THE ROLE OF SINGLE NUCLEOTIDE POLYMORPHISMS AND COPY NUMBER VARIATIONS IN THE SPECTRUM OF NON-ALCOHOLIC FATTY LIVER DISEASE

SHAMSUL BIN MOHD ZAIN

**THESIS SUBMITTED IN FULFILMENT
OF THE REQUIREMENTS
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY**

**FACULTY OF MEDICINE
UNIVERSITY OF MALAYA
KUALA LUMPUR**

2013

UNIVERSITI MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: (I.C/Passport No:)

Registration/Matric No:

Name of Degree:

Title of Project Paper/Research Report/Dissertation/Thesis ("this Work"):

Field of Study:

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature Date

Subscribed and solemnly declared before,

Witness's Signature Date

Name:
Designation:

I dedicate this thesis to my mother, who has taught me that diligence and perseverance are the keys to success.

Also to my fiancee, for being the one thing
that kept me sane for the last couple of months.

ABSTRACT

Non-alcoholic fatty liver disease (NAFLD) has been recognised as a growing public health problem worldwide. NAFLD is a spectrum ranging from simple steatosis to non-alcoholic steatohepatitis (NASH) and cirrhosis. Simple steatosis is generally benign condition whereas NASH can potentially progress to serious liver complications, including liver cancer. Genetic factors have been shown to play a significant role in the pathogenesis of NAFLD and may be responsible for the different phenotypes among individuals, including predisposition to severe liver disease. Genomic variability may present in various forms such as single nucleotide polymorphisms (SNPs), structural alterations (deletion, duplication and inversion), variable number of tandem repeats (VNTRs) and presence or absence of transportable element. Of all these variations, the commonest is the SNP. Copy number variation (CNV) is another type of mutation that could serve as a genetic marker for the NAFLD spectrum, and to date, there is no published report on CNV in NAFLD. This study aimed to investigate the association between SNPs of candidate genes with susceptibilty to NAFLD and the severity of the underlying liver disease. This study further identified genomic amplifications and deletions (CNVs) in the spectrum of NAFLD. A total of 144 biopsy-proven NAFLD patients and 198 controls without NAFLD were genotyped for polymorphisms of various candidate genes. Whole-genome array comparative genomic hybridization (aCGH) method was used to detect CNVs in a total of 40 patients with NASH and 40 age, gender, and ethnicity-matched controls. The polymorphisms of *PNPLA3*, *LEPR*, and *GCKR* render susceptibility to NAFLD (OR 2.23, 95% CI 1.60-3.11, $P < 0.0001$; OR 1.64, 95% CI 1.18-2.28, $P = 0.003$; and OR 1.49, 95% CI 1.09-2.05, $P = 0.012$, respectively) while SNPs of *AGTR1* render protection against the condition (OR 0.40, 95% CI 0.20-0.81, $P = 0.01$). The G allele of both the *PNPLA3* rs738409 and the *AGTR1* rs3772622 were associated with increased fibrosis score ($P < 0.05$). For *LEPR*

rs1137100, the G allele was associated with lower fibrosis score ($P < 0.05$). Analysis of gene-gene interaction revealed a strong interaction between the *AGTR1*, *LEPR* and *GCKR* genes with the *PNPLA3* gene (empirical $P = 0.007$, empirical $P = 0.001$ and empirical $P = 0.003$, respectively). Genome wide profiling identified five most frequently amplified CNVs located on the chromosome 14q11.2 (62.5%), 12p13.31 (50.0%), 11p15.4 (45.0%), 5p15.33 (42.5%), and 12p13.2 (40.0%). The most frequently deleted CNVs were located on chromosome 11q11 (35.0%), 12p13.2 (32.5%), 16q12.2 (32.5%), 4q13.2 (32.5%), 1q21.3 (32.5%), and 14q24.3 (32.5%). It can be hypothesised that these CNVs may be implicated in mechanisms that lead to progression to NASH. Of particular note was a 0.02 Mb deletion on chromosome 16q12.2. This region is rich in carboxylesterase 1 gene (*CES1*) that is important in cholesterol transportation leading to fatty liver. Downstream studies to validate findings are being pursued. Overall, this study showed the complex nature of NAFLD involving multiple gene interactions. The copy number findings are novel and could serve as potential genetic marker for the identification of the potentially progressive form of NAFLD.

ABSTRAK

Penyakit hati berlemak tanpa kehadiran alkohol (NAFLD) telah diiktiraf sebagai satu masalah kesihatan yang semakin meningkat di seluruh dunia. NAFLD merupakan satu spektrum yang bermula daripada steatosis biasa kepada steatohepatitis tanpa kehadiran alkohol (NASH) dan sirosis. Steatosis biasa, umumnya adalah keadaan benigna manakala NASH berpotensi berkembang menjadi komplikasi hati yang serius, termasuk kanser hati. Faktor genetik telah didapati memainkan peranan penting dalam patogenesis NAFLD dan mungkin bertanggungjawab terhadap perbezaan fenotaip-fenotaip di kalangan individu, termasuk kecenderungan terhadap penyakit hati yang lebih serius. Kebolehubahan genom boleh hadir dalam pelbagai bentuk seperti polimorfisme-polimorfisme nukleotid tunggal (SNPs), pengubahan struktur (penghapusan, salinan dan penyongsangan), pelbagai nombor pengulangan seiring (VNTRs) dan kehadiran atau ketiadaan elemen boleh angkut. Daripada semua variasi ini, yang paling lazim adalah SNP. Salinan variasi nombor (CNV) adalah satu lagi jenis mutasi yang boleh berperanan sebagai penanda genetik untuk spektrum NAFLD, dan sehingga kini, masih lagi tiada laporan penerbitan CNV untuk NAFLD. Tujuan kajian ini adalah untuk menyiasat perkaitan antara SNP calon gen-gen dengan kecenderungan terhadap NAFLD dan juga terhadap keterukan penyakit hati. Kajian ini seterusnya mengenalpasti amplifikasi dan penghapusan genom di dalam spektrum NAFLD. Sejumlah 144 pesakit NAFLD yang disahkan melalui biopsi dan 198 subjek kawalan tanpa NAFLD digenotaip untuk polimorfisme-polimorfisme pelbagai calon gen. Kaedah seluruh genom susunan perbandingan genomik penghibridan (aCGH) digunakan untuk mengesan CNV pada 40 pesakit NASH dan 40 subjek kawalan yang dipadankan dari segi umur, jantina dan bangsa. Polimorfisme-polimorfisme *PNPLA3*, *LEPR* dan *GCKR* menyebabkan kecenderungan terhadap NAFLD (OR 2.23, 95% CI 1.60-3.11, $P < 0.0001$; OR 1.64, 95% CI 1.18-2.28, $P = 0.003$; and OR 1.49, 95% CI 1.09-2.05, $P =$

0.012, masing-masing) manakala SNPs bagi *AGTR1* menyebabkan perlindungan daripada keadaan ini (OR 0.40, 95% CI 0.20-0.81, $P = 0.01$). Alel G untuk kedua-dua *PNPLA3* rs738409 dan *AGTR1* rs3772622 dikaitkan dengan kenaikan skor fibrosis ($P < 0.05$). Bagi *LEPR* rs1137100, alel G dikaitkan dengan penurunan skor fibrosis ($P < 0.05$). Analisis interaksi gen-gen mendedahkan perkaitan kuat antara gen-gen *AGTR1*, *LEPR* dan *GCKR* dengan gen *PNPLA3* (empirik $P = 0.007$, empirik $P = 0.001$ and empirik $P = 0.003$, masing-masing). Pemprofilan genom luas telah mengenalpasti lima CNV yang sering digandakan terletak di kromosom 14q11.2 (62.5%), 12p13.31 (50.0%), 11p15.4 (45.0%), 5p15.33 (42.5%), dan 12p13.2 (40.0%). CNVs yang sering dipadamkan terletak di kromosom 11q11 (35.0%), 12p13.2 (32.5%), 16q12.2 (32.5%), 4q13.2 (32.5%), 1q21.3 (32.5%), and 14q24.3 (32.5%). Boleh dihipotesiskan bahawa CNV ini mungkin diimplikasikan di dalam mekanisma yang menjurus kepada perkembangan NASH. Turut diberi perhatian adalah penghapusan 0.02 Mb pada kromosom 16q12.2. Kawasan ini kaya dengan gen carboxylesterase 1 (*CES1*) yang penting dalam pengangkutan kolesterol yang membawa kepada hati berlemak. Kajian berterusan untuk mengesahkan penemuan ini akan dilaksanakan. Pada keseluruhannya, kajian ini menunjukkan sifat kompleks NAFLD yang melibatkan interaksi gen berganda. Penemuan dari segi bilangan salinan merupakan penemuan baru dan berpotensi menjadi penanda genetik untuk mengenalpasti jenis NAFLD yang berpotensi untuk menjadi bertambah teruk.

ACKNOWLEDGEMENTS

In the name of Allah, the Most Gracious and the Most Merciful

Alhamdulillah, all praises to Allah for His blessings. This thesis is made possible with the guidance and help from several individuals who in one way or another contributed and extended their valuable assistance in the preparation and completion of this study.

Above all, my utmost gratitude to Professor Zahurin Mohamed, Head of Department of Pharmacology, also as a supervisor (genomics), whose guidance and encouragement I will never forget. She has been my inspiration during my Doctor of Philosophy (PhD) program. If she had not continuously encouraged me to pursue my PhD, we would not be reading this now. I would like to express my deepest appreciation to my clinical supervisor, Professor Rosmawati Mohamed who has introduced me to the field of hepatology, whose enthusiasm for the clinical aspect, and who has been a great reminder to always be positive and think positive, had brought quality to this study.

A brief word of thanks to Assistant Professor Agus Salim from National University of Singapore, for the insights in biostatistics he has shared. My sincere thanks also goes to the staff of the Oxford Gene Technology, Professor Colin Wheeler and Dr Douglas Hurd for the great assist in hands-on and data analysis during my training in Oxford. Not forgotten, thank you to the staff of the Sengenics Sdn Bhd, Dr Arif Anwar, Miss Nurul Shielawati Mohamed Rosli and Dr Rozaimi Razali for being accommodating to my queries.

I would like to acknowledge the financial, academic and technical support of the University of Malaya and its staff, particularly in the award of University Malaya Fellowship Scheme and Postgraduate Research Fund. The department and Medical

Biotechnology Laboratory facilities, have been indispensable. I thank my fellow labmates, Yamunah, Hwa Li and Aizat, for the moral support, friendship and memories. In one, I thank my laptop for not dying on me, and my motorcycle for transporting me to the faculty without breaking down too much.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

%	percentage
^{o}C	degree Celcius
±	plus minus
↑	increase
↓	decrease
Ɵ	no association
µg	microgram
µL	microliter
χ2	chi square
ABCB11	ATP-binding cassette, subfamily B, member 11
aCGH	array comparative genomic hybridization
ADIPOQ	adiponectin
ADIPOR1	adiponectin receptor 1
ADIPOR2	adiponectin receptor 2
AGTR1	angiotensin II type 1 receptor
AIDS	acquired immune deficiency syndrome
AKT1	v-akt murine thymoma viral oncogene homolog 1
ALT	alanine aminotransferase
ANOVA	analysis of variance
APOA5	apolipoprotein A-V
AST	aspartate aminotransferase
BCCP	Biotin Carboxyl Carrier Protein
BMI	body mass index
bp	base pair
CASP3	caspace 3
CCL3L1	chemokine (C-C motif) ligand 3-like 1

CD14	cluster of differentiation 14
CES1	carboxylesterase 1
CFTR/MRP	ATP-binding cassette, subfamily C, member 2
CHB	Han Chinese from Beijing
CI	confident interval
CLOCK	circadian locomoter output cycles protein kaput
cm	centimeter
CNA	copy number alteration
CNV	copy number variation
COL13A1	collagen type XIII alpha 1
COX-2	cyclooxygenase 2
CVD	cardiovascular disease
CYP2E1	cytochrome P450, family 2, subfamily E, polypetide 1
Defcr-1	defensin related cryptdin 1
DGV	Database of Genomic Variants
DNA	deoxyribonucleic acid
EFCAB4B	EF-hand calcium binding domain 4B
ENPP1	ectonucleotide pyrophosphatase/phosphodiesterase 1
FABP2	fatty acid binding protein 2
FATP5	fatty acid transport protein 5
FDFT1	farnesyl-diphosphate farnesyltransferase 1
FDR	false discovery rate
FOX	forkhead box
FTO	fat mass and obesity-associated protein
g	gravity force
g/day	gram per day
GCKR	glucokinase regulatory
GCLC	glutamate-cysteine ligase, catalytic subunit

gDNA	genomic deoxyribonucleic acid
GGT	gamma-glutamyl transpeptidase
GMDR	Generalized Multifactor Dimensionality Reduction
GOLD	Genetics of Obesity-related Liver Disease
GWAS	genome-wide association study
GYPA	glycophorin A
HbA1c	hemoglobin A1c
HCC	hepatocellular carcinoma
HDL	high density lipoprotein
HeFH	heterozygous familial hypercholestrolemia
HFE	hemochromatosis
HIV	human immunodeficiency virus
HOMA-IR	Homeostasis Model of Assessment - Insulin Resistance
HSCs	hepatic stellate cells
HU	Hounsfield units
IGH	immunoglobulin heavy chain
IL-6	interleukin-6
IL1B	interleukin 1, beta
iNOS	inducible nitric oxide synthase
IRS1	insulin receptor substrate 1
It1na	interlectin
IU/L	international units per liter
kb	kilobase
kg/m^2	kilogram per meter squared
KLF6	kruppel-like factor 6
LD	linkage disequilibrium
LDL	low density lipoprotein
LDLR	low density lipoprotein receptor

LEP	leptin
LEPR	leptin receptor
LNA	Locked Nucleic Acid
LTBP3	latent transforming growth factor beta binding protein 3
LYPLAL1	lysophospholiphase-like 1
MAF	minor allele frequency
Mb	mega base
MC4R	melanocortin-4 receptor
mg/dL	miligram per deciliter
MIF	macrophage migration inhibitory factor
miRNA	microRNA
mL	milliliter
MMP14	matrix metallopeptidase 14
MRI	magnetic resonance imaging
MTA1	metastasis associated 1
MTHFR	methylenetetrahydrofolate reducatase
MTTP	microsomal triglyceride transfer protein
NAFLD	non-alcoholic fatty liver disease
NAHR	non-allelic homologous recombination
NASH	non-alcoholic steatohepatitis
NASH[a]	non-alcoholic steatohepatitis without significant fibrosis
NASH[b]	non-alcoholic steatohepatitis with significant fibrosis
NCAN	neurocan
NCBI dbEST	National Center for Biotechnology Information database of Expressed Sequenced Tag
NHEJ	non-homologous end joining
NPY5R	neuropeptide Y receptor type 5
NR1I2	pregnane X receptor

OR	odds ratio
OR	olfactory receptor
PARK2	parkinson protein 2
PCR	polymerase chain reaction
PDGFA	platelet-derived growth factor alpha
PEMT	phosphatidylethanolamine N-methyl transferase
PNPLA3	patatin-like phospholipase domain-containing protein 3
PPARA	peroxisome proliferator-activated receptor alpha
PPARG	peroxisome proliferator-activated receptor gamma
PPARGC1A	peroxisome proliferators-activated receptor gamma coactivator 1 alpha
PPP1R3B	protein phosphatase 1, regulatory subunit 3b
PTGS2	prostaglandin-endoperoxide synthase 2
ROC	receiver operating characteristic
rpm	revolutions per minute
s	second
SD	standard deviation
SERPINA1	serpine peptidase inhibitor, clade A, member 1
SGCZ	sarcoglycan
SNPs	single nucleotide polymorphism
SOD2	superoxide dismutase 2
SREBF1	sterol regulatory element-binding transcription factor 1
STAT3	signal transducer and activator of transcription 3
TCF7L2	transcription factor 7-like 2
TGF-β 1	transforming growth factor-β 1
TNF-α	tumor necrosis factor alpha
TRAIL	tumor necrosis factor (TNF)-related apoptosis inducing ligand

TRIM12	tripartite motif-containing 12
TRIM34	tripartite motif-containing 34
UCP3	uncoupling protein 3
UMMC	University Malaya Medical Centre
UV	ultraviolet
VLDL	very low density lipoprotein
VNTRs	variable number of tandem repeats
WHO	World Health Organisation
ZNF267	zinc finger protein 267

CHAPTER ONE
INTRODUCTION

Non-alcoholic fatty liver disease (NAFLD) is one of the most common causes of chronic liver disease and has emerged as an important global health problem (L. Adams, Angulo, & Lindor, 2005; Angulo, 2002). The reported prevalence of NAFLD in the general population is as high as 35% (Cheung & Sanyal, 2010). NAFLD comprises a spectrum of diseases ranging from simple steatosis, which is benign fatty infiltration of the liver to its inflammatory counterpart non-alcoholic steatohepatitis (NASH−with or without fibrosis), cirrhosis (scarring and hardening of the liver), and hepatocellular carcinoma (Charlton, 2004; Malaguarnera, Di Rosa, Nicoletti, & Malaguarnera, 2009). Unlike simple steatosis which is usually benign, NASH can potentially progress, leading to liver-related mortality.

NAFLD represents the hepatic component of the metabolic syndrome and is strongly associated with obesity and insulin resistance (L. A. Adams, Waters, Knuiman, Elliott, & Olynyk, 2009; Bertrand, Lefevre, Prevot, & Perlemuter, 2011; Fatani, Itua, Clark, Wong, & Naderali, 2011; Vanni et al., 2010). The interplay between environmental and genetic factors promotes the development and progression of NAFLD. Life style modification has been shown to improve liver histology in NAFLD patients (Bhat, Baba, Pandey, Kumari, & Choudhuri, 2012; Suzuki et al., 2005). A familial aggregation study comparing overweight children with NAFLD, and overweight children without NAFLD reported that fatty liver was more significant in siblings of children with NAFLD (Schwimmer et al., 2009).

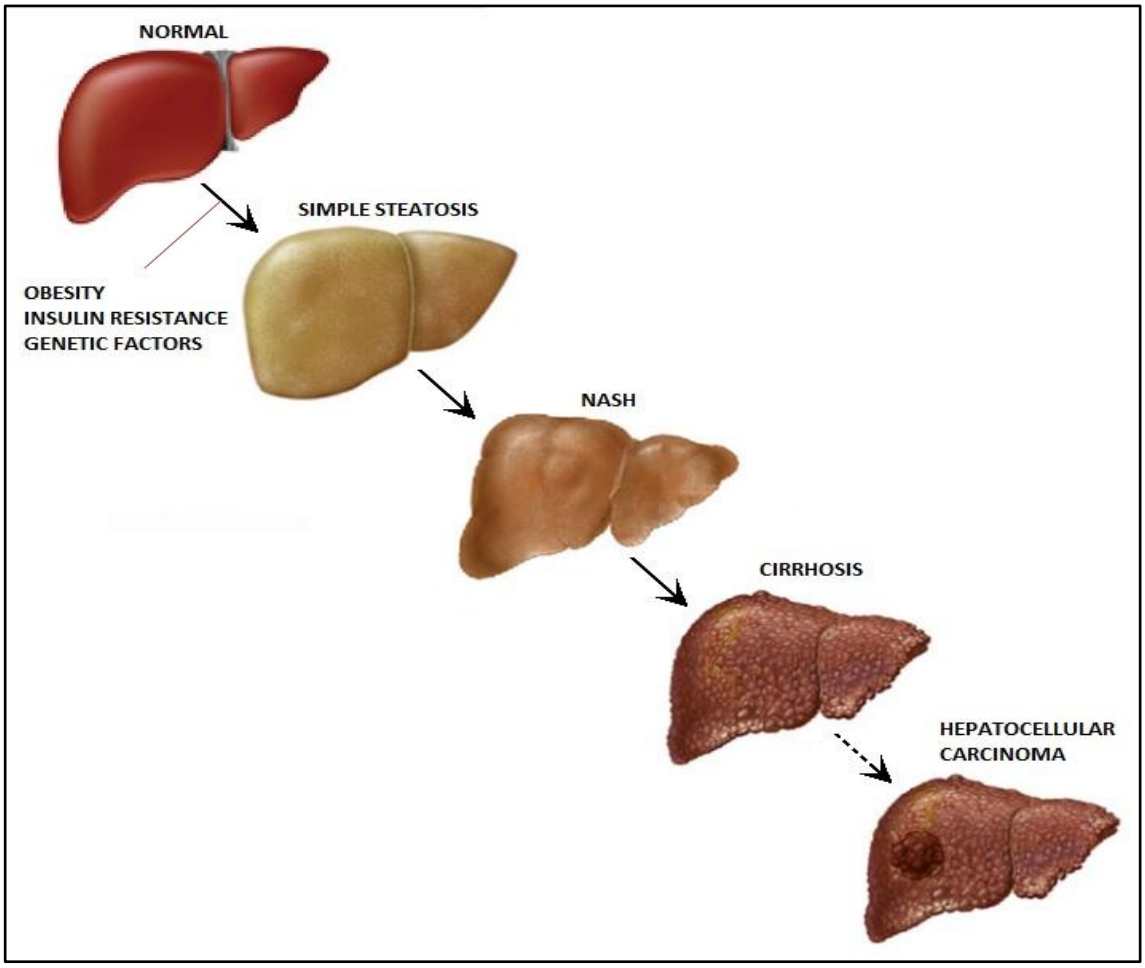


Figure 1.1 Spectrum of NAFLD

Adapted and modified from Trauner, Arrese, and Wagner (2010)

Genetic variations are responsible for different phenotype among individuals. This genetic makeup also determines the tendency of certain individuals to develop a disease. Genomic variability may present in various forms such as single nucleotide polymorphisms (SNPs), structural alterations (deletion, duplication and inversion), variable number of tandem repeats (VNTRs) and presence or absence of transportable element (Freeman et al., 2006). Of all these variations, SNPs are the commonest and preferred over others due to their high abundance along the human genome (SNPs with minor allele frequency > 0.1 occur once every ~600 kb), low mutation rate and the high accessibility of high-throughput genotyping (D. G. Wang et al., 1998). SNPs have been associated with many diseases including NAFLD.

Other than the more commonly studied SNPs, copy number variation (CNV) is another type of mutation that gives more insight into the human genome and may explain the missing heritability. To date there is no published report on CNV of NAFLD due to the nature of the study which is expensive and the characteristics of CNV with the likes of total number, genomic positions, gene content, frequency spectrum and patterns of linkage disequilibrium (LD) with one another remain unknown, which make the study capacity limited (K. K. Wong et al., 2007).

Linkage disequilibrium (LD) occurs when alleles at a given locus are not independent of alleles at a second locus. The alleles that pass from one generation to another are therefore as a result of non-random sampling. LD decays at each generation at a rate determined by the degree of recombination. There are five factors that contribute to linkage disequilibrium which are mutation, random drift, natural selection, inbreeding, and gene flow. Mutation leads to non-random associations between alleles at different loci, thus contributing to LD. However the effect of mutation is rather weak due to its weak effect on allele frequency change. Furthermore, the rate of recombination is typically higher than mutation and it will overtake the non-random associations. Random genetic drift from a gene pool of one generation to another affects the allele frequency change by random sampling variation. Thus, the smaller the population, the larger the effect of genetic drift. The expected value of disequilibrium among replicate populations is zero. Natural selection promotes LD by favoring a specific combination of alleles at different loci. This combination may function better as a group. Natural selection may lead to the evolution of supergenes. The recombination between individual genes in the supergene is reduced to maintain this strong linkage. Inbreeding affects the LD by reducing the rate of decay of LD through the reduction of the proportion of double heterozygotes formed by recombination. Gene flow or gene migration which is the transfer of alleles or genes from one population to another

affects the allele frequency change and therefore affects the LD. The rate of decay is determined by recombination when the gene flow is extensive. On the other hand, when the gene flow is small, the rate of decay is determined by the proportion of migrants.

It is increasingly important to learn that our biological process including susceptibility to disease and response to drugs depend not just on the genotypes, but also on the specific haplotypes. Haplotype is a combination of markers (alleles) at multiple locations on a single chromosome. Haplotyping is much more powerful than allele-by-allele analysis due to the application of multi markers rather than single marker only. Human genome can be divided into haplotypes block . A small fraction of SNPs in the block, referred as "tagSNP" is used to capture most haplotypes. However, different frequency of SNPs and haplotypes across populations generate different haplotype block structure and tagSNP. Different population therefore may have different haplotype block structure and tagSNP (Zhang, Calabrese, Nordborg, & Sun, 2002).

The main outcomes of the study were to assess, in the same set of study population, whether the genetic polymorphisms and copy numbers are associated with severity of NAFLD, and to evaluate the effects of these genetic variations on histological features of NAFLD. The Malaysian population comprises of three major ethnic subgroups, namely the Malays, Chinese and Indians, each of which are presumably of different genetic pool, hence presenting a good opportunity to study the ethnic difference of NAFLD susceptibility with regard to genetic polymorphisms. Meanwhile, the copy number study is novel and the new finding should serve as a potential new genetic marker for the progressive form of NAFLD.

1.1 Objectives

The aims of this study are:

1. To investigate the association between single nucleotide polymorphisms with NAFLD and with severity of NAFLD in the three major ethnic groups of Malaysian patients.
2. To study the association between candidate SNPs with parameters and histological features of NAFLD, namely steatosis, lobular inflammation, hepatocellular ballooning and fibrosis.
3. To evaluate the gene-gene interaction between the SNPs of interest.
4. To determine the profile of copy number variations in the spectrum of NAFLD.

1.2 Justification of study

The study of non-alcoholic fatty liver disease is relatively new. Eventhough the condition was first introduced in the 1930s, named in the 1960s and characterised histologically in 1980, the major breakthrough on the genetic association studies of NAFLD only started in 2000s. In 2008, a single variant of the *PNPLA3* gene rs738409 was found to be associated with hepatic fat content in a genome-wide association of the Dallas Heart Study. It then served as the starting point for more genetic association studies of NAFLD in various populations. Intriguingly, all yielded a postive association finding. Although most of the case-control association studies of *PNPLA3* rs738409 and its susceptibility to NAFLD have been reported in various populations, only few have reported on NAFLD disease severity and no data is available on its association with severity of non-alcoholic steatohepatitis (NASH). Published genetic reports were centered on the most common pathway of NAFLD; lipid metabolism and insulin resistance. Keeping in mind that NAFLD is a hepatic manifestation of metabolic syndrome that includes obesity and type 2 diabetes, investigation was also focused on

the association of SNPs that were previously reported in obesity and type 2 diabetes. This study reports the finding in a Malaysian population which comprises of three major ethnic groups, namely the Malays, Chinese and Indians. Each of the ethnic subgroups are presumably of different genetic pool, thus providing a good setting to study ethnic differences in their susceptibility to NAFLD. Interaction of genes that share the same pathway or function is another important aspect to study in order to understand the underlying pathway of NAFLD disease progression. This study is also the first attempt to identify CNVs associated with NAFLD. The identification of novel CNV markers especially in the NAFLD spectrum may provide hints towards NAFLD progression. Genetic tests may provide early insights into an individual's susceptibility to this disease despite life style modification. These NAFLD genetic studies may aid researchers or even clinicians to better understand the nature of the disease.

1.3 Research hypothesis

Hypothesis for this study is:

The null hypotheses (H_o):

1. There is no association between investigated SNPs with NAFLD and with severity of NAFLD in the three major ethnic groups of the Malaysian patients.
2. There is no association between investigated SNPs with parameters and histological features of NAFLD.
3. There is no gene-gene interaction between the investigared SNPs.
4. There is no difference between copy number variations at the different NAFLD spectrum.

The research hypotheses (H_1):

1. There is an association between investigated SNPs with NAFLD and with severity of NAFLD in the three major ethnic groups of the Malaysian patients.
2. There is an association between investigated SNPs with parameters and histological features of NAFLD.
3. There is a gene-gene interaction between investigated SNPs.
4. There is a difference between copy number variations at the different NAFLD spectrum

CHAPTER TWO
LITERATURE REVIEW

2.1 Background of NAFLD

Non-alcoholic fatty liver disease (NAFLD) took its name only in 1960s although it has been recognised as early as 1930s. However, research on NAFLD comes to peak after it was histopathologically charaterised in 1980. NAFLD has now becoming a growing global disease and a major concern on individuals health (Godlee, 2011). NAFLD cannot be explained by a single gene inheritance, such that occurs in Mendelian inheritance. It is rather a complex disease that involves multiple genes and their interactions with environmental factors (Hernaez, 2012).

The prevalence of NAFLD varies accross populations. Majority of patients with NAFLD possess metabolic disturbances such as obesity, diabetes mellitus and dyslipidemia (Chalasani et al., 2012). The prevalence of NAFLD therefore is not assessed in patients with NAFLD per se, but rather morbid obese patients undergoing bariatic surgery, patients with diabetes and patients with dyslipidemia. It was reported that the prevalence of NAFLD can reach up to 90% in morbidly obese patients undergoing bariatric surgery, of which 5% were at risk of developing cirrhosis (Boza et al., 2005; Haentjens et al., 2009; Machado, Marques-Vidal, & Cortez-Pinto, 2006). More than half of the diabetic patients showed faty infiltration on ultrasound, and from those who consented for liver biopsy, 87% confirmed NAFLD (Prashanth et al., 2009). Prevalence of NAFLD was reported to be 50% in individuals with dyslipidemia (Assy et al., 2000). NAFLD does not only affect adults but also children and adolescents. Data from a United States study documented that children and adolescents had 10% risk of NAFLD (Schwimmer et al., 2006). In the Asia-Pacific region, the prevalence is at least 10% while in some regions, one out of every three individuals appear to be at risk

(Amarapurkar et al., 2007; Fan et al., 2007). Even after applying a strict body mass index (BMI) cut-off points for Asians, as high as 15-21% of NAFLD subjects were revealed to be non-obese: normal body mass index (BMI 17.5-22.4 kg/m^2) or overweight (BMI 22.5–24.9 kg/m^2) (C. J. Liu, 2012).

The pathogenesis of NAFLD is based on the "two hits hypothesis". The "first hit" which is the development of steatosis, involves the accumulation of triglycerides in the liver due to insulin resistance. Insulin's role in this insult is by (i) promoting the adipocytes to secrete potentially harmful adipocytokines (ii) altering the rates of synthesis and transport of hepatic triglycerides (iii) increasing lipolysis (Bugianesi, McCullough, & Marchesini, 2005). Lipid accummulation, particularly diacylglycerol, together with mitochondrial dysfunction, endoplasmic reticulum stress and inflammation, could contributes to insulin resistance (Gariani, Philippe, & Jornayvaz, 2012). The triglycerides accummulation occurs when the rate of fatty acid uptake (hepatic fatty acid uptake and de novo synthesis) is greater than secretion (fatty acid oxidation and transport in the form of VLDL−very low density lipoprotein) (Fabbrini, Sullivan, & Klein, 2010).

Insulin resistance prepares the hepatocytes for the second insult. In type 2 diabetes mellitus-induced hyperglycemia patients with NAFLD, it was shown that parameters for oxidative stress such as hydrogen peroxide, malondialdehyde, interleukin-6 and tumor necrosis factor-alpha were found to be increased. Their levels were positively correlated with plasma glucose level (Shams, Al-Gayyar, & Barakat, 2011). The "second hit" event is particularly due to adipocytokines and oxidative stress, that further damages the liver and promotes the progression to steatohepatitis and fibrosis. The degree of hepatic inflammation and fibrosis is associated with the presence of metabolic traits (Rector, Thyfault, Wei, & Ibdah, 2008). Most of the NAFLD study

designs to date, including candidate gene studies, applied this hypothesis (Day, 2002; Day & James, 1998).

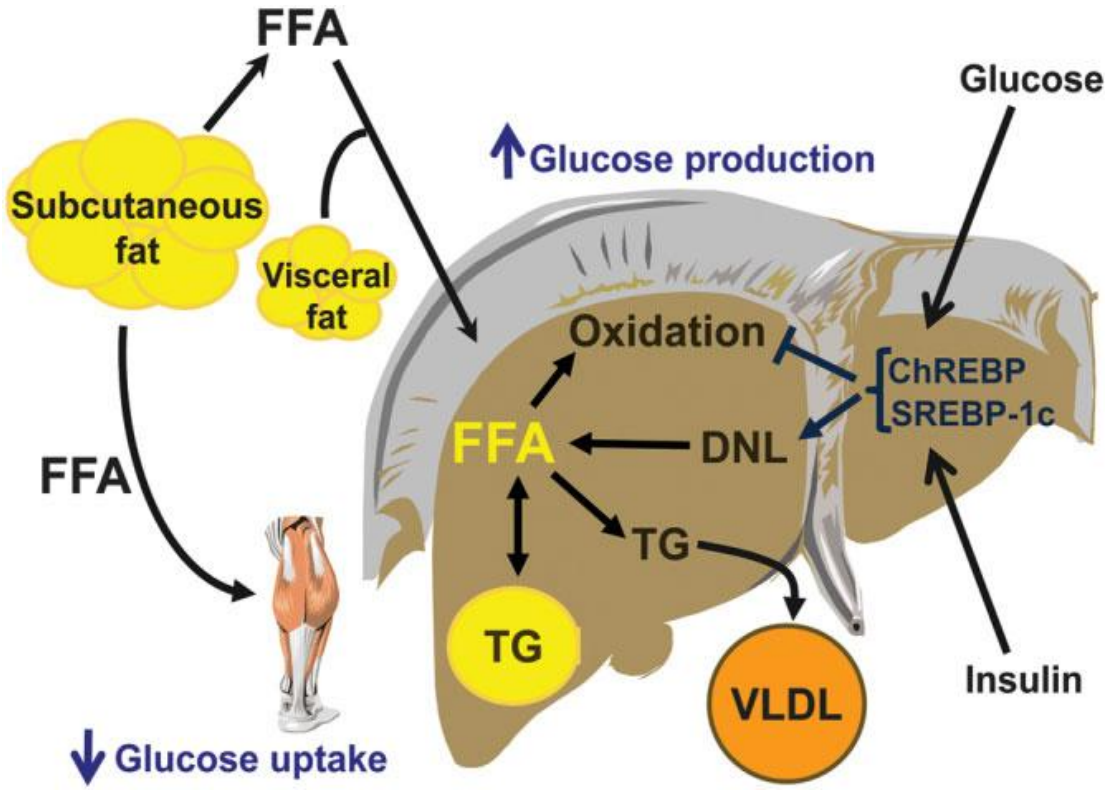


Figure 2.1 Physiological interrelationships among fatty acid metabolism, insulin resistance, dyslipidemia and intra-hepatic triglycerides content in NAFLD

Adapted from Fabrini et al. (2010)

Patients with NASH have a greater tendency to develop advanced fibrosis than those with simple steatosis (Jiang & Torok, 2008; Z. Li et al., 2003). NASH will subsequently lead to complications such as cirrhosis, hepatocellular carcinoma and liver failure (Matteoni et al., 1999). Powel et al. (1990) was the first to highlight the relationship between NASH and cirrhosis, and to report the association between NAFLD-associated cirrhosis with hepatocellular carcinoma (HCC). Studies have also documented the association between NAFLD and HCC in the absence of cirrhosis (Baffy, Brunt, & Caldwell, 2012). Patients with NAFLD-associated HCC displayed

more metabolic features than those with non-complicated by HCC (Ertle et al., 2011). A retrospective study that investigated patients that had undergone transplant due to NASH-associated cirrhosis, found that 24 out of the 40 patients has died while those with metabolic syndrome had a 1-year mortality rate of 42% (Heuer et al., 2012). Liver transplantation cases due to NASH have been rapidly growing from 3% to 19% for over a 10-year period (Agopian et al., 2012). Another underlying disease associated with NAFLD that is of major concern is cardiovascular disease (CVD). CVD has emerged as a highly death-related disease due to NAFLD. The prevalence of CVD is higher in patients with NAFLD than those without, and this finding is independent of the metabolic components (Targher et al., 2007). Several studies have conlcuded that CVD is the most common causes of death in patients with NAFLD and the prevalence ranges from 12.7% to 38% (Dam-Larsen et al., 2009; Ekstedt et al., 2006; N. Rafiq et al., 2009; Soderberg et al., 2010).

2.2 NAFLD and metabolic syndrome

NAFLD is closely associated with metabolic syndrome. Liver fat content was found to be significantly greater in subjects with metabolic syndrome than those without. Liver Hounsfield units (HU) which will decrease with every miligram of fat deposited in the liver, were negatively correlated with BMI, waist circumference and percent body fat (Dick, Lesser, Leipsic, Mancini, & Lear, 2013). It was reported that almost 90% patients with NAFLD presented with more than one criteria of metabolic syndrome and 33% with three or more criteria (Almeda-Valdes, Cuevas-Ramos, & Aguilar-Salinas, 2009). In a study of 1365 subjects in which the World Health Organisation (WHO) criteria was applied, the prevalence of metabolic syndrome was 49.8% in the pre-diabetic and 58.9% in the diabetic (Seo et al., 2012). In the adults study, NASH cases were reported to be from 40% to 100% in obese patients (Tarantino, Saldalamacchia,

Conca, & Arena, 2007). While in the children, it is about 15% to 25% (Povel, Boer, Reiling, & Feskens, 2011).

A Korean study revealed that BMI 28.9 kg/m^2 was useful to distinguish NASH from simple steatosis (J. W. Park, Jeong, Kim, Kim, & Park, 2007). Newly diagnosed diabetes patients have a greater prevalence of NAFLD (62%) compared to those with impaired glucose tolerance (43%) (Jimba et al., 2005). It was reported that in a Korean study, the risk of abnormal glucose metabolism is greater in NAFLD than in control group, with those assessed by impaired fasting glucose at 2.13 times risk, HOMA-IR 3.25 times risk, type 2 diabetes 7.63 times risk, BMI 3.35 times risk, triglyceride 3.05 times risk and fasting blood glucose 2.18 times risk (Chon et al., 2012). In the same year a Korean study revealed that the degree of NAFLD significantly affects the incidence rate of type 2 diabetes, in which normal at 7.0%, mild 9.8%, moderate to severe 17.8% (S. K. Park, Seo, Shin, & Ryoo, 2012). A group of newly diagnosed type 2 diabetes exhibited abnormal alanine aminotransferase (ALT) levels (Saligram, Williams, & Masding, 2012). Treatment with glucagon-like peptide receptor agonist in a well-defined obese, type 2 diabetes patients with fatty liver showed a reduction in hepatic fat content and this was correlated with HbA1c but not with the total body weight (Cuthbertson et al., 2012). About 20% to 80% of NAFLD cases were associated with dyslipidemia (Souza, Diniz Mde, Medeiros-Filho, & Araujo, 2012) which is characterized by increased triglyceride and low density lipoprotein (LDL) cholesterol levels, and decreased high density lipoprotein (HDL) cholesterol levels (Chatrath, Vuppalanchi, & Chalasani, 2012). Marchesini et al reported as high as 64% of patients with steatosis presented with hypertriglyceridemia (Marchesini et al., 2003).

Table 2.1 Criteria for clinical diagnosis of metabolic syndrome

National Cholesterol Education Program's Adult Treatment Panel III report (ATP III)	World Health Organization (WHO)	American Association of Clinical Endocrinologists (AACE)
• Abdominal obesity, waist circumference Men > 102 cm (> 40 in) Women > 88 cm (> 35in) • Triglycerides ≥ 150mg/dL • HDL cholesterol Men < 40 mg/dL Women < 50 mg/dL • Blood pressure ≥ 130/≥ 85 mmHg • Fasting glucose ≥ 100 mg/dL	Insulin resistance, identified by 1 of the following: • Type 2 diabetes • Impaired fasting glucose • Impaired glucose tolerance • Or for those with normal fasting glucose levels (< 110 mg/dL), glucose uptake below the lowest quartile for background population under investigation under hyperinsulinemic, euglycemic conditions Plus any 2 of the following: • Antihypertensive medication and/or high blood pressure (≥ 140 mmHg systolic or ≥ 90 mmHg diastolic) • Plasma triglycerides ≥ 150 mg/dL (≥ 1.7 mmol/L) • HDL cholesterol < 35 mg/dL (< 0.9 mmol/L) in men or < 39 mg/dL (1.0 mmol/L) in women • BMI > 30 kg/m^2 and/or waist:hip ratio > 0.9 in men, > 0.85 in women • Urinary albumin excretion rate ≥ 20 µg/min or albumin:creatinine ratio ≥ 30 mg/g	• Overweight/obesity BMI > 25 kg/m^2 • Elevated triglycerides ≥ 150 mg/dL (1.69 mmol/L) • Low HDL cholesterol Men < 40 mg/dL (1.04 mmol/L) Women < 50 mg/dL (1.29 mmol/L) • Elevated blood pressure ≥ 130/85 mmHg • 2-Hour postglucose challenge >140 mg/dL • Fasting glucose between 110 and 126 mg/dL • Other risk factors Family history of type 2 diabetes, hypertension, or CVD Polycystic ovary syndrome Senderaty lifestyle Advancing age Ethnic groups having high risk for type 2 diabetes or CVD

Adapted and modified from Grundy, Brewer, Cleeman, Smith, and Lenfant (2004)

2.3 Genetic background of NAFLD

Although NAFLD is considered a hepatic manifestation of metabolic syndrome, and the majority of patients with NAFLD present with metabolic risk factors, family-based genetic studies have shown that the presence of NASH and/or cryptogenic cirrhosis was not constantly associated with metabolic risk factors (Struben, Hespenheide, & Caldwell, 2000; Willner et al., 2001). Ethnic studies provided evidence that Asian patients with NAFLD presented with significantly lower body mass index (BMI) than other ethnic groups (Farrell, 2003). This was further supported by a study that reported increased prevalence of fatty liver in Asian Indian men presented with normal BMI (Petersen et al., 2006). A strong evidence of a genetic component is also reflected by the differences in the prevalence of NAFLD across different populations. In addition, a familial aggregation study has shown high heritability of NAFLD, with estimates of 59% in siblings and 78% in parents with NAFLD (Schwimmer et al., 2009). Intriguingly, well-defined genetic markers for obesity (*FTO*−fat mass and obesity-associated protein) and diabetes (*TCF7L2*–transcription factor 7-like 2) were not associated with NAFLD, indicating that fatty liver disease has its own genetic underpinnings (Hernaez, 2012). Based on this hypothesis, genetic studies can therefore be simplified into studies of genes involved in hepatic lipid metabolism, insulin resistance and oxidative stress.

The imbalance or alteration in the hepatic lipid pathway especially the hepatic triglycerides, will promote the development of hepatic steatosis (Hernaez, 2012). Some of the genes involved in hepatic lipid metabolism are phosphatidylethanolamine N-methyl transferase (*PEMT*), circadian locomoter output cycles protein kaput (*CLOCK*), fatty acid binding protein 2 (*FABP2*) and microsomal triglyceride transfer protein (*MTTP*). Study on six tagSNPs on the *CLOCK* gene found rs11932595 and rs6843722 to be associated with NAFLD (Sookoian, Castano, Gemma, Gianotti, & Pirola, 2007).

Significantly higher frequency of Met/Met at residue 175 (loss of function SNP) was observed in NAFLD subjects compared to controls (Song et al., 2005). Insulin resistance is important in the "first hit" mechanism. More importantly, it has a central role in metabolic syndrome. Key genes involved in insulin sensitivity are glucokinase regulatory (*GCKR*), peroxisome proliferator-activated receptor gamma (*PPARG*) and melanocortin-4 receptor (*MC4R*). A genome-wide association study (GWAS) consortium showed *GCKR* to be associated with glycemic traits such as fasting glucose and HOMA-IR (Speliotes et al., 2011). A recent study by Bhat et al. (2013) provided evidence that polymorphisms of Pro12Ala and C161T on *PPARG* gene increased risk of NAFLD in Asian Indians. Meanwhile, oxidative stress is an underlying cause of NAFLD progression to NASH and cirrhosis. Meta-analysis data on tumor necrosis factor alpha gene (*TNF-α*), which is known to be involved in oxidative stress concluded that promoter polymorphism at position -238 but not -308 contributes to risk of NAFLD (J. K. Wang, Feng, Li, Li, & Tao, 2012).

2.4 Single nucleotide polymorphism

Single nucleotide polymorphism or SNP (pronounced "snip") is one the most common genetic variations that occur at high density in the human genome. The ubiquity and the ease with which genotyping can be carried out, has facilitated the mushrooming studies of SNPs. By definition, SNP is a variation of the DNA sequence in which a single nucleotide of the sequence has been altered. These highly abundant SNPs occur in every 1000 bases in the human genome (Sachidanandam et al., 2001). These SNPs may result in phenotypic variation depending on the location in the human genome. SNPs that occur in the coding regions genes may alter the encoded protein products, thus changing the function of the proteins. These SNPs are important in the human pathophysiology and always been related to the dominant and recessive inherited monogenic disorders

and are generally analysed for diagnostic purposes. Alteration of the primary structure of a protein involved in drug metabolism, brings us to a new aspect of the study of pharmacogenomics and translational medicine. This type of SNP is frequently analysed to assess the risk of an individual for a particular disease and the choices of drugs administered. It is also likely that common diseases are influenced by SNPs in the regulatory regions of genes. However, most SNPs have no direct known impact on the phenotype of an individual. These SNPs are located in non-coding regions of the genome and are useful as markers in population genetics (Jorde et al., 2000).

The introduction of whole-genome SNP genotyping that has the ability to genotype millions of SNPs that cover the whole genome has made the production of high throughput genotyping data faster with greater power. This wide application of genotyping technology detects genomic regions in which the allele frequencies differ between patients and controls. The data generated would be useful to detect and isolate the disease-predisposing genes, and proteins encoded by these genes would potentially be valuable targets for developing new therapeutic drugs. High throughput genotyping requires advance laboratories and facilities, in addition to its high cost to run the genome-wide SNP genotyping. With that, an alternative to this whole-genome SNP mapping has been applied, which is the use of candidate genes, tagSNPs and SNP markers. Genotyping technology can be improved by effectively handling the key determinants which are cost, high throughput, assay design and accuracy. The present approach of improving the throughput and accuracy of the genotyping is by the use of microarray and DNA sequencing (Syvanen, 2001). The effectiveness of genome wide association studies is achievable by the completion of the International HapMap Project Phase I, II and III ("A haplotype map of the human genome," 2005; The international hapmap project," 2003). This approach has greatly increased the throughput and accuracy in a cost-effective way.

Affymetrix and Illumina are among the earliest platform used to generate this high-throughput genome-wide genotyping and have now been accessible in some standard laboratories. These platforms have now been used as standard technique in molecular genetics and clinical diagnostic laboratories. Genome-wide SNP genotyping is widely used to identify common genetic variant that underlies disease (Farrall & Morris, 2005; W. Y. Wang, Barratt, Clayton, & Todd, 2005). Although genome-wide SNP genotyping promises a good genotyping data, the experiments are extremely sensitive to the quality of DNA. Prior measurements of DNA are necessary and when the concentrations fall below the quality standard, counter measurements need to be taken such as DNA precipitation (Gibbs & Singleton, 2006). In 2008, the application of genome-wide SNP genotyping has found a non-synonymous variant of *PNPLA3* rs738409 to be highly associated with fat content in the liver (Romeo et al., 2008). This is a major breakthrough to the further studies of *PNPLA3* rs738409 both in molecular genetics and functional studies.

