insulin resistance; NASH, nonalcoholic steatohepatitis; SNP, single nucleotide polymorphism; PNPLA3, patatin-like phospholipase domain containing-3; APOC3, apolipoprotein C3; qRT-PCR, quantitative real time-polymerase chain reaction; WT, wild-type.

Introduction

Nonalcoholic fatty liver disease (NAFLD), affecting 20–34% of the population in Western countries [1,2], is the hepatic manifestation of the metabolic syndrome [3], and is associated with increased mortality due to liver disease [4,5].

Insulin resistance (IR) is a key factor in NAFLD pathophysiology and is deeply entangled with the progression of liver disease. Adipose tissue IR results in an excess of fatty acids supply to the liver [6], which induces lipotoxicity, oxidative stress, and apoptosis thus leading to steatohepatitis (NASH) [7]. NASH is considered the progressive form of NAFLD and is characterized by necroinflammatory changes and hepatocellular damage (ballooning), with potential evolution to fibrosis. IR predicts the severity of liver disease and is directly involved in the pathogenesis of liver damage [4,8,9].

Inherited factors play a major role in the susceptibility to NASH [10–13], and recently the rs738409 C >G single nucleotide polymorphism (SNP) of patatin-like phospholipase domain containing-3 (*PNPLA3*), encoding the I148M protein variant, was identified by a genomewide approach as a strong genetic determinant of liver fat content independently of body mass, dyslipidemia, and insulin resistance [14], and associated with NASH and progressive fibrosis in patients with steatosis [15–17]. In a large series of Caucasian patients with histological evaluation of liver damage, we have previously shown that also some genetic variants influencing insulin signaling and oxidative stress predispose to NASH and fibrosis in patients with steatosis [9,18,19].

Please cite this article in press as: Valenti L et al. 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 (2011), doi:10.1016/j.jhep.2011.03.035

Research Article

Interestingly, Petersen *et al.* have recently reported that the two common Apoliporotein C3 (*APOC3*) T-455C and C-482T promoter SNPs that hamper the regulation of the apolipoprotein C3 expression by insulin signaling via FOXO1 phosphorylation [20], predispose to liver fat accumulation in Indian men by altering lipid metabolism and IR [21]. However, the relationship with altered liver enzymes and liver damage was not assessed.

The aim of this study was to evaluate whether the T-455C and C-482T *APOC3* promoter SNPs influence the severity of liver damage (i.e. the presence of NASH and moderate/severe fibrosis) in Caucasian patients with NAFLD.

Patients and methods

Patients

We analyzed DNA from 758 unrelated patients with biopsy proven NAFLD diagnosed between January 1999 and January 2008. These included 163 Italian patients who underwent liver biopsy because of persistently abnormal liver enzymes/serum ferritin or a long lasting history of steatosis associated with severe metabolic abnormalities (82% overweight, 26% obese), 137 Italian severely obese patients who underwent percutaneous liver biopsy performed during bariatric surgery, 137 Italian children (100% obese after age correction), and 321 adult patients from Newcastle, UK, who were biopsied because of persistently abnormal liver enzymes (98% overweight, 84% obese, and 44% severely obese). Other causes of liver disease were excluded, including increased alcohol intake (>30/20 g/day for M/F), as confirmed by at least one family member or friend and carboxydesialylated transferrin determination, viral and autoimmune hepatitis, hereditary hemochromatosis, and alpha1-antitrypsin deficiency. We included bariatric patients because of the high, but not universal, risk of NAFLD and in particular of progressive NASH in severely obese patients [22], despite their overwhelming increase in body mass. In addition, *PNPLA3* genotype, a major determinant of liver fat in the general population, has been shown to predispose to liver damage in particular in morbidly obese subjects [23]. This makes these subjects an ideal group to test the effect of genetic factors specifically influencing liver steatosis and NASH progression independently of body mass [9,15,24]. Part of this group had previously been described [25]. The available demographic and clinical features are shown in Table 1. All patients included in the study had a fatty liver index >50 [26].

In addition, we tested 71 Italian family trios with both parents alive and an index child with biopsy-proven fibrotic NAFLD. We used transmission disequilibrium testing to look for preferential transmission of risk alleles to the affected children. This approach is not subjected to the potential confounding effects inherent in case–control studies and is significantly more powerful at detecting true associations [9,15,27,28].

Controls

The Italian control group included 316 geographically, age, and sex matched Italian subjects who were selected because of the lack of clinical and biochemical evidence of liver and metabolic disease and no alcohol abuse (<30/20 g/day in males/females). We excluded subjects with ALT >35/30 IU/ml in males/females, GGT >35 IU/ml, BMI >28, abdominal circumference >100 cm, glucose levels $\geqslant$100 mg/dl, triglycerides $\geqslant$150 mg/dl, HDL $\leqslant$45/55 in M/F, or a fatty liver index >35, a value with high specificity to rule out NAFLD in the general population [26].

