MOLECULAR BIOMARKERS AND GENETIC DETERMINANTS OF NON-ALCOHOLIC FATTY LIVER DISEASE

TAN HWA LI

FACULTY OF MEDICINE

UNIVERSITY OF MALAYA

KUALA LUMPUR

2019

MOLECULAR BIOMARKERS AND GENETIC DETERMINANTS OF NON-ALCOHOLIC FATTY LIVER DISEASE

TAN HWA LI

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

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

2019

UNIVERSITY OF MALAYA

ORIGINAL LITERARY WORK DECLARATION

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

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'sSignature Date:

Name:

Designation

MOLECULAR BIOMARKERS AND GENETIC DETERMINANTS OF NON-ALCOHOLIC FATTY LIVER DISEASE

ABSTRACT

Non-alcoholic fatty liver disease (NAFLD) is a chronic liver disease encompassing non-alcoholic fatty liver (NAFL) and non-alcoholic steatohepatitis (NASH). NAFL is relatively benign but NASH has a higher rate of progression to develop into liver cirrhosis and hepatocellular carcinoma. Stratification of NASH from NAFL is important for disease monitoring and medical interventions. Genetic variations such as single nucleotide polymorphisms (SNPs) and copy number variations (CNVs) contribute to the phenotypic variation and rate of disease progression. MicroRNAs (miRNAs) are small non-coding, stably expressed RNAs in the biofluids that regulate gene expression post-transcriptionally affecting disease phenotype. This study aimed to investigate the potential genetic variants and biomarkers involved in the development of NAFLD with focus to NASH. A total of 232 healthy controls and 249 biopsy-proven NAFLD patients were recruited. Quantitative polymerase chain reaction (qPCR) was used to genotype the SNPs on the glucokinase regulatory gene (*GCKR*) (rs1260326 and rs780094) and CNV 13q12.11. Sequence-specific primer (SSP) technique was applied for HLA typing. Differential expression of miRNAs was screened and validated in the healthy controls, NAFL and NASH patients. The diagnostic values of miRNAs were assessed using area under receiver operating characteristic (AUROC) curves. Both *GCKR* rs1260326 and rs780094 were associated with susceptibility to NASH (OR 1.55, 95 % CI 1.10–2.17, P = 0.013; and OR 1.56, 95 % CI 1.10–2.20, P = 0.012, respectively) and NASH with significant fibrosis (fibrosis score ≥ 2) (OR 1.50, 95 % CI 1.01–2.21, P = 0.044; and OR 1.52, 95 % CI 1.03–2.26, P = 0.038, respectively). Histological data showed a significant association of GCKR rs1260326 with higher

steatosis grade (OR 1.76, 95 % CI 1.08–2.85, *P* = 0.04). The CNV gain in the 13q12.11 was significantly associated with a greater risk of NAFLD (OR 2.22, 95% CI 1.42–3.46, *P* = 0.0004) and NASH (OR 2.33, 95% CI 1.47–3.68, *P* = 0.0003). HLA-DQB1*06 allele group was significantly associated with reduced risk of NASH with advanced fibrosis compared to NASH without advanced fibrosis (OR 0.272, 95% CI 0.14-0.55, *P* < 0.001). The DQB1*06 allele group was also significantly associated with lower risk of lobular inflammation (OR 0.84, 95% CI 0.39-0.97, *P* = 0.016) and a lower risk of hepatic fibrosis in NAFLD patients (OR 0.30, 95% CI 0.14-0.67, *P*< 0.001 suggesting its protective value in conferring a milder form of the disease. Increased expression of miR-122, miR-193, miR-192, miR-34a and miR-125b were demonstrated in NASH (*P* <0.001, *P* = 0.001, *P* = 0.015, *P* = 0.021 and *P* = 0.042, respectively). The miR-122 has the best diagnostic accuracy for NASH (AUROC = 0.73) and combination of miR-122 and AST gave the best prediction for NASH from NAFL (AUROC = 0.785). The miR-122, miR-193band miR-125b were significantly upregulated in significant fibrosis (P < 0.0001, *P* = 0.001 and *P* = 0.003, respectively) with miR-122 having the highest AUROC of 0.736. This study showed the potential of SNPs and CNVs in predicting the risk of NASH and the role of miRNAs as molecular biomarkers of NASH.

Keywords: NAFLD, NASH, genetic variations, microRNA, interaction of HLA

PENANDA BIOLOGI MOLEKUL DAN PENENTU GENETIK BAGI PENYAKIT HATI BERLEMAK

ABSTRAK

Penyakit hati berlemak tanpa kehadiran alkohol (NAFLD) ialah penyakit kronik yang merangkumi steatosis biasa (NAFL) dan radang hati tanpa kehadiran alkohol (NASH). NAFLsecara relatifnya adalah benigna tetapi NASH berpotensi berkembang menjadi sirosis hati dan barah hati. Stratifikasi NASH daripada NAFL adalah penting untuk memantau penyakit dan kawalan perubatan. Variasi-variasi genetik seperti polimorfisme nukleotid tunggal (SNPs), salinan variasi nombor (CNVs) dan polimorfisme antigen leukosit manusia (HLA) menyumbang kepada variasi fenotaip dan kadar perkembangan penyakit. MikroRNAs (miRNAs) merupakan RNA kecil yang tidak mengekod, diekspres stabil di dalam cecair biologi, yang mengawal ekspresi gen pasca-transkripsi, dengan itu mempengaruhi fenotaip penyakit. Kajian ini bertujuan untuk mengkaji potensi variasi-variasi genetik dan penanda-penanda biologi yang terlibat dalam perkembangan NAFLD, tumpuan diberikan kepada NASH. Seramai 232 subjek kawalan yang sihat dan 249 pesakit NAFLD yang disahkan melalui biopsi hati direkrut. Kaedah kuantitatif reaksi rantai*polymerase (qPCR)* digunakan untuk menggenotaip SNPs pada gen *glucokinase regulatory(GCKR)* (rs1260326 dan rs780094) dan CNV 13q12.11. Primer urutan specifik (SSP) telah digunakan untuk penjenisan HLA. Perbezaan ekspresi miRNAs telah disaring dan disahkan di kalangan subjek kawalan yang sihat, pesakit NAFL dan NASH. Nilai diagnostik miRNAs dinilai menggunakan graf *AUROC*. Kedua-dua *GCKR* rs1260326 dan rs780094 dikaitkan dengan kecenderungan terhadap NASH (OR 1.55, 95% CI 1.10-2.17, P = 0.013; dan OR 1.56, 95% CI 1.10-2.20, P = 0.012, masing-masing) dan NASH dengan fibrosis (OR 1.50, 95% CI 1.01-2.21, P = 0.044; dan OR 1.52, 95% CI 1.03-2.26, P = 0.038, masing-masing). Data histologi menunjukkan perkaitan signifikan antara

*GCKR*rs1260326 dengan steatosis bergred tinggi (OR 1.76, 95% CI 1.08-2.85, *P* = 0.04). Duplikasi CNV dalam 13q12.11 dikaitkan secara signifikan dengan risiko NAFLD (OR 2.22, 95% CI 1.42-3.46, *P* = 0.0004) dan NASH (OR 2.33, 95% CI 1.47-3.68, *P* = 0.0003). Alel HLA-DQB1*06 dikaitkan secara signifikan dengan risiko rendah NASH dengan fibrosis skor ≥ 3 berbanding NASH dengan fibrosis skor <3 (OR 0.04, 95% CI 0.003-0.40, *P* = 0.015). Alel DQB1*06 juga dikaitkan secara signifikan dengan risiko rendah radang dalam hati (OR 0.272, 95% CI 0.14-0.55, *P* < 0.001) dan fibrosis (OR 0.30, 95% CI 0.14-0.67, *P*< 0.001), menunjukkannilai ramalannya untukpenyakit hati berlemak yang kurang serius. Kenaikan regulasi miR-122, miR-193, miR-192, miR-34a, dan miR-125b telah dipamerkan dalam NASH (P < 0.001, *P* = 0.001, *P* = 0.015, *P* = 0.021 dan *P* = 0.042, masing-masing). MiR-122 mempunyai ketepatan diagnostik yang terbaik untuk NASH (AUROC = 0.73) dan gabungan miR-122 dan AST memberikan ramalan yang terbaik untuk NASH berbandingNAFL (AUROC = 0.785). Di kalangan pesakit yang mempunyai fibrosis signifikan, miR-122, miR-193b dan miR-125b mengalami kenaikan regulasi secara signifikan (P < 0.0001, *P* = 0.001 dan*P* = 0.003, masing-masing) dengan miR-122 mempunyai AUROC tertinggi iaitu 0.736. Kajian ini menunjukkan potensi SNPs dan CNVs dalam meramal risiko NASH dan peranan miRNAs sebagai penanda biologi NASH.

Keywords: NAFLD, NASH, variasi genetic, microRNA, interaksi dengan HLA

ACKNOWLEDGEMENTS

Firstly, I wish to extend my deepest gratitude to my supervisors. Professor Zahurin Mohamed has taken me as a student since my internship and has always given me inspiring ideas and provided chances for me to learn and acquire new skills and techniques during my postgraduate programme. I am deeply thankful to Professor Rosmawati Mohamed for her valuable clinical advice and critical review on my work from the clinical standpoint, allowing me to widen my research from various perspectives. I am also highly indebted to Dr. Shamsul Mohd Zain for his constant support, encouragement and guidance especially in data analyses and the proofreading of my works. This thesis would not be completed without their continuous motivation, guidance and support, both academically and mentally. I feel grateful to have the chance to work under their guidance to conduct such an interesting study.

Besides, I wish to express my sincere gratitude to the co-investigators and collaborators, Professor Sanjiv Mahadeva, Professor Chan Wah Kheong, Dr. Nik Raihan Nik Mustapha, Dr. Eng Hooi Sian, Dr. Sanjay Rampal, Professor Chan Kin Fah, and Mrs Manjit Kaur for all the guidance, help and support given during patients recruitment, histopathological assessment, and data analysis. Sincere thanks also goes to the University Malaya Health and Translational Medicine Centre (UMHTMC), Institute of Medical Research (IMR) and High Impact Research (HIR) Facility for lending their equipment for my laboratory works.

I truly appreciate the help from volunteers and participants who have given blood samples and especially the NAFLD patients. Special thanks also go to the Ministry of Education for awarding me the scholarship and also the University Malaya for the grants support.

I also want to extend my greatest thanks to my parents for their unconditional support, kindness and patience throughout my studies. I am also grateful to have my fiancé by my side, supporting me through ups and downs. I would also like to thank my colleagues in the Pharmacology Department for their companion and friendship that provided me with a pleasant learning and working environment.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

%	percentage
^{o}C	degree celcius
±	plus minus
µg	microgram
µL	microliter
ng	nanogram
χ2	chi square
AASLD	American Association for the Study of Liver Diseases
aCGH	array comparative genomic hybridization
AGTR1	angiotensin II type 1 receptor
ALT	alanine aminotransferase
APOC3	apolipoprotein C3
APRI	aspartate aminotransferase-to-platelet ratio index
Argo	Argonaute
AST	aspartate aminotransferase
AUROC	area under receiver operating characteristic
BARD	BMI, AST/ALT ratio, Diabetes
BMI	body mass index
bp	base pair
CHB	Han Chinese from Beijing
CI	confidence interval
CK18	cytokeratin 18
cm	centimeter
CNV	copy number variation

CRN	Clinical Research Network
CT	computed tomography
DNA	deoxyribonucleic acid
DGV	Database of Genomic Variants
EASD	European Association for the Study of Diabetes
EASL	European Association for the Study of the Liver
EASO	European Association for the Study of Obesity
EDRN	Early Detection Research Network
ELF	enhanced liver fibrosis
Exp5	exportin-5
F	fibrosis stage
FDFT1	farnesyl diphosphate farnesyl transferase 1
FDR	false discovery rate
FIB-4	fibrosis-4
FPR	fibrosis progression rate
GCK	glucokinase
GCKR	glucokinase regulatory
GCKRP	glucokinase regulatory protein
GGT	gamma-glutamyl transpeptidase
GMDR	Generalized Multifactor Dimensionality Reduction
GOLD	Genetics of Obesity-related Liver Disease
GWAS	genome-wide association study
HA	hyaluronic acid
HbA1c	hemoglobin A1c
HCC	hepatocellular carcinoma
HDL	high density lipoprotein

HLA	human leukocytes antigen
HOMA-IR	Homeostasis Model of Assessment – Insulin Resistance
HSC	hepatic stellate cells
IL-1β	interleukin-1β
IL-6	interleukin-6
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
LEPR	leptin receptor
LEPR	leptin receptor
LPIN1	lipin 1
LNA	locked nucleic acid
LYPLAL1	lysophospholipase-like 1
MAF	minor allele frequency
Mb	mega base
MFI	mean fluorescence intensity
mg/dL	milligram per deciliter
MRE	magnetic resonance elastography
mRNA	messenger RNA
MHC	major histocompatibility complex
min	minute
miR	microRNA
miRNA	microRNA

mL	mililiter
MRI	magnetic resonance imaging
NAFL	non-alcoholic fatty liver
NAFLD	non-alcoholic fatty liver disease
NAS	NASH Activity Score
NASH	non-alcoholic steatohepatitis
NCAN	neuroocan
NFS	NAFLD fibrosis score
NGS	next generation sequencing
NICE	National Institute of Clinical Excellence
NKT	natural killer T
NPV	negative predictive value
nm	nanometer
OELF	Original European Liver Fibrosis
OR	odds ratio
PCR	polymerase chain reaction
PEMT	phosphatidylethanolamine N-methyl transferase
PIIINP	procollagen type III amino-terminal peptide
PNPLA3	patatin-like phospholipase domain-containing protein 3
PPP1R3B	protein phosphatase1 regulatory inhibitor subunit 3B
PPV	positive predictive value
qRT-PCR	quantitative reverse transcription real-time PCR
RISC	RNA-induced silencing complex
RNA	ribonucleic acid
ROC	receiver under operating
rpm	revolutions per minute

s	second
SAF	Steatosis, Activity, Fibrosis
SD	standard deviation
SNP	single nucleotide polymorphism
SOD2	superoxide dismutase 2
TIMP1	tissue inhibitor of metalloproteinase I
TM6SF2	transmembrane 6 superfamily member 2
TNFα	tumor necrosis factor α
UK	United Kingdom
UMMC	University Malaya Medical Centre
US	ultrasonography
USA	United States of America
uv	ultraviolet
VCTE	vibration controlled transient elastography
VLDL	Very low density lipoprotein
XPO-4	exportin-4

LIST OF APPENDICES

CHAPTER 1: INTRODUCTION

1.1 Severity of NAFLD

Non-alcoholic fatty liver disease (NAFLD) has emerged as one of the most common chronic liver diseases worldwide, parallel to the increasing prevalence of obesity, type II diabetes, insulin resistance and hypertension(Younossi et al., 2011)(Younossi et al., 2011).NAFLD is characterised as the presence of steatosis or fat in individuals without significant alcohol consumption, encompassing a spectrum of liver disease spanning from NAFL (non-alcoholic fatty liver) to non-alcoholic steatohepatitis (NASH) (Chalasani et al., 2012). WhilstNAFL is defined as hepatic steatosis involving >5 % liver parenchyma without hepatocyte injury, the histological characterisation of NASH includes evidence of hepatocellular ballooning degeneration, inflammation and various stages of fibrosis (Chalasani et al., 2012).

With the current trends of high-fat diet and preponderance of sedentary lifestyle, it is of no surprise that NAFLD has been a global epidemic, affecting 25% global population (Younossi et al., 2016). It is also projected to be the leading cause of liver morbidity and mortality and the leading indication for liver transplantation in 20 years, posing as a great economic burden (Bertot & Adams, 2016). Despite the alarming fact that NAFLD patients are at higher risk of mortality, under-recording and missed diagnosis are very common in primary care, resulting in delayed intervention and treatment (Alexander et al., 2018).

Since NAFLD patients are usually asymptomatic, they are often discovered incidentally following abnormal liver enzymes. Imaging studies such as ultrasound, computed tomography (CT) and magnetic resonance imaging (MRI) have proved to be important modalities in diagnosing NAFLD but CT and MRI are not performed on a routine basis due to their high cost and inability to distinguish between various

NAFLDhistological features (Fitzpatrick & Dhawan, 2014). The definite diagnosis of NAFLD is based upon liver biopsy followed by the histopathological examination.

By being able to identify the histological features in NAFLD patients, clinicians are able to predict the prognosis and long term outcome of patients. Early studies suggested that compared to NAFL, NASH has an increased risk of progression to cirrhosis and hepatocellular carcinoma, thus, having a higher rate of mortality (Angulo, 2010; Ascha et al., 2010; Bugianesi et al., 2002). There is compelling evidence that patients with NAFL may still develop NASH and accompany with fibrosis progression (McPherson et al., 2015; Wong, Wong, et al., 2010). The rate of disease progression also varies with the underlying fibrosis stage, where patients with advanced fibrosis have a higher risk of progression compared to those without fibrosis. Patients with fibrosis stage 2 and above are categorised as significant fibrosis, as they are at higher risk for liver-related mortality (Thomas, 2017; von Roenn, 2018). Advanced fibrosis is used to describe fibrosis stage 3 and above, which is of most concern for clinicians as patients with advanced fibrosis are the one most in need of clinical surveillance, including for hepatocellular carcinoma(Kim, Kim, Talwalkar, Kim, & Ehman, 2013).

Stratification of NAFLD patients according to their histological findings will allow clinician to predict their risk of progression, thus easing decision making and patient selection for monitoring progression. The cost of patient management in secondary and tertiary care is overwhelming (Prinja, Bahuguna, Duseja, Kaur, & Chawla, 2018; Sattar, Forrest, & Preiss, 2014; Whalley, Puvanachandra, Desai, & Kennedy, 2007). Therefore selection of patients at risk of progressive liver disease is needed prior to referral. NAFLD patients are often referred for secondary care interventions involving different therapies such as insulin-sensitising agents (Bril &Cusi,

2017). It is also imperative to distinguish stage of fibrosis in patients in order to provide risk stratification and intervention in impeding disease progression in the NASH patients. Patients with evidence of cirrhosis will also require long term follow-up and entry into surveillance programs for detection of gastroesophageal varices and hepatic decompensation (Dyson, Anstee, & McPherson, 2014). Some patients will also be recruited in intense weight reduction program, clinical trials of drug therapy or be advised to undergo bariatric surgery (McCarty, Echouffo-Tcheugui, Lange, Haque, & Njei, 2018).

1.2 Non-invasive molecular biomarkers of NAFLD, NASH and advanced fibrosis

Repeated investigation of disease severity is required as part of the monitoring program in addition to determining the effectiveness of treatment trials. Albeit being the gold standard, liver biopsy is not a favourable modality in assessing the progression of the disease, due to its high cost and invasiveness. Liver biopsy is also limited by sampling error, procedure-related morbidity and observer variation (Rockey et al., 2009). A non-invasive test that diagnoses NASH or advanced fibrosis would be clinically beneficial to reduce the need for liver biopsy.

There has been an increasing recognition of microRNAs (miRNAs) as biomarkers due to their relatively stable expression in various biofluids (Peltier & Latham, 2008). MiRNAs are small non-coding RNAs which regulate gene expression post-transcriptionally thereby modulating various biological processes such as cellular development, differentiation, proliferation, apoptosis, and metabolism. MiRNAs are also found to be indicators of hepatocyte injury and inflammation, in modulating diverse spectrum of liver disease (Bala et al., 2012). The miRNAs expression profile in

NAFLD will be investigated in this study along with their differential expression in various stages of NAFLD.

In an attempt to identify advanced fibrosis, several clinical scoring systems have been validated, such as aspartate aminotransferase-to-platelet ratio index (APRI), aspartate aminotransferase/alanine aminotransferase (AST/ALT) ratio, BARD score, FIB-4 score, and NAFLD fibrosis score (Guha et al., 2008; Kim, Kim, Talwalkar, et al., 2013; McPherson, Stewart, Henderson, Burt, & Day, 2010; Xun et al., 2012). These algorithms are based on clinical or laboratory indices and therefore, are easily attainable.To complement the search of biomarkers for advanced fibrosis, the applicability of these scoring systems will also be investigated in this study.

1.3 Genetic variations as predictive risk factors

Evidence from familial aggregation, twin studies and interethnic differences in susceptibility has proven that interaction of both environmental and genetic factors contribute to the phenotypic expression of NAFLD (Anstee & Day, 2013). Studies on the field of genetics and epigenetics may shed light on the factors predisposing the development of NAFLD. It may also lead to the development of non-invasive biomarkers for early diagnosis of NAFLD, allowing early preventive and therapeutic strategies for those at high risk.

Inter-individual variation with regards to the severity of NAFLD and the risk of morbidity and mortality may be attributed to a combination of genetic and environmental factors. Discovery of genetic risk factors could prove useful for the clinical management of NAFLD patients, especially in allowing risk stratification of patients for efficient monitoring and surveillance programmes. It will also provide insights on the molecular mechanisms underlying disease progression, to suggest therapeutic targets as well as to

guide clinicians towards individualised therapy to improve clinical outcome. In this study, genetic variants including single nucleotide polymorphisms (SNPs), copy number variations (CNVs) and human leukocyte antigen (HLA) polymorphisms will be investigated in search of the genetic modifiers of NAFLD.

Genotyping of SNPs is highly recognised as a major tool for identifying disease susceptibility loci. There are over 1.4 million SNPs that have been detected across the genome, in coding, non-coding and untranslated regions (Sachidanandam et al., 2001). Following various genome-wide association studies (GWAS) and candidate association studies, various SNPs are found to be associated with phenotypic variability and disease susceptibility to NAFLD. Candidate genes for NAFLD study included genes influencing insulin resistance, fatty acid metabolisms, oxidative stress, fibrogenesis, endotoxin response, and cytokine activity (Duvnjak, Barsic, Tomasic, & Lerotic, 2009). The association of rs780094 and rs1260326 variants in glucokinase regulatory (*GCKR*) gene will be evaluated in this study due to its role in regulating glucokinase activity and controlling glucose metabolism.

CNVs are DNA copy-number alterations spanning from 1 kilo base to 1 mega base in length. They constitute a substantial fraction of genetic and phenotypic variability involving both duplications (copy number gains) and deletions (copy number losses) of DNA(Hegele, 2007). Sequencing of human genome found that 4.8-9.5% of genome contributes to CNVs (Zarrei, MacDonald, Merico, & Scherer, 2015) and approximately 85-95% of CNVs have been implicated in a corresponding change in gene expression in human and mice (Henrichsen et al., 2009; Stranger et al., 2007). CNV influences gene expression through disruption or duplication of CNVs regions-containing genes, where the over- or under-expression could then lead to phenotypic changes (Tang & Amon, 2013). Previous study in a small cohort identified several genomic regions that are potentially relevant to NASH(Zain et al., 2014). The

role of CNVR 13q12.11 in the susceptibility of NAFLD will be further assessed in the present study involving a larger set of samples.

HLA, the human version of the major histocompatibility complex (*MHC*) genes, is a large genetic region which encompasses over 200 highly polymorphic genes that encode for class I and class II antigens (Jin & Wang, 2003). This study investigates the potentials of HLA-DQA1 and HLA-DQB1 which are the two adjacent loci found on chromosome 6p21.3. These two loci encode the DQ α subunit and DQ β subunit respectively, forming a αβ heterodimer known as HLA-DQ, a component of the MHC class II antigens (Williams, 2001). Despite the extensive studies on the role of MHC II-mediated pathways in the immune processing, the involvement of HLA in hepatic inflammation and fibrosis in NAFLD has not been investigated. The plausible association was derived due to the central role of immune and inflammatory pathways in the pathogenesis of NAFLD. Expression of MHC II molecules on the surface of antigen-presenting cells would promote CD4+ helper T cell recognition. This subsequently triggers immune response including the production of inflammatory cytokines. These cytokines are relevant to the pathogenesis of NAFLD (Blum, Wearsch, & Cresswell, 2013; Mitchison, 2004).

1.4 Objectives

Assessment of suitable diagnostic biomarker for NAFLD

1) To investigate the serum miRNAs expression profiles in NAFLD including the differential expression in NAFL and NASH.
2) To explore the differential expression of miRNAs in various histological severity of NAFLD, namely steatosis, lobular inflammation, hepatocellular ballooning and fibrosis.
3) To validate the diagnostic performance of AST/ALT, BARD score, FIB-4 score, NAFLD fibrosis score and ELF in identification of advanced fibrosis.

Identification of genetic predictive risk factors of NAFLD

4) To investigate the association between two previously reported SNPs of the *GCKR* gene with the susceptibility to NAFLD and histological features.
5) To evaluate the association of CNV 13q12.11 with susceptibility to NAFLD and its histological features.
6) To determine the association of HLA-DQA1 and -DQB1 genes with the susceptibility to NAFLD and its histological features.

CHAPTER 2: LITERATURE REVIEW

2.1 Definition of NAFLD

Nonalcoholic fatty liver disease (NAFLD) is characterised by the presence of hepatic steatosis by imaging or by histology without secondary causes such as significant alcohol consumption (daily alcohol intake of <20g for women and <30 g for men), steatogenic medication, viral hepatitis and hereditary disorders (Chalasani et al., 2018). NAFLD is histologically categorised into NAFL and nonalcoholic steatohepatitis (NASH) where the former is the presence of hepatic steatosis alone without evidence of hepatocellular injury and the latter is the presence of inflammation and/or hepatocyte ballooning in addition to hepatic steatosis, with or without fibrosis (Benedict & Zhang, 2017).

2.2 Incidence and Prevalence of NAFLD

Limited data are available for the incidence of NAFLD. A study in Japanese population found that there are 308 (10%) new NAFLD cases among 3147 participants over a mean follow up period of 414 days (Hamaguchi et al., 2005). In England however, a retrospective study showed a lower incidence of 29 annual incidences per 100,000 population (Whalley et al., 2007). A longitudinal cohort study conducted in 77,425 individual without metabolic abnormalities at baseline found that 10,340 participants developed NAFLD, translating to an incidence rate of 29.7 per 1000 person per year (Chang et al., 2016). A recent meta-analysis estimated the pooled regional NAFLD incidence rate is 52.3 per 1000 person-year in Asia and 28.0 per 1000 person-year in Israel (Younossi et al., 2016).

Prevalence of NAFLD varies across different ethnicity and geographic groups. It is estimated that the global prevalence of NAFLD diagnosed by imaging is 25.24%, with the highest prevalence found in the Middle East (32%) and South America (31%), followed by Asia (27%), North America (24%), Europe (24%) and the lowest rate is reported from Africa (14%) (Younossi et al., 2016). The estimated prevalence of NAFLD in Asia ranges from 15% to 44%, with the global mean of 27% while biopsy-proven NASH prevalence estimates 64% among NAFLD patients(Younossi et al., 2016). Research carried out by National Health and Nutrition Examination Surveys in USA found that prevalence rate for chronic liver disease has increased from 11.78% (in 1988 to 1994) to 14.78% (in 2005 to 2008) (Younossi et al., 2011). While the prevalence of hepatitis B, hepatitis C and alcoholic liver disease remain generally stable, the prevalence of NAFLD had been increasing. NAFLD accounted for 46.8 % of chronic liver disease (in the year 1988 to 1944) but this figure had increased to an alarming proportion of 75.1% in the USA (in the year 2005 to 2008) (Younossi et al., 2011). Pooled database of existing primary care records from European territories also found trebling of prevalence rate of NAFLD from 2007 to 2014 (Alexander et al., 2018).

2.3 Risk factors

Being a complex disease, the development and progression of NAFLD are determined by the interaction of genetic modifiers and various environmental factors such as dietary pattern, lifestyle and intestinal microbiota (Anstee, Seth, & Day, 2016). Studies pointed out that obesity, type 2 diabetes, hypertension and dyslipidemia are the major risk factors for the development and progression of NAFLD (Souza, Diniz Mde, Medeiros-Filho, & Araujo, 2012).

2.3.1 Metabolic syndrome

High BMI and visceral obesity are well-recognised risk factors for NAFLD. In an earlier review paper, the prevalence of NAFLD was 24.5% in normal weight population, but 67% and 94% in overweight and obese patients, respectively (Bellentani, Bedogni, Miglioli, & Tiribelli, 2004). Most patients with NAFLD have higher abdominal obesity or BMI than those without NAFLD (Jakobsen, Berentzen, Sorensen, & Overvad, 2007).The prevalence of NAFLD ranging from 65% to 80% in obese patients compared to 16% in subjects with normal BMI without metabolic syndrome (Bellentani et al., 2000; Fabbrini, Sullivan, & Klein, 2010; Williams et al., 2011). The prevalence of NASH is also higher in individuals with morbid obesity (40%) and obesity (20%) compared to non-obese individuals (3%) (Fabbrini et al., 2010). Among morbidly obese patients who undergo bariatric surgery, the prevalence of NAFLD reaches up to 90.4% while NASH was found in 70% of morbidly obese patients where 43% have fibrosis and 2% have cirrhosis (Barros, Setubal, Martinho, Ferraz, & Gaudencio, 2016; Losekann et al., 2015). Obesity is also a risk factor for insulin resistance, thereby increasing the risk of NAFLD (Diehl, 2005).

Patients with type II diabetes mellitus are more prone to NAFLD with the prevalence ranging from 29.6% to 87.1% (Dai et al., 2017). This prevalence rate varies with the sample characteristics, such as obesity, hypertension and dyslipidemia. Meanwhile, it was estimated that approximately 20%-80% of NAFLD patients have dyslipidemia. Long term dyslipidemia altered the lipoprotein synthesis, thereby increasing triglycerides, low-density lipoprotein and very low-density lipoprotein (Zhang & Lu, 2015).

2.3.2 Age and gender

There is an increasing prevalence of NAFLD as age increases, where individuals younger than 20 years old has a prevalence of 20% as opposed to 40% in individuals who are older than 60 years old (Brea & Puzo, 2013). A prospective study also found that 42 (46%) out of 91 octogenarians (mean age of 85.6) have NAFLD (Kagansky et al., 2004). The incidence of NASH and cirrhosis are also greater in those of above 50 years old compared to younger age group (Sayiner, Koenig, Henry, & Younossi, 2016). Some studies also found that NASH patients who are older and with advanced fibrosis are more at risk of progressing to hepatocellular carcinoma (Ascha et al., 2010; Hashimoto et al., 2009).

2.3.3 Ethnic differences

Ethnicity also contributes to the differential prevalence of NAFLD, where highest prevalence of NAFLD was found in Hispanics (58.3%), followed by Caucasians (44.4%) and African Americans (35.1%) (Williams et al., 2011). A study in the United States discovered a higher degree of hepatocellular ballooning in Asians and higher rate of Mallory bodies in Hispanics (Mohanty et al., 2009). The Hispanics population has a higher occurrence of NASH and cirrhosis while African American population has a lower chance of developing liver failure (Kalia & Gaglio, 2016; Satapathy & Sanyal, 2015). Prevalence of NAFLD in Asia has been increasing and can also be found commonly in subjects with normal BMI (Kalia & Gaglio, 2016). In Malaysia, the prevalence of NAFLD was investigated among diabetic patients and Malays have the highest prevalence (61%), followed by Indians (52%) and Chinese (42%). This trend is in parallel with the higher prevalence of central obesity and fat calorie intake in the Malays (Chan et al., 2013).

2.3.4 Diet, smoking and lifestyle

High energy diet including high fat, refined grains and pastries, and sugar-laden beverages poses risk for development of metabolic syndrome which can lead to NAFLD (Satapathy & Sanyal, 2015). Researchers also observed the detrimental effect of grilled food and high fat intake on the risk of NAFLD (Miele et al., 2014). Besides that, both active and passive smokers are more predisposed to having NAFLD (Akhavan Rezayat et al., 2018; Okamoto et al., 2018). People with a sedentary lifestyle are also predisposed to develop NAFLD and NASH (Satapathy & Sanyal, 2015). By engaging in physical activity at least once a week, the risk of NAFLD would be halved (Miele et al., 2014). Due to the absence of an effective treatment for NAFLD to date, various clinical practice guidelines published by American Association for the Study of Liver Diseases (AASLD) and European Association for the Study of the Liver (EASL)–European Association for the Study of Diabetes (EASD)–European Association for the Study of Obesity (EASO) recommend weight reduction in managing patients with NAFLD (Chalasani et al., 2018). A randomised trial showed that NASH patients who underwent intensive lifestyle changes incorporating diet, physical activity and behaviour modification had an average of 9.3% weight loss and showed improvement in steatosis, necrosis and inflammation (Harrison, Fecht, Brunt, & Neuschwander-Tetri, 2009; Promrat et al., 2010).

2.3.5 Role of genetics in NAFLD

Familial aggregation, twin studies and interethnic differences in susceptibility to the disease have shown the heritable component of NAFLD (Guerrero, Vega, Grundy, & Browning, 2009; Makkonen, Pietilainen, Rissanen, Kaprio, & Yki-Jarvinen, 2009; Schwimmer et al., 2009). By conducting a non-hypothesis driven genome-wide

association study (GWAS), novel single nucleotide polymorphisms (SNPs) associated with the disease can be identified. This is often followed by cross-sectional candidate gene association studies to validate the implication of the gene variant in disease development and progression.

The first GWAS in NAFLD identified an association of increased hepatic triacylglycerol accumulation and the adiponutrin (*PNPLA3)* gene (Romeo et al., 2008). Subsequent studies across various ethnicities including in Malaysia further confirmed the role of PNPLA3 rs738409 variant in conferring susceptibility to NAFLD and NASH (R. Xu, Tao, Zhang, Deng, & Chen, 2015; Zain et al., 2012). The second GWAS was conducted on 236 biopsy-proven NAFLD patients and discovered gene variants associated with NAFLD severity and histological features. The rs2645424 in gene encoding farnesyl diphosphate farnesyl transferase 1 (*FDFT1*) is associated with higher NAFLD activity score while rs343062 is associated with a degree of fibrosis (Chalasani et al., 2010). In 2011, Speliotes et. al conducted the largest GWAS across 7176 individuals involving approximately 2.4 million SNPs genotyped. Following discovery of SNPs from 45 loci associated with hepatic steatosis, a validation phase carried out in 592 histologically proven NAFLD patients verified association of NAFLD with five SNPs: rs738408 in *PNPLA3*, rs2228603 in neurocan (*NCAN*), rs4240624 in protein phosphatase1 regulatory inhibitor subunit 3B (*PPP1R3B*), rs780094 in glucokinase regulator (*GCKR*) and rs12137855 in lysophospholipase-like 1 (*LYPLAL1*) (Speliotes et al., 2011). This leads to successive case-control association studies carried out in different cohorts in addition to in vitro and in vivo studies to validate the findings. Subsequently the association between hepatic triglyceride content and *NCAN* rs2228603 variant was abolished upon conditioning on the transmembrane 6 superfamily 2 human gene (*TM6SF2*) rs58542926 variant (Kozlitina et al., 2014). This exome wide association study on hepatic fat content found significant association of TM6SF2

variant with elevated liver enzymes in three independent populations and in vivo study further suggested that TM6SF2 activity is required for the normal very low density lipoprotein (VLDL) secretion (Kozlitina et al., 2014). Gene function analyses were also carried out in unravelling the molecular pathways implicated in the pathogenesis of NAFLD (Hou et al., 2018; Li et al., 2018). Besides investigating the SNPs discovered in GWAS, genes related to the underlying mechanisms such as genes influencing lipid metabolisms and liver injury were also explored (Macaluso, Maida, & Petta, 2015). However there is no other well-established genetic risk factor as widely validated as *PNPLA3* rs738409.

More recently, some studies suggest the use of SNPs in predicting the risk of NAFLD and NASH. For example, the genetic risk score combining *PNPLA3* rs738409, superoxide dismutase 2 (*SOD2*) rs4880, kruppel like factor 6 (*KLF6*) rs3750861, and lipin 1 (*LPIN1*) rs13412852 performed better than aspartate transaminase and age in predicting NASH, giving a sensitivity of 90% specificity of 36% and area under the receiver-operating characteristic curve (AUROC) of 0.80 (Nobili et al., 2014). A similar genetic risk score comprising of *PNPLA3* rs738409, *TM6SF2* rs58542926 and *KLF6* rs3750861 found that a score of 3-4 quadrupled the risk of NASH-cirrhosis (Vespasiani-Gentilucci et al., 2018).

2.4 Pathophysiology of NAFLD

A “two-hit” model was initially proposed in explaining the development of NASH from NAFL. Fatty infiltration in the liver was depicted as the first “hit”, followed by the “second hit” which is the free radicals causing oxidative stress to the liver resulting in cascades of inflammatory events (Day & James, 1998) Subsequently “multiple parallel hit” hypothesis was proposed suggesting the relative

benign condition of NAFL(Buzzetti, Pinzani, & Tsochatzis, 2016; Tilg & Moschen, 2010). The initial "hit" is described as insulin resistance consequently causing hepatic steatosis. Insulin resistance results in hepatic de novo lipogenesis which in turn reduces adipose tissue lipolysis, and subsequently increases fatty acid in the liver (Bugianesi, Moscatiello, Ciaravella, & Marchesini, 2010). The multiple "hits" leading to NASH could include unregulated hepatic adipose tissue lipolysis, endotoxins elevation, microbiota affecting energy metabolism, inflammation due to activation of aryl hydrocarbon receptor by trans-fatty acids and fructose, and imbalance of adipocytokines causing release of pro-inflammatory cytokines (Ibrahim, Kohli, & Gores, 2011; Larter, Chitturi, Heydet, & Farrell, 2010; Samuel, Petersen, & Shulman, 2010; Tilg & Moschen, 2010).

2.5 Histopathological classification of NAFLD

To enable assessment of histological features encompassing the whole spectrum of NAFLD, semi-quantitative system had been developed for the grading of NASH activity. Brunt et al. (1999) proposed grading of necro-inflammatory changes (mild, moderate and severe) and the staging of fibrosis according to the extent and location. In 2005, thorough univariate and multivariate analysis results in the development of NASH Activity Score (NAS) by the NASH Clinical Research Network (CRN) Pathology Committee for the grading of NASH activity and staging of fibrosis. The NAS score is the unweighted sum ranging from 0 to 8, comprising three histological components: steatosis (0-3), lobular inflammation (0-3) and ballooning degradation (0-2) (Kleiner et al., 2005). A NAS score of ≥ 5 is generally considered as NASH but lower NAS scores could also be found in NASH patients where steatosis and hepatocyte ballooning are present.

Table 0.1: NAS scoring system by Kleiner et al. (2005)

Histological feature	Score	Definition
Steatosis	0	<5 %
	1	5-33%
	2	34-66%
	3	>66%
Hepatocyte ballooning	0	None
	1	Few
	2	Many
Inflammation	0	None
	1	1-2 foci per x20 field
	2	2-4 foci per x20 field
	3	>4 foci per x20 field

A relatively new algorithm was developed by the Fatty Liver Inhibition of Progression Pathology Consortium comprising Steatosis, Activity, Fibrosis (SAF) scores based on morbidly obese patients undergoing liver biopsy before bariatric surgery (Bedossa et al., 2012). The SAF score summarises the main histological lesion including steatosis score (0-3); activity grade (0-4) which is the unweighted sum of hepatocyte ballooning (0-2) and lobular inflammation (0-2); and stage of fibrosis (0-4). The presence of NASH is indicated in patients with steatosis and showing at least grade 1 in both the features of activity (ballooning and lobular inflammation).