2.5 Single nucleotide polymorphism of candidate genes in NAFLD

Various genes and single nucleotide polymorphisms have been implicated in the development of NAFLD. Among these genes, genes that are involved in lipid metabolism, amino acid catabolism, inflammation, fatty acid transport and insulin pathway are strongly associated with NAFLD (Greco et al., 2008). The much studied gene in NAFLD, patatin-like phospholipase domain-containing protein 3 (*PNPLA3*) is not limited to molecular genetics in humans but is also functional in animals (Fuchs et al., 2012; Qiao et al., 2011). The role of obesity-related gene, leptin receptor (*LEPR*) has been well studied, particularly in obesity and metabolic syndrome (Stylianou et al., 2007; C. W. Yang et al., 2013). There is also evidence of influence of angiotensin II type 1 receptor (*AGTR1*) gene polymorphisms in Japanese patients with NAFLD

(Yoneda et al., 2009b). Study of tagSNPs of signal transducer and activator of transcription 3 (*STAT3*) showed both single marker and multimarker of the gene is associated with susceptibility to NAFLD (Sookoian et al., 2008). A GWAS of the NAFLD Database Study identified a variant in the farnesyl-diphosphate farnesyltransferase 1 gene (*FDFT1*) as a risk for NAFLD. The same GWAS showed variant in the collagen type XIII alpha 1 gene (*COL13A1*) to be associated with lobular inflammation (Chalasani et al., 2010). Peroxisome proliferator-activated receptor gamma (*PPARG*), sterol regulatory element-binding transcription factor 1 (*SREBF1*), tumour necrosis factor alpha (*TNF-α*), and caspace 3 (*CASP3*) are related to the discovery of a gene network involved in NAFLD using genome wide transcriptional profiling (Hui et al., 2008; Kohjima et al., 2007).

Table 2.2 Genome-wide association studies in NAFLD

GWAS	Gene	SNP	Steatosis	Inflammation	Fibrosis
Romeo et al., 2008 & Speliotes et al., 2010	PNPLA3	rs738409	↑↑		
Romeo et al., 2008	PNPLA3	rs6006460	↓		
Chalasani et al., 2010	FDFT1	rs2645424		↑	
	COL13A1	rs1227756			↑
	PDGFA	rs343064			↑
	LTBP3	rs1227756		↑	
	EFCAB4B	rs887304		↑	
Speliotes et al., 2010	NCAN	rs2228603	↑		
	LYPLAL1	rs12137855	↑		
	GCKR	rs780094	↑		
	PPP1R3B	rs4240624	↑		

↑: increase; ↓: decrease; *GWAS* genome-wide association study

A loss of functional mutation of the phosphatidylethanolamine N-methyl transferase (*PEMT*) gene confers susceptibility to NAFLD (Y. J. Zhou et al., 2010). Genetic polymorphism in the promoter region of the cyclooxygenase 2 (*COX-2*) confers susceptibility to NAFLD in the Chinese population (Cao, Yang, & Dong, 2010). In a Japanese population, inducible nitric oxide synthase (*iNOS*) polymorphisms were found to be associated with NAFLD. The rs1060822, which is in strong LD with rs2297510, rs2297511 and rs2797512 was associated with increased in fibrosis (Yoneda et al., 2009a). A multimarker study of rs2461823 and rs7643645 of the pregnane X receptor (*NR1I2*) proposed a combined effect of both SNPs on NAFLD severity. The rs7643645 is also associated with ALT level (Sookoian et al., 2010). A lower risk of NAFLD and steatosis grade were shown by polymorphisms in the tumor necrosis factor (*TNF*)-related apoptosis inducing ligand (*TRAIL*) (Yan et al., 2009). Kruppel-like factor 6 (*KLF6*) expression was proportionally increased with the severity of NAFLD. The functional *KLF6-IVS1-27*G>A polymorphism was shown to be a predictor of fibrotic NAFLD (Miele et al., 2008). Peroxisome proliferators-activated receptor gamma coactivator 1 alpha (*PPARGC1A*) polymorphism significantly affects severity of NAFLD (Yoneda et al., 2008).

Table 2.3 Genetic factors associated with NAFLD progression

Author	Gene	SNP	Steatosis	Inflammation	Fibrosis
Insulin sensitivity					
Carulli et al., 2009 & Dongiovanni et al., 2010	*ENPP1*	rs1044498	ϴϴ	ϴϴ	↑
Dongiovanni et al., 2010	*IRS1*	rs1801278	ϴ	ϴ	ϴ
Tokushige et al., 2009; Wang et al., 2008; Musso et al., 2008 & Wong et al., 2008	*ADIPOQ*	rs2241766	ϴϴϴ↑↑	↑	ϴ
Tokushige et al., 2009; Wang et al., 2008; Musso et al., 2008 & Wong et al., 2008		rs1501299	↓ϴϴ↑	ϴ↑ϴ	
Kotronen et al., 2009	*ADIPOR1*	rs6666089	ϴ		
Kotronen et al., 2009	*ADIPOR2*	rs767870	↓		
Dongiovanni et al., 2010 & Chen et al., 2008	*PPARA*	rs1800206	ϴ↑	ϴ	ϴ
Zhou et al 2010	*PPARG*	rs3856806	ϴ		
Dongiovanni et al., 2010; Gupta et al., 2010 & Rey et al., 2010		rs1801282	ϴϴϴ	ϴϴ	ϴϴ
Zhou et al., 2010	*PPARGC1A*	rs8192687	ϴ		
Musso et al., 2009	*TCF7L2*	rs7903146	↑	↑	↑
Yang et al., 2011	*GCKR*	rs780094	↑		
Haupt et al., 2009	*MC4R*	rs17782313	ϴ		
Zhou et al., 2010	*LEP*	rs7799039	ϴ		
Zain et al., 2013	*LEPR*	rs1137100	ϴ	ϴ	↓
		rs1137101	ϴ	ϴ	ϴ

↑: increase; ↓: decrease; ϴ no association

Table 2.3 (continued) Genetic factors associated with NAFLD progression

Author	Gene	SNP	Steatosis	Inflammation	Fibrosis
Reactive oxidant species, or cytokine					
Zhou et al., 2010; Wong et al., 2008; Aller et al., 2010; Hu et al., 2009; Tokushige et al., 2007 & Valenti et al., 2002	*TNF-α*	rs180062	ϴϴϴ	↑ϴϴ	↑ϴ
Wong et al., 2008; Tokushige et al., 2007 & Valenti et al., 2002		rs361525	↑↑↑ϴ	ϴϴ	ϴ
Wong et al., 2008 & Tokushige et al., 2007		rs1800630	ϴ	↑	ϴ
Tokushige et al., 2007		rs1799964	ϴ	↑	
Tokushige et al., 2007		rs1799724	ϴ	ϴ	
Tokushige et al., 2007	*TNF-β*	rs909253	ϴ	ϴ	
Carulli et al., 2009	*IL-6*	rs1800795	↑	↑	
Brun et al., 2006	*CD14*	rs2569190	ϴ	↑	
Hashemi et al., 2011 & Oliveira et al., 2010	*GCLC*	rs4140528	ϴ	↑	
El-Koofy et al., 2011, Namikawa et al., 2004 & Al-Serri et al., 2011	*SOD2*	rs4880	↑	ϴ↑↑	↑
Hernaez et al., 2011	*HFE*	rs1800562	ϴ	ϴ	
		rs1799945			
Aller et al., 2010 & Labruna et al., 2009	*UCP*	rs1800849	↑	↑	
Cao et al., 2010	*PTGS2*	rs689466	↑		
Iwata et al., 2011	*ABCB11*	rs2287622	ϴ	ϴ	ϴ
Akyildiz et al., 2010	*MIF*	rs755622	ϴ	ϴ	ϴ
Piao et al., 2003 & Varela et al., 2008	*CYP2E1*	rs2031920	ϴ↓	ϴ	
Valenti et al., 2006	*SERPINA1*	rs28929474	↑	↑	
Sazci et al., 2008	*MTHFR*	rs1801131	↑	↑	
Nozaki et al., 2004	*IL1B*	rs16944	↑	↑	
Sookoian et al., 2009	*CFTR/MRP*	rs17222723	↑	ϴ	
Sookoian et al., 2008	*STAT3*	rs6503695	↑	ϴ	
Zain et al., 2013 & Yoneda et al., 2009	*AGTR1*	rs3772622	ϴ↑	ϴ↑	↑
Miele et al., 2008	*KLF6*	rs3750861	↑		↑

↑: increase; ↓: decrease; ϴ no association Adapted and modified from Hernaez (2012)

NASH patients with fibrosis showed an over-expression of fibrogenic genes, including the leptin receptor and NASH patients without fibrosis showed an over-expression of proinflammatory and proapoptotic genes. The upregulated gene expression of prohibitin suggests mitochondrial dysfunction in NASH patients (Cayon, Crespo, Guerra, & Pons-Romero, 2008). Metabolic factors but not ALT are associated with histological severity in NAFLD patients (V. W. Wong et al., 2009). Visceral adipose tissue, HOMA-IR and serum uric acid levels are the main determinants of NAFLD in obese patients (Fenkci, Rota, Sabir, & Akdag, 2007). NAFLD is associated with carotid artery wall thickness (Intima media thickness by USG) in diet controlled type 2 diabetes (Targher et al., 2006). Variation in the promoter region of fatty acid transport protein 5 (*FATP5*) is associated with hepatic steatosis, higher ALT level, insulin resistance and dyslipidemia (Auinger et al., 2010). Apolipoprotein A-V gene (*APOA5*) is associated with lipid profile in a number of studies (Can Demirdogen et al., 2012; Hishida et al., 2012; S. Rafiq et al., 2012). The -55CT polymorphism of the uncoupling protein 3 (*UCP3*) promotes increased in insulin resistance, moderate to severe steatosis, and inflammation (Aller et al., 2010). Insulin sensitivity was also affected by polymorphisms in the transferrin receptor gene (*TFRC*) (Fernandez-Real et al., 2010).

Romeo et al. (2008) carried out genome-wide association association scan for hepatic content. One nonsynonymous variant stood out from the rest, rs738409, which is responsible for the substitution of an isoleucine to methionine at position 148 in *PNPLA3*. This variant showed evidence of association with NAFLD. *PNPLA3* is expressed in adipose and liver and can both hydrolyze liver fat and transfer fatty acids between lipids. It encodes a transmembrane protein whose expression is upregulated during fat cell differentiation. The expression is also reported to be higher in adipose tissue in obese compared to lean individuals. However, the precise role of *PNPLA3* in

liver fat metabolism remains to be determined. Romeo et al. (2008) also observed association between *PNPLA3* rs738409 variant and liver enzymes in Hispanics. Individuals of Hispanic descent carry an increased risk of NAFLD (45%) compared to individuals of European American (33%) or African American (24%) descent. Hispanics are also closely related to high risk of steatohepatitis and cirrhosis. The study also found that the T allele that is associated with decreased risk of NAFLD, and is commonly found in the African-American group, a group with the lowest risk of developing NAFLD (Karlsen, 2009; Romeo et al., 2008). This finding was replicated by another study that found that the association of *PNPLA3* and NAFLD remains strong even after the adjustment of body mass index (BMI), age and sex (Kotronen et al., 2009). It is reported that the hepatic and adipose tissue insulin sensitivity decrease with increase of hepatic fat level independent of the *PNPLA3* genotype. This therefore indicates that the GG *PNPLA3* genotype has no influence on insulin sensitivity but increase hepatic fat content. The serum ALT level which is associated with NAFLD, is found to be higher in G allele carriers (Kotronen, Westerbacka, Bergholm, Pietilainen, & Yki-Jarvinen, 2007).

Table 2.4 Association studies between *PNPLA3* rs738409 and NAFLD

First Author, Year	Population	Study Design & Sample size (n)	Diagnosis method	Age	NAFLD	NAFLD spectrum
Romeo et al., 2008	Mixed: Hispanics, African-Americans, Caucasian	Population-based, n = 3450	H-MRS	Adult	Yes	No
Sookoian et al., 2009	Caucasian	Hospital-based, n =266	US and LB	Adult	Yes	SS, NASH
Kotronen et al., 2009	Finnish	Hospital-based, n =291	H-MRS	Adult	Yes	No
Valenti et al., 2010	Caucasian	Hospital-based, n = 149	LB	Pediatric	Yes	SS, NASH
Valenti et al., 2010	Caucasian	Hospital-based, n = 753	LB	Adult	Yes	SS, NASH
Romeo et al., 2010	Caucasian	Hospital-based, n = 678	Morbid obese	Adult	Yes	No
Romeo et al., 2010	Caucasian	Hospital-based, n = 475	US. Over-weight/ Obese	Pediatric	Yes	No
Rotman et al., 2010	Caucasian	Mixed: Population and hospital-based, n = 894	LB	Adult	Yes	SS, NASH
Speliotes et al., 2010	Caucasian	Mixed: Population and hospital-based, n = 2128	LB	Adult	Yes	SS, NASH
Goran et al., 2010	Hispanic	Population-based, n = 327	MRI	Pediatric	Yes	No
Davis et al., 2010	Hispanic	Population-based, n = 153	H-MRS	Pediatric	Yes	No
Petit et al., 2010	Caucasian	Hospital-based, n = 218	H-MRS/ T2D	Adult	Yes	No
Kantartzis et al., 2010	Caucasian	Hospital-based, n = 330	H-MRS/ T2D	Adult	Yes	No
Lin et al., 2010	Asian	Population-based, n = 520	US	Pediatric	Yes	No
Hotta et al., 2010	Asian	Hospital-based, n = 831	LB	Adult	Yes	SS, NASH
Wagenknecht et al., 2011	Mixed: Hispanic American and African American	Population-based, n = 1214	CT scan	Adult	Yes	No
Cox, 2011	African American	Hospital-based, n = 442	CT scan/ T2D	Adult	Yes	No
Zain, 2012	Asian mixed: Malay, Chinese, Indian	Hospital-based, n = 342	LB	Adult	Yes	SS, NASH, NASH severity

CT computed tomography, *H-MRS* hydrogen magnetic resonance spectroscopy, *LB* liver biopsy, *MRI* magnetic resonance imaging *NASH* non-alcoholic steatohepatitis, *SS* simple steatosis, *T2D* type 2 diabetes, *US* ultrasonography

2.6 Copy number variation

There are usually thousands of papers published every year on SNPs studies while less so can be found for CNVs. CNV studies are expensive and the characteristics of CNV with the likes of total number, genomic positions, gene content, frequency spectrum and patterns of linkage disequilibrium (LD) with one another remain unknown, which make the study capacity limited (K. K. Wong et al., 2007). While the frequencies of most SNPs have been determined in various populations, the situation is different for copy number variants. Up until now, the frequencies are not well defined. Some may report CNV as copy number polymorphism (CNP), however the appropriate term should be variant as the term "polymorphism" is commonly used to refer to genetic variants that have minor allele frequency of at least 1% in a given population. CNV is defined as copy number change involving 1 kb or more of DNA fragments (Hegele, 2007). Large CNVs of more than 1 Mb have also been reported and they tend to be localized at regions such as centromeres, telomeres and heterochromatin (Giglio et al., 2001). Most of CNVs in humans are inherited variations, but some are formed de novo at a significant rate, both in the germ line and in somatic cells (Hastings, Lupski, Rosenberg, & Ira, 2009).

CNVs have now become one of the important genetic variabilities studied after SNPs, as it was estimated that almost 68% of the CNV regions were found to overlap with almost 300 proven disease-causing genes (Sebat et al., 2004). These CNVs are also accounted for by variability in phenotype, metabolism and disease susceptibility. Interestingly, some of the regions were associated with human microRNAs, thus reflecting the importance of microRNAs in human diversity (K. K. Wong et al., 2007). Approximately 12% of the human genome, is accounted for copy number variation (Redon et al., 2006). Due to numbers of CNVs overlapped the genes, changes in copy number therefore affect the expression levels of the gene thereby causing phenotypic

variability among individuals. Redundancy which are caused by these additional copies of genes alter the gene functions while other copies maintain the original function (Hastings et al., 2009). Some CNVs are essential to provide phenotypic variations in individuals. Nevertheless, some are disadvantageous. Copy number changes have been reported in cancer progression (Volik et al., 2006) and contribute to cancer susceptibility (Frank et al., 2007).

Copy number change happens when two separated DNA sequences are joined together thereby causing a change in chromosomal structure. A high number of structural changes show recurrent end-points where the break-points at a given locus are confined to a few genomic positions, while others show non-recurrent endpoints. The recurrent CNVs often coincide with segmental duplications unlike the non-recurrent CNVs. Despite that, non-recurrent events tend to occur in the vicinity of regions that are rich in segmental duplications (Lee et al., 2006; Stankiewicz et al., 2003). While the recurrent CNVs are likely to be formed by homologous recombination, the non-recurrent CNVs are not, due to limited short homology of 2-15 bp. The chromosomal structural changes of non-recurrent events are far more complex, in which it can occur from insertion of short sequences from elsewhere or include a mixture of duplications, triplications, inversions and deletions (J. M. Chen, Chuzhanova, Stenson, Ferec, & Cooper, 2005; Gajecka et al., 2008).

New duplicates that are formed will increase the mutation rate as they served as a substrate for additional mutations (Schrider & Hahn, 2010). The duplication and/or deletion are the results of mutational mechanisms, which in humans consist of the most common non-allelic homologous recombination (NAHR), followed by non-homologous end joining (NHEJ) and replication slippage, and the least retrotransposition (Conrad et al., 2010; Kidd et al., 2008). NAHR is accounted for by the formation of the largest proportion of CNVs while replication slippage was suggested to be the simplest

mechanism that is responsible for the smaller variants such as tandem repeats. NAHR between homologous chromosomes or between sister chromatids results in both a duplication and a deletion while recombination on the same chromatid causes only deletion (Turner et al., 2008). Unlike duplication by NAHR which requires similar sequences, no sequence homology are required by NHEJ. CNVs formed by NHEJ gives both deletion and insertion of DNA (Hastings et al., 2009). Retrotransposition on the other hand, results only in new duplications (Conrad et al., 2010).

2.7 Disease-related copy number variations

CNV is one of the sources of genetic variation not only in the humans but also in the primates and rodents. The mechanism by which CNV influences gene expression is by altering the gene through disruption or duplication of CNVs regions-containing genes. Interindividual variation in drug response, immune defense and disease resistance are the results of some CNVs encompassing the genes. In a study of differences in CNVs of gene encoding chemokine (C-C motif) ligand 3-like 1 (*CCL3L1*), an immunodeficiency virus 1 (HIV-1)-suppressive chemokine and ligand for HIV receptor, it was reported that fewer numbers of the *CCL3L1* copies as compared to the average increases susceptibility to HIV/AIDS (Freeman et al., 2006). A region that encompasses the putative tumor-suppressor genes was shown to be deleted in lung cancer and it is actually associated with copy-number loss in the individuals (K. K. Wong et al., 2007).

The mechanisms by which the progression of heterozygous familial hypercholestrolemia (HeFH) occurs, normally involves small insertions or deletions, or missense, splicing, or nonsense mutations within the LDL receptor gene. However approximately one third of the patients are clinically diagnosed as HeFH without the presence of such mutations. CNVs in the low density lipoprotein receptor gene (*LDLR*) partially contribute to heterozygous familial hypercholesterolemia in affected patients.

There were overlaps observed between the reported CNVs and the genes involved in metabolic syndrome. About 22 or 40% of the 55 investigated genes contained a CNV or indel that have been observed in the general population (Lanktree & Hegele, 2008).

The above-mentioned studies clearly depict the significant role of CNVs in some severe diseases such as cancer and HIV/AIDS. The role of copy number changes in NAFLD is however not well defined. To date, there is no published report yet on CNV in NAFLD. The closest study was to report on the association of CNVs with hepatic biomarkers in a population of more than 8000 Korean individuals (Kim, Cho, Yu, Sung, & Kim, 2010). One hundred CNVs were singnificantly associated with AST and 16 with ALT. A total of 39 genes were identified within the regions. Among them was neuropeptide Y receptor type 5 (*NPY5R*) which was associated with dyslipidemia. In addition, more reports on CNVs-related metabolic traits such as obesity and diabetes are continuously documented.

In a recent animal study that involved a construction of mouse strains with a deletion or duplication, it was shown that a duplication of Dp(11)17 resulted in obesity-opposing characteristics, namely, reduced weight, decreased total and LDL cholesterol and increased insulin sensitivity, which is independent of food intake or level of activity. The Dp(11)17 also displayed protection against metabolic syndrome compared to the wild type. These findings help us to understand the underlying physiology of weight control and energy metabolism contributed by CNVs (Lacaria et al., 2012). A significant correlation was observed between copy number and expression of CNV genes modulating metabolic traits in mice such as body weight, triglycerides and glucose. These genes which include interlectin (*It1na*), defensin related cryptdin 1 (*Defcr-1*), tripartite motif-containing 12 (*TRIM12*) and tripartite motif-containing 34 (*TRIM34*) were found in three CNVs mapped in the genome in chromosomes 1, 4 and 17 (Orozco et al., 2009).

Combined analyses in human studies have identified three CNVs associated with BMI, obesity, and other obesity-related traits. Several genes such as parkinson protein 2 (*PARK2*), glycophorin A (*GYPA*), and sarcoglycan (*SGCZ*) genes were found to be encompassed by these CNVs (Zhao et al., 2012). In the same year, a deletion at position 16p12.3 was reported in both European and Chinese populations, in which higher deletion frequency was observed in the European than that of the Chinese. The deletion in these populations was significantly associated with obesity (T. L. Yang et al., 2012). A putative association between CNV regions with type 2 diabetes was demonstrated in a Korean study (Bae et al., 2011). Several SNP-associated CNV markers were found to be associated with fasting insulin and HOMA-IR (Irvin et al., 2011).

CHAPTER THREE
MATERIALS AND METHODS

3.1 Materials

3.1.1 Blood Collection

Blood collection tubes; lavender and yellow top, syringe and needle, tourniquet, alcohol swab, cotton wool, sterile plaster, disposable latex examination gloves

3.1.2 DNA extraction

Absolute molecular grade ethanol (Merck, Germany), QIAamp DNA Blood Mini Kit (QIAGEN, Hilden, Germany); 1.5ml microcentrifuge tubes

3.1.3 Agarose gel electrophoresis

GenAgarose (Genaxis BIOTECHNOLOGY, USA); 10X TBE buffer (BST, Malaysia); 6X Loading Dye Solution (Fermentas, Canada); GelRed (Biotium); 100bp DNA ladder (Vivantis, UK)

3.1.4 Real-time polymerase chain reaction (RT-PCR)

TaqMan assay (Applied Biosystems), master mix (Applied Biosystems), MicroAmp reaction tubes (Applied Biosystems), MicroAmp cap strip (Applied Biosystems)

3.1.5 Sequenom MassARRAY

96-well plate and plate sealing film (for quality control process). Experiment was run as a service by Chinese University of Hong Kong

3.1.6 Array CGH

1 x TE pH8.0, Molecular Grade (Promega), Oligo aCGH Hybridization Kit (OGT), Human Cot-1 DNA® (Life Technologies), Agilent Wash 1 and 2 (OGT), CytoSure genomic DNA labelling kit (OGT),

3.1.7 Instruments

BIO-RAD UV Transluminator; KODAK EDAS 290 camera; Gel Electrophoresis System with Power Pack (TRANSLAB, Malaysia & WEALTEC ELITE 300); Vortex machine, Microcentrifuge (Labnet, NJ, USA); Microwave (SAMSUNG, Korea); Freezer (Fisher & Paykel, Auckland); Refrigerator (TOSHIBA HYBRID PLASMA), Weighing Machine, Ice-maker Machine (Scotsman AF 80), micropipettes (20µL, 200µL and 1000µL); Step One Real-Time PCR (Applied Biosystems), Real-Time PCR software (Applied Biosystems),magnetic stirrer, SureHyb hybridisation cassette (Agilent), gasket slides (OGT), glass dishes and rack, forcep, vacuum dessicator (SpeedVac®), PTC-200 Thermal Cycler (MJ Research), microarray scanner (Agilent), CytoSure Interpret software (OGT)

3.2 Methods

3.2.1 Subject recruitment

The study involved a total of 144 patients and 198 controls among adults. Subjects were recruited from the University Malaya Medical Centre (UMMC). The study protocol was approved by the Medical Ethics Committee of UMMC and all subjects provided written informed consent to participate. Participants with the following criteria were excluded from the study; current or previous alcohol consumption of > 10g/day (Ruhl & Everhart, 2005), hepatitis B or C infection, autoimmune hepatitis, were taking drugs

that are known to cause steatosis. All cases were consecutive patients recruited from the adult hepatology outpatient clinic suspected to have NAFLD based on increased echogenicity (compared to renal cortex) on ultrasound with or without abnormal ALT. All NAFLD patients were biopsy-proven and the liver biopsy specimens were on an average 1.5 cm long and contained at least six portal tracts. The biopsy specimens were evaluated according to the NASH Clinical Research Network (Brunt, Kleiner, Wilson, Belt, & Neuschwander-Tetri, 2011; Kleiner et al., 2005) by a qualified pathologist blinded to patients, clinical and laboratory data. Patients with NAFLD were classified into simple steatosis and NASH. NASH patients were further classified into NASH without significant fibrosis (fibrosis score of < 2) and NASH with significant fibrosis (fibrosis score of 2 or greater). The control group without fatty liver consists of those who were genetically unrelated healthy subjects with normal body mass index (BMI) < 25 kg/m^2, normal fasting plasma glucose < 110 mg/dL and normal lipid profile. NAFLD was excluded in the controls by ultrasonography according to the absence of the following criteria; diffuse increase in bright echoes in the liver parenchyma (exceeding that of renal cortex and spleen) with impaired visualization of the peripheral portal and hepatic vein borders, and/or loss of definition of the diaphragm and poor delineation of the portal and hepatic radicles (Sanyal, 2002). For CNV study, 40 NASH patients were run as comparative array against 40 age, gender and ethnicity-matched controls using whole-genome aCGH microarrays.

3.2.2 Clinical and laboratory assessments

Standard protocol was used to determine the body mass index (BMI), waist circumference and blood pressure of the subjects. The biochemical tests for the determination of haemoglobin A1c (HbA1c), total cholesterol, triglycerides, high-density lipoprotein cholesterol (HDL), low-density lipoprotein cholesterol (LDL),

alanine transferase (ALT), aspartate aminotransferase (AST), and gamma-glutamyl transpeptidase (GGT) level in all subjects were as standard clinical laboratory methods carried out in an accredited laboratory at UMMC.

3.2.3 Sample preparation and DNA extraction

Blood samples obtained from the subjects were collected into EDTA tubes. The EDTA tubes were centrifuged at 900 x *g* (3000 rpm) for 10 minutes. The buffy coat layer were separated and transferred into 1.5 mL centrifuge tubes. Genomic DNA was extracted from the concentrated lymphocytes of the buffy coat using the QiAamp DNA Mini Kit (Qiagen. Hilden, German) according to the manufacturer's protocol. Twenty µL of Qiagen protease was added into a 1.5 mL microcentrifuge tube. A volume of 200 µL of the concentrated lymphocytes was then added into the microcentrifuge tube and followed by the addition of AL buffer. The mixture was mixed thoroughly by pulse vortex and then spun down for 15 seconds. The mixture-containing tube was incubated at 56^{o}C in a heating block for 10 minutes. After that, the mixture was spun down for 15 seconds to remove any water vapors or droplets formed on the tube lid. A volume of 200 µL of ethanol (96-100%) was then added. The mixture was mixed thoroughly and spun down for 15 seconds again. Next, the mixture was transferred to the QiAamp Mini spin column without wetting the rim. The column was centrifuged at 6000 x *g* (8000 rpm) for three minutes. The collection tube containing the filtrate was discarded and changed with a new one. A volume of 500 µL of AW1 buffer was added into the spin column. The column was centrifuged at 6000 x *g* (8000 rpm) for one minute. The filtrate-containing collection tube was discarded and changed with a new one. A volume of 500 µL of AW2 buffer was then added into the spin column and centrifuged at 20000 x *g* (14000 rpm) for two minutes. The filtrate inside the collection tube was discarded without changing the collection tube. The spin column was centrifuged at 8000 rpm for

drying. The column was placed in a new 1.5 mL microcentrifuge tube. A volume of 150 µL of AE buffer was added into the column and incubated for 10 minutes at room temperature. The column was then centrifuged at 6000 x *g* (8000 rpm) for one minute and the QiAamp spin column was discarded. The end product inside the microcentrifuge tube is the DNA. The product was run under 1% gel electrophoresis and the quality of the DNA was checked using nanodrop 2000c (Thermo Scientific).

3.2.4 Agarose gel electrophoresis

Gel electrophoresis was carried out to determine the presence of DNA. The genomic DNA extraction product was run under 1% agarose gel electrophoresis. The size of the casting tray determines the volume of gel to be made. A small gel tray of 30 mL was prepared by adding 0.3 g of agarose powder and 30 mL of 0.5x TBE buffer into a conical flask. The mixture was then heated in a microwave until the agarose powder was totally dissolved in the buffer. The mixture was then allowed to cool down. In the meantime, the casting tray and comb were assembled. The mixture was quickly poured onto the casting tray and left to solidify. After the gel was ready, the gel tray was moved into the gel tank filled with 0.5x TBE buffer. The buffer amount was ensured to adequately cover the surface of the gel. A volume of 1 µL of the genomic DNA was mixed with 1.5 µL of loading dye and then loaded into the well. The loading dye was first prepared by diluting a volume of 1 µL of GelRed into 599 µL of 6x loading dye. The sample was run together with a 100 bp ladder at 110 V for about 20 to 30 minutes. After that, the product was viewed under ultraviolet (UV) light using BIO-RAD UV Transluminator 2000. The gel picture was captured using KODAK EDAS 290 gel documentation system and the file was saved.

3.2.5 DNA measurement

The quality of DNA was checked each time using nanodrop 2000c (Thermo Scientific) to ensure that 260/280 and 260/230 absorbance ratios exceeded 1.8 to indicate high quality DNA.

3.2.6 Selection of SNPs

SNPs were selected based on findings from reported genome-wide association studies, meta-analysis, gene-linkage studies, population candidate gene studies and knowledge of tagging SNPs using International HapMap database (hapmap.ncbi.nlm.nih.gov/).

3.2.7 Genotyping

3.2.7.1 Real-time polymerase chain reaction (RT-PCR)

The extracted genomic DNA was used for the genotyping. The PNPLA3 rs738409 variant and AGTR1 variants (rs3772622, rs3772627, rs3772630, rs3772633, and rs2276736) were genotyped using a predesigned Taqman SNP genotyping assay (Applied Biosystems, Foster City, CA, USA) on a Step One Real-Time PCR instrument (Applied Biosystems). The total reaction volume for each well was 10 µl containing 5 µl universal mastermix (Applied Biosystems), 0.5 µl assay mix, 3.5 µl distilled water and 1 µl genomic DNA. The plate was setup at 95 ^{o}C holding stage for 20 s, 45 cycles of 95 ^{o}C denaturation for 3 s and 60 ^{o}C annealing for 20 s and ran on a fast reaction (approximately 40 minutes for each run). Negative controls were introduced for every run to ensure the genotyping quality.

3.2.7.2 Sequenom MassARRAY

Twenty-nine SNPs from various genes that included *LEPR*, *GCKR*, *iNOS*, *TRAIL*, *SREBF1*, *CLOCK*, *COL13A1*, *FABP2*, *UCP3*, *PPARG*, *COX-2*, *NR1I2*, *KLF6*, *TFRC*, *PPARGC1A*, *FATP5*, *STAT3*, *APOA5*, *PEMT*, and *FDFT1* were genotyped at the University of Hong Kong, Genome Research Centre using the Sequenom MassARRAY technology platform with the iPLEX GOLD chemistry (Sequenom, San Diego, CA) according to the manufacturer's protocols. Briefly, MassARRAY AssayDesign software package (v4.0) was used to design the specific assays with proximal SNPs filtering. Quality of the PCR fragment amplification and extension primer specificity was checked prior to running the reaction. A volume of 1 μL of genomic DNA at 10-20 ng/μL were used in each amplification reaction. Residual nucleotides were dephosphorylated prior the iPLEX Gold reaction. Following a single-base extension, reaction products were desalted with SpectroClean resin (Sequenom, San Diego, CA) and 10 nL was spotted onto the SpectroCHIP using the MassARRAY Nanodispenser. MassARRAY Analyzer Compact MALDI-TOF mass spectrometer was used to determine the mass. For proper data acquisition and analysis, the MassARRAY® Typer 4.0 software was used. Genotypes were called after cluster analysis using the default setting of Gaussian mixture model. Inspection of the clusters were done to ensure a clear cluster separation with good signal to noise cut-off. A manual review was done to further clarify uncertain genotype calls. Assay with less than 80% call rate within the same SpectroChip was considered to have failed. A blank and five duplicates were introduced as quality controls. SpectroChip with more than 25% call rate in the blank control or with less than 99.5% concordance in duplicate checks along with more than 10% call rate in blank check were considered to have failed and the genotyping would be required to be repeated.

3.2.8 Statistical analysis

All values are presented as mean ± standard deviation for continuous data and as percentages for categorical data. Prior to analyzing, all variables were tested for normality. For normally distributed variables (Age, BMI, waist circumference, ALT, HDL cholesterol, LDL cholesterol, total cholesterol, triglycerides, and systolic blood pressure), independent *t*-test was performed to determine the comparison between two groups. For skewed variables; HbA1c, AST, GGT, diastolic blood pressure, steatosis grade, lobular inflammation, ballooning, and fibrosis, Mann-Whitney U test was applied. Analysis of Variance (ANOVA) and Kruskal-Wallis test were conducted for comparison of three groups, for normally distributed and skewed variables, respectively. A goodness-of fit $\chi 2$ test was used to assess whether each individual variant was in equilibrium at each locus in the population (Hardy-Weinberg equilibirum). A *P* value of less than 0.05 indicated a lack of agreement with the equilibrium.

Association of allele was performed using logistic regression. Likelihood tests indicated a significant effect of ethnicity but no significant effect of age and gender on the association. Multiple logistic regression was used to adjust the potential effect of ethnicity on susceptibility to NAFLD. The calibration and fit of the model were assessed using Hosmer Lemeshow goodness of fit and receiver operating characteristic (ROC) curves.

Univariate analysis of association of genotypes and histological ordinal variables was assessed using the Jonckheere-Terpstra test. Ordinal regression is perfomed for multivariate analysis of histological ordinal variables. To correct for testing for multiple histological parameters, the false discovery rate (FDR) was calculated using Benjamini-Hochberg procedure (Benjamini and Hochberg, 1995). For easier interpretation of the results, logistic regression was used to obtain the odds ratio. Analyses were performed

using SPSS 16.0 (Chicago IL) with a two sided $P < 0.05$ considered to be statistically significant.

The Generalized Multifactor Dimensionality Reduction (GMDR) method was employed to investigate the influence of gene-gene interaction on NAFLD (Lou et al., 2007). A 10-fold cross validation with exhaustive search that considered all possible combinations was selected. GMDR provides a cross-validation consistency score which is a measure of the degree of consistency with which the selected interaction is identified as the best model among all possibilities considered. The testing balanced accuracy generated is a measure of the degree by which the interaction accurately predicts case-control status. Testing accuracy is a measure of the strength of gene-gene interaction with a power of 80% at an accuracy of 0.58-0.60, given a sample of < 500 (G. B. Chen et al., 2011). Ethnicity was first determined as a covariate in the analysis and used in the gene-gene interaction analysis. Permutation testing based on 1000 shuffles was performed to obtain empirical p-value. Logistic regression model was performed to confirm the results of gene interaction analyses.

Linkage disequilibrium and haplotype analyses for the five SNPs were performed using Haploview 4.2 program. The odds ratio for the haploypes was calculated using R software version 2.11.1.

An estimated sample size of 135 cases and controls would provide 80% power at an α of 0.05 with the following assumptions: the allele frequency ranged from 0.24-0.45, the baseline risk for the Malaysian population was 0.17, and the minimum detectable odds ratio was 2.0.

3.2.9 Whole genome array CGH for CNV detection

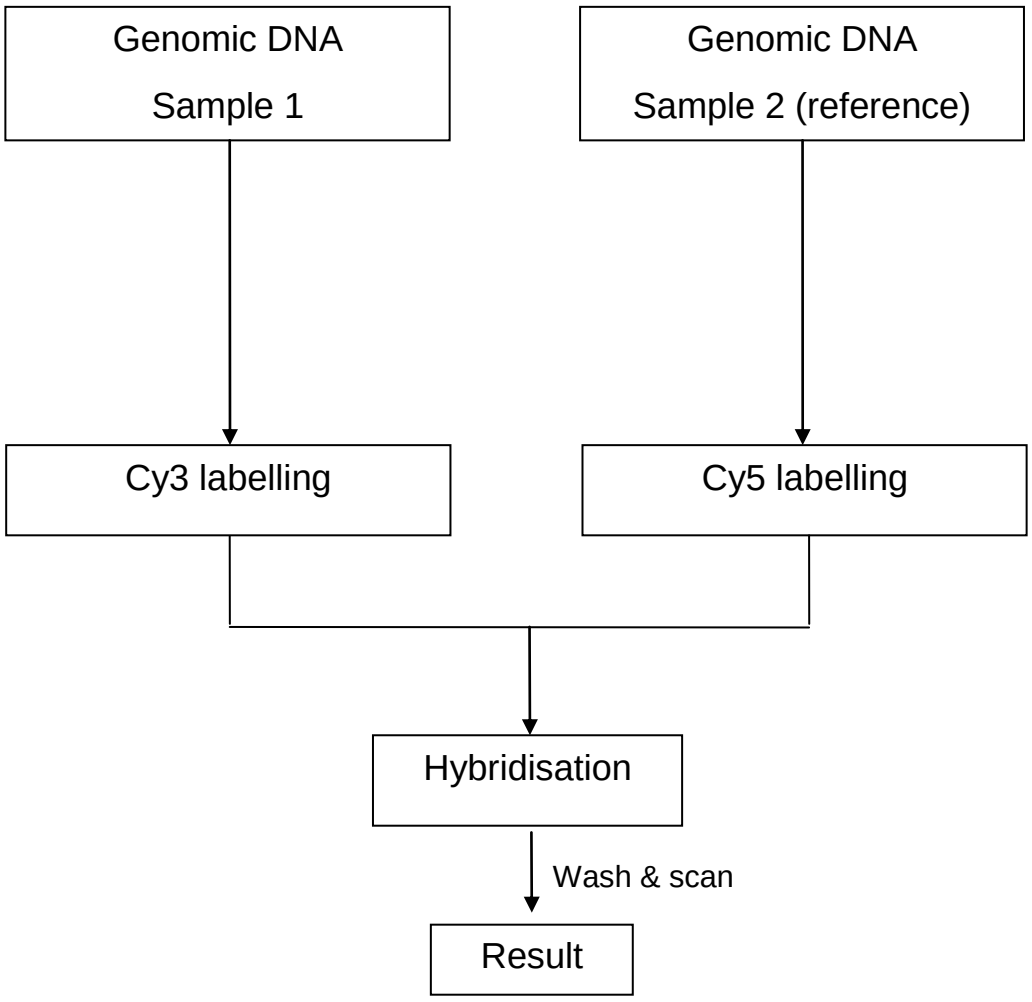


Figure 3.1 Overview of the array CGH procedure

3.2.9.1 Sample preparation for aCGH

Forty samples from patients with NASH were run as comparative array against 40 age, gender, and ethnicity-matched controls using whole-genome aCGH microarrays. The genomic DNA was prepared and purified to a standard in which the 260/280 absorbance ratio exceeds 1.8. Ideally the 260/230 absorbance ratio should also exceed 1.8. The final volume of the sample was 18 µl. At least 1.0 µg of DNA was required and similar starting amounts of sample and reference DNA were used.

3.2.9.2 Array CGH

Array-CGH was performed according to the protocol established by the manufacturer (Oxford Gene Technology). It was carried out using the SurePrint G3 Human CGH 2 x 400K array (Agilent Technologies, Santa Clare, CA, USA) for genome wide identification of putative disease-associated CNVs. Each oligonucleotide-based microarray slide contained 410,739 probes that enable profiling of molecular genomic imbalances with 5.3 kb average resolution. Each probe on the array is 60-mer in length and covers both non-coding and coding regions of the human genome. A total of 1.0 µg gDNA of patient's sample and 1.0 µg normal control gDNA were labeled with Cy3 and Cy5 dyes respectively using the CytoSure Genomic DNA labeling kit (Oxford Gene Technology, Oxford, UK). Probes were then purified using Microcon Centrifugation Filters, Ultracel YM-30 (Millipore, Billerica, MA, USA) and mixed thoroughly. This was followed by denaturation and pre-annealing with 50 µg of human Cot-1 DNA (Invitrogen, California). Hybridization of the mixture to the array slide was executed at a constant rotation at 65°C for 40 hours. The slide was then washed with Agilent wash buffer 1 and 2, and scanned immediately using an Agilent Microarray scanner (Agilent Technologies, Santa Clara, CA, USA). Data was extracted from scanned images using the Feature Extraction Software, version 10.7.3.1 (Agilent Technologies, USA). The

raw data obtained thereafter was uploaded into the CytoSure™ Interpret software version 4.2.5 (Oxford Gene Technology, Oxford, UK), normalized and converted into .cgh files. Data normalization software was used to improve the inconsistencies in dye incorporation. The data was segmented using a modified Circular Binary Segmentation (CBS) algorithm (Venkatraman & Olshen, 2007). Genomic aberrations were identified by applying a threshold of log2 ratio value of 0.3 for gains and 0.6 for losses. Chromosomal aberrations were reported in accordance to the human genome sequence assembly Build 37, Hg 19 (www.ncbi.nlm.nih.gov).

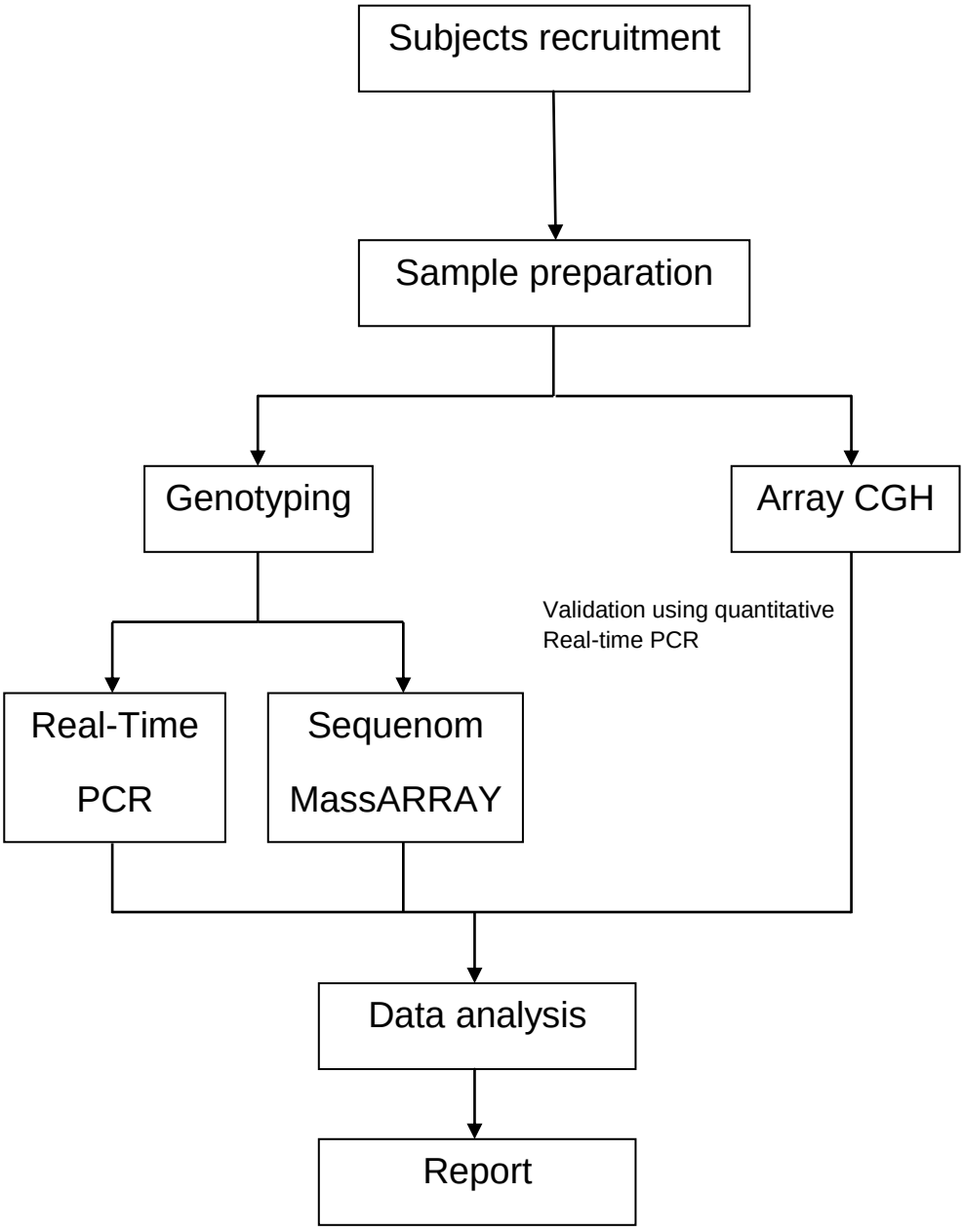


Figure 3.2 Summary of the methods

CHAPTER FOUR
RESULTS

4.1 Demographics and clinical data

The demographics and clinical characteristics of the NAFLD and control groups are described in Table 4.1. The differences between the two groups with regard to BMI, HbA1c, LDL cholesterol, total cholesterol, triglycerides, AST, ALT and GGT reflect the nature of the selection of the cases and controls.