The study protocol was approved by the Institutional Review Board of the Ospedale Maggiore "Ca' Granda" Policlinico IRCCS, Milan, Italy and by Local Research Ethics Committee in Newcastle. Informed written consent was obtained from each patient and control subject, and the study conforms to the ethical guidelines of the 1975 declaration of Helsinki.

Histological assessment

Tissue sections were stained with hematoxylin and eosin, impregnated with silver for reticulin framework, and stained with trichrome for collagen. One expert pathologist unaware of clinical and genetic data reviewed all biopsies for fibrosis stage at each clinical center. The severity of steatosis, features of NASH and fibrosis were assessed according to Kleiner *et al.* [29] NASH was considered to be present when steatosis, ballooning and lobular inflammation were present. The minimum biopsy size was 1.7 cm and the number of portal areas 10.

Genetic analysis

For the adult Italian samples, DNA was extracted from peripheral blood collected at the time of liver biopsy by the phenol–chloroform method. Success rate in extracting DNA was 100% for each study group. The *APOC3* T-455C (rs2854116) and C-482T (rs2854117) promoter SNPs were assessed by direct sequencing (forward primer: 5′-GAAACCCAGAGATGGAGGTG-3′; reverse primer: 5′-TCTCAGCCTTTCACACTGGA-3′) by personnel unaware of patients and controls clinical status. For the Italian families and UK cases, DNA was prepared from

Table 1. Demographic and clinical features of patients with NAFLD and controls.

	Italian controls	NAFLD patients			
		Italian adults	UK adults	Italian children	Italian obese
Number	316	180	321	137	120
Sex F	54 (18)	34 (19)	123 (38)	51 (37)	93 (78)
Age years	48 ± 12	48 ± 12	50 ± 13	10 ± 3	43 ± 11
BMI Kg/m^2	25.2 ± 3	27.1 ± 4	34.1 ± 5	25.7 ± 5	41.5 ± 8
LDL cholesterol mg/dl	117 ± 30	135 ± 45	123 ± 44	NA	131 ± 39
HDL cholesterol mg/dl	56 ± 13	44 ± 12	47 ± 13	NA	54 ± 14
Triglycerides mg/dl	90 ± 45	155 ± 86	245 ± 174	113 ± 69	129 ± 61
Glucose mg/dl	89 ± 10	98 ± 26	121 ± 59	82 ± 13	104 ± 33
Diabetes/IGT	0	43 (24)	117 (36)	68 (50)	33 (28)
HOMA-IR	2.9 ± 1.5	4.2 ± 3.2	6.8 ± 6.6	2.7 ± 1.4	NA
ALT UI/ml	23 ± 8	63 ± 43	81 ± 64	89 ± 61	30 ± 33
GGT UI/ml	23 ± 16	89 ± 98	128 ± 183	32 ± 19	26 ± 22
NASH	-	84 (47)	180 (56)	103 (75)	65 (54)
Fibrosis stage F2-F4	-	36 (20)	140 (44)	15 (11)	16 (13)

( ): % values, Na: not available, IGT: impaired glucose tolerance, NASH: nonalcoholic steatohepatitis.

Please cite this article in press as: Valenti L et al. 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 (2011), doi:10.1016/j.jhep.2011.03.035

JOURNAL OF HEPATOLOGY

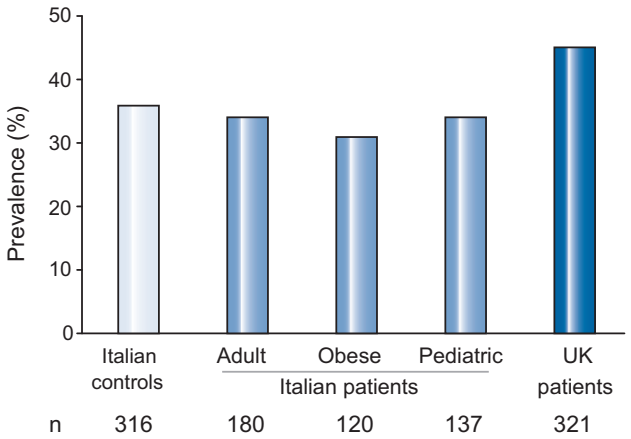

Fig. 1. Prevalence of homozygosity for the wild-type allele of the *APOC3* C-428T and T-455C SNPs (APOC3 WT) in 437 Italian patients with NAFLD and 316 Italian controls (p = n.s.). The prevalence of *APOC3* WT in 321 UK patients is also shown.

peripheral blood lymphocytes as described previously [30]. Genotyping used either the CVD35 multi-locus genotyping assay (Roche Molecular Diagnostics) [31] or by restriction fragment length polymorphism–PCR [32].

Valid genotypic data were obtained for over 99% of subjects analyzed.