Table 0.2: SAF scoring system by Bedossa et. al (2012)

Histological feature	Score	Definition
Steatosis	0	<5 %
	1	5-33% mild
	2	34-66% moderate
	3	>66% marked
Hepatocyte ballooning	0	None
	1	Few, Hepatocytes have rounded contours with clear reticular cytoplasm with similar size to normal hepatocytes
	2	Many, Cells are rounded with clear cytoplasm and twice as large as normal hepatocytes
Inflammation	0	None
	1	1-2 foci per x20 field
	2	>2 foci per x20 field

2.6 Histological scoring of fibrosis

A distinctive pattern of fibrosis is observed in NAFLD, usually beginning in the pericentral zone 3 region. This fibrosis can increase in density and complexity, leading to bridging fibrosis and cirrhosis. Various scoring systems have been used for histological staging of fibrosis. Some studies used Metavir, Ishak or Sheuer which were originally used for the scoring of fibrosis in chronic viral hepatitis (Asselah, Marcellin, & Bedossa, 2014). For the assessment of fibrosis in NAFLD, three scoring systems that gain more recognition include the Brunt system, NASH CRN system and SAF system. The criteria for staging of fibrosis according to the different systems are shown in Table 2.3, 2.4, 2.5. The NASH CRN system is a modification of the Brunt system by further dividing stage 1 fibrosis into three sub-stages. Unlike NASH CRN system where the fibrosis stage is reported separately from the NAFLD activity score (NAS), the SAF system incorporates fibrosis score into the final SAF score (Bedossa et al., 2012). Of these, NAS is the most validated system currently available, being widely accepted and used in clinical trials as an assessment of treatment endpoint, making it useful for comparative and interventional studies (Sanyal et al., 2011; Stal, 2015).

Table 0.3: Staging of fibrosis according to Brunt et. al (1999)

Stage	Criteria
1	zone 3 perisinusoidal fibrosis
2	periportal fibrosis with the zone 3 perisinusoidal fibrosis
3	bridging fibrosis
4	cirrhosis

Table 0.4: Staging of fibrosis according to Kleiner et. al (2005)

Stage	Criteria
1a	zone 3 perisinusoidal fibrosis with delicate collagen deposition
1b	zone 3 perisinusoidal fibrosis with dense collagen deposition
1c	portal or periportal fibrosis
2	portal fibrosis with the zone 3 perisinusoidal fibrosis
3	bridging fibrosis
4	cirrhosis

Table 0.5: Staging of fibrosis according to Bedossa et. al (2012)

Stage	Criteria
F0	No fibrosis
F1	perisinusoidal zone 3 or 1c portal fibrosis
F2	perisinusoidal and periportal fibrosis without bridging
F3	bridging fibrosis
F4	cirrhosis

2.7 Natural History and Prognosis of NAFLD

A study investigating the long term outcomes of 40 patients with NAFL found that none progressed to NASH or cirrhosis after a median 11 years of follow up (Teli, James, Burt, Bennett, & Day, 1995). Patients with NASH has a high tendency of disease progression, where the progression of liver fibrosis was found in one third of patients with NASH within an average follow up of 4.3 and 5.7 years respectively in two longitudinal studies (Fassio, Alvarez, Dominguez, Landeira, & Longo, 2004; Harrison, Torgerson, & Hayashi, 2003).

The prevalence of cirrhosis development was also higher in patients with NASH (10.8%) compared to patients with simple steatosis (0.7%) (Angulo, 2010). In terms of mortality, individuals with NAFL have similar long-term mortality as the general population while patients with NASH have increased liver-related mortality rate (0.9% vs 7.3%) (Angulo, 2010). While NASH has a higher liver-related mortality (OR 5.71)

compared to NAFL, this risk is exacerbated with the presence of advanced fibrosis (OR 10.06) (Musso, Gambino, Cassader, & Pagano, 2011).

Recent studies, however, have begged to differ from the central dogma that only NASH is a progressive disease while NAFL is a benign condition (Wong, Wong, et al., 2010). With a mean follow up of 3.7 years, 16 out of 25 NAFL patients were found to develop NASH, with six of them having bridging fibrosis. This ballooning and fibrosis progression is often accompanied by deterioration of metabolic risk factors, such as higher weight gain and a higher incidence of diabetes (Pais et al., 2013). In addition, arecent retrospective study with serial biopsies of median interval of 8.4 years found that fibrosis progression is equally likely to occur among patients with NAFL and patients with NASH (54% and 46%), irrespective of the diagnosis on baseline (Hagstrom, Elfwen, Hultcrantz, & Stal, 2018). Anyhow, in comparing the annual fibrosis progression rate (FPR) in patients with baseline stage 0 fibrosis, subjects with NASH have average progression of one stage over 7.1 years (FPR 0.07) while subjects with NAFL have slower progression, with an average progression of one stage per 14.3 years (FPR 0.14) (S. Singh et al., 2015).

More studies suggest that progression of fibrosis and severe liver disease is accounted by the stage of baseline fibrosis instead of NASH. An average 20 years of follow up on 646 biopsy-proven NAFLD patients found that the risk of liver decompensation or hepatocellular carcinoma increases according to the stage of fibrosis, with a hazard ratio of 1.9 in patients without fibrosis but 104.9 in patients with cirrhosis (F4). The risk of mortality is also independent of the presence of NASH (Hagstrom et al., 2017). Increased risk of mortality was found in patients with advanced fibrosis (hazard ratio 3.3), regardless of their NAS score(Ekstedt et al., 2015). This finding was also replicated in retrospective studies involving medical centers in United States,

Europe and Thailand (Dulai et al., 2017; Vilar-Gomez et al., 2018; von Roenn, 2018). It was also found that stage of fibrosis is strongly associated with risk of death and liver transplant, where the hazard ratio of fibrosis stage 1, 2, 3, 4 are 1.88, 2.89, 3.76, 10.9, respectively compared to stage 0 (Angulo et al., 2015).

2.8 Diagnosis of NAFLD

Patients with NAFLD may present fatigue or discomfort on the right upper quadrant while many can be asymptomatic. Patients are often overweight, and some may have normal liver function test result. In the Dionysos study, the prevalence of NAFLD is actually similar among subjects with elevated liver enzymes (25%) and subjects with normal liver enzymes (20%) (Bedogni et al., 2005), suggesting most cases of NAFLD may remain undiscovered in general setting due to their normal biochemistry test. Existing health records from UK, Italy, Spain and Netherlands showed that the prevalence rate of NAFLD is 1.9%, which is many folds lower than the estimated prevalence of NAFLD (20-30%) in the European population. This suggests under-recording at primary care, possibly due to missed diagnosis or lack of confidence in making the diagnosis (Alexander et al., 2018).

Screening of NAFLD is not routinely performed as there is a lack of effective interventions (Chalasani et al., 2018; Sattar et al., 2014). Presence of NAFLD is often discovered incidentally when patients present with elevated liver enzymes or when hepatic steatosis is being found on thoracic or abdominal imaging.

2.8.1 Imaging

In the diagnosis of NAFLD, ultrasonography is widely available and inexpensive to be carried out, in which increased echogenicity is often present. However, the diagnostic accuracy of ultrasonography declines with less than 20% hepatic steatosis (Khov, Sharma, & Riley, 2014). Magnetic resonance imaging such as magnetic resonance spectroscopy has good sensitivity and specificity in grading steatosis but is more expensive and not widely utilised for NAFLD (Dyson et al., 2014). Transient elastography is used to evaluate the degree of fibrosis in NAFLD patients but the interpretation can be obscured by abdominal fat (Chang, Goh, Ngu, Tan, & Tan, 2016). Magnetic resonance elastography (MRE) has higher accuracy in assessing fibrosis in NAFLD but is not widely available as a fibrosis screening utility in routine clinical practice (Dulai, Sirlin, & Loomba, 2016). These imaging modalities are often used to detect hepatic steatosis but are not reliable in the grading of steatohepatitis or staging of fibrosis.

2.8.2 Liver Biopsy

Due to the inability of imaging studies to reliably convey the degree of NASH and fibrosis, liver biopsy remains the gold standard in the diagnosis of NAFLD (Nalbantoglu & Brunt, 2014). It is able to provide a definite diagnosis, showing the degree of hepatic steatosis and level of necroinflammation, hepatocellular ballooning and staging of fibrosis (Sumida, Nakajima, & Itoh, 2014). This will then allow stratification of patients into different secondary care intervention as the likelihood of disease progression can be assessed.

Albeit the gold standard, liver biopsy is not indicated to all patients with the suspected disease, but only those who would benefit from the diagnosis, therapeutic guidance and prognostic perspectives. This is because liver biopsy is invasive, expensive and accompanied with risk of excessive bleeding (Sumida et al., 2014). It is normally performed in patients with elevated liver enzymes where the diagnosis remains uncertain (Chalasani et al., 2018). It also helps to rule out other concomitant liver diseases, such as hemochromatosis or autoimmune hepatitis (Spengler & Loomba, 2015). The American Association of Liver Disease recommends that liver biopsy should only be considered in NAFLD patients with increased risk to have NASH and advanced fibrosis (Chalasani et al., 2018). They also suggest the use of liver biopsy to exclude co-existing chronic liver disease and competing aetiologies for hepatic steatosis in patients with suspected NAFLD. Other literatures suggest that patients undergoing bariatric surgery should have an intraoperative liver biopsy due to the high prevalence of NASH among morbidly obese patients (Chavez-Tapia et al., 2010). The test can also be ordered upon clinical findings that suggest cirrhosis, such as thrombocytopenia, hypoalbuminemia and AST>ALT (Ratziu, Bellentani, Cortez-Pinto, Day, & Marchesini, 2010).

2.8.3 Molecular biomarkers

The inherent limitations of liver biopsy have led to a growing interest in developing and validating noninvasive biomarkers for NAFLD, NASH and advanced fibrosis. A good biomarker must possess good sensitivity to identify the disease, sufficient specificity to exclude individuals not having the disease, cost-effective, easy to be performed and reproducible (Obika & Noguchi, 2012). Identification of biomarkers has several different approaches: (1) selecting biochemical markers derived

from large case-control association studies, (2) creating algorithms including markers involved in inflammation and extracellular matrix turnover and (3) non-hypothesis driven discovery using microarray, proteomics and metabolomics (Fitzpatrick & Dhawan, 2014).

Circulating microRNAs (miRNAs) have recently emerged as a suitable candidate for disease biomarkers due to their role in modulating gene expression and them being relatively stable in various biofluids. While there is no established biomarkers available for identifying NASH from NAFL, increased epression of miR-34a, miR-122 and miR-192 were repeatedly observed in NASH patient compared to NAFLD patients (Cermelli, Ruggieri, Marrero, Ioannou, & Beretta, 2011; Pirola et al., 2015). The role of miRNAs in NAFLD is discussed in detail in chapter three. In the identification of advanced fibrosis, several panel marker tests have been derived by combining the variables found to be significant at univariate analysis, followed by multivariate analysis to yield a predictive algorithm. This includes NAFLD fibrosis risk score, FIB-4 score, BARD, Enhanced Liver Fibrosis (ELF), and the review of these scoring systems are discussed in depth in chapter four.

CHAPTER 3: CIRCULATING MIRNAS AS MOLECULAR BIOMARKER OF NAFLD

3.1 Introduction

One of the compelling challenges in the management of NAFLD patients is the asymptomatic nature of disease. The cases of NAFLD are often discovered incidentally following elevated liver enzymes but there are also substantial numbers of patients being presented with normal liver enzymes. Imaging studies such as ultrasound, computed tomography and magnetic resonance imaging have been proven to be important modalities in diagnosing NAFLD but not sufficient to distinguish between NAFL and NASH (Fitzpatrick & Dhawan, 2014).

Stratification of NAFLD patients into those with NASH and those with NAFL is crucial as the NASH patients should be referred to secondary care interventions involving lifestyle modification and/or strict diet restriction. NASH patients with fibrosis have a higher chance of disease progression, potentially leading to higher risk of morbidity and mortality such as cirrhosis, hepatocellular carcinoma and even cardiovascular diseases (Angulo, 2010; Fassio et al., 2004). It is therefore imperative to distinguish NASH from NAFL and identifying the stage of fibrosis in order to provide risk stratification and intervention in impeding disease progression in the NASH patients.

The definite diagnosis of NASH is based upon liver biopsy followed by the histopathological examination, yet it is risky and subjected to high cost, sampling error, and inter-observer variation (Rockey et al., 2009). Being an invasive procedure, its use in disease monitoring and treatment trials are limited. According to clinical practice guideline, liver biopsy should only be considered in NAFLD patients who are at increased risk of NASH and advanced fibrosis ($F \geq 3$) as this group of patients would

have a higher risk of developing liver cirrhosis and hepatocellular carcinoma (Chalasani et al., 2018). With the majority of patients having NAFL and low-grade NASH, a non-invasive test that diagnoses NASH or advanced fibrosis would be clinically beneficial to reduce the need for liver biopsy.

There has been an increasing recognition of miRNAs as biomarkers due to their relatively stable expression in various biofluids (Peltier & Latham, 2008). Unlike protein biomarker, miRNAs expression is tissue-specific, not subjected to degradation by protease and can be detected using PCR which implies a higher sensitivity (Turchinovich, Baranova, Drapkina, & Tonevitsky, 2018). MiRNAs are small non-coding RNAs which regulate gene expression post-transcriptionally thereby modulating various biological processes such as cellular development, differentiation, proliferation, apoptosis, and metabolism. The role of miRNAs in the progression of various common diseases heralds these molecules as novel markers to detect and classify disease severity, determine prognosis and predict response to existing therapies.

Lately, miRNAs have been found to be indicators of hepatocyte injury and inflammation, modulating a diverse spectrum of liver disease (Bala et al., 2012). Therefore, quantification of circulating miRNAs will unravel a new repertoire of biomarkers in the diagnosis and stratification of the disease. In NAFLD, an ongoing challenge is the search for novel non-invasive biomarkers that reflects the severity of liver disease, thus facilitating the detection of NAFLD at an early stage, as well as for monitoring the progression of patients with fibrosis.

Studies on the discovery and validation of miRNAs as diagnostic biomarkers have been on the risein the last few years. Most of these are validation studies, in whichmiRNAs profiling has been performed on a few selected miRNAs in terms of their relevance in the pathogenesis or based on their differential expression in liver

tissues from previous findings. Only two studies are multi-stage, with an initial discovery phase followed by validation phase in another cohort (Pirola et al., 2015; Tan, Ge, Pan, Wen, & Gan, 2014). Furthermore, only a few studies investigated the diagnostic performance of miRNAs in the ability to stratify NASH from NAFL and to discriminate the fibrosis stage(Cermelli et al., 2011; Liu et al., 2016; Lopez-Riera et al., 2018; Pirola et al., 2015; Tan et al., 2014).

Therefore the aim of this study was to first ascertain the potential of circulating miRNAs as biomarkers by comparing the miRNAs expression profile in NAFLD against matched healthy controls. This is followed by validation in a larger population to determine the potential relationship between selected circulating miRNAs and various stages of NAFLD including NAFL vs. NASH and their relative expression based on different histopathological severity.

3.2 Literature review

3.2.1 MiRNA biogenesis and function

MiRNAs are endogenous, non-coding single-stranded RNAs that are 19-24 nucleotides in length. MiRNAs exhibit high evolutionary conservation over a wide range of species despite the diversity in their expression profiles in different cell types at different developmental stages, suggesting important regulatory functions (Niwa & Slack, 2007). MiRNAs regulate post-transcriptional gene expression by binding to target mRNAs leading to their degradation or repression of translation and are thus implicated in several pathophysiological processes, such as cell proliferation, differentiation, metabolism, apoptosis, and cancer (Bartel, 2004; Sayed & Abdellatif, 2011).

The miRNA genes are transcribed in the nucleus by RNA polymerase II into primary miRNA transcripts (pri-miRNA), which are 5' capped (Cap) and 3' polyadenylated (AAA). Cleaving of pri-miRNA mediated by RNase III enzyme Drosha forms a 60-70 nucleotides long precursor miRNA (pre-miRNA) which is a hairpin structure that allows recognition by the nuclear export factor Exportin-5 (Exp5) to be exported to the cytoplasm (Lee et al., 2003; Lund, Guttinger, Calado, Dahlberg, & Kutay, 2004).Once in the cytoplasm, the terminal base pairs and loop of the pre-miRNA are cleaved off by the RNase III enzyme, Dicer, leaving a ~22 nucleotide double-stranded RNA referred to as the miRNA:miRNA* duplex (Bartel, 2004). The duplex is separated into two strands whereby one is degraded and the other strand forms the mature miRNA and is incorporated into the RNA-induced silencing complex (RISC) (Ameres, Martinez, & Schroeder, 2007). The RISC is a ribonucleoprotein that contains the mature miRNA which later recognises the complementary mRNA transcript and an Argonaute (Argo) protein which coordinates downstream gene-silencing events by interacting with other protein factors(Meister, 2013).

Translation of target mRNA can be inhibited by the RISC complex through 1) mRNA cleavage 2) mRNA destabilisation or 3) translational repression (Filipowicz, Bhattacharyya, & Sonenberg, 2008). Based on the degree of complementarity of miRNA and the 3'UTR of target miRNA, endonucleolytic cleavage of the target mRNA takes place when there is perfect complementarity whereas imperfect binding leads to mRNA destabilisation or translational repression (Huntzinger & Izaurralde, 2011; Pillai, Bhattacharyya, & Filipowicz, 2007). Inverse correlation in the levels of miRNAs and their target mRNAs suggests that miRNAs can induce mRNA destabilisation and subsequent degradation, possibly through the binding of the RISC complex followed by deadenylation of the target mRNA and degradation via an exonuclease (Lim et al., 2005; Piao, Zhang, Wu, & Belasco, 2010). Despite thousands of miRNAs having been discovered, only some miRNAs have a significant concentration in cells to exert posttranslational regulation (Witwer, 2015).

3.2.2 Circulating miRNA as biomarker

A biomarker is defined as cellular, biochemical or molecular alteration that can be quantified in biological media such as human tissues, cells or fluids(Mayeux, 2004). In clinical practice, it can provide information on disease prediction, diagnosis, prognosis and effectiveness of therapy.

The following are some of the criteria for an ideal disease biomarker: 1) easily obtained through non-invasive means, 2) stable under various conditions, 3) ease and inexpensive in quantification, 4) highly specific to disease of interest, and 5) allows early identification of disease before appearance of clinical symptoms (Weber et al., 2010). Previous studies revolved around the discovery of protein-based biomarkers but this was hindered by the challenges such as 1) complexity of protein composition in

various biological samples 2) posttranslational modifications of protein 3) low level of target protein in serum and plasma level, and 4) intricacies in developing a reliable high-affinity capture agents (Weber et al., 2010). These difficulties impede the discovery and development of protein-based biomarkers with high sensitivity and specificity as it is usually an expensive, time-consuming and laborious task.

The search for a non-invasive tool in clinical diagnosis has prompted the initial discovery of cell-free miRNAs in serum and plasma (Lawrie et al., 2008). This finding was further confirmed when distinct composition of circulating miRNAs was traced in various body fluids such as breast milk, colostrum, saliva, seminal fluid, tears, urine, amniotic fluid, cerebrospinal fluid, plasma, pleural fluid, and peritoneal fluid and thus,providing a non-invasive approach in diagnosis of various diseases (Wang, Zhang, Weber, Baxter, & Galas, 2010; Weber et al., 2010).

The most enigmatic feature which makes miRNA a potential biomarker is its remarkable stability. Despite harsh conditions including multiple freeze-thaw cycles, extended storage and extreme pH conditions, levels of circulating miRNAs are found to be stable, reproducible and consistent among individuals (Chen et al., 2008; Mitchell et al., 2008). In addition, they possess the ability to resist RNase mediated degradation (Gilad et al., 2008; Mitchell et al., 2008). This stability can be attributed to the microvesicular packaging of miRNAs prior to their secretion into the circulation (Mitchell et al., 2008). Under physiological and pathological conditions, miRNAs are actively secreted from the cell membrane after encapsulation in microvesicles/microparticles, exosomes (small membrane fragments 30-100nm derived from endosomal compartment), lipoproteins and ribonucleoprotein complexes or via apoptotic bodies (released as cells undergo apoptotic death) (Arroyo et al., 2011; Diehl

et al., 2012; Gallo, Tandon, Alevizos, & Illei, 2012; Mause & Weber, 2010; Vickers, Palmisano, Shoucri, Shamburek, & Remaley, 2011).

Detection and quantification of miRNAs are also easier than protein-based biomarkers. Development of synthetic complementary oligonucleotides offers great sensitivity and utilisation of qPCR and other DNA amplification methods facilitate high-throughput detection and quantification of miRNA expression (Andreasen et al., 2010; Schmittgen & Livak, 2008; Varkonyi-Gasic, Wu, Wood, Walton, & Hellens, 2007). Furthermore, miRNAs exhibit low complexity and absence of post-processing modifications with high conservation of sequence between humans and model organisms, making them ideal candidates for biomarkers investigation (Weber et al., 2010).

Being abundantly expressed in the liver, miRNAs regulate a vast array of functions and are implicated in the pathogenesis of various liver diseases including viral hepatitis, drug-induced liver injury, autoimmune hepatitis, hepatic fibrosis, and hepatocellular carcinoma (Enache et al., 2014; Szabo & Bala, 2013). The differential expressions of circulating miRNAs in the serum under various pathological conditions could be harnessed as a means to diagnose and determine the prognosis of the disease (Chen et al., 2008; Kroh, Parkin, Mitchell, & Tewari, 2010).

In addition to being localised in cells, extracellular circulating miRNAs can be found in lipid or lipoprotein complexes, apoptotic bodies, or secreted actively in microvesicles or exosomes (Wang et al., 2010). NAFLD is often accompanied by hepatocyte apoptosis and necrosis, in which these events release cellular miRNAs directly into the circulation, making miRNAs having better sensitivity compared to liver enzymes in the diagnosis of liver injury (Zhang et al., 2012)

3.2.3 Approaches in miRNAs profiling

With circulating miRNAs posing as promising biomarkers for various diseases, care must also be taken on the pre-analytical and analytical features that may affect serum miRNAs quantification. Column-based RNA extraction methods are modified and adapted to selectively capture and retain RNA fractions taking into account the small size and unique structure of miRNAs (Pritchard, Cheng, & Tewari, 2012). Advent of high-throughput profiling techniques has been useful in miRNA expression profiling: next-generation sequencing (NGS) technology, microarray chip technology and quantitative reverse transcription real-time PCR (qRT-PCR) (Dedeoglu, 2014).

Microarrays employ the design of probes by which the sequence information is limited to those found in public database combined through bioinformatics and extensive cloning experiments (Jaksik, Iwanaszko, Rzeszowska-Wolny, & Kimmel, 2015). A short length of mature miRNA sequences constrains microarray probe design, whereby often the entire miRNA sequence is used as probe. Hence the diverse range of miRNA probes melting temperatures results in difficulty in optimising an experimental condition to ensure simultaneous miRNA hybridisations (Lee et al., 2008). Despite being the most cost-efficient of the three methods, microarrays can only screen for the most abundant miRNAs due to their limited sensitivity. Thus the low amount of RNA in serum and plasma samples can lead to less reliable results (Chugh & Dittmer, 2012).

On the other hand, utilisation of high-throughput sequencing allows identification of novel miRNAs without the results being hampered by variability in melting temperatures, co-expression of similar miRNA family members, or post-transcriptional modifications (Git et al., 2010). Unlike microarray and qRT-PCR, there is no requirement of prior sequence information for probe design. However, the RNA ligation and PCR amplification steps are subjected to inherent biases and

sequencing is relatively laborious accompanied with higher costs and the lack in well-established tools for computational analysis (Bissels et al., 2009; Git et al., 2010; Pritchard, Cheng, et al., 2012). It also requires high RNA inputs which is hard to be obtained in a small sample of biofluids (Coenen-Stass et al., 2018).

The qRT-PCR technique is often considered as a "gold standard" in the detection and absolute quantification of miRNAs and has been widely applied to miRNA investigations (Roberts, Coenen-Stass, & Wood, 2014). This technique involves a reverse transcription reaction converting the miRNA into complementary DNA (cDNA) followed by amplification of the cDNA using PCR and quantification of the amplicons in real-time. It is well established with high sensitivity and specificity. However, when used as a discovery tool, the rapid increase in the number of miRNAs renders qPCR inefficient on a genomic scale, as it has a medium throughput due to the limited number of samples processed per day (Tan & Tan, 2017). Thus it is normally performed during the validation stage rather than the discovery stage. Nevertheless, qRT-PCR is more robust than microarray, has a moderate cost and requires less starting material than both microarray and sequencing (Git et al., 2010).

3.2.4 Challenges associated with circulating miRNAs

Cellular contamination of serum or plasma samples with various blood cell types is one of the common sources of bias in miRNA profiling (McDonald, Milosevic, Reddi, Grebe, & Algeciras-Schimnich, 2011). It is therefore, crucial to minimise the cellular content of the biofluid in miRNA profiling discovery projects by adhering to the standardisation of sampling procedure as suggested by The National Cancer Institute Early Detection Research Network (EDRN) (Blondal et al., 2013). In miRNA profiling studies, rigorous quality control of serum/plasma samples is critical as expression of

some miRNAs can be greatly affected by the degree of haemolysis, such as miR-21, miR-106a, miR-92a, miR-17, miR-29a, miR-451, and miR-16 (Kirschner et al., 2013; Kirschner et al., 2011; Yamada et al., 2014). As a part of the quality control, delta Cq (ΔCq) of (miR-23a – miR-451) can be employed where a score of 7-8 indicates a high possibility of haemolysis and the sample should be eliminated in the discovery and validation studies (Blondal et al., 2013).

One of the important analytical parameters in circulating miRNA studies is the selection of suitable endogenous controls which are genes stably expressed across the sample set. The qRT-PCR data are normalised using endogenous control to correct for variables such as the amount of starting template and enzymatic efficiencies (Davoren, McNeill, Lowery, Kerin, & Miller, 2008). Unfortunately, some of the well-defined endogenous controls used in tissue and cell culture show variable expression in serum and plasma. Several publications suggest the utilisation of small nuclear RNA U6 but it does not reflect the pattern of transcription and processing of miRNA molecules, and is especially downregulated in patients with liver fibrosis (Benz et al., 2013). Nevertheless, normalising the data against the geometric mean miRNAs detected in all samples or against synthetic exogenous miRNAs such as cel-miR-39 can also be performed to compensate for the lacking of reliable endogenous controls (Marabita et al., 2016). Normalisation using the global mean of miRNA expression data is normally used for large scale miRNAs profiling data (Marabita et al., 2016). Although spiked-in synthetic control miRNAs can be added as an inter-plate calibrator to correct for the deviations of the experimental procedures, it does not eliminate the deviations in sampling and quality of samples (Schwarzenbach, da Silva, Calin, & Pantel, 2015).

3.2.5 Aberrant expression of circulating miRNAs in NAFLD

In exploring the role of circulating miRNAs as putative biomarkers, there is only one study that performed global serum miRNA profiling whilst most other studies investigated the differential expression of individually selected miRNAs based on disease-associated pathways. The pioneering study that investigates the circulating miRNAs in NAFLD patients was conducted in 34 biopsy-proven patients and 19 healthy controls. MiR-122 levels and miR-16 were found to be upregulated (7.2 fold and 5.5 fold respectively) in NAFLD patients compared to the healthy controls. Meanwhile, mir-122 and mir-34a were upregulated in NASH patients compared to NAFL patients (Cermelli et al., 2011).

Subsequently, as the de novo discovery of circulating miRNAs emerges, deep sequencing was performed by Tan et. al (2014) followed by construction of a logistic regression model using a training cohort (90 healthy controls vs 90 NAFLD) and validation of diagnostic accuracy in an independent cohort (80 healthy controls vs 103 NAFLD). The results demonstrated high diagnostic accuracy of miRNA panel (miR-122, miR-1290 and miR-27b) for NAFLD with AUROC of 0.856. Despite the large sample size, the differential expression of miRNAs between NAFL and NASH, between different stages of fibrosis, and correlation with other histological features were not investigated.

A multistage cross-sectional study was also performed by Pirola et. al (2015) starting with 84 miRNAs expression profiling that were previously found to be associated with liver and heart disease, followed by an independent validation of selected miRNAs to explore the differential expression in serum samples and liver tissue. The result from validation cohort (19 healthy control vs 30 NAFL vs 47 NASH) identifies upregulation of miR-122, miR-192, miR-375, miR-19a and miR-19b, and

miR-125b in NAFLD patients compared to healthy controls (Pirola et al., 2015). In stratifying the disease severity, miR-122, miR-192 and miR-375 are able to distinguish NASH from NAFL whilst only miR-122 is able to distinguish significant fibrosis (F≥2).

In addition to the diagnostic performance in NAFLD, some studies also found the potential of miRNAs in stratifying NASH patients from NAFLD patients. To date, there are some miRNAs found to be upregulated in NASH compared to NAFL, such as miR-122, miR-34a, miR-192, miR-21, miR-27b, miR-22, and miR-375 (Becker et al., 2015; Cermelli et al., 2011; Liu et al., 2016; Pirola et al., 2015) whilst miR-16 and miR-30c are found to be downregulated (Lopez-Riera et al., 2018). Meanwhile, there are only a few studies which investigate the association of miRNAs with different histological severity. Being one of the earliest studies, Celikbilek et al (2014) found that miR-197 and miR-10b are negatively correlated with the degree of inflammation but not other histological features. However, this finding lacks confidence as miRNAs profiling was only performed on 20 NAFLD patients and 20 healthy controls (Celikbilek et al., 2014). MiR-122 and miR-34a are frequently found to be correlated with fibrosis, hepatocellular ballooning and inflammation (Cermelli et al., 2011; Liu et al., 2016; Miyaaki et al., 2014; Yamada et al., 2013). Other miRNAs that are found to be dysregulated include miR-21 (inflammation) (Becker et al., 2015), miR-192 (inflammation and ballooning) (Liu et al., 2016), miR-16, miR-22, and miR-29a (fibrosis) (Liu et al., 2016; Lopez-Riera et al., 2018).

3.3 Methodology

3.3.1 Subject recruitment

Consecutive histologically-confirmed NAFLD patients and control subjects without NAFLD were included in this study. Validation of ethnicity was performed by volunteers, affirming no mixed m marriages for more than two generations. The written informed consent was provided and signed by the volunteers prior to their participation. The protocol had been approved by the Medical Ethics Committee of UMMC.

After the finding of increased echogenicity in the liver as compared to the renal cortex on ultrasound, all patients with NAFLD underwent a liver biopsy. Histopathological examination was carried out on the liver biopsy samples and the histological scoring was made based on recommendations by NASH Clinical Research Network (Brunt et al., 2011; Kleiner et al., 2005). All biopsy specimens have an average of 1.5 cm long, and contained at least six portal tracts. Upon histopathological examination, tissues were scored for steatosis grade (0–3), lobular inflammation (0–3), ballooning (0–2) and fibrosis (0-4). The exclusion criteria of this study includes: alcohol consumption of more than daily intake of 10 g(Ruhl & Everhart, 2005), coexisting liver diseases such as autoimmune liver diseases, viral hepatitis B and hepatitis C infection, Wilson's disease, primary biliary cirrhosis, α-1-antitrypsin deficiency, and consumption of drug that would precipitate steatosis.

Genetically unrelated healthy individuals with BMI less than 25 kg/m^2 were recruited as controls.They also had fasting plasma glucose less than 110 mg/dL with normal liver enzymes and without dyslipidaemia.By performing liver ultrasonography on the heathy controls, they were excluded from the study upon finding of the followings: (i) diffuse increase in bright echoes in the liver parenchyma with slightly impaired visualization of the peripheral portal and hepatic vein borders; (ii) slight

diffuse increase in bright homogeneous echoes in the liver parenchyma with normal visualization of the diaphragm and portal and hepatic vein borders, and normal hepatorenal echogenicity contrast; (iii) marked increase in bright echoes at a shallow depth with deep attenuation, impaired visualisation of the diaphragm, and marked vascular blurring (Sanyal & American Gastroenterological, 2002).

Blood samples were taken on the date of liver biopsy for patients and on the day of recruitment for healthy controls. The biochemical tests were performed based on the standard operating procedure in UMMC hospital clinical laboratory. The blood biochemistry profiling and the anthropometric data were thenrecorded: age, sex, height, weight, HbA1c, aspartate aminotransferase (AST), serum alanine aminotransferase (ALT), gamma-glutamyl transpeptidase (GGT),total cholesterol,high density lipoprotein (HDL) cholesterol, low density lipoprotein (LDL) cholesterol, and serum triglyceride level. The body mass index (BMI) was calculated as weight/height^2 (kg/m^2). Measurement of clinical parameters such as systolic and diastolic blood pressure, waist circumference and pulse rate were also carried out.

3.3.2 Study design

This study was divided into two phases: (i) Biomarker discovery phase and (ii) Biomarker selection and validation phase

3.3.2.1 Phase I: Biomarker discovery phase

In search for the potential biomarkers, analysis of the miRNA expression profiles was established from serum samples from a total of 10 healthy controls and 9 NASH

patients. This is followed by a comparison of differential miRNA expression patterns to identify the dysregulated miRNAs for further analysis in Phase II.

3.3.2.2 Phase II: Biomarker Validation phase

A total of 13 candidate miRNA markers identified from the profiling experiment in Phase I was selected based on literature and the differential expression fold change. The expression of these putative markers was then validated in an expanded cohort of NAFLD patients (n=122) and healthy controls (n=50).

3.3.3 miRNA Isolation

Total RNA extraction from serum was performed using miRCURY™ RNA Isolation Kits (Exiqon, Denmark) and RNA spike-in kit according to the manufacturer's instruction. In brief, serum samples were centrifuged at 3000 x g for 5 minutes before transferring 200 µL of supernatant to new tube and 60 µL of Lysis solution BF and 1 µg of MS2 RNA carrier (Roche) was added. The mixture was vortexed for 5 seconds and incubated for 3 minutes at room temperature. A total of 20 µL of Protein Precipitation Solution BF was added and vortexed for 5 seconds followed by incubation for 1 minute at room temperature. After the mixture was centrifuged for 3 min at 11,000 x g, the clear supernatant was transferred into a new collection tube with 270 µL of Isopropanol added and vortexed for 5 seconds. The mixture was then transferred and loaded onto microRNA Mini Spin Column BF in a collection tube followed by incubation for 2 minutes at room temperature. The tube was centrifuged for 30 seconds at 11,000 x g and the flow-through was discarded. The same centrifugation conditions were used for the two subsequent washing steps using 100 µL Wash Solution 1 BF, and 700 µL Wash Solution 2 BF respectively. The column was again washed with 250 µL Wash Solution

2 BF by centrifuging at 11,000 x g for 2 minutes. The small RNA was eluted from the column in 50µL RNase-free water by centrifuging at 11,000g for 1 minute following incubation of 1 minute at room temperature.

3.3.4 Phase I (Biomarker Discovery Phase)

Total RNA was reverse transcribed in 80 µL reaction volume using the miRCURY LNA™ Universal RT cDNA synthesis kit (Exiqon). Following dilution, the cDNA template was subjected to miRNA profiling by quantitative real-time PCR (qRT-PCR) using microRNA Ready-to-Use PCR, Human panel I and panel II (v2) (Exiqon) which covers a total of 752 human miRNAs that were selected by manufacturer based on previous published association of differentially expressed miRNAs in the context of various diseases, cancers and commonly found in human serum and plasma.

3.3.4.1 Reverse transcription for cDNA synthesis

Due to low RNA yields in serum and the use of RNA carrier, the concentration of purified RNA could not be reliably measured. Therefore, fixed volumes of eluted RNA (16 µL for 768 reactions) were used for miRNA expression assays as previously reported (Andreasen et al., 2010; Komatsu et al., 2011). UniSp6 RNA template was added as synthetic spike in to be used for subsequent quality control steps (Blondal et al., 2013). Total RNA from serum was reverse transcribed using miRCURY LNA™ Universal RT cDNA synthesis kit (Exiqon) as follows:

5x Reaction buffer	16 µL
Nuclease-free water	36 µL
Enzyme mix	8 µL
UniSp6 RNA template (spike in)	4 µL
Total RNA template	16 µL
Total volume	80 µL

In thermal cycler, the mixture was incubated for 60 minutes at 42ºC and heat-inactivated for 5 minutes at 95ºC followed by immediate cooling at 4ºC and stored at -80ºC.