Table 4.1 Demographics and clinical data of the subjects

Characteristics	*n* (%) or mean ± SD		
	Control (*n* = 198)	NAFLD (*n* = 144)	*P* value
Gender			0.084
Males	85 (43)	77 (53)	–
Females	113 (57)	67 (47)	–
Ethnicity			0.044
Malays	80 (41)	59 (41)	–
Chinese	54 (27)	54 (38)	–
Indians	64 (32)	31 (21)	–
Age (years)	53.1 ± 11.5	51.2 ± 12.0	0.136
BMI (kg/m^2)	22.7 ± 2.6	28.7 ± 4.4	<0.0001
HbA1c (%)	5.7 ± 0.8	6.6 ± 1.7	<0.0001
HDL cholesterol (mg/dl)	49.5 ± 12.9	48.5 ± 12.7	0.310
LDL cholesterol (mg/dl)	89.5 ± 22.4	117.1 ± 40.0	<0.0001
Total cholesterol (mg/dl)	176.4 ± 26.9	196.7 ± 44.0	<0.0001
Triglycerides (mg/dl)	118.4 ± 32.3	155.0 ± 62.7	<0.0001
AST (IU/L)	21.8 ± 9.5	42.9 ± 25.4	<0.0001
ALT (IU/L)	36.0 ± 16.6	83.0 ± 48.5	<0.0001
GGT (IU/L)	44.0 ± 25.4	111.6 ± 115.5	<0.0001

Data are expressed as mean ± SD for continuous data and as percentage for categorical data. *P* values obtained using Mann-Whitney U test except gender used Chi-square test.

ALT alanine transferase, *AST* aspartate aminotransferase, *BMI* body mass index, *GGT* gamma glutamyl transpeptidase, *HbA1c* haemoglobin A1c, *HDL* high-density lipoprotein, *LDL* low-density lipoprotein, *NAFLD* non-alcoholic fatty liver disease

Table 4.2 shows the demographics and clinical data of the patients with simple steatosis and with NASH. The BMI, HbA1c, waist circumference, triglycerides, systolic and diastolic blood pressure were significantly higher in the NASH group ($P < 0.05$) as compared to the simple steatosis group.

Table 4.2 Demographics and clinical data of the NAFLD patients

Characteristics	*n* (%) or mean ± SD		
	Simple steatosis (*n* = 33)	NASH (*n* = 111)	*P* value
Gender, n (%)			0.461
Males	20 (61)	57 (51)	–
Females	13 (39)	54 (49)	–
Age (years)	50.7 ± 11.8	51.2 ± 12.1	0.820
BMI (kg/m^2)	26.7 ± 3.9	29.2 ± 4.4	0.003
HbA1c (%)*	6.1 ± 1.3	6.7 ± 1.7	0.021
Waist circumference (cm)	89.2 ± 11.2	95.2 ± 10.4	0.005
HDL cholesterol (mg/dl)	50.2 ± 15.1	48.0 ± 11.9	0.391
LDL cholesterol (mg/dl)	114.7 ± 42.8	117.8 ± 39.3	0.698
Total cholesterol (mg/dl)	191.6 ± 44.9	198.2 ± 43.8	0.448
Triglycerides (mg/dl)	124.6 ± 42.0	164.1 ± 65.1	0.001
AST (IU/L)*	37.6 ± 21.3	44.5 ± 26.3	0.139
ALT (IU/L)	71.9 ± 50.0	86.4 ± 47.8	0.134
GGT (IU/L)*	99.4 ± 106.6	115.3 ± 118.2	0.132
Systolic blood pressure (mmHg)	125.2 ± 13.0	134.2 ± 14.1	0.001
Diastolic blood pressure (mmHg)*	78.2 ± 9.1	83.8 ± 9.7	0.003

Data are expressed as mean ± SD for continuous data and as percentage for categorical data. **P*-values obtained using Mann-Whitney U test, all other comparisons used independent *t*-test. Gender used Chi-square test.

ALT alanine transferase, *AST* aspartate aminotransferase, *BMI* body mass index, *GGT* gamma glutamyl transpeptidase, *HbA1c* haemoglobin A1c, *HDL* high-density lipoprotein, *LDL* low-density lipoprotein, *NAFLD* non-alcoholic fatty liver disease, *NASH* non-alcoholic steatohepatitis

4.2 Single nucleotide polymorphism

4.2.1 Single nucleotide polymorphism: *PNPLA3*

4.2.1.1 Genotypes and allele frequencies of *PNPLA3* rs738409 polymorphism

The association tests of *PNPLA3* rs738409 for the different NAFLD spectrum are presented in Table 4.3. The frequency of the risk allele G is significantly higher in the NASH (48%) and in the NAFLD (45%) when each is compared to the control group (24%). The genotype distribution was in Hardy-Weinberg equilibrium for all the patients and controls of each ethnic subgroup.

There was a significant association between the risk allele G and NAFLD in the pooled subjects (OR 2.23, 95% CI 1.60-3.11, $P < 0.0001$) as well as after stratification into the three ethnic subgroups, (OR 1.93, 95% CI 1.11-3.36, $P = 0.02$; OR 3.50, 95% CI 1.67-7.33, $P = 0.001$ and OR 2.00, 95% CI 1.21-3.21, $P = 0.007$, for Chinese, Indian and Malay, respectively). The power for the significant associations is shown to be 99%, 89%, 72%, 96%, 99%, 77%, 99%, and 63%, for rs738409 with NAFLD, rs738409 with NAFLD in Malay, rs738409 with NAFLD in Chinese, rs738409 with NAFLD in Indian, rs738409 with NASH, rs738409 with NASH without significant fibrosis, rs738409 with NASH with significant fibrosis, and rs738409 with NASH with significant fibrosis compared to simple steatosis, respectively.

PNPLA3 gene polymorphism at position 148 influences insulin secretion levels and obesity. Thus, adjustment for type 2 diabetes and BMI was performed. However, the results remain significant (OR 2.29, 95% CI 1.63-3.23, $P < 0.0001$ and OR 2.19, 95% CI 1.42-3.37, $P < 0.0001$, for type 2 diabetes and obesity, respectively)

Table 4.3 Association tests of *PNPLA3* rs738409 in different NAFLD spectrum

NAFLD spectrum	G allele frequency	*P* value	OR (CI)
Control as reference			
NAFLD vs. control			
Overall	0.45 vs. 0.24	<0.0001	2.23 (1.60-3.11)*
Malays	0.42 vs. 0.24	0.007	2.00 (1.21-3.21)
Chinese	0.49 vs. 0.32	0.020	1.93 (1.11-3.36)
Indians	0.44 vs. 0.18	0.001	3.50 (1.67-7.33)
Simple steatosis vs. control	0.35 vs. 0.24	0.272	1.37 (0.78-2.42)*
All NASH vs. control	0.48 vs. 0.24	<0.0001	2.58 (1.80-3.69)*
NASH[a] vs. control	0.39 vs. 0.24	0.023	1.86 (1.09-3.18)*
NASH[b] vs. control	0.52 vs. 0.24	<0.0001	2.99 (1.99-4.50)*
Simple steatosis as reference			
All NASH vs. simple steatosis	0.48 vs. 0.35	0.083	1.63 (0.94-2.81)*
NASH[a] vs. simple steatosis	0.39 vs. 0.35	0.600	1.20 (0.61-2.38)*
NASH[b] vs. simple steatosis	0.52 vs. 0.35	0.035	1.85 (1.05-3.26)*
NASHa as reference			
NASH[b] vs NASH[a]	0.52 vs 0.39	0.088	1.62 (0.93-2.81)*

**P* values adjusted for ethnicity

CI confident interval, *OR* odds ratio, *NAFLD* non-alcoholic fatty liver disease, *NASH* non-alcoholic steatohepatitis, *NASH[a]* non-alcoholic steatohepatitis without significant fibrosis, *NASH[b]* non-alcoholic steatohepatitis with significant fibrosis

4.2.1.2 *PNPLA3* rs738409 polymorphism and liver histology

The NAFLD patients were then divided into simple steatosis and NASH group, following which association analysis was performed (Table 4.3). The G allele is shown to be associated with NASH (OR 2.64, 95% CI 1.85-3.75, $P < 0.0001$) but not with simple steatosis. NASH patients were further stratified into two groups; NASH without significant fibrosis (fibrosis score < 2) and NASH with significant fibrosis (fibrosis score ≥ 2) and went on to compare these two groups of patients with that of the simple steatosis group. Results revealed that the G allele frequency was significantly higher in the NASH group with significant fibrosis compared to group with simple steatosis (OR 1.85, 95% CI 1.05-3.26, $P = 0.035$), but the G allele frequency was similar between the NASH group without significant fibrosis and the simple steatosis group, suggesting that the G allele is associated with severe stage of NASH.

Investigation was carried out on the association of *PNPLA3* rs738409 with the histological features of NAFLD; steatosis, lobular inflammation, hepatocellular ballooning and fibrosis. It was shown that there is a significant association between the GG genotype with fibrosis stage (P = 0.038) but not with the other histological features (patients with genotype GG compared to CC, had mean score 2.03 vs. 1.52) (Table 4.4).

Table 4.4 Comparison of various clinical and histological parameters between the *PNPLA3* rs738409 genotypes among NAFLD patients

Characteristics	NAFLD, *n* = 144 (mean ± SD)			
	CC (*n* = 48)	CG (*n* = 63)	GG (*n* = 33)	*P* value
Age (years)	50.9 ± 12.2	52.1 ± 12.0	49.5± 12.1	0.585
BMI (kg/m^2)	28.8 ± 4.4	28.5 ± 4.7	28.7 ± 3.8	0.936
HbA1c (%)*	6.5 ± 1.8	6.6 ± 1.6	6.6 ± 1.6	0.759
Waist circumference (cm)	94.3 ± 10.2	93.9 ± 12.0	92.9 ± 9.8	0.838
HDL cholesterol (mg/dl)	49.4 ± 13.91	47.7 ± 13.6	48.7 ± 8.6	0.768
LDL cholesterol (mg/dl)	117.9 ± 37.1	119.9 ± 40.2	110.7 ± 44.1	0.558
Total cholesterol (mg/dl)	203.9 ± 41.0	195.0 ± 44.3	189.6 ± 47.3	0.327
Triglycerides (mg/dl)	170.3 ± 81.8	155.0 ± 51.5	132.8 ± 41.8	0.029
AST (IU/L)*	43.8 ± 25.1	40.7 ± 23.2	45.9 ± 29.8	0.858
ALT (IU/L)	86.1 ± 50.3	79.2 ± 40.8	86.0 ± 59.2	0.705
GGT (IU/L)*	121.0 ± 117.7	106.4 ± 114.9	107.9 ± 116.2	0.276
Systolic BP (mmHg)	130.4 ± 13.3	133.4 ± 13.9	132.3 ± 16.6	0.548
Diastolic BP (mmHg)*	81.5 ± 10.2	83.4 ± 8.8	82.2 ± 11.3	0.477
Steatosis grade*	1.8 ± 0.8	1.9 ± 0.8	1.7 ± 0.8	0.434
Lobular inflammation*	1.2 ± 0.6	1.3 ± 0.6	1.4 ± 0.6	0.472
Ballooning*	1.1 ± 0.6	1.2 ± 0.6	1.3 ± 0.6	0.396
Fibrosis*	1.5 ± 1.0	1.7 ± 1.0	2.0 ± 0.9	0.038

**P* values obtained using Kruskal-Wallis test, all other comparisons used ANOVA

ALT alanine transferase, *AST* aspartate aminotransferase, *BMI* body mass index, *GGT* gamma glutamyl transpeptidase, *HbA1c* haemoglobin A1c, *HDL* high-density lipoprotein, *LDL* low-density lipoprotein, *NAFLD* non-alcoholic fatty liver disease

Investigation on the magnitude of the effect of the risk allele G with the histological features was further carried out. Patients bearing at least a single risk alle G of rs738409 were more likely to have a fibrosis score ≥ 2 (OR 1.95, 95% CI 1.17-3.26, P = 0.013). The association remains significant after the adjustment of other histological

features (P = 0.012). However, there was no association between the risk allele and steatosis, lobular inflammation or hepatocellular ballooning (Table 4.5).

Table 4.5 Association of G allele of rs738409 with histological features in NAFLD patients

Histology	Univariate *P* value[a]	multivariate *p* value[b] (FDR *q* value[g])	OR (95% CI)
Steatosis			
> 33% vs. < 33%	0.91	0.892 (0.446)	0.88 (0.52-1.47)[c]
Lobular inflammation			
≥ 2 foci vs. < 2 foci	0.237	0.202 (0.135)	1.13 (0.64-1.99)[d]
Hepatocellular ballooning			
≥ 1 vs. < 1	0.176	0.168 (0.168)	1.07 (0.61-1.87)[e]
Fibrosis			
≥ 2 vs. < 2	0.013	0.012 (0.024)	1.95 (1.17-3.26)[f]

OR odds ratio, *CI* confident interval
[a]Jonckheere-Terpstra test
[b]Ordinal regression
[c, d, e, f]Multivariate logistic regression
[g]False discovery rate, $q < 0.05$ is significant

4.2.1.3 Analysis of various clinical parameters with *PNPLA3* rs738409 genotypes

Next, the clinical parameters between the *PNPLA3* rs738409 genotypes among patients with NAFLD were compared (Table 4.4). Patients with the GG genotype have a significantly lower level of plasma triglycerides compared to those with the CC genotype (P = 0.029). The levels of triglyceride between the three groups: simple steatosis, NASH without significant fibrosis and NASH with significant fibrosis, were then compared (Figure 4.1). A significant difference between the three groups (P = 0.002) was observed. Compared to those with simple steatosis, the triglycerides level is significantly higher in the NASH group with significant fibrosis (P = 0.006) and without significant fibrosis ($P < 0.0001$). The triglyceride level is lower in the NASH group with significant fibrosis compared to the NASH group without significant fibrosis, but this did not achieve statistical significance.

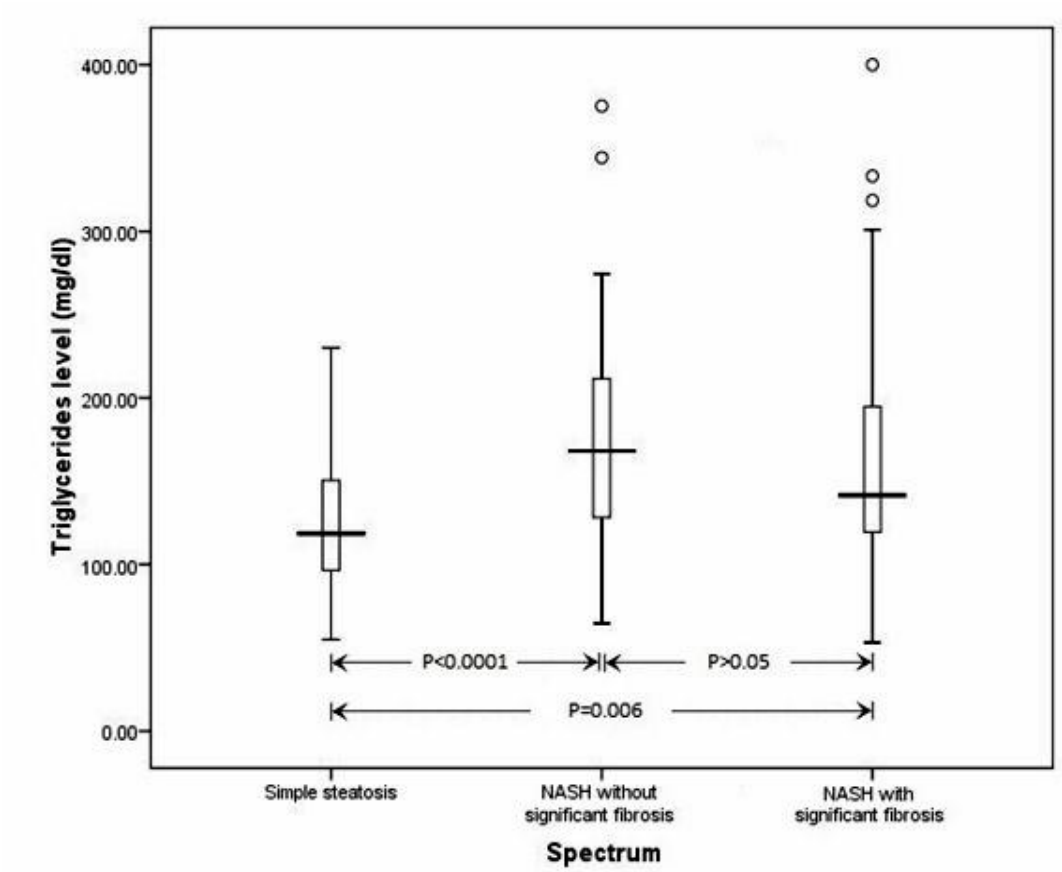


Figure 4.1 The comparison of plasma triglycerides level in the NAFLD spectrum

The boxplot showed the extreme values, 25th, 50th (median) and 75th percentiles and outliers

4.2.2 Single nucleotide polymorphism: *AGTR1*

4.2.2.1 Genotypes and allele frequencies of *AGTR1* polymorphisms

The genotypes of each SNP were in Hardy-Weinberg equilibrium for both NAFLD cases and controls, for pooled subjects as well as after stratification by ethnicity. The association tests between NAFLD and control subjects are shown in Table 4.6. None of the SNPs were associated with susceptibility to NAFLD. However, after ethnic stratification, in the Indian ethnic subgroup, the rs2276736, rs3772630 and rs3772627 were found to be protective against NAFLD (OR 0.40, 95% CI 0.20-0.81, P = 0.010; OR 0.43, 95% CI 0.22-0.86, P = 0.016; and OR 0.46, 95% CI 0.23-0.91, P = 0.026, respectively). The sample power for the significant findings is shown to be > 81%.

Next, the NAFLD group was divided into simple steatosis and NASH, however, no significant association was found between the five SNPs with susceptibility to NASH in the pooled subjects (Table 4.6). Results by ethnicity group showed that in the Indian ethnic subgroup, the three SNPs (rs2276736, rs3772630 and rs3772627) were also protective against NASH (OR 0.42, 95% CI 0.21-0.86, P = 0.017; OR 0.46, 95% CI 0.22-0.92, P = 0.029; and OR 0.49, 95% CI 0.24-0.98, P = 0.045, respectively).

Table 4.6 Association tests of *AGTR1* SNPs and NAFLD

NAFLD spectrum	All ethnicities*			Malay			Chinese			Indian		
	MAF	OR (CI)	*P*	MAF	OR (CI)	*P*	MAF	OR (CI)	*P*	MAF	OR (CI)	*P*
rs2276736 (T > C)												
Control as reference	0.42	1.00		0.35	1.00		0.43	1.00		0.49	1.00	
NAFLD	0.38	0.82 (0.48-1.37)	0.445	0.36	1.03 (0.61-1.15)	0.917	0.44	1.08 (0.62-1.88)	0.778	0.27	0.40 (0.20-0.81)	0.010
Simple steatosis	0.41	0.99 (0.55-1.79)	0.968	0.33	0.93 (0.37-2.31)	0.873	0.50	1.35 (0.62-2.93)	0.450	0.25	0.35 (0.07-1.81)	0.209
NASH	0.36	0.78 (0.46-1.34)	0.366	0.36	1.05 (0.61-1.81)	0.848	0.42	0.97 (0.53-1.78)	0.925	0.30	0.42 (0.21-0.86)	0.017
rs3772622 (A > G)												
Control as reference	0.38	1.00		0.38	1.00		0.41	1.00		0.34	1.00	
NAFLD	0.42	1.17 (0.87-1.59)	0.301	0.45	1.30 (0.81-2.07)	0.275	0.40	0.97 (0.58-1.61)	0.896	0.40	1.32 (0.69-2.53)	0.409
Simple steatosis	0.38	0.97 (0.57-1.65)	0.912	0.33	0.83 (0.35-1.97)	0.666	0.38	0.91 (0.43-1.94)	0.804	0.5	1.96 (0.45-8.55)	0.371
NASH	0.43	1.22 (0.88-1.69)	0.226	0.48	1.46 (0.88-2.42)	0.142	0.41	0.99 (0.58-1.71)	0.981	0.39	1.24 (0.62-2.46)	0.546
rs3772633 (A > G)												
Control as reference	0.23	1.00		0.23	1.00		0.30	1.00		0.19	1.00	
NAFLD	0.28	1.21 (0.86-1.70)	0.282	0.28	1.28 (0.75-2.18)	0.370	0.34	1.23 (0.70-2.16)	0.475	0.19	1.04 (0.49-2.21)	0.922
Simple steatosis	0.24	0.99 (0.53-1.84)	0.965	0.17	0.66 (0.21-2.08)	0.473	0.35	1.27 (0.58-2.80)	0.550	0.12	0.61 (0.07-5.34)	0.656
NASH	0.30	1.31 (0.91-1.88)	0.150	0.31	1.45 (0.83-2.55)	0.190	0.34	1.20 (0.65-2.23)	0.564	0.22	1.23 (0.57-2.66)	0.597

* Results based on combining results across ethnicities

CI confident interval, *MAF* minor allele frequency, *NAFLD* non-alcoholic fatty liver disease, *NASH* non-alcoholic steatohepatitis, *OR* odds ratio

Table 4.6 (continued) Association tests of *AGTR1* SNPs and NAFLD

NAFLD spectrum	All ethnicities*			Malay			Chinese			Indian		
	MAF	OR (CI)	*P*	MAF	OR (CI)	*P*	MAF	OR (CI)	*P*	MAF	OR (CI)	*P*
rs3772630 (A > G)												
Control as reference	0.42	1.00		0.35	1.00		0.44	1.00		0.49	1.00	
NAFLD	0.39	1.03 (0.76-1.38)	0.859	0.37	1.10 (0.67-1.81)	0.695	0.45	1.08 (0.62-1.88)	0.779	0.31	0.43 (0.22-0.86)	0.016
Simple steatosis	0.42	1.01 (0.73-1.39)	0.951	0.38	1.11 (0.46-2.67)	0.813	0.50	1.31 (0.60-2.87)	0.504	0.25	0.35 (0.07-1.81)	0.209
NASH	0.38	1.09 (0.67-1.78)	0.728	0.37	1.10 (0.65-1.88)	0.719	0.43	0.99 (0.54-1.80)	0.971	0.31	0.46 (0.22-0.92)	0.029
rs3772627 (T > C)												
Control as reference	0.42	1.00		0.38	1.00		0.44	1.00		0.48	1.00	
NAFLD	0.38	0.62 (0.33-1.15)	0.128	0.36	0.92 (0.56-1.52)	0.742	0.45	1.08 (0.63-1.85)	0.783	0.31	0.46 (0.23-0.91)	0.026
Simple steatosis	0.41	0.63 (0.38-1.04)	0.073	0.33	0.84 (0.34-2.06)	0.697	0.50	1.29 (0.60-2.76)	0.516	0.25	0.37 (0.07-1.92)	0.236
NASH	0.37	0.71 (0.28-1.82)	0.473	0.36	0.94 (0.55-1.61)	0.831	0.43	0.99 (0.55-1.78)	0.971	0.31	0.49 (0.24-0.98)	0.045

* Results based on combining results across ethnicities
CI confident interval, *MAF* minor allele frequency, *NAFLD* non-alcoholic fatty liver disease, *NASH* non-alcoholic steatohepatitis, *OR* odds ratio

Among the three protective SNPs, rs2276736 showed the lowest *P* value (OR 0.40, 95% CI 0.20-0.81, *P* = 0.010). To investigate whether the genotypes of SNP rs2276736 were associated with the clinical and histological parameters, the age, BMI, HbA1c, waist circumference, HDL cholesterol, LDL cholesterol, total cholesterol, triglycerides, AST, ALT, GGT, systolic blood pressure, diastolic blood pressure, steatosis, lobular inflammation, ballooning and fibrosis grade were compared between the different genotypes (Table 4.7). None of the histological parameters was significantly different among the genotypes. However, both BMI and waist circumference were found to be lower in the CC genotypes, further strengthening the case for the protective effect of the SNP.

Table 4.7 Comparison of various clinical and histological parameters between the *AGTR1* rs2276736 genotypes among NAFLD patients

Characteristics	NAFLD, *n* = 144 (mean ± SD)			
	TT (*n* = 51)	TC (*n* = 78)	CC (*n* = 15)	*P* value
Age (years)	51.5 ± 11.4	51.3 ± 11.8	48.7 ± 15.4	0.718
BMI (kg/m^2)	29.6 ± 4.5	28.6 ± 4.3	26.0 ± 3.6	0.017
HbA1c (%)*	6.6 ± 1.7	6.7 ± 1.8	5.9 ± 1.0	0.385
Waist circumference (cm)	96.0 ± 9.7	93.8 ± 11.1	86.8 ± 11.4	0.016
HDL cholesterol (mg/dl)	48.7 ± 12.0	48.7 ± 13.9	47.2 ± 8.1	0.917
LDL cholesterol (mg/dl)	115.5 ± 39.7	116.6 ± 41.1	125.2 ± 36.9	0.706
Total cholesterol (mg/dl)	194.3 ± 45.3	197.3 ± 44.1	202.1 ± 41.1	0.827
Triglycerides (mg/dl)	163.5 ± 68.7	152.2 ± 58.5	140.9 ± 62.7	0.399
AST (IU/L)*	42.9 ± 26.0	42.7 ± 25.4	43.7 ± 24.7	0.976
ALT (IU/L)	80.5 ± 48.5	83.1 ± 48.2	91.3 ± 52.4	0.753
GGT (IU/L)*	122.3 ± 123.9	106.0 ± 116.1	105.7 ± 81.3	0.424
Systolic BP (mmHg)	133.3 ± 15.1	130.7 ± 14.4	130.3 ± 18.1	0.589
Diastolic BP (mmHg)*	83.6 ± 9.0	82.4 ± 10.2	79.7 ± 10.5	0.328
Steatosis grade*	1.8 ± 0.7	1.8 ± 0.8	2.2 ± 0.9	0.148
Lobular inflammation*	1.2 ± 0.5	1.3 ± 0.6	1.5 ± 0.6	0.229
Ballooning*	1.2 ± 0.6	1.2 ± 0.6	1.3 ± 0.7	0.802
Fibrosis*	1.8 ± 0.9	1.7 ± 1.0	1.6 ± 1.1	0.624

**P* values obtained using Kruskal-Wallis test, all other comparisons used ANOVA

ALT alanine transferase, *AST* aspartate aminotransferase, *BMI* body mass index, *GGT* gamma glutamyl transpeptidase, *HbA1c* haemoglobin A1c, *HDL* high-density lipoprotein, *LDL* low-density lipoprotein, *NAFLD* non-alcoholic fatty liver disease

4.2.2.2 *AGTR1* polymorphisms and liver histology

Of the five SNPs, rs3772622 was the only SNP to show association with liver histology in the NAFLD patients (Table 4.8). The GG genotype of this SNP is significantly associated with increased fibrosis (P = 0.009) (Figure 4.2). It has to be noted that the rs3772622 is not among the SNPs that have been found to be protective. To investigate whether the G allele of the rs3772622 is associated with risk of fibrosis, univariate and multivariate tests were performed. The G allele of this SNP was shown to be associated with presence of fibrosis (P = 0.003). Subjects with the G allele has a propensity of 2.18 times to develop fibrosis score of ≥ 2. The association remains significant (P = 0.003) after adjustment of other histological attributes (Table 4.8). Investigation of the remaining SNPs and their association with histological features of NAFLD were also performed, however none of them showed a significant association.

Table 4.8 Association of G allele of rs3772622 with histological features in NAFLD patients

Histology	Univariate *P* value[a]	multivariate *p* value[b] (FDR *q* value[g])	OR (95% CI)
Steatosis			
> 33% vs. < 33%	0. 589	0.546 (0.728)	0.98 (0.61-1.56)[c]
Lobular inflammation			
≥ 2 foci vs. < 2 foci	0. 953	0.251 (0.502)	0.89 (0.55-1.45)[d]
Hepatocellular ballooning			
≥ 1 vs. < 1	0. 395	0.744 (0.744)	2.08 (0.84-5.14)[e]
Fibrosis			
≥ 2 vs. < 2	0. 003	0.003 (0.012)	2.18 (1.32-3.60)[f]

OR odds ratio, *CI* confident interval
[a]Jonckheere-Terpstra test
[b]Ordinal regression
[c, d, e, f]Multivariate logistic regression
[g]False discovery rate, $q < 0.05$ is significant

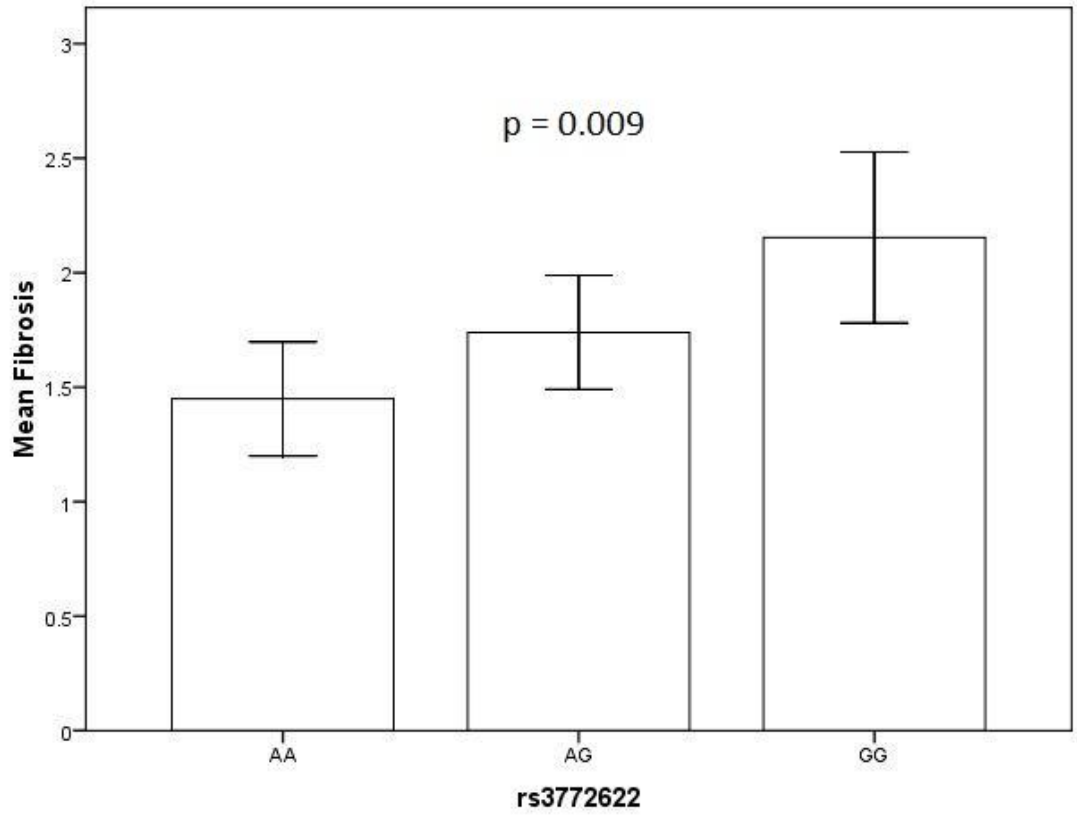


Figure 4.2 Mean fibrosis score among the genotypes of rs3772622

P value calculated from comparison between three groups using Kruskal-Wallis test

4.2.2.3 Linkage disequilibrium and haplotype analysis of *AGTR1* polymorphisms

Linkage disequilibrium (LD) analysis revealed a strong LD between the five SNPs in pooled subjects and even after ethnic stratification, except between rs3772633 and rs3772622. Variants rs2276736 and rs3772630 are in complete LD in the Chinese and Malays, however variants rs2276736 and rs3772627 are in complete LD in the Chinese only (Figure 4.3).

Four haplotypes with frequencies of above 5% are presented in Table 4.9. In pooled subjects, haplotypes ATATG, GCGCA and ATATA are more frequent than the others (35.1%, 21.8% and 20.9%, respectively). Similar results are observed in the Malays (37.8%, 22.7% and 23.2%, respectively). Meanwhile in the Chinese, the most frequent haplotypes are ATATG (34.8%) and GCGCA (25.9%). The Indians on the

other hand has high frequency of haplotypes ATATG (31.4%), ATATA (22.3%) and ACGCA (24.7%). Each of the following haplotypes, ATATG, GCGCA, ATATA and ACGCA is significantly different among the ethnic groups (P < 0.0001, for each haplotype respectively). Haplotype analysis revealed that haplotype ACGCA was protective against NAFLD in the Indians (OR 0.42, 95% CI 0.15-0.95, P = 0.03). Since rs2276736, rs3772630 and rs3772627 are presented in similar structure and some were in complete LD, the haplotype among the three SNPs was further analysed. It was found that the Indians presenting with the CGC haplotype were protected (OR 0.41, 95% CI 0.21-0.84, P = 0.015) against NAFLD.

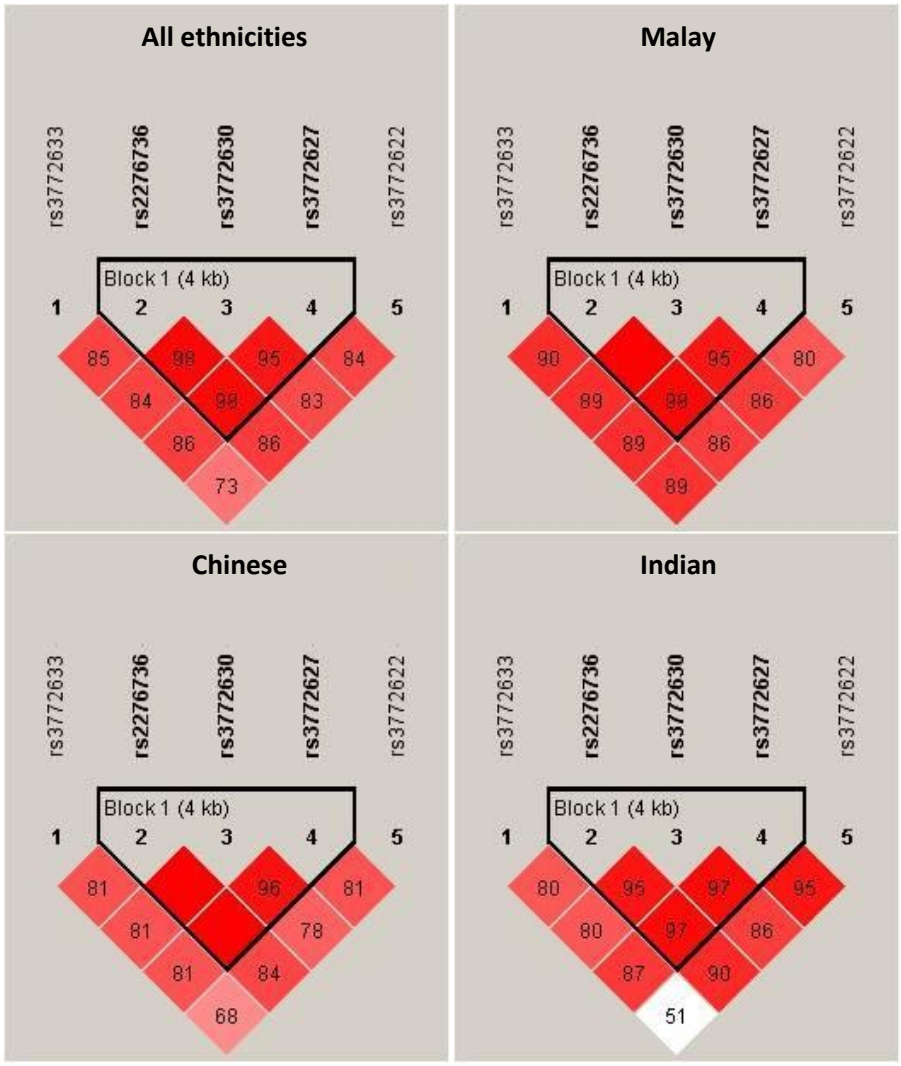


Figure 4.3 Linkage disequilibrium among five SNPs of AGTR1 gene

Table 4.9 Haplotype frequencies of *AGTR1* polymorphisms

Haplotype	Ethnicity															
	Chinese (n=108)				Indian (n=95)				Malay (n=139)				Total (n=342)			
	Case (%)	Ctrl (%)	*P*	OR[a](95% CI)	Case (%)	Ctrl (%)	*P*	OR[a] (95% CI)	Case (%)	Ctrl (%)	*P*	OR[a] (95% CI)	Case (%)	Ctrl (%)	*P*[c]	OR[a] (95% CI)
ATATG	33	37	0.60	Referent	33	31	0.75	Referent	43	34	0.15	Referent	37	34	0.37	Referent
GCGCA	26	26	0.99	1.07 (0.54-2.10)	12	18	0.28	0.59 (0.20-1.68)	24	22	0.59	0.74 (0.38-1.44)	22	22	0.40	0.83 (0.54-1.28)
ATATA	16	18	0.67	1.02 (0.45-2.32)	31	18	0.05	1.62 (0.59-3.28)	18	27	0.09	0.54 (0.29-1.03)	20	22	0.62	0.86 (0.48-1.54)
ACGCA	15	14	0.84	1.26 (0.53-2.96)	15	29	0.03	0.42 (0.15-0.95)	8	12	0.34	0.53 (0.22-1.28)	13	18	0.23	0.66 (0.34-1.29)
Others[b]	10	5	-	-	9	4	-	-	7	5	-	-	8	4	-	-
Total	100	100	-	-	100	100	-	-	100	100	-	-	100	100	-	-

[c]*P* values based on combining results across ethnicities
CI confidence interval, *Ctrl* control, *OR* odds ratio, [a] OR estimated odds ratio by R, [b] Haplotypes with total frequencies below 5% in all subjects

4.2.3 Single nucleotide polymorphism: *LEPR*

4.2.3.1 Genotypes and allele frequencies of *LEPR* polymorphisms

Table 4.10 shows the allele frequencies and association between *LEPR* rs1137100 and rs1137101 with NAFLD. The genotypes of each SNPs were first checked and ensured to be in Hardy-Weinburg equilibirum for both NAFLD cases and controls, for pooled subjects as well as after stratification by ethnicity. There is a significant association observed for both rs1137100 and rs1137101 with susceptibility to NAFLD (OR 1.64, 95% CI 1.18-2.28, $P = 0.003$; and OR 1.61, 95% CI 1.11-2.34, $P = 0.013$, respectively). The association is also observed for rs1137100 in the Malays after ethnic stratification (OR 2.18, 95% CI 1.21-3.95, $P = 0.010$). However, no association is found in all the three ethnic subgroups studied for rs1137101.

Patients with NAFLD were then stratified into simple steatosis ($n = 33$) and NASH ($n = 111$) following which, association analysis was performed (Table 4.10). Both SNPs were associated with NASH (OR 1.49, 95% CI 1.05-2.12, $p = 0.026$; and OR 1.57, 95% CI 1.05-2.35, $P = 0.029$, respectively). However, only rs1137100 is found to be associated with simple steatosis (OR 2.27, 95% CI 1.27-4.08, $P = 0.006$). NASH patients were further stratified into two groups; NASH without significant fibrosis (fibrosis score < 2) and NASH with significant fibrosis (fibrosis score ≥ 2). A significant association with NASH without significant fibrosis was observed for rs1137100 (OR 2.18, 95% CI 1.24-3.84, $P = 0.007$). The power for the significant associations is shown to be 92%, 85%, 60%, 85%, 66%, 79%, and 93%, for rs1137100 with NAFLD, rs1137101 with NAFLD, rs1137100 with NAFLD in Malay, rs1137100 with simple steatosis, rs1137100 with NASH, rs1137101 with NASH, and rs1137100 with NASH without significant fibrosis, respectively.

Table 4.10 Association tests of *LEPR* polymorphisms in different NAFLD spectrum

NAFLD spectrum	rs1137100			rs1137101		
	G allele frequency	*P* value	OR (CI)	G alelle frequency	*P* value	OR (CI)
Control as reference						
NAFLD vs. control						
Overall	0.64 vs. 0.51	0.003	1.64 (1.18-2.28)	0.82 vs. 0.72	0.013	1.61 (1.11-2.34)
Malays	0.71 vs. 0.58	0.010	2.18 (1.21-3.95)	0.85 vs. 0.79	0.199	1.53 (0.80-2.92)
Chinese	0.77 vs. 0.78	0.861	0.94 (0.47-1.87)	0.89 vs. 0.84	0.358	1.42 (0.68-2.97)
Indians	0.26 vs. 0.19	0.278	1.48 (0.73-3.01)	0.65 vs. 0.55	0.203	1.51 (0.80-2.84)
Simple steatosis vs. control	0.71 vs. 0.51	0.006	2.27 (1.27-4.08)	0.83 vs. 0.72	0.113	1.74 (0.88-3.47)
All NASH vs. control	0.61 vs. 0.51	0.026	1.49 (1.05-2.12)	0.82 vs. 0.72	0.029	1.57 (1.05-2.35)
NASH[a] vs. control	0.73 vs. 0.51	0.007	2.18 (1.24-3.84)	0.86 vs. 0.72	0.069	1.91 (0.95-3.82)
NASH[b] vs. control	0.55 vs. 0.51	0.264	1.25 (0.84-1.86)	0.79 vs.0.72	0.114	1.44 (0.78-1.52)
Simple steatosis as reference						
All NASH vs. simple steatosis	0.61 vs. 0.71	0.163	0.64 (0.34-1.20)	0.82 vs. 0.83	0.810	0.91 (0.65-1.86)
NASH[a] vs. simple steatosis	0.73 vs. 0.71	0.823	0.91 (0.41-2.05)	0.86 vs. 0.83	0.857	1.09 (0.44-2.72)
NASH[b] vs. simple steatosis	0.55 vs. 0.71	0.053	0.51 (0.26-1.01)	0.79 vs 0.83	0.511	0.77 (0.34-1.70)
NASH[a] as reference						
NASH[b] vs NASH[a]	0.55 vs. 0.73	0.079	0.56 (0.29-1.07)	0.79 vs 0.86	0.436	0.73 (0.33-1.60)

CI confident interval, *OR* odds ratio, *NAFLD* non-alcoholic fatty liver disease, *NASH* non-alcoholic steatohepatitis, *NASH[a]* non-alcoholic steatohepatitis without significant fibrosis; *NASH[b]* non-alcoholic steatohepatitis with significant fibrosis

The polymorphism of the rs1137100 was earlier found to be associated with simple steatosis (fatty liver) in the Malays. The analysis was continued by comparing the distribution of various clinical and histological parameters among the three major ethnic subgroups (Table 4.11). The parameters that are associated with NAFLD such as BMI, AST and ALT enzymes were found to be the highest in the Malays as compared to the Chinese and Indians ($P < 0.05$).

Table 4.11 Comparison of various clinical and histological parameters between three major ethnic groups among NAFLD patients

Characteristics	NAFLD, *n* = 144 (Mean ± SD)			
	Malay (*n* = 59)	Chinese (*n* = 54)	Indian (*n* = 31)	*P* value
Age (years)	48.1 ± 12.7	53.5 ± 12.1	52.8 ± 9.5	0.038
BMI (kg/m^2)	30.0 ± 4.3	27.4 ± 4.7	28.5 ± 3.2	0.006
HbA1c (%)*	6.9 ± 1.8	5.9 ± 0.9	7.0 ± 2.1	0.003
Waist circumference (cm)	95.8 ± 10.5	90.1 ± 11.7	96.7 ± 8.2	0.004
HDL cholesterol (mg/dl)	49.9 ± 13.9	50.5 ± 10.5	42.4 ± 12.3	0.007
LDL cholesterol (mg/dl)	115.8 ± 46.8	115.0 ± 34.5	123.4 ± 35.2	0.616
Total cholesterol (mg/dl)	197.4 ± 49.9	195.4 ± 38.1	197.8 ± 42.8	0.962
Triglycerides (mg/dl)	164.9 ± 65.7	137.8 ± 44.7	166.2 ± 77.4	0.037
AST (IU/L)*	48.4 ± 27.9	38.5 ± 22.7	40.3 ± 23.5	0.091
ALT (IU/L)	95.5 ± 54.9	74.5 ± 42.8	74.3 ± 40.4	0.037
GGT (IU/L)*	140.8 ± 122.9	89.9 ± 114.9	93.9 ± 90.6	0.005
Systolic BP (mmHg)	135.6 ± 14.8	126.7 ± 14.1	132.5 ± 15.0	0.006
Diastolic BP (mmHg)*	84.5 ± 10.4	80.3 ± 9.1	82.7 ± 9.5	0.054
Steatosis grade*	1.9 ± 0.8	1.7 ± 0.8	1.7 ± 0.7	0.288
Lobular inflammation*	1.3 ± 0.6	1.3 ± 0.6	1.4 ± 0.5	0.845
Ballooning*	1.2 ± 0.6	1.2 ± 0.6	1.4 ± 0.5	0.113
Fibrosis*	1.6 ± 0.9	1.5 ± 0.9	2.3 ± 1.1	0.003

**P* values obtained using Kruskal-Wallis test, all other comparisons used ANOVA

ALT alanine transferase, *AST* aspartate aminotransferase, *BMI* body mass index, *GGT* gamma glutamyl transpeptidase, *HbA1c* haemoglobin A1c, *HDL* high-density lipoprotein, *LDL* low-density lipoprotein, *NAFLD* non-alcoholic fatty liver disease

4.2.3.2 *LEPR* polymorphisms and liver histology

The distribution of various clinical and histological parameters of NAFLD among the genotypes of *LEPR* rs1137100 and rs1137101 were then investigated. However, only rs1137100 showed a significant difference between the genotypes with fibrosis stage ($P = 0.004$) but not with other clinical and histological features (Table 4.12). Genotype GG is significantly associated with decreased fibrosis score (patients with genotype GG compared to AA, had mean score of 1.48 vs. 2.29). For the purpose of determining the magnitude of effect of allele G with fibrosis score, univariate and multivariate analysis were performed (Table 4.13). A single allele G of rs1137100 in an individual is found to be protective against fibrosis (OR 0.47, 95% CI 0.28-0.78, $P = 0.001$). The association remained significant even after the adjustment of other histological features ($P < 0.001$).

However, no association was found between the G allele of rs1137100 with steatosis, lobular inflammation and hepatocellular ballooning.