PNPLA3 genotype, determined as previously described [15], was available in 731/758 patients (96%).

Gene expression analysis

In 60 patients submitted to bariatric surgery for whom adequate samples were available, RNA was isolated by the Trizol reagent (Lifetech, Carlsbad, CA), digested with DNaseI, and quality evaluated by measuring the 260/280 nm absorbance ratio ($\geqslant$1.8) and by electrophoresis. First-strand cDNA was synthesized using equal amounts (0.5 μg) of total RNA, with the SuperScript VILO cDNA synthesis kit (Invitrogen, Carlsbad, CA). The *APOC3* mRNA levels were analyzed by qRT-PCR with SYBR Green chemistry (Fluocycle II SYBR green mix, Euroclone, Pero, Italy). All the reactions were performed in triplicate with the ABI 7500fast quantitative real time (qRT)-PCR system (Lifetech, Carlsbad, CA). Primers are available upon request. Results were normalized for *β-actin* and *18S*, which were chosen as a control because of stable expression among different samples.

Statistical analysis

Results are expressed as means ± standard deviation and considered significant when p <0.05 (two-tailed). Transmission Disequilibrium Tests (TDT) were performed using the TDTHAP package (http://www-gene.cimr.cam.ac.uk/clayton/software) running in the R software environment (R version 2.7.1). The R package was used to determine preferential transmission of alleles and statistical significance. Mean values were compared by *t*-test and frequencies by Fisher's exact test, and gene expression was correlated by Spearman's rho test. The association between the *APOC3* SNPs and the presence of NASH and moderate/severe fibrosis (stage >1) was evaluated by logistic regression analysis adjusted for confounders (age, BMI, ALT, glucose and ferritin levels, considered as continuous variables). Given a prevalence of homozygosity for the wild-type alleles of 35%, a prevalence of NASH of 50% and of fibrosis >1 of 30%, our sample had a 99% and 95% power of detecting an OR of 1.5 for NASH and fibrosis >1, respectively, with a significance of 5%. Analyses were carried out with JMP 6.0 statistical analysis software (SAS Institute Inc, Cary, NC).

Results

The frequency distribution of the *APOC3* T-455C and C-482T SNPs was in Hardy–Weinberg equilibrium in all groups tested. The prevalence of homozygosity for the wild-type alleles at both loci (henceforth *APOC3* WT), previously associated with lower ApoC3 levels, faster lipid clearance, insulin sensitivity, and protection from steatosis in Indians [21], was not significantly different between Italian patients and controls (Fig. 1). We also did not detect any significant difference in the frequency distribution of the two SNPs analyzed separately between Italian patients and controls (not shown in details). We did not find any protective effect of *APOC3* WT on the risk of NAFLD either in men or women, or in subjects who were not overweight compared with those with BMI >25 (p >0.3 for all comparisons).

In contrast, the prevalence of homozygosity for the *PNPLA3* 148M allele, previously associated with NASH [14,15], was significantly higher in Italian patients than in controls (59/437, 13.5% *vs.* 12/316, 3.8%; p <0.0001). Stratification for *PNPLA3* genotype did affect the lack of significant association between APOC3 WT and the presence of NAFLD in Italian subjects (not shown).

Table 2. Effect of APOC3 WT on clinical and metabolic parameters in 316 Italian healthy controls, 437 Italian, and 321 UK patients with NAFLD.

	Italian controls			Italian patients			UK patients		
	APOC3		p	*APOC3*		p	*APOC3*		p
	WT (n = 114)	POL (n = 202)		WT (n = 145)	POL (n = 292)		WT (n = 127)	POL (n = 194)	
Age years	48 ± 13	48 ± 13	0.78	34 ± 19	35 ± 19	0.51	51 ± 13	49 ± 13	0.17
Gender F	13 (11)	41 (20)	0.05	57 (39)	121 (41)	0.69	77 (40)	46 (36)	0.53
BMI Kg/m^2	25.1 ± 3	25.2 ± 3	0.75	30.8 ± 8	30.8 ± 9	0.98	34.4 ± 6	33.8 ± 5	0.36
LDL mg/dl	115 ± 30	123 ± 23	0.03	138 ± 46	131 ± 40	0.20	123 ± 43	123 ± 45	1
HDL mg/dl	58 ± 14	55 ± 12	0.03	48 ± 14	47 ± 13	0.32	46 ± 14	48 ± 13	0.30
Triglycerides mg/dl	88 ± 44	94 ± 50	0.36	135 ± 71	135 ± 80	0.98	268 ± 212	229 ± 142	0.07
Glucose mg/dl	88 ± 9	89 ± 11	0.43	96 ± 22	95 ± 30	0.60	125 ± 59	118 ± 59	0.30
HOMA-IR	2.0 ± 1.8	1.8 ± 0.9	0.20	2.8 ± 1	3.0 ± 2	0.18	2.8 ± 1	3.0 ± 2	0.18
Diabetes	0	0	1	51 (35)	93 (32)	0.32	54 (44)	63 (33)	0.056
Hypertension	0	0	1	47 (32)	94 (32)	0.96	47 (32)	94 (32)	0.96
ALT IU/ml	24 ± 8	22 ± 8	0.28	68 ± 58	60 ± 51	0.21	75 ± 58	85 ± 68	0.15
GGT IU/ml	23 ± 16	24 ± 15	0.80	66 ± 73	59 ± 79	0.49	139 ± 191	122 ± 178	0.43

APOC3 WT/POL: absence or presence of polymorphic allele at position −455 and −482 of the *APOC3* gene, respectively.