3.3.4.2 Quantitative real-time PCR (qRT-PCR) reaction

The qRT-PCR array profiling was performed for each sample in two 384-well qPCR plates which contained sequence-specific primers suspended in each well (Figure 3.1 and Figure 3.2). Each plate also contains an inter-plate calibrator in triplicate and an empty negative control. For two entire 384-well qPCR plates, the following mastermix was synthesised:

2x SYBR® Green master mix	4000 µL
Nuclease-free water	3920 µL
cDNA template	80 µL
Total volume	8000 µL

Each well was added with 10 µL of mastermix and the plate was centrifuged briefly for 1500 x g for 1 minute after applying the optical sealing. The qRT-PCR cycling was carried out using a LightCycler® 480 Real-Time PCR System(Roche). PCR reactions were initiated with a 10-minute denaturation step at 95°C followed by 45 amplification cycles of 95°C for 10 seconds and 60°C for 60 seconds with ramp-rate of 1.6°C/s.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
A	hsa-miR-379	hsa-miR-217	hsa-miR-337-5p	hsa-miR-328	hsa-miR-374b*	hsa-miR-143	hsa-miR-623	hsa-miR-520c-3p	hsa-miR-557	hsa-miR-218	hsa-miR-136	hsa-miR-127-5p	hsa-miR-140-5p	hsa-miR-31*	hsa-miR-20b*	hsa-miR-325	hsa-miR-509-3-5p	hsa-miR-210	hsa-miR-199b-5p	hsa-miR-194	hsa-let-7g	hsa-miR-203	hsa-miR-181a*	hsa-miR-934
B	hsa-miR-551b	UniSp6 CP	hsa-miR-524-3p	hsa-miR-7	hsa-miR-486-5p	hsa-miR-30c	hsa-miR-301b	hsa-miR-128	hsa-miR-329	hsa-miR-224	hsa-miR-487b	hsa-miR-130a	hsa-miR-138	hsa-miR-26a-2*	hsa-miR-378	hsa-miR-381	hsa-miR-671-5p	hsa-miR-521	hsa-miR-221	hsa-miR-142-5p	hsa-miR-132	hsa-miR-424	UniSp3 IPC	hsa-miR-374a
C	hsa-miR-532-5p	hsa-miR-99a	hsa-miR-92a-1*	hsa-miR-125b	hsa-miR-185	hsa-miR-25	hsa-miR-524-5p	hsa-miR-20a	hsa-miR-765	hsa-miR-24	hsa-miR-369-5p	hsa-miR-425	hsa-miR-590-5p	hsa-miR-760	hsa-miR-574-3p	hsa-miR-130b	hsa-let-7e	hsa-miR-133b	hsa-miR-542-5p	hsa-miR-23a	hsa-miR-193b	UniSp3 IPC	hsa-miR-518c*	hsa-miR-204
D	hsa-miR-933	SNORD49A	hsa-miR-452	hsa-miR-215	hsa-miR-141	hsa-miR-374b	hsa-miR-668	hsa-miR-33a	hsa-miR-101	hsa-miR-30c-2*	hsa-miR-331-3p	hsa-miR-340	hsa-miR-196a	hsa-miR-888	hsa-miR-330-3p	hsa-miR-570	hsa-miR-518c	hsa-miR-200a	hsa-miR-188-5p	hsa-miR-26a	hsa-miR-99b	hsa-miR-431	hsa-miR-23b	hsa-miR-367
E	hsa-miR-505	hsa-miR-18a	hsa-miR-92a	hsa-miR-500a	hsa-miR-887	hsa-miR-491-3p	hsa-miR-423-3p	hsa-miR-126	hsa-miR-622	hsa-miR-376b	hsa-miR-302c	hsa-miR-185*	hsa-miR-339-5p	hsa-miR-873	hsa-miR-323-3p	hsa-miR-181d	hsa-miR-125a-5p	hsa-miR-129-5p	hsa-miR-492	hsa-miR-519d	UniSp3 IPC	hsa-miR-302d	hsa-miR-346	hsa-miR-151-3p
F	hsa-miR-493	hsa-miR-423-5p	hsa-miR-99a*	hsa-miR-10a	hsa-miR-202	hsa-miR-10b	hsa-miR-503	hsa-miR-890	hsa-miR-30d	hsa-miR-514	hsa-miR-16	hsa-miR-150	hsa-miR-654-5p	hsa-miR-545	hsa-miR-29b-2*	hsa-miR-491-5p	hsa-miR-92b	hsa-miR-665	hsa-miR-506	hsa-miR-363	hsa-miR-663	hsa-miR-651	hsa-miR-342-3p	hsa-miR-432
G	hsa-miR-154*	hsa-miR-27a	hsa-miR-376c	hsa-miR-940	hsa-miR-22*	hsa-miR-34c-5p	hsa-miR-885-5p	hsa-miR-320a	hsa-miR-18b	hsa-miR-187	hsa-miR-516b	hsa-miR-302c*	hsa-miR-548b-3p	hsa-miR-186	hsa-miR-199a-5p	hsa-miR-155	hsa-miR-107	hsa-miR-302b	hsa-miR-662	hsa-miR-30a	hsa-miR-302d*	hsa-miR-484	hsa-miR-337-3p	hsa-miR-494
H	hsa-miR-371-3p	hsa-miR-103	hsa-miR-144	hsa-miR-184	hsa-miR-631	hsa-miR-519a	hsa-miR-211	hsa-miR-802	hsa-let-7f	hsa-miR-625*	hsa-miR-34a	hsa-miR-744	hsa-miR-518e	hsa-miR-29b	hsa-miR-658	hsa-miR-572	hsa-let-7a	hsa-miR-30e	hsa-miR-433	hsa-miR-660	hsa-let-7c	hsa-miR-28-5p	hsa-miR-324-5p	hsa-miR-219-5p
I	hsa-miR-19b	hsa-miR-526b	hsa-miR-720	hsa-miR-30b	hsa-miR-637	hsa-miR-422a	hsa-miR-199a-3p	hsa-miR-335	hsa-miR-134	hsa-miR-21	hsa-miR-129-3p	hsa-miR-26b	hsa-miR-214	hsa-miR-32	hsa-miR-324-3p	hsa-miR-488	hsa-miR-371-5p	hsa-miR-455-5p	hsa-miR-891a	hsa-miR-549	hsa-miR-205	hsa-miR-518b	hsa-miR-361-5p	hsa-miR-454
J	hsa-miR-15a	hsa-miR-191	hsa-miR-608	hsa-miR-576-5p	hsa-miR-497	hsa-miR-19a	hsa-miR-187*	hsa-miR-620	hsa-let-7i	hsa-miR-501-5p	hsa-miR-652	hsa-miR-1979	hsa-miR-30e*	hsa-miR-181c	hsa-miR-499-5p	hsa-miR-548c-3p	hsa-miR-152	hsa-miR-93	hsa-miR-490-3p	hsa-miR-29c	hsa-miR-372	hsa-miR-133a	hsa-miR-124	hsa-miR-190
K	hsa-miR-302a	hsa-miR-595	hsa-miR-602	hsa-miR-223	hsa-miR-627	hsa-miR-34b	hsa-miR-410	hsa-miR-17	hsa-miR-376a	hsa-miR-877	hsa-miR-512-5p	hsa-miR-449a	hsa-miR-498	hsa-miR-148b	hsa-miR-127-3p	hsa-miR-598	hsa-miR-96	hsa-let-7d	hsa-miR-135b	hsa-miR-495	hsa-miR-299-5p	hsa-miR-34c-3p	hsa-miR-596	hsa-miR-126*
L	hsa-miR-145	SNORD38B	hsa-miR-516a-5p	hsa-miR-421	hsa-miR-96*	hsa-miR-362-5p	hsa-miR-615-3p	hsa-miR-550a	hsa-miR-766	hsa-miR-200b	hsa-miR-298	hsa-miR-193a-5p	hsa-miR-449b	hsa-miR-520d-5p	hsa-miR-192	hsa-miR-29a	hsa-miR-18a*	hsa-miR-383	hsa-miR-9	hsa-miR-202*	hsa-miR-363*	hsa-miR-147b	hsa-miR-197	hsa-miR-597
M	hsa-miR-326	hsa-miR-15b	hsa-miR-105	hsa-miR-196b	hsa-miR-296-5p	hsa-miR-20b	hsa-miR-147	hsa-miR-198	hsa-miR-375	hsa-miR-517a	hsa-miR-361-3p	hsa-miR-21*	hsa-miR-220a	hsa-miR-518f	hsa-miR-222	hsa-miR-617	hsa-miR-154	hsa-miR-708	hsa-let-7b	hsa-miR-95	hsa-miR-517c	hsa-miR-151-5p	hsa-miR-502-5p	hsa-miR-345
N	hsa-miR-509-3p	U6	hsa-miR-382	hsa-miR-373	hsa-miR-200c	hsa-miR-9*	hsa-miR-181b	hsa-miR-628-3p	hsa-miR-195	hsa-miR-183	hsa-miR-135a	hsa-miR-30b*	hsa-miR-146b-5p	hsa-miR-301a	hsa-miR-1	hsa-miR-299-3p	hsa-miR-142-3p	hsa-miR-338-3p	hsa-miR-584	hsa-miR-377	hsa-miR-216a	hsa-miR-206	hsa-miR-921	hsa-miR-513a-5p
O	hsa-miR-140-3p	hsa-miR-181a	hsa-miR-122	hsa-miR-106a	hsa-miR-182	hsa-miR-370	hsa-let-7d*	hsa-miR-425*	hsa-miR-450a	hsa-miR-411	hsa-miR-216b	hsa-miR-106b	hsa-miR-886-3p	hsa-miR-510	hsa-miR-212	hsa-miR-525-5p	hsa-miR-589	hsa-miR-576-3p	hsa-miR-583	hsa-miR-483-3p	hsa-miR-582-5p	hsa-miR-886-5p	hsa-miR-33b	hsa-miR-193a-3p
P	hsa-miR-153	Blank (H20)	hsa-miR-409-3p	hsa-miR-22	hsa-miR-629	hsa-miR-365	hsa-miR-429	hsa-miR-98	hsa-miR-518a-3p	hsa-miR-137	hsa-miR-508-3p	hsa-miR-539	hsa-miR-148a	hsa-miR-146a	hsa-miR-139-5p	hsa-miR-373*	hsa-miR-149	hsa-miR-642a	hsa-miR-31	hsa-miR-451	hsa-miR-100	hsa-miR-27b	hsa-miR-523	Blank (H20)

Figure 0.1: miRNA Ready-to-Use PCR, Human panel I array plate for biomarker discovery phase

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
A	hsa-miR-499-3p	hsa-miR-219-1-3p	hsa-miR-543	hsa-miR-1245	hsa-miR-522	hsa-miR-571	hsa-miR-323-5p	hsa-miR-592	hsa-miR-487a	hsa-miR-1249	hsa-miR-25*	hsa-miR-922	hsa-miR-124*	hsa-miR-1264	hsa-miR-504	hsa-miR-138-1*	hsa-miR-502-3p	hsa-miR-490-5p	hsa-miR-567	hsa-miR-18b*	hsa-miR-125a-3p	hsa-miRPlus-A1027	hsa-miR-129*	hsa-miR-148a*
B	hsa-miR-412	Blank (H20)	hsa-miRPlus-C1066	hsa-miR-548o	hsa-miR-143*	hsa-miR-513a-3p	hsa-miR-943	hsa-miR-196b*	hsa-miR-200b*	hsa-miR-551b*	hsa-miR-942	hsa-miR-515-3p	hsa-miRPlus-C1076	hsa-miR-653	hsa-miR-24-2*	hsa-miR-1539	hsa-miR-27b*	hsa-miR-1909	hsa-miR-770-5p	hsa-miR-548l	hsa-miR-376a*	hsa-miR-1247	hsa-miR-520b	hsa-miR-936
C	hsa-miR-28-3p	hsa-miR-875-5p	hsa-miR-551a	hsa-miR-1183	hsa-miR-758	hsa-miR-1244	hsa-miR-566	hsa-miR-1256	hsa-miR-516b*	hsa-miR-548c-5p	hsa-miR-496	hsa-miR-876-3p	hsa-miR-532-3p	hsa-miR-654-3p	hsa-miR-659	hsa-miR-135b*	hsa-miR-641	hsa-miR-616*	hsa-miR-489	hsa-miR-335*	hsa-miR-892a	hsa-miR-10b*	hsa-miR-122*	hsa-miR-100*
D	hsa-miR-769-3p	UniSp6 CP	hsa-miR-519c-5p	hsa-miR-938	hsa-miR-937	hsa-miR-1911*	hsa-miR-640	(hsa-miR-220b)	hsa-miR-29c*	hsa-miR-449b*	hsa-let-7f-1*	hsa-miR-297	hsa-miRPlus-A1031	hsa-miR-144*	hsa-miR-106a*	hsa-miR-1254	hsa-miR-580	hsa-miR-1252	hsa-miR-643	hsa-miR-30c-1*	hsa-miR-221*	hsa-miR-103-as	hsa-miR-411*	hsa-miR-519e*
E	hsa-miR-338-5p	hsa-miRPlus-C1089	hsa-miR-323b-5p	hsa-miR-548i	hsa-miR-541	hsa-miR-1272	hsa-miR-1205	hsa-miR-1266	hsa-miR-431*	hsa-miR-621	hsa-miR-556-5p	hsa-miR-1267	hsa-miR-141*	hsa-miR-1269	hsa-miR-501-3p	hsa-miR-15b*	hsa-miR-146b-3p	hsa-miR-222*	hsa-miR-601	hsa-miR-924	hsa-miR-29a*	hsa-let-7a-2*	hsa-miR-520f	hsa-miR-101*
F	hsa-miR-520a-3p	hsa-miR-423-5p	hsa-miR-517*	hsa-miR-380	hsa-miR-1296	hsa-miR-2053	hsa-miR-920	hsa-miR-1471	hsa-miR-19b-2*	hsa-miR-10a*	hsa-miR-106b*	hsa-miR-365*	hsa-miR-619	hsa-miR-518f*	hsa-miR-17*	hsa-miR-2113	hsa-miR-877*	hsa-miR-362-3p	hsa-miR-3180-3p	hsa-miR-7-1*	hsa-miR-767-5p	hsa-miR-562	hsa-miRPlus-D1033	hsa-miR-300
G	hsa-miR-604	hsa-miR-130b*	hsa-miR-149*	hsa-miR-1271	hsa-miR-520h	hsa-miR-769-5p	hsa-miR-612	hsa-miR-1237	hsa-miR-527	hsa-miR-1260	hsa-miR-182*	hsa-miR-632	hsa-miR-508-5p	hsa-miR-671-3p	hsa-miR-941	hsa-miR-23b*	hsa-miR-591	hsa-miR-26b*	hsa-miR-519b-3p	hsa-miR-30d*	hsa-miR-526a	hsa-miR-675b	hsa-miR-520e	hsa-miR-646
H	hsa-miR-519e	hsa-miR-103	hsa-miR-26a-1*	hsa-miR-1908	hsa-miRPlus-D1061	hsa-miR-103-2*	hsa-miR-586	hsa-miR-1179	hsa-miR-488*	hsa-miR-208b	hsa-miR-192*	hsa-miR-588	hsa-miR-181a-2*	hsa-miR-1914	hsa-miR-573	hsa-miR-548d-5p	hsa-miR-194*	hsa-miR-1972	hsa-miR-556-3p	hsa-miR-664	hsa-let-7b*	hsa-miR-195*	hsa-miR-125b-1*	Blank (H20)
I	hsa-miR-519c-3p	hsa-miR-554	hsa-miR-558	hsa-miR-1243	(hsa-miR-1201)	hsa-miR-1185	hsa-miR-512-3p	hsa-miR-587	hsa-miR-603	hsa-miR-1184	hsa-miR-20a*	hsa-miR-1263	hsa-miR-455-3p	hsa-miR-582-3p	hsa-miR-409-5p	hsa-miR-452*	hsa-miR-19b-1*	hsa-miR-610	hsa-miR-511	hsa-miR-200c*	hsa-let-7a*	hsa-miR-135a*	hsa-miR-520a-5p	hsa-miR-1468
J	hsa-miRPlus-C1100	hsa-miR-191	hsa-miR-145*	hsa-miR-548h	hsa-miR-378*	hsa-miR-553	hsa-miR-181c*	hsa-miR-624*	hsa-miR-578	hsa-miR-448	hsa-miR-875-3p	hsa-miR-661	hsa-miR-876-5p	hsa-miR-1913	(hsa-miR-220c)	hsa-miR-218-1*	hsa-miR-744*	hsa-miR-548m	hsa-miR-138-2*	hsa-miR-561	hsa-miR-99b*	hsa-miR-23a*	hsa-miR-191*	Blank (H20)
K	hsa-miR-1204	hsa-miR-548j	hsa-miR-555	hsa-miR-1224-3p	hsa-miR-649	hsa-miR-663b	hsa-miR-634	hsa-miR-1248	hsa-miR-889	hsa-miR-1227	hsa-miR-517b	hsa-miR-1255b	hsa-miR-330-5p	hsa-miR-1238	hsa-miR-188-3p	hsa-miR-589*	hsa-miR-125b-2*	hsa-miR-16-2*	hsa-miR-515-5p	hsa-miR-340*	UniSp3 IPC	hsa-miR-34a*	hsa-miR-342-5p	hsa-miR-639
L	hsa-let-7i*	SNORD38B	hsa-miR-645	(hsa-miR-1259)	hsa-miR-33a*	hsa-miR-609	hsa-miR-34b*	hsa-miR-579	hsa-miR-379*	hsa-miR-1914*	hsa-miR-614	(hsa-miR-1974)	hsa-miR-93*	hsa-miR-219-2-3p	hsa-miR-616	hsa-miR-1912	hsa-miR-2110	hsa-miR-626	hsa-miR-200a*	hsa-miR-432*	hsa-miR-636	hsa-miR-888*	hsa-miR-450b-3p	Blank (H20)
M	hsa-miR-520d-3p	hsa-miR-1265	hsa-miR-1203	hsa-miR-548k	hsa-miR-548a-5p	hsa-miR-1253	hsa-miR-615-5p	hsa-miR-1236	hsa-miR-1208	hsa-miR-302e	hsa-miR-1206	hsa-miR-1270	hsa-miR-525-3p	hsa-miR-1200	hsa-miR-29b-1*	hsa-miR-33b*	hsa-miR-223*	hsa-miR-380*	hsa-miR-891b	hsa-miR-424*	hsa-miR-339-3p	UniSp3 IPC	hsa-miR-647	hsa-miR-183*
N	hsa-miR-92b*	U6	hsa-miR-146a*	hsa-miR-1911	hsa-miR-593*	hsa-miR-628-5p	hsa-miR-767-3p	hsa-miR-518d-3p	hsa-miR-24-1*	hsa-miR-544	hsa-miR-193b*	hsa-miR-577	hsa-miR-541*	hsa-miR-208a	hsa-miR-629*	hsa-miR-1207-5p	hsa-miR-630	hsa-miR-384	hsa-miR-885-3p	hsa-miR-675*	hsa-let-7g*	hsa-miR-526b*	UniSp3 IPC	Blank (H20)
O	hsa-miR-552	hsa-miR-513c	hsa-miR-331-5p	hsa-miR-1182	hsa-miR-611	hsa-miR-1181	hsa-miR-638	hsa-miR-1258	hsa-miR-650	hsa-miR-1178	hsa-miR-600	hsa-miR-599	hsa-miR-520g	hsa-miR-564	hsa-miR-132*	hsa-miR-450b-5p	hsa-miR-485-3p	hsa-let-7f-2*	hsa-miR-155*	hsa-miR-105*	hsa-miR-486-3p	hsa-miR-320b	hsa-miR-296-3p	hsa-miR-7-2*
P	hsa-miR-550a*	SNORD49A	hsa-miR-593	hsa-miR-605	hsa-miR-493*	hsa-miR-1538	hsa-miR-454*	hsa-miR-224*	hsa-miR-30a*	hsa-miR-607	hsa-miR-214*	hsa-miR-190b	hsa-miR-505*	hsa-miR-708*	hsa-miRPlus-C1070	hsa-miR-302a*	hsa-miR-130a*	hsa-miR-1537	hsa-let-7c*	hsa-miR-148b*	hsa-miR-15a*	Blank (H20)	hsa-miR-92a-2*	Blank (H20)

Figure 0.2: miRNA Ready-to-Use PCR, Human panel II array plate for biomarker discovery phase

3.3.4.3 Quality Control

After each PCR run, melting curve analysis was performed to assess the melting temperature (T_m) of the PCR amplicon to verify the specificity of the amplification reaction. PCR reactions that gave rise to multiple melting curve peaks or single peaks with T_m incoherent with the specifications for the corresponding assay were flagged and removed from the dataset. No template control was included in the plate as negative control to detect RNA contamination in the reverse transcription step. Only assays with 5 Cq's less than the negative control, and with Cq<37 were included in the data analysis. Endogenous miRNA assays were carried out to assess the quality of RNA isolated from cell-free serum. To ensure the absence of PCR inhibition, cDNA synthesis control (UniSp6) was added during the reaction to evaluate the efficiency of RT whereby a raw Cq value of more than 22 indicates inhibition of cDNA synthesis and reverse transcription would be repeated. To assess haemolysis the level of miR-451 (miRNA with high abundance in red blood cells) and miR-23a (miRNA with relatively stable expression in serum independent of hemolysis) were assessed in RNA samples. Samples with ΔCq (miR-23a-3p – miR-451a) above 8.0 indicate risk of hemolysis and were excluded from subsequent analysis (Blondal et al., 2013).

3.3.4.4 Relative quantitation of miRNA expression

Analysis of the amplification curves was performed using Roche LC software in the determination of Cq (by the second derivative method) and for melting curve analysis. qRT-PCR results were exported to GenEx (MultiD Analyses) and normalised using inter-plate calibrators on the miRNA Ready-to-Use panels. Data normalization was performed based on the global mean expression, as suggested by the NormFinder software (Andersen, Jensen, & Orntoft, 2004; Mestdagh et al., 2009). NormFinder was

also used to identify the reference miRNA which yields the lowest stability value to be selected as the normalizer for Phase II. ΔCq was calculated using the following mathematical formula: ΔCq=Cq average - Cq sample. ΔΔCq was then calculated by subtracting ΔCq of the test group from ΔCq of reference. The fold change of miRNA gene expression was calculated by the equation $2^{-\Delta\Delta Cq}$.. The miRNA values represented in figures and table are log transformed to assume normal distribution.

3.3.5 Phase II (Biomarker Validation Phase)

Total RNA was reverse transcribed in 20 µL reaction volume using the miRCURY LNA™ Universal RT cDNA synthesis kit (Exiqon). Following dilution, the cDNA template was subjected to miRNA profiling by quantitative real-time PCR (qRT-PCR) on a total of 13 human miRNAs that were selected from Phase I based on respective fold change and literature review.

3.3.5.1 Reverse transcription for cDNA synthesis

Fixed volumes of eluted RNA (4 µL) were reverse transcripted with UniSp6 RNA template added as synthetic spike in to be used for subsequent quality control steps (Blondal et al., 2013). Total RNA from serum was reverse transcribed using miRCURY LNA™ Universal RT cDNA synthesis kit (Exiqon) as follows:

5x Reaction buffer	4 µL
Nuclease-free water	9 µL
Enzyme mix	2 µL
UniSp6 RNA template (spike in)	1 µL
Total RNA template	4 µL
Total volume	20 µL

In thermal cycler, the mixture was incubated for 60 minutes at 42ºC and heat-inactivated for 5 minutes at 95ºC followed by immediate cooling at 4ºC and stored at -80ºC.

3.3.5.2 Quantitative real-time PCR (qRT-PCR) reaction

The qRT-PCR array profiling was performed for each sample in 96-well qPCR plates using StepOnePlus™ Real-Time PCR System (Applied Biosystem). Inter-plate control calibrator was included on each plate and all reactions were performed in triplicate. MiR-24-3p was used as endogenous control for quantitative PCR analysis of miRNA expression given the relative stability of its expression in all samples as demonstrated by the NormFinder during Phase I of the study which was also used as normalizer in other study (Marabita et al., 2015). Each individual well on the qPCR plates contains the following mixture:

2x SYBR® Green master mix	5 µL
PCR primer mix	1 µL
cDNA template	4 µL
Total volume	10 µL

PCR reactions were initiated with a 10-minute denaturation step at 95°C followed by 40 amplification cycles of 95°C for 10 seconds and 60°C for 60 seconds with ramp-rate of 1.6°C/s.

3.3.5.3 Quality Control

After each PCR run, melting curve analysis was performed to assess the T_m of the PCR amplicon to verify the specificity of the amplification reaction. PCR reactions that gave rise to multiple melting curve peaks or single peaks with T_m incoherent with the specifications for the corresponding assay were repeated. No template control was included in the plate as negative control to detect RNA contamination in the reverse transcription step. Endogenous miRNA assays were carried out to assess the quality of RNA isolated from cell-free serum. To ensure the absence of PCR inhibition, cDNA synthesis control (UniSp6) was added during the reaction to evaluate the efficiency of RT whereby a raw Cq value of more than 22 indicate inhibition of cDNA synthesis and RT would be repeated. To assess haemolysis the level of miR-451 (miRNA with high abundance in red blood cells) and miR-23a (miRNA with relatively stable expression in serum independent of hemolysis) were assessed in RNA samples. Samples with ΔCq (miR-23a-3p – miR-451a) above 8.0 indicates risk of hemolysis and was excluded from subsequent analysis.

3.3.5.4 Relative quantitation of miRNA expression

The qRT-PCR results were exported to GenEx (MultiD Analyses AB, Goteborg, Sweden) and normalised using inter-plate calibrators. MiRNA expression data were normalised against the expression of miR-24-3p, as suggested by the NormFinder software during Phase I. Relative quantification was calculated using the following mathematical formula: ΔCq=Cq average - Cq sample. ΔΔCq was then calculated by subtracting ΔCq of the testgroup from ΔCq of reference. The fold change of miRNA gene expression was calculated by the equation $2^{-\Delta\Delta Cq}$ (Schmittgen

&Livak, 2008). The expression data was converted to a logarithmic scale for graphical representations and further statistical analyses(Pacifici, Delbue, Kadri, & Peruzzi, 2014)

3.3.6 Statistical analysis

All demographic and clinical parameters were presented as mean ± standard deviation for continuous data and as percentages for categorical data. Categorical data were compared using Chi-square (χ^2) test. For two groups comparisons, independent t-test was performed to compare normally distributed variables while Mann-Whitney U test was employed for skewed variables. Analysis of Variance (ANOVA) and Kruskal-Wallis tests were conducted for comparison of means and medians between three groups for normally and non-normally distributed variables, respectively.

Statistical tests were performed using GenEx version 6 (MultiD Analyses) and SPSS version 16.0 (IBM). The qPCR data were log-transformed and analysed. Relative expression levels of miRNAs were compared between two groups using independent t-test followed by strict Benjamini-Hochberg correction for multiple testing to control for false discovery rate (FDR) (Benjamini & Hochberg, 1995b). The predictability of surrogate marker in the diagnosis of NAFLD was depicted in a receiver operating characteristic (ROC) curve. For the evaluation of diagnostic performance of selected miRNAs, area under ROC curve (AUROC) and 95% confidence intervals (CI) were calculated. Multivariate ROC analysis was carried out via logistic regression to examine the diagnostic performance of combined miRNAs in predicting the severity of NAFLD. Correlation between level of miRNAs with liver enzymes, BMI, serum triglycerides and age were assessed using linear regression and Spearman's correlation test, for normally distributed variable and skewed variables, respectively.

3.4 Results

3.4.1 Patient characteristics

The demographics and clinical parameters of healthy volunteers (n=10) and NASH patients (n=9) in the screening cohort are showed in Table 3.1. The subjects were matched for age and sex. Among the NASH patients, two had stage 1 fibrosis, four had stage 2 fibrosis, two had stage 3 fibrosis, and one had stage 4 fibrosis. Table 3.2 shows the demographic and clinical parameters of healthy subjects (n=50) and NAFLD patients (n=122) in the validation cohort. The NAFLD patients in the validation cohort were further stratified according to the severity of the disease: non-alcoholic fatty liver (NAFL) (n=25) and non-alcoholic steatohepatitis (NASH) (n=97) with the demographics and clinical parameters are shown in Table 3.3. Among the NASH patients, 14 patients had no fibrosis, 53 patients have stage 1 fibrosis, six patients had stage 2 fibrosis, 23 patients had stage 3 fibrosis and one patient had stage 4 fibrosis.

Table 0.1: Demographic and clinical data of the subjects in screening phase

Characteristics	***n* (%) or Mean ± SD**		***P* value**
	Control (*n*=10)	**NASH (*n*=9)**	
Gender			
Males	4 (40)	4 (44)	0.845
Females	6 (60)	5 (56)	
Age (years)	30.5±2.3	32.5±2.9	0.375
BMI (kg/m^2)	21.2±2.0	31.3±2.9	<0.0001
HbA1c (%)	5.4±0.2	6.5±1.9	<0.0001
HDL cholesterol (mg/dl)	64.9±13.0	43.2±7.9	<0.0001
LDL cholesterol (mg/dl)	110.4±21.9	130.2±53.0	<0.0001
Total cholesterol (mg/dl)	186.9±29.5	203.8±60.6	<0.0001
Triglycerides (mg/dl)	58.9±16.8	193.4±50.6	<0.0001
AST (IU/L)	21.6±3.4	63.2±18.3	<0.0001
ALT (IU/L)	21.9±9.8	129.4±53.0	<0.0001
GGT (IU/L)	19.1±7.4	126.3±53.5	<0.0001

Data are expressed in mean ± SD for continuous data and percentage for categorical data.

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

Table 0.2: Demographic and clinical data of the subjects in validation phase

Characteristics	*n* (%) or Mean ± SD		*P* value
	Control (*n*=50)	NAFLD (*n*=122)	
Gender			
Males	24 (48)	61 (50)	0.812
Females	26 (52)	61 (50)	
Age (years)	47.7±9.8	50.1±11.4	0.089
BMI (kg/m^2)	21.1±2.5	29.3±4.1	<0.001
HbA1c (%)	5.6±0.3	6.3±1.2	<0.001
HDL cholesterol (mg/dl)	62.1±18.0	46.0±9.0	<0.001
LDL cholesterol (mg/dl)	122.9±30.2	116.0±40.3	0.225
Total cholesterol (mg/dl)	195.9±31.6	192.4±43.4	0.557
Triglycerides (mg/dl)	80.2±35.0	150.9±65.5	<0.001
AST (IU/L)	22.0±5.9	51.0±28.0	<0.001
ALT (IU/L)	23.5±13.0	82.0±46.6	<0.001
GGT (IU/L)	23.2±14.6	106.9±96.5	<0.001

Data are expressed in mean ± SD for continuous data and percentage for categorical data.

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

Table 0.3: Demographic and clinical data of the patients in validation phase

Characteristics	*n* (%) or Mean ± SD		*P* value
	NAFL (*n*=25)	NASH (*n*=97)	
Gender			
Males	18 (72)	43 (44)	0.014
Females	7 (28)	54 (56)	
Age (years)	47.3±10.9	50.8±11.5	0.161
BMI (kg/m^2)	27.1±3.2	29.2±5.3	0.003
HbA1c (%)	6.2±1.5	6.4±1.2	0.250
Waist circumference (cm)	93.7±6.7	97.6±10.2	0.104
HDL cholesterol (mg/dl)	43.8±8.0	46.6±9.2	0.169
LDL cholesterol (mg/dl)	122.0±41.9	114.5±40.0	0.405
Total cholesterol (mg/dl)	206.2±34.5	188.9±44.9	0.075
Triglycerides (mg/dl)	182.1±93.1	142.9±54.0	0.008
AST (IU/L)	34.2±13.6	55.4±29.1	0.001
ALT (IU/L)	62.8±87.0	33.2±48.3	0.035
GGT (IU/L)	114.1±143.7	105.0±81.0	0.192

Data are expressed in mean ± SD for continuous data and percentage for categorical data.

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; *NAFL,* non-alcoholic fatty liver; *NASH,* non-alcoholic steatohepatitis;

3.4.2 MiRNA differential expression in discovery and validation phase

The differential expression of circulating miRNA profile in NASH patients and healthy controls was investigated. Out of the total 752 miRNAs assessed in sera from nine NASH patients and 10 healthy subjects (Phase I), a distinct expression pattern was observed after strict Benjamini-Hochberg correction for multiple testing. Table 3.4 shows a list of 14 significantly increased miRNAs levels and five significantly decreasedmiRNAs levels (P_c< 0.05) with respective fold change. The most remarkable increases were demonstrated by miR-122-5p, miR-193b-3p, and miR-885-5p (fold

change of 13.55, 11.19 and 7.47 respectively) in the NASH patients compared to healthy controls.

Out of the 19 miRNAs, 12 miRNAs that were increased by fold change >2 (miR-122-5p, miR-193b-3p, miR-885-5p, miR-34a-5p, miR-125b-5p, miR-194-5p, miR-215, miR-99a-5p, miR-192-5p, miR-505-3p, miR-148a-3p, miR-378a-3p) and one miRNA decreasedby fold change < 0.30 (miR-103a-3p) were chosen for biomarker validation in phase II.A hierarchical clustering of the 19 miRNAs with the highest differential expression between NASH and controls is shown in Figure 3.3.

Table 0.4: MiRNAs significantly increased and decreased (P_c< 0.05) in NASH serum samples compared to healthy controls.

MiRNA	*P* value	P_c value	Fold change
miR-122-5p	$3.02x10^{-5}$	0.003	13.55
miR-193b-3p	$1.61x10^{-4}$	0.005	11.19
miR-885-5p	$3.67 x10^{-3}$	0.037	7.47
miR-34a-5p	$1.61x10^{-3}$	0.028	5.35
miR-125b-5p	$9.01x10^{-5}$	0.004	5.20
miR-194-5p	$4.29x10^{-3}$	0.041	4.95
miR-215	$2.21x10^{-3}$	0.034	4.85
miR-99a-5p	$1.59x10^{-4}$	0.005	4.17
miR-192-5p	$3.74x10^{-4}$	0.009	4.09
miR-22-5p	$3.48x10^{-3}$	0.040	3.25
miR-505-3p	$6.91x10^{-5}$	0.004	3.21
miR-148a-3p	$2.02x10^{-5}$	0.003	3.06
miR-378a-3p	$2.22x10^{-3}$	0.032	2.97
miR-27b-3p	$8.27x10^{-4}$	0.018	1.70
miR-18a-5p	$3.58x10^{-3}$	0.038	0.57
miR-339-5p	$5.37x10^{-3}$	0.049	0.56
miR-151a-5p	$2.73x10^{-3}$	0.034	0.51
miR-199a-3p	$2.38x10^{-3}$	0.031	0.44
miR-103a-3p	$8.88x10^{-4}$	0.017	0.28

P_c, corrected *P* value using Benjamini-Hochberg multiple testing correction.Fold change < 1 indicates decreased level, fold change > 1 indicates increased level

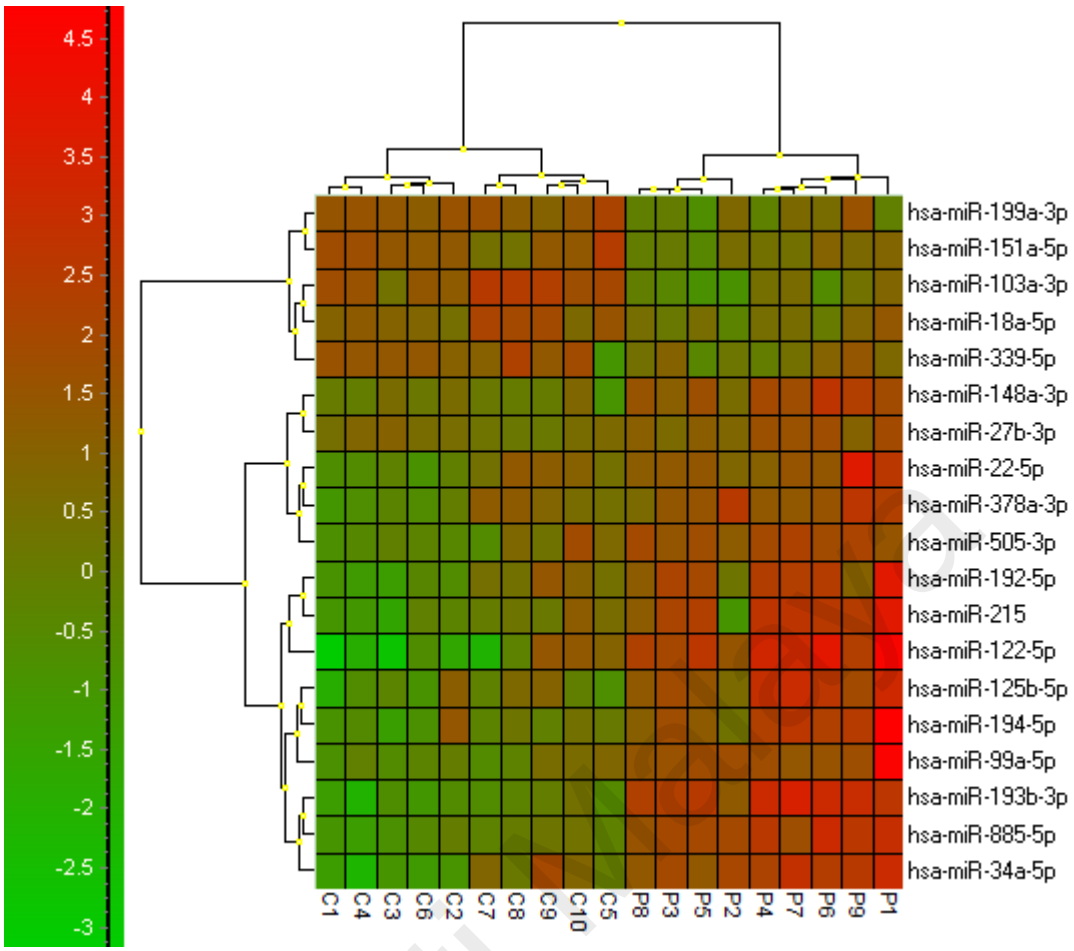
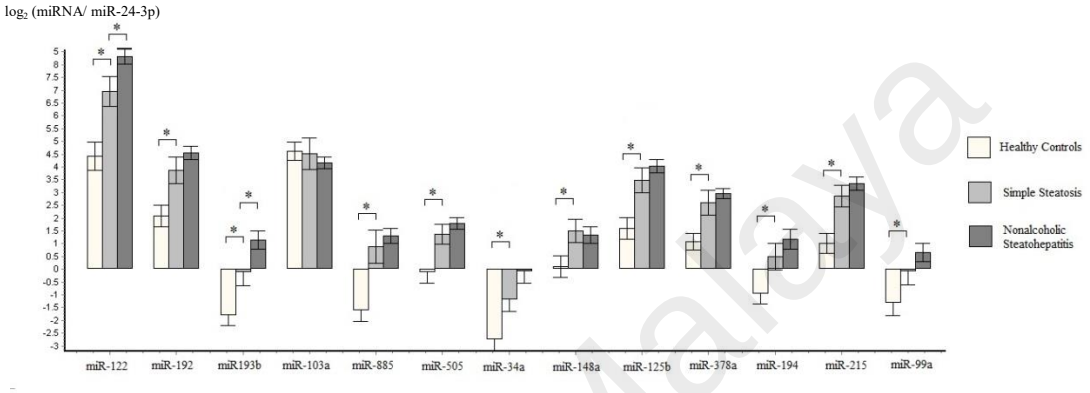


Figure 0.3: Heat-map and unsupervised hierarchical clustering of the 19 differentially expressed miRNAs in serum of 9 NASH patients as compared with 10 healthy control subjects. Red colour represents an expression level above mean, green colour represents expression lower than the mean.

To examine whether a corresponding serum miRNA profile in NASH could be demonstrated in a larger sample of well-characterised patients and controls, a total of 13 markers identified from phase I were quantified in sera among 50 healthy subjects, 25 NAFL patients, and 97 NASH patients. Concordance expression was demonstrated in the validation set when the expression levels of all miRNAs were significantly altered in NASH compared to healthy controls ($P_c < 0.05$).

Despite the primary investigation in phase I was to discover biomarkers in NASH, it is also of interest to investigate if the miRNAs are found to be differentially expressed in patients with NAFL prior to the development of NASH. As demonstrated

in Figure 3.4, when the miRNAs expression level of NAFL was compared to healthy controls, 12 miRNAs were increased (P_c< 0.01) except for miR-103a-3p. In determining the miRNAs expressions with respect to different stages of NAFLD, only two miRNAs (miR-122-5p, miR-193b-3p) were significantly increased in NASH compared to NAFL (P_c< 0.01).



*P_c value < 0.01

Figure 0.4: qRT-PCR validation of increased and decreased miRNAs in healthy controls, non-alcoholic fatty liver (NAFL) and nonalcoholic steatohepatitis (NASH).

The association between the expression level of these miRNAs and the histological components in all the NAFLD patients (n=122) was further explored by grouping the independent variables according to degree of fatty infiltration, lobular inflammation, hepatocellular ballooning, and liver fibrosis (Figure 3.5). Significant upregulation of serum miR-122-5p, miR-193b-3p and miR-125b-5p were observed in patients with significant fibrosis (F ≥ 2) (n=30) compared to patients without significant fibrosis (F < 2) (n=92) (P_c< 0.001, P_c = 0.01 and P_c = 0.04, respectively). Increased levels of miR-122-5p and miR-193b-3p were observed in patients with hepatocellular ballooning (Ballooning ≥1) compared to patients without hepatocellular ballooning (Ballooning = 0) (P_c< 0.001 and P_c = 0.001, respectively). The expressions of

miR-122-5p and miR-193b-3p were also associated with higher degree of lobular inflammation (Inflammation ≥2 foci per x200 field) (P_c< 0.001). There was no differential expression of miRNA observed when the patients were grouped according to the degree of steatosis (≤ 33% fatty infiltration vs >33% fatty infiltration). Since the stage of fibrosis is of clinical importance, the differential expressions of miR-122-5p, miR-193b-3p and miR-125b-5p according to different stage of fibrosis are shown in Figure 3.6. Level of miR-122-5p increased steadily from stage 0 to stage 2 fibrosis but dropped slightly in stage 3 fibrosis. MiR-193b-3p was highly increased in stage 1 to 3 fibrosis compared to no fibrosis. Expression of miR-125b-5p was increased in significant fibrosis (fibrosis score ≥2) and decreased slightly in stage 3 fibrosis.