Table 4.12 Comparison of various clinical and histological parameters between the *LEPR* rs1137100 genotypes among NAFLD patients

Characteristics	NAFLD, *n* = 144 (mean ± SD)			
	AA (*n* = 21)	AG (*n* = 63)	GG (*n* = 60)	*P* value
Age (years)	52.9 ± 10.1	51.1 ± 12.3	50.5 ± 12.5	0.735
BMI (kg/m^2)	28.2 ± 2.9	28.8 ± 5.1	28.7 ± 4.1	0.887
HbA1c (%)*	7.3 ± 2.3	6.5 ± 1.6	6.4 ± 1.4	0.243
Waist circumference (cm)	95.2 ± 8.3	92.6 ± 12.4	94.7 ± 9.9	0.477
HDL cholesterol (mg/dl)	47.5 ± 16.7	48.4 ± 13.3	49.0 ± 10.5	0.907
LDL cholesterol (mg/dl)	114.8 ± 40.2	118.2 ± 40.4	116.8 ± 40.2	0.942
Total cholesterol (mg/dl)	192.3 ± 45.2	202.5 ± 42.7	192.2 ± 44.8	0.380
Triglycerides (mg/dl)	153.7 ± 77.3	167.5 ± 66.9	142.4 ± 49.7	0.084
AST (IU/L)*	37.5 ± 21.8	45.9 ± 28.5	41.6 ± 22.9	0.598
ALT (IU/L)	71.0 ± 36.2	85.0 ± 49.7	85.3 ± 51.0	0.469
GGT (IU/L)*	96.7 ± 98.0	112.5 ± 116.0	116.0 ± 121.8	0.558
Systolic BP (mmHg)	133.8 ± 13.3	130.4 ± 15.4	130.0 ± 15.2	0.654
Diastolic BP (mmHg)*	83.1 ± 9.9	82.4 ± 11.0	82.5 ± 8.7	0.877
Steatosis grade*	1.8 ± 0.7	1.8 ± 0.8	1.8 ± 0.8	0.920
Lobular inflammation*	1.3 ± 0.5	1.3 ± 0.6	1.3 ± 0.6	0.937
Ballooning*	1.4 ± 0.5	1.2 ± 0.6	1.2 ± 0.6	0.426
Fibrosis*	2.3 ± 1.5	1.8 ± 0.8	1.5 ± 1.0	0.004

**P* values obtained using Kruskal-Wallis test, all other comparisons used ANOVA

ALT alanine transferase, *AST* aspartate aminotransferase, *BMI* body mass index, *GGT* gamma glutamyl transpeptidase, *HbA1c* haemoglobin A1c, *HDL* high-density lipoprotein, *LDL* low-density lipoprotein, *NAFLD* non-alcoholic fatty liver disease

Table 4.13 Association of *LEPR* rs1137100 G allele with histological features among NAFLD patients

Histology	Univariate *P* value[a]	multivariate p-value[b] (FDR *q*-value[g])	OR (95% CI)
Steatosis			
> 33% vs. < 33%	0.717	0.880 (1.000)	0.84 (0.52-1.35)[c]
Lobular inflammation			
≥ 2 foci vs. < 2 foci	0.722	0.270 (0.540)	1.01 (0.62-1.65)[d]
Hepatocellular ballooning			
≥ 1 vs. < 1	0.456	0.901 (0.901)	0.77 (0.33-1.79)[e]
Fibrosis			
≥ 2 vs. < 2	0.001	<0.001 (0.002)	0.47 (0.28-0.78)[f]

OR, odds ratio; CI, confident interval
[a]Jonckheere-Terpstra test
[b]Ordinal regression
[c, d, e, f]Multivariate logistic regression
[g]False discovery rate, $q < 0.05$ is significant

4.2.4 Single nucleotide polymorphism: *GCKR*

4.2.4.1 Genotypes and allele frequencies of *GCKR* polymorphisms

The genotype distribution of each SNP rs1260326 C>T and rs780094 C>T was in Hardy-Weinberg equilibrium for both NAFLD cases and controls, as well as after stratification by ethnicity. Assessment in the pooled subjects showed that the frequency of risk allele T was higher in patients with NAFLD compared to controls (OR 1.49, 95% CI 1.09-2.05, P = 0.012; and OR 1.51, 95% CI 1.09–2.09, P = 0.013, for rs1260326 and rs780094, respectively). Higher frequency of T allele was also observed in the Indian NAFLD patients than that of the controls (OR 2.64, 95% CI 1.28-5.43, P = 0.009 and OR 4.35, 95% CI 1.93–9.81, P < 0.0001, for rs1260326 and rs780094, respectively) but this was not seen in the Malays and Chinese (Table 4.14). Data thus indicate an association of T allele with susceptibility to NAFLD. To investigate the next objective which is whether T allele is associated with NAFLD spectrum, patients with NAFLD were grouped into simple steatosis (n = 33) and NASH (n = 111), following which NASH patients were further stratified into NASH with no significant fibrosis (fibrosis score < 2, n = 37) and NASH with significant fibrosis (fibrosis score ≥ 2, n = 74). Significant association was replicated in the NASH (OR 1.55, 95% CI 1.10–2.17, P = 0.013 and OR 1.56, 95% CI 1.10–2.22, P = 0.012, for rs1260326 and rs780094, respectively) and NASH with significant fibrosis (OR 1.49, 95% CI 1.01–2.21, P = 0.044 and OR 1.52, 95% CI 1.03–2.26, P = 0.038, for rs1260326 and rs780094, respectively). No significant finding was observed between both SNPs and simple steatosis and NASH without significant fibrosis (Table 4.14). The power for the significant associations is shown to be 75%, 83%, 77%, 95%, 82%, 83%, 48% and 57%, for rs1260326 with NAFLD, rs780094 with NAFLD, rs1260326 with NAFLD in Indian, rs780094 with NAFLD in Indian, rs1260326 with NASH, rs780094 with NASH,

rs1260326 with NASH with significant fibrosis, rs780094 with NASH with significant fibrosis, respectively.

GCKR polymorphisms confer risk of type 2 diabetes. Adjustment of type 2 diabetes however did not affect the significant results (OR 1.47, 95% CI 1.05–2.08, P = 0.027 and OR 1.70, 95% CI 1.21–2.38, P = 0.002, for rs1260326 and rs780094, respectively).

Table 4.14 Association tests of *GCKR* polymorphisms in different NAFLD spectrum

NAFLD spectrum	rs1260326			rs780094		
	G allele frequency	*P* value	OR (CI)	G alelle frequency	*P* value	OR (CI)
Control as reference						
NAFLD vs. control						
Overall	0.47 vs 0.37	0.012	1.49 (1.09-2.05)	0.46 vs. 0.35	0.013	1.51 (1.09-2.09)
Malays	0.47 vs 0.43	0.466	1.20 (0.74-1.95)	0.46 vs. 0.42	0.497	1.19 (0.72-1.98)
Chinese	0.53 vs 0.44	0.501	1.20 (0.71-2.04)	0.44 vs. 0.44	0.893	1.04 (0.61-1.76)
Indians	0.37 vs 0.19	0.009	2.64 (1.28-5.43)	0.45 vs. 0.20	< 0.0001	4.35 (1.93-9.81)
Simple steatosis vs. control	0.45 vs 0.37	0.276	1.35 (0.79-2.29)	0.45 vs. 0.35	0.162	1.48 (0.86-2.54)
All NASH vs. control	0.49 vs 0.37	0.013	1.55 (1.10-2.17)	0.47 vs. 0.35	0.012	1.56 (1.10-2.22)
NASH[a] vs. control	0.51 vs 0.37	0.051	1.68 (0.99-2.82)	0.49 vs. 0.35	0.080	1.60 (0.95-2.71)
NASH[b] vs. control	0.46 vs 0.37	0.044	1.49 (1.01-2.21)	0.45 vs.0.35	0.038	1.52 (1.03-2.26)
Simple steatosis as reference						
All NASH vs. simple steatosis	0.49 vs 0.45	0.720	1.11 (0.64-1.91)	0.47 vs 0.45	0.840	1.04 (0.57-1.83)
NASH[a] vs. simple steatosis	0.51 vs. 0.45	0.452	1.30 (0.66-2.58)	0.49 vs. 0.45	0.640	1.19 (0.58-2.46)
NASH[b] vs. simple steatosis	0.46 vs. 0.45	0.810	1.07 (0.61-1.89)	0.45 vs 0.45	0.908	0.97 (0.53-1.77)
NASH[a] as reference						
NASH[b] vs NASH[a]	0.46 vs. 0.51	0.506	0.82 (0.46-1.47)	0.49 vs 0.45	0.579	0.84 (0.46-1.54)

CI confident interval, *OR* odds ratio, *NAFLD* non-alcoholic fatty liver disease, *NASH* non-alcoholic steatohepatitis, *NASH[a]* non-alcoholic steatohepatitis without significant fibrosis; *NASH[b]* non-alcoholic steatohepatitis with significant fibrosis

4.2.4.2 *GCKR* polymorphisms and liver histology

None of the NAFLD-associated parameters were different among the genotypes of *GCKR* rs1260326 and rs780094. However, comparison of histological attributes among the genotypes revealed a significantly higher score of hepatic steatosis in patients homozygous for the T allele (homozygous TT compared to homozygous CC, had a mean score of 2.18 *vs* 1.63, $P = 0.008$) (Table 4.15). Both univariate and multivariate analysis were applied in order to determine the magnitude of effect of the risk allele T with the NAFLD histological features (Table 4.16). Carrier of allele T exhibited 1.76 risk to develop hepatic steatosis than the non-carrier (OR 1.76, 95% CI 1.08–2.85, $P = 0.004$) and this significant finding remained after adjustment with other histological features ($P = 0.04$). This significant finding however was not seen with lobular inflammation, hepatocellular ballooning and fibrosis.

Table 4.15 Comparison of various clinical and histological parameters between the *GCKR* rs1260326 genotypes among NAFLD patients

Characteristics	NAFLD, *n* = 144 (mean ± SD)			
	CC (*n* = 41)	CT (*n* = 70)	TT (*n* = 33)	*P* value
Age (years)	50.8 ± 12.5	51.8 ± 11.5	50.0 ± 12.8	0.774
BMI (kg/m^2)	28.8 ± 4.2	28.9 ± 4.6	28.0 ± 4.2	0.597
HbA1c (%)*	7.0 ± 1.9	6.4 ± 1.6	6.4 ± 1.4	0.136
Waist circumference (cm)	95.1 ± 10.1	93.6 ± 11.3	92.7 ± 11.1	0.626
HDL cholesterol (mg/dl)	49.4 ± 13.1	47.6 ± 12.6	49.3 ± 12.8	0.714
LDL cholesterol (mg/dl)	113.6 ± 34.0	122.3 ± 44.4	110.6 ± 36.6	0.309
Total cholesterol (mg/dl)	194.8 ± 40.3	199.2 ± 48.2	193.9 ± 39.6	0.802
Triglycerides (mg/dl)	169.2 ± 77.8	145.8 ± 54.0	157.0 ± 57.1	0.160
AST (IU/L)*	38.3 ± 21.7	43.8 ± 26.0	46.9 ± 28.0	0.518
ALT (IU/L)	74.7 ± 36.2	84.5 ± 52.0	90.4 ± 53.8	0.365
GGT (IU/L)*	88.1 ± 70.7	132.8 ± 143.8	96.0 ± 82.7	0.222
Systolic BP (mmHg)	133.9 ± 14.8	132.4 ± 14.3	126.9 ± 16.2	0.109
Diastolic BP (mmHg)*	83.0 ± 10.7	82.6 ± 9.5	81.9 ± 9.8	0.775
Steatosis grade*	1.6 ± 0.7	1.7 ± 0.7	2.2 ± 0.8	0.008
Lobular inflammation*	1.4 ± 0.5	1.2 ± 0.6	1.4 ± 0.6	0.086
Ballooning*	1.2 ± 0.7	1.2 ± 0.5	1.3 ± 0.7	0.672
Fibrosis*	1.8 ± 1.0	1.7 ± 1.0	1.6 ± 0.9	0.570

**P* values obtained using Kruskal-Wallis test, all other comparisons used ANOVA
ALT alanine transferase, *AST* aspartate aminotransferase, *BMI* body mass index, *GGT* gamma glutamyl transpeptidase, *HbA1c* haemoglobin A1c, *HDL* high-density lipoprotein, *LDL* low-density lipoprotein, *NAFLD* non-alcoholic fatty liver disease

Table 4.16 Association of *GCKR* rs1260326 T allele with histological features among NAFLD patients

Histology	Univariate *P* value[a]	multivariate p-value[b] (FDR *q*-value[g])	OR (95% CI)
Steatosis			
> 33% vs. < 33%	0. 004	0.040 (0.016)	1.76 (1.08-2.85)[c]
Lobular inflammation			
≥ 2 foci vs. < 2 foci	0. 750	0.639 (0.750)	1.18 (0.73-1.91)[d]
Hepatocellular ballooning			
≥ 1 vs. < 1	0. 739	0.281 (0.985)	1.05 (0.47-2.33)[e]
Fibrosis			
≥ 2 vs. < 2	0. 414	0.859 (0.828)	0.81 (0.51-1.29)[f]

OR, odds ratio; CI, confident interval
[a]Jonckheere-Terpstra test
[b]Ordinal regression
[c, d, e, f]Multivariate logistic regression
[g]False discovery rate, $q < 0.05$ is significant

4.2.5 Single nucleotide polymorphism: *iNOS*

A SNP of *iNOS* gene rs1060822 T/C was genotyped in this study. The rs1060822 is located in the coding region on chromosome 17. The rs1060822 is a synonymous coding SNP where different alleles encode for the same amino acid glycine. Table 4.17 shows the association test between NAFLD and control subjects with NAFLD. No allelic association with NAFLD was observed (OR 0.96, 95% CI 0.61-1.50, $P = 0.839$). The genotypes deviated from Hardy-Weinberg equilibrium and thus analysis by ethnicity stratification and by comparison of parameters among the genotypes did not proceed further.

Table 4.17 Allelic association of *iNOS* polymorphism

SNP	C allele frequency	*P* value	OR (CI)
rs1060822 (T > C) NAFLD vs. control	0.76 vs. 0.77	0.839	0.96 (0.61-1.50)

CI confident interval, *OR* odds ratio, *NAFLD* non-alcoholic fatty liver disease

4.2.6 Single nucleotide polymorphism: *TRAIL*

A SNP of *TRAIL* gene rs1131568 A/G was genotyped in this study. The rs1131568 is an untranslated-3 SNP, located in the non-coding region on chromosome 3. Table 4.18 shows the association test between NAFLD and control subjects with NAFLD. The genotypes were in Hardy-Weinberg equilibrium. No allelic association with NAFLD was observed in the pooled subjects (OR 1.05, 95% CI 0.76-1.65, *P* = 0.760) nor after stratification by ethnicity (OR 0.99, 95% CI 0.76-2.55, *P* = 0.955; OR 1.39, 95% CI 0.46-1.48, *P* = 0.286; and OR 0.91, 95% CI 0.49-1.69, *P* = 0.766, for Malay, Chinese and Indian, respectively).

Table 4.18 Allelic association of *TRAIL* polymorphism

SNP	G allele frequency	*P* value	OR (CI)
rs1131568 (A > G)			
NAFLD vs. control			
Overall	0.69 vs. 0.67	0.760	1.05 (0.76-1.65)
Malays	0.72 vs. 0.72	0.955	0.99 (0.76-2.55)
Chinese	0.68 vs. 0.72	0.286	1.39 (0.46-1.48)
Indians	0.65 vs. 0.66	0.766	0.91 (0.49-1.69)

CI confident interval, *OR* odds ratio, *NAFLD* non-alcoholic fatty liver disease

Analysis of comparison of parameters between the *TRAIL* rs1131568 genotypes among NAFLD patients is shown in Table 4.19. Only diastolic blood pressure is different among the genotypes where the highest is observed in those with GG genotype (*P* = 0.033).

Table 4.19 Comparison of various clinical and histological parameters between the *TRAIL* rs1131568 genotypes among NAFLD patients

Characteristics	NAFLD, *n* = 144 (mean ± SD)			
	CC (*n* = 19)	CG (*n* = 52)	GG (*n* = 73)	*P* value
Age (years)	45.9 ± 12.7	51.5 ± 12.0	51.9± 11.5	0.151
BMI (kg/m^2)	29.1 ± 5.2	27.7 ± 3.7	29.3 ± 4.6	0.130
HbA1c (%)*	6.3 ± 1.6	6.4 ± 1.7	6.8 ± 1.6	0.133
Waist circumference (cm)	91.8 ± 11.6	91.9 ± 9.8	95.6 ± 11.1	0.115
HDL cholesterol (mg/dl)	44.4 ± 8.5	48.7 ± 14.0	49.9 ± 12.6	0.259
LDL cholesterol (mg/dl)	122.1 ± 36.7	117.5 ± 43.7	113.5 ± 35.7	0.667
Total cholesterol (mg/dl)	197.6 ± 45.0	197.4 ± 49.3	194.6 ± 38.3	0.908
Triglycerides (mg/dl)	150.3 ± 83.5	165.4 ± 60.5	148.8 ± 57.9	0.314
AST (IU/L)*	41.7 ± 26.9	40.1 ± 19.4	45.5 ± 29.1	0.887
ALT (IU/L)	80.9 ± 46.8	83.2 ± 48.0	83.6 ± 50.9	0.978
GGT (IU/L)*	89.8 ± 84.6	124.9 ± 130.0	108.8 ± 113.5	0.696
Systolic BP (mmHg)	129.9 ± 20.8	130.4 ± 13.9	132.2 ± 14.3	0.752
Diastolic BP (mmHg)*	80.7 ± 14.2	80.4 ± 8.0	84.5 ± 9.7	0.033
Steatosis grade*	1.9 ± 0.7	1.8 ± 0.8	1.8 ± 0.8	0.698
Lobular inflammation*	1.5 ± 0.5	1.4 ± 0.5	1.2 ± 0.6	0.176
Ballooning*	1.3 ± 0.5	1.2 ± 0.6	1.2 ± 0.6	0.698
Fibrosis*	1.9 ± 1.0	1.7 ± 0.9	1.7 ± 1.0	0.637

**P* values obtained using Kruskal-Wallis test, all other comparisons used ANOVA

ALT alanine transferase, *AST* aspartate aminotransferase, *BMI* body mass index, *GGT* gamma glutamyl transpeptidase, *HbA1c* haemoglobin A1c, *HDL* high-density lipoprotein, *LDL* low-density lipoprotein, *NAFLD* non-alcoholic fatty liver disease

4.2.7 Single nucleotide polymorphism: *SREBF1*

A total of two SNPs of *SREBF1* gene rs1186803 G/A and rs2297508 C/G were genotyped in this study. Both are untranslated-3 SNPs, located in the non-coding region on chromosome 13. Table 4.20 shows the association test between NAFLD and control subjects with NAFLD. The genotypes were in Hardy-Weinberg equilibrium. No allelic association with NAFLD was observed for rs1186803 ($P > 0.05$). However, there is a difference between cases and controls in the Indian for rs2297508 (OR 3.56, 95% CI 1.41-7.98, $P = 0.007$).

Table 4.20 Allelic association of *SREBF1* polymorphisms

SNP	Risk allele frequency	*P* value	OR (CI)
rs1186803 (G > A)			
NAFLD vs. control			
Overall	0.77 vs. 0.78	0.847	0.96 (0.66-1.41)
Malays	0.73 vs. 0.82	0.097	0.61 (0.34-1.10)
Chinese	0.81 vs. 0.75	0.341	1.40 (0.70-2.81)
Indians	0.84 vs. 0.76	0.263	1.51 (0.76-3.04)
rs2297508 (C > G)			
NAFLD vs. control			
Overall	0.78 vs. 0.77	0.375	1.19 (0.81-1.75)
Malays	0.73 vs. 0.81	0.108	0.61 (0.34-1.11)
Chinese	0.81 vs. 0.75	0.341	1.40 (0.70-2.81)
Indians	0.89 vs. 0.79	0.007	3.56 (1.41-7.98)

CI confident interval, *OR* odds ratio, *NAFLD* non-alcoholic fatty liver disease

Analysis of comparison of parameters between the genotypes among NAFLD patients was therefore performed in the *SREBF1* rs2297508 (Table 4.21). Diastolic blood pressure and GGT were found higher in the CC genotype ($P < 0.05$). A marginal association was found in the systolic blood pressure ($P = 0.05$).

Table 4.21 Comparison of various clinical and histological parameters between the *SREBF1* rs2297508 genotypes among NAFLD patients

Characteristics	NAFLD, *n* = 144 (mean ± SD)			
	CC (*n* = 5)	CG (*n* = 54)	GG (*n* = 85)	*P* value
Age (years)	45.7 ± 10.6	51.1 ± 12.1	51.3 ± 11.5	0.538
BMI (kg/m^2)	30.1 ± 5.2	28.7 ± 4.8	28.5 ± 4.1	0.633
HbA1c (%)*	6.1 ± 0.7	7.0 ± 2.1	6.4 ± 1.3	0.612
Waist circumference (cm)	94.4 ± 11.6	94.8 ± 10.8	95.0 ± 11.0	0.641
HDL cholesterol (mg/dl)	56.1 ± 20.5	46.7 ± 10.1	49.2 ± 13.3	0.189
LDL cholesterol (mg/dl)	134.9 ± 61.7	120.5 ± 38.2	112.1 ± 37.1	0.233
Total cholesterol (mg/dl)	224.6 ± 67.0	198.3 ± 45.4	192.8 ± 34.5	0.183
Triglycerides (mg/dl)	158.2 ± 74.3	163.2 ± 68.2	149.5 ± 56.0	0.465
AST (IU/L)*	57.2 ± 16.9	40.9 ± 26.2	43.4 ± 25.5	0.085
ALT (IU/L)	121.5 ± 44.7	76.2 ± 50.4	84.5 ± 47.5	0.094
GGT (IU/L)*	240.8 ± 164.6	86.8 ± 70.7	119.0 ± 129.5	0.039
Systolic BP (mmHg)	142.5 ± 12.5	133.3 ± 13.9	129.2 ± 14.1	0.050
Diastolic BP (mmHg)*	92.7 ± 6.9	82.4 ± 8.9	81.9 ± 10.0	0.031
Steatosis grade*	1.9 ± 0.7	1.8 ± 0.8	1.8 ± 0.8	0.642
Lobular inflammation*	1.3 ± 0.8	1.4 ± 0.6	1.4 ± 0.6	0.726
Ballooning*	1.0 ± 0.6	1.3 ± 0.6	1.2 ± 0.6	0.419
Fibrosis*	1.7 ± 1.0	1.8 ±1.1	1.7 ± 0.9	0.967

**P* values obtained using Kruskal-Wallis test, all other comparisons used ANOVA

ALT alanine transferase, *AST* aspartate aminotransferase, *BMI* body mass index, *GGT* gamma glutamyl transpeptidase, *HbA1c* haemoglobin A1c, *HDL* high-density lipoprotein, *LDL* low-density lipoprotein, *NAFLD* non-alcoholic fatty liver disease

4.2.8 Single nucleotide polymorphism: *CLOCK*

A total of four SNPs of *CLOCK* gene rs1193259 A/G, rs4864548 G/A, rs6843722 A/C and rs6850524 G/C were genotyped in this study. The rs4864548 is a non-coding SNP, located in the locus region near-gene-5 on chromosome 4. The following SNPs rs1193259, rs6843722 and rs6850524 are in the intronic regions on chromosome 4. Table 4.22 shows the association test between NAFLD and control subjects with NAFLD. The genotypes were in Hardy-Weinberg equilibrium. No allelic association with NAFLD was observed for all four SNPs in the pooled ($p > 0.05$) nor after stratification by ethnicity ($P > 0.05$).

Table 4.22 Allelic association of *CLOCK* polymorphisms

SNP	Risk allele frequency	*P* value	OR (CI)
rs1193259 (A > G)			
NAFLD vs. control			
Overall	0.19 vs. 0.20	0.920	0.98 (0.67-1.44)
Malays	0.19 vs. 0.11	0.098	1.75 (0.90-3.41)
Chinese	0.08 vs. 0.15	0.114	0.48 (0.19-1.20)
Indians	0.37 vs. 0.36	0.880	1.05 (0.57-1.93)
rs4864548 (G > A)			
NAFLD vs. control			
Overall	0.60 vs. 0.57	0.928	1.02 (0.74-1.39)
Malays	0.58 vs. 0.63	0.446	0.83 (0.52-1.34)
Chinese	0.62 vs. 0.61	0.882	1.05 (0.62-1.86)
Indians	0.56 vs. 0.51	0.451	1.27 (0.68-2.39)
rs6843722 (A > C)			
NAFLD vs. control			
Overall	0.58 vs. 0.57	0.966	0.99 (0.73-1.35)
Malays	0.56 vs. 0.58	0.409	0.82 (0.51-1.31)
Chinese	0.62 vs. 0.60	0.885	1.04 (0.59-1.84)
Indians	0.53 vs. 0.49	0.603	1.18 (0.64-2.17)
rs6850524 (G > C)			
NAFLD vs. control			
Overall	0.30 vs. 0.29	0.670	1.07 (0.77-1.49)
Malays	0.33 vs. 0.29	0.464	1.17 (0.71-1.94)
Chinese	0.29 vs. 0.25	0.540	1.16 (0.65-2.07)
Indians	0.24 vs. 0.30	0.425	0.75 (0.37-1.51)

CI confident interval, *OR* odds ratio, *NAFLD* non-alcoholic fatty liver disease

Although none of the SNPs showed significance, lowest P-value was recorded by *CLOCK* rs6850524. The clinical and histological parameters were compared between genotypes of this SNP among NAFLD patients (Table 4.23). No parameters were significantly different among the genotypes ($P > 0.05$).

Table 4.23 Comparison of various clinical and histological parameters between the *CLOCK* rs6850524 genotypes among NAFLD patients

Characteristics	NAFLD, *n* = 144 (mean ± SD)			
	GG (*n* = 72)	GC (*n* = 56)	CC (*n* = 16)	*P* value
Age (years)	49.9 ± 11.3	51.9 ± 12.8	53.9 ± 12.3	0.407
BMI (kg/m^2)	29.3 ± 4.4	28.7 ± 4.3	28.1 ± 4.1	0.332
HbA1c (%)*	6.7 ± 1.8	6.4 ± 1.5	6.4 ± 1.4	0.769
Waist circumference (cm)	95.3 ± 11.2	93.3 ± 10.5	89.9 ± 10.4	0.192
HDL cholesterol (mg/dl)	47.8 ± 10.6	47.5 ± 12.7	55.2 ± 18.4	0.189
LDL cholesterol (mg/dl)	117.1 ± 44.0	117.9 ± 35.2	114.1 ± 37.6	0.233
Total cholesterol (mg/dl)	196.9 ± 45.6	195.3 ± 41.4	200.2 ± 48.5	0.183
Triglycerides (mg/dl)	150.0 ± 59.6	158.2 ± 67.7	165.0 ± 59.4	0.465
AST (IU/L)*	42.0 ± 25.8	43.5 ± 24.2	44.9 ± 26.5	0.815
ALT (IU/L)	79.8 ± 44.7	86.5 ± 51.2	85.5 ± 52.2	0.727
GGT (IU/L)*	95.3 ± 90.9	119.0 ± 129.1	159.0 ± 152.0	0.701
Systolic BP (mmHg)	133.5 ± 16.0	128.3 ± 15.0	132.2 ± 10.2	0.177
Diastolic BP (mmHg)*	83.7 ± 10.5	80.9 ± 8.9	83.1 ± 10.0	0.386
Steatosis grade*	1.8 ± 0.8	1.9 ± 0.8	1.8 ± 0.9	0.814
Lobular inflammation*	1.3 ± 0.6	1.3 ± 0.6	1.4 ± 0.5	0.788
Ballooning*	1.3 ± 0.6	1.1 ± 0.6	1.0 ± 0.6	0.094
Fibrosis*	1.9 ± 1.0	1.6 ± 0.9	1.6 ± 0.9	0.215

**P* values obtained using Kruskal-Wallis test, all other comparisons used ANOVA

ALT alanine transferase, *AST* aspartate aminotransferase, *BMI* body mass index, *GGT* gamma glutamyl transpeptidase, *HbA1c* haemoglobin A1c, *HDL* high-density lipoprotein, *LDL* low-density lipoprotein, *NAFLD* non-alcoholic fatty liver disease

4.2.9 Single nucleotide polymorphism: *COL13A1*

A SNP of *COL13A1* gene rs1227756 G/A was genotyped in this study. The rs1227756 is a non-coding SNP, located in the intronic region on chromosome 10. Table 4.24 shows the association test between NAFLD and control subjects with NAFLD. No allelic association with NAFLD was observed (OR 0.80, 95% CI 0.58-1.09, *P* = 0.159). The genotypes deviated from Hardy-Weinberg equilibrium and thus analysis by ethnic stratification and by comparison of parameters among the genotypes did not proceed further.

Table 4.24 Allelic association of *COL13A1* polymorphism

SNP	A allele frequency	*P* value	OR (CI)
rs1227756 (G > A)			
NAFLD vs. control	0.29 vs. 0.35	0.159	0.80 (0.58-1.09)

CI confident interval, *OR* odds ratio, *NAFLD* non-alcoholic fatty liver disease

4.2.10 Single nucleotide polymorphism: *FABP2*

A SNP of *FABP2* gene rs1799883 A/G was genotyped in this study. The polymorphism of rs1799883 is a missense mutation, substituting the amino acid threonine to alanine. The rs1799883 is located on chromosome 10. Table 4.25 shows the association test between NAFLD and control subjects with NAFLD. The genotypes were in Hardy-Weinberg equilibrium. No allelic association with NAFLD was observed in the pooled subjects (OR 1.14, 95% CI 0.84-1.77, P = 0.296) nor after stratification by ethnicity (OR 1.27, 95% CI 0.80-2.52, P = 0.216; OR 1.70, 95% CI 0.91-3.18, P = 0.096; and OR 0.87, 95% CI 0.47-1.63, P = 0.661, for Malay, Chinese and Indian, respectively).

Table 4.25 Allelic association of *FABP2* polymorphism

SNP	G allele frequency	*P* value	OR (CI)
rs1799883 (A > G)			
NAFLD vs. control			
Overall	0.81 vs. 0.77	0.296	1.14 (0.84-1.77)
Malays	0.85 vs. 0.83	0.216	1.27 (0.80-2.52)
Chinese	0.81 vs. 0.71	0.096	1.70 (0.91-3.18)
Indians	0.71 vs. 0.74	0.661	0.87 (0.47-1.63)

CI confident interval, *OR* odds ratio, *NAFLD* non-alcoholic fatty liver disease

Analysis of comparison of clinical and histological parameters between the *FABP2* rs1799883 genotypes among NAFLD patients is shown in Table 4.26. None of the parameters were different between the genotypes ($P > 0.05$).

Table 4.26 Comparison of various clinical and histological parameters between the *FABP2* rs1799883 genotypes among NAFLD patients

Characteristics	NAFLD, *n* = 144 (mean ± SD)			
	AA (*n* = 8)	AG (*n* = 40)	GG (*n* = 96)	*P* value
Age (years)	49.5 ± 9.5	54.9 ± 10.3	49.5 ± 12.3	0.060
BMI (kg/m^2)	26.6 ± 2.8	28.8 ± 4.9	28.8 ± 4.3	0.402
HbA1c (%)*	5.8 ± 0.8	6.4 ± 1.5	6.7 ± 1.7	0.140
Waist circumference (cm)	88.3 ± 6.3	92.3 ± 12.6	94.6 ± 10.2	0.235
HDL cholesterol (mg/dl)	49.2 ± 13.8	49.8 ± 14.4	48.1 ± 11.9	0.780
LDL cholesterol (mg/dl)	117.9 ± 31.2	125.0 ± 39.4	112.1 ± 38.9	0.238
Total cholesterol (mg/dl)	195.5 ± 43.1	207.3 ± 40.8	191.2 ± 43.8	0.151
Triglycerides (mg/dl)	135.0 ± 77.7	158.6 ± 63.2	155.1 ± 61.3	0.625
AST (IU/L)*	34.1 ± 15.2	39.3 ± 20.9	45.3 ± 27.7	0.491
ALT (IU/L)	71.6 ± 32.4	77.5 ± 43.2	86.3 ± 52.2	0.509
GGT (IU/L)*	62.9 ± 32.1	118.8 ± 131.1	113.8 ± 114.8	0.640
Systolic BP (mmHg)	130.6 ± 10.2	131.1 ± 15.6	131.4 ± 15.2	0.988
Diastolic BP (mmHg)*	81.7 ± 6.5	82.0 ± 10.5	82.8 ± 10.0	0.758
Steatosis grade*	1.8 ± 0.7	1.7 ± 0.7	1.9 ± 0.8	0.418
Lobular inflammation*	1.8 ± 0.5	1.3 ± 0.6	1.3 ± 0.6	0.067
Ballooning*	1.5 ± 0.5	1.1 ± 0.6	1.2 ± 0.6	0.209
Fibrosis*	2.1 ± 0.4	1.7 ± 1.0	1.7 ± 1.0	0.213

**P* values obtained using Kruskal-Wallis test, all other comparisons used ANOVA

ALT alanine transferase, *AST* aspartate aminotransferase, *BMI* body mass index, *GGT* gamma glutamyl transpeptidase, *HbA1c* haemoglobin A1c, *HDL* high-density lipoprotein, *LDL* low-density lipoprotein, *NAFLD* non-alcoholic fatty liver disease

4.2.11 Single nucleotide polymorphism: *UCP3*

A SNP of *UCP3* gene rs1800849 C/T was genotyped in this study. The rs1800849 is an untranslated-5 SNP, located in the non-coding region on chromosome 11. Table 4.27 shows the association test between NAFLD and control subjects with NAFLD. The genotypes were in Hardy-Weinberg equilibrium. No allelic association with NAFLD was observed in the pooled subjects (OR 1.33, 95% CI 0.94-1.90, P = 0.109) nor after stratification by ethnicity (OR 1.57, 95% CI 0.94-2.64, P = 0.086; OR 1.61, 95% CI 0.84-3.08, P = 0.149; and OR 0.67, 95% CI 0.31-1.46, P = 0.352, for Malay, Chinese and Indian, respectively).

Table 4.27 Allelic association of *UCP3* polymorphism

SNP	T allele frequency	*P* value	OR (CI)
rs1800849 (C > T)			
NAFLD vs. control			
Overall	0.30 vs. 0.24	0.109	1.33 (0.94-1.90)
Malays	0.34 vs. 0.24	0.086	1.57 (0.94-2.64)
Chinese	0.31 vs. 0.23	0.149	1.61 (0.84-3.08)
Indians	0.20 vs. 0.26	0.352	0.67 (0.31-1.46)

CI confident interval, *OR* odds ratio, *NAFLD* non-alcoholic fatty liver disease

Analysis of comparison of clinical and histological parameters between the *UCP3* rs1800849 genotypes among NAFLD patients is shown in Table 4.28. None of the parameters were different between the genotypes (P > 0.05).

Table 4.28 Comparison of various clinical and histological parameters between the *UCP3* rs1800849 genotypes among NAFLD patients

Characteristics	NAFLD, *n* = 144 (mean ± SD)			
	CC (*n* = 69)	CT (*n* = 62)	TT (*n* = 13)	*P* value
Age (years)	50.6 ± 11.6	51.3 ± 12.1	51.6 ± 13.9	0.925
BMI (kg/m^2)	28.6 ± 4.7	28.3 ± 3.7	30.7 ± 5.3	0.224
HbA1c (%)*	6.4 ± 1.5	6.7 ± 1.7	7.1 ± 2.1	0.543
Waist circumference (cm)	93.3 ± 11.6	93.4 ± 9.4	96.8 ± 13.2	0.606
HDL cholesterol (mg/dl)	49.4 ± 13.6	47.5 ± 12.3	49.8 ± 8.3	0.780
LDL cholesterol (mg/dl)	116.5 ± 40.2	116.3 ± 37.5	112.1 ± 40.3	0.238
Total cholesterol (mg/dl)	195.5 ± 42.5	195.6 ± 44.7	198.9 ± 44.2	0.151
Triglycerides (mg/dl)	150.4 ± 64.0	153.6 ± 57.2	188.6 ± 73.9	0.625
AST (IU/L)*	41.4 ± 26.4	43.4 ± 20.7	51.0 ± 40.4	0.298
ALT (IU/L)	78.5 ± 50.9	87.8 ± 46.2	86.3 ± 52.7	0.548
GGT (IU/L)*	104.6 ± 109.6	119.5 ± 117.9	120.6 ± 153.2	0.569
Systolic BP (mmHg)	129.4 ± 14.8	132.3 ± 14.9	136.5 ± 16.2	0.244
Diastolic BP (mmHg)*	83.0 ± 9.9	81.2 ± 9.5	86.4 ± 13.3	0.327
Steatosis grade*	1.7 ± 0.7	1.8 ± 0.8	2.3 ± 0.9	0.129
Lobular inflammation*	1.3 ± 0.6	1.4 ± 0.6	1.3 ± 0.5	0.563
Ballooning*	1.2 ± 0.6	1.2 ± 0.6	1.3 ± 0.5	0.909
Fibrosis*	1.6 ± 0.9	1.8 ± 1.1	1.8 ± 1.0	0.568

**P* values obtained using Kruskal-Wallis test, all other comparisons used ANOVA

ALT alanine transferase, *AST* aspartate aminotransferase, *BMI* body mass index, *GGT* gamma glutamyl transpeptidase, *HbA1c* haemoglobin A1c, *HDL* high-density lipoprotein, *LDL* low-density lipoprotein, *NAFLD* non-alcoholic fatty liver disease

4.2.12 Single nucleotide polymorphism: *PPARG*

A SNP of *PPARG* gene rs1801282 C/G was genotyped in this study. The polymorphism of rs1801282 is a missense mutation, substituting the amino acid proline to alanine. The rs1801282 is located on chromosome 3. Table 4.29 shows the association test between NAFLD and control subjects with NAFLD. The genotypes were in Hardy-Weinberg equilibrium. No allelic association with NAFLD was observed in the pooled subjects (OR 1.02, 95% CI 0.52-1.99, P = 0.966) nor after stratification by ethnicity (OR 1.73, 95% CI 0.50-5.97, P = 0.385; OR 3.12, 95% CI 0.31-11.00, P = 0.331; and OR 0.80, 95% CI 0.30-2.08, P = 0.642, for Malay, Chinese and Indian, respectively).

Table 4.29 Allelic association of *PPARG* polymorphism

SNP	G allele frequency	*P* value	OR (CI)
rs1801282 (C > G)			
NAFLD vs. control			
Overall	0.06 vs. 0.06	0.966	1.02 (0.52-1.99)
Malays	0.05 vs. 0.03	0.385	1.73 (0.50-5.97)
Chinese	0.03 vs. 0.01	0.331	3.12 (0.31-11.00)
Indians	0.12 vs. 0.14	0.642	0.80 (0.30-2.08)

CI confident interval, *OR* odds ratio, *NAFLD* non-alcoholic fatty liver disease

Analysis of comparison of clinical and histological parameters between the *PPARG* rs1801282 genotypes among NAFLD patients is shown in Table 4.30. Ballooning score was found different between the genotypes (P = 0.040).

Table 4.30 Comparison of various clinical and histological parameters between the *PPARG* rs1801282 genotypes among NAFLD patients

Characteristics	NAFLD, *n* = 144 (mean ± SD)			
	CC (*n* = 128)	CG (*n* = 16)	GG (*n* = 0)	*P* value
Age (years)	50.7 ± 12.1	53.4 ± 10.8	-	0.392
BMI (kg/m^2)	28.7 ± 4.5	28.6 ± 3.5	-	0.938
HbA1c (%)*	6.6 ± 1.7	6.4 ± 1.2	-	0.979
Waist circumference (cm)	93.6 ± 11.0	95.1 ± 9.3	-	0.609
HDL cholesterol (mg/dl)	48.7 ± 13.0	48.1 ± 10.2	-	0.866
LDL cholesterol (mg/dl)	115.5 ± 38.2	121.3 ± 42.7	-	0.564
Total cholesterol (mg/dl)	195.1 ± 43.5	201.7 ± 42.4	-	0.563
Triglycerides (mg/dl)	156.4 ± 64.9	144.4 ± 38.2	-	0.475
AST (IU/L)*	43.1 ± 26.3	42.7 ± 19.7	-	0.687
ALT (IU/L)	84.1 ± 50.9	75.6 ± 30.2	-	0.519
GGT (IU/L)*	113.2 ± 121.0	105.5 ± 76.7	-	0.507
Systolic BP (mmHg)	131.0 ± 15.0	133.3 ± 15.3	-	0.548
Diastolic BP (mmHg)*	82.5 ± 10.1	83.2 ± 9.1	-	0.940
Steatosis grade*	1.8 ± 0.8	1.6 ± 0.7	-	0.326
Lobular inflammation*	1.3 ± 0.6	1.1 ± 0.6	-	0.228
Ballooning*	1.3 ± 0.6	0.9 ± 0.4	-	0.040
Fibrosis*	1.8 ± 0.9	1.4 ± 1.2	-	0.119

**P* values obtained using Kruskal-Wallis test, all other comparisons used ANOVA

ALT alanine transferase, *AST* aspartate aminotransferase, *BMI* body mass index, *GGT* gamma glutamyl transpeptidase, *HbA1c* haemoglobin A1c, *HDL* high-density lipoprotein, *LDL* low-density lipoprotein, *NAFLD* non-alcoholic fatty liver disease

4.2.13 Single nucleotide polymorphism: *COX-2*

A SNP of *COX-2* gene rs20417 G/C was genotyped in this study. The rs20417 is a non-coding SNP, located in the locus region near-gene-5 on chromosome 1. Table 4.31 shows the association test between NAFLD and control subjects with NAFLD. No allelic association with NAFLD was observed (OR 0.62, 95% CI 0.28-1.38, $P = 0.245$). The genotypes deviated from Hardy-Weinberg equilibrium and thus analysis by ethnic stratification and by comparison of parameters among the genotypes did not proceed further.

Table 4.31 Allelic association of *COX-2* polymorphism

SNP	C allele frequency	*P* value	OR (CI)
rs20417 (G > C)			
NAFLD vs. control	0.01 vs. 0.04	0.245	0.62 (0.28-1.38)

CI confident interval, *OR* odds ratio, *NAFLD* non-alcoholic fatty liver disease

4.2.14 Single nucleotide polymorphism: *NR1I2*

A total of two SNPs of *NR1I2* gene rs2461823 A/G and rs7643645 A/G were genotyped in this study. Both are non-coding SNPs, located in the intronic regions on chromosome 3. Table 4.32 shows the association test between NAFLD and control subjects with NAFLD. The genotypes were in Hardy-Weinberg equilibrium. No allelic association with NAFLD was observed for both SNPs in the pooled subjects ($P > 0.05$) nor after stratification by ethnicity ($P > 0.05$).

Table 4.32 Allelic association of *NR1I2* polymorphisms

SNP	G allele frequency	*P* value	OR (CI)
rs2461823 (A > G)			
NAFLD vs. control			
Overall	0.74 vs. 0.60	0.291	1.18 (0.87-1.61)
Malays	0.65 vs. 0.65	0.986	1.01 (0.60-1.67)
Chinese	0.65 vs. 0.57	0.226	1.41 (0.81-2.44)
Indians	0.63 vs. 0.57	0.477	1.23 (0.69-2.20)
rs7643645 (A > G)			
NAFLD vs. control			
Overall	0.43 vs. 0.45	0.201	0.84 (0.65-1.10)
Malays	0.43 vs. 0.47	0.509	0.86 (0.55-1.35)
Chinese	0.44 vs. 0.43	0.896	1.04 (0.62-1.72)
Indians	0.44 vs. 0.43	0.945	1.02 (0.58-1.79)

CI confident interval, *OR* odds ratio, *NAFLD* non-alcoholic fatty liver disease

Analysis of comparison of clinical and histological parameters between the *NR1I2* rs2461823 genotypes among NAFLD patients is shown in Table 4.33. None of the parameters were different between the genotypes ($P > 0.05$).

Table 4.33 Comparison of various clinical and histological parameters between the *NR1I2* rs2461823 genotypes among NAFLD patients

Characteristics	NAFLD, *n* = 144 (mean ± SD)			
	AA (*n* = 18)	AG (*n* = 69)	GG (*n* = 57)	*P* value
Age (years)	51.9 ± 13.6	50.5 ± 11.0	51.6 ± 12.8	0.835
BMI (kg/m^2)	29.1 ± 4.1	28.9 ± 4.9	28.2 ± 3.8	0.602
HbA1c (%)*	6.6 ± 1.6	6.5 ± 1.6	6.6 ± 1.8	0.934
Waist circumference (cm)	95.7 ± 11.6	93.3 ± 11.7	93.9 ± 9.7	0.723
HDL cholesterol (mg/dl)	50.0 ± 14.2	48.1 ± 12.2	48.6 ± 13.0	0.852
LDL cholesterol (mg/dl)	117.4 ± 39.6	116.4 ± 42.4	117.9 ± 37.8	0.976
Total cholesterol (mg/dl)	193.3 ± 39.5	198.4 ± 46.1	195.9 ± 43.3	0.888
Triglycerides (mg/dl)	141.4 ± 41.6	156.6 ± 67.1	157.4 ± 63.1	0.621
AST (IU/L)*	35.4 ± 23.5	46.4 ± 26.0	41.0 ± 24.8	0.104
ALT (IU/L)	67.2 ± 51.8	85.8 ± 43.0	84.3 ± 53.4	0.333
GGT (IU/L)*	86.4 ± 72.1	129.8 ± 135.8	97.6 ± 96.3	0.408
Systolic BP (mmHg)	134.5 ± 16.8	129.5 ± 13.5	133.5 ± 16.1	0.264
Diastolic BP (mmHg)*	82.3 ± 10.6	81.9 ± 8.8	83.3 ± 10.9	0.724
Steatosis grade*	1.4 ± 0.6	1.9 ± 0.8	1.8 ± 0.8	0.094
Lobular inflammation*	1.3 ± 0.6	1.3 ± 0.6	1.3 ± 0.6	0.857
Ballooning*	1.2 ± 0.6	1.3 ± 0.6	1.2 ± 0.5	0.909
Fibrosis*	2.0 ± 1.2	1.7 ± 1.0	1.6 ± 0.9	0.569

**P* values obtained using Kruskal-Wallis test, all other comparisons used ANOVA

ALT alanine transferase, *AST* aspartate aminotransferase, *BMI* body mass index, *GGT* gamma glutamyl transpeptidase, *HbA1c* haemoglobin A1c, *HDL* high-density lipoprotein, *LDL* low-density lipoprotein, *NAFLD* non-alcoholic fatty liver disease

4.2.15 Single nucleotide polymorphism: *KLF6*

A SNP of *KLF6* gene rs3750861 C/T was genotyped in this study. The rs3750861 is a non-coding SNP, located in the intronic region on chromosome 10. Table 4.34 shows the association test between NAFLD and control subjects with NAFLD. The genotypes were in Hardy-Weinberg equilibrium. No allelic association with NAFLD was observed in the pooled subjects (OR 1.09, 95% CI 0.61-1.96, *P* = 0.778) nor after stratification by ethnicity (OR 1.36, 95% CI 0.58-3.20, *P* = 0.485; OR 1.33, 95% CI 0.28-6.27, *P* = 0.716; and OR 0.98, 95% CI 0.36-2.72, *P* = 0.974, for Malay, Chinese and Indian, respectively).

Table 4.34 Allelic association of *KLF6* polymorphism

SNP	T allele frequency	*P* value	OR (CI)
rs3750861 (C > T)			
NAFLD vs. control			
Overall	0.07 vs. 0.07	0.778	1.09 (0.61-1.96)
Malays	0.09 vs. 0.06	0.485	1.36 (0.58-3.20)
Chinese	0.04 vs. 0.03	0.716	1.33 (0.28-6.27)
Indians	0.10 vs. 0.10	0.974	0.98 (0.36-2.72)

CI confident interval, *OR* odds ratio, *NAFLD* non-alcoholic fatty liver disease

Analysis of comparison of clinical and histological parameters between the *KLF6* rs3750861 genotypes among NAFLD patients is shown in Table 4.35. Systolic blood pressure was different between the genotypes in which the highest was in those with TT genotype (*P* = 0.004).