Please cite this article in press as: Valenti L et al. 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 (2011), doi:10.1016/j.jhep.2011.03.035

Research Article

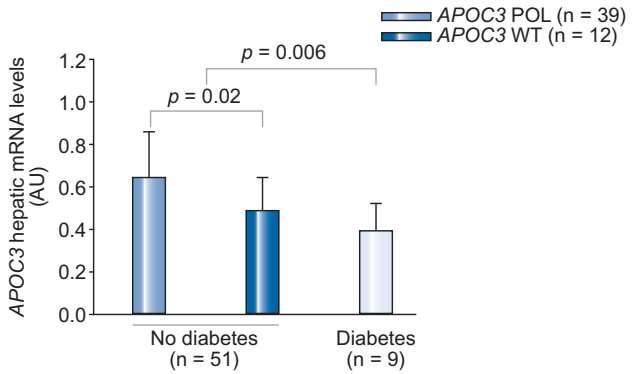

Fig. 2. Effect of *APOC3* WT genotype on APOC3 expression (hepatic mRNA levels) in Italian obese patients with NAFLD, but without diabetes. *APOC3* WT/POL: absence or presence of polymorphic allele at position −455 and −482 of the *APOC3* gene, respectively.

Interestingly, the prevalence of the putative protective *APOC3* WT genotype was higher in UK than in Italian patients (Fig. 1), despite higher severity of liver disease in UK patients (Table 1).

In the family study conducted in the 71 Italian trios, the −455 wild type T allele, previously reported to protect from NAFLD [21], showed borderline significance at the TDT with the transmission of NAFLD (informative families: 66; transmitted in 43/66 cases, 65%; $p = 0.02$), whereas the wild-type −482 T allele was not significantly associated with NAFLD transmission (informative families: 61; transmitted in 25/61 cases, 41%; $p = 0.2$). In the same family study, *PNPLA3* genotype has been previously shown to predispose to NAFLD [15].

We next evaluated whether the *APOC3* WT genotype influenced insulin resistance (Table 2). We did not find any significant association between *APOC3* genotype and metabolic data in Italian patients, whereas *APOC3* WT was not significantly associated with higher triglycerides and prevalence of diabetes in UK patients. In control subjects, *APOC3* WT was associated with higher HDL (58 ± 14 *vs.* 55 ± 12; $p = 0.03$) and lower LDL cholesterol levels (115 ± 30 *vs.* 123 ± 33; $p = 0.03$). No significant association was found between the two single SNPs and the metabolic and clinical parameters were evaluated.

In obese patients with NAFLD, hepatic *APOC3* mRNA levels were not significantly associated with age, sex, BMI, ALT, GGT, and triglyceride levels, the presence of severe steatosis, NASH, and fibrosis >1, but were significantly correlated with total cholesterol and HDL levels (rho = 0.34, $p = 0.019$ and rho = 0.30, $p = 0.034$, respectively). In addition, hepatic *APOC3* mRNA levels were significantly downregulated in patients with type 2 diabetes (0.39 ± 0.14, n = 9 *vs.* 0.58 ± 0.25, n = 51; $p = 0.006$; Fig. 2). The *APOC3* WT genotype was significantly associated with reduced hepatic *APOC3* mRNA levels in patients without type 2 diabetes (0.48 ± 0.16, n = 12 *vs.* 0.64 ± 0.27, n = 39; $p = 0.02$; Fig. 2).

The main aim of the study was to assess whether the T-455C and C-482T promoter SNPs influence the severity of histologically determined liver damage in patients with NAFLD. There was no significant association between *APOC3* WT genotype and steatosis severity, the presence of NASH, and the presence and severity of liver fibrosis in Italian and UK patients (Table 3, Fig. 3, and not shown). The associations tested were not significant in any subgroup of patients evaluated (adult patients, severely obese, and children). In contrast, *PNPLA3* genotype was strongly associated with NASH and the severity of liver damage (stage >1) in the

Table 3. Prevalence of steatosis >1, NASH, and fibrosis stage >1 in Italian and UK patients with NAFLD according to the *APOC3* and *PNPLA3* genotypes.