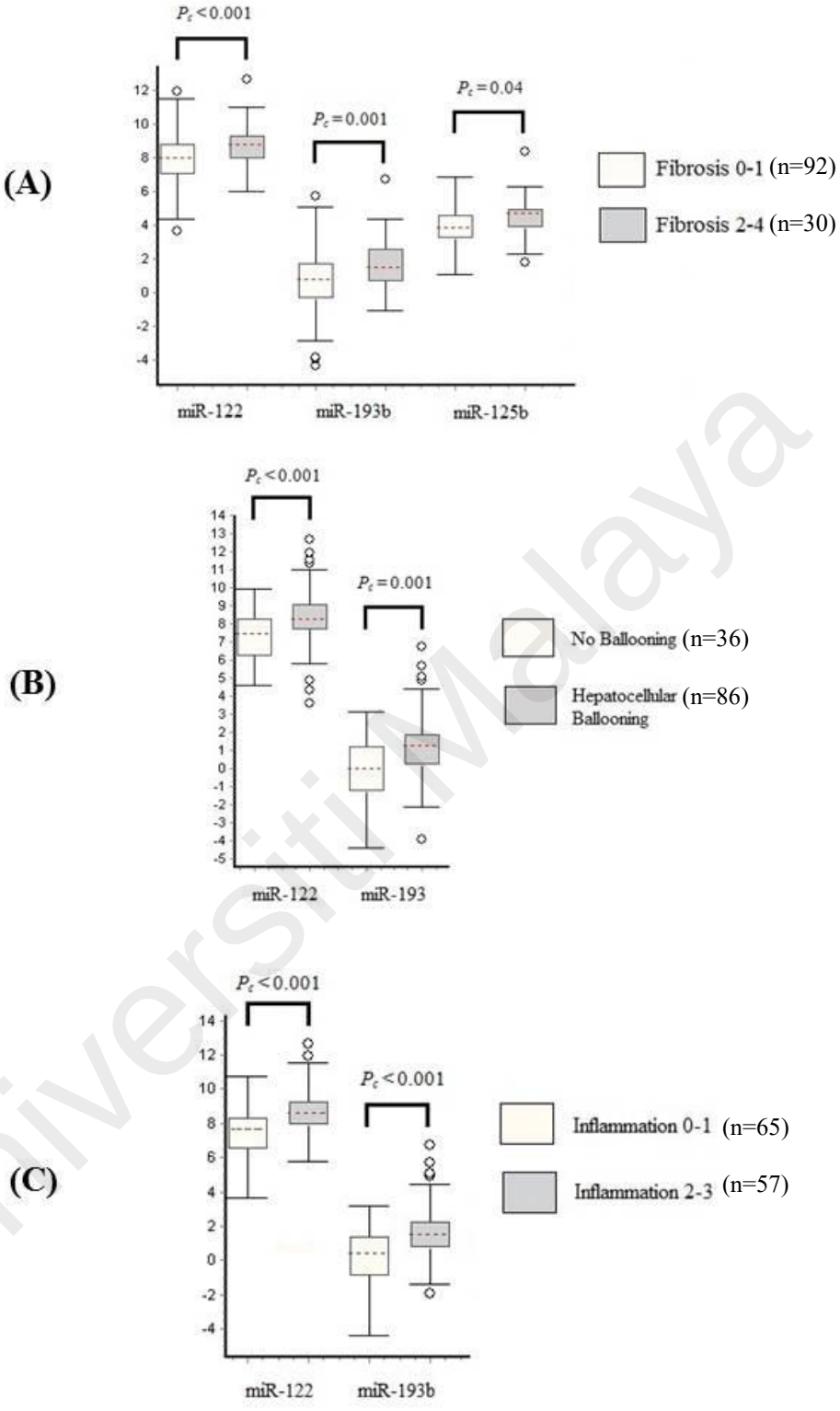


Figure 0.5: Comparison of serum levels of miRNAs in different stages of histological features in NAFLD patients. (A) Fibrosis; (B) Hepatocellular ballooning; (C) Lobular inflammation

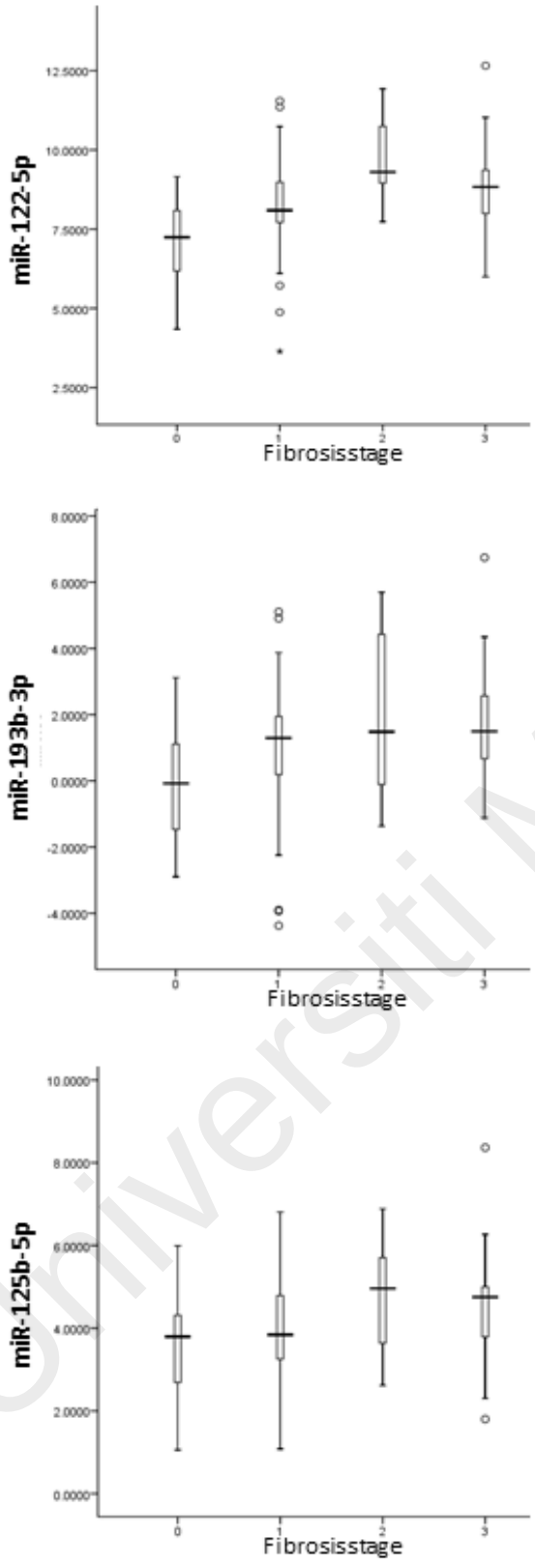
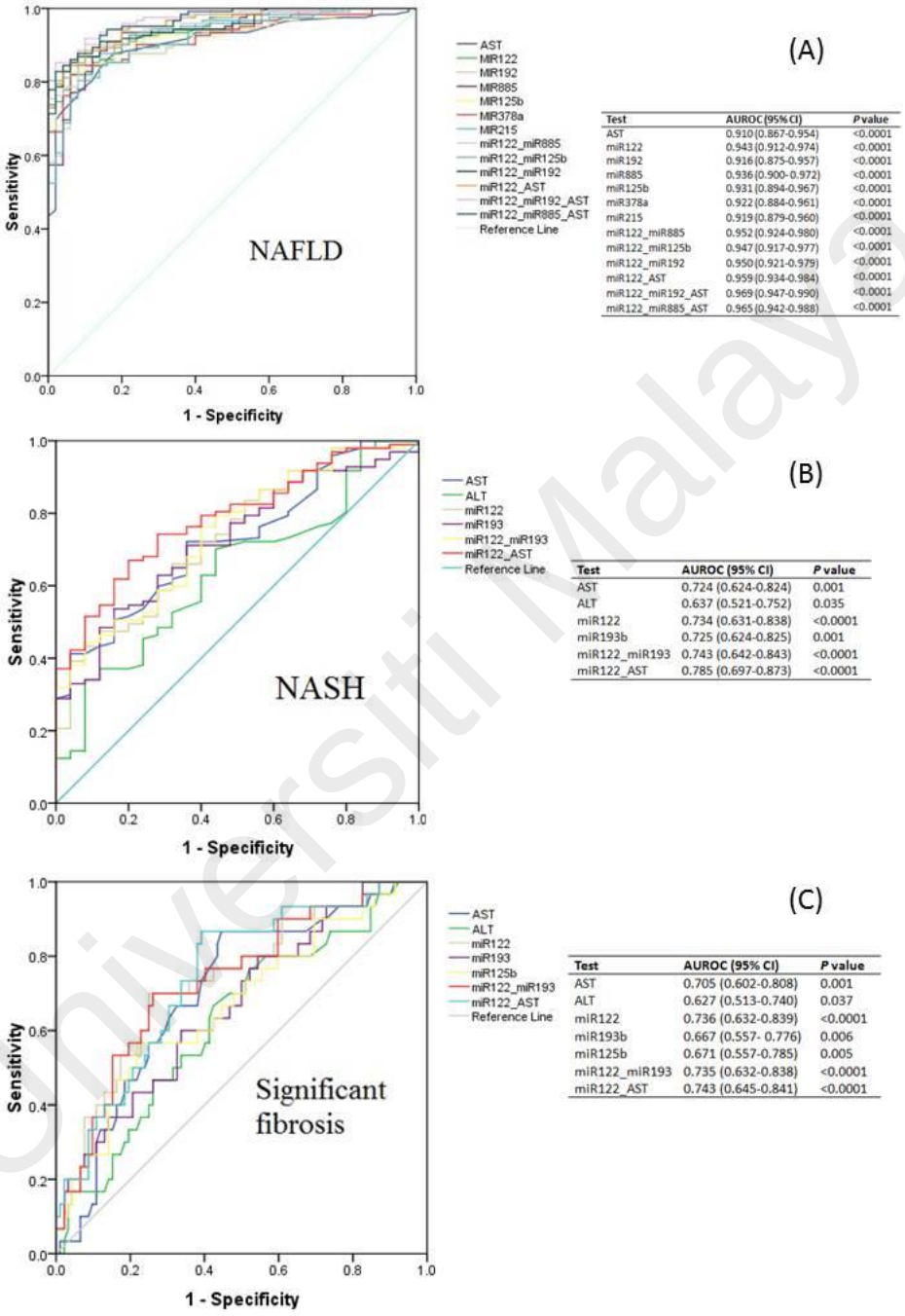


Figure 0.6: Comparison of serum levels of miRNAs in different stages of fibrosis: F0 (n=14), F1 (n=53), F2 (n=6), F3 (n=23)

Receiver operating characteristic (ROC) curves were used for the evaluation on the performance of these miRNAs in predicting the incidence of NAFLD, NASH and stage of fibrosis. The tested serum miRNAs provided excellent discrimination between NAFLD patients and healthy controls with area under ROC (AUROC) ranged from 0.823 to 0.943, most being superior to AST in distinguishing NAFLD. The miRNAs with AUROC higher than 0.9 for NAFLD are demonstrated in Figure 3.7. Multivariate ROC analysis showed increased AUROC (0.969) upon incorporation of miR-122-5p, miR-192-5p and AST into the model. Nevertheless, it is of clinical importance to distinguish NASH patients from NAFL. The accuracy of miR-122-5p and miR-193b-3p for discriminating NASH from NAFL showed AUROC of 0.734 and 0.725 respectively, performing slightly better than AST (AUROC of 0.724). Combination of miR-122-5p and miR-193b-3p gave better diagnostic accuracy in predicting NASH (AUROC of 0.743) compared to the use of either miRNAs alone. Out of all models, combined utility of miR-122-5p and AST gave the best prediction for NASH from NAFL, with AUROC of 0.785. With regards to the detection of significant fibrosis, miR-122-5p, miR-193b-3p and miR-125b-5p demonstrated AUROC of 0.736, 0.667 and 0.671 respectively, with only miR-122-5p performing better than AST which has an AUROC of 0.705. A slight increase of AUROC was observed when combining miR-122-5p and AST in a single model (AUROC = 0.743) but not with other miRNAs.

Spearman's rank correlation test on the NAFLD-related clinical parameters and levels of miRNAs expression showed significant associations of liver enzymes (AST, ALT and GGT) with most miRNAs, except for miR-103a-3p (Table 3.5). In overall, there was no correlation of miRNAs with serum triglycerides levels and only miRNA-215 showed a positive correlation with BMI. One of the attributes of good biomarkers is that its expression does not change appreciably with age or gender. No significant correlation was observed between the miRNAs concentration and age.

Similarly, there was no difference in the expression of biomarkers between male and female patients.



Test	AUROC (95% CI)	*P* value
AST	0.910 (0.867-0.954)	<0.0001
miR122	0.943 (0.912-0.974)	<0.0001
miR192	0.916 (0.875-0.957)	<0.0001
miR885	0.936 (0.900-0.972)	<0.0001
miR125b	0.931 (0.894-0.967)	<0.0001
miR378a	0.922 (0.884-0.961)	<0.0001
miR215	0.919 (0.879-0.960)	<0.0001
miR122_miR885	0.952 (0.924-0.980)	<0.0001
miR122_miR125b	0.947 (0.917-0.977)	<0.0001
miR122_miR192	0.950 (0.921-0.979)	<0.0001
miR122_AST	0.959 (0.934-0.984)	<0.0001
miR122_miR192_AST	0.969 (0.947-0.990)	<0.0001
miR122_miR885_AST	0.965 (0.942-0.988)	<0.0001

Test	AUROC (95% CI)	*P* value
AST	0.724 (0.624-0.824)	0.001
ALT	0.637 (0.521-0.752)	0.035
miR122	0.734 (0.631-0.838)	<0.0001
miR193b	0.725 (0.624-0.825)	0.001
miR122_miR193	0.743 (0.642-0.843)	<0.0001
miR122_AST	0.785 (0.697-0.873)	<0.0001

Test	AUROC (95% CI)	*P* value
AST	0.705 (0.602-0.808)	0.001
ALT	0.627 (0.513-0.740)	0.037
miR122	0.736 (0.632-0.839)	<0.0001
miR193b	0.667 (0.557- 0.776)	0.006
miR125b	0.671 (0.557-0.785)	0.005
miR122_miR193	0.735 (0.632-0.838)	<0.0001
miR122_AST	0.743 (0.645-0.841)	<0.0001

Figure 0.7: Receiver operating characteristic (ROC) curves with corresponding area under ROC curves for comparison of the diagnostic accuracy of circulating miRNAs and classic serum biomarkers in predicting disease severity: (A) NAFLD; (B) NASH; (C) Significant fibrosis

Table 0.5: Correlation between significantly increased miRNAs in NAFLD patients and the clinical and NAFLD-related biochemical parameters

	Spearman's rank correlation test r (*P* value)			
Circulating miRNAs	**AST (U/L)**	**ALT (U/L)**	**GGT (U/L)**	**BMI (kg/m^2)**
miR-122-5p	0.45 (<0.0001)	0.41 (<0.0001)	0.34 (<0.0001)	-
miR-192-5p	0.29 (0.001)	0.25 (0.006)	0.27 (0.002)	-
miR-193b-3p	0.50 (<0.0001)	0.46 (<0.0001)	0.35 (<0.0001)	-
miR-103a-3p	-	-	-	-
miR-885-5p	0.33 (<0.0001)	0.36 (<0.0001)	0.22 (0.016)	-
miR-505	0.21 (0.021)	0.26 (0.004)	-	-
miR-34a-5p	0.40 (<0.0001)	0.35 (<0.0001)	0.32 (<0.0001)	-
miR-148a	0.24 (0.007)	0.21 (0.022)	-	-
miR-125b-5p	0.31 (0.001)	0.28 (0.002)	0.27 (0.002)	-
miR-378a	0.21 (0.018)	0.24 (0.008)	0.21 (0.023)	-
miR-194-5p	0.27 (0.03)	0.30 (0.001)	0.21 (0.02)	-
miR-215	0.29 (0.001)	0.26 (0.048)	0.29 (0.001)	0.20 (0.025)
miR-99a-5p	0.31 (0.001)	0.30 (0.003)	0.35	-

ALT, alanine aminotransferase*; AST*, aspartate aminotransferase*; BMI*, body mass index*;*

GGT, gamma-glutamyl transpeptidase

3.5 Discussion

This study explored the potential of circulating miRNAs in the diagnosis of NAFLD and revealed that except for miR-103a-3p, all miRNAs validated (miR-122-5p, miR-193b-3p, miR-885-5p, miR-34a-5p, miR-125b-5p, miR-194-5p, miR-215, miR-99a-5p, miR-192-5p, miR-505-3p, miR-148a-3p, miR-378a-3p) are significantly increased in NAFL patients compared to healthy controls. Since NAFL represents the early stage of the disease, the findings indicate that miRNAs may serve as a great utility in the screening of the disease in the general population before the appearance of clinical symptoms. The applicability of miRNAs as a biomarker to be used in general population was also been demonstrated when miR-122 and miR-34a are found to be increased in 92 NAFLD patients out of 403 Japanese subjects who underwent medical check-up (Yamada et al., 2013).

Beside the identification of NAFLD, the ability of miRNAs to discriminate between NASH and NAFL is of great concern due to the higher risk of disease progression in NASH patients. MiR-122-5p and miR193b-3p were able to distinguish NASH from NAFL patients. This will provide practical insight for the clinician in decision making, in identifying patients in need of referral to secondary care. In addition to the discrimination of NASH, proceeding studies also found that patients with a higher stage of fibrosis are at higher risk of developing cirrhosis and HCC (Alkhouri & McCullough, 2012). Stratification of patients according to their stage of fibrosis is crucial to facilitate surveillance for the development of oesophageal varices and HCC as well as enrolment into various treatment trials (Chalasani et al., 2012). Interestingly, there is a high prospect of miR-122-5p, miR-193b-3p and miR-125b-5p to be utilised as a stratifying tool due to their higher expression levels in patients with significant fibrosis (fibrosis score ≥ 2).

Elevated level of circulating miR-122-5p in NAFLD was consistently being reported in other literature, with goodcorrelation with liver enzymes activity and liver histologic stage (Cermelli et al., 2011; Zhang et al., 2010). In agreement with other studies that performed stratification on the severity of NAFLD patients, present study found upregulation of miR-122-5p in NAFL compared to healthy controls, as well as upregulation in NASH compared to NAFL (Cermelli et al., 2011; Pirola et al., 2015; Zhang et al., 2010). The expression of miR-122-5p is also the highest among all the miRNAs tested, as it is expressed almost exclusively by hepatocytes, making up to 70% of the total pool of hepatic miRNAs (Chang 2004).However, high expression of miR-122-5p has also been consistently reported in patients with chronic hepatitis, drug-induced liver injury and HCC (Lewis et al., 2011; Xu et al., 2011; Zhang et al., 2010). This suggests that the increase in serum miR-122-5p is common to chronic liver disease of all aetiologies, instead of being specific to NAFLD.

In terms of the histological severity, upregulation of miR-122-5p was found in patients with a higher grade of inflammation and hepatocellular ballooning, consistent with the findings from other studies (Akuta et al., 2016; Auguet et al., 2016; Becker et al., 2015; Liu et al., 2016). In the assessment according to the stage of fibrosis, Miyaaki H et al. (2013) found a lower level of miR-122-5p in patients with significant fibrosis ($F \geq 2$) compared to those without significant fibrosis ($F < 2$). On the other hand, Pirola et al. (2015) demonstrated higher expression of miR-122-5p in patients with fibrosis compared to those without, but did not stratify the patients according to a different stage of fibrosis. Like the more recent studies, present finding supports miR-122-5p has a biphasic change with the progression of fibrosis (Akuta et al., 2016; Lopez-Riera et al., 2018). The expression was the highest in fibrosis stage 2 but likely to decrease in stage 3, which may limit the use of miR-122-5p in the identification of advanced fibrosis ($F \geq 3$) in NAFLD patients. The lower level of miR-122-5p in F3 compared to F2 was probably due to the

constant replacement of hepatocytes by extracellular matrix during fibrogenesis. Since these main miR-122-producing hepatic cells are reduced, the expression of miR-122-5p in the serum subsequently decreases (Chang et al., 2004).

Being another liver-enriched miRNA, increased levels of miR-192-5p were also being reported in NASH (Becker et al., 2015; Liu et al., 2016; Lopez-Riera et al., 2018; Pirola et al., 2015) and was upregulated in hepatocellular ballooning (Pirola et al., 2015). A recent study suggested that the ratio of miR-192/miR-197 and miR-192/30c is useful in the discrimination of NASH patients from those with NAFL (Lopez-Riera et al., 2018). Despite upregulation of miR-192-5p in NAFLD in the present study, the differential expression of this miRNA in NASH compared to NAFL did not achieve statistical significance after correction for multiple testing. Another miRNA that is consistently reported to be upregulated in NAFLD is miR-34a-5p (Cermelli et al., 2011; Liu et al., 2016; Yamada et al., 2013). Levels of miR-34a-5p and miR-193b-3p were found to be largely undetectable in serum of healthy controls but significantly increased in patients with NASH. A recent study also reported a similar trend in the plasma level of miR-34a-5p in NASH patients, but like the study performed in the Chinese patients (Liu et al., 2016), the positive correlation of miR-34a-5p with fibrosis severity was not replicated in the present study (Cermelli et al., 2011).

Majority of the miRNA profiling studies in NAFLD validated the diagnostic performance of individually selected miRNAs based on previous literature review, thus limiting the discovery of new miRNAs as potential disease biomarkers. One of the novel findings in this study was the performance of miR-193b-3p in distinguishing NASH from NAFL, as well as its significant association with hepatocellular ballooning, degree of inflammation and stage of fibrosis. The diagnostic value of serum miR-193b-3p has not been explored in NAFLD but studies had reported it is

down-regulated in the liver of NASH patients and also in HCC tissues (Estep et al., 2010; Xu et al., 2010). This suggests a possible utility of miR-193b-3p as a diagnostic biomarker for NASH but further validation in other independent cohort is required.

Higher expression of miR-125b-5p was also observed in patients with significant fibrosis compared to those without. *In vitro* and *in vivo* studies have discovered the role of miR-125b in promoting hepatic stellate cells activation during hepatic fibrogenesis (You et al., 2018). *In-silico* pathway analysis suggests the role of miR-125b-5p in disease pathways associated with metabolic disorders, hypertension and blood coagulation (Pirola et al., 2015). Up-regulation of miR-125b-5p is reported in NAFLD but not NASH and is implicated in a state of hepatic inflammation (Pirola et al., 2015). In our study, miR-125b-5p was not found to be associated with hepatic inflammation after multiple correction testing and its association with NASH compared to NAFL was also of borderline significance (P =0.042). Of note, miR-125b-5p has been found to be associated with the progression of liver fibrosis in mice (Murakami et al., 2011) and also suppressing HCC cell growth *in vitro* and *in vivo*(Liang et al., 2010), suggesting its plausible role as a prognostic biomarker.

In terms of the diagnostic performances of miRNAs, the AUROC showed that miR-122-5p outperforms liver enzymes in the identification of NAFLD, similar to the findings in another two studies (Cermelli et al., 2011; Tan et al., 2014). On the other hand, a cohort involving 871 participants also found an upregulation of miR-122 and miR-885 in ultrasound-proven fatty liver, but unlike our study their miRNAs diagnostic accuracy is not better than liver enzymes (Raitoharju et al., 2016). This could be due to liver biopsy not being performed in their study and ultrasound could potentially miss hepatic steatosis of <5 %. Their large cohort also included participants regardless of their drinking habit, which may also contribute to the discrepancy. In stratifying NASH from NAFL, The

AUROC showed that miR-122-5p gives satisfying diagnostic accuracy in identifying NASH from NAFL, consistent with the results in other studies (Becker et al., 2015; Cermelli et al., 2011; Liu et al., 2016; Pirola et al., 2015). Despite miR-122-5p having relatively higher AUROC than liver enzymes in identifying NASH, the multivariate ROC analysis incorporating miR-122-5p and AST gave the best AUROC (0.785) in the identification of NASH. Pairing of miR-122-5p and AST also resulted in the highest diagnostic accuracy in identifying significant fibrosis (≥2) compared to using any single miRNA alone. However, the present study also found that the miRNAs do not give good diagnostic accuracy in identification of advanced fibrosis (≥F3). Other algorithms such as APRI and FIB-4 score are found to outperform miRNAs in the identification of significant or advanced fibrosis (Liu et al., 2016; Lopez-Riera et al., 2018).

One of the limitations of this study is that the NAFLD patients were recruited from a tertiary referral centre, thus their histological finding may not represent those found in a general setting. The healthy controls were also scrutinised for the absence of fatty liver and those who have dyslipidaemia and/or elevated liver enzymes were not recruited, thus a higher AUROC of liver enzymes in the diagnosis of NAFLD was shown in this study. The present study also lacks long term outcome data where the prognostication is unknown. Another limitation of this study was the involvement of only NASH patients and healthy controls in the screening phase. This may hinder the identification of miRNAs that best discriminate NASH from NAFL, contributing to the higher diagnostic accuracy of miRNAs validated in detecting NAFLD from healthy controls rather than distinguishing NASH from NAFL. The number of patients in the screening phase was also relatively low since the quantification of 752 miRNAs was carried out. Despite this, the patients samples used for the screening were all of NAS score higher than or equals to 5, with mixture of fibrosis scores. This represents a wide spectrum of NASH in terms of the histological parameters and reduces the chances of

excluding important miRNAs in the validation phase.One limitation of miRNAs study was the absence of consensus in selecting a suitable house-keeping gene for the data normalisation in miRNA profiling studies. However, since the expression information of a large number of miRNAs is available in the screening phase, global normalisation was employed as it outperforms other normalisation strategy with reduced technical variation (Mestdagh et al., 2009). The screening phase also allows the identification of housekeeping gene which is found to be expressed stably across all the samples, and hence being chosen to be the reference gene in the validation phase.

One of the advantages of this study was the multi-stage experiment design. Unlike most studies (Becker et al., 2015; Celikbilek et al., 2014; Cermelli et al., 2011; Liu et al., 2016; Lopez-Riera et al., 2018; Miyaaki et al., 2014; Pirola et al., 2015; Xu et al., 2011; Yamada et al., 2013) which performed miRNA testing directly on individually selected miRNAs, this study screened a large number of serum miRNAs to first identify potential diagnostic markers, following by a validation on an independent training set of cohort. The only pilot study on large scale miRNA profiling in NAFLD was performed using sequencing (Y. Tan et al., 2014) while present study utilised qRT-PCR in the characterisation of miRNAs during both screening and validation phase. It has higher sensitivity than microarray, whilst being more cost effective and require only a little sample of biofluids compared to NGS. Though it was not a *de-novo* discovery study, the initial screening phase was conducted on a panel of 752 miRNas commonly found in the human, instead of a selection of a few miRNAs based on existing literature. In addition, the samples were scrutinised based on strict quality control procedures. Since human erythrocytes contain high levels of certain miRNAs, samples that underwent haemolysis were excluded from this study (Pritchard, Kroh, et al., 2012). Most studies did not perform testing on haemolysis and may result in discrepancies, such as high level of miR-16 being found in NAFLD could be due to haemolysis as they are highly abundant

in erythrocytes (Cermelli et al., 2011; Liu et al., 2016; Lopez-Riera et al., 2018). In addition, samples in this study were also tested to confirm the absence of PCR inhibitors which could otherwise alter the reliability of the results.

Overall, circulating miRNAs in biofluids pose as a potential non-invasive biomarker due to their relative stability, ease of sample collection and availability of qRT-PCR which is cost effective and highly sensitive. Despite the appeal of miRNA as promising diagnostic biomarkers, it is still in infancy stage towards clinical application. It requires further evaluation in terms of diagnostic specificity and reproducibility and also the establishment of a reliable normalisation strategy and sample handling protocol. Future studies could consider including larger patient cohort with liver disease of different aetiology to further verify the specificity of these miRNAs in different liver pathological conditions. Longitudinal studies involving serial quantification of miRNAs can also be performed to assess if the changes in miRNAs expression reflect disease progression.

3.6 Conclusion

In conclusion, the present miRNAs profiling study demonstrated the potential of miR-122-5p and miR-193b-3p in the stratification of NASH from NAFL, as well as in identification of significant fibrosis. Whilst miR-122-5p is a widely recognised liver-related miRNA, the potential of miR-193b-3p as a diagnostic biomarker in NAFLD awaits further evaluation and validation. The validated miRNAs could be used to complement or improve the current diagnostic algorithm, in stratifying NAFLD patients at higher risk of disease progression.

CHAPTER 4: NON-INVASIVE SCORING SYSTEMS FOR THE DIAGNOSIS OF ADVANCED FIBROSIS IN NAFLD

4.1 Introduction

NAFLD is the hepatic manifestation of the metabolic syndrome with increasing prevalence worldwide, paralleling the epidemic of obesity and diabetes, with the global estimated prevalence at 24%, 27% in the Asian and 22.7 % in the Malaysian population (Goh, Ho, & Goh, 2013; Younossi et al., 2016). The global epidemic of obesity and diabetes continue to exacerbate metabolic conditions, thus, increasing the clinical and economic burden of NAFLD (Younossi et al., 2018).

The apparent challenge, however, is that the disease is often asymptomatic until the late stage of NAFLD which is accompanied by liver disease-associated cutaneous stigmata or features of hepatic decompensation (Ahmed, 2015). Patients with NAFLD have a higher mortality rate compared to general population, mainly attributed to cardiovascular disease, malignancy or liver related mortality (Bertot & Adams, 2016). Among NAFLD patients, those with advanced fibrosis (stage 3 and 4) exhibit the worst prognosis independent of their NAS score, with hazard ratio ranging from 3.3-5.7 (Ekstedt et al.,2015).

It is therefore, imperative to identify subgroups of patients with NAFLD that are at high risk for adverse outcomes who will require additional workup and surveillance, so that interventions can be targeted to patients at greatest need. Since NAFLD is often discovered incidentally based on elevated liver enzymes in the primary care setting, the next step would be to risk-stratify the patients prior to referral to secondary or tertiary care, with focus on the presence of advanced fibrosis. Since NAFLD is often discovered incidentally based on elevated liver enzymes in the primary

care setting, the next step would be to risk-stratify the patients prior to referral to secondary or tertiary care, with focus on the presence of advanced fibrosis.

The current gold standard of liver biopsy offers a vast array of information including the degree of steatosis, severity of necroinflammation, hepatocellular ballooning, and fibrosis stage (Nalbantoglu & Brunt, 2014) . Notwithstanding that, liver biopsy, being an invasive procedure, has other limitations such as sampling errors, does not permit longitudinal monitoring of fibrosis progression and it is not feasible to be performed on all NAFLD patients.

This has led to the emergence of several non-invasive screening tests utilising anthropometric data and easily available clinical parameters such as AST/ALT ratio, APRI (AST to platelet ratio index), BARD (BMI, AST/ALT ratio, Diabetes) , FIB-4 (Fibrosis-4) and NFS (NAFLD fibrosis score). They have been validated against the liver biopsy with variable accuracy in different populations with the capability to identify or rule out advanced fibrosis in NAFLD patients (McPherson et al., 2010). Since these scoring methods are calculated based on simple clinical parameters, they are easy to be carried out on a routine basis without incurring a huge cost. Besides these simple scoring systems, complex predictive models comprising panels of fibrosis markers such as ELF (Enhanced Liver Fibrosis Panel) and FibroTest are also gaining recognition. Regardless of the wide validation of some of these non-invasive tests, only a few studies were carried out in Asia where some of these scores are less applicable in Asian cohort.

The aim of this study was to compare and validate the diagnostic accuracy and clinical utility of non-invasive tests including AAR, APRI, BARD, FIB-4, and NFS in Malaysian cohort. As a commercial panel marker, the performance of ELF was also compared to the performances of other simple biomarkers to assess if the inclusion of

extracellular matrix turnover markers would give additional benefits in the diagnosis of advanced fibrosis. This study can help to inform policymakers of a simplified testing algorithm for a public health approach, particularly in the community setting, to identify not only those at risk of liver-related mortality, but also cardiovascular deaths and malignancy.

4.2 Literature review

4.2.1 Importance of identification of fibrosis

Advanced fibrosis (fibrosis stage 3-4) is often the most important histological feature that determines the progression of NASH to hepatocellular carcinoma. A recent meta-analysis showed that increases in fibrosis stage increases the risk of liver-related mortality, where stage 3 fibrosis has a mortality rate ratio (MRR) of 16.7 and stage 4 with MRR of 42.3, as opposed to NAFLD patients without fibrosis (Dulai et al., 2017). Identification of patients with advanced fibrosis is therefore of greatest concern for clinicians as these patients require careful clinical surveillance and they are in greatest need of emerging pharmacology therapy.

Patients with metabolic syndrome are at higher risk of steatohepatitis and should be considered for liver biopsy (Chalasani et al., 2018). However with the rising prevalence of diabetes mellitus and obesity, non-invasive markers are more feasible as a first-line staging approach in stratifying patients with higher risk of advanced fibrosis from the large population of NAFLD. Considering the absence of treatment for NAFLD, identification of non-significant fibrosis (fibrosis stage 0-1) in the community level facilitates the early implementation of dietary and lifestyle intervention as well as the referral of certain patients to secondary care (Guha et al., 2008). In the secondary care setting, it will serve as a valuable adjunct to biopsy, especially in the serial surveillance of disease progression. Results from the scoring systems can also be used to select the patients of greatest need for emerging therapeutic interventions to be enrolled in clinical trials (Bedossa & Patel, 2016). In addition, studies demonstrated that fibrosis regression reflects the recovery of liver injury in NAFLD and antifibrotic agents has been delivering promising results in various phases of clinical trials (Brenner, Galluzzi, Kepp,

& Kroemer, 2013; Ellis & Mann, 2012; Sumida & Yoneda, 2018). Hence serial monitoring of fibrosis stage during therapy is crucial for evaluation of treatment efficacy.

4.2.2 Non-invasive serum biomarkers for fibrosis

Indirect serum markers encompass routine biochemical tests such as transaminase level, platelet count, albumin level, and sometimes taking into consideration of parameters potentially related to liver fibrosis such as age or diabetes. These panels usually have dual cut-offs: a high cut-off with high specificity and a low cut-off with high sensitivity (Buzzetti, Lombardi, De Luca, & Tsochatzis, 2015). By utilising both cut-offs, the number of false positive and false negative could be minimised but there will also be a proportion of patients falling under the indeterminate range of values and will require further evaluation. The EASL-EASD-EASO recommend the use of widely validated scoring systems such as NFS, FIB-4 and ELF as the first-line triage to identify patients at low risk of advanced fibrosisThe use of non-invasive scoring system has also been expanded to help predict advanced fibrosis in diabetic patients (Singh, Le, Lopez, & Alkhouri, 2018).

Studies involving the assessments of some of the common serum markers for fibrosis in NAFLD are summarised in Table 4.1. All studies used liver biopsy as a gold standard in confirming the stage of fibrosis, except for Kim et al. (2013) who used liver biopsy and magnetic resonance elastography (MRE) as standard. Various scoring systems were used with majority of the studies employing the NASH CRN histological scoring system that was modified by Kleiner or the earlier Brunt classification where both scoring systems are comparable at stage 3-4 fibrosis (Brunt, Janney, Di Bisceglie, Neuschwander-Tetri, & Bacon, 1999; Kleiner et al., 2005). A study by Wahl et al. (2012)

used ISHAK scoring method as the study comprises of patients with chronic liver disease with various aetiologies (Ishak et al., 1995). A study in France population used Metavir scoring system (Bedossa & Poynard, 1996) meanwhile a study in morbidly obese Spanish patients used Matteoni classification (Matteoni et al., 1999).

Table 0.1: Non-invasive scoring systems for diagnosis of fibrosis in NAFLD patients

AST/ALT Ratio													
Reference	**Population**	**Diagnosis**	**Scoring**	**Age**	**Number of patients**	**Prevalence of advanced fibrosis (%)**	**Cut off**	**Sensitivity**	**Specificity**	**PPV**	**NPV**	**AUROC (95% CI)**	**Prediction of disease severity**
(Shah et al., 2009)	USA	Liver biopsy	Kleiner	48±12	541	23.1	N/A	N/A	N/A	N/A	N/A	0.74 (0.69-0.79)	F3-F4
(McPherson et al., 2010)	UK	Liver biopsy	Kleiner	51±12	145	18.6	0.8	74	78	44	93	0.83 (0.74-0.91)	F3-F4
							1	52	90	55	89		
(Wong, Vergniol, et al., 2010)	French and Chinese	Liver biopsy	Kleiner	51±11	246	22.8	0.8	40	78	38	81	0.66 (0.58-0.74)	F3-F4
							1	21	90	39	78		
(Sumida et al., 2012)	Japan	Liver biopsy	Kleiner	52±15	576	16.3	0.8	66	76	26	95	0.871	F3-F4
							1	48	92	44	94		
(Xun et al., 2012)	Chinese	Liver biopsy	Kleiner	37±10	152	15.8	0.8	42	79	27	88	0.67 (0.56-0.78)	F3-F4
							1	25	87	26	86		
(McPherson, Anstee, Henderson, Day, & Burt, 2013)	UK	Liver biopsy	Kleiner	54±11	70 (Normal ALT)	24.3	0.8	94	44	35	96	0.81 (0.67-0.92)	F3-F4
							1	82	72	48	93		
				48±13	235 (Elevated ALT)	17.4	0.8	59	86	46	91	0.79 (0.71-0.86)	
							1	37	95	60	88		
(Kim, Kim, Talwalkar, et al., 2013)	USA	Liver biopsy and MRE	Kleiner	54±13	325	32.0	0.8	77	61	48	84	0.71 (0.62 - 0.78)	F3-F4
							1	53	79	53	78		
(Perez-Gutierrez et al., 2013)	Latin	Liver biopsy	Brunt	49±13	243	11.1	0.8	66	62	19	93	0.67 (0.57-0.77)	F3-F4

Table 4.1, continued

Reference	Population	Diagnosis	Scoring	Age	Number of patients	Prevalence of advanced fibrosis (%)	Cut off	Sensitivity	Specificity	PPV	NPV	AUROC (95% CI)	Prediction of disease severity
(Mohamed, Nabih, ElShobaky, & Khattab, 2014)	Egypt	Liver biopsy	Kleiner	41±6	76	25.0	0.8	90	25	28	88	0.84 (0.72-0.96)	F3-F4
(Treeprasertsuk et al., 2017)	Thailand	Liver biopsy	N/A	41±13	139	6.5	0.8	44	65	8	94	0.57	F3-F4
(De Silva et al., 2018)	South Asian	Liver biopsy	Kleiner	44*	90	35.6	0.8	50	76	41	82	0.72 (0.60-0.84)	F3-F4
	White			52*	79	29.1	0.8	70	82	62	87	0.87 (0.78-0.95)	
APRI													
Reference	**Population**	**Diagnosis**	**Scoring**	**Age**	**Number of patients**	**Prevalence of advanced fibrosis (%)**	**Cut off**	**Sensitivity**	**Specificity**	**PPV**	**NPV**	**AUROC (95% CI)**	**Prediction of disease severity**
(Cales et al., 2009)	France	Liver biopsy	Metavir	51±11	235	18.7	0.5	66	91	73	88	0.87 (0.81-0.92)	F2-4
(Shah et al., 2009)	USA	Liver biopsy	Kleiner	48±12	541	23.1	N/A	N/A	N/A	N/A	N/A	0.73 (0.68-0.78)	F3-F4
(McPherson et al., 2010)	UK	Liver biopsy	Kleiner	51±12	145	18.6	1	27	89	37	84	0.67 (0.54-0.80)	F3-F4
(Wong, Vergniol, et al., 2010)	French and Chinese	Liver biopsy	Kleiner	51±11	246	22.8	0.5	65	72	42	87	0.74 (0.67-0.82)	F3-F4
							1.5	6	97	40	77		
(Sumida et al., 2012)	Japan	Liver biopsy	Kleiner	52±15	576	16.3	1	67	81	31	95	0.82	F3-F4
(Xun et al., 2012)	Chinese	Liver biopsy	Kleiner	37±10	152	15.8	0.5	79	50	23	93	0.74 (0.62-0.86)	F3-F4
							1	42	88	40	89		
							1.5	25	96	55	87		
(Kim, Kim, Talwalkar, et al., 2013)	USA	Liver biopsy and MRE	Kleiner	54 ±13	312	N/A	0.5	87	43	43	87	0.74 (0.66-0.81)	F3-F4
							1.5	30	94	72	73		

Table 4.1, continued

Reference	Population	Diagnosis	Scoring	Age	Number of patients	Prevalence of advanced fibrosis (%)	Cut off	Sensitivity	Specificity	PPV	NPV	AUROC (95% CI)	Prediction of disease severity
(Perez-Gutierrez et al., 2013)	Latin	Liver biopsy	Brunt	49 ±13	243	11.1	1	37	86	26	91	0.66 (0.55-0.77)	F3-F4
(Mohamed et al., 2014)	Egypt	Liver biopsy	Kleiner	41±6	76	25.0	1	21	93	50	78	0.91 (0.84-0.97)	F3-F4
	South Asian			44*	90	35.6	1	18	91	40	82	0.68 (0.56-0.81)	F3-F4
(De Silva et al., 2018)	White	Liver biopsy	Kleiner	52*	79	41.8	1	52	91	71	82	0.78 (0.65-0.91)	
BARD													
References	**Population**	**Diagnosis**	**Scoring**	**Age**	**Patient**	**Prevalence of advanced fibrosis (%)**	**Cut off**	**Sensitivity**	**Specificity**	**PPV**	**NPV**	**AUROC (95% CI)**	**Prediction of disease severity**
(Harrison, Oliver, Arnold, Gogia, & Neuschwander-Tetri, 2008)	USA (caucasian african american hispanic asian pacific islander)	Liver biopsy	Brunt	49*	827	22.0	2	N/A	N/A	43	96	0.81 (9.2 - 31.9)	F3-F4
(Shah et al., 2009)	USA	Liver biopsy	Kleiner	48±12	541	23.1	N/A	N/A	N/A	N/A	N/A	0.70(0.64-0.75)	F3-F4
(Fujii et al., 2009)	Japan	Liver biopsy	N/A	59*	122	37.7	2	N/A	N/A	59	77	0.73	F3-F4
(Wong, Vergniol, et al., 2010)	French and Chinese	Liver biopsy	Kleiner	51±11	246	22.8	2	62	66	36	85	0.69 (0.61-0.77)	F3-F4