Table 4.35 Comparison of various clinical and histological parameters between the *KLF6* rs3750861 genotypes among NAFLD patients

Characteristics	NAFLD, *n* = 144 (mean ± SD)			
	CC (*n* = 126)	CT (*n* = 16)	TT (*n* = 2)	*P* value
Age (years)	50.98 ± 12.2	52.5 ± 11.5	59.0 ± 7.1	0.572
BMI (kg/m^2)	28.7 ± 4.4	27.7 ± 3.7	34.8 ± 4.0	0.602
HbA1c (%)*	6.6 ± 1.7	6.2 ± 1.1	6.5 ± 0.6	0.890
Waist circumference (cm)	93.8 ± 11.2	92.4 ± 8.2	104.5 ± 4.9	0.336
HDL cholesterol (mg/dl)	48.7 ± 12.4	46.1 ± 15.7	53.6 ± 8.5	0.606
LDL cholesterol (mg/dl)	116.7 ± 39.6	118.5 ± 44.9	135.3 ± 24.8	0.800
Total cholesterol (mg/dl)	196.5 ± 44.4	196.4 ± 43.7	212.4 ± 32.8	0.881
Triglycerides (mg/dl)	155.7 ± 63.4	152.9 ± 61.6	127.0 ± 11.9	0.807
AST (IU/L)*	42.9 ± 25.4	45.1 ± 27.0	25.5 ± 7.8	0.488
ALT (IU/L)	82.7 ± 47.2	89.8 ± 60.5	51.0 ± 1.4	0.553
GGT (IU/L)*	107.4 ± 106.1	151.0 ± 175.9	39.5 ± 6.3	0.457
Systolic BP (mmHg)	131.1 ± 15.0	130.8 ± 10.4	166.5 ± 2.1	0.004
Diastolic BP (mmHg)*	82.3 ± 9.9	82.0 ± 8.3	99.0 ± 1.4	0.096
Steatosis grade*	1.8 ± 0.8	1.8 ± 0.8	2.5 ± 0.7	0.443
Lobular inflammation*	1.3 ± 0.6	1.3 ± 0.6	1.5 ± 0.7	0.867
Ballooning*	1.2 ± 0.6	1.3 ± 0.6	1.0 ± 0.0	0.815
Fibrosis*	1.7 ± 1.0	1.8 ± 0.8	2.0 ± 0.0	0.737

**P* values obtained using Kruskal-Wallis test, all other comparisons used ANOVA

ALT alanine transferase, *AST* aspartate aminotransferase, *BMI* body mass index, *GGT* gamma glutamyl transpeptidase, *HbA1c* haemoglobin A1c, *HDL* high-density lipoprotein, *LDL* low-density lipoprotein, *NAFLD* non-alcoholic fatty liver disease

4.2.16 Single nucleotide polymorphism: *TFRC*

A SNP of *TFRC* gene rs3817672 G/A was genotyped in this study. The polymorphism of rs3817672 is a missense mutation, substituting the amino acid glycine with serine. The rs3817672 is located on chromosome 3. Table 4.36 shows the association test between NAFLD and control subjects with NAFLD. The genotypes were in Hardy-Weinberg equilibrium. No allelic association with NAFLD was observed in the pooled subjctes (OR 0.81, 95% CI 0.53-1.16, P = 0.245) nor after stratification by ethnicity (OR 0.99, 95% CI 0.56-1.75, P = 0.963; OR 1.08, 95% CI 0.49-2.39, P = 0.841; and OR 0.65, 95% CI 0.34-1.25, P = 0.200, for Malay, Chinese and Indian, respectively).

Table 4.36 Allelic association of *TFRC* polymorphism

SNP	A allele frequency	*P* value	OR (CI)
rs3817672 (G > A)			
NAFLD vs. control			
Overall	0.22 vs. 0.26	0.245	0.81 (0.53-1.16)
Malays	0.24 vs. 0.24	0.963	0.99 (0.56-1.75)
Chinese	0.16 vs. 0.15	0.841	1.08 (0.49-2.39)
Indians	0.28 vs. 0.38	0.200	0.65 (0.34-1.25)

CI confident interval, *OR* odds ratio, *NAFLD* non-alcoholic fatty liver disease

Analysis of comparison of clinical and histological parameters between the *TFRC* rs3817672 genotypes among NAFLD patients is shown in Table 4.37. None of the parameters were different between the genotypes (P > 0.05).

Table 4.37 Comparison of various clinical and histological parameters between the *TFRC* rs3817672 genotypes among NAFLD patients

Characteristics	NAFLD, *n* = 144 (mean ± SD)			
	GG (*n* = 87)	GA (*n* = 50)	AA (*n* = 7)	*P* value
Age (years)	51.1 ± 11.4	50.6 ± 13.5	55.0 ± 10.3	0.661
BMI (kg/m^2)	28.6 ± 4.5	28.6 ± 3.9	30.1 ± 5.8	0.602
HbA1c (%)*	6.5 ± 1.6	6.7 ± 1.9	6.0 ± 0.8	0.816
Waist circumference (cm)	94.2 ± 11.2	93.1 ± 10.0	94.6 ± 9.3	0.854
HDL cholesterol (mg/dl)	48.1 ± 13.9	49.2 ± 10.9	48.3 ± 10.7	0.894
LDL cholesterol (mg/dl)	116.4 ± 41.3	117.9 ± 39.1	119.9 ± 34.1	0.962
Total cholesterol (mg/dl)	195.5 ± 45.1	198.5 ± 42.7	201.4 ± 43.5	0.883
Triglycerides (mg/dl)	154.7 ± 61.3	152.7 ± 64.9	174.1 ± 70.1	0.701
AST (IU/L)*	43.3 ± 26.4	43.7 ± 24.3	32.9 ± 10.1	0.661
ALT (IU/L)	79.7 ± 41.2	91.3 ± 60.8	64.9 ± 19.9	0.242
GGT (IU/L)*	106.6 ± 114.9	124.8 ± 123.0	79.5 ± 45.6	0.410
Systolic BP (mmHg)	130.4 ± 14.7	132.5 ± 15.0	139.5 ± 18.2	0.251
Diastolic BP (mmHg)*	82.2 ± 10.2	82.5 ± 9.1	87.1 ± 11.1	0.490
Steatosis grade*	1.8 ± 0.8	1.8 ± 0.8	2.1 ± 0.7	0.423
Lobular inflammation*	1.3 ± 0.6	1.4 ± 0.5	1.3 ± 0.5	0.669
Ballooning*	1.2 ± 0.6	1.2 ± 0.6	1.1 ± 0.4	0.899
Fibrosis*	1.7 ± 1.0	1.8 ± 1.0	1.6 ± 1.1	0.770

**P* values obtained using Kruskal-Wallis test, all other comparisons used ANOVA

ALT alanine transferase, *AST* aspartate aminotransferase, *BMI* body mass index, *GGT* gamma glutamyl transpeptidase, *HbA1c* haemoglobin A1c, *HDL* high-density lipoprotein, *LDL* low-density lipoprotein, *NAFLD* non-alcoholic fatty liver disease

4.2.17 Single nucleotide polymorphism: *PPARGC1A*

A SNP of *PPARGC1A* gene rs3755863 G/A was genotyped in this study. The rs3755863 is located in the coding region on chromosome 4. The rs3755863 is a synonymous coding SNP where different alleles encode for the same amino acid threonine. Table 4.38 shows the association test between NAFLD and control subjects with NAFLD. The genotypes were in Hardy-Weinberg equilibrium. No allelic association with NAFLD was observed in the pooled subjects (OR 1.35, 95% CI 0.99-1.86, P = 0.061) nor after stratification by ethnicity (OR 1.61, 95% CI 0.95-2.73, P = 0.076; OR 1.40, 95% CI 0.82-2.41, P = 0.222; and OR 0.99, 95% CI 0.54-1.83, P = 0.984, for Malay, Chinese and Indian, respectively).

Table 4.38 Allelic association of *PPARGC1A* polymorphism

SNP	A allele frequency	*P* value	OR (CI)
rs3755863 (G > A)			
NAFLD vs. control			
Overall	0.49 vs. 0.41	0.061	1.35 (0.99-1.86)
Malays	0.55 vs. 0.45	0.076	1.61 (0.95-2.73)
Chinese	0.49 vs. 0.41	0.222	1.40 (0.82-2.41)
Indians	0.35 vs. 0.35	0.984	0.99 (0.54-1.83)

CI confident interval, *OR* odds ratio, *NAFLD* non-alcoholic fatty liver disease

Analysis of comparison of clinical and histological parameters between the *PPARGC1A* rs3755863 genotypes among NAFLD patients is shown in Table 4.39. Fibrosis score was found significantly higher in GG genotype (P = 0.046).

Table 4.39 Comparison of various clinical and histological parameters between the *PPARGC1A* rs3755863 genotypes among NAFLD patients

Characteristics	NAFLD, *n* = 144 (mean ± SD)			
	GG (*n* = 39)	GA (*n* = 69)	AA (*n* = 36)	*P* value
Age (years)	50.0 ± 12.8	52.2 ± 11.0	50.4 ± 13.2	0.604
BMI (kg/m^2)	28.1 ± 3.3	29.1 ± 5.1	28.3 ± 4.0	0.602
HbA1c (%)*	6.8 ± 1.8	6.5 ± 1.7	6.4 ± 1.5	0.660
Waist circumference (cm)	95.0 ± 10.2	93.9 ± 11.6	92.2 ± 10.4	0.544
HDL cholesterol (mg/dl)	46.7 ± 12.1	49.3 ± 13.5	48.9 ± 11.9	0.571
LDL cholesterol (mg/dl)	118.8 ± 39.2	114.6 ± 40.6	120.2 ± 40.5	0.763
Total cholesterol (mg/dl)	194.7 ± 45.4	196.4 ± 42.5	199.8 ± 46.1	0.880
Triglycerides (mg/dl)	145.4 ± 66.8	155.7 ± 65.0	164.7 ± 52.3	0.411
AST (IU/L)*	43.2 ± 28.0	42.1 ± 23.7	44.2 ± 26.0	0.915
ALT (IU/L)	82.1 ± 44.0	83.8 ± 54.3	82.9 ± 42.0	0.983
GGT (IU/L)*	90.6 ± 99.8	121.8 ± 125.8	115.5 ± 110.6	0.368
Systolic BP (mmHg)	130.0 ± 14.7	132.2 ± 13.2	132.2 ± 18.7	0.742
Diastolic BP (mmHg)*	81.5 ± 11.5	83.0 ± 9.1	82.8 ± 9.5	0.950
Steatosis grade*	1.8 ± 0.7	1.7 ± 0.8	2.1 ± 0.8	0.089
Lobular inflammation*	1.4 ± 0.5	1.3 ± 0.6	1.3 ± 0.6	0.590
Ballooning*	1.3 ± 0.5	1.2 ± 0.6	1.1 ± 0.6	0.587
Fibrosis*	2.1 ± 1.0	1.6 ± 1.0	1.5 ± 0.9	0.046

**P* values obtained using Kruskal-Wallis test, all other comparisons used ANOVA

ALT alanine transferase, *AST* aspartate aminotransferase, *BMI* body mass index, *GGT* gamma glutamyl transpeptidase, *HbA1c* haemoglobin A1c, *HDL* high-density lipoprotein, *LDL* low-density lipoprotein, *NAFLD* non-alcoholic fatty liver disease

4.2.18 Single nucleotide polymorphism: *FATP5*

A SNP of *FATP5* gene rs56225452 C/T was genotyped in this study. The rs56225452 is a non-coding SNP, located in the locus region near-gene-5 on chromosome 19. Table 4.40 shows the association test between NAFLD and control subjects with NAFLD. The genotypes were in Hardy-Weinberg equilibrium. No allelic association with NAFLD was observed in the pooled subjects (OR 0.90, 95% CI 0.53-1.51, *P* = 0.681) nor after stratification by ethnicity (OR 0.92, 95% CI 0.41-2.06, *P* = 0.835; OR 3.12, 95% CI 0.31-11.02, *P* = 0.331; and OR 1.02, 95% CI 0.46-2.27, *P* = 0.952, for Malay, Chinese and Indian, respectively).

Table 4.40 Allelic association of *FATP5* polymorphism

SNP	T allele frequency	*P* value	OR (CI)
rs56225452 (C > T)			
NAFLD vs. control			
Overall	0.09 vs. 0.10	0.681	0.90 (0.53-1.51)
Malays	0.09 vs. 0.09	0.835	0.92 (0.41-2.06)
Chinese	0.03 vs. 0.01	0.331	3.12 (0.31-11.02)
Indians	0.18 vs. 0.18	0.952	1.02 (0.46-2.27)

CI confident interval, *OR* odds ratio, *NAFLD* non-alcoholic fatty liver disease

Analysis of comparison of clinical and histological parameters between the *FATP5* rs56225452 genotypes among NAFLD patients is shown in Table 4.41. Difference was observed in triglycerides level (*P* = 0.028) and fibrosis score (*P* = 0.018) between the genotypes.

Table 4.41 Comparison of various clinical and histological parameters between the *FATP5* rs56225452 genotypes among NAFLD patients

Characteristics	NAFLD, *n* = 144 (mean ± SD)			
	CC (*n* = 121)	CT (*n* = 19)	TT (*n* = 4)	*P* value
Age (years)	50.5 ± 12.0	53.7 ± 10.6	52.0 ± 19.7	0.566
BMI (kg/m^2)	28.9 ± 4.4	27.5 ± 4.3	25.5 ± 0.9	0.193
HbA1c (%)*	6.5 ± 1.6	7.2 ± 2.3	7.3 ± 1.3	0.144
Waist circumference (cm)	94.0 ± 11.2	92.9 ± 9.9	91.0 ± 7.0	0.822
HDL cholesterol (mg/dl)	49.1 ± 12.0	45.6 ± 17.0	49.9 ± 11.4	0.550
LDL cholesterol (mg/dl)	116.4 ± 38.7	113.6 ± 43.6	116.2 ± 16.2	0.966
Total cholesterol (mg/dl)	195.7 ± 43.5	198.4 ± 46.1	184.0 ± 19.8	0.868
Triglycerides (mg/dl)	151.1 ± 56.0	188.1 ± 72.8	109.7 ± 21.1	0.028
AST (IU/L)*	41.9 ± 24.7	47.7 ± 27.2	59.0 ± 50.3	0.617
ALT (IU/L)	81.2 ± 48.9	94.2 ± 49.9	92.0 ± 52.8	0.553
GGT (IU/L)*	105.6 ± 110.2	153.5 ± 148.3	133.7 ± 152.5	0.596
Systolic BP (mmHg)	130.9 ± 15.0	132.6 ± 16.2	135.0 ± 11.5	0.826
Diastolic BP (mmHg)*	82.4 ± 9.9	84.1 ± 9.6	79.0 ± 14.9	0.669
Steatosis grade*	1.8 ± 0.8	1.7 ± 0.8	2.0 ± 0.0	0.550
Lobular inflammation*	1.3 ± 0.6	1.1 ± 0.5	1.7 ± 0.6	0.164
Ballooning*	1.2 ± 0.6	1.0 ± 0.7	1.7 ± 0.6	0.143
Fibrosis*	1.7 ± 1.0	1.4 ± 0.8	3.3 ± 0.6	0.018

**P* values obtained using Kruskal-Wallis test, all other comparisons used ANOVA

ALT alanine transferase, *AST* aspartate aminotransferase, *BMI* body mass index, *GGT* gamma glutamyl transpeptidase, *HbA1c* haemoglobin A1c, *HDL* high-density lipoprotein, *LDL* low-density lipoprotein, *NAFLD* non-alcoholic fatty liver disease

4.2.19 Single nucleotide polymorphism: *STAT3*

A total of two SNPs of *STAT3* gene rs6503695 T/C and rs9891119 A/C were genotyped in this study. Both are non-coding SNPs, located in the intronic regions on chromosome 17. Table 4.42 shows the association test between NAFLD and control subjects with NAFLD. The genotypes of rs6503695 were in Hardy-Weinberg equilibrium but not for rs9891119 . No allelic association with NAFLD was observed for both SNPs in the pooled subjects ($P > 0.05$) nor after stratification by ethnicity ($P > 0.05$).

Table 4.42 Allelic association of *STAT3* polymorphism

SNP	C allele frequency	*P* value	OR (CI)
rs6503695 (T > C)			
NAFLD vs. control			
Overall	0.49 vs. 0.42	0.064	1.35 (0.98-1.86)
Malays	0.48 vs. 0.44	0.486	1.20 (0.71-2.03)
Chinese	0.49 vs. 0.44	0.499	1.20 (0.71-2.05)
Indians	0.50 vs. 0.36	0.075	1.77 (0.94-3.32)
rs9891119 (A > C)			
NAFLD vs. control	0.24 vs. 0.26	0.194	0.84 (0.64-1.09)

CI confident interval, *OR* odds ratio, *NAFLD* non-alcoholic fatty liver disease

Analysis of comparison of clinical and histological parameters between the *STAT3* rs6503695 genotypes among NAFLD patients is shown in Table 4.43. No difference in parameters was observed between the genotypes ($P > 0.05$).

Table 4.43 Comparison of various clinical and histological parameters between the *STAT3* rs6503695 genotypes among NAFLD patients

Characteristics	NAFLD, *n* = 144 (mean ± SD)			
	TT (*n* = 34)	TC (*n* = 76)	CC (*n* = 31)	*P* value
Age (years)	50.6 ± 10.8	50.8 ± 12.5	51.7 ± 12.0	0.929
BMI (kg/m^2)	28.4 ± 4.3	28.7 ± 4.5	28.9 ± 4.4	0.872
HbA1c (%)*	6.5 ± 1.8	6.7 ± 1.8	6.3 ± 1.0	0.464
Waist circumference (cm)	93.8 ± 8.8	93.5 ± 11.4	94.3 ± 11.7	0.954
HDL cholesterol (mg/dl)	49.6 ± 11.0	48.3 ± 12.4	48.5 ± 15.0	0.882
LDL cholesterol (mg/dl)	115.2 ± 35.4	117.6 ± 41.2	114.1 ± 37.3	0.913
Total cholesterol (mg/dl)	195.1 ± 40.3	198.8 ± 44.6	189.2 ± 43.4	0.577
Triglycerides (mg/dl)	154.2 ± 64.8	159.2 ± 61.6	145.3 ± 63.2	0.581
AST (IU/L)*	40.9 ± 23.5	45.1 ± 25.2	40.6 ± 28.9	0.488
ALT (IU/L)	78.2 ± 50.0	89.7 ± 52.2	72.4 ± 37.3	0.553
GGT (IU/L)*	113.7 ± 138.0	119.9 ± 106.3	92.2 ± 116.8	0.083
Systolic BP (mmHg)	127.0 ± 11.6	132.2 ± 16.1	133.5 ± 15.0	0.158
Diastolic BP (mmHg)*	79.7 ± 8.6	83.4 ± 10.0	83.4 ± 10.8	0.169
Steatosis grade*	1.8 ± 0.7	1.9 ± 0.8	1.7 ± 0.8	0.757
Lobular inflammation*	1.4 ± 0.5	1.2 ± 0.6	1.4 ± 0.6	0.465
Ballooning*	1.3 ± 0.5	1.2 ± 0.7	1.1 ± 0.5	0.424
Fibrosis*	1.6 ± 0.8	1.7 ± 0.9	1.9 ± 1.2	0.568

**P* values obtained using Kruskal-Wallis test, all other comparisons used ANOVA

ALT alanine transferase, *AST* aspartate aminotransferase, *BMI* body mass index, *GGT* gamma glutamyl transpeptidase, *HbA1c* haemoglobin A1c, *HDL* high-density lipoprotein, *LDL* low-density lipoprotein, *NAFLD* non-alcoholic fatty liver disease

4.2.20 Single nucleotide polymorphism: *APOA5*

A SNP of *APOA5* gene rs662799 G/A was genotyped in this study. The rs662799 is a non-coding SNP, located in the locus region near-gene-5 on chromosome 11. Table 4.44 shows the association test between NAFLD and control subjects with NAFLD. The genotypes were in Hardy-Weinberg equilibrium. One sample from the controls was with no call, thus in total 144 cases and 197 controls were genotyped. No allelic association with NAFLD was observed in the pooled subjects (OR 1.39, 95% CI 0.98-1.95, P = 0.063) nor after stratification by ethnicity (OR 1.18, 95% CI 0.71-1.96, P = 0.519; OR 1.32, 95% CI 0.72-2.44, P = 0.359; and OR 2.06, 95% CI 0.92-4.63, P = 0.080, for Malay, Chinese and Indian, respectively).

Table 4.44 Allelic association of *APOA5* polymorphism

SNP	A allele frequency	*P* value	OR (CI)
rs662799 (G > A)			
NAFLD vs. control			
Overall	0.76 vs. 0.70	0.063	1.39 (0.98-1.95)
Malays	0.69 vs. 0.65	0.519	1.18 (0.71-1.96)
Chinese	0.77 vs. 0.72	0.359	1.32 (0.72-2.44)
Indians	0.87 vs. 0.74	0.080	2.06 (0.92-4.63)

CI confident interval, *OR* odds ratio, *NAFLD* non-alcoholic fatty liver disease

Analysis of comparison of clinical and histological parameters between the *APOA5* rs662799 genotypes among NAFLD patients is shown in Table 4.45. Difference was observed in triglycerides level in which highest level was shown in those with GG genotype (P = 0.001).

Table 4.45 Comparison of various clinical and histological parameters between the *APOA5* rs662799 genotypes among NAFLD patients

Characteristics	NAFLD, *n* = 144 (mean ± SD)			
	GG (*n* = 8)	GA (*n* = 54)	AA (*n* = 82)	*P* value
Age (years)	49.9 ± 9.9	50.8 ± 12.4	51.2 ± 11.8	0.958
BMI (kg/m^2)	29.4 ± 3.8	28.5 ± 4.2	28.7 ± 4.6	0.875
HbA1c (%)*	6.7 ± 1.7	6.3 ± 1.4	6.7 ± 1.8	0.471
Waist circumference (cm)	95.9 ± 4.7	93.3 ± 9.3	93.9 ± 12.1	0.823
HDL cholesterol (mg/dl)	47.4 ± 16.3	48.7 ± 12.7	48.7 ± 12.5	0.968
LDL cholesterol (mg/dl)	111.1 ± 42.6	114.6 ± 41.3	117.5 ± 37.2	0.863
Total cholesterol (mg/dl)	198.1 ± 43.0	198.8 ± 45.6	193.6 ± 42.0	0.786
Triglycerides (mg/dl)	199.4 ± 48.6	172.4 ± 64.6	139.3 ± 57.6	0.001
AST (IU/L)*	41.9 ± 21.6	44.2 ± 27.9	42.4 ± 24.5	0.935
ALT (IU/L)	88.2 ± 41.0	85.7 ± 48.9	81.0 ± 50.0	0.832
GGT (IU/L)*	98.0 ± 43.7	133.2 ± 139.4	99.4 ± 102.1	0.436
Systolic BP (mmHg)	135.0 ± 12.3	131.4 ± 14.6	130.8 ± 15.6	0.779
Diastolic BP (mmHg)*	87.9 ± 10.4	80.8 ± 9.7	83.2 ± 9.9	0.098
Steatosis grade*	2.0 ± 0.8	1.8 ± 0.8	1.8 ± 0.8	0.785
Lobular inflammation*	1.6 ± 0.5	1.3 ± 0.5	1.3 ± 0.6	0.445
Ballooning*	1.1 ± 0.7	1.2 ± 0.6	1.3 ± 0.6	0.489
Fibrosis*	1.6 ± 1.4	1.6 ± 0.9	1.8± 1.0	0.650

**P* values obtained using Kruskal-Wallis test, all other comparisons used ANOVA

ALT alanine transferase, *AST* aspartate aminotransferase, *BMI* body mass index, *GGT* gamma glutamyl transpeptidase, *HbA1c* haemoglobin A1c, *HDL* high-density lipoprotein, *LDL* low-density lipoprotein, *NAFLD* non-alcoholic fatty liver disease

4.2.21 Single nucleotide polymorphism: *PEMT*

A SNP of *PEMT* gene rs7946 C/T was genotyped in this study. The polymorphism of rs7946 is a missense mutation, substituting the amino acid valine with methionine. The rs7946 is located on chromosome 17. Table 4.46 shows the association test between NAFLD and control subjects with NAFLD. The genotypes were in Hardy-Weinberg equilibrium. Two samples from the controls and one from cases were with no call, thus in total 143 cases and 196 controls were genotyped. No allelic association with NAFLD was observed in the pooled subjects (OR 0.91, 95% CI 0.65-1.26, P = 0.551) nor after stratification by ethnicity (OR 1.12, 95% CI 0.67-1.87, P = 0.663; OR 1.15, 95% CI 0.55-2.42, P = 0.706; and OR 0.78, 95% CI 0.43-1.42, P = 0.409, for Malay, Chinese and Indian, respectively).

Table 4.46 Allelic association of *PEMT* polymorphism

SNP	T allele frequency	*P* value	OR (CI)
rs7946 (C > T)			
NAFLD vs. control			
Overall	0.28 vs. 0.32	0.551	0.91 (0.65-1.26)
Malays	0.28 vs. 0.26	0.663	1.12 (0.67-1.87)
Chinese	0.19 vs. 0.17	0.706	1.15 (0.55-2.42)
Indians	0.43 vs. 0.50	0.409	0.78 (0.43-1.42)

CI confident interval, *OR* odds ratio, *NAFLD* non-alcoholic fatty liver disease

Analysis of comparison of clinical and histological parameters between the *PEMT* rs7946 genotypes among NAFLD patients is shown in Table 4.47. There was no difference in parameters observed between the genotypes (P > 0.05)

Table 4.47 Comparison of various clinical and histological parameters between the *PEMT* rs7946 genotypes among NAFLD patients

Characteristics	NAFLD, *n* = 144 (mean ± SD)			
	CC (*n* = 73)	CT (*n* = 55)	TT (*n* = 12)	*P* value
Age (years)	51.3 ± 12.1	49.2 ± 12.3	56.9 ± 7.2	0.117
BMI (kg/m^2)	28.8 ± 4.3	28.3 ± 4.5	29.2 ± 4.6	0.716
HbA1c (%)*	6.6 ± 1.6	6.5 ± 1.7	6.9 ± 1.8	0.466
Waist circumference (cm)	93.1 ± 10.7	94.5 ± 11.4	94.0 ± 10.6	0.778
HDL cholesterol (mg/dl)	48.9 ± 10.5	47.6 ± 14.2	51.9 ± 17.5	0.553
LDL cholesterol (mg/dl)	117.1 ± 36.1	113.6 ± 39.5	120.0 ± 53.7	0.821
Total cholesterol (mg/dl)	198.2 ± 41.1	191.1 ± 43.8	202.1 ± 56.5	0.572
Triglycerides (mg/dl)	159.9 ± 67.4	150.5 ± 60.7	146.5 ± 39.0	0.627
AST (IU/L)*	43.9 ± 27.6	41.4 ± 24.6	44.3 ± 17.7	0.655
ALT (IU/L)	80.7 ± 44.5	85.4 ± 55.6	84.9 ± 47.5	0.867
GGT (IU/L)*	120.5 ± 134.1	101.0 ± 91.7	117.1 ± 113.1	0.639
Systolic BP (mmHg)	130.2 ± 15.2	131.5 ± 15.3	136.4 ± 12.8	0.411
Diastolic BP (mmHg)*	83.5 ± 9.8	80.9 ± 10.7	84.1 ± 7.3	0.139
Steatosis grade*	1.9 ± 0.8	1.7 ± 0.7	1.9 ± 0.9	0.547
Lobular inflammation*	1.3 ± 0.6	1.3 ± 0.6	1.5 ± 0.5	0.453
Ballooning*	1.3 ± 0.6	1.2 ± 0.6	1.3 ± 0.5	0.573
Fibrosis*	1.7 ± 0.9	1.6 ± 1.0	2.3 ± 1.1	0.207

**P* values obtained using Kruskal-Wallis test, all other comparisons used ANOVA

ALT alanine transferase, *AST* aspartate aminotransferase, *BMI* body mass index, *GGT* gamma glutamyl transpeptidase, *HbA1c* haemoglobin A1c, *HDL* high-density lipoprotein, *LDL* low-density lipoprotein, *NAFLD* non-alcoholic fatty liver disease

4.2.22 Single nucleotide polymorphism: Other candidate genes

Farnesyl-diphosphate farnesyltransferase 1 gene (*FDFT1*) rs2645424 C/T and inducible nitric oxide synthase (*iNOS*) rs2797512 A/C are non-coding SNPs in the intronic region on chromosome 8 and 9. The genotype distribution of both SNPs deviated from Hardy-Weinberg equilibrium. The association test did not proceed since more than 30% of the samples failed to be genotyped.

4.3 Gene-gene interaction analysis

4.3.1 *AGTR1* and *PNPLA3* gene interaction

The overall results, pooling all subjects, indicate a lack of evidence for an association between *AGTR1* gene and NAFLD, which is in contrast with the positive findings by Yoneda et al. (2009b). The *PNPLA3* study showed a strong association of *PNPLA3* variant with the occurence of NAFLD (Zain et al., 2012a), and therefore attempt was made to find out whether there is any gene-gene interaction between *PNPLA3* and *AGTR1*. The results with the two-locus and three-locus models are shown in Table 4.48. The two-locus model has the highest testing accuracy while the three-locus model has the best cross-validation consistency. Investigation of the gene-gene interaction on the occurence of NAFLD, identified a significant effect of *PNPLA3* and *AGTR1* interaction (empirical P = 0.007). The best GMDR model suggested for the interaction was the three-locus model. The significant interaction was confirmed by logistic regression (P = 0.017).

Table 4.48 Best fitted gene-gene interaction model between *AGTR1* and *PNPLA3*

Locus number	Model	Cross-validation consistency	Testing accuracy (%)	*P* value*
2	*AGTR1* (rs2276736), *PNPLA3* (rs738409)	8/10	63.68	0.038
3	*AGTR1* (rs3772627), *AGTRI* (rs3772630), *PNPLA3* (rs738409)	9/10	62.88	0.007

**P* values based on 1000 permutations. Analysis of GMDR with adjustment of ethnicity

Genotype association in additive model revealed that the risk genotypes (*AGTR1* rs3772627 TT, *AGTR1* rs3772630 AA and *PNPLA3* rs738409 GG) versus one risk factor (*AGTR1* rs3772627 TC, *AGTR1* rs3772630 AG and *PNPLA3* rs738409 CG) versus the protective genotypes (*AGTR1* rs3772627 CC, *AGTR1* rs3772630 GG and *PNPLA3* rs738409 CC), conferred enhanced risk for NAFLD (OR 2.23, 95% CI 1.16-4.31, $P = 0.017$). Since the protective effect was seen in the Indian ethnic subgroup, the association was additionally tested specifically in this group. The risk for NAFLD in this group was dramatically enhanced (OR 12.41, 95% CI 1.61-55.42, $P = 0.016$) following genotypes combinations. The association of *AGTR1* SNPs with NAFLD following stratification by *PNPLA3* genotypes was further evaluated. The results however indicated the absence of significant association.

4.3.2 *LEPR* and *PNPLA3* gene interaction

Functionally, both *LEPR* and *PNPLA3* genes play important roles in lipid regulation. An analysis was performed to find out whether there is an interaction between the two genes on susceptibility to NAFLD. The results of the gene-gene interaction analysis are demonstrated in Table 4.49. There are two best combinations (two-locus and three-locus) yielded by the analysis. Both two-locus and three-locus models have a perfect cross-validation consistency and similar testing accuracy. Investigation of the gene-gene interaction on the susceptibility to NAFLD identified a significant effect of *LEPR* rs1137100 and *PNPLA3* rs738409 interaction (empirical $P = 0.001$ and empirical $P = 0.011$, for rs1137100 and rs1137101 respectively). The best GMDR model suggested for the interaction was the two-locus model. The significant interaction was confirmed by logistic regression ($p < 0.0001$).

Table 4.49 Best fitted gene-gene interaction model between *LEPR* and *PNPLA3*

Locus number	Model	Cross-validation consistency	Testing accuracy (%)	*P* value*
2	*LEPR* (rs1137100), *PNPLA3* (rs738409)	10/10	63.3	0.001
3	*LEPR* (rs1137100), *LEPR* (rs1137101), *PNPLA3* (rs738409)	10/10	63.0	0.011

[a]*P* value based on 1000 permutations. Analysis of GMDR with adjustment of ethnicity

The effects of combined alleles from each variant of the two different genes, *LEPR* rs1137100 and *PNPLA3* rs738409 were further evaluated in additive model, and tested as follows; two risk factors vs. one risk factor vs. no risk factor. The risk of NAFLD is first to be noted as 1.64 for *LEPR* rs1137100 and 2.23 for *PNPLA3* rs738409. Although not shown in the table, it is noteworthy that the risk of NAFLD increased to 3.73 with the combined effect of risk allele of both SNPs (OR 3.73, 95% CI 1.84-7.55, $P < 0.0001$).

4.3.3 *GCKR* and *PNPLA3* gene interaction

Variants of both the *GCKR* and *PNPLA3* genes (Zain et al., 2012a) have been postulated to confer risk of NAFLD in this study population and involved in hepatic fat content. Hence, the interaction between the two genes on risk of NAFLD was investigated (Table 4.50). Two best models were derived from the interaction. Having a perfect cross-validation consistency, the three-locus model (*GCKR* rs1260326, *GCKR* rs780094, *PNPLA3* rs738409) was suggested to be the best model (empirical $P = 0.003$).

Table 4.50 Best fitted gene-gene interaction model between *GCKR* and *PNPLA3*

Locus number	Model	Cross-validation consistency	Testing accuracy (%)	*P* value*
2	*GCKR* (rs780094), *PNPLA3* (rs738409)	6/10	61.0	0.001
3	*GCKR* (rs1260326), *GCKR* (rs780094), *PNPLA3* (rs738409)	10/10	59.0	0.003

[a]*P* value based on 1000 permutations. Analysis of GMDR with adjustment of ethnicity

This finding was confirmed by performing the additive logistic regression model that takes into account all risk alleles of three SNPs. Analysis yielded an odds ratio of 3.87 with the combined effect of the three SNPs (OR 3.87, 95% CI 1.29-11.61, *P* = 0.016) which is a greater risk of NAFLD than either SNP alone: 1.49 risk for *GCKR* rs1260326, 1.51 risk for *GCKR* rs780094 and 2.23 risk for *PNPLA3* rs738409.

4.4 Copy number variation

Analysis of copy number on the basis of log ratio and probe incidence filtering, exhibited a total of 521 copy number alterations (CNAs) ($P < 0.05$). Molecular genomic profiling successfully identified five genomic regions; 14q11.2 (n = 25, 62.5%), 12p13.31 (n = 20, 50.0%), 11p15.4 (n = 18, 45.0%), 5p15.33 (n = 17, 42.5%), and 12p13.2 (n = 16, 40.0%) that were most frequently amplified among NASH patients (Table 4.51). The most frequently deleted CNAs were located on region 11q11 (n = 14, 35.0%), 12p13.2 (n = 13, 32.5%), 16q12.2 (n = 13, 32.5%), 4q13.2 (n = 13, 32.5%), 1q21.3 (n = 13, 32.5%), and 14q24.3 (n = 13, 32.5%) (Table 4.52). There are 11 amplified CNAs present in at least 1/3 of the samples. On the other hand, only 6 deleted CNAs were present in at least 1/3 of the samples. CNAs involving chromosome 11q11, 12p13.2 and 16q12.2 were informative both as gain and loss events. They exhibited almost similar number of samples for gain and loss events which may suggest a generally unstable nature of these regions. There was one CNA that showed a clear-cut gain, 16p12.2 (n = 11, 27.5%), which could be involved in the development of the disease and could be considered as a potential biomarker candidate for NAFLD. More than 50% of genomic regions were informative as gains only, however, it appears that the regions were presented with not more than 27.5% (n = 11) (Table 4.53). Meanwhile, less than 10% of the genomic regions were informative as losses only. These regions bear a frequency of not more than 7.5% (n = 3) (Table 4.54). The percentages in the parentheses indicate the frequency of each event. The most commonly amplified region, 14q11.2 contains a few olfactory receptor (*OR*) and T cell receptor family genes whereas forkhead box genes (*FOX*) are located on the second most amplified chromosome 12p13.31.

Table 4.51 Top 5% most amplified CNAs

CYTOBAND	NASH Gain (%)	NASH Loss (%)	CNV COVERAGE	CNV in DGV	Size (Mb)	Start	End
14q11.2	62.5	15.0	100	Yes	0.02	19376762	20420849
12p13.31	50.0	17.5	100	Yes	0.08	9637323	9718846
11p15.4	45.0	30.0	100	Yes	0.02	8500104	8598491
5p15.33	42.5	7.5	100	Yes	0.10	449199	673013
12p13.2	40.0	32.5	100	Yes	0.04	11218244	11256467
16q12.2	37.5	32.5	100	Yes	0.02	55832511	55853358
1q44	35.0	15.0	100	Yes	0.07	248601802	248672268
6p21.32	35.0	27.5	100	Yes	0.17	32413439	32592800
8p23.1	32.5	5.0	100	Yes	0.41	7169490	7786708
11q11	32.5	35.0	100	Yes	0.08	55368154	55450788
13q21.1	32.5	15.0	100	Yes	0.02	57760478	57783723
3q29	30.0	12.5	100	Yes	0.10	195354124	195456540
3q26.1	30.0	20.0	100	Yes	0.07	162514534	162619141
6p25.3	30.0	20.0	100	Yes	0.08	299363	378956
4q13.2	27.5	32.5	100	Yes	0.09	69392545	69483277
7q31.1	27.5	25.0	100	Yes	0.01	111227035	111333959
16p12.2	27.5	0	100	Yes	0.02	21570472	21585470
22q11.23	27.5	27.5	100	Yes	0.05	24347959	24395353
5q35.3	25.0	20.0	100	Yes	0.02	180410254	180429788
10q11.22	25.0	2.5	100	Yes	0.09	46968072	47551468
12p11.21	25.0	5.0	100	Yes	0.41	31277996	31346297
14q32.33	25.0	17.5	0	No	0.03	106531557	106559103
15q11.2	25.0	12.5	100	Yes	0.12	20172544	22409391
22q13.1	25.0	22.5	100	Yes	0.03	39359112	39385485
1p31.1	22.5	7.5	100	Yes	0.03	72768855	72795480

CNV copy number variation, *DGV* Database of Genomic Variants, *NASH* non-alcoholic steatohepatitis

Table 4.52 Top 5% most deleted CNAs

CYTOBAND	NASH Gain (%)	NASH Loss (%)	CNV COVERAGE	CNV in DGV	Size (Mb)	Start	End
11q11	32.5	35.0	100	Yes	0.08	55368154	55450788
12p13.2	40.0	32.5	100	Yes	0.04	11218244	11256467
16q12.2	37.5	32.5	100	Yes	0.02	55832511	55853358
4q13.2	27.5	32.5	100	Yes	0.09	69392545	69483277
1q21.3	20.0	32.5	100	Yes	0.02	152569857	152586281
14q24.3	20.0	32.5	100	Yes	0.02	78329728	78368494
11p15.4	45.0	30.0	100	Yes	0.02	8500104	8598491
6p21.32	35.0	27.5	100	Yes	0.17	32413439	32592800
22q11.23	27.5	27.5	100	Yes	0.05	24347959	24395353
7q31.1	27.5	25.0	100	Yes	0.01	111227035	111333959
1q24.2	12.5	25.0	100	Yes	0.01	169227144	169241333
22q13.1	25.0	22.5	100	Yes	0.03	39359112	39385485
3q26.1	30.0	20.0	100	Yes	0.07	162514534	162619141
6p25.3	30.0	20.0	100	Yes	0.08	299363	378956
5q35.3	25.0	20.0	100	Yes	0.02	180410254	180429788
16p11.2	20.0	20.0	100	Yes	0.80	32573808	32651084
1q21.2	12.5	20.0	100	Yes	0.20	149041933	149241652
12p13.31	50.0	17.5	100	Yes	0.08	9637323	9718846
14q32.33	25.0	17.5	0	No	0.03	106531557	106559103
15q11.1-q11.2	20.0	17.5	0	No	1.98	20432851	22409391
6p22.1	17.5	17.5	100	Yes	0.05	29854870	29961707
20p13	15.0	17.5	100	Yes	0.02	4769772	4785367
14q11.2	62.5	15.0	100	Yes	0.02	19376762	20420849
1q44	35.0	15.0	100	Yes	0.07	248601802	248672268
13q21.1	32.5	15.0	100	Yes	0.02	57760478	57783723

CNV copy number variation, *DGV* Database of Genomic Variants, *NASH* non-alcoholic steatohepatitis

Table 4.53 Top 5% copy number gain only

CYTOBAND	NASH Gain (%)	CNV COVERAGE	CNV in DGV	Size (Mb)	Start	End
16p12.2	27.5	100	Yes	0.02	21570472	21585470
12q24.33	17.5	0	No	0.02	131797898	131815396
7q36.3	15.0	0	No	0.02	158243089	158279162
15q22.31	15.0	0	No	0.31	66254326	66562391
19q13.41	15.0	100	Yes	0.05	52186797	52240114
22q13.2	15.0	0	No	0.05	42907649	42957173
2q37.1	12.5	0	No	0.06	234502414	234563295
7q36.1	12.5	59	Yes	0.33	151555495	151900894
7p14.1	12.5	0	No	0.02	38311203	38329752
9p21.1	12.5	0	No	0.01	28610932	28739593
11p15.1	12.5	0	No	0.20	17357360	17559681
21p11.2-p11.1	12.5	0	No	0.14	10861051	11000494
2q35	10.0	67	Yes	0.02	220364652	220386647
2p25.3	10.0	0	No	0.06	17019	78843

CNV copy number variation, *DGV* Database of Genomic Variants, *NASH* non-alcoholic steatohepatitis

Table 4.54 Top 5% copy number loss only

CYTOBAND	NASH Loss (%)	CNV COVERAGE	CNV in DGV	Size (Mb)	Start	End
3p26.1	7.5	0	No	0.04	4240523	4276049
9p24.2	7.5	0	No	0.15	2901047	3047813

CNV copy number variation, *DGV* Database of Genomic Variants, *NASH* non-alcoholic steatohepatitis

CHAPTER FIVE
DISCUSSION

5.1 Single nucleotide polymorphism

5.1.1 Single nucleotide polymorphism: *PNPLA3*

In the present study, the association of *PNPLA3* rs738409 G allele with susceptibility to NAFLD as has been shown in a genome-wide association study was confirmed (Romeo et al., 2008). More importantly, this study showed a positive association of the risk allele G with risk of NASH and with fibrosis stage but not with other histological parameters. Thus, the G allele is not only associated with occurence of NASH but also with presence of fibrosis.

The risk allele G was found to be significantly associated with susceptibility to NAFLD after stratification into the three major ethnic subgroups, the Chinese, Indian and Malay. The Malaysian Indians migrated to Malaysia from Southern India in the 1800s (Periasamy, 2007). Studies have suggested that Indians from India were proto-Asian origin with West Eurasian admixture, hence giving them the genetic affinity towards both Asian and European (Bamshad et al., 2001; Jorde & Wooding, 2004). Compared to the other ethnic groups, the Malaysian Indians appear to have a high prevalence of NAFLD as is so with the Europeans. The Malaysian Chinese on the other hand, are descendance from the Hans from Southern China (Hock, 2007) and it is noted that the results of the Chinese subgroup in this study are consistent with the finding in the Taiwanese Hans (Lin et al., 2011) as proven by positive associations in both populations. The results of this study are consistent with the results of the previous studies of different populations including Hispanics (Romeo et al., 2008), Argentinian (Sookoian et al., 2009), Germans (Kantartzis et al., 2009), Italian (Romeo et al., 2010a; Romeo et al., 2010b; Valenti et al., 2010), Japanese (Hotta et al., 2010), Taiwanese (Lin

et al., 2011), and African Americans (Cox et al., 2011). The present study thus confirms earlier findings that the rs738409 plays a significant role in susceptibility to NAFLD.

The finding of an association between the *PNPLA3* variant with NASH in this population was also successfully replicated as was observed in several studies (Hotta et al., 2010; Rotman, Koh, Zmuda, Kleiner, & Liang, 2010; Sookoian et al., 2009). There are more patients with NASH in this study population compared to those with simple steatosis, which is probably reflective of the patients who are seen at the UMMC, which is a tertiary referral centre.

A novel finding in this study is that the G allele is associated with the severity of NASH. Most of the NAFLD biopsy-proven studies to date only looked at two spectrum of NAFLD: simple steatosis and NASH (Hotta et al., 2010; Rotman et al., 2010; Sookoian et al., 2009). Prediciting steatohepatitis is more important as NASH can potentially progress and treatment should be considered at an early stage of NASH. Compared to those with simple steatosis, the G allele is significantly higher in patients with NASH with significant fibrosis but not in NASH patients without significant fibrosis. Furthermore, the G allele frequency in NASH with significant fibrosis is higher than that of NASH without significant fibrosis. It is also revealed that the G allele is associated with a higher fibrosis score but not with the other histological features of NAFLD. The association of rs738409 G allele with higher fibrosis score was also observed in other studies (Hotta et al., 2010; Rotman et al., 2010; Valenti et al., 2010). One Asian study found that the G allele is associated with fibrosis score but not with steatosis (Hotta et al., 2010). The results of the present study were in concordance with that of Hotta et al, and in addition, there was no association with lobular inflammation and hepatocellular ballooning.

There have been conflicting reports of the correlation between presence of G allele and serum transaminase levels (Hotta et al., 2010; Kotronen et al., 2009; Lin et al., 2011; Romeo et al., 2008; Romeo et al., 2010a; Romeo et al., 2010b; Sookoian et al., 2009; Valenti et al., 2010). The rs738409 G allele was shown to be associated with increased levels of both serum AST and ALT in the Argentinian (Sookoian et al., 2009), Italian (Romeo et al., 2010a; Romeo et al., 2010b), Japanese (Hotta et al., 2010), and the Taiwanese (Lin et al., 2011). In the Hispanics (Romeo et al., 2008) and in an Italian population (Valenti et al., 2010), the G allele was found to be related with increased serum ALT only. The G allele was found to be associated with increased serum AST level only in the Finnish population (Kotronen et al., 2009). However the serum level of AST/ALT was found not to be significantly associated with the G allele in the African Americans (Cox et al., 2011; Romeo et al., 2008), European Americans (Romeo et al., 2008), Germans (Kantartzis et al., 2009), and in the present study (Zain et al., 2012a). The differences in these findings are probably due to the differences in the recruited subjects in the various studies, ranging from adults with NAFLD (Hotta et al., 2010; Kotronen et al., 2009; Romeo et al., 2008; Sookoian et al., 2009; Valenti et al., 2010), either obese or diabetic adults (Kantartzis et al., 2009), obese adults (Romeo et al., 2010b), and obese children (Lin et al., 2011; Romeo et al., 2010a).