		Italian patients			UK patients			Overall series		
		APOC3		*p*	*APOC3*		*p*	*APOC3*		*p*
Overall series (n = 758)		WT (n = 145)	POL (n = 292)		WT (n = 127)	POL (n = 194)		WT (n = 272)	POL (n = 486)	
	Steatosis >1	69 (48)	137 (47)	0.89	NA	NA	NA	NA	NA	NA
	NASH	60 (41)	121 (41)	0.99	65 (52)	115 (60)	0.17	125 (46)	236 (49)	0.53
	Fibrosis >1	21 (14)	46 (16)	0.72	53 (42)	87 (45)	0.64	74 (27)	133 (27)	0.96
PNPLA3 148 I/I (n = 313)		WT (n = 60)	POL (n = 114)		WT (n = 59)	POL (n = 80)		WT (n = 119)	POL (n = 194)	
	Steatosis >1	21 (35)	33 (29)	0.49	NA	NA	NA	NA	NA	NA
	NASH	16 (27)	30 (26)	0.99	23 (40)	42 (53)	0.12	39 (33)	72 (37)	0.47
	Fibrosis >1	8 (13)	8 (7)	0.18	20 (34)	28 (35)	1	28 (24)	36 (19)	0.31
PNPLA3 148 I/M (n =323)		WT (n = 57)	POL (n = 133)		WT (n = 58)	POL (n = 75)		WT (n = 115)	POL (n = 208)	
	Steatosis >1	32 (56)	78 (59)	0.75	NA	NA	NA	NA	NA	NA
	NASH	25 (45)	58 (44)	0.97	34 (61)	44 (60)	0.96	59 (52)	102 (50)	0.72
	Fibrosis >1	7 (12)	27 (20)	0.17	27 (47)	37 (49)	0.86	34 (30)	64 (31)	0.82
PNPLA3 148 M/M (n = 94)		WT (n = 25)	POL (n = 33)		WT (n = 8)	POL (n = 28)		WT (n = 277)	POL (n = 498)	
	Steatosis >1	15 (60)	25 (72)	0.39	NA	NA	NA	NA	NA	NA
	NASH	19 (76)	28 (85)	0.50	6 (75)	22 (78)	0.83	25 (76)	50 (82)	0.59
	Fibrosis >1	6 (24)	8 (24)	1	4 (50)	18 (64)	0.46	10 (30)	26 (43)	0.27

( ): % values, Na: not available, *APOC3* WT/POL: absence or presence of polymorphic allele at position −455 and −482 of the *APOC3* gene, respectively. The presence of NASH could not be assessed in six UK patients with advanced fibrosis (burn-out NASH). *PNPLA3* genotype was available in 731/758 patients (96%). $p \leqslant 0.001$ for the prevalence of NASH and fibrosis >1 across *PNPLA3* genotypes in both Italian and UK patients.

Please cite this article in press as: Valenti L et al. 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 (2011), doi:10.1016/j.jhep.2011.03.035

JOURNAL OF HEPATOLOGY

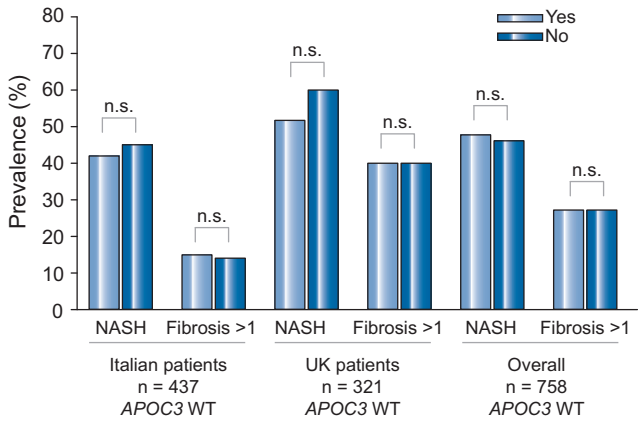

Fig. 3. Effect of homozygosity for the wild-type allele of the *APOC3* C-428T and T-455C SNPs (APOC3 WT) on histologically evaluated liver damage in 775 Italian and UK patients with NAFLD. Yes: *APOC3* WT genotype, NO: presence of *APOC3* SNPs.

same patients (Table 3; p <0.0001 and p = 0.001, respectively). In particular, *PNPLA3* genotype was associated with NASH and fibrosis >1 in Italian patients (p = 0.05 and p = 0.02, respectively), UK patients (p = 0.003 and p = 0.01, respectively), Italian children (p <0.0001 and p = n.s., likely due to their young age [16], respectively), and morbid obese patients (p = 0.01 and p = 0.1, respectively). Furthermore, no significant association was found between the two single *APOC3* SNPs evaluated and the histological parameters evaluated in both the populations and the subgroups studied.