Table 4.1, continued

Reference	Population	Diagnosis	Scoring	Age	Number of patients	Prevalence of advanced fibrosis (%)	Cut off	Sensitivity	Specificity	PPV	NPV	AUROC (95% CI)	Prediction of disease severity
(Raszeja-Wyszomirska et al., 2010)	Polish-Caucasians	Liver biopsy	N/A	48±12	104	14.4	2	87	73	35	97	0.82 (3.64-82.56)	F3-F4
(Ruffillo et al., 2011)	Argentina	Liver biopsy	Brunt	49*	238	15.5	2	51	77	45	81	0.67 (0.51-0.77)	F3-F4
(McPherson et al., 2010)	UK	Liver biopsy	Kleiner	51±12	145	18.6	2	89	44	27	95	0.77 (0.68-0.87)	F3-F4
(Cichoz-Lach et al., 2012)	Polish-caucasians	Liver biopsy	Kleiner	43±14	126	21.4	2	89	89	69	97	0.87 (0.79–0.92)	F3-F4
(Sumida et al., 2012)	Japan	Liver biopsy	Kleiner	52±15	576	16.3	2	80	65	22	97	0.77	F3-F4
(Xun et al., 2012)	Chinese	Liver biopsy	Kleiner	37±10	152	15.8	2	42	79	27	88	0.64 (0.51-0.77)	F3-F4
(McPherson et al., 2013)	UK	Liver biopsy	Kleiner	54±11	70 (Normal ALT)	24.3	2	94	26	29	93	0.71 (0.57-0.85)	F3-F4
				48±13	235 (Elevated ALT)	17.4	2	83	57	28	94	0.78 (0.57-0.85)	
(Kim, Kim, Talwalkar, et al., 2013)	USA	Liver biopsy and MRE	Kleiner	54 ±13	325	32.0	2	77	61	48	85	0.72 (0.63-0.79)	F3-F4
(Perez-Gutierrez et al., 2013)	Latin	Liver biopsy	Brunt	49 ±13	243	11.1	2	76	43	15	93	0.65 (0.52-0.77)	F3-F4

Table 4.1, continued

Reference	Population	Diagnosis	Scoring	Age	Number of patients	Prevalence of advanced fibrosis (%)	Cut off	Sensitivity	Specificity	PPV	NPV	AUROC (95% CI)	Prediction of disease severity
(Demir et al., 2013)	German	Liver biopsy	Brunt Kleiner	45±13	159	11.9	2	74	77	30	96	0.82 (0.71-0.92)	F3-F4
(Mohamed et al., 2014)	Egypt	Liver biopsy	Kleiner	41±6	76	25.0	2	95	21	29	92	0.56 (0.42-0.70)	F3-F4
(Treeprasertsuk et al., 2017)	Thailand	Liver biopsy	N/A	41±13	139	6.5	2	78	40	8	96	0.58	F3-F4
(De Silva et al., 2018)	South Asian	Liver biopsy	Kleiner	44*	90	35.6	2	50	84	50	84	0.73 (0.61-0.85)	F3-F4
	White			52*	79	29.1	2	70	84	64	87	0.83 (0.73-0.94)	
ELF													
References	**Population**	**Diagnosis**	**Scoring**	**Age**	**Number of Patient**	**Prevalence of advanced fibrosis (%)**	**Cut off**	**Sensitivity**	**Specificity**	**PPV**	**NPV**	**AUROC (95% CI)**	**Prediction of disease severity**
(Guha et al., 2008)	UK	Liver biopsy	Kleiner	49±13	192	22.9	10.3576	80	90	71	94	0.90 (0.84 -0.96)	F3-F4
							9.8932	70	80	70	80	0.82 (0.75-0.88)	F2-F4
							9.793	60	80	81	79	0.76 (0.69-0.83)	F1-F4
(Nobili et al., 2009)	Italian Paediatric	Liver biopsy	Brunt	14±3	112	7.1	9.28	88	81	90	77	0.92 (0.86-0.97)	F1-F4
							10.18	94	93	70	99	0.98 (0.96-1.00)	F2-F4
							10.51	100	98	80	100	0.99 (0.97-1.00)	F3-F4
(Alkhouri et al., 2011)	Italian Paediatric	Liver biopsy	Kleiner	10.5*	111	N/A	8.49	76.9	97	N/A	N/A	0.92 (0.87-0.98)	F1-F3
							9.28	21	100	N/A	N/A	0.97 (0.94-0.99)	F2-F4
(Wahl et al., 2012)	Chronic liver disease (22 with NAFLD)	Liver biopsy	ISHAK	N/A	102	20.6	8.99	86	70	N/A	N/A	0.87 (0.78-0.96)	F2-F6
							9.39	100	77	N/A	N/A	0.93 (0.88-0.99)	F5-F6

Table 4.1, continued

References	Population	Diagnosis	Scoring	Age	Number of Patient	Prevalence of advanced fibrosis (%)	Cut off	Sensitivity	Specificity	PPV	NPV	AUROC (95% CI)	Prediction of disease severity
(Miele et al., 2017)	Italian	Liver Biopsy	Kleiner	46	82	18.3	9.8	87	93	72	97	0.95 (0.88-1.00)	F3-F4
(Lopez et al., 2017)	Spanish morbid obese	Liver biopsy	Matteoni	44±12	57	N/A	8.72	71	74	74	72	0.74	NASH and/or fibrosis
FIB-4													
References	**Population**	**Diagnosis**	**Scoring**	**Age**	**Patient**	**Prevalence of advanced fibrosis (%)**	**Cut off**	**Sensitivity**	**Specificity**	**PPV**	**NPV**	**AUROC (95% CI)**	**Prediction of disease severity**
(Shah et al., 2009)	USA	Liver biopsy	Kleiner	48±12	541	23.1	1.3	74	71	43	90	0.80 (0.76-0.85)	F3-F4
							2.67	33	98	80	83		
(McPherson et al., 2010)	UK	Liver biopsy	Kleiner	51±12	145	18.6	1.3	85	65	36	95	0.86 (0.78-0.94)	F3-F4
							3.25	26	98	75	85		
(Wong, Vergniol, et al., 2010)	French and Chinese	Liver biopsy	Kleiner	51±11	246	22.8	1.3	65	80	51	88	0.8 (0.74-0.87)	F3-F4
							2.67	21	96	59	79		
(Sumida et al., 2012)	Japan	Liver biopsy	Kleiner	52±15	576	16.3	1.45	90	64	24	98	0.87	F3-F4
							3.25	48	95	53	94		
(Xun et al., 2012)	Chinese	Liver biopsy	Kleiner	37±10	152	15.8	1.3	67	67	28	92	0.76 (0.64-0.88)	F3-F4
							2.67	38	96	64	89		
							3.25	21	97	56	87		

Table 4.1, continued

References	Population	Diagnosis	Scoring	Age	Number of Patient	Prevalence of advanced fibrosis (%)	Cut off	Sensitivity	Specificity	PPV	NPV	AUROC (95% CI)	Prediction of disease severity
(McPherson et al., 2013)	UK	Liver biopsy	Kleiner	54±11	70 Normal ALT	24.3	1.3	82	77	52	92	0.86 (0.77-0.96)	F3-F4
							3.25	35	98	85	83		
				48±13	235 Elevated ALT	17.4	1.3	81	72	37	95	0.85 (0.79-0.92)	F3-F4
							3.25	20	98	67	86		
(Kim, Kim, Talwalkar, et al., 2013)	USA	Liver biopsy and MRE	Kleiner	54.±13	312	N/A	1.3	87	56	49	90	0.83 (0.76-0.89)	F3-F4
							3.25	49	93	77	79		
(Perez-Gutierrez et al., 2013)	Latin	Liver biopsy	Brunt	49 ±13	243	11.1	3.25	53	87	37	94	0.74 (0.65-0.84)	F3-F4
(Demir et al., 2013)	German	Liver biopsy	Kleiner	45±13	165	12.1	1.45	85	88	49	98	0.95 (0.91-1.00)	F3-F4
							3.25	40	100	100	92		
(Mohamed et al., 2014)	Egypt	Liver biopsy	Kleiner	41±6	76	25.0	1.3	84	87	67	94	0.94 (0.88-0.99)	F3-F4
							2.67	63	93	75	88		
(Treeprasertsuk et al., 2017)	Thailand	Liver biopsy	N/A	41±13	139	6.5	1.3	56	84	19	96	0.74	F3-F4
(De Silva et al., 2018)	South Asian	Liver biopsy	Kleiner	44*	90	30.0	3.25	18	100	100	79	0.85	F3-F4
	White			52*	79	29.1		43	98	91	81	0.93	F3-F4

Table 4.1, continued

NFS													
References	**Population**	**Diagnosis**	**Scoring**	**Age**	**Number of Patient**	**Prevalence of advanced fibrosis (%)**	**Cut off**	**Sensitivity**	**Specificity**	**PPV**	**NPV**	**AUROC (95% CI)**	**Prediction of disease severity**
(Angulo et al., 2007)	US, UK, Australia, Italy	Liver biopsy	Kleiner	48±14	253	29.2	-1.455	77	71	52	88	0.82 (0.76- 0.88)	F3-F4
							0.676	51	98	90	85		
(Qureshi, Clements, & Abrams, 2008)	UK (majority of non-Hispanic White)	Liver biopsy	Kleiner	41±9	331 morbidly obese	13.6	-1.455	96	N/A	N/A	98	N/A	F3-F4
							0.676	N/A	84	33	N/A		
							-1.455	85	N/A	N/A	87	N/A	F2-F4
							0.676	N/A	88	57	N/A		
				40±8	221 elevated ALT	15.8	-1.455	97	N/A	N/A	99	N/A	F3-F4
							0.676	N/A	87	39	N/A		
							-1.455	87	N/A	N/A	87	N/A	F2-F4
							0.676	N/A	91	68	N/A		
(Shah et al., 2009)	USA	Liver biopsy	Kleiner	48±12	541	23.1	N/A	N/A	N/A	N/A	N/A	0.77 (0.72-0.82)	F3-F4
(Cales et al., 2009)	France	Liver biopsy	Metavir	51±11	235	18.7	N/A	61	96	87	86	0.88 (0.83-0.94)	F2-F4
(Fujii et al., 2009)	Japan	Liver biopsy	N/A	59*	122	37.7	N/A	N/A	N/A	59	89	0.84	F3-F4
(McPherson et al., 2010)	UK	Liver biopsy	Kleiner	51±12	145	18.6	-1.46	78	58	30	92	0.81 (0.71-0.91)	F3-F4
							0.676	33	98	79	86		
(Ruffillo et al., 2011)	Argentina	Liver biopsy	Brunt	49*	238	15.5	-1.46	54	73	43	81	0.68 (0.57-0.78)	F3-F4
							0.676	14	100	100	76		
(Wong, Vergniol, et al., 2010)	French and Chinese	Liver biopsy	Kleiner	51±11	246	22.8	-1.46	73	70	44	89	0.75 (0.67-0.83)	F3-F4
							0.676	18	96	61	79		

Table 4.1, continued

References	Population	Diagnosis	Scoring	Age	Number of Patient	Prevalence of advanced fibrosis (%)	Cut off	Sensitivity	Specificity	PPV	NPV	AUROC (95% CI)	Prediction of disease severity
(Cichoz-Lach et al., 2012)	Polish-caucasians	Liver biopsy	Kleiner	43±14	126	21.4	-1.46	96	53	36	98	0.919 (0.84–0.97)	F3-F4
							0.676	89	90	71	97		
(Sumida et al., 2012)	Japan	Liver biopsy	Kleiner	52±15	576	16.3	-1.46	92	63	24	98	0.863	F3-F4
							0.676	33	96	50	92		
(Xun et al., 2012)	Chinese	Liver biopsy	Kleiner	37±10	152	15.8	-1.46	38	86	33	88	0.65 (0.52-0.79)	F3-F4
							0.676	8	100	100	85		
(Demir et al., 2013)	German	Liver biopsy	Kleiner	45±13	120	13.3	-1.455	75	93	63	96	0.96 (0.92-0.99)	F3-F4
							0.676	19	100	100	89		
(Perez-Gutierrez et al., 2013)	Latin	Liver biopsy	Brunt	49±13	243	11.1	0.676	53	87	26	95	0.72 (0.60-0.83)	F3-F4
(McPherson et al., 2013)	UK (Normal ALT)	Liver biopsy	Kleiner	54±11	70 (Normal ALT)	24.3	-1.46	82	51	35	90	0.85 (0.74-0.96)	F3-F4
							0.676	47	98	88	85		
	UK (Elevated ALT)			48±13	235 (Elevated ALT)	17.4	-1.46	71	65	29	92	0.80 (0.72-0.88)	F3-F4
							0.676	27	97	65	87		
(Kim, Kim, Talwalkar, et al., 2013)	USA	Liver biopsy and MRE	Kleiner	54 ±13	289	N/A	-1.46	91	43	46	90	0.80 (0.72-0.86)	F3-F4
							0.676	57	85	66	79		
(Mohamed et al., 2014)	Egypt	Liver biopsy	Kleiner	41±6	76	25.0	0.676	79	95	83	93	0.92(0.83-1.00)	F3-F4
							-1.455	90	85	65	96		
(Aykut et al., 2014)	Turkish	Liver biopsy	Kleiner	46±9	88	30.7	N/A	52	89	N/A	N/A	0.66 ±0.059	F2-F4
								63	82	N/A	N/A	0.74 ±0.059	F3-F4
								67	72	N/A	N/A	0.68±0.108	F4

Table 4.1, continued

References	Population	Diagnosis	Scoring	Age	Number of Patient	Prevalence of advanced fibrosis (%)	Cut off	Sensitivity	Specificity	PPV	NPV	AUROC (95% CI)	Prediction of disease severity
(Treeprasertsuk et al., 2017)	Thailand	Liver biopsy	N/A	41±13	139	6.5	-1.455	89	48	11	98	0.66	F3-F4
(De Silva et al., 2018)	South Asian	Liver biopsy	Kleiner	44*	90	35.6	0.676	9	98	67	76	0.86 (0.76-0.95)	F3-F4
	White			52*	79	29.1		48	100	100	82	0.95 (0.89-1)	

*Median age

N/A data not available

ALT, alanine transferase; *APRI*, aspartate aminotransferase to platelet ratio index; *AST*, aspartate aminotransferase; *BARD*, BMI, AST/ALT ratio, Diabetes; *ELF*, Enhanced Liver Fibrosis; *FIB-4*, Fibrosis-4; *MRE*, magnetic resonance elastography; *NFS*, NAFLD fibrosis score; *UK*, United Kingdom; *USA*, United States of America

4.2.3 AST/ALT (Aspartate aminotransferase/alanine aminotransferase) ratio

Being the simplest predictive model for advanced fibrosis, the AST/ALT ratio is a test that is easily available and calculated in routine laboratories. Elevated levels of both AST and ALT are often observed among NAFLD patients, with more prominent increase in ALT level compared to AST in the absence of cirrhosis (McPherson et al., 2010). Towards the progression to cirrhosis, AST may become higher than ALT resulting in AST/ALT ratio >1 (Anstee, Targher, & Day, 2013; Kaswala, Lai, & Afdhal, 2016). Most studies used a cut-off value of < 0.8 to exclude advanced fibrosis and a value of > 1 to rule in advanced fibrosis. Its sensitivity ranges from 40% to 94% using a low cut-off and a specificity of 90%-95% using a high cut-off, with AUROC ranges from 0.57 to 0.84. Goh et al. suggests that interpretation of AST/ALT should be done with caution in elderly patients because as age increases, ALT declines progressively whilst AST remains stable (Goh et al., 2015). This can cause higher AST/ALT ratio in patients with or without advanced fibrosis. To enhance its accuracy, AST/ALT ratio is often incorporated with other non-invasive scores such as BARD, NFS and FIB-4 score (Angulo et al., 2007; Harrison et al., 2008; Sterling et al., 2006).

4.2.4 AST to Platelet Ratio Index (APRI)

APRI is calculated based on the elevation of AST and the decrease in platelet count as liver disease progresses and changes in splenic blood flow. APRI was first derived in a study involving patients with hepatitis C and was less commonly being used in NAFLD (Wai et al., 2003). Nevertheless, the test was validated in several studies involving NAFLD patients. Some utilised dual cut-off where >0.5 indicates presence of significant fibrosis (≥F2) (Cales et al., 2009) and >1 or >1.5 indicate presence of advanced fibrosis (≥F3)(Kim, Kim, Talwalkar, et al., 2013; Xun et al.,

2012). The sensitivity of APRI is relatively low (65-79%) when a low cut-off is used and even lower (18-67%) when high cut-off is used. The negative predictive value (NPV) ranges from 77-93% but the positive predictive value (PPV) ranges from only 23 to 73% , suggesting APRI may be more useful to rule out patients unlikely to have advanced fibrosis but less applicable in predicting the presence of advanced fibrosis.

Study by Cales et al.(2009) is the only study that used Metavir system to stage fibrosis. This system was initially used in staging of chronic hepatitis C and was less being used as a histological reference in NAFLD compared to the Kleiner system (Kleiner et al., 2005). For Kleiner system, presence of either perisinusoidal or portal or periportal fibrosis is cateogorised under F1 stage whilst a mixture of them is cateogorised as F2 stage. On the other hand, Metavir system defines F2 as septal fibrosis (Bedossa & Poynard, 1996). This causes the result derived from both system to be comparable only in F3 (bridging fibrosis) and F4 (cirrhosis).

4.2.5 BARD

BARD score was derived from a logistic regression analysis from 827 NAFLD patients, encompassing body mass index (BMI), AST/ALT ratio and presence of type II diabetes mellitus (Harrison et al., 2008). The sensitivity for the detection of advanced fibrosis ranges from 42 % to 95% while specificity ranges from 21% to 89% by using a cut-off point of ≥ 2, with AUROC ranges from 0.56 to 0.87. Most studies reported high NPVs (77-97%) but BARD score has a low sensitivity in studies involving Chinese or South Asian cohort (De Silva et al., 2018; Wong, Vergniol, et al., 2010; Xun et al., 2012). Since most NAFLD patients have diabetes and obesity, it could easily result in score of ≥ 2, limiting its utility in clinical practise (Dyson et al., 2014).

4.2.6 Enhanced Liver Fibrosis (ELF) Test

ELF test is an algorithm that consists of three extracellular matrix components, namely hyaluronic acid (HA), procollagen type III amino-terminal peptide (PIIINP), and tissue inhibitor of metalloproteinase I (TIMP-1). HA is a component of extracellular matrix synthesised by hepatic stellate cells and degraded by hepatic sinusoids that was found to correlate well with fibrosis stage (Chwist et al., 2014; Santos et al., 2005). PIIINP is released during synthesis and deposition of collagen or degradation of existing collagen fibrils and was able to discriminate between NAFL and NASH or advanced fibrosis (Tanwar et al., 2013). TIMP-1 inhibits activities of matrix metalloproteinases and was correlated with severity of cirrhosis (Busk et al., 2014). The Original European Liver Fibrosis (OELF) test was first characterised and validated in a mixed cohort of chronic liver disease patients, of which 61 of them were diagnosed with NAFLD (Rosenberg et al., 2004). ELF is a modified algorithm by removing the age factor from the OELF test and was first validated in a total of 192 adult NAFLD patients for the staging of liver fibrosis (Guha et al., 2008). Elimination of age facilitated automated analysis of samples independent of the demographic data, thereby reducing transcription errors and facilitating population screening. The ELF panel demonstrated better performance in distinguishing advanced fibrosis (stage 3-4) with AUROC of 0.90 compared to its performance in distinguishing fibrosis from no fibrosis (AUROC 0.76)(Guha et al., 2008). The performance of ELF was also assessed in paediatric patients and showed higher level of performance (AUC>0.95 for advanced fibrosis and AUC>0.92 for presence of any fibrosis)(Alkhouri et al., 2011; Nobili et al., 2009).

ELF score was recently endorsed by the National Institute of Clinical Excellence (NICE) in United Kingdom in the diagnosis and management of NAFLD (Glen, Floros, Day, Pryke, & Guideline Development, 2016). Patients with ELF score >10.51 are diagnosed with advanced fibrosis while patients with ELF < 10.51 are recommended to

screen for advanced fibrosis using ELF every 3 years (Glen et al., 2016). Compared to other imaging modalities such as vibration controlled transient elastography (VCTE) and magnetic resonance elastography (MRE), ELF score is also found to be more cost effective (Cheah, McCullough, & Goh, 2017). Despite the fairly high accuracy in distinguishing different stages of liver fibrosis, ELF requires specialised testing which are not readily available in normal biochemical blood tests and will incur extra costs compared to the other tests.

4.2.7 FIB-4

The initial development of FIB-4 score was to assess fibrosis in hepatitis C/HIV (Vallet-Pichard et al., 2007). Using readily available clinical parameters such as ALT, AST, platelet count, and age, the score is easy to be calculated and was first validated in NAFLD using a cohort of 541 patients from the NASH CRN network (Shah et al., 2009). The AUROC for FIB-4 ranges from 0.76 to 0.95, and comparable to NFS. A FIB-4 score of ≥ 2.67 has PPV of 59-80% whilst score of ≤ 1.30 has NPV of 88-95%. Interestingly in multiple studies comparing the diagnostic performance of non-invasive tests, FIB-4 outperforms other tests by having higher AUROC (Kim, Kim, Talwalkar, et al., 2013; McPherson et al., 2013; McPherson et al., 2010; Mohamed et al., 2014; Shah et al., 2009; Sumida & Yoneda, 2018; Treeprasertsuk et al., 2017; Wong, Vergniol, et al., 2010; Xun et al., 2012). This was supported by a recent meta-analysis which found the highest diagnostic accuracy associated with FIB-4, followed by NFS and finally BARD (AUROC 0.8496, 0.8355, 0.7625, respectively)(Sun et al., 2016). In addition to the diagnostic accuracy, multiple retrospective and longitudinal studies found that higher FIB-4 score is associated with higher rate of mortality (Angulo et al., 2013; Kim, Kim, Kim, & Therneau, 2013; Xun et al., 2014).

4.2.8 NAFLD Fibrosis Score (NFS)

NFS is an algorithm developed in a multicentre study involving 480 patients in the discovery phase and 253 patients in the validation phase. A regression formula is derived from age, presence of diabetes, BMI, platelet count, albumin and AST/ALT ratio. Two cut-offs were established, where a score of <-1.455 is used to exclude advanced fibrosis and score of >0.676 indicates presence of advanced fibrosis. Guideline from the American Association for the Study of Liver Diseases (AASLD) advocates the use ofNFS to identify NAFLD patients with higher likelihood of bridging fibrosis and/or cirrhosis (Chalasani et al., 2018).

The NFS since then was validated in various studies with AUROC ranging from 0.65 to 0.96, sensitivity from 38% to 97% and NPV from 81% to 98% when the low cut -off value is applied. When high cut-off value is used, the specificity ranges from 84% to 100% and PPV from 26% to 100%. NFS can be calculated from routinely available biochemical tests, with good diagnostic performance but is limited where 20-58% fall in the indeterminate score between -1.455 and 0.676 (Musso et al., 2011). The applicability of this scoring system was validated in the Western and Middle East cohort, showing high PPV ranging from 83% to 100% when the high cut off value (>0.676) is used to predict advanced fibrosis, similar to the study by Angulo et al. which has a PPV of 90% (De Silva et al., 2018; Demir et al., 2013; Mohamed et al., 2014; Ruffillo et al., 2011). However the validation of NFS is limited to Asian population. In a Chinese cohort, high NPVs was found but there were only two patients exceeding the high cut-off (>0.676)(Wong et al., 2008). This accentuates that NFS may not be as applicable in the Asian cohort where metabolic syndromes may occur at a lower BMI.

4.3 Methodology

4.3.1 Subjects recruitment

Consecutive recruitment of 122 NAFLD patients were performed in University Malaya Medical Centre (UMMC) based on presence of increased liver echogenicity (compared to renal cortex) at ultrasound examination. This group of patients is of the same group as previously described in Chapter 3. The exclusion criteria includes (1) alcohol consumption of more than 20 g per day, (2) presence of coexisting liver disease such as autoimmune hepatitis, chronic viral hepatitis, primary biliary cirrhosis, Wilson's disease, biliary obstruction, hemochromatosis, α1-antitrypsin deficiency and drug induced liver disease. Written informed consent was obtained from each patient. The study protocol was approved by the Medical Ethics Committee of UMMC.

4.3.2 Clinical evaluation and Biochemistry profiling

Anthropometric measurements were taken including body weight, body height and waist circumference. Body mass index (BMI) was calculated as weight divided by height squared (kg/m^2). Fasting blood samples were taken and subjected to laboratory testing for ALT, AST, GGT, LDL-cholesterol, HDL-cholesterol, total cholesterol, triglycerides, glycosylated haemoglobin (HbA1c), albumin and platelet count. Due to the constraint in samples availability, the ELF test was performed on only 91 patients. The component assays for HA, PIIINP and TIMP-1 were performed using the ELFTM Test ADVIA Centaur® Kit (Siemens Healthcare Diagnostics, Germany) on the ADVIA Centaur CP immunochemical analyser (Siemens Healthcare Diagnostics, Germany) according to the manufacturer's instructions. The AST/ALT ratio, APRI,

BARD, FIB-4 and NFS were calculated based on the anthropometric measurements and biochemistry profiles. The calculation of the non-invasive scores is as follow:

$$\text{AST/ALT ratio} = \frac{\text{AST (IU/L)}}{\text{ALT (IU/L)}}$$

$$\text{APRI} = \frac{\frac{\text{AST (IU/L)}}{\text{Upper limit of AST (IU/L)}}}{\text{Platelet count (x10}^9\text{/L)}} \times 100$$

BARD = Sum of score of three variables (BMI ≥ 28 = 1 point, AST/ALT ratio ≥ 0.8 = 2 points, diabetes =1 point)

ELF score = 2.278 + 0.851 ln(HA(ng/mL)) + 0.751 ln(PIIINP(ng/mL)) + 0.394 ln(TIMP1(ng/mL))

$$\text{FIB-4 score} = \frac{\text{Age (year)x AST(IU/L)}}{\text{Platelet count (x10}^9\text{/L)} \times \sqrt{\text{ALT (IU/L)}}}$$

NFS = −1.675 + 0.037 × age (year) + 0.094 × BMI (kg/m2) + 1.13 × impaired fasting glucose/diabetes (yes = 1, no = 0) + 0.99 × AST/ALT ratio–0.013 × platelet (×109/L)–0.66 × albumin (g/dL)

4.3.3 Histological Assessment

On the same day when blood collection was carried out, liver biopsy was performed on NAFLD patients followed by histopathological examination. All biopsy specimens were at least of 1.5cm length consisting of at least six portal tracts. The histological grading and staging of NAFLD is based on recommendations by NASH Clinical Research Network, with grading of macrovesicular steatosis from 0 to 3, necroinflammatory activity from 0 to 3, and staging of fibrosis from 0 to 4 (Kleiner et al., 2005). The patients were categorised as having significant fibrosis when they have

stage 2 fibrosis and above whereas stage 3 fibrosis and above is indicated as advanced fibrosis.

4.3.4 Statistical analysis

For the demographic and clinical parameters, continuous variables were presented as mean ± standard deviation for normally distributed data or median (interquartile range) for non-normally distributed data. The categorical variables were presented as frequency (percentage). Statistical analyses were performed using SPSS version 16.0 (IBM Corp, Chicago, IL, USA). Categorical data were compared using Chi square ($\chi 2$) test. For two groups comparisons, independent t-test was performed to compare normally distributed variables while Mann-Whitney U test was employed for skewed variables.

The diagnostic accuracies of all the non-invasive tests in predicting advanced fibrosis were calculated using area under ROC curve and the 95% confidence interval (CI) were determined. The sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) were calculated based on the cut-off values on previous publications.

4.4 Results

A total of 122 NAFLD patients were reviewed. The mean age of the patients at time of liver biopsy was 50.0 (±11.4 SD) years. BMI-defined obesity (≥25 kg/m^2) was observed in 36% of the patients and 48% of the patients were diabetic. All patients were included in the clinical scoring tests. NAFL was found in 25 patients (20%) whilst 97 patients (80%) were non-alcoholic steatohepatitis (NASH), a severe stage of NAFLD. Thirty-nine patients (32%) had fibrosis score=0, 53 patients (43%) had fibrosis score=1, six patients (6%) had fibrosis score=2, and 24 patients (20%) had advanced fibrosis or cirrhosis (Kleiner fibrosis score 3-4). Due to the availability of sample, serum samples of 91 NAFLD patients were subjected to ELF testing, of which 34 (37%) patients were without fibrosis, 38 patients (42%) had stage 1 fibrosis, 4 patients (4%) had stage 2 fibrosis and 15 patients (16%) had advanced fibrosis.

Since the scoring systems aim to identify patients with advanced fibrosis, the clinical and biochemistry parameters of patients with no/mild fibrosis (fibrosis=0-2) were compared against patients with advanced fibrosis (stage 3-4) as in Table 4.2. Patients with advanced fibrosis were significantly much older (P = 0.001) with greater levels of HbA1c (P=0.003), AST (P=0.006) and GGT (P=0.005), but lower levels of platelet count (P = 0.002), LDL cholesterol (P = 0.012) and total cholesterol (P = 0.042). As for the scoring systems, patients with advanced fibrosis exhibited higher AST/ALT ratio (P=0.005), APRI (P <0.0001), BARD score (P=0.001), NFS (P< 0.0001), and FIB-4 score (P< 0.0001). For the ELF test, patients with advanced fibrosis had significantly higher level of HA (P = 0.002), PIINP (P = 0.012) but not TIMP-1.

Table 0.2: Demographic and clinical data of the NAFLD patients

Characteristics	*n* (%) or Mean ± SD		*P* value
	Fibrosis 0-2 (*n*=98)	Fibrosis 3-4 (*n*=24)	
Gender			
Males	51 (52)	10 (42)	0.362
Females	47 (48)	14 (58)	
Age (years)	51.0(17.0)	59.0(10.0)	0.001
BMI (kg/m^2)	28.6(4.7)	29.5(6.0)	0.475
BMI category			
Normal (<25)	15 (15)	3 (12)	0.539
Overweight (25-29.9)	50 (51)	10 (42)	
Obese (≥ 30)	33 (34)	11 (46)	
HbA1c (%)	5.9(1.0)	6.9(1.8)	0.003
Diabetes	41 (42)	17 (71)	0.011
Waist circumference (cm)	95.0(12.4)	98.9(16.5)	0.062
HDL cholesterol (mg/dL)	46.0(12.9)	46.4(14.7)	0.629
LDL cholesterol (mg/dL)	122.8±3.8	101.3±34.7	0.012
Total cholesterol (mg/dL)	196.4±43.7	176.3±38.9	0.042
Triglycerides (mg/dL)	141.7(70.9)	128.4(50.9)	0.230
Albumin (d/dL)	4.3±0.3	4.1±0.4	0.054
HA (ng/mL)	23.5 (23.2)	57.3 (64.5)	0.002
PIINP (ng/mL)	8.6 (4.5)	15.0 (6.6)	0.012
TIMP-1 (ng/mL)	193.8 ±60.8	197.5±41.2	0.824
Platelet (x10^9/L)	278.6±64.4	234.7±48.8	0.002
AST (IU/L)	38.0(31.0)	57.5(39.0)	0.001
ALT (IU/L)	66.5(63.0)	77.0(74.0)	0.228
GGT (IU/L)	71.0(74.0)	116.0(129.0)	<0.0001
AST/ALT ratio	0.6(0.2)	0.7(0.2)	0.005
APRI	0.4(0.4)	0.7(0.5)	<0.0001
BARD score	1.0(1.0)	2.0(2.0)	0.001
ELF	8.7±0.9	9.7±1.1	<0.0001
FIB-4 score	0.8(0.6)	1.7(0.7)	<0.0001
NFS	-2.5±1.2	-1.0±1.1	<0.0001

Data are expressed in mean ± SD for normally distributed continuous data; median (interquartile range) for non-normally distributed continuous data and percentage for categorical data.

ALT, alanine transferase; *APRI*, aspartate aminotransferase to platelet ratio index; *AST*, aspartate aminotransferase; *BARD*, BMI, AST/ALT ratio, Diabetes; *BMI*, body mass index; *ELF*, Enhanced Liver Fibrosis; *FIB-4*, Fibrosis-4; *GGT*, gamma-glutamyl transpeptidase; *HA*, hyaluronic acid; *HbA1c*, haemoglobin A1c; *HDL*, high-density lipoprotein; *LDL*, low-density lipoprotein; *NAFLD*, non-alcoholic fatty liver disease; *NFS*, NAFLD fibrosis score; *PIIINP*, procollagen type III amino-terminal peptide; *TIMP-1*, tissue inhibitor of metalloproteinase I

The AUROCs were calculated to compare the diagnostic performance of the scoring systems and the result is demonstrated in Figure 4.1. The FIB-4 score has the best diagnostic accuracy for advanced fibrosis (AUROC 0.857; 95% CI 0.78 – 0.94), followed by NFS (AUROC 0.836; 95% CI 0.75 – 0.92), APRI (AUROC 0.759; 95% CI 0.66 – 0.86),ELF (AUROC 0.757; 95% CI 0.61-0.91), BARD score (AUROC 0.702; 95% CI 0.58 –0.82), and AST/ALT ratio (AUROC 0.687; 95% CI 0.57 – 0.80). Since NFS and FIB-4 score gave the highest AUROC, the effects of these predictors were combined to yield an AUROC of 0.843; 95% CI 0.76-0.93, where the effect did not outperform the use of FIB-4 score alone. Some literature also tested the diagnostic utility of these scoring systems in identification of significant fibrosis (stage 2-4), hence the AUROCs were also calculated and shown in Table 4.3 and Figure 4.1. The NFS has the highest AUROC of 0.785 (95%CI 0.69-0.88), followed by FIB-4 with similar AUROC of 0.782 (95%CI 0.69–0.88), ELF (AUROC 0.747 95%CI0.61-0.88), APRI (AUROC 0.744, 95%CI0.65–0.84), BARD (AUROC 0.668 95%CI 0.55-0.78),and lastly AST/ALT ratio (AUROC 0.603 95% CI 0.48-0.72).

The sensitivity, specificity, positive predictive values (PPV), and negative predictive value (NPV) of each test were calculated based on previously published cut-offs as shown in Table 4.3. For tests having dual cut-off, Figure 4.2 shows the percentage of patients correctly identified (true negative and true positive), falsely identified (false positive and false negative) and also proportion of patients within the indeterminate range (higher than the lower cut-off but lower than the higher cut-off).

Table 0.3: Comparison of the performance of non-invasive scores for the diagnosis of advanced fibrosis and significant fibrosis in NAFLD patients

Advanced fibrosis						
Test	**AUROC (95% CI)**	**Cut-off**	**Sensitivity (%)**	**Specificity (%)**	**PPV (%)**	**NPV (%)**
AST/ALT ratio	0.687 (0.57-0.80)	0.8	37.5	83.7	36.0	84.5
		1	16.7	91.8	33.3	81.8
APRI	0.759 (0.66-0.86)	0.5	83.3	59.2	33.3	93.5
		1	37.5	91.8	52.9	85.7
BARD score	0.702 (0.58-0.82)	2	70.8	63.3	32.1	89.9
NFS	0.836 (0.75-0.92)	-1.455	62.5	77.6	40.5	89.4
		0.676	4.2	99.0	50.0	80.8
FIB-4 score	0.857 (0.78-0.94)	1.3	79.2	84.7	55.9	94.3
		2.67	8.3	96.9	40.0	81.2
		3.25	4.2	98.0	33.3	80.7
ELF	0.757 (0.61-0.91)	9.8	40.0	84.2	33.3	87.7
		10.51	20.0	94.7	42.9	85.7

Table 4.3,continued

Significant fibrosis						
Test	**AUROC (95% CI)**	**Cut-off**	**Sensitivity (%)**	**Specificity (%)**	**PPV (%)**	**NPV (%)**
AST/ALT ratio	0.603 (0.48-0.72)	0.8	30.0	82.6	36.0	78.4
		1	13.3	91.3	33.3	76.4
APRI	0.744 (0.65-0.84)	0.5	80.0	60.9	40.0	90.3
		1	30.0	91.3	52.9	80.0
BARD score	0.668 (0.55-0.78)	2	66.7	64.1	37.7	85.5
NFS	0.785 (0.69-0.88)	-1.455	56.7	78.3	45.9	84.7
		0.676	3.3	98.9	50.0	75.8
FIB-4 score	0.782 (0.69-0.88)	1.3	66.7	84.8	58.8	88.6
		2.67	6.7	96.7	40.0	76.1
		3.25	3.3	97.8	33.3	75.6
ELF	0.747 (0.61-0.88)	9.8	42.1	86.1	44.4	84.9
		10.51	15.8	94.4	42.9	81.0

ALT, alanine transferase; *APRI,* aspartate aminotransferase to platelet ratio index; *AST*, aspartate aminotransferase; *BARD*, BMI, AST/ALT ratio, Diabetes; *ELF*, Enhanced Liver Fibrosis; *FIB-4*, Fibrosis-4; *NFS*, NAFLD fibrosis score

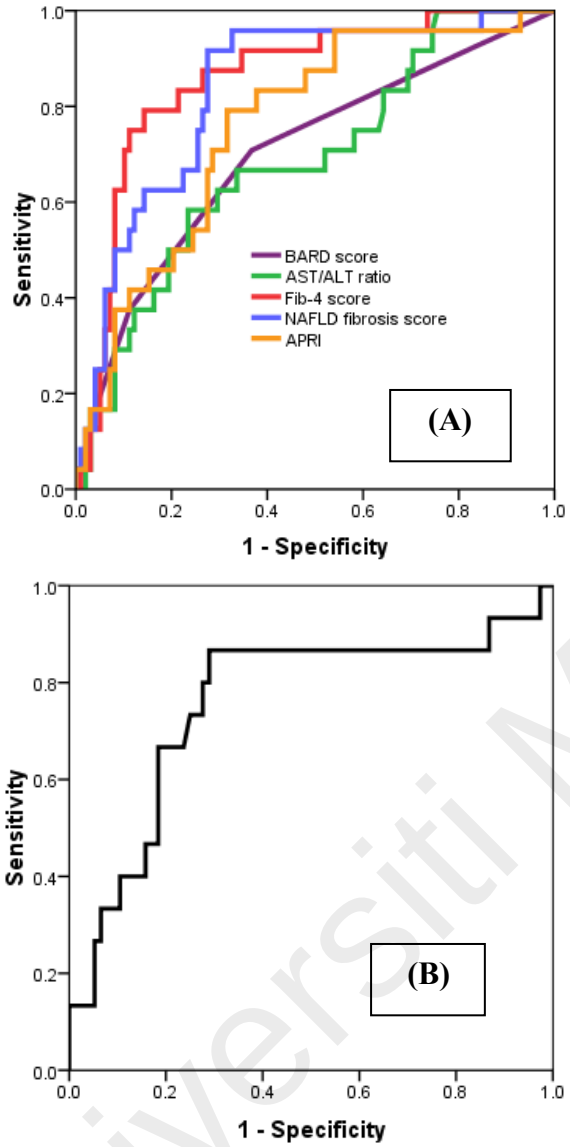


Figure 0.1(A): Receiver operating characteristic (ROC) curves for the noninvasive scores for diagnosis of advanced fibrosis (stage 3-4) in 122 patients (B) Receiver operating characteristic (ROC) curves for ELF for diagnosis of advanced fibrosis (stage 3-4) in 91 patients

ALT, alanine transferase; *APRI*, aspartate aminotransferase to platelet ratio index; *AST*, aspartate aminotransferase; *BARD*, BMI, AST/ALT ratio, Diabetes; *ELF*, Enhanced Liver Fibrosis; *FIB-4*, Fibrosis-4

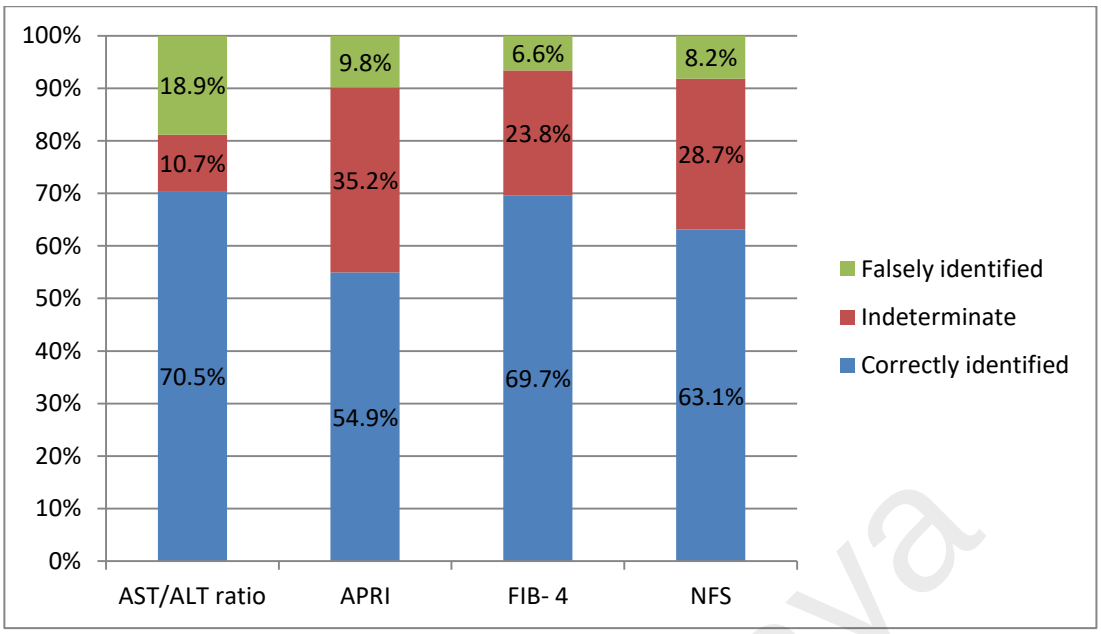


Figure 4.2: Utility of non-invasive scores having dual cut-off

Table 0.4: Percentage of patients avoiding liver biopsy using non-invasive tests

Test	Cut-off	Patients avoiding liver biopsy	False negative result
AST/ALT ratio	0.8	99/122 (81%)	15 (15%)
APRI	0.5	62/122 (51%)	4 (6%)
BARD score	2	69/122 (57%)	7 (10%)
NFS	-1.455	85/122 (70%)	9 (11%)
FIB-4 score	1.3	88/122 (72%)	5 (6%)
ELF	9.8	76/91 (84%)	12 (16%)

ALT, alanine transferase; *APRI*, aspartate aminotransferase to platelet ratio index; *AST*, aspartate aminotransferase; *BARD*, BMI, AST/ALT ratio, Diabetes; *ELF*, Enhanced Liver Fibrosis; *FIB-4*, Fibrosis-4; *NFS*, NAFLD fibrosis score

The FIB-4 score has low sensitivity (8.3%) and PPV (40%) when the high cut-off (>2.67) was used to distinguish advanced fibrosis from other NAFLD patients. A higher cut-off of 3.25 was used in a study by McPherson et al. (2013) but this resulted in a lower sensitivity (4%) and PPV (33%) in our cohort. There were only five (%) patients having score of more than 2.67 and of these, only two had advanced fibrosis, resulting in low sensitivity and PPV. When the low cut-off (<1.3) was used to rule out advanced fibrosis, 88 patients (72%) were identified of which 83(94%) were correctly staged. In addition, out of the 98 patients without advanced

fibrosis, 83 (85%) were correctly excluded using the lower cut-off, resulting in high specificity (85%) and NPV (94%). If liver biopsy was not performed on patients with FIB-4 score lower than the lower cut-off (1.3), biopsy could be avoided in 72% patients with only five (5.7%) of patients with advanced fibrosis being under-staged. Overall, by using a score of <1.3 to rule out advanced fibrosis and >2.67 to predict advanced fibrosis, 69.7% were correctly identified, 23.8% were in the indeterminate region and only 6.6% were falsely identified. Among the patients with indeterminate score, 17 had advanced fibrosis and 12 without.