Two studies reported a decreased level of plasma triglycerides in the NAFLD subjects harboring the G allele (Hotta et al., 2010; Speliotes, Butler, Palmer, Voight, & Hirschhorn, 2010). This finding is in agreement with the present study, in which decreased level of plasma triglycerides was found amongst the NAFLD patients with GG genotype despite a significantly higher level of triglyceride in the NAFLD patients as compared to controls, suggesting that genetic susceptibility of G allele plays a much more important role than triglyceride levels in the pathogenesis of NAFLD. In the present study, the triglyceride levels were further analyzed in the NAFLD spectrum.

The triglyceride levels are higher in NASH without significant fibrosis compared to NASH with significant fibrosis, suggesting that the triglyceride level subsides along with the degree of hepatocellular injury (Jou, Choi, & Diehl, 2008).

5.1.2 Single nucleotide polymorphism: *AGTR1*

Angiotensin II appears to have an important role in the progression to non-alcoholic steatohepatitis (NASH) (Oakley et al., 2009; Yoshiji et al., 2001). Angiotensin II acts on angiotensin II type 1 receptor to activate hepatic stellate cells (HSCs). The activated HSCs will express the transforming growth factor-β 1 (TGF-β 1), a profibrogenic cytokine, as a counter response to liver injury (X. Li, Meng, Wu, Zhang, & Yang, 2007; Yoshiji et al., 2001). The preventive role of *AGTR1* blocker in the progression to NASH was succesfully observed in the animal and human studies (Fujita et al., 2007; Hirose et al., 2007; Moreno et al., 2010; Yokohama et al., 2004). The deletion of angiotensin II type 1 receptor in animal studies has been shown to reduce hepatic steatosis suggesting *AGTR1* is an important regulator of hepatic steatosis (Nabeshima, Tazuma, Kanno, Hyogo, & Chayama, 2009).

In the present study, data indicated that the five previously reported SNPs of the angiotensin II type I receptor (*AGTR1*) gene (rs3772622, rs3772627, rs3772630, rs3772633, and rs2276736) were not associated with susceptibility to non-alcoholic fatty liver disease (NAFLD) and non-alcoholic steatohepatitis (NASH). This is in contrast with the Japanese study which showed that these *AGTR1* polymorphisms were significantly associated with the occurence of NAFLD and NASH (Yoneda et al., 2009b).

Among the SNPs studied, rs2276736, rs3772630 and rs3772627 were found to be protective against NAFLD and NASH in the Indian ethnic subgroup. Interestingly,

these three SNPs share the same linkage disequilibrium structure. It was revealed that haplotype ACGCA is protective against NAFLD in the Indians. The study by Yoneda et al (2009b) reported that the protective haplotype in the Japanese was ATATG. The discrepancy in the findings may be explained by ethnic differences between the populations studied. The three major ethnic subgroups in Malaysia, namely the Malay, Chinese and Indian are of the south Asian origin with the Chinese are of the Hans descendant of Southern China (Hock, 2007), Indians from Southern India (Periasamy, 2007), whereas the Malays are the indigenous ethnic group. There is a supportive evidence that the Japanese are from the north Asian (Northern China and Korea) origin but not the south Asian (Southern China and Southeast Asia) based on the classical marker polymorphisms, Y chormosome and mitochondrial DNA (Hammer et al., 2006; Omoto & Saitou, 1997). The difference in the association studies between the two population probably reflects the difference in the genetic pools of the Malaysian and the Japanese.

The results from this study indicated that there was no difference in levels of liver enzymes between patients with simple steatosis and those with NASH. This is in agreement with the results from Sorrentino et al (Sorrentino et al., 2004) and Mofrad et al (Mofrad et al., 2003), but is in contrast with the findings of several other studies which showed differences in levels between the two groups (Abrams, Kunde, Lazenby, & Clements, 2004; Dixon, Bhathal, & O'Brien, 2001; Lima, Mourao, Diniz, & Leite, 2005; Ong et al., 2005). The differences in these findings is probably due to the differences in the recruitment of subjects for the various studies, in which inclusion criteria ranges from morbid obese NAFLD to adults with NAFLD. Furthermore, a significantly higher triglycerides levels were observed in patients with NASH compared to those with simple steatosis. Triglyceride storage in hepatocytes marks the intensity of

exposure to potentially toxic fatty acids. As NAFLD progresses, excess free fatty acids are compensated with increase in triglycerides production (Jou et al., 2008).

Analysis of interaction between *PNPLA3* and *AGTR1* gene, revealed a strong interaction between *AGTR1* (rs3772627), *AGTRI* (rs3772630) and *PNPLA3* (rs738409) SNPs on NAFLD susceptibility. Compared to the Malays and Chinese, the Indians presented with the greatest susceptibilty to NAFLD in the *PNPLA3* study (Zain et al., 2012a). Thus the *PNPLA3* can possibly mask the effect of *AGTR1* on NAFLD among the Indians. Although both genes encode different products (*PNPLA3* encodes adiponutrin while *AGTR1* encodes angiotensin II type 1 receptor), the strong interaction observed suggested a possible interaction as both genes are involved in lipid accumulation. The action of angiotensin II on angiotensin II type 1 receptor (*AGTR1*) promotes lipid accumulation in the liver while lack of *AGTR1* reduces hepatic triglycerides (Nabeshima et al., 2009). *PNPLA3* on the other hand is highly upregulated in response to feeding and downregulated in fasting state, displaying its role in lipid storage in the liver (Hoekstra et al., 2010). Of note, *AGTR1* and *PNPLA3* are correlatively upregulated in a state of excess lipid.

This study further showed that the SNP rs3772622 was associated with occurence of fibrosis but not with the other histological parameters of NAFLD. The G allele of this SNP is associated with increased fibrosis score. The Japanese also found that SNP rs3772622 was associated with the presence of fibrosis (Yoneda et al., 2009b), but in contrast to our findings, the G allele was found to be associated with decreased fibrosis score. The difference is probably explained by a higher minor allele frequency in this study population than that of the Japanese. A greater systolic blood pressure mean value (136.00 mmHg) observed in the GG genotype compared to the AA genotype (128.65 mmHg) of the SNP rs3772633 could also contribute to the discrepancy. Data of the previous report showed that hypertensive patients have

significantly higher fibrosis score than that of non-hypertensive (Corey et al., 2009). The strong linkage disequilibrium observed between the five SNPs in the pooled subjects and even after ethnic stratification indicates a potential tagSNP selection based on the pairwise LD. This approach will be useful in future in order to minimize the genotyping cost yet presenting the information that could reflect the role of the other relevant SNPs within the gene.

5.1.3 Single nucleotide polymorphism: *LEPR*

This study showed significant association between both the *LEPR* rs1137100 and rs1137101 with susceptibility to NAFLD. Results suggest that in patients with NAFLD, the G allele of rs1137100 is associated with a less severe form of liver disease. To the best of knowledge, the current study is the first to report an interaction between *LEPR* and *PNPLA3* genes with increased risk of NAFLD compared to either gene alone.

The data to date, on the relationship between the *LEPR* gene variants, rs1137100 and rs1137101, and obesity is conflicting. No significant association was found between rs1137100 with obesity (Furusawa et al., 2010; Yiannakouris et al., 2001). Although several studies have reported an association between rs1137101 with obesity (Mattevi, Zembrzuski, & Hutz, 2002; Yiannakouris et al., 2001), an equal number of negative studies have also been reported (Constantin et al., 2010; Pyrzak, Wisniewska, Kucharska, Wasik, & Demkow, 2009). A recent meta-analysis of the *LEPR* tagSNPs did not show an overall association with obesity (Bender et al., 2011). The discrepancy of these findings may be explained by the differences in allele frequency across ethnic groups. Ethnicity could act as a significant effect modifier thereby explaining the conflicting observations (Marchini, Cardon, Phillips, & Donnelly, 2004).

In the present study, polymorphisms on the *LEPR*, rs1137100 and rs1137101 were shown to be associated with increased risk of NAFLD and NASH. The rs1137100 was also shown to be associated with simple steatosis and NASH without significant fibrosis. The G allele frequency of rs1137100 in the control subjects was 0.78 for the Chinese and 0.19 for the Indian, which was almost the same as that reported in the International HapMap database (hapmap.ncbi.nlm.nih.gov/) but higher than that in other populations (such as European and African American).

An interesting finding from this study is the association of the G allele with NASH without significant fibrosis but not with NASH with significant fibrosis, suggesting that the polymorphism of *LEPR* rs1137100 is associated with lower fibrosis score and patients with G allele are protected against higher stage of fibrosis ≥ 2. These findings suggest that rs1137100, specifically the G allele, is associated with a milder form of liver disease in patients with NAFLD. To date, there has been no published report of an association between *LEPR* rs1137100 G allele and liver fibrosis. The present study is the first to provide this association finding and the availability of histological profile of the liver samples made this possible, which is a major strength of this study. Hepatic fibrosis score was found to be higher in subjects with metabolic syndrome (Ryan et al., 2005). In a Taiwanese Han population, lower metabolic risk was reported in those bearing the GG genotype (C. W. Yang et al., 2013). A Japanese study in obese children showed a significantly lower total and LDL cholesterol in subjects with GG genotype (Okada, Ohzeki, Nakagawa, Sugihara, & Arisaka, 2010). Leptin level has been shown to correlate with metabolic syndrome profile such as adipocity and insulin resistance (Silha et al., 2003; Stylianou et al., 2007). It was reported that leptin level was significantly greater in subjects with AA genotype than G allele carriers (Wauters et al., 2001). Hence, these studies suggest that *LEPR* rs1137100 G allele is associated with a lower metabolic risk.

This study indicated nearly three-fold increased risk of NAFLD by the combined effect of the SNPs of *LEPR* with *PNPLA3* compared to either gene alone. A common variant in the *PNPLA3* gene (rs738409) was found to be associated with hepatic fat content and confers risk of NAFLD in the Dallas Heart Study of different ancestries; Hispanics, European-Americans and African Americans (Romeo et al., 2008). Subsequently, a number of replicated studies on *PNPLA3* in various populations have been carried out and interestingly, all yielded a positive association including that of this study population (Cox et al., 2011; Hotta et al., 2010; Kantartzis et al., 2009; Lin et al., 2011; Romeo et al., 2010a; Romeo et al., 2010b; Sookoian et al., 2009; Valenti et al., 2010; Zain et al., 2012a). The effect size of NAFLD is increased with a single variant of *LEPR* or *PNPLA3*. The risk of NAFLD is however increased dramatically with the co-existence of both variants of *LEPR* and *PNPLA3*, suggesting a possible positive interaction between the two genes. Investigation of gene interaction between *LEPR* and *PNPLA3* is therefore a good approach that allows us to understand the genetic interaction of the two genes that encode for products that are important in lipid regulation, and within same study population.

Analysis of interaction between *LEPR* and *PNPLA3* gene, revealed a strong interaction between *LEPR* (rs1137100) and *PNPLA3* (rs738409) SNPs on NAFLD susceptibility. The pleiotropic actions of leptin via leptin receptor help suppress food intake, a mechanism involved in appetite control and body weight regulation (Zastrow et al., 2003). *PNPLA3* exerts its effect in response to food intake, in which it is highly upregulated in response to feeding and downregulated in fasting state, displaying its role in lipid storage in the liver (Hoekstra et al., 2010). Both genes are correlatively upregulated in a state of excess lipid.

5.1.4 Single nucleotide polymorphism: *GCKR*

The present study identified a common intronic SNP rs780094 and a loss of function SNP rs1260326 (C/T change results in a proline to leucine substitution in the amino sequence of the encoded protein) on *GCKR* that occur more frequently in patients with NAFLD compared to controls. These findings are thus in the same direction of that reported by a Genetics of Obesity-related Liver Disease (GOLD) consortium comprising of a meta-analysis of four genome-wide association studies (GWAS) (Speliotes et al., 2011). This study also report an association between rs1260326 T allele with increased grade of hepatic steatosis.

Investigation of NAFLD-associated SNPs was performed in a Malaysian heterogeneous population comprised of three major ethnic groups, namely the Malay, Chinese and Indian. The pooled result indicates increased odds of NAFLD by 1.49 for rs1260326 and 1.51 for rs780094. This finding was supported by a GWAS consortium that exhibited 1.45 risk of NAFLD for *GCKR* rs780094. It has to be noted that both SNPs rs1260326 and rs780094 were in strong linkage disequilibrium (Speliotes et al., 2011). Genetic origin of the Malaysian Chinese and Indians are evidenced from the immigration of these people into Malaysia from Southern China (Hock, 2007) and Southern India (Periasamy, 2007) in the late 18^{th} century and early 19^{th} century. However, only the Indian showed significant association with NAFLD. Despite relatively low sample size in the Indian subjects, this study was able to show the power ranges from 77% to 94%. This association can also be explained by the genetic admixture of this subpopulation with West Eurasian resulting in genetic affinity towards both Asian and European (Bamshad et al., 2001; Jorde & Wooding, 2004). The prevalence of NAFLD in European ancestral was reported higher than that of the Asian ancestral (Bellentani & Marino, 2009) thus supporting this finding. It is noteworthy that the minor allele frequency in the Indian controls is similar to that reported in the

International HapMap database (hapmap.ncbi.nlm.nih.gov/) for Gujarati Indians (19%). The pooled result is also in accordance to a report by Z. Yang et al. (2011), but failed to be replicated in the Chinese subjects. The difference in this finding could be due to difference in population structure. It was proposed that Shanghai subpopulation (Chinese subjects from Z. Yang et al., 2011 study) shares similar population structure with that of the Han Chinese from Beijing (CHB) reported in International HapMap database, however differs from Southeast Asia (Malaysian Chinese subjects) (J. Chen et al., 2009). This explains the difference in minor allele frequency in the control subjects between the Malaysian Chinese (44%) and Chinese subjects from Z. Yang et al. (2011) study (53%) and CHB (57%). Studies at a metapopulation level such as Chinese population therefore invite a better understanding of population structure that is an important part of evolutionary genetics. Of note, in the study by Z. Yang et al. (2011) the diagnosis of NAFLD was based on ultrasound but not definitively histological-proven by liver biopsy. In a cohort of nine years, from 2000 to 2009, 1621 subjects who met the criteria participated in a consecutive health screening program in a day-care medical center−the Megah Medical Centre, Petaling Jaya, adjacent to the capital city of Kuala Lumpur, Malaysia (Goh, Ho, & Goh, 2012). The proportion of subjects in this study was 12.1% Malays, 78.3% Chinese and 9.6% Indians with a reported prevalence of NAFLD of 32.5%, 19.9% and 33.5%, respectively. This study indicates that the Indians, although being the minority has the highest prevalence of NAFLD, while Chinese the lowest. The data in this larger cohort therefore reflects the findings in the present study with regards to ethnic differences.

The present study provide the first data on association between *GCKR* SNPs with NASH. The availability of histological profile of each patient contributes to the strength of this study. The *GCKR* rs1260326 and rs780094 T allele were found to be associated with susceptibility to NASH. Subsequently, association between the risk

allele with NASH with significant fibrosis was observed. Although it was first demonstrated to be associated with histologic NAFLD that could include NAFLD with inflammation and fibrosis in the recent GWAS (Speliotes et al., 2011), the authors do not identify how many patients with NAFLD are simple steatosis and how many are NASH. Prediction of steatohepatitis (which in most cases encompasses advanced fibrosis) in NAFLD is more important as it is potentially progressive (Wilson & Chalasani, 2007). Another interesting finding from this study is that *GCKR* rs1260326 T allele was significantly associated with higher grade steatosis. This finding was supported by Santoro et al (Santoro et al., 2012) in which *GCKR* rs1260326 was found to be associated with hepatic fat content in individuals of Caucasian, African American and Hispanic ancestry. The paradox of this finding may be because of the role of *GCKR* in insulin resistance (Shen et al., 2013), which is important in the "first hit" mechanism towards development of steatosis and triglycerides accumulation in the liver (Bugianesi et al., 2005).

Concerning the importance of *GCKR* and *PNPLA3* genes in the development of hepatic fat content (Romeo et al., 2008; Santoro et al., 2012), the present study reports an interaction between these two genes on susceptibility to NAFLD. The effect size contributed by the combined genes would result in a greater risk of NAFLD than either gene alone. Intriguingly, genes that share similar pathway or outcome as *PNPLA3* have been shown to interact between each other (Zain et al., 2012b; Zain et al., 2013). These data therefore indicate the complex nature of NAFLD involving multiple gene interactions. More studies need to clarify the genetic underpinnings of NAFLD.

5.1.5 Single nucleotide polymorphism: Other candidate genes

The present study also investigated a number of SNPs in various candidate genes that includes *iNOS*, *TRAIL*, *SREBF1*, *CLOCK*, *COL13A1*, *FABP2*, *UCP3*, *PPARG*, *COX-2*, *NR1I2*, *KLF6*, *TFRC*, *PPARGC1A*, *FATP5*, *STAT3*, *APOA5*, and *PEMT* and *FDFT1*. None of the SNPs in these genes were found to be associated with susceptibility to NAFLD in this study population except for *SREBF1* rs2297508, which was found to be a risk for the Indian ethnic group. These genes and their related SNPs were chosen for analysis because during the time this study was conducted, they were among the recent SNPs and tagSNPs found to be associated with NAFLD in other populations. This study therefore intended to use these SNPs as a preliminary screening for NAFLD in Malaysian population.

5.2 Copy number variation

Genome wide CNV profiling was carried out using high-resolution aCGH to determine the genomic aberrations in a total of 40 NASH samples that were normalised against age-, gender- and ethnicity-matched controls. The response of the liver to stress during the second hit process and the progressive nature of NASH, contribute to the selectively chosen NASH patients in this CNV study. Apart from evaluation of biochemical parameters and ultrasonogaphy, magnetic resonance imaging (MRI) was carried out on all 40 controls for the genome wide profiling in order to rule out fatty liver. Given that the present study is first to investigate the genome wide profile of copy number in NAFLD, the total number, genomic positions, gene content, frequency spectrum and patterns of linkage disequilibrium (LD) of the genomic CNAs are challenging. The use of ultra-dense aCGH technology allows CNAs discovery at high resolution down to approximately 3 kb and hence provides great confidence in detection of CNVs. The use of 60mer probes provides high sensitivity and specificity to

accurately detect both known and *de novo* CNVs as compared to shorter oligonucleotide probes (Curtis et al., 2009).

In this study, the most amplified 14q11.2 region is enriched with both T cell receptor family and olfactory receptor (*OR*) family genes, whereas an abundance of *OR* family genes were found in the most frequently deleted 11q11 region. Surprisingly, a search from database of Expressed Sequence Tags, NCBI dbEST, revelead many hits of *OR* gene expression not only in the olfactory system, but also in other tissues such as liver, lung, kidney, prostate, testis, brain and many others including tumor tissues. The families of these receptors play an important role during embryogenesis, and are expressed in various parts of the embryo suggesting that they are not only involved in the olfactory system but also throughout many other organs as these organs are assembled (Dreyer, 1998). In the early 2000s, a complete database on human olfactory subgenome was completed after comprehensive data mining using highly automated data mining system (Glusman, Yanai, Rubin, & Lancet, 2001). Glusman et al. (2001) reported the presence of 906 potential coding regions for *OR* genes that cover almost all human chromosomes with the exception of chromosome 20 and Y, in which 2/3 of them have not been reported. Subsequently, a new set of database called the Olfactory Receptor Microarray Database (ORMD) that houses microarray gene expression data of the ORs was developed (N. Liu, Crasto, & Ma, 2007). These achievements clearly stated the importance of *OR*s not only in the olfactory sytem, but throughout organs that remain elusive. T cell receptors on the other hand, are essential to fight against cancer (Guitart et al., 2012) and viral infection (Pasetto et al., 2012). Region 14q11.2 also contained matrix metallopeptidase 14 gene (*MMP14*) in which expression was increased to 2-3 folds in fibrotic liver in comparison to normal liver (X. Zhou et al., 2004). The expression was also found higher in cancerous portion compared to non-concerous portion of hepatocellular carcinoma tissues (Harada et al., 1998).

Carboxylesterase 1 gene (*CES1*) which is important in the transportation of cholesterol was located in the chromosome 16q12.2. This region shared relatively similar high frequency of gain and loss events (37.5% gain, 32.5% loss) suggesting simultaneous involvement of multiple loci from this chromosome. In human study, expression of *CES1* was higher in NAFLD-proven liver tissue compared to non-NAFLD (Ashla et al., 2010). Positive correlation of *CES1* was observed with triglyceride lipase activities and adiposity suggesting the lipolysis role of *CES1* thereby contributing to hepatic steatosis (Nagashima et al., 2011). In animal study, *CES1* knockout mice exhibited gain in weight, hepatic steatosis and hyperinsulinemia, thus supporting the role of *CES1* in the regulation of fatty acids (Quiroga et al., 2012). As can be predicted therefore, feeding with fatty acids and cholesterol-rich diet upregulates *CES1* (Dolinsky et al., 2003). A recent mapping approach of 16q12.2 locus has found variants associated with BMI (Peters et al., 2013). Similar trend of gain and loss events was observed in the p-arm of chromosome 16p11.2 (20% gain, 20% loss). Among genes of interest that could be found was zinc finger protein 267 (*ZNF267*). Zinc finger proteins are among the most abundant proteins in the human genome with a diverse function including DNA recognition, transcriptional activation, protein folding and assembly, and lipid binding (Laity, Lee, & Wright, 2001). The expression of *ZNF267* was highly upregulated in HCC compared to non-cancerous tissue. Furthermore, the expression was found to be greater when induced by reactive oxygen species (Schnabl, Valletta, Kirovski, & Hellerbrand, 2011).

Aberration in the region 14q32.33 (25% gain, 17.5 loss) was not found in the Database of Genomic Variants (DGV). This chromosomal region could serve as a potential copy number marker for NAFLD. This region is rich in immunoglobulin heavy chain gene (*IGH*) which is overexpressed in patients with NASH (Puri et al., 2008). More importantly, v-akt murine thymoma viral oncogene homolog 1 (*AKT1*) and

metastasis associated 1 (*MTA1*) genes were concomitantly located in the same region. Deletion of *AKT1* in animal model was shown to regulate hepatic metabolism by insulin (Lu et al., 2012), which is a contributing factor to NAFLD development. In a study consisting of 506 HCC patients who underwent hepatic resection, 17% of them were positive for *MTA1*. Those patients who were MTA-positive HCCs had poor survival rate (Ryu et al., 2008). It is also noteworthy to concentrate on the clear-cut gain on chromosome 16p12.2. At least 1/4 of the samples were amplified in this region, suggesting the importance of extended investigation such as chromosomal region mapping.

Downstream studies to validate findings are to be pursued. Findings observed in genome wide analysis are basically in data output. Validation using quantitative Real-Time PCR is necessary to provide the "real-time" picture of the events. But that is not the main concern here. Commonly, quantitative validation will ensure similar findings. What could be the next intriguing step to be taken is the downstream approach of copy number analysis. Current data are CNV-based, in which the genes of interest were "fished out" based on the high frequency of chromosomal aberrations found and is known to create bias. It will be a lot more interesting to investigate on the basis of gene. It is important to verify the CNV regions and their impact on NAFLD progression. In the context of exploring the biological link between these CNVs and NAFLD progression, functional annotation analysis using David Pathway (david.abcc.ncifcrf.gov/) should be performed between genes residing at the loci and NAFLD known genes, especially *PNPLA3*, in the search of similarity in terms of cellular components, biological processes and molecular functions. Principal component analysis that further emphasizes on annotations such as gender is another approach to be taken.

5.3 Limitations and strenghts of study

5.3.1 Limitations

There are several limitations to the present study. Most studies of NAFLD have been case-control association studies that only examined the association of NAFLD per se but not with the spectrum of NAFLD that comprises of simple steatosis, NASH without significant fibrosis and NASH with significant fibrosis. Studies on NAFLD spectrum have been limited since this requires valid ascertainment and scoring of the different NAFLD stages, for which liver biopsy is the gold standard in providing such accurate diagnosis. Such biopsies are limited to situations in which there are definite clinical indications for biopsy, and this reflects the relatively small sample size in this study. Unlike the diagnosis of NASH, diagnosis of fatty liver (steatosis) can be made based on clinical, radiological and laboratory findings, which therefore makes it slightly easier to carry out studies on NAFLD. Due to ethical considerations, no liver biopsies were performed for the controls leading to a possibility of misclassification of controls. However this likelihood was reduced by using the stated selection criteria for the controls. These criteria includes normality in BMI, fasting plasma glucose, lipid profile and liver enzymes. The difference in the distribution of these parameters between the two extremes of the NAFLD disease spectrum, provides confidence in the selection process of NAFLD cases and controls. Mild steatosis among the controls however cannot be ruled out and this may result in a bias towards the null. Some of the resulting associations between SNPs and NAFLD spectrum were smaller than hypothesized (OR less than 2.0) and would have resulted in a loss of power for the relevant analysis. The small sample size after stratification of the subjects into three ethnic groups, and NAFLD patients into simple steatosis and NASH contribute to this limitation. This yields an unmatched case-control study. However, the potential for confounding can be reduced by adjustment during the analysis. More importantly, the power after study was

provided and most of the analyses achieved the desired power of > 80%. The causal pathway between rs738409 and NASH has been postulated to be mediated through triglycerides. As this study conditions on BMI, the case control association of rs738409 and NASH is vulnerable to confounding bias betweeen the pathway of triglycerides and NASH. Thus residual confounding from collider stratification bias due to conditioning of a downstream covariate might refute some of the association between rs738409 and NASH. However the associations between rs738409 and the different grades of NAFLD among the cases are not likely to be confounded and provide valid estimates. The limited knowledge on CNVs in NAFLD is another concern. Unlike SNPs where most of the frequencies have been determined in various populations, frequencies of CNVs are not well-defined. Considering the Malay subjects used in this CNVs study, CNAs identified in this study may be different from the future report from Western studies, or even another Asian studies.

5.3.2 Strengths

A major strength of this study was the ability to compare the genetic association among three asian ethnic groups in Malaysia, which are also the main ethnic groups in Asia (Crossette, 2011). This study suggest that ethnicity modifies the relationship between genetic polymorphisms and NAFLD. It appears that the role of genetic polymorphisms on NAFLD is amplified among Indians compared to Chinese and Malays. The present study investigated the NAFLD spectrum that ranges from simple steatosis to NASH without significant fibrosis to NASH with significant fibrosis. This study is also the first to report the genetic interaction between the *AGTR1*, *LEPR*, and *GCKR*, and the much reported *PNPLA3* gene on NAFLD susceptibility. Reports on the copy number findings are novel and could serve as potential genetic marker for the identification of the potentially progressive form of NAFLD. It is also learnt that this is the first study on

NAFLD in Malaysia. Data from the present study could assist researchers and practitioners in downstream studies not limited to Malaysian population, but also to the other populations.

CHAPTER SIX
CONCLUSION AND FUTURE STUDIES

6.1 Conclusion

Genetic association study of susceptibility to NAFLD indicates a positive association between the candidate genes *PNPLA3*, *LEPR* and *GCKR* with susceptibility to NAFLD in a Malaysian population comprising three major ethnic groups, namely, the Malays, Chinese and Indians. *AGTR1* polymorphisms on the other hand, render a protection against NAFLD in the Indians. Contrary to reports in other populatuions, various other gene polymorphisms were not found to be associated with NAFLD susceptibility, suggesting variability in genetics across populations. This study also provides new finding on the association with NASH severity. While most of the studies have attempted to predict the risk or susceptibility to NAFLD, it is clinically more important to identify the predictor for the risk of NASH as it is potentially progressive. The G allele of the widely studied *PNPLA3* rs738409 is associated with severity of NASH. The Malaysian Indians are genetically unique such that results differ between the Malays and Chinese. NAFLD is genetically a complex disease as delineated by multiple gene interactions, that may indirectly share similar pathway in cellular components, biological processes and molecular functions. There is a paucity in information with regard to CNV in NAFLD, and hence, current data provides some interestingly new information which points towards the complex nature of CNV in NAFLD. A larger sample size and more ethnically diverse samples are needed to replicate these findings.

6.2 Future studies

A broad aspect of NAFLD covering not only genome but also proteome, remains a continuous plan for future studies. SNPs and CNVs are parts of genome that provide meaningful however static information. A functional study that offers more dynamic information such as proteomic study is one of the great aspects to look into. Although several studies on proteomics of NAFLD have been conducted, the novel approach can still be extracted out. The uniqueness of Biotin Carboxyl Carrier Protein (BCCP) tag technology ensures that only correctly folded, functional and full-length proteins are presented on the array. The tag also provides a single point of attachment which negates the problems of protein unfolding, random orientation and non-specific binding that can occur with other protein attachment methods. Another scope that can be explored is profiling of microRNA (miRNA). Reported miRNA in NAFLD are mostly coming from studies involving liver tissues. This however does not eliminate the use of biopsy, which is the main objective in most of the NAFLD studies−to predict NASH in the absence of liver biopsy. That is when a new technology called Locked Nucleic Acid (LNA) comes in handy. LNA technology can be used for a wide range of samples including biofluids such as serum. The first challenge of microRNAs studies is that miRNAs are made up of short sequences of approximately 22 nueclotides and hence, conventional DNA-based methods that have been commonly used are not sensitive enough to detect these sequences with any reliability. Secondly, closely-related miRNA family members differ by as short as one nucleotide, emphasizing the need for high specificity and sensitivity to discriminate between single nucleotide mismatches. LNA technology provides increased affinity and specificity of the probes, inhibitors and primers for their miRNA targets, thereby addressing both challenges described above.

REFERENCES

A haplotype map of the human genome. (2005). *Nature, 437*(7063), 1299-1320.

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*(2), 475-483.

Adams, L. A., Waters, O. R., Knuiman, M. W., Elliott, R. R., & Olynyk, J. K. (2009). Nafld as a risk factor for the development of diabetes and the metabolic syndrome: An eleven-year follow-up study. *Am J Gastroenterol, 104*(4), 861-867.

Adams, LA., Angulo, & Lindor, KD. (2005). Nonalcoholic fatty liver disease. *CMAJ, 172*(7), 899-905.

Agopian, V. G., Kaldas, F. M., Hong, J. C., Whittaker, M., Holt, C., Rana, A.,...Busuttil, R. W. (2012). Liver transplantation for nonalcoholic steatohepatitis: The new epidemic. *Ann Surg, 256*(4), 624-633.

Aller, R., De Luis, D. A., Izaola, O., Gonzalez Sagrado, M., Conde, R., Alvarez, T.,...Velasco, M. C. (2010). Role of -55CT polymorphism of UCP3 gene on non alcoholic fatty liver disease and insulin resistance in patients with obesity. *Nutr Hosp, 25*(4), 572-576.

Almeda-Valdes, P., Cuevas-Ramos, D., & Aguilar-Salinas, C. A. (2009). Metabolic syndrome and non-alcoholic fatty liver disease. *Ann Hepatol, 8 Suppl 1*, S18-24.

Amarapurkar, D. N., Hashimoto, E., Lesmana, L. A., Sollano, J. D., Chen, P. J., & Goh, K. L. (2007). How common is non-alcoholic fatty liver disease in the asia-pacific region and are there local differences? *J Gastroenterol Hepatol, 22*(6), 788-793.

Angulo, P. (2002). Nonalcoholic fatty liver disease. *N Engl J Med, 346*(16), 1221-1231.

Ashla, A. A., Hoshikawa, Y., Tsuchiya, H., Hashiguchi, K., Enjoji, M., Nakamuta, M.,...Shiota, G. (2010). Genetic analysis of expression profile involved in retinoid metabolism in non-alcoholic fatty liver disease. *Hepatol Res, 40*(6), 594-604.

Assy, N., Kaita, K., Mymin, D., Levy, C., Rosser, B., & Minuk, G. (2000). Fatty infiltration of liver in hyperlipidemic patients. *Dig Dis Sci, 45*(10), 1929-1934.

Auinger, A., Valenti, L., Pfeuffer, M., Helwig, U., Herrmann, J., Fracanzani, A. L.,...Rubin, D. (2010). A promoter polymorphism in the liver-specific fatty acid transport protein 5 is associated with features of the metabolic syndrome and steatosis. *Horm Metab Res, 42*(12), 854-859.

Bae, J. S., Cheong, H. S., Kim, J. H., Park, B. L., Park, T. J., Kim, J. Y.,...Shin, H. D. (2011). The genetic effect of copy number variations on the risk of type 2 diabetes in a korean population. *PLoS One, 6*(4), e19091.

Baffy, G., Brunt, E. M., & Caldwell, S. H. (2012). Hepatocellular carcinoma in non-alcoholic fatty liver disease: An emerging menace. *J Hepatol, 56*(6), 1384-1391.

Bamshad, M., Kivisild, T., Watkins, W. S., Dixon, M. E., Ricker, C. E., Rao, B. B.,...Jorde, L. B. (2001). Genetic evidence on the origins of indian caste populations. *Genome Res, 11*(6), 994-1004.

Bellentani, S., & Marino, M. (2009). Epidemiology and natural history of non-alcoholic fatty liver disease (nafld). *Ann Hepatol, 8 Suppl 1*, S4-8.

Bender, N., Allemann, N., Marek, D., Vollenweider, P., Waeber, G., Mooser, V.,...Bochud, M. (2011). Association between variants of the leptin receptor gene (LEPR) and overweight: A systematic review and an analysis of the colaus study. *PLoS One, 6*(10), e26157.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate - a practical and powerful approach to multiple testing. *J R Stat Soc Series B Stat Methodol, 57*(1), 289-300.

Bertrand, J. B., Lefevre, H., Prevot, S., & Perlemuter, G. (2011). [nonalcoholic fatty liver disease in a severely obese adolescent. An arguable liver biopsy]. *Arch Pediatr, 18*(1), 62-65.

Bhat, G., Baba, C. S., Pandey, A., Kumari, N., & Choudhuri, G. (2012). Life style modification improves insulin resistance and liver histology in patients with non-alcoholic fatty liver disease. *World J Hepatol, 4*(7), 209-217.

Bhatt, S. P., Nigam, P., Misra, A., Guleria, R., Luthra, K., Pandey, R. M.,...Pasha, M. A. (2013). Association of peroxisome proliferator activated receptor-gamma gene with non-alcoholic fatty liver disease in asian Indians residing in North India. *Gene, 512*(1), 143-147.

Boza, C., Riquelme, A., Ibanez, L., Duarte, I., Norero, E., Viviani, P.,...Arrese, M. (2005). Predictors of nonalcoholic steatohepatitis (nash) in obese patients undergoing gastric bypass. *Obes Surg, 15*(8), 1148-1153.

Brunt, E. M., Kleiner, D. E., Wilson, L. A., Belt, P., & Neuschwander-Tetri, B. A. (2011). Nonalcoholic fatty liver disease (nafld) activity score and the histopathologic diagnosis in nafld: Distinct clinicopathologic meanings. *Hepatology, 53*(3), 810-820.

Bugianesi, E., McCullough, A. J., & Marchesini, G. (2005). Insulin resistance: A metabolic pathway to chronic liver disease. *Hepatology, 42*(5), 987-1000.

Can Demirdogen, B., Sahin, E., Turkanoglu Ozcelik, A., Bek, S., Demirkaya, S., & Adali, O. (2012). Apolipoprotein A5 polymorphisms in turkish population: Association with serum lipid profile and risk of ischemic stroke. *Mol Biol Rep, 39*(12), 10459-10468.

Cao, M. B., Yang, Y. X., & Dong, L. (2010). [relationship between single nucleotide polymorphisms in the promoter of COX-2 gene and hereditariness to nafld]. *Zhonghua Gan Zang Bing Za Zhi, 18*(10), 773-777.

Cayon, A., Crespo, J., Guerra, A. R., & Pons-Romero, F. (2008). [gene expression in obese patients with non-alcoholic steatohepatitis]. *Rev Esp Enferm Dig, 100*(4), 212-218.

Chalasani, N., Guo, X., Loomba, R., Goodarzi, M. O., Haritunians, T., Kwon, S.,...Rotter, J. I. (2010). Genome-wide association study identifies variants associated with histologic features of nonalcoholic fatty liver disease. *Gastroenterology, 139*(5), 1567-1576, 1576 e1561-1566.

Chalasani, N., Younossi, Z., Lavine, J. E., Diehl, A. M., Brunt, E. M., Cusi, K.,...Sanyal, A. J. (2012). The diagnosis and management of non-alcoholic fatty liver disease: Practice guideline by the american gastroenterological association, american association for the study of liver diseases, and american college of gastroenterology. *Gastroenterology, 142*(7), 1592-1609.

Charlton, M. (2004). Nonalcoholic fatty liver disease: A review of current understanding and future impact. *Clin Gastroenterol Hepatol, 2*(12), 1048-1058.

Chatrath, H., Vuppalanchi, R., & Chalasani, N. (2012). Dyslipidemia in patients with nonalcoholic fatty liver disease. *Semin Liver Dis, 32*(1), 22-29.

Chen, G. B., Xu, Y., Xu, H. M., Li, M. D., Zhu, J., & Lou, X. Y. (2011). Practical and theoretical considerations in study design for detecting gene-gene interactions using mdr and gmdr approaches. *PLoS One, 6*(2), e16981.

Chen, J. M., Chuzhanova, N., Stenson, P. D., Ferec, C., & Cooper, D. N. (2005). Intrachromosomal serial replication slippage in trans gives rise to diverse genomic rearrangements involving inversions. *Hum Mutat, 26*(4), 362-373.

Chen, J., Zheng, H., Bei, J. X., Sun, L., Jia, W. H., Li, T.,...Liu, J. (2009). Genetic structure of the Han Chinese population revealed by genome-wide SNP variation. *Am J Hum Genet, 85*(6), 775-785.

Cheung, O., & Sanyal, A. J. (2010). Recent advances in nonalcoholic fatty liver disease. *Curr Opin Gastroenterol, 26*(3), 202-208.

Chon, C. W., Kim, B. S., Cho, Y. K., Sung, K. C., Bae, J. C., Kim, T. W.,...Joo, K. J. (2012). Effect of nonalcoholic fatty liver disease on the development of type 2 diabetes in nonobese, nondiabetic Korean men. *Gut Liver, 6*(3), 368-373.

Conrad, D. F., Pinto, D., Redon, R., Feuk, L., Gokcumen, O., Zhang, Y.,...Hurles, M. E. (2010). Origins and functional impact of copy number variation in the human genome. *Nature, 464*(7289), 704-712.

Constantin, A., Costache, G., Sima, A. V., Glavce, C. S., Vladica, M., & Popov, D. L. (2010). Leptin G-2548A and leptin receptor Q223R gene polymorphisms are not associated with obesity in Romanian subjects. *Biochem Biophys Res Commun, 391*(1), 282-286.

Corey, K. E., Shah, N., Misdraji, J., Abu Dayyeh, B. K., Zheng, H., Bhan, A. K.,...Chung, R. T. (2009). The effect of angiotensin-blocking agents on liver fibrosis in patients with hepatitis C. *Liver Int, 29*(5), 748-753.

Cox, A. J., Wing, M. R., Carr, J. J., Hightower, R. C., Smith, S. C., Xu, J.,...Freedman, B. I. (2011). Association of PNPLA3 SNP rs738409 with liver density in african americans with type 2 diabetes mellitus. *Diabetes Metab, 37*(5), 452-455

Crossette, B. (2011). State of the world population 2011: People and possiblities in a world of 7 billion.

Curtis, C., Lynch, A. G., Dunning, M. J., Spiteri, I., Marioni, J. C., Hadfield, J.,...Caldas, C. (2009). The pitfalls of platform comparison: DNA copy number array technologies assessed. *BMC Genomics, 10*, 588.

Cuthbertson, D. J., Irwin, A., Gardner, C. J., Daousi, C., Purewal, T., Furlong, N.,...Kemp, G. J. (2012). Improved glycaemia correlates with liver fat reduction in obese, type 2 diabetes, patients given glucagon-like peptide-1 (GLP-1) receptor agonists. *PLoS One, 7*(12), e50117.

Dam-Larsen, S., Becker, U., Franzmann, M. B., Larsen, K., Christoffersen, P., & Bendtsen, F. (2009). Final results of a long-term, clinical follow-up in fatty liver patients. *Scand J Gastroenterol, 44*(10), 1236-1243.

Davis, J. N., Le, K. A., Walker, R. W., Vikman, S., Spruijt-Metz, D., Weigensberg, M. J.,...Goran, M. I. (2010). Increased hepatic fat in overweight Hispanic youth influenced by interaction between genetic variation in PNPLA3 and high dietry carbohydrate and sugar consumption. *Am J Clin Nutr, 92*(6), 1522-1527.

Day, C. P. (2002). Pathogenesis of steatohepatitis. *Best Pract Res Clin Gastroenterol, 16*(5), 663-678.

Day, C. P., & James, O. F. (1998). Steatohepatitis: A tale of two "Hits"? *Gastroenterology, 114*(4), 842-845.

Dick, T. J., Lesser, I. A., Leipsic, J. A., Mancini, G. B., & Lear, S. A. (2013). The effect of obesity on the association between liver fat and carotid atherosclerosis in a multi-ethnic cohort. *Atherosclerosis, 226*(1), 208-213.

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*(1), 91-100.

Dolinsky, V. W., Gilham, D., Hatch, G. M., Agellon, L. B., Lehner, R., & Vance, D. E. (2003). Regulation of triacylglycerol hydrolase expression by dietary fatty acids and peroxisomal proliferator-activated receptors. *Biochim Biophys Acta, 1635*(1), 20-28.

Dreyer, W. J. (1998). The area code hypothesis revisited: Olfactory receptors and other related transmembrane receptors may function as the last digits in a cell surface code for assembling embryos. *Proc Natl Acad Sci U S A, 95*(16), 9072-9077.

Ekstedt, M., Franzen, L. E., Mathiesen, U. L., Thorelius, L., Holmqvist, M., Bodemar, G.,...Kechagias, S. (2006). Long-term follow-up of patients with nafld and elevated liver enzymes. *Hepatology, 44*(4), 865-873.

Ertle, J., Dechene, A., Sowa, J. P., Penndorf, V., Herzer, K., Kaiser, G.,...Canbay, A. (2011). Non-alcoholic fatty liver disease progresses to hepatocellular carcinoma in the absence of apparent cirrhosis. *Int J Cancer, 128*(10), 2436-2443.

Fabbrini, E., Sullivan, S., & Klein, S. (2010). Obesity and nonalcoholic fatty liver disease: Biochemical, metabolic, and clinical implications. *Hepatology, 51*(2), 679-689.

Fan, J. G., Saibara, T., Chitturi, S., Kim, B. I., Sung, J. J., & Chutaputti, A. (2007). What are the risk factors and settings for non-alcoholic fatty liver disease in asia-pacific? *J Gastroenterol Hepatol, 22*(6), 794-800.

Farrall, M., & Morris, A. P. (2005). Gearing up for genome-wide gene-association studies. *Hum Mol Genet, 14 Spec No. 2*, R157-162.

Farrell, G. C. (2003). Non-alcoholic steatohepatitis: What is it, and why is it important in the asia-pacific region? *J Gastroenterol Hepatol, 18*(2), 124-138.

Fatani, S., Itua, I., Clark, P., Wong, C., & Naderali, E. K. (2011). The effects of diet-induced obesity on hepatocyte insulin signaling pathways and induction of non-alcoholic liver damage. *Int J Gen Med, 4*, 211-219.

Fenkci, S., Rota, S., Sabir, N., & Akdag, B. (2007). Ultrasonographic and biochemical evaluation of visceral obesity in obese women with non-alcoholic fatty liver disease. *Eur J Med Res, 12*(2), 68-73.

Fernandez-Real, J. M., Mercader, J. M., Ortega, F. J., Moreno-Navarrete, J. M., Lopez-Romero, P., & Ricart, W. (2010). Transferrin receptor-1 gene polymorphisms are associated with type 2 diabetes. *Eur J Clin Invest, 40*(7), 600-607.

Frank, B., Bermejo, J. L., Hemminki, K., Sutter, C., Wappenschmidt, B., Meindl, A.,...Burwinkel, B. (2007). Copy number variant in the candidate tumor suppressor gene MTUS1 and familial breast cancer risk. *Carcinogenesis, 28*(7), 1442-1445.

Freeman, J. L., Perry, G. H., Feuk, L., Redon, R., McCarroll, S. A., Altshuler, D. M.,...Lee, C. (2006). Copy number variation: New insights in genome diversity. *Genome Res, 16*(8), 949-961.

Fuchs, C. D., Claudel, T., Kumari, P., Haemmerle, G., Pollheimer, M. J., Stojakovic, T.,...Trauner, M. (2012). Absence of adipose triglyceride lipase protects from hepatic endoplasmic reticulum stress in mice. *Hepatology, 56*(1), 270-280.

Fujita, K., Yoneda, M., Wada, K., Mawatari, H., Takahashi, H., Kirikoshi, H.,...Nakajima, A. (2007). Telmisartan, an angiotensin ii type 1 receptor blocker, controls progress of nonalcoholic steatohepatitis in rats. *Dig Dis Sci, 52*(12), 3455-3464.

Furusawa, T., Naka, I., Yamauchi, T., Natsuhara, K., Kimura, R., Nakazawa, M.,...Ohashi, J. (2010). The Q223R polymorphism in LEPR is associated with obesity in pacific islanders. *Hum Genet, 127*(3), 287-294.

Gajecka, M., Gentles, A. J., Tsai, A., Chitayat, D., Mackay, K. L., Glotzbach, C. D.,...Shaffer, L. G. (2008). Unexpected complexity at breakpoint junctions in phenotypically normal individuals and mechanisms involved in generating balanced translocations t(1;22)(p36;q13). *Genome Res, 18*(11), 1733-1742.

Gariani, K., Philippe, J., & Jornayvaz, F. R. (2012). Non-alcoholic fatty liver disease and insulin resistance: From bench to bedside. *Diabetes Metab*.

Gibbs, J. R., & Singleton, A. (2006). Application of genome-wide single nucleotide polymorphism typing: Simple association and beyond. *PLoS Genet, 2*(10), e150.

Giglio, S., Broman, K. W., Matsumoto, N., Calvari, V., Gimelli, G., Neumann, T.,...Zuffardi, O. (2001). Olfactory receptor-gene clusters, genomic-inversion polymorphisms, and common chromosome rearrangements. *Am J Hum Genet, 68*(4), 874-883.

Glusman, G., Yanai, I., Rubin, I., & Lancet, D. (2001). The complete human olfactory subgenome. *Genome Res, 11*(5), 685-702.

Godlee, Fiona. (2011). Non-alcoholic fatty liver disease. *BMJ, 343*.

Goh, S. C., Ho, E.L.M, & Goh, K. L. (2012). Prevalence and risk factors of non-alcoholic fatty liver disease in a multiracial suburban asian population in malaysia. *Hepatol Int*.

Greco, D., Kotronen, A., Westerbacka, J., Puig, O., Arkkila, P., Kiviluoto, T.,...Yki-Jarvinen, H. (2008). Gene expression in human nafld. *Am J Physiol Gastrointest Liver Physiol, 294*(5), G1281-1287.