Stratification for the *PNPLA3* genotype did not reveal any association between *APOC3* genotype and liver damage in any of the categories evaluated in both Italian and UK patients (Table 3). Since the presence of diabetes was associated with reduced *APOC3* expression independently of *APOC3* WT genotype, we also explored whether the presence of diabetes might influence the association between *APOC3* WT and liver damage. Although the results of this analysis should be interpreted with caution, we only observed a nominally significant lower prevalence of fibrosis >1 in *PNPLA3* 148M/M patients with *APOC3* WT genotype than in those carrying *APOC3* promoter SNPs (3/22, 14% *vs.* 16/44, 36%; p = 0.05).

Discussion

In this study, we evaluated whether *APOC3* promoter region SNPs previously reported to influence lipid metabolism, insulin resistance, and NAFLD susceptibility [21], affect the histological severity of liver damage in a large series of Caucasian patients with biopsy-proven NAFLD. Our data indicate that *APOC3* SNPs do not influence the severity of liver damage in NAFLD patients of European descent.

Furthermore, we did not detect any association between *APOC3* SNPs and the presence of NAFLD in Italian subjects, whereas *PNPLA3* genotype, a strong genetic determinant of steatosis [14,15], was strongly associated with NAFLD in the same case–control series. Although a borderline significance for the transmission of the −455T allele was observed in the family study, since the same allele was previously reported to protect from steatosis [21], and the effect was not consistent for the two genetic variants studied, we believe that this formal association is likely due to chance. These results are in line with very recent findings obtained in the US general population, where *APOC3* genotype was not found associated with liver fat content or with insulin resistance (HOMA-IR) [33].

APOC3 SNPs were also not associated with higher insulin resistance and dyslipidemia in Italian patients with NAFLD, although an association was detected between the "protective" *APOC3* WT genotype, previously associated with lower APOC3 levels and faster lipid clearance [21], and less atherogenic lipid profile (lower LDL and higher HDL cholesterol) in Italian control subjects without steatosis, as well as in the US general population [33]. This phenotype is consistent with the protective effect of null *APOC3* mutations on atherosclerosis due to lower LDL and triglycerides and higher HDL levels [34]. However, possibly due to the confounding effect of unhealthy lifestyle and other genetic factors, the association was lost in patients with NAFLD. Indeed, it has been reported that *APOC3* genotype interacts with dietary factors in the regulation of lipid metabolism [35]. Thus, further studies are required to evaluate the relevance of the association between APOC3 promoter SNPs, dyslipidemia, and atherogenesis.

Importantly, although the results of this exploratory analysis should be interpreted with caution, we were able to confirm in a subset of obese patients with NAFLD included in this study that the *APOC3* WT genotype influenced APOC3 hepatic expression, associated with cholesterol levels, suggesting that the lack of association between *APOC3* WT and the metabolic phenotype was not explained by a lack of biological activity of the genetic variants under study. Interestingly, the effect was only evident in patients without diabetes, which *per se* was associated with a strong downregulation of APOC3, possibly explained by hyperinsulinemia [20] and/or pharmacological therapy with insulin, insulin sensitizing, and lipid lowering drugs.

The main result of this study is that *APOC3* genotype was not associated with liver enzymes, and with the histological severity of liver damage (severity of steatosis, presence of NASH, and of fibrosis of moderate/severe degree) in Italian as well as in UK patients, and in the whole cohort. We also did not detect any significant influence on the specific components of the NAFLD activity score [29] and fibrosis stage evaluated as ordinal variables (not shown in details). Stratification for the I148M *PNPLA3* mutation, associated with the susceptibility to NASH [15–17,36], did not reveal any *PNPLA3* genotype-specific association between *APOC3* genotype and liver damage, and NAFLD susceptibility as well.

A major strength of this study is the large number of patients evaluated with available liver histology, which still represents the gold standard for evaluation of liver damage, the complete clinical characterization, the evaluation of the biological activity of *APOC3* genotype on gene expression in a subset of patients, and the availability of *PNPLA3* genotype as a positive control and for the stratification of patients. The study was well powered to detect an effect of *APOC3* WT genotype on the risk of NASH and moderate/severe fibrosis. To our knowledge, this is the first report of an evaluation of the association between *APOC3* genotype and histologically determined liver damage.

Limitations of this study include the lack of evaluation of *APOC3* SNP prevalence in UK control subjects. In addition, we could not assess the effect of *APOC3* genotype on ApoC3 serum levels, and we also cannot exclude that *APOC3* genotype influences liver damage in Indian (Asian) patients with NAFLD, in whom it was reported as a strong effect on lipid metabolism and insulin resistance [21], as the effect of the genetic variant may be ethnic specific requiring additional genetic factors for phenotypic expression. Furthermore,

Please cite this article in press as: Valenti L et al. 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 (2011), doi:10.1016/j.jhep.2011.03.035

our data suggest that diabetes or insulin sensitizing/lipid lowering drugs may affect the expression of *APOC3* genotype in European patients with metabolic diseases.