The NFS has a low cut-off of –1.455 and a high cut-off of 0.676. There were only two (2%) patients in our cohort had NFS of >0.676, resulting in low sensitivity (4%) and PPV (50%). Eighty five (70%) patients had NFS of less than -1.455 and 76 of them were correctly staged, causing high NPV of 89%. However, its specificity (78%) was not that high, being only able to rule out advanced fibrosis in 76 patients out of 98 patients without advanced fibrosis. NFS could help 70% of the patients with score <-1.455 to avoid liver biopsy, with only 11% of them being misclassified. By using a dual cut-off system, NFS was able to correctly identify 63.1% and falsely identify 8.2% with 28.7% of the patients within the indeterminate range. The indeterminate range consists of 14 patients with advanced fibrosis and 21without.

Previous studies on APRI test suggest using a cut-off point of >1 to indicate presence of advanced fibrosis and a cut-off point of <0.5 to indicate absence of fibrosis. A total of 62 (%) patients had a APRI score of <0.5, of which 58 (94%) were correctly ruled out for having advanced fibrosis whilst four subjects were falsely identified, resulting in the highest NPV (94%) among all the other tests but a sensitivity of 59% as it was only able to exclude 58 patients out of the 98 who did

not have advanced fibrosis. APRI was also the score with the highest sensitivity (83%) if the lower cut-off was used but the sensitivity decreased to 38% using the higher cut-off. The higher cut-off point of 1 correctly identify nine (53%) of the 98 patients with advanced fibrosis and misclassified eight of them. If liver biopsy was avoided in patients below the lower cut-off, 62 (51%) of patients could be benefited. If the dual cut-off system was implemented, 55% would be appropriately classified, 10% patients miscategorised and a total of 43 (35%) subjects had APRI values in the indeterminate range.

The BARD score has a single cut-off value of 2, where a score ≥ 2 indicates presence of advanced fibrosis and a score of <2 rule out advanced fibrosis. It correctly identified 17 out of 24 patients with advanced fibrosis, giving a satisfactory sensitivity of 71% but out of 53 patients with score ≥ 2, only 17 had advanced fibrosis, resulting in a low PPV (32%). Nevertheless, out of 69 patients with a score <2, 62 patients did not have advanced fibrosis, giving a high NPV (90%) and out of 98 patients without advanced fibrosis, 62 had score <2, giving a specificity of 63%. If BARD score was used, it can help 69 (57%) patients to avoid liver biopsy, with only 7 (10%) of them being misclassified. Overall, BARD correctly identified 65% patients and the remaining 35% weremisclassified.

The AST/ALT ratio has lower cut-off of <0.8 to rule out advanced fibrosis and ≥ 1 to identify presence of advanced fibrosis. The sensitivity was very low (17%) when the high cut-off was used, with only four patients with advanced fibrosis being identified from 24 patients. The PPV was also not satisfactory (33%), with only four true positive out of 12 patients who had a score of ≥ 1. When a low cut-off was used, it correctly ruled out advanced fibrosis in among 82 patients out of the 97 with a score of <0.8, having good NPV(85%). Being able to exclude the 82

patients from 98 patients with advanced fibrosis, AST/ALT ratio also has a good specificity (84%). By using AST/ALT ratio, 97 (%) patients were able to avoid liver biopsy but 15 (15%) of them were incorrectly classified, of which one patient had cirrhosis but having AST/ALT ratio of 0.71. If AST/ALT ratio was used, it was able to classify 71% of the patients correctly, whilst 19% were not managed appropriately and 11% fell within the indeterminaterange.

The ELF test was carried out in 91 NAFLD patients. By using a cut-off value of 9.8 as suggested by the manufacturer to identify advanced fibrosis, a total of 76 patients (84%) patients got to avoid liver biopsy with 64 patients correctly ruled out of advanced fibrosis but false negative in eight patients having stage 3 fibrosis and one with cirrhosis. The NICE guideline advocates the use of 10.51 as a cut-off to diagnose advanced fibrosis but this cut-off gave a low PPV of 43%, where out of seven patients having ELF above 10.51, only three had advanced fibrosis. The sensitivity was also low (20%), being only able to identify 3 out of all 15 patients with advanced fibrosis. Of all the ELF components, elevated TIMP-1 was not associated with the occurrence of advancedfibrosis.

4.5 Discussion

Among the composite scores validated, NFS and FIB-4 have higher AUROC for advanced fibrosis. All the non-invasive tests included in this study demonstrated AUROC ranging from 0.697 to 0.857 in predicting advanced fibrosis (F3-4). The diagnostic accuracy of these tests in identifying significant fibrosis (F2-4) are less satisfactory, with AUROC ranging from 0.603 to 0.785. This is consistent with the other studies where most of these tests were initially developed for advanced fibrosis. It is found that AST/ALT ratio, APRI, BARD score, NFS and FIB-4 score had high NPV (>84%) especially APRI and FIB-4 which had a NPV of 94%. Consistent with the findings in other studies (De Silva et al., 2018; Sumida et al., 2012; Treeprasertsuk et al., 2017; Wong et al., 2008), high NPV using the low cut-off point suggests their potential to be used clinically to rule out advanced fibrosis in patients with NAFLD.

In the absence of an established treatment, the purpose of ruling out advanced fibrosis in NAFLD patients is to avoid performing liver biopsy in patients who have lower risk of progressive disease. Due to the increasing number of NALFD patients being referred to liver clinics for further evaluation, these non-invasive tests should also be assessed based on their ability to substantially reduce the number of patients who undergo liver biopsy if it is implemented. The European Association for the Study of the Liver (EASL) clinical practice guidelines also endorse the use of non-invasive methods as the first line triage to identify patients with lower risk of advanced fibrosis or cirrhosis, especially in patients with metabolic syndrome or diabetes (European Association for the Study of the Liver, 2016). Being the scoring system that has the easiest calculation, AST/ALT ratio can help the most number of patients (81%) to avoid liver biopsy when the low cut-off (<0.8) is used. The drawback of this test is the lower

accuracy (85%) in excluding advanced fibrosis, where out of the 97 patients with score lower than 0.8, 14 patients have stage 3 fibrosis and 1 patient has cirrhosis (false negative).

In contrast if FIB-4 is used, there are 72% of patients having FIB-4 score below lower cut-off (<1.3) where they get to avoid liver biopsy, with a high chance (94%) that they would have been identified correctly. The mere 6% misclassification includes 5 cases of stage 3 fibrosis having scores below 1.3. In this study, FIB-4 score has the highest diagnostic performance, not only that it has the highest AUROC (0.86) and NPV (94%), it is easily derived from laboratory parameters that are inexpensive and readily available such as age, ALT, AST and platelet count. As compared to ELF, it is inexpensive and does not involve measurement of matrix turnover markers which require specialist testing. Despite having a high NPV (94%), FIB-4 score has only moderate sensitivity (79%), showing it is more useful to exclude NAFLD patients with advanced fibrosis rather than to correctly identify patients with advanced fibrosis. FIB-4 score is also more superior compared to other tests based on the proportion of liver biopsy correctly avoided (94%). Other studies also found FIB-4 to be outperforming other tests in terms of AUROC and sensitivity and NPV (Kim, Kim, Talwalkar, et al., 2013; McPherson et al., 2013; Shah et al., 2009; Sumida et al., 2012; Treeprasertsuk et al., 2017). In addition, the diagnostic performance and NPV of FIB-4 was also found to be consistently better compared to other scoring systems when performed in patients with normal ALT (McPherson et al., 2013). This demonstrates the potential application of FIB-4 in screening of advanced fibrosis in general population where NAFLD patients may have normal ALT level.

In contrast to the NPV, the PPV for each test were modest ranging from 32% to 56%, suggesting these tests are insufficient to diagnose advanced fibrosis, especially when the result would cause serious repercussions. Therefore it would seem appropriate to consider liver biopsy in all patients who have a value above the lower cut-off. Despite both Fib-4 and NFS showing high AUROC in identification of advanced fibrosis, the limitation here is the applicability of the high cut-off in Asian population. NFS is also a test widely being validated and is endorsed by American Association for the Study of Liver Diseases (AASLD) to exclude the patients having lower risk for advanced fibrosis (Chalasani et al., 2018). It gives high NPV (89%) for ruling out advanced fibrosis when the lower cut-off (-1.455) is used, consistent with findings in other studies(Kim, Kim, Kim, et al., 2013; McPherson et al., 2013; Wong, Vergniol, et al., 2010; Xun et al., 2012). Despite this, NFS is found to be inferior to FIB-4 not only in its diagnostic accuracy (AUROC 0.836 vs 0.857) but also in the NPV (89% vs 94%). Interestingly, a study by Angulo et al. (2007) showed that NFS >0.676 has AUROC with PPV of 90% in detecting advanced fibrosis. This is however in contrast with our finding that the PPV is only 50%, but are consistent with studies by (De Silva et al., 2018; Fujii et al., 2009; Sumida et al., 2012; Treeprasertsuk et al., 2017; Wong, Vergniol, et al., 2010). This is because there are only 2 out of 122 (2%) NAFLD patients in this study who have NFS above 0.676, similar to studies in Chinese cohorts (1%) and Japanese cohort (8%) where NFS above the upper cut-off (>0.676) is uncommon (Sumida et al., 2012; Wong, Vergniol, et al., 2010; Xun et al., 2012). This limits the applicability of NFS to predict advanced fibrosis in Asian population as metabolic syndromes and NASH tend to develop at a lower BMI, which is one of the factors taken into account in the calculation of the score. In fact, there is no significant difference in the BMI value in our study cohort between patients with advanced fibrosis and patients without. Consistent with our findings, studies in Asians also showed mean BMI of 26-29 in patients with

advanced fibrosis (De Silva et al., 2018; Sumida et al., 2012; Wong, Vergniol, et al., 2010; Xun et al., 2012), which is lower than the mean BMI of 32-36 in the Western and Middle East cohort (Angulo et al., 2007; Mohamed et al., 2014). Interestingly, NFS was superior to FIB-4 and APRI in predicting overall mortality rate in two separate NAFLD cohorts, one of which also have a low prevalence (3.2%) of patients having NFS above 0.676. Therefore larger longitudinal studies are warranted in Asian population in accessing its applicability in predicting mortality. Based on the AUROC curves, the diagnostic accuracy of both NFS and FIB-4 are consistently better than other tests. In improving the performance, new thresholds for NFS can be considered for Asian patients that would maximise its sensitivity and specificity.

One of the aspects that should be considered in non-invasive testing with dual cut-off is the proportion of patients where identification can be made with confidence and the remaining proportion of patients whom test results are in the indeterminate range. Four of the tests including BARD, NFS, FIB-4 and APRI have high NPV (>84%). Despite APRI having the highest NPV (94%) and highest sensitivity (83%) among all the tests, there are 43 (35%) patients having the APRI score which falls in the indeterminate range and liver biopsy has to be carried out to confirm the stage of fibrosis. The AST/ALT ratio has the smallest proportion of patients (11%) in the indeterminate range but with the highest rate of false positive and false negative (19%). Despite having 24% of patients in the indeterminate region, FIB-4 test has the lowest rate of false negative and false positive of only 7%. BARD does not have a dual cut-off value and thus having the highest proportion of patients (35%) being mistakenly identified. Besides this, BARD utilised a sum of score taking BMI, AST/ALT ratio and diabetes into account and its diagnostic accuracy was shown to be lower in Asian cohort (De Silva et al., 2018; Fujii et al., 2009; Wong, Vergniol, et al., 2010; Xun et al., 2012). Our study is consistent with findings from Fujii et al. that presence of advanced fibrosis

is associated with the presence of diabetes and elevated AST/ALT ratio but not associated with increased BMI.

The advantages of using AST/ALT ratio, APRI, BARD score, NFS and FIB-4 score is that they can be easily calculated from readily available clinical laboratory parameters. Implementation of these tests into daily clinical practice is relatively simple and will not incur extra costs, as the calculation could be carried out easily and rapidly by entering the relevant details into a pre-designed excel spreadsheet. In comparison, ELF is a commercial panel that is more costly as the component assays such as HA, PIIINP and TIMP-1 are not carried out routinely in the lab. In addition, the performance of the ELF panel in this study is moderate in its ability to distinguish advanced fibrosis. This could be due to the absence of significant upregulation of TIMP-1 in patients with advanced fibrosis compared to those without. The role of TIMP-1 in fibrogenesis requires further studies as their functional role is yet to be confirmed using mice models(Thiele et al., 2017). The validated studies on the utility of ELF in NAFLD patients are limited to only four adult studies and two paediatric studies. Of which two of the studies did not use Kleiner histological scoring system, making the results hard to be comparable (Kleiner et al., 2005; Lopez et al., 2017; Wahl et al., 2012). There is also evidence of higher level of performance of ELF panel in paediatric patients because unlike the adults, children are not likely to have the effects of aging and borderline comorbidities on extrahepatic extracellular matrix turnover and organ fibrogenesis(Alkhouri et al., 2011; Nobili et al., 2009).

One limitation of these scoring systems is their inability to estimate individual fibrosis stage like liver biopsy. This is however not achievable by any other non-invasive biomarkers to date. Furthermore these biomarkers are also not dedicated to be used to distinguish NASH from NAFL. It was found that serum marker of apoptosis

such as CK18 was able to identify NASH but again its applicability is limited in the Malaysian population (Chan, Sthaneshwar, Nik Mustapha, & Mahadeva, 2014). FIB4 is potentially useful to evaluate the likelihood of having advanced fibrosis among subjects with NAFLD, thus aid to identify patients who should undergo liver biopsy as patients with advanced fibrosis are more in need of close monitoring and screening for hepatocellular carcinoma.

To date liver specific treatments available for NASH are limited and this causes ignorance in patients and poses a challenge to physicians as to when to perform liver biopsy. Consequently some patients with advanced fibrosis may be overlooked and develop severe liver complications. Risk stratification of patients is needed and to reduce the number of patients undergoing liver biopsy, staging of disease should be aided with an algorithm involving simple non-invasive tests, in addition to clinical judgement. For example, FIB-4 score can be used as a first-line investigation to exclude advanced fibrosis, followed by a more sophisticated test such as FibroScan to diagnose advanced fibrosis in patients with elevated FIB-4 score. Nevertheless, such algorithm will need careful evaluation and validation in future studies.

One of the limitation in this study is the patients come from a tertiary care setting, having a more severe disease spectrum where the prevalence of advanced fibrosis is 20%. This is because liver biopsy is required as a reference standard and cannot be performed in a general population. This finding may not represent NAFLD patients in the community as a higher proportion will have milder liver disease, which will further increase the NPV of the scoring systems. In addition, liver biopsy is used in the staging of disease in this study which can be subjected to sampling error and interobserver variability. Nevertheless the findings of present study are consistent with

previous reports supporting the utility of these scoring systems for advanced fibrosis in NAFLD.

The advantage of this study is that it is made up of prospectively recruited cohort where all the clinical and biochemical data are evaluated at time of histological evaluation. This is able to give a more robust and reliable data as compared to retrospective studies. In assessing the value of these scoring systems, further longitudinal studies can be considered in defining their utility to test if the scores are correlated with the changes in the liver histological parameters over a long term follow up period.

4.6 Conclusion

The scoring systems validated in this study are able to non-invasively stratify patients without advanced fibrosis, thereby avoiding liver biopsies in substantial number of patients and reserving liver biopsy only for patients who are more in need of appropriate therapeutic strategies. FIB-4 outperforms other non-invasive tests in terms of diagnostic accuracy, NPV, lower percentage of patients with indeterminate results in addition to the ease of calculation. Consistent with other studies performed in the Asian population, this study shows that scoring systems such as BARD and NFS that include BMI as parameter are less applicable in Asian population where advanced fibrosis can occur at a lower BMI.

CHAPTER 5: GLUCOKINASE REGULATORY (*GCKR*) GENE VARIANT AS A PREDICTIVE RISK FACTOR FOR NON-ALCOHOLIC FATTY LIVER DISEASE (NAFLD)

5.1 Introduction

NAFLD has been an emerging public health concern as it is the leading cause of chronic liver disease worldwide. The prevalence of NAFLD increases at an alarming rate from 15% in 2005 to 25% in 2010 and it is expected to be increasing, parallel with the increased consumption of energy-dense food and reduced physical activity (Younossi et al., 2016). Interestingly, the prevalence of NAFD in Asia is estimated to be 27%, placing the third among other continents after Middle East (32%) and South America (30%) whereas Africa has the lowest prevalence (13%) (Younossi et al., 2016).

Being a disease with a wide clinical spectrum, NAFLD can range from the relatively benign NAFL to NASH, advanced fibrosis, and end-stage liver disease such as cirrhosis and hepatocellular carcinoma. Patients with NAFLD have an overall higher risk of morbidity and mortality deriving from liver-related and cardiovascular disease compared to the general population (Musso et al., 2011). The risk of liver-related mortality is more prominent in patients with NASH (5 times the risk) and NASH patients with advanced fibrosis (10 times the risk) (Musso et al., 2011). The prognosis of individual NAFLD patients is highly variable, where some develop progressive fibrosis and some can have fibrosis regression (Wong, Wong, et al., 2010). While overall NAFL has a relatively indolent rate of progression compared to NASH, heterogeneity is also observed, where approximately a quarter of NAFL patients develop bridging fibrosis over a short duration (Singh et al., 2015). The

inter-individual variation in the susceptibility to NAFLD and fibrotic progression despite similar environmental risk factors implies the contribution of genetic factors.

Familial clustering of NAFLD, interracial and ethnic difference in prevalence of NAFLD have prompted the investigation of genetic variants as possible aetiological factors (Schwimmer et al., 2009). Several GWAS have provided insights into a possible molecular pathway contributing to the pathogenesis of NAFLD (Chalasani et al., 2010; Romeo et al., 2008; Speliotes et al., 2011). Since NAFLD is the hepatic manifestation of metabolic syndrome and spiking prevalence of dyslipidaemia is found among NASH patients (Younossi et al., 2016), the SNPs on genes encoding proteins in the lipogenesis pathways may have a putative role in modulating progression of NAFLD. The candidate gene association study that was carried out previously in Malaysia cohort found a strong significant association between the patatin-like phospholipase domain-containing protein 3 (*PNPLA3*) gene rs738409 variant with higher risk of NAFLD and NASH (Zain et al., 2012). This study has now been extended to perform the candidate gene association study on the two SNPs found on the glucokinase regulatory protein (*GCKR*) gene, which are the rs1260326 and rs780094.

Being a phosphorylating enzyme, the glucokinase (GCK) regulates glucose metabolism in the liver and induce hepatic lipogenesis (Peter et al., 2011). The GCKR gene encodes for glucokinase regulatory protein (GCKRP), which is a substance that binds allosterically to GCK and modulates its activity. There is a non-synonymous *GCKR* variant (rs1260326), which is characterised by a C to T substitution. at amino acid position 446, it results in the proline-to-leucine substitution (P446L). Association of this SNP with triglyceride and fasting plasma glucose levels have been reported in other GWAS and candidate gene association studies(Orho-Melander et al., 2008;

Saxena et al., 2007; Vaxillaire et al., 2008). Meanwhile, there is a non-functional *GCKR* variant (rs780094) found to be in strong linkage disequilibrium with rs1260326 (HapMap CEU r^2=0.93, CHB r^2=0.82) (Diabetes Genetics Initiative of Broad Institute, 2007). Likewise, this variant was reported to be associated with serum triglycerides level in the GWAS analyses performed in the Finnish, Swedish (Saxena et al., 2007) and Danish populations (Sparso et al., 2008). Due to the consistent findings on the association of *GCKR* variants with elevated serum triglycerides level (Bi et al., 2010; Speliotes et al., 2011; Yang et al., 2011), and because accumulation of triglycerides in hepatocytes mark the first hit in pathogenesis of NAFLD, it was hypothesised that *GCKR* variants are associated with the susceptibility to NAFLD.

The aim of this study is to explore the association between *GCKR* variants rs780094 and rs1260326 and NAFLD and also to investigate if association is modified by race. Given that at time of publication, the candidate gene association studies of these SNPs were only carried out in patients with CT-measured hepatic steatosis, this study aims to determine the association of these GCKR variants with the histological severity of NAFLD based on histopathological assessment. In addition, the combined effect of *GCKR* and *PNPLA3* gene variants on the susceptibility to NAFLD is also determined.

5.2 Literature review

5.2.1 PNPLA3

The first GWAS performed in 3,383 ethnically diverse NAFLD subjects analyzed a total of 12,138 SNPs and detected the strong association of *PNPLA3* rs738409 with hepatic fat accumulation. The association remained significant even after adjustment for BMI, diabetes status, alcohol intake and ethnicity (Romeo et al., 2008). The association of *PNPLA3* rs738409 C>G SNP with the susceptibility to NAFLD was subsequently widely replicated in studies of various cohorts, showing significant association of this variant with higher degree of hepatic steatosis, elevated liver enzymes and also increased risk of NASH and fibrosis progression (Romeo et al., 2008; Sookoian & Pirola, 2011; Valenti, Al-Serri, et al., 2010; Valenti, Alisi, & Nobili, 2012). Studies assessing liver histology also found the risk G allele being associated with higher degree of necroinflammation and fibrosis (Sookoian & Pirola, 2011; Valenti, Alisi, et al., 2010; Verrijken et al., 2013; R. Xu et al., 2015). The effect of this allele in paediatric patients is stronger than those found in adults (Valenti, Alisi, et al., 2010). The influence of *PNPLA3* rs738409 was also investigated in the Malaysia cohort reported to be significantly associated with NASH severity and presence of fibrosis but no other histological features (Zain et al., 2012).

The *PNPLA3* gene is located at chromosome 22, with the rs738409 encoding an amino acid substitution from isoleucine to methionine at position 148 (I148M). This SNP appears to be result in decreased adiponutrin activity in hydrolysing triglyceride in the adipose tissues, promoting the development of steatosis (Pirazzi et al., 2012). The inactivation of hepatic glycerolipid hydrolysis also reduces lipid efflux from liver to adipocytes. There is some evidence that NAFLD patients who carry the variant allele show a poorer response to pharmacotherapy. predisposing the NAFLD patients to a poorer response to pharmacotherapy treatment (Wang, Cao, Chen, & Pan, 2018).

5.2.2 GCKR

Despite the discovery of the role of *PNPLA3* rs738409 in NAFLD, a major component of disease heritability remained unclear. A largescale GWAS involving the genotyping of 2.4 million SNPs was then carried out in 7176 volunteers drawn from multiple studiesof Europeans which forms the Genetics of Obesity-related Liver Disease (GOLD) consortium (Speliotes et al., 2011). Among 46 SNPs identified to be associated with computed tomography measured hepatic steatosis, the *PNPLA3* rs738409, *NCAN* rs2228603, *GCKR* rs780094, and *LYPLAL1* rs12137855 variant were found to be associated with NAFLD following the validation in 592 biopsy-proven NASH patients (Speliotes et al., 2011). In addition, significant association was also observed between the risk T allele of *GCKR* and higher levels of LDL-cholesterol and triglycerides, lower fasting insulin, lower fasting glucose, and lower HOMA-IR, which could have contributed to the increased risk of NAFLD (Speliotes et al., 2011).

Following the GWAS, multiple cross-sectional candidate gene association studieswere performed to validate the association of the *GCKR* SNPs with susceptibility to NAFLD. Analysis on the data from the third National Health and Nutrition Examination Survey involving 1825 non-Hispanic white, 1442 non-Hispanic black and 1537 Mexican American found significant association of*PNPLA3* rs738409 and *GCKR* rs780094 with ultrasound measured hepatic steatosis along with elevated level of ALT (Hernaez et al., 2013). Interestingly the variants in *PNPLA3* and *GCKR* were associated with non-Hispanic whites but not with non-Hispanic black, showing specific genetic influences to the susceptibility to NAFLD across different ancestries (Hernaez et al., 2013). Meanwhile, a case-control study in the Chinese population attempted to quantify the degree of hepatic steatosis using ultrasound but failed to find a significant association between the *GCKR* rs780094 with the degree of fatty infiltration in the liver.

Despite this, the T allele is significantly associated with susceptibility to NAFLD and also increased fasting triglyceride and elevated C-reactive protein (Yang et al., 2011).

Significant association of *GCKR* rs780094 with biopsy-proven NAFLD independent of age, sex, BMI, and status of diabetes mellitus was observed in the Japanese (Kitamoto et al., 2014). This study did not find any association of the SNP with any histological traits but found that the SNP is associated with decreased plasma glucose and increased triglycerides level (Kitamoto et al., 2014). Another study performed in 366 biopsy-proven NAFLD patients found a significant association of *GCKR* rs780094 T allele with the presence of fibrosis and higher serum triglyceride level (Petta et al., 2014). The association of *GCKR* rs780094 however, was not replicated in a study carried out in morbidly obese (average BMI is 50 kg/m^2) patients despite the significant association of this SNP with elevated triglycerides level (Gorden et al., 2013). Similarly,in another cohort of severely obese patients (BMI $\geq$ 32kg/m^2) performing bariatric surgery in Taiwan, the association of *GCKR* rs780094 could not be established (Tai et al., 2016).

The effects of SNPs also vary across different ethnicity. A wide cohort study measuring hepatic steatosis using computed tomography found the association of *GCKR* rs780094 with NAFLD in African Americans but not in the Hispanic Americans (Palmer et al., 2013).

Besides NAFLD, the *GCKR* rs780094 variant was also found to be associated with various metabolic parameters in the Atherosclerosis Risk in Communities (ARIC) Study (Bi et al., 2010). When race-stratified analyses were performed, the risk T allele was associated with lower fasting glucose and insulin level, lower insulin resistance, lower risk of diabetes, higher triglyceride level, and higher prevalence of metabolic syndrome in white participants. Among the African American participants however, the

T allele was only associated with reduced insulin level and higher level of triglycerides (Bi et al., 2010).

There were only a few studies that investigated the relationship of non-synonymous variant of *GCKR* (rs1260326) in NAFLD. In a cohort of 455 obese children and adolescent, the risk T allele was found to be associated with NAFLD, higher level of triglycerides and very low-density lipoprotein (Santoro et al., 2012). Significant association of both *GCKR* SNPs with NAFLD was also replicated in Uygur adult population (Cai et al., 2019) but both studies did not perform liver biopsy and hence the histological parameters were not investigated. Measurement of liver fat content using proton magnetic resonance spectroscopy among 380 patients with type II diabetes also found that the *GCKR* rs1260326 C>T is associated with higher liver fat content, independent of their age, BMI and serum triglyceride level (Petit et al., 2016). Another study which performed fine mapping revealed that PNPLA rs738409 and *GCKR* rs1260326 are the strongest associating variants in the risk of hepatic steatosis in African Americans (Palmer et al., 2013) .

5.2.3 Combined effects of *PNPLA3* and *GCKR*

Santoro et al. (2012) investigated the combined effect of *PNPLA3* s738409, *GCKR* rs1260326 and *APOC3* rs2854116 SNPs on the NAFLD among children and adolescents. The additive effect of *PNPLA3* and *GCKR* SNPs was observed by which it explains 15%, 32% and 39% of the hepatic fat content variability in Hispanics, Caucasians, and African Americans respectively. Joint effect of *GCKR* rs1260326, *PNPLA3* rs738409 and TM6SF2 rs58542926 was also demonstrated in paediatric NAFLD patients, where it is associated with higher hepatic fat content as measured by MRI (Goffredo et al., 2016). Based on the significant association of

PNPLA3 and *GCKR* with the susceptibility to NAFLD, a recent study attempted to investigate the possibility of NAFLD diagnosis by using the combination of *PNPLA3* rs738409,*GCKR* rs780094, *APOC3* rs2854116, *ABCA1* rs4149267, and *LPL* rs13702in a regression model with a comprehensive index composing of six independent anthropometric parameters and serum biomarkers (Yang et al., 2018). The study, however, suggests the diagnostic value of *APOC3 rs2854116* instead of other SNPs.

5.3 Methodology

5.3.1 Subject recruitment

Consecutive histologically-confirmed NAFLD patients and control subjects without NAFLD were included in this study. The patients recruited for this study is not of the same group as patients described in Chapter 3 and Chapter 4. Validation of ethnicity was performed by volunteers, affirming no mixed m marriages for more than two generations. The written informed consent was provided and signed by the volunteers before their participation. The protocol had been approved by the Medical Ethics Committee of UMMC.

After the finding of increased echogenicity in the liver as compared to the renal cortex on ultrasound, all patients with NAFLD underwent a liver biopsy. Histopathological examination was carried out on the liver biopsy samples and the histological scoring was made based on recommendations by NASH Clinical Research Network (Brunt et al., 2011; Kleiner et al., 2005). All biopsy specimens have an average of 1.5 cm long, and contained at least six portal tracts. Upon histopathological examination, tissues were scored for steatosis grade (0–3), lobular inflammation (0–3), ballooning (0–2) and fibrosis (0-4). The exclusion criteria of this study includes: alcohol consumption of more than daily intake of 10 g (Ruhl & Everhart, 2005), coexisting liver diseases such as autoimmune liver diseases, viral hepatitis B and hepatitis C infection, Wilson's disease, primary biliary cirrhosis, α-1-antitrypsin deficiency and consumption of drug that would precipitate steatosis.

Genetically unrelated healthy individuals with BMI less than 25 kg/m^2 were recruited as control. They also had fasting plasma glucose less than 110 mg/dL with normal liver enzymes and without dyslipidaemia. By performing liver ultrasonography on the heathy controls, they were excluded from the study upon finding of the

followings: (i) diffuse increase in bright echoes in the liver parenchyma with slightly impaired visualisation of the peripheral portal and hepatic vein borders; (ii) slight diffuse increase in bright homogeneous echoes in the liver parenchyma with normal visualization of the diaphragm and portal and hepatic vein borders, and normal hepatorenal echogenicity contrast; (iii) marked increase in bright echoes at a shallow depth with deep attenuation, impaired visualisation of the diaphragm, and marked vascular blurring (Sanyal & American Gastroenterological, 2002).

Blood samples were taken on the date of liver biopsy for patients and on the day of recruitment for healthy controls. The biochemical tests were performed based on the standard operating procedure in UMMC hospital clinical laboratory. The blood biochemistry profiling and the anthropometric data were thenrecorded: age, sex, height, weight, HbA1c, aspartate aminotransferase (AST), serum alanine amonitransferase (ALT), gamma-glutamyl transpeptidase (GGT),total cholesterol,high density lipoprotein (HDL) cholesterol, low density lipoprotein (LDL) cholesterol, and serum triglyceride level. The body mass index (BMI) was calculated as weight/height^2 (kg/m^2). Measurement of clinical parameters such as systolic and diastolic blood pressure, waist circumference and pulse rate were also carried out.

5.3.2 DNA extraction

Blood samples obtained were centrifuged at 1500 rpm for 10 minutes and the buffy coat layer was isolated. Extraction of genomic DNA was performed usingthe QiAamp DNA Mini Kit (Qiagen, Hilden, German). First, 20 µl of QIAGEN Protease was pipetted into the bottom of a 1.5 ml microcentrifuge tube. After that, 200 µl of buffy coat was added followed by 200 µl of Buffer AL to the sample. Then the mixture was mixed by pulse vortexing for 15 seconds. After incubation of the tube at 56°C for

10 minutes, the tube was centrifuged briefly. Then 200 µl of ethanol was added to the sample and mixed again by pulse vortexing for 15 seconds, followed by a brief centrifuge. The mixture was then carefully applied to the QIAamp Mini spin column without wetting the rim. Then the cap was closed and centrifuged at 6000 x g (8000 rpm) for 1 minute. Next, the spin column was placed in a clean 2 ml collection tube while the tube containing the filtrate was discarded. After that, 500 µl Buffer AW1 was added without wetting the rim and centrifuged at 6000 x g (8000 rpm) for 1 minute. After discarding the collection tube, the spin column was placed in a new collection tube and 500 µl Buffer AW2 was added and centrifuged at full speed (20,000 x g; 14,000 rpm) for 3 minutes. The collection tube was discarded and the spin column was placed in a clean 1.5 ml microcentrifuge tube and 200 µl Buffer AE was added. The tube was incubated at room temperature for 1 minute and centrifuged at 6000 x g (8000 rpm) for 1 minute.

5.3.3 Genotyping assay

Genotyping of *GCKR* rs1260326 and rs780094 werecarried out using the Sequenom MassARRAY technology platform integrated with iPLEX GOLD chemistry (Sequenom, San Diego, CA). The specific assays with proximal SNPs filtering were designed using MassARRAY AssayDesign software package (v4.0).

For each amplification reaction, a total of 1µL of genomic DNA (10-20 ng/µL) was used. Desaltation of reaction products were carried out using SpectroClean resin (Sequenom, San Diego, CA) after the single-base extension. The MassARRAY Nanodispenser was used to spot 10 nL of reaction products onto the SpectroCHIP and the MassARRAY Analyzer Compact MALDI-TOF mass spectrometer was used to determine the mass. Data acquisition and analysis was performed using the

MassARRAY® Typer 4.0 software. After cluster analysis, assignations of genotypes werecarried out using the default setting of Gaussian mixture model. Inspection of clusters was also needed in order to result in a clear cluster separation with good signal to noise cut-off. The genotype calls which are uncertain were clarified by reviewing them manually.

For quality control, the assay was performed in five duplicates in addition to a blank assay. Failing of quality control was indicated if any of the following happeded: i) Assay with < 80% call rate within the same SpectroChip; (ii) < 99.5% concordance in duplicate checks; Iii) > 25% call rate in the blank control; iv) > 10% call rate in blank checks.

5.3.4 Statistical Analysis

The Hardy Weinberg equilibrium (HWE) was used in the assessment of genotype distribution using χ^2-test. Agreement with HWE will be indicated with a p value of more than 0.05.Power analysis was carried out using Quanto by assuming a gene-only effect. A total of 116 NAFLD cases and 160 controls provide 80% power at α of 0.05 with the following assumptions: the allele frequency ranges from 0.30 to 0.40, baseline risk for the Malaysian population is 0.17 and the minimum detectable odds ratio was 2.0.

The NAFLD cases (n=144) are further classified, based on biopsy, to NAFL (n = 33), NASH (n = 111), NASH with no significant fibrosis (fibrosis score < 2, n = 37), and NASH with significant fibrosis (fibrosis score ≥ 2, n = 74). Continuous data were tested usingan independent t-test for normally distributed variables and Mann-Whitney U test for non-normally distributed variables.

Genotype was coded as 0, 1, or 2 that corresponds to the number of minor alleles carried by each individual. The association between genotype and spectrum of NAFLD was evaluated using an additive model of inheritance. Regression analysis demonstrated a significant association with ethnicity but not with age and gender. Multiple logistic regressions adjustedfor potential confounding by ethnicity on susceptibility to NAFLD.

Association of histological ordinal variables and genotypes were determined using univariate analysis of Jonckheere-Terpstra test. Multivariate analysis of histological ordinal variables was performed using ordinal regression. Statistical analyses were performed using SPSS 16.0 (Chicago IL) with a two-sided $p < 0.05$ considered to be statistically significant.

The association between the clinical parameters and genotypes were compared using Kruskal-Wallis test for skewed variables and Analysis of Variance (ANOVA) for normally-distributed variables. Data are expressed as mean and standard deviation (SD) unless otherwise stated.

Generalized Multifactor Dimensionality Reduction (*GMDR*) method was used to determine the influence of GCKR and PNPLA3 gene-gene interaction on NAFLD. Possible interactions were assessed using 10-fold cross validation with exhaustive search considering ethnicity as covariate. The parameters outcome of GMDR analysis includes the crossvalidation consistency, testing balanced accuracy, and empirical p-values (Lou et al., 2007). The finding was then confirmed by performing additive logistic regression model that takes into account all risk alleles of the SNPs.

5.4 Results

5.4.1 Study Subjects

Table 5.1 outlines the demographic and clinical data of the subjects which consisted of 144 NAFLD patients and 198 controls. The demographic and clinical data of the NAFLD patients who were stratified into NAFL and NASH are described in Table 5.2.