Grundy, S. M., Brewer, H. B., Jr., Cleeman, J. I., Smith, S. C., Jr., & Lenfant, C. (2004). Definition of metabolic syndrome: Report of the national heart, lung, and blood institute/american heart association conference on scientific issues related to definition. *Circulation, 109*(3), 433-438.

Guitart, J., Weisenburger, D. D., Subtil, A., Kim, E., Wood, G., Duvic, M.,...Kim, Y. H. (2012). Cutaneous gammadelta t-cell lymphomas: A spectrum of presentations with overlap with other cytotoxic lymphomas. *Am J Surg Pathol, 36*(11), 1656-1665.

Haentjens, P., Massaad, D., Reynaert, H., Peeters, E., Van Meerhaeghe, A., Vinken, S.,...Velkeniers, B. (2009). Identifying non-alcoholic fatty liver disease among asymptomatic overweight and obese individuals by clinical and biochemical characteristics. *Acta Clin Belg, 64*(6), 483-493.

Hammer, M. F., Karafet, T. M., Park, H., Omoto, K., Harihara, S., Stoneking, M.,...Horai, S. (2006). Dual origins of the Japanese: Common ground for hunter-gatherer and farmer Y chromosomes. *J Hum Genet, 51*(1), 47-58.

Harada, T., Arii, S., Mise, M., Imamura, T., Higashitsuji, H., Furutani, M.,...Imamura, M. (1998). Membrane-type matrix metalloproteinase-1(MT1-MTP) gene is overexpressed in highly invasive hepatocellular carcinomas. *J Hepatol, 28*(2), 231-239.

Hastings, P. J., Lupski, J. R., Rosenberg, S. M., & Ira, G. (2009). Mechanisms of change in gene copy number. *Nat Rev Genet, 10*(8), 551-564.

Hegele, R. A. (2007). Copy-number variations add a new layer of complexity in the human genome. *CMAJ, 176*(4), 441-442.

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

Heuer, M., Kaiser, G. M., Kahraman, A., Banysch, M., Saner, F. H., Mathe, Z.,...Treckmann, J. W. (2012). Liver transplantation in nonalcoholic steatohepatitis is associated with high mortality and post-transplant complications: A single-center experience. *Digestion, 86*(2), 107-113.

Hirose, A., Ono, M., Saibara, T., Nozaki, Y., Masuda, K., Yoshioka, A.,...Onishi, S. (2007). Angiotensin II type 1 receptor blocker inhibits fibrosis in rat nonalcoholic steatohepatitis. *Hepatology, 45*(6), 1375-1381.

Hishida, A., Morita, E., Naito, M., Okada, R., Wakai, K., Matsuo, K.,...Hamajima, N. (2012). Associations of apolipoprotein a5 (APOA5), glucokinase (GCK) and glucokinase regulatory protein (GCKR) polymorphisms and lifestyle factors with the risk of dyslipidemia and dysglycemia in Japanese - a cross-sectional data from the j-micc study. *Endocr J, 59*(7), 589-599.

Hock, S. S. (2007). Chinese migration. Singapore: ISEAS Publishing.

Hoekstra, M., Li, Z., Kruijt, J. K., Van Eck, M., Van Berkel, T. J., & Kuiper, J. (2010). The expression level of non-alcoholic fatty liver disease-related gene PNPLA3 in hepatocytes is highly influenced by hepatic lipid status. *J Hepatol, 52*(2), 244-251.

Hotta, K., Yoneda, M., Hyogo, H., Ochi, H., Mizusawa, S., Ueno, T.,...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*(1), 172.

Hui, Y., Yu-Yuan, L., Yu-Qiang, N., Wei-Hong, S., Yan-Lei, D., Xiao-Bo, L.,...Yong-Jian, Z. (2008). Effect of peroxisome proliferator-activated receptors-gamma and co-activator-1alpha genetic polymorphisms on plasma adiponectin levels and susceptibility of non-alcoholic fatty liver disease in Chinese people. *Liver Int, 28*(3), 385-392.

Irvin, M. R., Wineinger, N. E., Rice, T. K., Pajewski, N. M., Kabagambe, E. K., Gu, C. C.,...Arnett, D. K. (2011). Genome-wide detection of allele specific copy number variation associated with insulin resistance in African Americans from the Hypergen Study. *PLoS One, 6*(8), e24052.

Jiang, J., & Torok, N. (2008). Nonalcoholic steatohepatitis and the metabolic syndrome. *Metab Syndr Relat Disord, 6*(1), 1-7.

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*(9), 1141-1145.

Jorde, L. B., Watkins, W. S., Bamshad, M. J., Dixon, M. E., Ricker, C. E., Seielstad, M. T.,...Batzer, M. A. (2000). The distribution of human genetic diversity: A comparison of mitochondrial, autosomal, and Y-chromosome data. *Am J Hum Genet, 66*(3), 979-988.

Jorde, L. B., & Wooding, S. P. (2004). Genetic variation, classification and 'race'. *Nat Genet, 36*(11 Suppl), S28-33.

Jou, J., Choi, S. S., & Diehl, A. M. (2008). Mechanisms of disease progression in nonalcoholic fatty liver disease. *Semin Liver Dis, 28*(4), 370-379.

Kantartzis, K., Peter, A., Machicao, F., Machann, J., Wagner, S., Konigsrainer, I.,...Stefan, N. (2009). Dissociation between fatty liver and insulin resistance in humans carrying a variant of the patatin-like phospholipase 3 gene. *Diabetes, 58*(11), 2616-2623.

Karlsen, T. H. (2009). Genome-wide association studies reach hepatology. *J Hepatol, 50*(6), 1278-1280.

Kidd, J. M., Cooper, G. M., Donahue, W. F., Hayden, H. S., Sampas, N., Graves, T.,...Eichler, E. E. (2008). Mapping and sequencing of structural variation from eight human genomes. *Nature, 453*(7191), 56-64.

Kim, H. Y., Cho, S., Yu, J., Sung, S., & Kim, H. (2010). Analysis of copy number variation in 8,842 Korean individuals reveals 39 genes associated with hepatic biomarkers ast and alt. *BMB Rep, 43*(8), 547-553.

Kleiner, D. E., Brunt, E. M., Van Natta, M., Behling, C., Contos, M. J., Cummings, O. W.,...Sanyal, A. J. (2005). Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology, 41*(6), 1313-1321.

Kohjima, M., Enjoji, M., Higuchi, N., Kato, M., Kotoh, K., Yoshimoto, T.,...Nakamuta, M. (2007). Re-evaluation of fatty acid metabolism-related gene expression in nonalcoholic fatty liver disease. *Int J Mol Med, 20*(3), 351-358.

Kotronen, A., Johansson, L. E., Johansson, L. M., Roos, C., Westerbacka, J., Hamsten, A.,...Yki-Jarvinen, H. (2009). A common variant in PNPLA3, which encodes adiponutrin, is associated with liver fat content in humans. *Diabetologia, 52*(6), 1056-1060.

Kotronen, A., Westerbacka, J., Bergholm, R., Pietilainen, K. H., & Yki-Jarvinen, H. (2007). Liver fat in the metabolic syndrome. *J Clin Endocrinol Metab, 92*(9), 3490-3497.

Lacaria, M., Saha, P., Potocki, L., Bi, W., Yan, J., Girirajan, S.,...Gu, W. (2012). A duplication CNV that conveys traits reciprocal to metabolic syndrome and protects against diet-induced obesity in mice and men. *PLoS Genet, 8*(5), e1002713.

Laity, J. H., Lee, B. M., & Wright, P. E. (2001). Zinc finger proteins: New insights into structural and functional diversity. *Curr Opin Struct Biol, 11*(1), 39-46.

Lanktree, M., & Hegele, R. A. (2008). Copy number variation in metabolic phenotypes. *Cytogenet Genome Res, 123*(1-4), 169-175.

Lee, J. A., Inoue, K., Cheung, S. W., Shaw, C. A., Stankiewicz, P., & Lupski, J. R. (2006). Role of genomic architecture in PLP1 duplication causing pelizaeus-merzbacher disease. *Hum Mol Genet, 15*(14), 2250-2265.

Li, X., Meng, Y., Wu, P., Zhang, Z., & Yang, X. (2007). Angiotensin II and aldosterone stimulating NF-kappaB and AP-1 activation in hepatic fibrosis of rat. *Regul Pept, 138*(1), 15-25.

Li, Z., Yang, S., Lin, H., Huang, J., Watkins, P. A., Moser, A. B.,...Diehl, A. M. (2003). Probiotics and antibodies to tnf inhibit inflammatory activity and improve nonalcoholic fatty liver disease. *Hepatology, 37*(2), 343-350.

Lima, M. L., Mourao, S. C., Diniz, M. T., & Leite, V. H. (2005). Hepatic histopathology of patients with morbid obesity submitted to gastric bypass. *Obes Surg, 15*(5), 661-669.

Lin, Y. C., Chang, P. F., Hu, F. C., Yang, W. S., Chang, M. H., & Ni, Y. H. (2011). A common variant in the PNPLA3 gene is a risk factor for non-alcoholic fatty liver disease in obese taiwanese children. *J Pediatr, 158*(5), 740-744.

Liu, C. J. (2012). Prevalence and risk factors for non-alcoholic fatty liver disease in asian people who are not obese. *J Gastroenterol Hepatol, 27*(10), 1555-1560.

Liu, N., Crasto, C. J., & Ma, M. (2007). Integrated olfactory receptor and microarray gene expression databases. *BMC Bioinformatics, 8*, 231.

Lou, X. Y., Chen, G. B., Yan, L., Ma, J. Z., Zhu, J., Elston, R. C.,...Li, M. D. (2007). A generalized combinatorial approach for detecting gene-by-gene and gene-by-environment interactions with application to nicotine dependence. *Am J Hum Genet, 80*(6), 1125-1137.

Lu, M., Wan, M., Leavens, K. F., Chu, Q., Monks, B. R., Fernandez, S.,...Birnbaum, M. J. (2012). Insulin regulates liver metabolism in vivo in the absence of hepatic AKT and FOXO1. *Nat Med, 18*(3), 388-395.

Machado, M., Marques-Vidal, P., & Cortez-Pinto, H. (2006). Hepatic histology in obese patients undergoing bariatric surgery. *J Hepatol, 45*(4), 600-606.

Malaguarnera, M., Di Rosa, M., Nicoletti, F., & Malaguarnera, L. (2009). Molecular mechanisms involved in nafld progression. *J Mol Med (Berl), 87*(7), 679-695.

Marchesini, G., Bugianesi, E., Forlani, G., Cerrelli, F., Lenzi, M., Manini, R.,...Rizzetto, M. (2003). Nonalcoholic fatty liver, steatohepatitis, and the metabolic syndrome. *Hepatology, 37*(4), 917-923.

Marchini, J., Cardon, L. R., Phillips, M. S., & Donnelly, P. (2004). The effects of human population structure on large genetic association studies. *Nat Genet, 36*(5), 512-517.

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*(6), 1413-1419.

Mattevi, V. S., Zembrzuski, V. M., & Hutz, M. H. (2002). Association analysis of genes involved in the leptin-signaling pathway with obesity in brazil. *Int J Obes Relat Metab Disord, 26*(9), 1179-1185.

Miele, L., Beale, G., Patman, G., Nobili, V., Leathart, J., Grieco, A.,...Reeves, H. L. (2008). The kruppel-like factor 6 genotype is associated with fibrosis in nonalcoholic fatty liver disease. *Gastroenterology, 135*(1), 282-291 e281.

Mofrad, P., Contos, M. J., Haque, M., Sargeant, C., Fisher, R. A., Luketic, V. A.,...Sanyal, A. J. (2003). Clinical and histologic spectrum of nonalcoholic fatty liver disease associated with normal alt values. *Hepatology, 37*(6), 1286-1292.

Moreno, M., Gonzalo, T., Kok, R. J., Sancho-Bru, P., van Beuge, M., Swart, J.,...Bataller, R. (2010). Reduction of advanced liver fibrosis by short-term targeted delivery of an angiotensin receptor blocker to hepatic stellate cells in rats. *Hepatology, 51*(3), 942-952.

Nabeshima, Y., Tazuma, S., Kanno, K., Hyogo, H., & Chayama, K. (2009). Deletion of angiotensin II type 1 receptor reduces hepatic steatosis. *J Hepatol, 50*(6), 1226-1235.

Nagashima, S., Yagyu, H., Takahashi, N., Kurashina, T., Takahashi, M., Tsuchita, T.,...Ishibashi, S. (2011). Depot-specific expression of lipolytic genes in human adipose tissues--association among CES1 expression, triglyceride lipase activity and adiposity. *J Atheroscler Thromb, 18*(3), 190-199.

Oakley, F., Teoh, V., Ching, A. Sue G., Bataller, R., Colmenero, J., Jonsson, J. R.,...Mann, D. A. (2009). Angiotensin II activates I kappaB kinase phosphorylation of RelA at Ser 536 to promote myofibroblast survival and liver fibrosis. *Gastroenterology, 136*(7), 2334-2344 e2331.

Okada, T., Ohzeki, T., Nakagawa, Y., Sugihara, S., & Arisaka, O. (2010). Impact of leptin and leptin-receptor gene polymorphisms on serum lipids in Japanese obese children. *Acta Paediatr, 99*(8), 1213-1217.

Omoto, K., & Saitou, N. (1997). Genetic origins of the Japanese: A partial support for the dual structure hypothesis. *Am J Phys Anthropol, 102*(4), 437-446.

Ong, J. P., Elariny, H., Collantes, R., Younoszai, A., Chandhoke, V., Reines, H. D.,...Younossi, Z. M. (2005). Predictors of nonalcoholic steatohepatitis and advanced fibrosis in morbidly obese patients. *Obes Surg, 15*(3), 310-315.

Orozco, L. D., Cokus, S. J., Ghazalpour, A., Ingram-Drake, L., Wang, S., van Nas, A.,...Lusis, A. J. (2009). Copy number variation influences gene expression and metabolic traits in mice. *Hum Mol Genet, 18*(21), 4118-4129.

Park, J. W., Jeong, G., Kim, S. J., Kim, M. K., & Park, S. M. (2007). Predictors reflecting the pathological severity of non-alcoholic fatty liver disease: Comprehensive study of clinical and immunohistochemical findings in younger asian patients. *J Gastroenterol Hepatol, 22*(4), 491-497.

Park, S. K., Seo, M. H., Shin, H. C., & Ryoo, J. H. (2012). The clinical availability of non-alcoholic fatty liver disease as an early predictor of type 2 diabetes mellitus in Korean men: 5-years' prospective cohort study. *Hepatology*.

Pasetto, A., Frelin, L., Aleman, S., Holmstrom, F., Brass, A., Ahlen, G.,...Chen, M. (2012). TCR-redirected human T cells inhibit hepatitis C virus replication: Hepatotoxic potential is linked to antigen specificity and functional avidity. *J Immunol, 189*(9), 4510-4519.

Periasamy, M. (2007). Indian migration into malaya and singapore during the british period. *Biblioasia, 3,* 8.

Peters, U., North, K. E., Sethupathy, P., Buyske, S., Haessler, J., Jiao, S.,...Kooperberg, C. (2013). A systematic mapping approach of 16q12.2/FTO and BMI in more than 20,000 African Americans narrows in on the underlying functional variation: Results from the Population Architecture using Genomics and Epidemiology (PAGE) study. *PLoS Genet, 9*(1), e1003171.

Petersen, K. F., Dufour, S., Feng, J., Befroy, D., Dziura, J., Dalla Man, C.,...Shulman, G. I. (2006). Increased prevalence of insulin resistance and nonalcoholic fatty liver disease in asian-Indian men. *Proc Natl Acad Sci U S A, 103*(48), 18273-18277.

Petit, J. M., Guiu, B., Masson, D., Duvillard, R., Jooste, V., Buffier, P.,...Verges, B. (2010). Specifically PNPLA3-mediated accumulation of liver fat in obese patients with type 2 diabetes. *J Clin Endocrinol Metab, 95*(12), E430-E436.

Povel, C. M., Boer, J. M., Reiling, E., & Feskens, E. J. (2011). Genetic variants and the metabolic syndrome: A systematic review. *Obes Rev, 12*(11), 952-967.

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*(1), 74-80.

Prashanth, M., Ganesh, H. K., Vima, M. V., John, M., Bandgar, T., Joshi, S. R.,...Shah, N. S. (2009). Prevalence of nonalcoholic fatty liver disease in patients with type 2 diabetes mellitus. *J Assoc Physicians India, 57*, 205-210.

Puri, P., Mirshahi, F., Cheung, O., Natarajan, R., Maher, J. W., Kellum, J. M.,...Sanyal, A. J. (2008). Activation and dysregulation of the unfolded protein response in nonalcoholic fatty liver disease. *Gastroenterology, 134*(2), 568-576.

Pyrzak, B., Wisniewska, A., Kucharska, A., Wasik, M., & Demkow, U. (2009). No association of LEPR Gln223Arg polymorphism with leptin, obesity or metabolic disturbances in children. *Eur J Med Res, 14 Suppl 4*, 201-204.

Qiao, A., Liang, J., Ke, Y., Li, C., Cui, Y., Shen, L.,...Chang, Y. (2011). Mouse patatin-like phospholipase domain-containing 3 influences systemic lipid and glucose homeostasis. *Hepatology, 54*(2), 509-521.

Quiroga, A. D., Li, L., Trotzmuller, M., Nelson, R., Proctor, S. D., Kofeler, H.,...Lehner, R. (2012). Deficiency of carboxylesterase 1/esterase-x results in obesity, hepatic steatosis, and hyperlipidemia. *Hepatology, 56*(6), 2188-2198.

Rafiq, N., Bai, C., Fang, Y., Srishord, M., McCullough, A., Gramlich, T.,...Younossi, Z. M. (2009). Long-term follow-up of patients with nonalcoholic fatty liver. *Clin Gastroenterol Hepatol, 7*(2), 234-238.

Rafiq, S., Venkata, K. K., Gupta, V., Guru, V. D., Spurgeon, C. J., Parameshwaran, S.,...Chandak, G. R. (2012). Evaluation of seven common lipid associated loci in a large Indian sib pair study. *Lipids Health Dis, 11*(1), 155.

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*(2), 185-192.

Redon, R., Ishikawa, S., Fitch, K. R., Feuk, L., Perry, G. H., Andrews, T. D.,...Hurles, M. E. (2006). Global variation in copy number in the human genome. *Nature, 444*(7118), 444-454.

Romeo, S., Kozlitina, J., Xing, C., Pertsemlidis, A., Cox, D., Pennacchio, L. A.,...Hobbs, H. H. (2008). Genetic variation in PNPLA3 confers susceptibility to nonalcoholic fatty liver disease. *Nat Genet, 40*(12), 1461-1465.

Romeo, S., Sentinelli, F., Cambuli, V. M., Incani, M., Congiu, T., Matta, V.,...Baroni, M. G. (2010a). The 148m allele of the PNPLA3 gene is associated with indices of liver damage early in life. *J Hepatol, 53*(2), 335-338.

Romeo, S., Sentinelli, F., Dash, S., Yeo, G. S., Savage, D. B., Leonetti, F.,...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*(1), 190-194.

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*(3), 894-903.

Ruhl, C. E., & Everhart, J. E. (2005). Joint effects of body weight and alcohol on elevated serum alanine aminotransferase in the united states population. *Clin Gastroenterol Hepatol, 3*(12), 1260-1268.

Ryan, M. C., Wilson, A. M., Slavin, J., Best, J. D., Jenkins, A. J., & Desmond, P. V. (2005). Associations between liver histology and severity of the metabolic

syndrome in subjects with nonalcoholic fatty liver disease. *Diabetes Care, 28*(5), 1222-1224.

Ryu, S. H., Chung, Y. H., Lee, H., Kim, J. A., Shin, H. D., Min, H. J.,...Kim, K. W. (2008). Metastatic tumor antigen 1 is closely associated with frequent postoperative recurrence and poor survival in patients with hepatocellular carcinoma. *Hepatology, 47*(3), 929-936.

Sachidanandam, R., Weissman, D., Schmidt, S. C., Kakol, J. M., Stein, L. D., Marth, G.,...Altshuler, D. (2001). A map of human genome sequence variation containing 1.42 million single nucleotide polymorphisms. *Nature, 409*(6822), 928-933.

Saligram, S., Williams, E. J., & Masding, M. G. (2012). Raised liver enzymes in newly diagnosed type 2 diabetes are associated with weight and lipids, but not glycaemic control. *Indian J Endocrinol Metab, 16*(6), 1012-1014.

Santoro, N., Zhang, C. K., Zhao, H., Pakstis, A. J., Kim, G., Kursawe, R.,...Caprio, S. (2012). Variant in the glucokinase regulatory protein (GCKR) gene is associated with fatty liver in obese children and adolescents. *Hepatology, 55*(3), 781-789.

Sanyal, A. J. (2002). Aga technical review on nonalcoholic fatty liver disease. *Gastroenterology, 123*(5), 1705-1725.

Schnabl, B., Valletta, D., Kirovski, G., & Hellerbrand, C. (2011). Zinc finger protein 267 is up-regulated in hepatocellular carcinoma and promotes tumor cell proliferation and migration. *Exp Mol Pathol, 91*(3), 695-701.

Schrider, D. R., & Hahn, M. W. (2010). Gene copy-number polymorphism in nature. *Proc Biol Sci, 277*(1698), 3213-3221.

Schwimmer, J. B., Celedon, M. A., Lavine, J. E., Salem, R., Campbell, N., Schork, N. J.,...Sirlin, C. B. (2009). Heritability of nonalcoholic fatty liver disease. *Gastroenterology, 136*(5), 1585-1592.

Schwimmer, J. B., Deutsch, R., Kahen, T., Lavine, J. E., Stanley, C., & Behling, C. (2006). Prevalence of fatty liver in children and adolescents. *Pediatrics, 118*(4), 1388-1393.

Sebat, J., Lakshmi, B., Troge, J., Alexander, J., Young, J., Lundin, P.,...Wigler, M. (2004). Large-scale copy number polymorphism in the human genome. *Science, 305*(5683), 525-528.

Seo, H. I., Cho, Y. K., Lee, W. Y., Rhee, E. J., Sung, K. C., Kim, B. S.,...Jin, W. (2012). Which metabolic syndrome criteria best predict the presence of non-alcoholic fatty liver disease? *Diabetes Res Clin Pract, 95*(1), 19-24.

Shams, M. E., Al-Gayyar, M. M., & Barakat, E. A. (2011). Type 2 diabetes mellitus-induced hyperglycemia in patients with NAFLD and normal LFTs: Relationship to lipid profile, oxidative stress and pro-inflammatory cytokines. *Sci Pharm, 79*(3), 623-634.

Shen, Y., Wu, L., Xi, B., Liu, X., Zhao, X., Cheng, H.,...Mi, J. (2013). GCKR variants increase triglycerides while protecting from insulin resistance in chinese children. *PLoS One, 8*(1), e55350.

Silha, J. V., Krsek, M., Skrha, J. V., Sucharda, P., Nyomba, B. L., & Murphy, L. J. (2003). Plasma resistin, adiponectin and leptin levels in lean and obese subjects: Correlations with insulin resistance. *Eur J Endocrinol, 149*(4), 331-335.

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

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*(10), 1266-1271.

Sookoian, S., Castano, G., Gemma, C., Gianotti, T. F., & Pirola, C. J. (2007). Common genetic variations in clock transcription factor are associated with nonalcoholic fatty liver disease. *World J Gastroenterol, 13*(31), 4242-4248.

Sookoian, S., Castano, G., Gianotti, T. F., Gemma, C., Rosselli, M. S., & Pirola, C. J. (2008). Genetic variants in STAT3 are associated with nonalcoholic fatty liver disease. *Cytokine, 44*(1), 201-206.

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*(10), 2111-2116.

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), 1-8.

Sorrentino, P., Tarantino, G., Conca, P., Perrella, A., Terracciano, M. L., Vecchione, R.,...Lobello, R. (2004). Silent non-alcoholic fatty liver disease-a clinical-histological study. *J Hepatol, 41*(5), 751-757.

Souza, M. R., Diniz Mde, F., Medeiros-Filho, J. E., & Araujo, M. S. (2012). Metabolic syndrome and risk factors for non-alcoholic fatty liver disease. *Arq Gastroenterol, 49*(1), 89-96.

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*(3), 904-912.

Speliotes, E. K., Yerges-Armstrong, L. M., Wu, J., Hernaez, R., Kim, L. J., Palmer, C. D.,...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*(3), e1001324.

Stankiewicz, P., Shaw, C. J., Dapper, J. D., Wakui, K., Shaffer, L. G., Withers, M.,...Lupski, J. R. (2003). Genome architecture catalyzes nonrecurrent chromosomal rearrangements. *Am J Hum Genet, 72*(5), 1101-1116.

Struben, V. M., Hespenheide, E. E., & Caldwell, S. H. (2000). Nonalcoholic steatohepatitis and cryptogenic cirrhosis within kindreds. *Am J Med, 108*(1), 9-13.

Stylianou, C., Galli-Tsinopoulou, A., Farmakiotis, D., Rousso, I., Karamouzis, M., Koliakos, G.,...Nousia-Arvanitakis, S. (2007). Ghrelin and leptin levels in obese adolescents. Relationship with body fat and insulin resistance. *Hormones (Athens), 6*(4), 295-303.

Suzuki, A., Lindor, K., St Saver, J., Lymp, J., Mendes, F., Muto, A.,...Angulo, P. (2005). Effect of changes on body weight and lifestyle in nonalcoholic fatty liver disease. *J Hepatol, 43*(6), 1060-1066.

Syvanen, A. C. (2001). Accessing genetic variation: Genotyping single nucleotide polymorphisms. *Nat Rev Genet, 2*(12), 930-942.

Tarantino, G., Saldalamacchia, G., Conca, P., & Arena, A. (2007). Non-alcoholic fatty liver disease: Further expression of the metabolic syndrome. *J Gastroenterol Hepatol, 22*(3), 293-303.

Targher, G., Bertolini, L., Padovani, R., Poli, F., Scala, L., Zenari, L.,...Falezza, G. (2006). Non-alcoholic fatty liver disease is associated with carotid artery wall thickness in diet-controlled type 2 diabetic patients. *J Endocrinol Invest, 29*(1), 55-60.

Targher, G., Bertolini, L., Padovani, R., Rodella, S., Tessari, R., Zenari, L.,...Arcaro, G. (2007). Prevalence of nonalcoholic fatty liver disease and its association with cardiovascular disease among type 2 diabetic patients. *Diabetes Care, 30*(5), 1212-1218.

The international hapmap project. (2003). *Nature, 426*(6968), 789-796.

Trauner, M., Arrese, M., & Wagner, M. (2010). Fatty liver and lipotoxicity. *Biochim Biophys Acta, 1801*(3), 299-310.

Turner, D. J., Miretti, M., Rajan, D., Fiegler, H., Carter, N. P., Blayney, M. L.,...Hurles, M. E. (2008). Germline rates of de novo meiotic deletions and duplications causing several genomic disorders. *Nat Genet, 40*(1), 90-95.

Valenti, L., Al-Serri, A., Daly, A. K., Galmozzi, E., Rametta, R., Dongiovanni, P.,...Day, C. P. (2010). Homozygosity for the patatin-like phospholipase-3/adiponutrin I148M polymorphism influences liver fibrosis in patients with nonalcoholic fatty liver disease. *Hepatology, 52*(4), 1274-1280.

Valenti, L., Alisi, A., Galmozzi, E., Bartuli, A., Del Menico, B., Alterio, A.,...Nobili, V. (2010). I148M patatin-like phospholipase domain-containing 3 gene variant and severity of pediatric nonalcoholic fatty liver disease. *Hepatology, 51*(4), 1209-1217.

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*(5), 320-330.

Venkatraman, E. S., & Olshen, A. B. (2007). A faster circular binary segmentation algorithm for the analysis of array cgh data. *Bioinformatics, 23*(6), 657-663.

Volik, S., Raphael, B. J., Huang, G., Stratton, M. R., Bignel, G., Murnane, J.,...Collins, C. (2006). Decoding the fine-scale structure of a breast cancer genome and transcriptome. *Genome Res, 16*(3), 394-404.

Wagenknecht, L. E., Palmer, N. D., Bowden, D. W., Rotter, J. I., Norris, J. M., Ziegler, J.,...Langefeld, C. D. (2011). Association of PNPLA3 with non-alcoholic fatty liver disease in a minority cohort: the Insulin Resistance Atherosclerosis Family Study. *Liver Int, 31*(3), 412-416.

Wang, D. G., Fan, J. B., Siao, C. J., Berno, A., Young, P., Sapolsky, R.,...Lander, E. S. (1998). Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. *Science, 280*(5366), 1077-1082.

Wang, J. K., Feng, Z. W., Li, Y. C., Li, Q. Y., & Tao, X. Y. (2012). Association of tumor necrosis factor-alpha gene promoter polymorphism at sites -308 and -238 with non-alcoholic fatty liver disease: A meta-analysis. *J Gastroenterol Hepatol, 27*(4), 670-676.

Wang, W. Y., Barratt, B. J., Clayton, D. G., & Todd, J. A. (2005). Genome-wide association studies: Theoretical and practical concerns. *Nat Rev Genet, 6*(2), 109-118.

Wauters, M., Mertens, I., Chagnon, M., Rankinen, T., Considine, R. V., Chagnon, Y. C.,...Bouchard, C. (2001). Polymorphisms in the leptin receptor gene, body composition and fat distribution in overweight and obese women. *Int J Obes Relat Metab Disord, 25*(5), 714-720.

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*(10), 2957-2961.

Wilson, S., & Chalasani, N. (2007). Noninvasive markers of advanced histology in nonalcoholic fatty liver disease: Are we there yet? *Gastroenterology, 133*(4), 1377-1378; discussion 1378-1379.

Wong, K. K., deLeeuw, R. J., Dosanjh, N. S., Kimm, L. R., Cheng, Z., Horsman, D. E.,...Lam, W. L. (2007). A comprehensive analysis of common copy-number variations in the human genome. *Am J Hum Genet, 80*(1), 91-104.

Wong, V. W., Wong, G. L., Tsang, S. W., Hui, A. Y., Chan, A. W., Choi, P. C.,...Chan, H. L. (2009). Metabolic and histological features of non-alcoholic fatty liver disease patients with different serum alanine aminotransferase levels. *Aliment Pharmacol Ther, 29*(4), 387-396.

Yan, X., Xu, L., Qi, J., Liang, X., Ma, C., Guo, C.,...Gao, L. (2009). Strail levels and trail gene polymorphisms in Chinese patients with fatty liver disease. *Immunogenetics, 61*(8), 551-556.

Yang, C. W., Li, C. I., Liu, C. S., Bau, D. T., Lin, C. H., Lin, W. Y.,...Lin, C. C. (2013). The joint effect of cigarette smoking and polymorphisms on LRP5, LEPR, near MC4R and SH2B1 genes on metabolic syndrome susceptibility in taiwan. *Mol Biol Rep, 40*(1), 525-533.

Yang, T. L., Guo, Y., Li, S. M., Li, S. K., Tian, Q., Liu, Y. J.,...Deng, H. W. (2012). Ethnic differentiation of copy number variation on chromosome 16p12.3 for association with obesity phenotypes in european and chinese populations. *Int J Obes (Lond)*.

Yang, Z., Wen, J., Tao, X., Lu, B., Du, Y., Wang, M.,...Hu, R. (2011). Genetic variation in the GCKR gene is associated with non-alcoholic fatty liver disease in chinese people. *Mol Biol Rep, 38*(2), 1145-1150.

Yiannakouris, N., Yannakoulia, M., Melistas, L., Chan, J. L., Klimis-Zacas, D., & Mantzoros, C. S. (2001). The Q223R polymorphism of the leptin receptor gene is significantly associated with obesity and predicts a small percentage of body weight and body composition variability. *J Clin Endocrinol Metab, 86*(9), 4434-4439.

Yokohama, S., Yoneda, M., Haneda, M., Okamoto, S., Okada, M., Aso, K.,...Nakamura, K. (2004). Therapeutic efficacy of an angiotensin II receptor antagonist in patients with nonalcoholic steatohepatitis. *Hepatology, 40*(5), 1222-1225.

Yoneda, M., Hotta, K., Nozaki, Y., Endo, H., Tomeno, W., Watanabe, S.,...Nakajima, A. (2009a). Influence of inducible nitric oxide synthase polymorphisms in Japanese patients with non-alcoholic fatty liver disease. *Hepatol Res, 39*(10), 963-971.

Yoneda, M., Hotta, K., Nozaki, Y., Endo, H., Uchiyama, T., Mawatari, H.,...Nakajima, A. (2009b). Association between angiotensin II type 1 receptor polymorphisms and the occurrence of nonalcoholic fatty liver disease. *Liver Int, 29*(7), 1078-1085.

Yoneda, M., Hotta, K., Nozaki, Y., Endo, H., Uchiyama, T., Mawatari, H.,...Nakajima, A. (2008). Association between ppargc1a polymorphisms and the occurrence of nonalcoholic fatty liver disease (NAFLD). *BMC Gastroenterol, 8*, 27.

Yoshiji, H., Kuriyama, S., Yoshii, J., Ikenaka, Y., Noguchi, R., Nakatani, T.,...Fukui, H. (2001). Angiotensin-II type 1 receptor interaction is a major regulator for liver fibrosis development in rats. *Hepatology, 34*(4 Pt 1), 745-750.

Zain, S. M., Mohamed, R., Mahadeva, S., Cheah, P. L., Rampal, S., Basu, R. C.,...Mohamed, Z. (2012a). A multi-ethnic study of a PNPLA3 gene variant and its association with disease severity in non-alcoholic fatty liver disease. *Hum Genet, 131*(7), 1145-1152.

Zain, S. M., Mohamed, Z., Mahadeva, S., Cheah, P. L., Rampal, S., Chin, K. F.,...Mohamed, R. (2012b). The impact of LEPR variants on risk of non-alcoholic fatty liver disease and its interaction with pnpla3 variant. *J Gastroenterol Hepatol*.

Zain, S. M., Mohamed, Z., Mahadeva, S., Rampal, S., Basu, R. C., Cheah, P. L.,...Mohamed, R. (2013). Susceptibility and gene interaction study of the angiotensin II type 1 receptor (AGTR1) gene polymorphisms with non-alcoholic fatty liver disease in a multi-ethnic population. *PLoS One*.

Zastrow, O., Seidel, B., Kiess, W., Thiery, J., Keller, E., Bottner, A.,...Kratzsch, J. (2003). The soluble leptin receptor is crucial for leptin action: Evidence from clinical and experimental data. *Int J Obes Relat Metab Disord, 27*(12), 1472-1478.

Zhang, K., Calabrese, P., Nordborg, M., & Sun, F. (2002). Haplotype block structure and its applications to association studies: Power and study designs. *Am J Hum Genet, 71*(6), 1386-1394.

Zhao, W., Wineinger, N. E., Tiwari, H. K., Mosley, T. H., Broeckel, U., Arnett, D. K.,...Sun, Y. V. (2012). Copy number variations associated with obesity-related traits in African Americans: A joint analysis between genoa and hypergen. *Obesity (Silver Spring), 20*(12), 2431-2437.

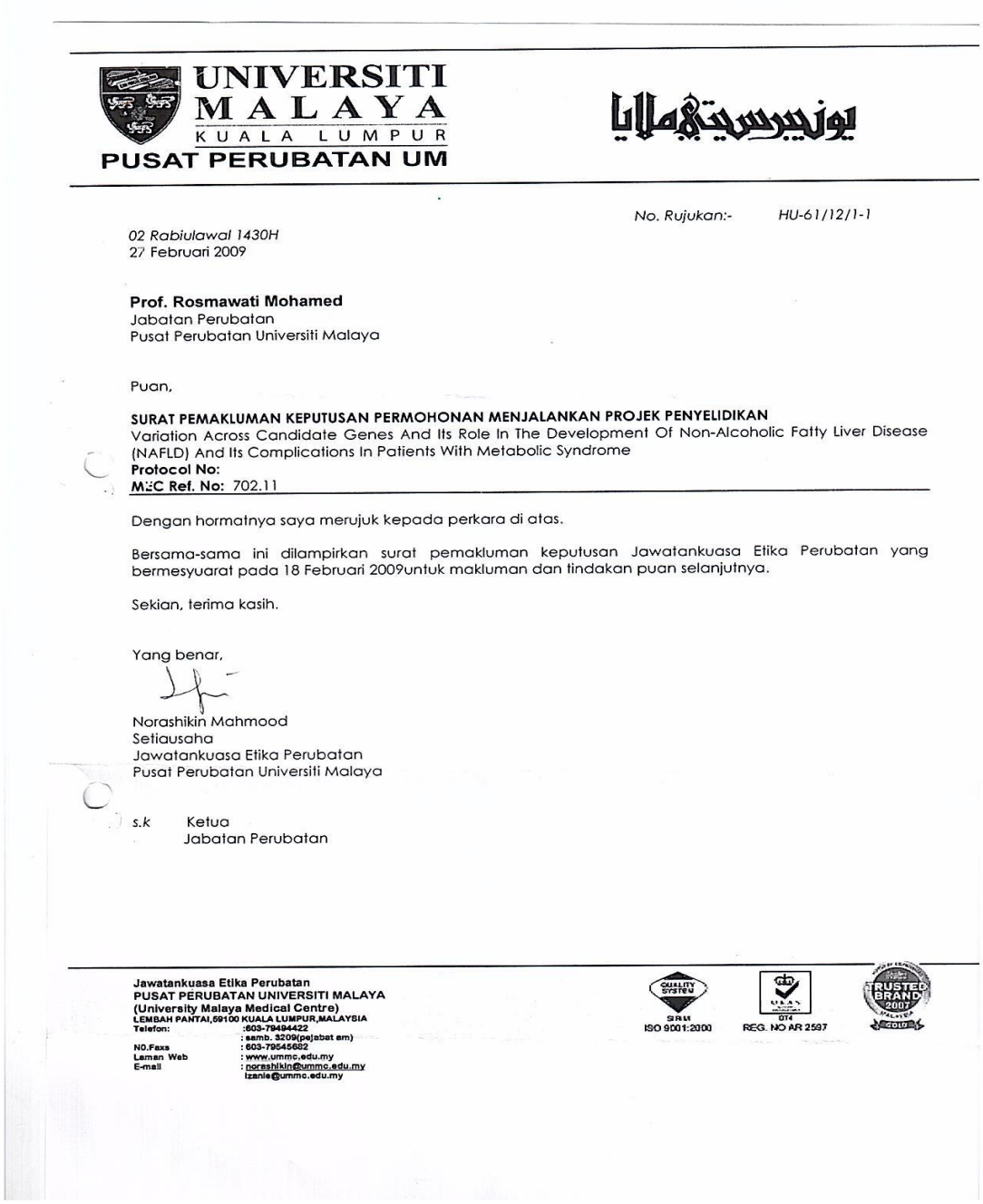
Zhou, X., Hovell, C. J., Pawley, S., Hutchings, M. I., Arthur, M. J., Iredale, J. P.,...Benyon, R. C. (2004). Expression of matrix metalloproteinase-2 and -14 persists during early resolution of experimental liver fibrosis and might contribute to fibrolysis. *Liver Int, 24*(5), 492-501.

Zhou, Y. J., Li, Y. Y., Nie, Y. Q., Yang, H., Zhan, Q., Huang, J.,...Huang, H. L. (2010). Influence of polygenetic polymorphisms on the susceptibility to non-alcoholic fatty liver disease of Chinese people. *J Gastroenterol Hepatol, 25*(4), 772-777.

APPENDICES

APPENDIX A

I) Ethic approval form



UNIVERSITI MALAYA
KUALA LUMPUR
PUSAT PERUBATAN UM

يونيورسيتي ملايا

No. Rujukan:- HU-61/12/1-1

02 Rabiulawal 1430H
27 Februari 2009

Prof. Rosmawati Mohamed
Jabatan Perubatan
Pusat Perubatan Universiti Malaya

Puan,

SURAT PEMAKLUMAN KEPUTUSAN PERMOHONAN MENJALANKAN PROJEK PENYELIDIKAN
Variation Across Candidate Genes And Its Role In The Development Of Non-Alcoholic Fatty Liver Disease (NAFLD) And Its Complications In Patients With Metabolic Syndrome
Protocol No:
MEC Ref. No: 702.11

Dengan hormatnya saya merujuk kepada perkara di atas.

Bersama-sama ini dilampirkan surat pemakluman keputusan Jawatankuasa Etika Perubatan yang bermesyuarat pada 18 Februari 2009untuk makluman dan tindakan puan selanjutnya.

Sekian, terima kasih.

Yang benar,

Norashikin Mahmood
Setiausaha
Jawatankuasa Etika Perubatan
Pusat Perubatan Universiti Malaya

s.k Ketua
Jabatan Perubatan

Jawatankuasa Etika Perubatan
PUSAT PERUBATAN UNIVERSITI MALAYA
(University Malaya Medical Centre)
LEMBAH PANTAI,59100 KUALA LUMPUR,MALAYSIA
Telefon: :603-79494422
: samb. 3209(pejabat am)
NO.Faxs : 603-79545682
Laman Web : www.ummc.edu.my
E-mail : norashikin@ummc.edu.my
izanie@ummc.edu.my

SIRIM ISO 9001:2000 074 REG. NO AR 2597

UNIVERSITI MALAYA
KUALA LUMPUR
PUSAT PERUBATAN UM

JAWATANKUASA ETIKA PERUBATAN
PUSAT PERUBATAN UNIVERSITI MALAYA
ALAMAT: LEMBAH PANTAI, 59100 KUALA LUMPUR, MALAYSIA
TELEFON: 03-79494422 FAKSIMILI: 03-79545682

NAME OF ETHICS COMMITTEE/IRB: Medical Ethics Committee, University Malaya Medical Centre **ADDRESS:** LEMBAH PANTAI 59100 KUALA LUMPUR	**ETHICS COMMITTEE/IRB REFERENCE NUMBER:** 702.11
PROTOCOL NO: **TITLE:** Variation Across Candidate Genes And Its Role In The Development Of Non-Alcoholic Fatty Liver Disease (NAFLD) And Its Complications In Patients With Metabolic Syndrome	
PRINCIPAL INVESTIGATOR: Prof. Rosmawati Mohamed **TELEPHONE:** **KOMTEL:**	**SPONSOR:**

The following item [✓] have been received and reviewed in connection with the above study to be conducted by the above investigator.

[✓] Borang Permohonan Penyelidikan	Ver date: 6 Feb 09
[✓] Study Protocol	Ver date:
[ ] Investigator Brochure	Ver date:
[✓] Patient Information Sheet	Ver date:
[✓] Consent Form	Ver date:
[ ] Questionnaire	
[✓] Investigator(s) CV's (Prof. Rosmawati Mohamed)	

and have been [✓]

[✓] Approved
[ ] Conditionally approved (identify item and specify modification below or in accompanying letter)
[ ] Rejected (identify item and specify reasons below or in accompanying letter)

Comments:

i. *Investigator is required to follow instructions, guidelines and requirements of the Medical Ethics Committee.*
ii. *Investigator is required to report any protocol deviations/violations through the Clinical Investigation Centre and provide annual/closure reports to the Medical Ethics Committee.*

Date of approval: 18th FEBRUARY 2009

s.k Ketua
Jabatan Perubatan

Timbalan Dekan (Penyelidikan)
Fakulti Perubatan, Universiti Malaya

Setiausaha
Jawatankuasa Penyelidikan Pusat Perubatan
Fakulti Perubatan, Universiti Malaya

.....
PROF. LOOI LAI MENG
Chairman
Medical Ethics Committee

UNIVERSITI MALAYA
KUALA LUMPUR
PUSAT PERUBATAN UM

JAWATANKUASA ETIKA PERUBATAN
PUSAT PERUBATAN UNIVERSITI MALAYA
ALAMAT: LEMBAH PANTAI, 59100 KUALA LUMPUR, MALAYSIA
TELEFON: 03-79494422 FAKSIMILI: 03-79545682

MEDICAL ETHICS COMMITTEE COMPOSITION, UNIVERSITY MALAYA MEDICAL CENTRE

Date: 18th FEBRUARY 2009

Member (Title and Name)	**Occupation (Designation)**	**Male/Female (M/F)**	**Tick (✓) if present when above items were reviewed**
Chairperson: Prof. Looi Lai Meng	Representative Dean/Director	Female	
Deputy Chairperson: Prof. Kulenthran Arumugam	Consultant Medical Education Research and Development Unit (MeRDU)	Male	✓
Secretary (non-voting): Norashikin Mahmood	Science Officer, Medical Development Unit	Female	✓
Members: 1. Prof. Jamiyah Hassan	Deputy Chairman (Professional)	Female	
2. Prof. Mohd Hussain Habil	Head Department of Psychological Medicine	Male	✓
3. Assoc. Prof. Mohamed Ibrahim Noordin	Head Department of Pharmacy, FOM	Male	✓
4. Prof. Tan Chong Tin	Representative of Head Department of Medicine	Male	✓
5. Assoc. Prof. George Lee Eng Geap	Representative of Head Department of Surgery	Male	✓
6. Assoc. Prof. Grace Xavier	Representative of Dean Faculty of Law	Female	
7. Tuan Haji Amrahi b. Buang	Senior Manager, Department of Pharmacy, UMMC	Female	✓
8. YBhg. Datin Aminah Pit Abdul Rahman	Public Representative	Female	✓
9. Madam Ong Eng Lee	Public Representative.	Female	✓

Comments: The MEC of University Malaya Medical Centre is operating according to ICH GCP guideline and the Declaration of Helsinki. Members no. 6, 8 & 9 are representatives from Faculty of Law in the University of Malaya and the public, respectively. They are independent of the hospital or trial site.

.....

PROF. LOOI LAI MENG
Chairman
Medical Ethics Committee

II) Informed consent form

A Clinical, Biochemical & Molecular Study of Multiethnic Metabolic Syndrome Subjects

A. Personal Particulars:

i. Name :En/ Puan/ Cik

ii. Age & Sex ____ yrs ____ months Male/Female

iii. Local Address

iv. Permanent Address

v. Occupation

vi. IC number

B. Informed consent

I, _____ (Full Name) have been explained about the Study of Metabolic Syndrome (MS) in my first language _____

today on _____ (Date) by Dr. _____ (Clinician).

1. I have read the information sheet called *Subject Information Sheet* explaining this study and I have been given a copy for my reference.
2. I agree to be clinically examined by a Clinician of Project and permit for my clinical data to be used as part of the project.
3. I am fully aware that being a study subject in this project does not involve or necessitate me undergoing any invasive methods or procedures other than during my blood/tissue collection.
4. I agree to give a sample of my blood/biopsy tissue to this project for screening which will be collected as per the normal medical practice. My participation in this project is voluntary and I have not been induced with money or gift or coerced for doing the same.