In conclusion, our data suggest that *APOC3* SNPs do not predispose to progressive liver damage in Caucasian patients with NAFLD.

Conflict of interest

The authors who have taken part in this study declared that they do not have anything to disclose regarding funding or conflict of interest with respect to this manuscript.

Financial support

The work was supported by the following grants: FIRST Università di Milano 2007, 2008 (L.V., S.F.); Ricerca corrente Ospedale Maggiore Policlinico 2006 and 2008 (L.V., S.F.); and Centro per lo Studio delle Malattie del Fegato e del Metabolismo.

Acknowledgments

We thank Enrico Galmozzi and Dario Tavazzi for technical assistance, Erika Fatta, Cristina Bertelli, Paolo Maggioni, Daniela Bignamini, and Rosa Lombardi for clinical assistance.

References

[1] Browning JD, Szczepaniak LS, Dobbins R, Nuremberg P, Horton JD, Cohen JC, et al. Prevalence of hepatic steatosis in an urban population in the United States: impact of ethnicity. Hepatology 2004;40:1387–1395.
[2] Bellentani S, Saccoccio G, Masutti F, Croce LS, Brandi G, Sasso F, et al. Prevalence of and risk factors for hepatic steatosis in Northern Italy. Ann Intern Med 2000;132:112–117.
[3] Marchesini G, Brizi M, Bianchi G, Tomassetti S, Bugianesi E, Lenzi M, et al. Nonalcoholic fatty liver disease: a feature of the metabolic syndrome. Diabetes 2001;50:1844–1850.
[4] Ekstedt M, Franzen LE, Mathiesen UL, Thorelius L, Holmqvist M, Bodemar G, et al. Long-term follow-up of patients with NAFLD and elevated liver enzymes. Hepatology 2006;44:865–873.
[5] Bugianesi E, Leone N, Vanni E, Marchesini G, Brunello F, Carucci P, et al. Expanding the natural history of nonalcoholic steatohepatitis: from cryptogenic cirrhosis to hepatocellular carcinoma. Gastroenterology 2002;123:134–140.
[6] Gastaldelli A, Harrison SA, Belfort-Aguilar R, Hardies LJ, Balas B, Schenker S, et al. Importance of changes in adipose tissue insulin resistance to histological response during thiazolidinedione treatment of patients with nonalcoholic steatohepatitis. Hepatology 2009;50:1087–1093.
[7] Day CP. From fat to inflammation. Gastroenterology 2006;130:207–210.
[8] Marchesini G, Bugianesi E, Forlani G, Cerrelli F, Lenzi M, Manini R, et al. Nonalcoholic fatty liver, steatohepatitis, and the metabolic syndrome. Hepatology 2003;37:917–923.
[9] Dongiovanni P, Valenti L, Rametta R, Daly AK, Nobili V, Mozzi E, et al. Genetic variants regulating insulin receptor signaling are associated with the severity of liver damage in patients with nonalcoholic fatty liver disease. Gut 2010;59:267–273.
[10] Wilfred de Alwis NM, Day CP. Genetics of alcoholic liver disease and nonalcoholic fatty liver disease. Semin Liver Dis 2007;27:44–54.
[11] Schwimmer JB, Celedon MA, Lavine JE, Salem R, Campbell N, Schork NJ, et al. Heritability of nonalcoholic fatty liver disease. Gastroenterology 2009;136:1585–1592.
[12] Guerrero R, Vega GL, Grundy SM, Browning JD. Ethnic differences in hepatic steatosis: an insulin resistance paradox? Hepatology 2009;49:791–801.
[13] Makkonen J, Pietilainen KH, Rissanen A, Kaprio J, Yki-Jarvinen H. Genetic factors contribute to variation in serum alanine aminotransferase activity independent of obesity and alcohol: a study in monozygotic and dizygotic twins. J Hepatol 2009;50:1035–1042.
[14] Romeo S, Kozlitina J, Xing C, Pertsemlidis A, Cox D, Pennacchio LA, et al. Genetic variation in PNPLA3 confers susceptibility to nonalcoholic fatty liver disease. Nat Genet 2008;40:1461–1465.
[15] Valenti L, Al-Serri A, Daly AK, Galmozzi E, Rametta R, Dongiovanni P, et al. Homozygosity for the PNPLA3/adiponutrin I148M polymorphism influences liver fibrosis in patients with nonalcoholic fatty liver disease. Hepatology 2010;51:1209–1217.
[16] Valenti L, Alisi A, Galmozzi E, Bartuli A, Del Menico B, Alterio A, et al. I148M patatin-like phospholipase domain-containing 3 gene variant and severity of pediatric nonalcoholic fatty liver disease. Hepatology 2010;52:1274–1280.
[17] Rotman Y, Koh C, Zmuda JM, Kleiner DE, Liang TJ. The association of genetic variability in patatin-like phospholipase domain-containing protein 3 (PNPLA3) with histological severity of nonalcoholic fatty liver disease. Hepatology 2010;52:894–903.