Table 0.1: Demographic and clinical data of the subjects

Characteristics	***n* (%) or Mean ± SD**		***P* value**
	Control (*n*=198)	**NAFLD (*n*=144)**	
Gender			
Males	85 (43)	77 (53)	0.054
Females	113 (57)	67 (47)	
Ethnicity			
Malays	80 (41)	59 (41)	0.044
Chinese	54 (27)	54 (38)	
Indians	64 (32)	31 (21)	
Age (years)	53.1±11.5	51.1±12.0	
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	<0.0001
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	0.001
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 in mean ± SD for continuous data and percentage for categorical data.
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 0.2Demographic and clinical data of the NAFLD patients

Characteristics	***n* (%) or Mean ± SD**		
	NAFL (*n*=33)	**NASH (*n*=111)**	***P* value**
Gender			0.461
Males	20 (61)	57 (51)	
Females	13 (39)	54 (49)	
Age (years)	50.7 ± 11.8	51.2 ± 12.1	0.82
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 in mean ± SD for continuous data and percentage for categorical data.
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; *NAFL*, non-alcoholic fatty liver; *NAFLD*, non-alcoholic fatty liver disease*; NASH*, non-alcoholic steatohepatitis

5.4.2 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. Overall, the frequency of risk allele Twas 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). The association between the T allele and NAFLD was also significant among the Indians (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 not in the Malays and Chinese (Figure 5.1, Table 5.3).

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 1.02–2.22, p = 0.012, for rs1260326 and rs780094, respectively) and NASH with significant fibrosis (OR 1.50, 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). These associations was also significant in the Indians (OR 3.15, 95% CI 1.46–6.78, p = 0.003 and OR 3.86, 95% CI 1.72–8.66, p = 0.001, for rs1260326 and rs780094, respectively) but not in the Malays and Chinese.

Both *GCKR* SNPs are not significantly associated with NAFL and NASH with no significant fibrosis (Table 5.3). Since *GCKR* is involved in the pathway of glucose metabolism, the analysis was expanded by categorizing the NAFLD patients into diabetic NAFLD (n = 54) and non-diabetic NAFLD (n = 90). However, no association was observed between these SNPs and diabetes among these NAFLD patients.

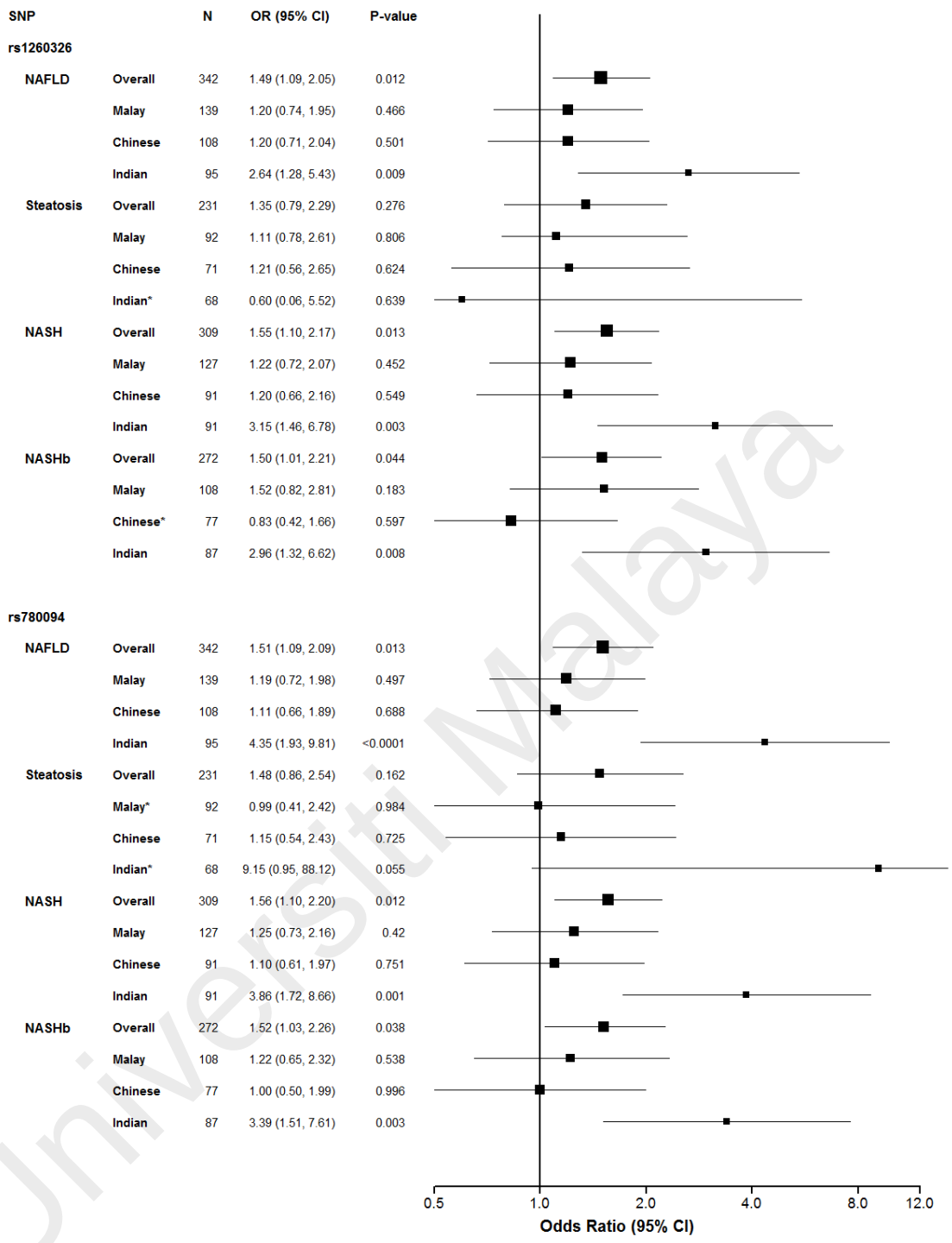


Figure 5.1: The association between *GCKR* polymorphisms with NAFLD spectrum.

*The horizontal line representing the 95 % CI of the point estimate has been truncated due to space limitation.

NASHb Non-alcoholic steatohepatitis with fibrosis

Table 0.3Association tests of *GCKR* polymorphisms in different NAFLD spectrum

NAFLD spectrum	rs1260326			rs780094		
	T allele frequency	*P* value	OR (CI)	T 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	0.688	1.11 (0.66-1.89)
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)
NAFL vs. control						
Overall	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)
Malays	0.46 vs 0.43	0.806	1.11 (0.78-2.61)	0.42 vs 0.42	0.984	0.99 (0.41-2.42)
Chinese	0.53 vs 0.44	0.624	1.21 (0.56-2.65)	0.47 vs 0.44	0.725	1.15 (0.54-2.43)
Indians	0.13 vs 0.19	0.639	0.60 (0.06-5.52)	0.50 vs 0.20	0.055	9.15 (0.95-88.12)
NASH vs. control						
Overall	0.48 vs 0.37	0.013	1.55 (1.10-2.17)	0.46 vs. 0.35	0.012	1.56 (1.10-2.22)
Malays	0.48 vs 0.43	0.452	1.22 (0.72-2.07)	0.47 vs 0.42	0.420	1.25 (0.73-2.16)
Chinese	0.53 vs 0.44	0.549	1.20 (0.66-2.16)	0.46 vs 0.44	0.751	1.10 (0.61-1.97)
Indians	0.41 vs 0.19	0.003	3.15 (1.46-6.78)	0.44 vs 0.20	0.001	3.86 (1.72-8.66)
NASH[a] vs. control	0.51 vs 0.37	0.051	1.68 (1.00–2.82)	0.49 vs. 0.35	0.080	1.60 (0.95-2.71)
NASH[b] vs. control						
Overall	0.46 vs 0.37	0.044	1.50 (1.01-2.21)	0.45 vs.0.35	0.038	1.52 (1.03-2.26)
Malays	0.54 vs 0.43	0.183	1.52 (0.82-2.81)	0.46 vs 0.42	0.538	1.22 (0.65-2.32)
Chinese	0.43 vs 0.44	0.597	0.83 (0.42-1.66)	0.43 vs 0.44	0.996	1.00 (0.50-1.99)
Indians	0.39 vs 0.19	0.008	2.96 (1.32-6.62)	0.43 vs 0.20	0.003	3.39 (1.51-7.61)

Table 5.3, continued

NAFLD spectrum	rs1260326			rs780094		
	T allele frequency	*P* value	OR (CI)	T alelle frequency	*P* value	OR (CI)
NAFL as reference						
All NASH vs. NAFL	0.48 vs 0.45	0.720	1.11 (0.64-1.91)	0.45 vs 0.45	0.940	1.02 (0.57-1.83)
NASH[a] vs. NAFL	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. NAFL	0.46 vs. 0.45	0.810	1.07 (0.61-1.89)	0.45 vs 0.45	0.908	0.97 (0.53-1.76)
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, confidence interval*; NAFL*, non-alcoholic fatty liver*; NAFLD*, non-alcoholic fatty liver disease; *NASH*, non-alcoholic steatohepatitis; *NASH[a]*, non-alcoholic steatohepatitis with no or mild significant fibrosis*; NASH[b]*, non-alcoholic steatohepatitis with significant fibrosis*; OR*, odds ratio.

5.4.3 *GCKR* polymorphisms and liver histology

There was an association between hepatic steatosis grade and T allele (homozygous TT compared to homozygous CC, had a mean score of 2.18 *vs* 1.63, p = 0.008). Following multivariate adjustment, carriers of allele T had exhibited 76% higher odds of developing hepatic steatosis of higher grade compared to non-carriers (OR 1.76; 95% CI 1.08–2.85, p = 0.004). This association remained significant even after adjustment for other histological features (p = 0.04) (Table 5.4). There was no association between T allele and lobular inflammation, hepatocellular ballooning and fibrosis.

Table 0.4: Association of T allele of rs1260326 with histological features in NAFLD patients

Histology	Univariate *P* value[a]	multivariate *P* value[b] (FDR *q* value[d])	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)[c]
Hepatocellular ballooning			
≥ 1 vs. < 1	0. 739	0.281 (0.985)	1.05 (0.47-2.33)[c]
Fibrosis			
≥ 2 vs. < 2	0. 414	0.859 (0.828)	0.81 (0.51-1.29)[c]

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

5.4.4 Combined effect of*GCKR* and *PNPLA3*gene with risk of NAFLD

Variants of both the *GCKR* and *PNPLA3*genes (Zain et al., 2012)have been postulated to confer risk of NAFLD in the current study population. Hence, the interaction between the two risk genes on the occurrence of NAFLD was investigated. Two best models for the interaction were derived. In order to obtain a perfect

cross-validation consistency, the three-locus model (*GCKR* rs1260326, *GCKR* rs780094, *PNPLA3* rs738409)would be the best model (empirical $p = 0.003$). The combined effect of the three SNPs (OR 4.14, 95% CI 1.41-12.18, $p = 0.010$) confers a greater risk for NAFLD than either SNP alone: 1.49 risk for *GCKR*rs1260326, 1.51 risk for *GCKR* rs780094 and 2.23 risk for *PNPLA3* rs738409.The combined effect of *GCKR* and *PNPLA3* showed no significant association with histological parameters ($p > 0.05$).

5.4.5 Statistical power

All significant findings were checked for study power, whereby estimated power of 75%, 82%, 77%, 94%, and 76% were obtained for the association between rs1260326 with NAFLD, rs780094 with NAFLD, rs1260326 with NAFLD in Indian, rs780094 with NAFLD in Indian, and rs1260326 with NASH, respectively.

5.5 Discussion

A positive association is observed between a common intronic SNP (rs780094) and a loss of function SNP (rs1260326) on *GCKR* with NAFLD. This association is modified by race with a strong and significant association seen in the Indians and a weak and non-significant association seen in the Malays and Chinese. The rs1260326 T allele is also associated with increased hepatic steatosis severity, and there is an interaction between the *GCKR* and *PNPLA3* genes in conferring higher risk to NAFLD.

Previous candidate gene association studies have reported the positive association between *GCKR* SNPs with NAFLD. Adding to this evidence base, the present study highlights the association between *GCKR* rs1260326 and rs780094 T allele with NASH and significant fibrosis. This is similar to the findings in the pioneering GWAS where the *GCKR* rs780094 is associated with histology-proven NAFLD from the NASH Clinical Network Research sample (Speliotes et al., 2011). However, the association of the SNPs with other histological parameters were not determined in the GWAS report. Following the publication of this study, the association of the *GCKR* variants and NASH was not replicated in the severely obese patients in Taiwan and the morbidly obese patients from Pennsylvania (Gorden et al., 2013; Tai et al., 2016). Despite not being statistically significant, the author claimed that the direction of effect and effect size of *GCKR* rs780094 was similar to the GOLD consortium study and the non-statistically significant result could be due to the lack of power to identify the SNPs of the moderate effect size (Gorden et al., 2013).

Another interesting finding from this study is that *GCKR* rs1260326 T allele is significantly associated with higher grade of hepatic steatosis. This is consistent with a study among obese children and adolescents in the United States among Caucasians, African Americans, and Hispanics (Santoro et al., 2012). The *GCKR* gene encodes

GCKRP which a regulatory enzyme in the modulation of hepatic GCK activity, thereby regulating de novo lipogenesis by controlling the glucose influx in hepatocytes (Grimsby et al., 2000; Hayward et al., 1998). The common missense loss-of-function variant (rs1260326) encodes for the P446L variant which increases GCK activity by downward regulation of fructose 6-phosphate (Beer et al., 2009). Increased GCK activity promotes hepatic glycolytic influx and subsequently enhances hepatic glucose metabolism. During this process, the concentration of malonyl coenzyme A, which is the substrate for de novo lipogenesis increases (Donnelly et al., 2005). This inhibits carnitine-palmitoyltransferase and thereby blocks fatty acid oxidation, and increases fat deposition in the liver (Beer et al., 2009; Doganay et al., 2014). As a result, de novo lipogenesis is induced and hepatic triglyceride level increases (Peter et al., 2011; Rees et al., 2012).

The finding of gene interaction between the *GCKR* and *PNPLA3* genes on susceptibility to NAFLD may be related to the role of the two genes in the development of hepatic fat content (Romeo, Cohen, & Hobbs, 2006; Santoro et al., 2012). A synergistic additive interaction was seen, whereby the effect size contributed by the combined genes resulted in a greater risk of NAFLD than either gene alone. Recent study carried out resequencing of GWAS-identified candidate genes (*GCKR, PPP1R3B, NCAN, LYPLAL1,* and *TM6SF*2) and found that the *PNPLA3* variant is the strongest genetic determinant of NAFLD, followed by *GCKR*(Di Costanzo et al., 2018). Association of NAFLD with *PNPLA3* rs738409 C>G is fitted better with a dominant model of inheritance whereas the effects of *GCKR* rs1260326 C>T is best explained using a recessive model of inheritance (Di Costanzo et al., 2018). Intriguingly, genes that share similar pathway or outcome as the *PNPLA3* gene have been shown to interact with each other (Zain, Mohamed, Mahadeva, Cheah, et al., 2013; Zain, Mohamed,

Mahadeva, Rampal, et al., 2013). The aetiology of NAFLD is complex and more studies are required to clarify the genetic underpinnings of NAFLD.

The present study population is multiracial and comprised of Malays, Chinese and Indians. The Malaysian Chinese and Indians migrated from Southern China (Saw, 2007) and Southern India (Periasamy, 2007) in the late 18^{th} century and early 19^{th} century, respectively. The association between the *GKCR* variants and NAFLD and NASH is modified by race, being stronger and more significant in Indians compared to Malays and Chinese. Compared to the Malays and Chinese, there is genetic admixture of the Indian subpopulation with West Eurasian resulting in their genetic affinity towards both Asian and European (Bamshad et al., 2001; Jorde & Wooding, 2004). The association of *GCKR* variants with NAFLD are more frequently demonstrated in subjects with European descent compared to patients of other ancestries(Hernaez et al., 2013; Palmer et al., 2013; Speliotes et al., 2011), thus explaining the strong significant association of the *GCKR* variants in the Indians.

When the subjects are stratified according to their race, significant association of *GCKR* variants and NAFLD is not observed within the Chinese patients unlike the findings reported in the China population (Yang et al., 2011). This may be attributed to the difference in genetic variability of Southeast Asian Chinese compared to Chinese from Shanghai and Beijing Han Chinese (CHB) (Chen et al., 2009). The minor allele frequency in Malaysia Chinese control subjects (44%) is different to Chinese from Shanghai (53%) (Yang et al., 2011) and the CHB (57%). There are also methodological differences in case definition where liver biopsy was not carried out in this other study (Yang et al., 2011). Similar to current findings, the *GCKR* variants are not associated with risk of NAFLD in a study carried out in Taiwan involving severely obese Chinese patients (Tai et al., 2016).

The strength of present study is the availability of histological profile of the biopsy-proven NAFLD patients that provides a definite diagnosis of the disease. This allows a complete investigation of the *GCKR* variants and NAFLD spectrum: NAFL, NASH with no significant fibrosis and NASH with significant fibrosis (Chalasani et al., 2018). It is very important to profile NAFLD according to the severity of the condition because unlike NAFL which has a relatively benign prognosis, NASH and hepatic fibrosis signify greater risk of progression to cirrhosis and end-stage liver disease (Adams et al., 2005; Matteoni et al., 1999; Musso et al., 2011). In addition, this study presents ethnic specific comparisons of the association between *GCKR* polymorphisms and NAFLD among three major Asian ethnic groups, Malays, Chinese, Indians. This allows the investigation of how ethnicity modifies the association between *GCKR* variants and NAFLD spectrum.

A limitation of this study is that liver biopsy was not carried out in healthy controls due to ethical concerns. This might result in misclassification of the controls but the controls in this study had met strict selection criteria that included normal BMI, fasting plasma glucose, lipid profile and liver enzymes. These criteria thus minimised the possibility of misclassification in this study.

At time of publication, this was the first report on the genetic interaction between *GCKR* rs1260326 and rs780094 together with *PNPLA3* rs738409 on the increased susceptibility to NAFLD. Additive effect of the *PNPLA3* rs738409 and *GCKR* rs1260326 was previously reported in predisposition of NAFLD in children and adolescent, but investigation was only carried out on one genetic variant of *GCKR*(Santoro et al., 2012). Consecutively similar findings were reported by Goffredo et al. (2016) showing the combined effect of *GCKR* rs1260326 and *PNPLA3* rs738409 in paediatric patients.

5.6 Conclusion

A positive significant association between G*CKR* rs1260326 and rs780094 with NAFLD and NASH is demonstrated. These two SNPs are also associated with higher severity of NASH. The *GCKR* variants are strongly associated in Indians but not Chinese and Malays suggesting different genetic influence in modifying the development of NAFLD across different ancestry group. In addition, the rs1260326 variant is associated with higher grade of hepatic steatosis and there is a combined effect of the *GCKR* variants and *PNPLA3* rs738409 in conferring a higher risk of NAFLD.

CHAPTER 6: COPY NUMBER VARIATION IN EXPORTIN-4 (*XPO4*) GENEAND ITS ASSOCIATION WITH HISTOLOGICAL SEVERITY OF NON-ALCOHOLIC FATTY LIVER DISEASE

6.1 Introduction

Upon the development of the Human Genome Project, there has been mounting evidence showing the contribution of single nucleotide polymorphisms (SNPs) as the major source of human genetic variation modifying disease of complex aetiology (International HapMap et al., 2010). Subsequently, rapid developments in high-throughput next generation sequencing (NGS) and a variety of genotyping platforms have been emerging (Nielsen, Paul, Albrechtsen, & Song, 2011). The advancement of technologies has extended the scale of resolution and depth of coverage on the whole genome, unravelling the rare variants located in the genes predisposing to the development of non-alcoholic fatty liver disease (NAFLD) (Dai et al., 2017; Gerhard et al., 2013). NAFLD encompasses a spectrum of disease ranging from fatty infiltration of the liver which is also known as non-alcoholic fatty liver (NAFL) to the more severe form non-alcoholic steatohepatitis (NASH) which is characterised by the presence of liver injury such as inflammation and hepatocellular ballooning with or without fibrosis (Chalasani et al., 2018). Compared to NAFL, NASH gains higher clinical attention as it has a more rapid fibrotic progression, and with the presence of fibrosis, it leads to higher risk of morbidity and mortality (Dulai et al., 2017; von Roenn, 2018). The non-hypothesis driven GWAS approaches have only discovered a fraction of the genetic component of NAFLD, while the candidate gene association studies in deciphering the relationship between genetic variants to the severity of NAFLD, except for the patatin-like phospholipase domain containing 3 (*PNPLA3*) which is consistently found to be associated with the susceptibility to NAFLD in various cohorts (Bruschi, Tardelli, Claudel, & Trauner, 2017; Salameh et al., 2016; Xu et al., 2015)

Recent studies have reported the role of structural genetic variation such as copy number variations (CNVs) in regulating gene expression and disease phenotype (Ceylan et al., 2019; Gorker et al., 2018; Mace, Kutalik, & Valsesia, 2018; Shahba et al., 2018). CNVs including the gain or loss of DNA constitute a substantial proportion of genetic variability affecting segments ranging from several dozens of bases to megabases (MacDonald, Ziman, Yuen, Feuk, & Scherer, 2014). Construction of CNV map by compiling the published data on healthy individuals across various ethnicities have shown that approximately 4-8-9.5% of the human genome contributes to CNVs, with over 1000 genes mapped within or close to CNV-affected regions (Pinto, Marshall, Feuk, & Scherer, 2007; Zarrei et al., 2015). Upon discovery of widespread common CNVs among healthy individuals, there has been increasing attention on the impacts of CNVs in the common complex diseases, in addition to, Mendelian disorders (Ionita-Laza, Rogers, Lange, Raby, & Lee, 2009). To date, there have only been twoGWAS investigating the relationship of CNV with NAFLD (Royo et al., 2013; Zain et al., 2014). Both studies suggest the role of CNVs in conferring risk of NAFLD development, thereby contributing to a whole new level of complexity in the search of molecular determinants of NAFLD.

Several genomic regions that are potentially relevant to the development of NASH have been identified in the pioneering study (Zain et al., 2014). However, the validation of the findings in a larger sample size is yet to be carried out at time of publication. In the present study, the CNV region (CNVR) 13q12.11 was investigated because (i) CNV gain at locus 13q is one of the highly aberrant regions found in the hepatocellular carcinoma (HCC) samples (Jia et al., 2011); (ii) a tumor suppressor gene named exportin-4 (*XPO4*) is located in this CNV region, and *XPO4* was previously found to be involved in the pathogenesis of HCC (Zender et al., 2008); and (iii)

differential expressions of *XPO4* mRNA were discovered in liver cirrhosis and chronic hepatitis B compared to healthy controls (Zhang et al., 2014).

Despite the relationship of *XPO4* with development of liver diseases including viral hepatitis, cirrhosis and HCC, the association of *XPO4* expression in NAFLD is yet to be determined. Given the potential of the CNV and XPO4 in conferring higher risk of liver diseases, the present study attempts to validate the association between the CNV 13q12.11 and susceptibility to NAFLD. Since the patients included are biopsy-proven, the current study also aims to investigate the association between this CNV and disease severity, histological parameters and the blood biochemistry profiles. The relationship of CNV with the serum XPO4 level was also investigated.

6.2 Literature review

6.2.1 CNVs and disease association

CNVs may contribute to the genetic factor underlying the disease or result in no phenotypic effect. High-resolution mapping of CNVs found that about 100 genes can be homozygously deleted without exerting conspicuous phenotypic consequences (Zarrei et al., 2015). The challenge with CNVs is thus the discrimination of CNVs which affect vital biological function from those that are benign (Nowakowska, 2017). Interestingly, the regions of chromosomal amplification are often found to carry oncogenes whereas the tumor suppressor genes are often harboured in the deleted regions (Chin & Gray, 2008).

6.2.2 CNVs in NAFLD

As opposed to SNP-association studies which are equipped with well-developed resources, incorporation of CNVs in GWAS is still at infancy stage. The first CNVs study carried out in NAFLD involved 17 biopsy-proven individuals, 10 of which are diagnosed with NAFLD upon histological examination (Royo et al., 2013). The genome-wide profiling by array comparative genomic hybridisation (aCGH) found differences in the pattern of CNVs located at 3q29, 6p2, 11q11 and 22q11 in patients with hepatic steatosis compared to those without. The study also found structural variation in the samples derived from NAFLD, including gains on chromosome region 6q16.3 and losses on chromosome regions 7p21.3 and 10q21.3. However, this study was limited by the number of samples and a definite conclusion about the association could not be established (Royo et al., 2013). The second GWAS of CNV was carried out in a larger sample size comprising of 49 NAFLD patients and 49 matched healthy

controls in Malaysia (Zain, 2014). A total of 224 CNVs were found exclusively in NASH patients, 11 CNVs were uniquely found in NAFL patients while 22CNVs were shared among all the NAFLD patients. Following integrative analysis and functional enrichment of CNVs, the study suggested four CNVs having higher propensity in influencing the development of NASH. This includes two rare CNVs (12q24.33 and 13q12.11) with Database of Genomic Variants (DGV) coverage of less than 50%; and two novel CNVs (21p11.1–11.2 and 12q13.2) with DGV coverage of 0% (Zain, 2014).

6.2.3 Chromosome region 13q12.11

One of the earlier discoveries was the identification of homozygous deletion on 13q12.11 in the hepatocellular carcinoma cell line, where the loss of heterozygosity in the 13q12.11 region was significantly associated with early onset of hepatocellular carcinoma (Chen, Yeh, Chen, Chen, & Jou, 2005). A study investigating the frequency of somatic copy number alteration was carried out in cancer patients and discovered that the loss of chromosome region 13q12.11 spanning 60 kilobase is significantly associated with the development of extrahepatic metastasis while there is no loss found in samples that developed gastrointestinal metastasis (Roessler et al., 2015). Besides this, the gain in 13q12.11-q14.2 is also found to be one of the most frequent aberrations correlated with liver metastasis of rectal adenocarcinoma (Zhou et al., 2013).

6.2.4 XPO4 with liver diseases

XPO4 encodes a nuclear export protein by which the substrate EIF5A2 is often found abundantly in human tumors (Clement, Johansson, Wolff, & Park, 2006). It was initially identified as a tumor suppressor gene in the mouse model with liver cancer where the cellular proliferation is inhibited when there is XPO4 re-expression in *XPO4*-deficient tumor cells (Zender et al., 2008). Subsequently, the expressions of XPO4 in HCC cell lines were quantified and XPO4 protein expressions in biopsy samples were analysed using immunohistochemistry. The research found decreased expression of XPO4 in HCC cell lines and likewise, the XPO4 mRNA is also decreased in most of the tumor tissues (Liang et al., 2011). With the correlation of XPO4 expression with tumor size and histopathological classification, downregulation of XPO4 is significantly associated with poor prognosis and could serve as an independent prognostic factor in predicting the survival of HCC patients (Liang et al., 2011).

In a study carried out in patients with HCC, higher survival rates are found in patients with overexpression of XPO4 in carcinoma tissue, suggesting high expression of XPO4 is indicative of a better prognosis and increased survival. On the other hand, overexpression of XPO4 in paracancerous liver tissue tends to have a worse prognosis (Zhang et al., 2013). The author also suggested the role of XPO4 in mediating the progression of HCC and could be a potential biomarker in the evaluation of the prognosis.

6.3 Methodology

6.3.1 Study subjects

A total of 249 consecutive biopsy-proven NAFLD patients from the University Malaya Medical Centre (UMMC) are included in this study. These patients include the patients previously described in Chapter 3, 4 and 5. Presence of NAFLD was confirmed through liver histology upon detection of symptoms or signs that are attributable to liver disease through imaging or upon discovery of abnormal liver biochemistry (Chalasani et al., 2018).Histopathological examination was performed.NAFLD stages were evaluated according to the NASH Clinical Research Network (CRN) criteria (Brunt et al., 1999; Kleiner et al., 2005). All liver biopsy specimens were on average 1.5 cm long and contained at least six portal tracts. There was no evidence of Hepatitis B nor Hepatitis C infection, autoimmune hepatitis, history of alcohol consumption > 10g/day[28], exposure to drugs known to cause steatosis or Wilson's disease reported in any subjects. Based on the NASH CRN, NAFLD patients were grouped into NAFL (n = 32) and NASH (n = 217), the latter was further stratified into NASH without significant fibrosis (fibrosis score < 2, n = 114) and NASH with significant fibrosis (fibrosis score ≥ 2, n = 103). For subsequent analysis, NASH was also grouped into NASH without advanced fibrosis (fibrosis score ≤ 2, n = 167) and NASH with advanced fibrosis (fibrosis score ≥ 3, n = 50) (Chalasani et al., 2018).

All controls (n = 232) were genetically unrelated healthy subjects, confirmed to have normal liver function and had no indication of fatty liver as determined by the following parameters: body mass index (BMI) < 23 kg/m^2, fasting plasma glucose < 110 mg/dL, and normal lipid profile. NAFLD was actively excluded in the controls by ultrasonography according to the absence of the following criteria: (i) slight diffuse

increase in bright homogeneous echoes in the liver parenchyma with normal visualisation of the diaphragm and portal and hepatic vein borders, and normal hepatorenal echogenicity contrast; (ii) diffuse increase in bright echoes in the liver parenchyma with slightly impaired visualisation of the peripheral portal and hepatic vein borders; and (iii) marked increase in bright echoes at a shallow depth with deep attenuation, impaired visualisation of the diaphragm and marked vascular blurring (Sanyal & American Gastroenterological, 2002). All experimental protocols were approved by the responsible Medical Ethics Committee of UMMC (ethics reference number: 702.11) and the methods were carried out in accordance with the approved guidelines. Written informed consent was obtained from all the patients prior to recruitment into the study.

6.3.2 Biochemical and clinical assessments

Anthropometric data such as height and weight for the determination of body mass index (BMI, kg/m^2), and waist circumference, were determined using standard protocols. Measurement of blood pressure (mmHg) was according to a standard recommendation and clinical practice guidelines. The biochemical tests for the determination of hemoglobin A1c (HbA1c), high-density lipoprotein cholesterol (HDL), low-density lipoprotein cholesterol (LDL), total cholesterol, triglycerides, alanine transferase (ALT), aspartate aminotransferase (AST), and gamma-glutamyl transpeptidase (GGT) levels were according to standard clinical laboratory methods carried out in an accredited laboratory at UMMC.

6.3.3 Measurement of serum XPO4 levels

From the serum samples available, random subjects that represented different copy number status from different disease stage (42 controls– 7 losses, 23 neutral and 12 gains; 19 NAFL– 2 losses, 9 neutral and 8 gains; and 34 NASH– 5 losses, 18 neutral and 11 gains) were selected. XPO4 evaluation was performed on an aliquot of serum collected after overnight fasting at the time of sampling and stored at -80^{o}C. Serum XPO4 levels were determined using a commercial enzyme-linked immunosorbent assay (ELISA) kit (SunRed Biotech, Shanghai, China) according to the manufacturer's recommendations, the lowest limit of detection being 0.172 ng/mL. Firstly, serum samples were centrifuged for 20 min at 2000 rpm and supernatant were removed. A total of 40 µL of serum sample was used for each test well, and then 10 µL of XPO4 antibody and 50 µL of Streptavidin-HRP was added. After sealing the membrane, it is shaken gently and incubated at 37°C for 60 mins. After that the membrane was removed and the liquid was drained. A total of 50 µL of chromogen solution A and 50 µL of chromogen solution B was added into each well, followed by gentle mixing. It is then incubated for 10 mins at 37°C away from light. Then 50µL of Stop solution was added into each well to stop the reaction. Measurement was finally taken by setting the blank well as zero and the optical density (OD) at 450 nm wavelengths was measured.

6.3.4 CNVR 13q12.11 genotyping

Genomic DNA was extracted from the blood samples using the QiAamp DNA Mini Kit (Qiagen, Hilden, Germany). The extracted DNA with good quality (OD_{260}/OD_{280} = 1.8-2.0) was diluted to a final concentration of 5 ng/µL. The Applied Biosystems protocols that use a duplex TaqMan real-time quantitative polymerase chain reaction (qPCR) method were employed to call for CNV (13q12.11: Assay

Hs03857719_cn) for every sample. Basically, each reaction (20 µL) contained 10 µL master mix, 1 µL TaqMan Copy Number Assay, 1 µL TaqMan Copy Number Reference Assay, 4 µL nuclease-free water, and 4 µL genomic DNA. All reactions were run in quadruplicate with PCR cycling conditions as follows: 1 PCR cycle at 95 °C for 10 min, followed by 40 cycles at 95 °C for 15 sec and 60 °C for 1 min. Negative controls were introduced for every run to ensure genotyping quality.

6.3.5 CNV validation and meta-analysis

Validation of the previous CNV typing was done on the available samples using qPCR as above. To add strength to the findings, results of the discovery study and replication study were meta-analysed.

6.3.6 Statistical analysis

All statistical tests were performed using SPSS version 16.0 (IBM Corp., Chicago, IL, USA), unless otherwise mentioned. Data were presented as percentage or mean ± standard deviation (S.D). Categorical and continuous variables were compared between NAFLD patients and controls using Pearson's χ^2 test, independent t-test and Mann-Whitney U test as appropriate. Odds ratios and 95% confidence interval (CI) for the findings were computed using logistic regression. Multivariate analysis revealed that gender was a contributing factor for NAFLD, and hence, gender was adjusted in the subsequent analysis. We also included age in the adjustment despite the fact that it was matched between NAFLD patients and controls as age is a known risk factor for NAFLD. Parameter comparisons among CNV status were tested using Analysis of Variance (ANOVA) and Kruskal-Wallis as appropriate. Subsequently, Analysis of

Covariance (ANCOVA) using the general linear model was applied with age, gender and ethnicity as covariates. The *P* values were corrected for multiple testing using the false discovery rate (FDR) method from the Benjamini-Hochberg procedure (Benjamini & Hochberg, 1995a). Linear regression was used to assess the correlation between genetic variants and clinical parameters for normally distributed variables; otherwise Spearman's correlation test was adopted. Variables were log-transformed to achieve normality. Sensitivity analysis was performed according to the approach by Mefford and Witte (Mefford & Witte, 2012). A two-sided *P* value of < 0.05 is considered to be statistically significant.

Meta-analysis was conducted using the Review Manager (RevMan 5.3) of the Cochrane Collaboration utilising a Mantel-Haenszel test to estimate the pooled ORs and corresponding 95% CIs by assuming either fixed or random effect meta-analysis, where appropriate.

Power calculations were performed using Quanto power calculator version 1.2.4, with the following assumptions: the CNV frequency was 0.19, the baseline risk for the Malaysian population was 0.17 (Zain, 2013; Zain, Mohamed, Mahadeva, Cheah, et al., 2013) and the detectable odds ratio ranged from 1.5-2.0.

6.4 Results

6.4.1 Study subjects.

Subject demographic and clinical data are shown in Table 6.1. There were a total of 249 NAFLD patients: 110 (44%) were Malays, 86 (35%) Chinese and 53 (21%) Indians. Out of the 232 controls, 79 (34%) were Malays, 97 (42%) Chinese and 56 (24%) Indians. NAFLD patients and controls significantly differed ($P < 0.05$) in gender, BMI, HbA1c, liver enzymes and lipid profiles, but not in age. The differences in the parameters simply reflect the nature of the subjects according to the criteria set. NAFLD patients were further grouped into NAFL and NASH (Table 6.2). Significantly higher ($P < 0.05$) levels of histological parameters, waist circumference, BMI, HbA1c, and liver enzymes were observed in patients with NASH compared to those with NAFL.

Table 0.1:Demographic and clinical data of the subjects

Characteristics	***n* (%) or Mean ± SD**			
	Control (*n* = 232)	**NAFLD (*n* = 249)**	***P* value***	***P* value****
Gender			0.043	-
Males	96 (41)	126 (51)	-	-
Females	136 (59)	123 (49)	-	-
Ethnicity			0.073	-
Malay	79 (34)	110 (44)	-	-
Chinese	97 (42)	86 (35)	-	-
Indian	56 (24)	53 (21)	-	-
Age (years)	49.6 ± 13.1	50.6 ± 11.6	0.307	0.102
BMI (kg/m2)	21.6 ± 1.7	29.2 ± 4.5	<0.0001	<0.0001
HbA1c (%)	5.3 ± 0.4	6.5 ± 1.6	<0.0001	<0.0001
HDL cholesterol (mg/dl)	49.3 ± 11.9	46.9 ± 11.5	0.30	0.014
LDL cholesterol (mg/dl)	83.2 ± 17.2	116.9 ± 41.0	<0.0001	<0.0001
Total cholesterol (mg/dl)	174.0 ± 34.0	195.0 ± 44.0	<0.0001	<0.0001
Triglycerides (mg/dl)	105.6 ± 36.5	154.8 ± 68.7	<0.0001	<0.0001
AST (IU/L)	20.3 ± 5.1	46.7 ± 27.1	<0.0001	<0.0001
ALT (IU/L)	24.4 ± 6.2	83.0 ± 48.2	<0.0001	<0.0001
GGT (IU/L)	36.5 ± 17.3	106.4 ± 103.0	<0.0001	<0.0001
Serum XPO4 (ng/mL)[a]	24.6 ± 18.5	16.3 ± 14.8	0.038	0.048

[a]Number of subjects (n) was not equal to overall subjects. *P* values obtained using Mann-Whitney U test except for gender and ethnicity used $\chi 2$ test. ***P* values obtained using ANCOVA with age, gender and ethnicity as covariates.
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, XPO4 exportin-4

Table 0.2: Demographic and clinical data of the NAFLD patients

Characteristics	*n* (%) or Mean ± SD		
	NAFL (*n*=32)	NASH (*n*=217)	*P* value
Gender, n (%)			0.461
Males	18 (56)	108 (50)	
Females	14 (44)	109 (50)	
Age (years)*	50.4 ± 10.4	50.7 ± 11.8	0.906
BMI (kg/m2)*	27.4 ± 4.4	29.5 ± 4.4	0.012
HbA1c (%)	6.1 ± 1.5	6.6 ± 1.5	0.010
Waist circumference (cm)*	92.6 ± 11.4	96.5 ± 10.8	0.047
HDL cholesterol (mg/dl)*	47.0 ± 14.6	46.8 ± 11.0	0.942
LDL cholesterol (mg/dl)*	113.4 ± 43.1	117.4 ± 40.7	0.594
Total cholesterol (mg/dl)*	196.3 ± 40.7	194.8 ± 44.6	0.854
Triglycerides (mg/dl)	155.6 ± 86.2	153.5 ± 66.0	0.802
AST (IU/L)	33.4 ± 17.0	48.7 ± 27.8	0.001
ALT (IU/L)	60.7 ± 39.8	86.4 ± 48.6	0.001
GGT (IU/L)	101.1 ± 128.3	107.2 ± 99.0	0.019
Systolic blood pressure (mmHg)	132.7 ± 13.6	137.6 ± 18.7	0.144
Diastolic blood pressure (mmHg)	84.9 ± 11.5	85.7 ± 11.8	0.756
Serum XPO4 (ng/mL)[a]	20.8 ± 19.9	13.8 ± 10.2	0.141

[a]Number of subjects (n) was not equal to overall subjects. **P* values obtained using independent t-test, all other comparisons used Mann-Whitney U 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, XPO4 exportin-4

6.4.2 Association of CNV 13q12.11 with NAFLD

On the basis of the GWAS discovery, we followed up the CNVR on chromosome 13q12.11 in 481 case-control replication samples, in whom we detected eight subjects (1.7%) with homozygous deletions, 55 subjects (11.4%) with heterozygous deletions, 293 subjects (60.9%) with two copies, 99 subjects (20.6%) with three copies, 20 subjects (4.2%) with four copies, and six subjects (1.2%) with five copies. Overall, 134 NAFLD patients (53.8%) were copy number neutral, 33 (13.3%) had deletions and 82 (32.9%) had duplications. As for the controls, 159 (68.6%) were copy number neutral, 30 (12.9%) had copy number losses and 43 (18.5%) had copy number gains.