5. I have been assured that being a study subject in this project will in no way cause interference in the regular treatment that I have to undergo for my existing medical condition at the hospital.

6. I am fully aware that being a study subject in this project does not involve or necessitate me to take any new or trial drug in the form of tablets, injections or syrups.

7. I give permission for the authorized staff from the Project team to look at my medical records to obtain necessary information. I have been assured that all my personal and health related information will be kept confidential and all data will be anonymised.

8. I understand that I will be informed if any of the results from the screening done using my blood/tissue sample as part of the project that is important to my health.

9. I understand that I will not benefit financially if this project leads to the development of a new treatment or medical test.

10. I am free to withdraw my Consent for this study at any time of the project without giving a reason and without my medical treatment or legal rights being affected.

11. I am aware and fully agree that the blood sample I have given and the clinical data about me can be stored for possible use to meet future project requirements.

12. I understand that the Project is a noble project that aims towards the study and better understanding of Metabolic Syndrome and for general betterment of the society.

Name of Subject (Capitals)	
Signature & Date	
Name of Project staff obtaining Consent	
Signature & Date	
Independent Witness Name	
Signature & Date	

III) Patient information sheet

A Clinical Biochemical & Molecular Study of Multiethnic Metabolic Syndrome Subjects

Please read the following information carefully, do not hesitate to discuss any question you may have with your Doctor.

What is Metabolic Syndrome?

Metabolic syndrome which is also known as metabolic syndrome X, syndrome X, insulin resistance syndrome or Reaven's syndrome is a combination of medical disorders that increase the risk of developing cardiovascular disease and diabetes.

How do you identify Metabolic Syndrome?

There are several guidelines available for identification of Metabolic Syndrome However, the one which is most accepted and followed in this study is the updated NCEP ATP III (National Cholesterol Education Program Adult Treatment Panel III) criteria for Metabolic Syndrome. The criteria are:

- Men who have waist circumference above 90 cm/30.5 inches and women above 80 cm/31.4 inches
- Blood pressure above 130/85 mm Hg or having normal BP but on medication.
- Fasting glucose above 110 mg/dl (> 6.1 mmol/l)
- Serum Triglycerides above 150 mg/dl (≥1.69 mmol/l)
- High Density Cholesterol less than 40mg/dl (< 1.03 mmol/l) in men and less than 50mg/dl (1.29 mmol/l) in women.

How common is Metabolic Syndrome?

With the presence of Metabolic Syndrome a person's chances of developing cardiovascular disease and diabetes increases manifolds. It is a highly common yet silent condition which affects many people and has high prevalence rate. Some studies estimate the prevalence in the USA to be up to 25% of the population. Malaysia is a developing country with its people due to various life style and genetic factors are at a high risk of developing Metabolic Syndrome. The prevalence of metabolic syndrome among patients with premature vascular disease may be as high as 50%.With the changing lifestyle patterns it is estimated that the number of people afflicted with Metabolic Syndrome could be much higher than what has been anticipated. It is also associated with other diseases like increased uric acid levels in body, fatty liver (especially in concurrent obesity) which can progress to fatty liver disease, polycystic ovarian disease in women. Hence, identification, prevention and treatment of Metabolic Syndrome will help you to ensure a healthy and disease free life. Non alcoholic fatty liver disease (NAFLD) represent the hepatic manifestation of the metabolic syndrome which is characterized by obesity, Type 2 DM, dyslipidaemia, hypertension with insulin resistance. Liver enzymes (ALT, AST) are mildly elevated. Patients with NAFLD typically present in the fourth or fifth decade of life. NAFLD is characterized by excess fat in the liver of people who drink little or no alcohol and have no viral or autoimmune liver disease. In some people with NAFLD, fat accumulation is associated with liver cell injury and inflammation, a condition named nonalcoholic steatohepatitis (NASH)

What is the purpose of this study?

- To determine the possible causes of Metabolic syndrome in all the three ethnicities of Malaysian population namely the Malay, Chinese and Indian.
- To obtain information to help develop methods to diagnose and treat Metabolic Syndrome.

What are the procedures to be followed?

- Please ask the clinical coordinator to explain about this study or project. After you clarify all the doubts, make sure that you have understood the purpose of the study and your role in this project. Then sign the Informed Consent to participate in this study / project.
- You would be asked routine questions about your lifestyle like diet, smoking,

exercise, family history etc. through a questionnaire.

- A routine clinical examination by the clinician in the Out Patient Clinic
- A routine blood sample from your arm vein will be collected only once using sterile, disposable syringe for the blood test.
- Liver biopsy would be done wherein little part of liver tissue will be obtained.

Note : You are not required to take any medicines or come for repeated follow up

Who should not enter the study?

- There is no specific /general risk involved to the subjects as part of this project. Those persons who satisfy the laid down inclusive criteria for selection of subject can participate and be part of this project.
- Those below the age of 18 years need to get informed consent form signed by their parents or guardians

What are the benefits of the study?

(a) To you as the subject

- The result/report concerning yourself can be intimated to your treating Doctor who could use the specific information to treat you better.
- Your family members can be screened at an early age which will help the Doctors to make early diagnosis, advice preventive measures and prevent early complications including death.

(b) To the investigator

This is a noble project that aims towards establishing a clinical, molecular and biochemical profile of Metabolic syndrome amongst Malaysian population.

Are there any possible drawbacks?

- There are no drawbacks for the subjects who are part of the project
- The project does not involve any invasive procedure or administration of any new trial drugs. A control or study subject in this project does not necessitate you to take any new drugs such as tablets, injections or syrups for project purpose.
- Your personal and health related information will be kept confidential and all data will be anonymised to prevent identification of your data in the project.

Can I refuse to take part in the study?

- Yes, you have complete freedom to refuse from taking part of this study.
- Giving your blood sample for testing is voluntary on your part and you will not be induced with money or gift for doing the same.
- You are free to withdraw your consent without giving a reason. Your on-going medical treatment will be unaffected.

How can I participate in the study?

You have to ask the clinical coordinator who will explain the study procedure to you in detail. You will have to sign an informed consent form. After this it will be assessed if you fulfill the study criteria (both inclusion as well as exclusion criteria) and then the decision regarding your participation will be made.

If you need further information please contact /call

1. Prof.Rosmawati Mohamed.79492077

2. Dr.Roma Choudhury Basu 79492886

APPENDIX B

Detail methods of aCGH

I) Labelling the target

Reaction tubes were prepared as a mixture of 18 µL of genomic DNA sample (or reference), 10 µL of random primer and 10 µL of reaction buffer. The reactions were mixed and centrifuged for 15 s, followed by denaturation in a PCR block with a heated lid at 99°C for 20 min. Immediately after, the tubes were placed on ice for 5 min. On ice, the followings were added to the tubes: 10 µL of dCTP labelling mix, 1 µL of Cy3-dCTP (for sample) / 1 µL of Cy5-dCTP (for reference), and 1 µL of Klenow. The mixture was centrifuged for 15 s and incubated at 37°C for 2 h. Following that, the mixture was incubated at 65°C for 10 min, then placed on ice for 5 min. The mixture was centrifuged for 15 s.

II) Purifying the labelled target

Two purification columns were prepared by vortexing the resin briefly. The cap was loosened one-quarter turn and the bottom closure was snapped off. The column was placed in a collection tube and spun for 1 min at 2,000 x *g*. The spin column was removed from the tube and the eluate was discarded. The column was placed in a fresh tube. The lid was removed from the column and the sample was added to the centre of the resin and centrifuged at 2,000 x *g* for 1 min. The final volume was ~45 µl. The absorbance was measured and the amounts of target that should be obtained were as follows: minimum 6 µg of DNA and 250 pmol of dye (for Cy3) and minimum 5 µg of DNA and 200 pmol of dye (for Cy5).

III) Hybridisation of arrays with labelled target

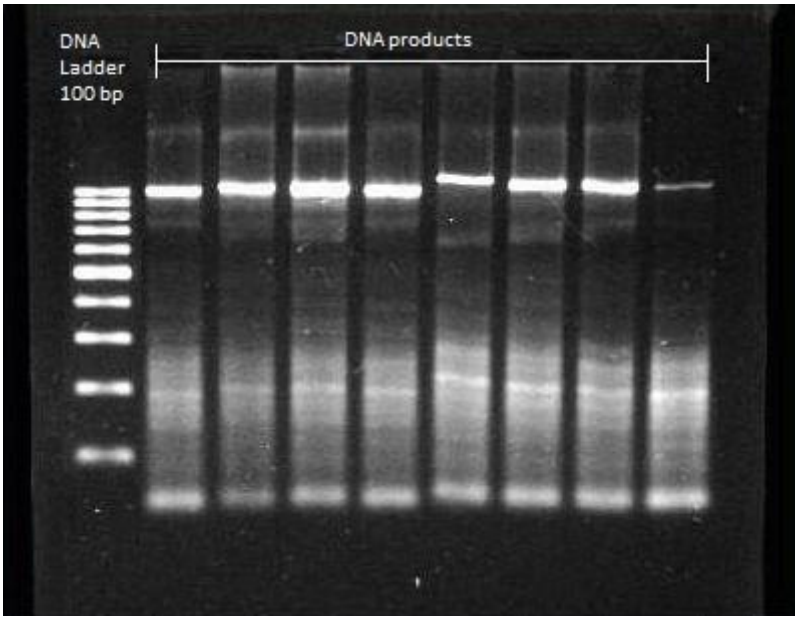
The 10x Blocking Agent was prepared by adding 1,350 µl of water to the 10x Blocking Agent tube. The Blocking Agent was left at room temperature for 60 min. The hybridisation was set up. Each sample was combined with its corresponding reference. A 1 x TE (pH 8.0) was added to give a final volume of 158 µl each. The following hybridisation mix: 158 µl of Cy3 and Cy5-labelled genomic DNA, 50 µl of Cot-1 (1 mg/ml), 52 µl of Agilent 10x Blocking Agent, and 260 µl of Agilent 2x HiRPM Hybridization Buffer. The target was denatured at 94°C for 3 min. The mixture was incubated at 37°C for 30 min and centrifuged for 10 s. The Agilent SureHyb Gasket was placed into an Agilent Chamber base or MaiTai cassette. Immediately, 500 µl of the hybridisation mix was pipetted onto the gasket slide. OGT array was placed onto the Gasket slide. The hybridisation chamber was placed in the hybridisation oven and hybridised at 65°C for 22 h in a light-tight container. The slides were fitted vertically and the chamber was rotated at a speed of 20 rpm.

IV) Washing and scanning of arrays

Wash 2 and one glass dish were pre-warmed to 37°C. Wash 1 was placed in two glass dishes (one for disassemble and one for first wash). The slide was placed in the disassemble bath, the gasket slide was gently prised from the CytoSure array under the surface of the buffer. The slide was quickly transferred to the Wash 1 bath. A stirring flea was placed in the Wash 1 and stirred at room temperature for 5 min. The rack was removed and blotted on a paper towel. The set up was quickly placed into the pre-warmed Wash 2 bath and stirred at 37°C for exactly 1 min. The rack was removed, blotted and blow dried using air. The slide was scanned immediately. The data was then ready to be analysed.

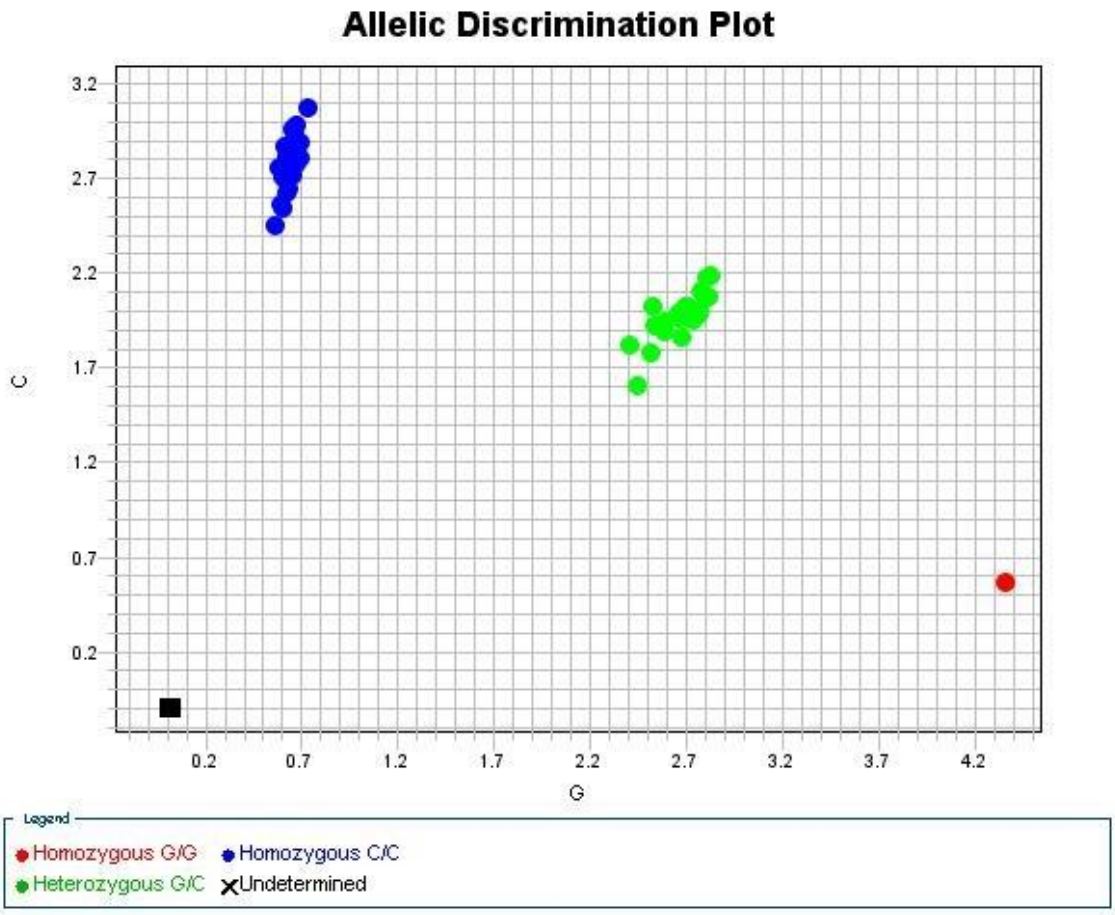
APPENDIX C

Gel picture of the extracted DNA products

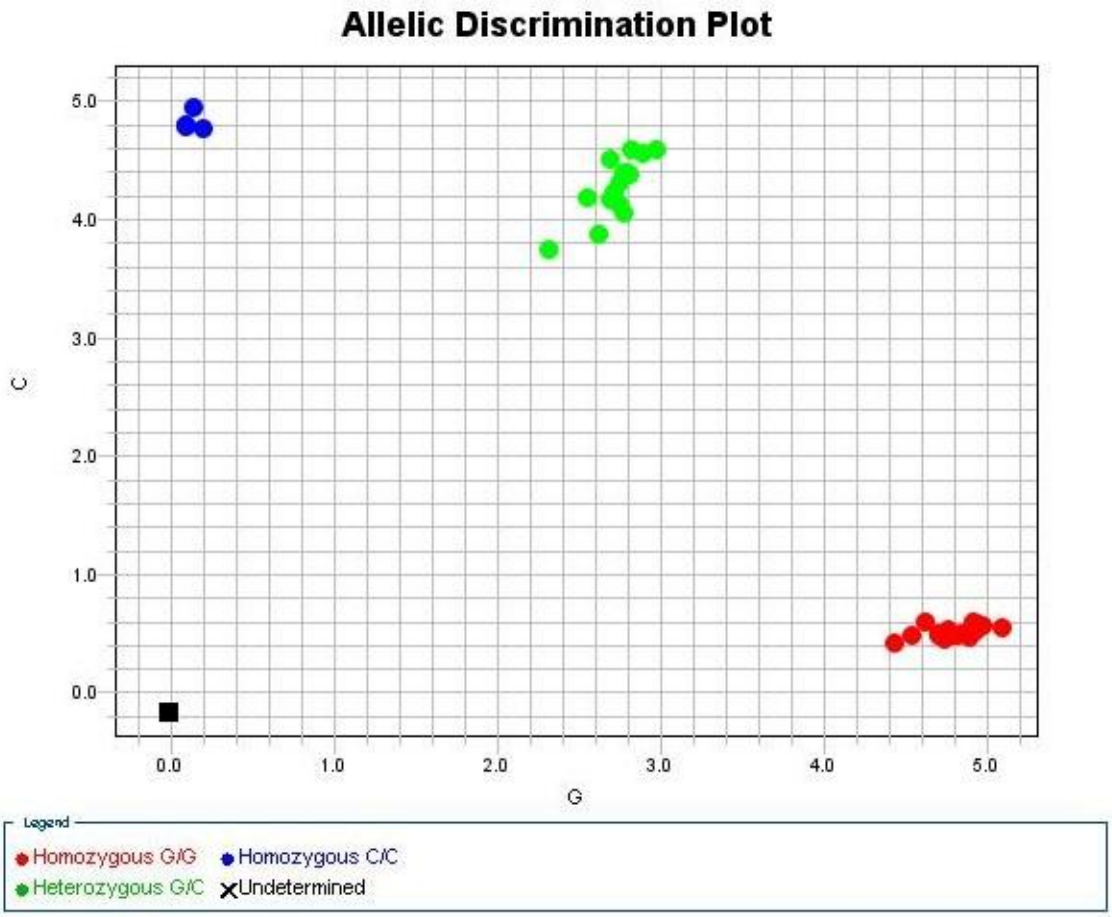


APPENDIX D

i) Allelic discrimination plot of the PNPLA3 rs738409 in controls



ii) Allelic discrimination plot of the PNPLA3 rs738409 in cases



APPENDIX E

Hardy-Weinberg equilibrium of the SNPs in various candidate genes

Gene	SNP(s)	Ethnicity	HWE	
			Case	Control
PNPLA3	rs738409	Overall	0.17	0.36
		Malay	0.33	0.13
		Chinese	0.10	0.30
		Indian	0.52	0.43
AGTR1	rs3772633	Overall	0.34	0.70
		Malay	0.37	0.86
		Chinese	0.84	0.41
		Indian	0.34	0.84
	rs3772627	Overall	0.08	0.33
		Malay	0.40	0.72
		Chinese	0.25	0.33
		Indian	0.11	0.82
	rs2276736	Overall	0.06	0.44
		Malay	0.40	0.92
		Chinese	0.14	0.50
		Indian	0.16	0.80
	rs3772630	Overall	0.19	0.52
		Malay	0.91	0.92
		Chinese	0.25	0.67
		Indian	0.11	0.80
	rs3772622	Overall	0.84	0.23
		Malay	0.56	0.52
		Chinese	0.80	0.09
		Indian	0.44	0.76
LEPR	rs1137100	Overall	0.50	0.12
		Malay	0.07	0.06
		Chinese	0.11	0.79
		Indian	0.07	0.31
	rs1137101	Overall	0.46	0.15
		Malay	0.17	0.80
		Chinese	0.65	0.09
		Indian	0.39	0.37

P > 0.05 indicates SNP in HWE

Hardy-Weinberg equilibrium of the SNPs in various candidate genes (cont.)

Gene	SNP(s)	Ethnicity	HWE	
			Case	Control
iNOS	rs1060822	Overall	< 0.05	< 0.05
		Malay	< 0.05	< 0.05
		Chinese	< 0.05	< 0.05
		Indian	0.40	0.16
TRAIL	rs1131568	Overall	0.09	0.31
		Malay	0.14	0.31
		Chinese	0.68	0.61
		Indian	< 0.05	0.76
SREBF1	rs1186803	Overall	0.55	0.45
		Malay	0.21	0.78
		Chinese	0.16	0.16
		Indian	< 0.05	0.70
	rs2297508	Overall	0.14	0.09
		Malay	0.21	< 0.05
		Chinese	0.41	< 0.05
		Indian	0.48	0.41
CLOCK	rs1193259	Overall	0.61	0.08
		Malay	0.42	0.99
		Chinese	0.50	0.20
		Indian	0.33	0.14
	rs4864548	Overall	0.44	0.25
		Malay	0.66	0.59
		Chinese	0.65	0.22
		Indian	0.12	0.08
	rs6843722	Overall	0.28	0.55
		Malay	0.60	0.52
		Chinese	0.92	0.38
		Indian	0.11	0.21
	rs6850524	Overall	0.46	0.97
		Malay	0.36	0.45
		Chinese	0.21	0.54
		Indian	0.86	0.70

P > 0.05 indicates SNP in HWE

Hardy-Weinberg equilibrium of the SNPs in various candidate genes (cont.)

Gene	SNP(s)	Ethnicity	HWE	
			Case	Control
COL13A1	rs1227756	Overall	< 0.05	0.07
		Malay	0.13	< 0.05
		Chinese	< 0.05	< 0.05
		Indian	0.99	0.24
FABP2	rs1799883	Overall	0.12	0.86
		Malay	0.77	0.07
		Chinese	0.32	0.76
		Indian	0.15	0.25
UCP3	rs1800849	Overall	0.84	0.57
		Malay	0.79	0.36
		Chinese	0.47	0.57
		Indian	0.82	0.14
PPARG	rs1801282	Overall	0.48	0.74
		Malay	0.68	0.77
		Chinese	0.83	0.95
		Indian	0.47	0.78
COX-2	rs20417	Overall	< 0.05	< 0.05
		Malay	< 0.05	< 0.05
		Chinese	< 0.05	< 0.05
		Indian	< 0.05	< 0.05
NR1I2	rs2461823	Overall	0.16	0.86
		Malay	0.67	0.32
		Chinese	0.43	0.58
		Indian	0.18	0.27
	rs7643645	Overall	0.38	< 0.05
		Malay	0.85	< 0.05
		Chinese	0.67	0.29
		Indian	0.12	0.27
KLF6	rs3750861	Overall	0.10	0.87
		Malay	< 0.05	0.55
		Chinese	0.78	0.83
		Indian	0.54	0.64

P > 0.05 indicates SNP in HWE

Hardy-Weinberg equilibrium of the SNPs in various candidate genes (cont.)

Gene	SNP(s)	Ethnicity	HWE	
			Case	Control
TFRC	rs3817672	Overall	0.69	0.64
		Malay	0.79	0.65
		Chinese	0.16	0.82
		Indian	0.60	0.74
PPARGC1A	rs3755863	Overall	0.81	0.66
		Malay	0.73	0.06
		Chinese	0.89	0.87
		Indian	0.79	0.25
FATP5	rs56225452	Overall	0.08	0.96
		Malay	0.34	0.70
		Chinese	0.83	0.95
		Indian	0.23	0.37
STAT3	rs6503695	Overall	0.35	0.74
		Malay	0.19	0.21
		Chinese	0.68	0.38
		Indian	0.72	0.89
	rs9891119	Overall	< 0.05	< 0.05
		Malay	< 0.05	< 0.05
		Chinese	< 0.05	< 0.05
		Indian	< 0.05	< 0.05
APOA5	rs662799	Overall	0.58	0.21
		Malay	0.33	0.23
		Chinese	0.82	0.61
		Indian	0.40	0.07
PEMT	rs7946	Overall	0.70	0.12
		Malay	0.40	0.33
		Chinese	0.41	0.59
		Indian	0.64	0.32
GCKR	rs1260326	Overall	0.77	0.66
		Malay	0.50	0.96
		Chinese	0.60	0.78
		Indian	0.57	0.31
	rs780094	Overall	0.26	0.85
		Malay	0.22	0.64
		Chinese	0.85	0.20
		Indian	0.34	0.25

P > 0.05 indicates SNP in HWE

APPENDIX F

I) Comparison of various clinical and histological parameters between the SREBF1 rs1186803 genotypes among NAFLD patients

Characteristics	NAFLD, *n* = 144 (mean ± SD)			
	CC (*n* = 6)	CG (*n* = 53)	GG (*n* = 85)	*p* value
Age (years)	45.7 ± 10.6	51.1 ± 12.1	51.3 ± 11.9	0.538
BMI (kg/m^2)	30.3 ± 5.2	28.7 ± 4.8	28.5 ± 4.1	0.633
HbA1c (%)*	6.1 ± 0.7	7.0 ± 2.1	6.4 ± 1.3	0.612
Waist circumference (cm)	94.3 ± 11.6	94.8 ± 10.4	93.0 ± 11.1	0.641
HDL cholesterol (mg/dl)	56.1 ± 20.6	46.8 ± 10.1	49.2 ± 13.3	0.189
LDL cholesterol (mg/dl)	134.9 ± 61.8	120.3 ± 38.2	112.1 ± 37.1	0.233
Total cholesterol (mg/dl)	224.6 ± 67.0	198.3 ± 45.4	192.2 ± 39.5	0.183
Triglycerides (mg/dl)	158.3 ± 74.3	163.2 ± 68.2	149.5 ± 56.0	0.465
AST (IU/L)*	57.2 ± 16.8	40.9 ± 26.2	43.4 ± 25.5	0.085
ALT (IU/L)	121.5 ± 44.7	76.4 ± 50.4	84.5 ± 47.5	0.094
GGT (IU/L)*	240.8 ± 164.7	86.8 ± 70.7	119.0 ± 129.5	0.039
Systolic BP (mmHg)	142.5 ± 12.5	133.3 ± 16.0	129.2 ± 14.1	0.050
Diastolic BP (mmHg)*	92.8 ± 6.9	82.3 ± 9.5	81.9 ± 10.0	0.031
Steatosis grade*	1.5 ± 0.6	1.8 ± 0.8	1.8 ± 0.8	0.642
Lobular inflammation*	1.3 ± 0.8	1.4 ± 0.6	1.3 ± 0.6	0.726
Ballooning*	1.0 ± 0.6	1.3 ± 0.6	1.2 ± 0.6	0.419
Fibrosis*	1.7 ± 1.0	1.8 ±1.1	1.7 ± 0.9	0.967

**P* values obtained using Kruskal-Wallis test, all other comparisons used ANOVA

ALT alanine transferase, *AST* aspartate aminotransferase, *BMI* body mass index, *GGT* gamma glutamyl transpeptidase, *HbA1c* haemoglobin A1c, *HDL* high-density lipoprotein, *LDL* low-density lipoprotein, *NAFLD* non-alcoholic fatty liver disease

II) Comparison of various clinical and histological parameters between the CLOCK rs1193259 genotypes among NAFLD patients

Characteristics	NAFLD, *n* = 144 (mean ± SD)			
	AA (*n* = 96)	AG (*n* = 42)	GG (*n* = 6)	*p* value
Age (years)	51.6 ± 12.0	50.7 ± 12.1	45.0 ± 11.3	0.548
BMI (kg/m^2)	28.6 ± 4.8	28.8 ± 3.5	28.8 ± 1.7	0.949
HbA1c (%)*	6.5 ± 1.4	6.8 ± 2.1	6.6 ± 2.5	0.634
Waist circumference (cm)	93.8 ± 11.7	94.0 ± 9.2	93.5 ± 8.4	0.991
HDL cholesterol (mg/dl)	49.6 ± 12.7	46.5 ± 13.1	43.2 ± 6.4	0.287
LDL cholesterol (mg/dl)	115.7 ± 39.6	119.9 ± 40.8	123.0 ± 48.4	0.812
Total cholesterol (mg/dl)	194.4 ± 42.6	201.7 ± 46.4	200.8 ± 49.3	0.661
Triglycerides (mg/dl)	150.0 ± 58.0	167.4 ± 74.7	149.3 ± 21.7	0.323
AST (IU/L)*	41.9 ± 23.5	44.8 ± 29.7	48.3 ± 25.5	0.846
ALT (IU/L)	80.3 ± 45.5	88.1 ± 53.6	97.8 ± 71.1	0.570
GGT (IU/L)*	115.7 ± 39.7	201.7 ± 46.9	200.8 ± 49.3	0.198
Systolic BP (mmHg)	131.4 ± 15.8	132.5 ± 12.5	126.2 ± 21.8	0.728
Diastolic BP (mmHg)*	81.7 ± 10.2	84.1 ± 9.2	82.5 ± 5.0	0.386
Steatosis grade*	1.8 ± 0.8	1.9 ± 0.8	1.5 ± 0.6	0.613
Lobular inflammation*	1.3 ± 0.5	1.3 ± 0.7	1.3 ± 0.5	0.964
Ballooning*	1.2 ± 0.6	1.2 ± 0.5	1.5 ± 0.6	0.601
Fibrosis*	1.6 ± 1.0	1.9 ± 0.9	2.0 ± 1.4	0.181

**P* values obtained using Kruskal-Wallis test, all other comparisons used ANOVA

ALT alanine transferase, *AST* aspartate aminotransferase, *BMI* body mass index, *GGT* gamma glutamyl transpeptidase, *HbA1c* haemoglobin A1c, *HDL* high-density lipoprotein, *LDL* low-density lipoprotein, *NAFLD* non-alcoholic fatty liver disease

III) Comparison of various clinical and histological parameters between the CLOCK rs4864548 genotypes among NAFLD patients

Characteristics	NAFLD, *n* = 144 (mean ± SD)			
	GG (*n* = 26)	GA (*n* = 64)	AA (*n* = 54)	*p* value
Age (years)	50.0 ± 13.0	51.6 ± 11.8	51.2 ± 12.0	0.843
BMI (kg/m^2)	28.5 ± 3.8	28.4 ± 4.3	29.1 ± 4.8	0.617
HbA1c (%)*	6.1 ± 1.2	6.8 ± 1.9	6.5 ± 1.5	0.358
Waist circumference (cm)	92.1 ± 10.3	94.0 ± 10.7	94.5 ± 11.4	0.613
HDL cholesterol (mg/dl)	50.6 ± 17.7	48.2 ± 11.1	47.9 ± 11.4	0.698
LDL cholesterol (mg/dl)	125.0 ± 41.8	114.4 ± 41.0	116.0 ± 38.0	0.499
Total cholesterol (mg/dl)	204.9 ± 45.4	193.9 ± 47.2	195.8 ± 38.9	0.538
Triglycerides (mg/dl)	151.6 ± 48.9	158.4 ± 74.2	152.3 ± 53.7	0.821
AST (IU/L)*	45.5 ± 26.5	40.8 ± 21.7	44.1 ± 28.8	0.829
ALT (IU/L)	93.3 ± 58.1	80.3 ± 44.2	80.8 ± 48.2	0.448
GGT (IU/L)*	137.7 ± 133.5	119.7 ± 122.9	87.8 ± 90.4	0.282
Systolic BP (mmHg)	129.8 ± 12.5	130.5 ± 13.7	133.9 ± 17.7	0.384
Diastolic BP (mmHg)*	82.3 ± 8.9	82.1 ± 9.2	83.5 ± 11.1	0.886
Steatosis grade*	1.8 ± 0.8	1.8 ± 0.8	1.9 ± 0.8	0.728
Lobular inflammation*	1.3 ± 0.5	1.3 ± 0.6	1.3 ± 0.5	0.859
Ballooning*	1.1 ± 0.6	1.2 ± 0.6	1.3 ± 0.6	0.204
Fibrosis*	1.6 ± 1.0	1.7 ± 1.0	1.8 ± 1.0	0.741

**P* values obtained using Kruskal-Wallis test, all other comparisons used ANOVA

ALT alanine transferase, *AST* aspartate aminotransferase, *BMI* body mass index, *GGT* gamma glutamyl transpeptidase, *HbA1c* haemoglobin A1c, *HDL* high-density lipoprotein, *LDL* low-density lipoprotein, *NAFLD* non-alcoholic fatty liver disease

IV) Comparison of various clinical and histological parameters between the CLOCK rs6843722 genotypes among NAFLD patients

Characteristics	NAFLD, *n* = 144 (mean ± SD)			
	AA (*n* = 30)	AC (*n* = 62)	CC (*n* = 52)	*p* value
Age (years)	50.8 ± 12.8	50.9 ± 12.0	51.5 ± 11.8	0.954
BMI (kg/m^2)	28.4 ± 3.7	28.3 ± 4.4	29.2 ± 4.8	0.523
HbA1c (%)*	6.3 ± 1.5	6.7 ± 1.8	6.5 ± 1.5	0.328
Waist circumference (cm)	92.0 ± 9.9	94.1 ± 10.8	94.7 ± 11.6	0.552
HDL cholesterol (mg/dl)	49.4 ± 17.3	48.7 ± 11.1	47.6 ± 11.4	0.828
LDL cholesterol (mg/dl)	127.5 ± 42.8	112.4 ± 39.7	116.7 ± 38.3	0.228
Total cholesterol (mg/dl)	206.2 ± 46.2	192.3 ± 46.2	196.5 ± 39.0	0.351
Triglycerides (mg/dl)	151.4 ± 46.5	158.0 ± 75.2	153.6 ± 53.4	0.876
AST (IU/L)*	45.4 ± 26.1	40.7 ± 21.8	44.2 ± 29.1	0.757
ALT (IU/L)	91.8 ± 56.1	81.3 ± 44.7	80.0 ± 48.3	0.521
GGT (IU/L)*	133.3 ± 127.3	119.7 ± 124.8	88.1 ± 91.4	0.277
Systolic BP (mmHg)	130.8 ± 12.5	130.3 ± 13.7	133.9 ± 17.9	0.418
Diastolic BP (mmHg)*	82.4 ± 8.6	81.8 ± 9.4	83.6 ± 11.2	0.773
Steatosis grade*	1.8 ± 0.8	1.8 ± 0.8	1.9 ± 0.8	0.742
Lobular inflammation*	1.3 ± 0.5	1.3 ± 0.6	1.3 ± 0.5	0.961
Ballooning*	1.2 ± 0.6	1.1 ± 0.6	1.3 ± 0.6	0.174
Fibrosis*	1.7 ± 0.9	1.7 ± 1.0	1.8 ± 1.0	0.867

**P* values obtained using Kruskal-Wallis test, all other comparisons used ANOVA

ALT alanine transferase, *AST* aspartate aminotransferase, *BMI* body mass index, *GGT* gamma glutamyl transpeptidase, *HbA1c* haemoglobin A1c, *HDL* high-density lipoprotein, *LDL* low-density lipoprotein, *NAFLD* non-alcoholic fatty liver disease

V) Comparison of various clinical and histological parameters between the NR1I2 rs7643645 genotypes among NAFLD patients

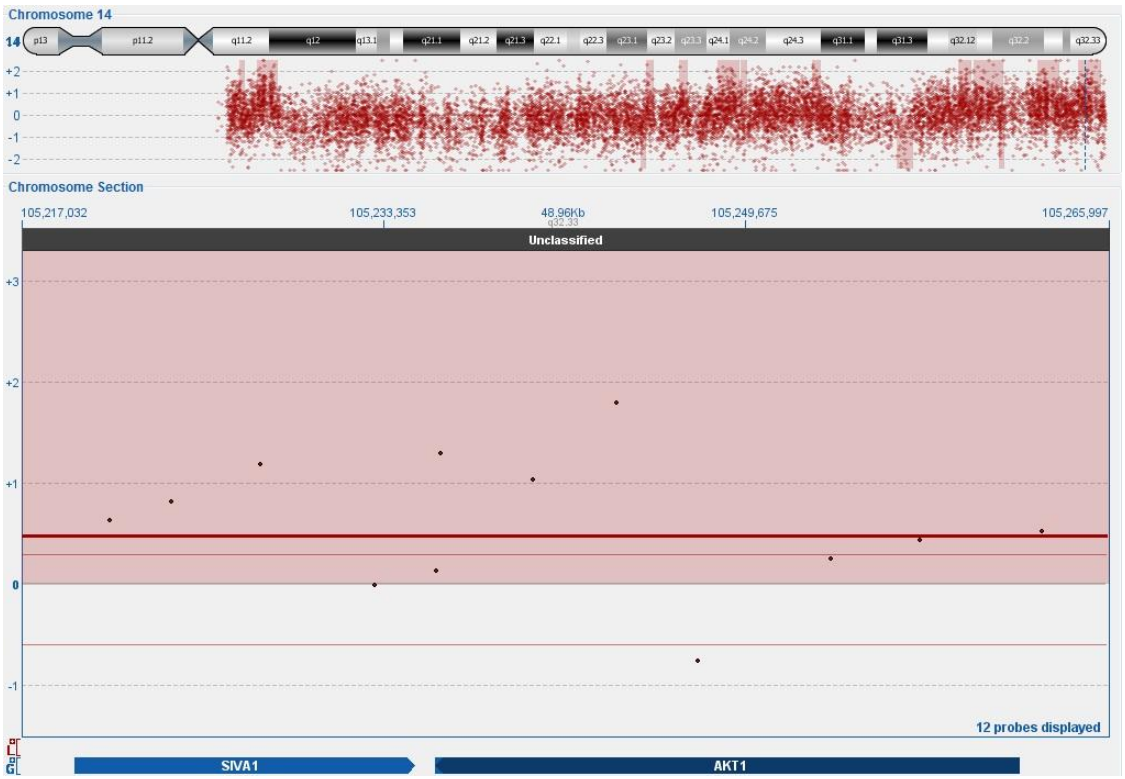
Characteristics	NAFLD, *n* = 144 (mean ± SD)			
	AA (*n* = 50)	AG (*n* = 62)	GG (*n* = 32)	*p* value
Age (years)	52.0 ± 12.6	49.5 ± 12.4	52.8 ± 10.4	0.352
BMI (kg/m^2)	27.9 ± 3.7	29.1 ± 4.6	29.0 ± 4.8	0.343
HbA1c (%)*	6.6 ± 1.9	6.5 ± 1.6	6.6 ± 1.5	0.688
Waist circumference (cm)	93.3 ± 10.3	93.9 ± 11.5	94.6 ± 10.7	0.869
HDL cholesterol (mg/dl)	48.0 ± 12.6	48.1 ± 13.5	50.1 ± 11.5	0.718
LDL cholesterol (mg/dl)	120.5 ± 41.1	118.8 ± 40.4	108.8 ± 37.8	0.399
Total cholesterol (mg/dl)	198.9 ± 45.3	199.0 ± 44.7	188.8 ± 40.8	0.518
Triglycerides (mg/dl)	148.8 ± 55.1	162.4 ± 69.6	150.4 ± 60.0	0.466
AST (IU/L)*	44.2 ± 27.9	44.2 ± 24.3	38.3 ± 23.4	0.452
ALT (IU/L)	85.7 ± 50.8	88.0 ± 52.7	69.2 ± 32.7	0.183
GGT (IU/L)*	87.4 ± 74.3	129.3 ± 138.8	115.1 ± 115.6	0.659
Systolic BP (mmHg)	131.5 ± 15.8	130.2 ± 14.3	134.5 ± 15.6	0.452
Diastolic BP (mmHg)*	82.4 ± 10.5	82.0 ± 9.5	83.8 ± 9.9	0.803
Steatosis grade*	1.9 ± 0.8	1.9 ± 0.8	1.7 ± 0.7	0.472
Lobular inflammation*	1.3 ± 0.6	1.3 ± 0.6	1.2 ± 0.6	0.651
Ballooning*	1.2 ± 0.6	1.3 ± 0.6	1.2 ± 0.6	0.387
Fibrosis*	1.7 ± 1.0	1.7 ± 0.9	1.86 ± 1.1	0.784

**P* values obtained using Kruskal-Wallis test, all other comparisons used ANOVA

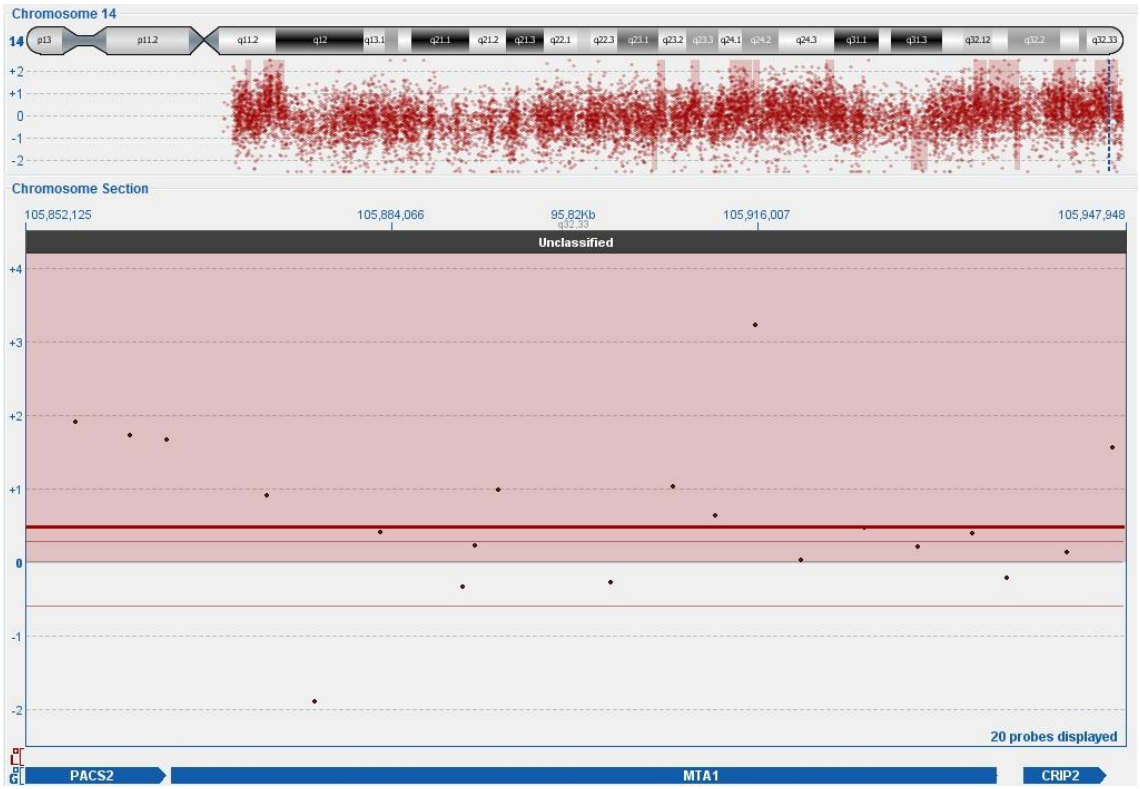
ALT alanine transferase, *AST* aspartate aminotransferase, *BMI* body mass index, *GGT* gamma glutamyl transpeptidase, *HbA1c* haemoglobin A1c, *HDL* high-density lipoprotein, *LDL* low-density lipoprotein, *NAFLD* non-alcoholic fatty liver disease

APPENDIX G

I) Amplified chromosome 14q32.33 region encompassing AKT1 gene



II) Amplified chromosome 14q32.33 region encompassing MTA1 gene



APPENDIX H

I) PUBLICATIONS

1. Zain SM, Mohamed R, Mahadeva S, Cheah PL, Rampal S, Basu RC, Mohamed Z. A multi-ethnic study of a PNPLA3 gene variant and its association with disease severity in non-alcoholic fatty liver disease. *Hum Genet*. 2012;131:1145-1152 (Tier 1, impact factor 4.633)

2. Zain SM, Mohamed Z, Mahadeva S, Cheah PL, Rampal S, Chin KF,Basu RC, Tan HL, Mohamed R. Impact of leptin receptor gene variants on risk of non-alcoholic fatty liver disease and its interaction with adiponutrin gene. *J Gastroenterol Hepatol* 2013: 28:873-9 (Tier 2, Impact Factor 3.325)

3. Zain SM, Mohamed Z, Mahadeva S, Rampal S, Basu RC, et al. (2013) Susceptibility and gene interaction study of the angiotensin II type 1 receptor (AGTR1) gene polymorphisms with non-alcoholic fatty liver disease in a multi-ethnic population. *PLOS ONE* 2013;8(3):e58538 (Tier 1, Impact Factor 3.73)

4. Tan HL*, Zain SM*, Mohamed R, Rampal S, Chin KF, Basu RC, Cheah PL, Mahadeva S, Mohamed Z . Association of glucokinase regulatory gene polymorphisms with risk and severity of non-alcoholic fatty liver disease: An interaction study with adiponutrin gene. *J Gastroenterol* 2013. DOI 10.1007/s00535-013-0850 (Tier 1, Impact Factor 3.788)

H. L. Tan and S. M. Zain contributed equally to this project and should be considered co-first authors

II) PROCEEDINGS

1. Shamsul Mohd Zain, Zahurin Mohamed, Sanjiv Mahadeva, Sanjay Rampal, Roma Choudhury Basu, Agus Salim, Rosmawati Mohamed. Identification of novel copy number variation biomarkers for non-alcoholic fatty liver disease spectrum. Malaysian Society of Pharmacology and Physiology (MSPP) Conference 2012, Penang, Malaysia. National (Poster-Award).

2. Shamsul Mohd Zain, Zahurin Mohamed, Sanjiv Mahadeva, Sanjay Rampal, Roma Choudhury Basu, Agus Salim, Rosmawati Mohamed. Susceptibility and gene interaction study of the angiotensin ii type 1 receptor gene with non-alcoholic fatty liver disease in a multi-ethnic population. Malaysian Society of Pharmacology and Physiology (MSPP) Conference 2012, Penang, Malaysia. National (Oral).

3. SM Zain, Z Mohamed, S Mahadeva, PL Cheah, RC Basu, R Mohamed. Does polymorphism in patatin-like phospholipase 3 gene (PNPLA3) confer risk of nonalcoholic fatty liver disease (NAFLD) in multiethnic Malaysian population? 4th Asia Pacific ISSX Meeting 2011, Tainan, Taiwan. International (Poster).

4. SM Zain, R Mohamed, S Mahadeva, PL Cheah, RC Basu, Z Mohamed. Angiotensin II type 1 receptor (AGTR1) polymorphisms and the risk of non-alcoholic steatohepatitis in a Malaysian population. The Ninth JSH Single Topic Conference 2010, Tokyo, Japan. International (Poster-Award).

III) Poster awards

Ninth JSH Single Topic Conference

"NASH 2010"

November 18th – 19th 2010 Tokyo Japan

Travel Award

Abstract No.: 10050

Dr. Rosmawati Mohamed

Institution: Faculty of Medicine, University of Malaya, Kuala Lumpur

Title: Angiotensin II type 1 receptor (AGTR1) polymorphisms and the risk of non-alcoholic steatohepatitis in a Malaysian population

This is to certify that the organizing committee hereby recognizes the excellence of your paper in the poster presentation for the Ninth JSH Single Topic Conference, held on November 18th – 19th 2010, at Tokyo, Japan.

Please accept our congratulations and appreciation for your efforts to make this conference a success.

Sumio Watanabe
Chairman of
The Organizing Committee

Kazuhiko Koike
Director General
The Japan Society of Hepatology

Poster awards (continued)

26th Scientific Meeting of Malaysian Society of Pharmacology & Physiology

"Translational research for sustainable wellness: From perception to reality"

18th – 20th May 2012
Equatorial Hotel,
Penang, Malaysia

UNIVERSITI SAINS MALAYSIA

Certificate of

The Best

Poster Presentation

3rd Prize

Shamsul Mohd. Zain

26th Scientific Meeting of Malaysian Society of Pharmacology & Physiology 2012

Prof. Munavvar Zubaid Abdul Sattar
Organising Chairman