[18] Valenti L, Canavesi E, Galmozzi E, Dongiovanni P, Rametta R, Maggioni P, et al. Beta-globin mutations are associated with parenchymal siderosis and fibrosis in patients with non-alcoholic fatty liver disease. J Hepatol 2010;53:927–933.
[19] Dongiovanni P, Rametta R, Fracanzani AL, Benedan L, Borroni V, Maggioni P, et al. 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 2010;10:102.
[20] Altomonte J, Cong L, Harbaran S, Richter A, Xu J, Meseck M, et al. Foxo1 mediates insulin action on apoC-III and triglyceride metabolism. J Clin Invest 2004;114:1493–1503.
[21] Petersen KF, Dufour S, Hariri A, Nelson-Williams C, Foo JN, Zhang XM, et al. Apolipoprotein C3 gene variants in nonalcoholic fatty liver disease. N Engl J Med 2010;362:1082–1089.
[22] Mathurin P, Hollebecque A, Arnalsteen L, Buob D, Leteurtre E, Caiazzo R, et al. Prospective study of the long-term effects of bariatric surgery on liver injury in patients without advanced disease. Gastroenterology 2009;137:532–540.
[23] Romeo S, Sentinelli F, Dash S, Yeo GS, Savage DB, Leonetti F, et al. Morbid obesity exposes the association between PNPLA3 I148M (rs738409) and indices of hepatic injury in individuals of European descent. Int J Obes (Lond) 2010;34:190–194.
[24] Valenti L. HFE genotype, parenchymal iron accumulation, and liver fibrosis in patients with nonalcoholic fatty liver disease. Gastroenterology 2010;138:905–912.
[25] Valenti L, Fracanzani AL, Dongiovanni P, Santorelli G, Branchi A, Taioli E, et al. Tumor necrosis factor alpha promoter polymorphisms and insulin resistance in nonalcoholic fatty liver disease. Gastroenterology 2002;122:274–280.
[26] Bedogni G, Bellentani S, Miglioli L, Masutti F, Passalacqua M, Castiglione A, et al. The Fatty Liver Index: a simple and accurate predictor of hepatic steatosis in the general population. BMC Gastroenterol 2006;6:33.
[27] Miele L, Beale G, Patman G, Nobili V, Leathart J, Grieco A, et al. The Kruppel-like factor 6 genotype is associated with fibrosis in nonalcoholic fatty liver disease. Gastroenterology 2008;135:282–291, e281.
[28] Spielman RS, McGinnis RE, Ewens WJ. Transmission test for linkage disequilibrium: the insulin gene region and insulin-dependent diabetes mellitus (IDDM). Am J Hum Genet 1993;52:506–516.
[29] Kleiner DE, Brunt EM, Van Natta M, Behling C, Contos MJ, Cummings OW, et al. Design and validation of a histological scoring system for nonalcoholic fatty liver disease. Hepatology 2005;41:1313–1321.
[30] Daly AK, King BP, Leathart JB. Genotyping for cytochrome P450 polymorphisms. Methods Mol Biol 2006;320:193–207.
[31] Cheng S, Pallaud C, Grow MA, Scharf SJ, Erlich HA, Klitz W, et al. A multilocus genotyping assay for cardiovascular disease. Clin Chem Lab Med 1998;36:561–566.
[32] Hegele RA, Connelly PW, Hanley AJ, Sun F, Harris SB, Zinman B. Common genomic variation in the APOC3 promoter associated with variation in plasma lipoproteins. Arterioscler Thromb Vasc Biol 1997;17:2753–2758.
[33] Kozlitina J, Boerwinkle E, Cohen JC, Hobbs HH. Dissociation between APOC3 variants, hepatic triglyceride content and insulin resistance. Hepatology 2011;53:467–474.
[34] Pollin TI, Damcott CM, Shen H, Ott SH, Shelton J, Horenstein RB, et al. A null mutation in human APOC3 confers a favorable plasma lipid profile and apparent cardioprotection. Science 2008;322:1702–1705.
[35] Olivieri O, Martinelli N, Sandri M, Bassi A, Guarini P, Trabetti E, et al. Apolipoprotein C-III, n-3 polyunsaturated fatty acids, and "insulin-resistant" T-455C APOC3 gene polymorphism in heart disease patients: example of gene–diet interaction. Clin Chem 2005;51:360–367.
[36] Speliotes EK, Butler JL, Palmer CD, Voight BF, Hirschhorn JN. PNPLA3 variants specifically confer increased risk for histologic nonalcoholic fatty liver disease but not metabolic disease. Hepatology 2010;52:904–912.

Please cite this article in press as: Valenti L et al. 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 (2011), doi:10.1016/j.jhep.2011.03.035