As shown in Table 6.3, the frequency of CNV gain was significantly higher in the NAFLD patients compared to the controls (adjusted OR 2.32, 95% CI 1.49-3.61, P = 0.0002). However, there was no significant difference between the two groups for CNV loss. CNV gain was also significantly associated with NASH (adjusted OR 2.43, 95% CI 1.54-3.83, P = 0.0001) but not with NAFL. When further stratified into NASH with no significant fibrosis (fibrosis score < 2) and NASH with significant fibrosis (fibrosis score ≥ 2), both groups were found to be significantly associated with CNV gain (adjusted OR 1.87, 95% CI 1.08-3.21, P = 0.029 and adjusted OR 3.21, 95% CI 1.90-5.72, $P = 5.94 \times 10^{-5}$, respectively) with a stronger effect observed in the latter group. We evaluated the association of CNV gain with NASH at the severe stage (fibrosis score ≥ 3; bridging fibrosis and cirrhosis) and found that patients with CNV gain possess 2.55 higher risk for advanced NASH (P = 0.012).

Table 0.3:Association tests of CNV gain with different NAFLD stages

NAFLD spectrum	CNV gain frequency	Unadjusted		Adjusted for age, gender and ethnicity	
		p-value	OR (CI)	p-value	OR (CI)
Control as reference					
NAFLD vs. control	0.33 vs. 0.19	0.0002	2.26 (1.47-3.50)	0.0002	2.32 (1.49-3.61)
NAFL vs. control	0.28 vs. 0.19	0.244	1.66 (0.71-3.92)	0.240	1.69 (0.70-4.08)
All NASH vs. control	0.34 vs. 0.19	0.0002	2.37 (1.51-3.70)	0.0001	2.43 (1.54-3.83)
NASH[a] vs. control	0.31 vs. 0.19	0.013	1.96 (1.15-3.33)	0.029	1.87 (1.08-3.21)
NASH[b] vs. control	0.37 vs. 0.19	0.0001	2.93 (1.70-5.04)	<0.0001	3.21 (1.90-5.72)
NASH[c] vs. control	0.34 vs. 0.19	0.0003	2.41 (1.50-3.88)	0.0016	2.22 (1.35-3.63)
NASH[d] vs. control	0.34 vs. 0.19	0.022	2.25 (1.13-4.48)	0.012	2.55 (1.27-5.21)
NAFL as reference					
All NASH vs. NAFL	0.34 vs. 0.28	0.410	1.42 (0.61-3.30)	0.590	1.27 (0.53-3.04)*
NASH[a] vs. NAFL	0.31 vs. 0.28	0.717	1.18 (0.49-2.86)	0.984	0.98 (0.38-2.51)*
NASH[b] vs. NAFL	0.37 vs. 0.28	0.216	1.76 (0.72-4.30)	0.288	1.68 (0.66-4.20)*
NASH[c] vs. NAFL	0.34 vs. 0.28	0.397	1.45 (0.62-3.41)	0.650	1.25 (0.54-2.96)*
NASH[d] vs. NAFL	0.34 vs. 0.28	0.554	1.35 (0.50-3.64)	0.901	1.06 (0.36-3.21)*
NASH[a] as reference					
NASH[b] vs NASH[a]	0.37 vs 0.31	0.184	1.49 (0.83-2.70)		
NASH[c] as reference					
NASH[d] vs NASH[c]	0.34 vs 0.34	0.842	0.93 (0.47-1.86)		

**P* values additionally adjusted for BMI, waist circumference and HbA1c.
CI confidence 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, *NASH[c]* early non-alcoholic steatohepatitis, *NASH[d]* advanced non-alcoholic steatohepatitis

When the subjects were stratified by ethnicity, the frequency of CNV gain was found to be relatively high in the Malays (case: 43%, control: 27%), followed by the Chinese (case: 39%, control: 19%) and Indians (case: 27%, control: 18%). The CNV gain was associated with risk of NAFLD in the Malays (adjusted OR 2.02, 95% CI 1.02-3.99, P = 0.043) and Chinese (adjusted OR 2.80, 95% CI 1.32-5.95, P = 0.007) but not in the Indians. CNV gain was also associated with risk of NASH in the Chinese (adjusted OR 3.29, 95% CI 1.52-7.13, P = 0.003) (Table 6.4). Our replication study has a power of 98%, 22% and 98% with α of 0.05 to detect associations with duplications at 13q12.11, for association with NAFLD, NAFL and NASH, respectively.

In order to further investigate the effect of the CNV gain, we first validated the 67 (68%) available samples from the 98 discovery samplesusing qPCR (Zain, 2014). The frequency of CNV gain (20.7%) in the control discovery samples was relatively similar to that of the controls from the replication samples (18.5). Results from the discovery samples indicated that CNV gain at 13q12.11 was associated with risk of NAFLD (OR 5.88, 95% CI 1.94-17.82, P = 0.002) and NASH (OR 5.11, 95% CI 1.67-15.67, P = 0.004), but not with NAFL. Then, we performed a meta-analysis including both discovery and replication studies (Figure 6.1), which confirmed the significant association with NAFLD (OR 2.68, 95% CI 1.79-4.02, P < 0.0001) and NASH (OR 2.64, 95% CI 1.75-4.00, P < 0.0001).

Table 0.4: Association tests of CNV gain with different NAFLD stages according to ethnicity

NAFLD spectrum	Malay		Chinese		Indian	
	***P* value**	**OR (CI)**	***P* value**	**OR (CI)**	***P* value**	**OR (CI)**
NAFLD vs. control	0.043	2.02 (1.02-3.99)	0.007	2.80 (1.32-5.95)	0.259	1.75 (0.66-4.65)
NAFL vs. control	0.125	2.77 (0.75-10.22)	0.669	0.62 (0.07-5.45)	0.665	1.55 (0.21-11.20)
NASH vs. control	0.062	1.95 (0.97-3.94)	0.003	3.29 (1.52-7.13)	0.263	1.78 (0.65-4.86)
NASH vs. NAFL	0.473	0.60 (0.15-2.42)*	0.164	5.16 (0.51-52.10)*	0.722	1.41 (0.21-9.57)*

**P* values additionally adjusted for BMI, waist circumference and HbA1c.
CI confidence interval, *OR* odds ratio, *NAFLD* non-alcoholic fatty liver disease, *NASH* non-alcoholic steatohepatitis

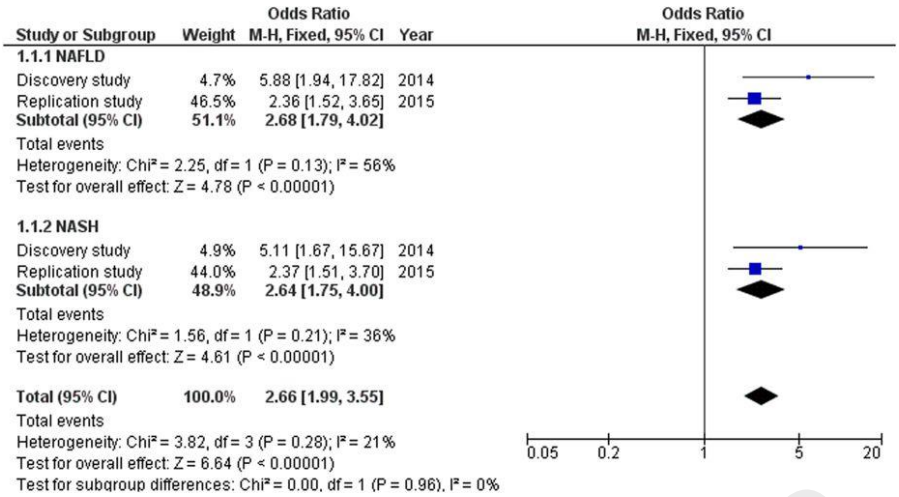


Figure 0.1: Association analysis between CNV gain and NAFLD

Estimations of odds ratios (OR) and 95% confidence intervals (CI) in each study are displayed as closed squares and horizontal lines, respectively. The size of the black squares reflects the weight of the study in the meta-analysis. The diamond represents the combined OR, calculated using a fixed-effect model, with its 95%CI.*NAFLD* non-alcoholic fatty liver disease, *NASH* non-alcoholic steatohepatitis

6.4.3 Comparison of clinical parameters by different copy number status

We compared clinical parameters in NAFLD patients with different copy number status (Table 6.5). Significant differences in the serum were found with ALT (P = 0.038), GGT (P = 0.007) and triglycerides levels (P = 0.046); both serum ALT and triglyceride are indicators for NAFLD severity. Higher copy numbers were associated with higher serum ALT (copy number deletion vs. neutral vs. duplication: 72.1 vs. 82.3 vs. 90.3 IU/L) and triglycerides (136.0 vs. 154.1 vs 160.9 mg/dL). However, after adjustments for the covariates such as age, gender and ethnicity, the differences were no longer significant (Table 6.5). To investigate the magnitude and direction of the effect, using linear regression, we revealed positive correlations between CNV gains and log-transformed serum ALT (P = 3.20 x 10^{-5}) and triglycerides (P = 0.004). Spearman's correlation also showed a similar trend (P = 4.01 x 10^{-5} and P = 0.003, respectively).

6.4.4 Sensitivity analysis

Because triglyceride is a NAFLD-associated risk factor(Chen et al., 2006), and there was a significant correlation between the CNV and serum triglyceride, we wanted to exclude the possibility of collider stratification bias. We therefore, performed a sensitivity analysis by excluding those patients with serum triglyceride levels that were above normal. The associations remained unchanged for NAFLD (OR 2.10, 95% CI 1.27-3.48, P = 0.004), NASH (OR 2.23, 95% CI 1.32-3.75, P = 0.003), and NASH with significant fibrosis (OR 3.02, 95% CI 1.61-5.69, P = 0.001) while NAFL was still not significantly associated. NASH with non-significant fibrosis lost significance. This again suggests that the CNV gain at locus 13q12.11 is associated with more severe degrees of NASH.

6.4.5 Serum XPO4 levels and indices of liver damage

Serum XPO4 levels were shown to be significantly lower in NAFLD patients (16.3 ng/mL) compared to the controls (24.6 ng/mL) (P = 0.038; Table 6.1). There was a reduction in XPO4 levels going from controls (24.6 ng/mL) to NAFL (20.8 ng/mL) and NASH (13.8 ng/mL) (P = 0.043; Table 6.2). As expected, the serum XPO4 levels were not different between controls and NAFL but were significantly different between controls and NASH (P = 0.014). We then assessed the effect of CNV 13q12.11 on serum XPO4 levels: there was a suggestion of a decrease in levels with extra copies of the DNA segment, although this was not statistically significant (CNV losses– 23.0 ng/mL; CNV neutral– 20.4 ng/mL; and CNV gains– 18.0 ng/mL; Table 6.5).

Table 0.5: Comparison of various clinical and histological parameters between the CNV status among NAFLD patients

Characteristics	CNV status, *n*=249 (Mean ± SD)				
	Losses (*n*=33)	Neutral (*n*=134)	Gains (*n*=82)	p-value*	p-value**
Age (years)[a]	50.2 ± 11.3	51.1 ± 11.7	50.2± 12.0	0.835	0.961
BMI (kg/m^2)[a]	28.3 ± 4.4	29.0 ± 4.8	29.8 ± 3.7	0.235	0.249
HbA1c (%)	6.4 ± 1.3	6.5 ± 1.5	6.5 ± 1.7	0.875	0.947
Waist circumference (cm)[a]	93.5 ± 10.2	96.2 ± 12.1	96.3 ± 9.1	0.404	0.393
HDL cholesterol (mg/dl)[a]	49.4 ± 9.2	46.2 ± 11.6	47.3 ± 10.9	0.316	0.345
LDL cholesterol (mg/dl)[a]	113.6 ± 36.9	115.3 ± 38.0	119.2 ± 44.1	0.697	0.692
Total cholesterol (mg/dl)[a]	185.4 ± 49.6	193.4 ± 41.9	200.1 ± 45.7	0.250	0.238
Triglycerides (mg/dl)	136.0 ± 58.2	154.1 ± 69.6	160.9 ± 66.8	0.046	0.190
AST (IU/L)	40.3 ± 25.7	46.2 ± 26.4	50.8 ± 27.7	0.063	0.167
ALT (IU/L)	72.1 ± 43.4	82.3 ± 52.0	90.3 ± 44.6	0.038	0.179
GGT (IU/L)	88.1 ± 110.4	102.3 ± 102.3	125.9 ± 106.2	0.007	0.206
Systolic BP (mmHg)	134.3 ± 18.8	137.8 ± 19.1	136.4 ± 16.7	0.732	0.574
Diastolic BP (mmHg)	85.3 ± 10.2	85.7 ± 10.9	85.5 ± 13.9	0.844	0.946
Steatosis grade	1.7 ± 0.7	1.9 ± 0.7	2.0 ± 0.7	0.105	0.115
Lobular inflammation	1.4 ± 0.5	1.4 ± 0.6	1.4 ± 0.6	0.922	0.932
Ballooning	1.2 ± 0.6	1.1 ± 0.7	1.2 ± 0.6	0.490	0.439
Fibrosis	1.7 ± 1.0	1.3 ± 1.1	1.6 ± 1.1	0.127	0.107
Serum XPO4 (ng/mL)	23.0 ± 16.1	20.4 ± 18.0	18.0 ± 16.3	0.304	0.663

**P* values[a] obtained using ANOVA, while Kruskal-Wallis test used for the other comparisons. ***P* values obtained using ANCOVA with age, gender and ethnicity as covariates.

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

6.5 Discussion

As an extension of the previous findings in the Malaysia NAFLD cohort (Zain et al., 2014), this study is carried out in a larger cohort of NAFLD patients with well-defined histological characteristics. The association between the CNV 13q12.11 in the *XPO4* gene and NAFLD is validated in the present study. Being a member of the importin β family, the *XPO4* gene mediates the nuclear-cytoplasmic transport of protein cargoes including SMAD3, EIF5A1 and EIF5A2 where the loss of XPO4 could be pro-tumorigenic as it promotes the nuclear accumulation of the proteins (Lipowsky et al., 2000; Zender et al., 2010). An oncogenomics-based *in vivo* RNAi screening had demonstrated the role of *XPO4* as a tumor-suppressing gene (Zender et al., 2008). In addition, overexpression of *XPO4* was correlated with better prognosis and survival rate among patients with HCC (Liang et al., 2011; Zhang et al., 2013). Likewise, the lower expressions of *XPO4* are also found in cirrhotic livers and liver with chronic hepatitis B as compared to the healthy controls (Zhang et al., 2014).

The current study highlights the significant association of CNV gains in *XPO4* with severity of NAFLD, and the association remained significant between the CNV gains and NASH with greater severity (NASH with significant fibrosis and advanced fibrosis). Lower serum XPO4 levels were quantified in patients with NAFLD as compared to healthy controls. Therefore, CNV in the *XPO4* gene could be a potential predictor for the histological severity of NAFLD. The result also suggests a significant association of *XPO4* CNV duplication with histological severity of NAFLD especially with that of NASH. However since there is no association between the CNV in NAFL and NASH, this variant cannot be used as a distinctive marker in discriminating NASH from the relatively benign NAFL.

This CNV association study also demonstrated a significant association of CNV gains with increased serum ALT and triglyceride levels. Although there is inadequate evidence to explain the mechanism of the variants in modulating these biochemical parameters, it can be speculated that the alteration is related to the involvement of *XPO4* in the signal transduction and nuclear export of SMAD family member 3 (Smad3) protein. Smad3 is a crucial transforming growth factor beta 1 (TGF-β1) signaling molecule that has a multifaceted regulatory effect on metabolic homeostasis (Tan, Chong, Tan, & Tan, 2012). The higher level of triglycerides could be due to the downstream effects of Smad3 resulting in the aggravation of insulin resistance and adiposity (Tan et al., 2011; Yadav et al., 2011). The effect of Smad3 can also be observed in Smad3-knockout mice, where they exhibit improved insulin sensitivity and β-oxidation which ameliorate glucotoxicity and lipotoxicity in the liver (Tan et al., 2011). The Smad3 was also found to be involved in modulating the activation of hepatic stellate cells, inducing the production of concomitant extracellular matrix (Khimji, Shao, & Rockey, 2008; Schnabl et al., 2001). Consequently, Smad3-overexpressing mice were found to have elevated serum ALT compared to the control mice during the liver injury (Niu et al., 2016). However, the effects of the CNV on these biochemical parameters warrant confirmation and validation in other populations with larger sample size.

Down-regulated of *XPO4* gene were found in HCC tissues compared to normal hepatic tissue (Liang et al., 2011; Zhang et al., 2013). Similarly, lower expression of XPO4 was also demonstrated in HCC compared to liver cirrhosis and chronic hepatitis B (Zhang et al., 2014). One of the interesting findings in the present study is the association between decreasing level of serum XPO4 with increasing disease severity. This trend may be attributed to the role of *XPO4* as a tumor suppressor gene. XPO4 mediates TGF-β1 which results in the recruitment and phosphorylation of Smad3; thereby the binding of Smad3 to DNA modulates transcriptional events (Shi &

Massague, 2003). Notably, as the degree of phosphorylation increases, there will be a decreased expression of proteins involved in signal transduction, such as XPO4 (Ardito, Giuliani, Perrone, Troiano, & Lo Muzio, 2017). Current results also suggest that serum XPO4 levels may have an inverse relationship with the number of copies, although this was not significant. The potential functional effect of CNV in modulating the serum XPO4 levels require further study, particularly since there seems to be a trend to a decrease in levels as the copy number increases. Interestingly, despite most CNVs being gene dosage insensitive (Zhou, Lemos, Dopman, & Hartl, 2011b), approximately 10% of the CNV duplications in the human genome are found to be dosage reversed (Beckmann, Estivill, & Antonarakis, 2007; Stranger et al., 2007). This may be due to reduced transcription and gene silencing as a consequence of the gene duplication (Zhou, Lemos, Dopman, & Hartl, 2011a).

There are several limitations in this study. One of which is the relatively small number of NAFL patients which may have resulted in the negative association in this sub-group. Despite this, the observed association was consistent with the GWAS study (Zain, 2014) suggesting that the association holds true. In addition, the genome-wide scan carried out by Royo et al. (2013) in a small cohort of NAFL patients did not find an association of CNV 13q12.11 with NAFL. The low number of NAFL patients is also due to the fact that UMMC is a tertiary referral center where more severely affected patients, such as those with NASH and/or fibrosis, are more likely to be recruited. Besides that, association comparison analysis was not performed in homozygous deletion, subjects with four copies and subjects with five copies due to limitation in the sample size. In the discovery phase using aCGH, the CNV 13q12.11 was found uniquely only in NASH but not in the healthy controls (Zain, 2014). However, the present study showed doubling in the frequency of CNV gain in NASH (33.6%) compared to healthy controls (18.5%). This could be due to the following reasons: (i)

relatively low sample in the genome-wide discovery study, (ii) the signal ratio between a case and control sample is normalised and converted to a log2 ratio in the discovery study in assigning the copy number call. The detection of CNV could have been observed in the control but the copy number call is not assigned when the ratio does not reach significant level, and (iii) unlike Mendelian disease with high penetrance and clear patterns of inheritance, NAFLD is a polygenic disease with varying degrees of penetrance. Future validation studies in a prospective setting are warranted. The strength of this study is the consistent findings with the discovery GWAS (Zain, 2014). Furthermore, meta-analysis of the data from the discovery and replication studies were conducted in confirming the association of CNVs with NASH (Figure 6.1). At time of publication, this was also the first study to investigate the association of the CNV with biopsy proven NASH along with its histological severity.

6.6 Conclusion

This study demonstrates and confirms the association of CNV gains at locus 13q12.11 with susceptibility to NASH. Lower serum XPO4 levels are observed in patients with NASH compared to those with NAFL and despite not achieving statistical significance, there is a suggestive decline in levels of serum XPO4 with extra copy number. This study could provide insight to future studies in deciphering the molecular pathways of *XPO4* to enhance our understanding of its role in NAFLD progression. This study also needs to be replicated in a larger cohort and across various ethnicities.

CHAPTER 7: ASSOCIATION OF HUMAN LEUKOCYTE ANTIGEN (HLA) WITH SUSCEPTIBILITY TO NAFLD

7.1 Introduction

In attempts to search for genetic variants that are associated with susceptibility to NAFLD, genotyping arrays and genome wide association studies (GWAS) have accelerated our understanding on genomic variations by being able to establish robust disease loci, some of which provide novel insight to the pathogenic mechanisms. While PNPLA3 variants is the most reproducible SNP associated with the risk of NAFLD (Trepo, Romeo, Zucman-Rossi, & Nahon, 2016), GWAS in hepatology have unravelled strong association between SNPs in the human leukocyte antigen (HLA) region with susceptibility to various liver diseases, including primary biliary cirrhosis, primary sclerosing cholangitis, drug induced liver injury and chronic hepatitis B (Jones & Donaldson, 2003; Kamatani et al., 2009; Karlsen et al., 2010; Lonjou et al., 2008)

The human leucocyte antigens (HLAs) are encoded by the major histocompatibility complex (MHC) located in chromosome 6p21. Being among the most dense of the genes and located in the polymorphic part of the human genome, the MHC region contains a diversity of genes encoding molecules involved in the immune and inflammatory response (Gough & Simmonds, 2007). The progression of NAFLD to NASH is mediated by "multiple hit", involving the interaction of genetic and environmental factors (Takaki, Kawai, & Yamamoto, 2013). In NAFLD, fatty acid accumulation results in lipotoxicity and insulin resistance, and together with gut-derived endotoxins, trigger a cascade of inflammatory response by releasing proinflammatory cytokines in the liver (Fang, Chen, Wang, & Liang, 2018). With the flux of excessive free fatty acids or other pathogenic factors from the gut into the liver, Kupffer cells act as

antigen presenting cells, where the factors are phagocytosed and presented through pattern recognition receptors. Following activation of Kupffer cells, cytokines such as TNF-α will be released, leading to NASH progression (Fang et al., 2018). Since the expression of MHC class II was found on Kupffer cells (Kolios, Valatas, & Kouroumalis, 2006), examination of the disease association with HLA variants may shed some light on the pathogenesis of NAFLD.

The HLA-DQA1 and HLA-DQB1 are the two adjacent loci found on chromosome 6p21.3 which encode the DQ α subunit and DQ β subunit respectively, forming a αβ heterodimer known as HLA-DQ, a component of the MHC class II antigens (Williams, 2001). Despite the current knowledge on the MHC class II mediated pathways in immune processing; the involvement of HLA in hepatic inflammation and fibrosis has not been elucidated. Fine mapping revealed the association of HLA-DQB1 alleles with susceptibility to the various liver diseases including persistent hepatitis infection and hepatitis induced hepatocellular carcinoma (Duggal et al., 2013; Jiang et al., 2013; Kumar et al., 2011; Mbarek et al., 2011). As opposed to the established evidence of association found between HLA alleles with autoimmune disease and infection, studies on disease association of HLA alleles with fatty liver is still at a primordial stage. To date there are only three published studies investigating the association of HLA alleles with susceptibility to NAFLD, and all were conducted in Caucasians (Amzolini et al., 2015; Celikbilek, Selcuk, & Yilmaz, 2011; Doganay et al., 2014). However, the association of the HLA alleles with the histological features of NAFLD was not addressed in these previous studies. This study aims to investigate the association between HLA-DQA1 and HLA-DQB1 with susceptibility to NAFLD and to histological severity in a Malaysian population.

7.2 Literature review

7.2.1 Human Leukocyte Antigens (HLA) system

HLA antigens are cell surface proteins found on nucleated cells controlled by genes on the short arm of chromosome six (6p21.3). As a component of the human MHC, HLA is highly polymorphic and comprising ~4 Mbps, or 0.1% of the genome (Mangalam, Taneja, & David, 2013). These HLA genes encode for proteins involved in the immune defence system. Nomenclature of HLA alleles start with the locus followed by an asterisk (*), a two digit number indicating the antigen specificity, and the assigned allele number.

The HLA system is composed of three regions: Class I, Class II and Class III. The class I region encompasses the *HLA-A, HLA-B* and *HLA-C* genes encoding for the heavy chain of class I molecules. In class II region, there is a series of subregions such as DP, DQ and DR, each containing A and B genes encoding α and β chains respectively. The DQ family is composed of *DQA1* and *DQB1* gene, with each code for two non-covalently associated glycosylated polypeptide α and β chains, forming a heterodimer HLA-DQ molecule (Mangalam et al., 2013). The Class III region does not encode HLA molecules but is composed of genes for complement components and tumour necrosis factors (Choo, 2007).

Class II molecules are expressed only on B lymphocytes, activated T lymphocytes and antigen presenting cells (such as monocytes, macrophages and dendritic cells). During immune response, endocytosis of extracellular exogenous proteins occurs, followed by degradation in the acidic endosomal compartment. Then the class II molecule-peptide complex is transported to the cell surface and recognised by T–cell receptor of CD4+ lymphocytes (Holling, Schooten, & van Den Elsen, 2004).

The extracellular portion of heterodimer HLA-DQ is made of two domains (α1 and α2, or β1 and β2) anchoring on the membrane by a short transmembrane region and a cytoplasmic domain. The α1 and β1 domains form an antigen binding groove which is able to accommodate long peptides (Engelhard, 1994). The polymorphisms of HLA genes are mostly found in the antigen-binding groove (Klein & Sato, 2000). The variation in amino acid sequences affects the fine shape of the groove and thus alters the peptide-binding specificity of HLA molecules (Falk, Rotzschke, Stevanovic, Jung, & Rammensee, 1991).

7.2.2 Serologic typing of HLA antigen

Traditionally serological techniques were employed for HLA antigen typing. However this technique is limited by the availability of suitable antisera to recognize HLA antigens and it requires preparation of viable lymphocytes (Mahdi, 2013). High percentage of incorrect and inconclusive assignments was also obtained when serology was performed when compared to molecular typing (Paunic, Gragert, Madbouly, Freeman, & Maiers, 2012; Schaffer & Olerup, 2001).

7.2.3 Polymerase chain reaction (PCR)

Since the last decade, DNA-based typing techniques have gained popularity in clinical applications to replace serological techniques due to their higher accuracy and success rate (Zetterquist et al., 1997). Initial application of molecular techniques was in Class II typing and more recently used to determine Class I alleles (Loiseau, al-Daccak, David, & Colombani, 1991). PCR-based methods include sequence specific oligonucleotides (SSO) and sequence specific primer (SSP). PCR-based techniques

yields better accuracy and precision in confirming or refuting phenotypic homozygosity compared to serological techniques. Furthermore, PCR can be easily performed in small samples as opposed to lymphocytes preparation for serological typing which can be difficult due to poor cell viability or expression (Ayed, Jendoubi, Makhlouf, Sfar, & Abdallah, 2003). By performing PCR, effort required in subsequent sequencing can be greatly reduced.

3.1.1.1 Sequence specific oligonucleotides (SSO)

This technique involves immobilisation of amplified DNA on nylon membrane, followed by hybridisation of oligonucleotide probes to the amplified DNA. The oligonucleotide probes is a panel of synthetic oligonucleotide sequences corresponding to variable regions of the gene that are designed based on the HLA sequence database (Dunckley, 2012). The probes are normally labelled with horseradish peroxidase or biotin which helps in probe detection (Dalva & Beksac, 2014). The probes are then detected with enhanced chemiluminescence system. More recently, HLA-typing using oligonucleotide probes attached to fluorescently coded microsphere in the discrimination of HLA class I and class II alleles (Dunckley, 2012).

3.1.1.2 Sequence specific primer (SSP)

This technique utilise polymorphic DNA sequence as amplification primer. Only the alleles containing sequences complementary to the primer will be amplified and identified using gel electrophoresis (Erlich, 2012). The development of HLA class II allele typing is more established than class I typing. The class I polymorphism is located in the two domains (α1 and α2) and thus require amplification of the two exons. In

addition more polymorphic sequences are found in class I compared to class II, requiring more probes or primers, making the development of molecular strategies more challenging (Choo, 2007).

7.2.4 HLA and disease association

Recent studies show the association of the susceptibility of diseases with certain HLA types, including autoimmune diseases, infectious diseases, cancer and adverse drug responses (Gough & Simmonds, 2007; Holoshitz, 2013). For instance, narcolepsy was found to be associated with HLA-DQB1*0602 (Ollila, Fernandez-Vina, & Mignot, 2015) and susceptibility of celiac disease was significantly associated with HLA-DQB1*02 (Bosca-Watts et al., 2018; Vader et al., 2003). Some HLA alleles are associated with the severity of disease. For example, linear regression found that the frequency of HLA-DQB1*0602 is associated with the severity of the narcolepsy symptoms, age of symptoms onset and sleep latency (Watson, Ton, Koepsell, Gersuk, & Longstreth, 2010). Among patients with rheumatoid arthritis, individuals with homozygous risk allele HLA-DRB1*04 are at higher risk of having nodular disease (Weyand, Hicok, Conn, & Goronzy, 1992). As one of the autoimmune disease, type 1 diabetes is one of the risk factor for NAFLD (Barros, Santos, Pizarro, del Melo, & Gomes, 2017; Regnell & Lernmark, 2011). Approximately 40-50% of familial aggregation of type 1 diabetes was found to be associated with genes in HLA, with the major genetic determinants being polymorphisms in the class II genes HLA-DQ and HLA-DR (Noble & Valdes, 2011; van Belle, Coppieters, & von Herrath, 2011). There is also an increasing recognition on the value of genetic testing of HLA in diagnosis of disease. For instance in the diagnosis of coeliac disease, the European Society of Paediatric Gastroenterology, Hepatology and Nutrition proposed the typing for HLA-DQ2 and HLA-DQ8 in symptomatic children with suspected coeliac disease (Husby et al., 2012).

7.2.5 HLA and liver diseases

The association of HLA with liver diseases is yet to be widely validated. Due to the expensive cost of HLA typing, the association of HLA with susceptibility of liver diseases is often discovered during genome wide association studies (GWAS) where the single nucleotide polymorphisms (SNPs) are located within the HLA locus. Using GWAS, Singer et al. (2010) found significant association of drug induced liver injury with SNPs from MHC class II region. Further fine mapping discovered strong association of liver injury with HLA haplotype HLA-DRB1*150 - HLA-DQB1*0602 - HLA-DRB5*0101 - HLA-DQA1*0102 (Singer et al., 2010). Primary biliary cirrhosis was also found to be strongly associated with the HLA-DQB1 locus (Kumar et al., 2011).By conducting GWAS in a Japanese cohort, Mbarek et al. (2011) discovered significant association between two SNPs (rs2856718 and rs7453920) within the HLA-DQ locus with the susceptibility to chronic hepatitis B. Subsequent HLA typing showed that DQA1*0102-DQB1*0604 and DQA1*0101-DQB1*0501 are the protective haplotypes while DQA1*0102-DQB1*0303 and DQA1*0301-DQB1*0601 are the risk haplotypes (Mbarek et al., 2011). Another GWAS carried out in a Japanese cohort found that rs2596542 on the MICA gene located on chromosome 6p21.33 is not associated with susceptibility to chronic hepatitis C but significantly associated with the progression to hepatocellular carcinoma (Kumar et al., 2011). Study in multiple cohort also found DQB1*03:01 being significantly associated with spontaneous resolution of HCV infection (Duggal et al., 2013). More recently, rs9275319 at HLA-DQ was found to be associated with hepatitis B virus related hepatocellular carcinoma (HCC) in a Chinese population (Jiang et al., 2013).

7.2.6 HLA and NAFLD

To date, there have been only three studies that have been performed to investigate the association between HLA and NAFLD. Serotyping and low resolution HLA typing in HLA Class I and Class II antigens revealed significant increase of HLA-A24, HLA-B15, HLA-B65, HLA-DR15, HLA-DR16, HLA-DQ3 and HLA-DQ5 in the ultrasonography-diagnosed NAFLD patients compared to controls (Amzolini et al., 2015; Celikbilek et al., 2011). Another study which explores HLA-DQB1 alleles at high resolution showed positive association of HLA-DQB1*06:04 with development of NAFLD while HLA-DQB1*03:02 is associated with reduced risk (Doganay et al., 2014). Although there is an incorporation of biopsy-proven NAFLD patients in the latter study, the association is only confined to NAFLD in general but failed to address the association with susceptibility to NASH and histological severity, most probably due to the low sample size.

7.3 Methodology

7.3.1 Subject recruitment

Consecutive histologically-confirmed NAFLD patients and control subjects without NAFLD were included in this study. The patients included in this study overlapped with the subjects described in previous chapters. Validation of ethnicity was performed by volunteers, affirming no mixed m marriages for more than two generations. The written informed consent was provided and signed by the volunteers before their participation. The protocol had been approved by the Medical Ethics Committee of UMMC.

After the finding of increased echogenicity in the liver as compared to the renal cortex on ultrasound, all patients with NAFLD underwent a liver biopsy. Histopathological examination was carried out on the liver biopsy samples and the histological scoring was made based on recommendations by NASH Clinical Research Network (Brunt et al., 2011; Kleiner et al., 2005). All biopsy specimens have an average of 1.5 cm long, and contained at least six portal tracts. Upon histopathological examination, tissues were scored for steatosis grade (0–3), lobular inflammation (0–3), ballooning (0–2) and fibrosis (0-4). The exclusion criteria of this study includes: alcohol consumption of more than daily intake of 10 g (Ruhl & Everhart, 2005), coexisting liver diseases such as autoimmune liver diseases, viral hepatitis B and hepatitis C infection, Wilson's disease, primary biliary cirrhosis, α-1-antitrypsin deficiency and consumption of drug that would precipitate steatosis.

Genetically unrelated healthy individuals with BMI less than 25 kg/m^2 were recruited as control. They also had fasting plasma glucose less than 110 mg/dL with normal liver enzymes and without dyslipidaemia. By performing liver ultrasonography

on the heathy controls, they were excluded from the study upon finding of the followings: (i) diffuse increase in bright echoes in the liver parenchyma with slightly impaired visualisation of the peripheral portal and hepatic vein borders; (ii) slight diffuse increase in bright homogeneous echoes in the liver parenchyma with normal visualisation of the diaphragm and portal and hepatic vein borders, and normal hepatorenal echogenicity contrast; (iii) marked increase in bright echoes at a shallow depth with deep attenuation, impaired visualisation of the diaphragm, and marked vascular blurring (Sanyal & American Gastroenterological, 2002).

Blood samples were taken on the date of liver biopsy for patients and on the day of recruitment for healthy controls. The biochemical tests were performed based on the standard operating procedure in UMMC hospital clinical laboratory. The blood biochemistry profiling and the anthropometric data were then recorded: age, sex, height, weight, HbA1c, aspartate aminotransferase (AST), serum alanine amonitransferase (ALT), gamma-glutamyl transpeptidase (GGT), total cholesterol, high density lipoprotein (HDL) cholesterol, low density lipoprotein (LDL) cholesterol, and serum triglyceride level. The body mass index (BMI) was calculated as weight/height2 (kg/m2). Measurement of clinical parameters such as systolic and diastolic blood pressure, waist circumference and pulse rate were also carried out.

7.3.2 DNA extraction

Blood samples obtained were centrifuged at 1500 rpm for 10 minutes and the buffy coat layer was isolated. Extraction of genomic DNA was performed usingthe QiAamp DNA Mini Kit (Qiagen, Hilden, German). First, 20 µl of QIAGEN Protease was pipetted into the bottom of a 1.5 ml microcentrifuge tube. After that, 200 µl of buffy coat was added followed by 200 µl of Buffer AL to the sample. Then the mixture

was mixed by pulse vortexing for 15 seconds. After incubation of the tube at 56°C for 10 minutes, the tube was centrifuged briefly. Then 200 µl of ethanol was added to the sample and mixed again by pulse vortexing for 15 seconds, followed by a brief centrifuge. The mixture was then carefully applied to the QIAamp Mini spin column without wetting the rim. Then the cap was closed and centrifuged at 6000 x g (8000 rpm) for 1 minute. Next, the spin column was placed in a clean 2 ml collection tube while the tube containing the filtrate was discarded. After that, 500 µl Buffer AW1 was added without wetting the rim and centrifuged at 6000 x g (8000 rpm) for 1 minute. After discarding the collection tube, the spin column was placed in a new collection tube and 500 µl Buffer AW2 was added and centrifuged at full speed (20,000 x g; 14,000 rpm) for 3 minutes. The collection tube was discarded and the spin column was placed in a clean 1.5 ml microcentrifuge tube and 200 µl Buffer AE was added. The tube was incubated at room temperature for 1 minute and centrifuged at 6000 x g (8000 rpm) for 1 minute.

7.3.3 HLA Typing

7.3.3.1 DNA Amplification

The concentration of genomic DNA was adjusted to 20 ng/µl using DNAase free water (Qiagen). For every single sample, 13.8 µl of D-mix is mixed with 4 µl of Amplification primer and 0.2 µl of Taq polymerase. The amplification mixture was prepared as needed based on the number of samples performed and subsequently 18 µl of the amplification mixture was added to 2 µl of DNA for each assay. The amplification was performed using thermal cycler (Biometra) as follows:

Table 0.1: Amplification cycle

Temperature (°C)	Incubation Time	Number of cycles
96	03:00	1
96	00:20	5
60	00:20	
72	00:20	
96	00:10	30
60	00:15	
72	00:20	
72	10:00	1
4	indefinitely	1

The amplified DNA is then analyzed by gel electrophoresis to confirm amplification product prior to hybridisation assay.

7.3.3.2 Agarose gel electrophoresis

Gel electrophoresis was carried out to determine the presence of DNA by running the genomic DNA extraction product under 1% agarose gel electrophoresis. A 1% gel is made of mixing 1 g of agarose powder in 100 ml of TBE buffer. 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, followed by heating in the microwave to dissolve the agarose powder. The mixture was then carefully 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. A volume of 1 µL of the genomic DNA was mixed by pipetting with 1.5µL 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.

7.3.3.3 Denaturation/Neutralisation

A clean 96-well plate was placed in an iced tray holder. Each amplified DNA sample was transferred into the 8-well strips. Then 2.5 µl of denaturation buffer wasadded and mixed thoroughly. The sampleswere then incubated at room temperature for 10 minutes. Next, 5µl Neutralisation Buffer was added and mixed thoroughly.

For every single reaction, 4 µl of Bead Mixture and 34 µl of Hybridisation Buffer was required, and the amount of Hybridisation Mixture was prepared based on the number of samples performed. A total of 38µl of this mixture was added to each well. The PCR tubes are sealed and vortexed thoroughly at low speed. The PCR strips are then placed into pre-warmed thermal cycler at 60°C, and incubated for 15 minutes. The PCR strips were removed and 100 µl Wash Buffer was added quickly to each well. The strips were covered and centrifuged for 5 minutes at 1000-1300xg. After that, the strips were placed in tray holder and wash buffer was removed. Another two more washing steps were then carried out.

7.3.3.4 Labeling

A total of 50µl of 1x SAPE solution was added to each well. The PCR strips were placed and tray and sealed, then vortexed thoroughly at low speed. Then the strips were placed in pre-heated thermal cycler at 60°C and incubated for 5 minutes. Then 100 µl of Wash Buffer was added quickly to each well, followed by centrifuging for 5 minutes at 1000-1300xg. The strips were placed on tray holder and supernatant were removed. Then 70 µl of Wash buffer was added to each well and mixed gently by pipetting, and kept in the dark at 4°C. Upon reading with LABScan™ 100, the mixtures were transferred to reading plate.

7.3.3.5 Data Acquisition and calculation

The reading plate was inserted in the low cytometry platform LABScan™100 (One Lambda, Inc.) for data acquisition. The mean fluorescence intensity (MFI) was generated by the Luminex® Data Collector software. The percent positive value is calculated by the software and compared against the pre-determined cut off values for each test probe. A positive attribute was assigned to probe having percent positive above the cut-off and a negative attribute to those below the cut-off. The MFI of each probe is normalised against the positive control MFI and is expressed as a percentage of the positive control MFI. The assignment of the HLA typing is performed using HLA Fusion 3.0 software (One Lambda, Inc.) based on the reaction pattern compared to patterns associated with published HLA gene sequences.

7.3.4 Statistical analysis

All demographic and clinical parameters were presented as mean ± standard deviation for normally distributed continuous data and median (standard deviation) for non-normally distributed continuous data and as percentages for categorical data. Categorical data were compared using Chi square (χ^2) test. Continuous data were tested using independent *t*-test for normally distributed variables and Mann–Whitney *U* test for skewed variables. NAFLD patients (n=191) were further stratified into three groups based on assessment of the biopsies; NAFL, NASH without significant fibrosis (fibrosis score < 2), and NASH with significant fibrosis (fibrosis score ≥2). Analysis of Variance (ANOVA) and Kruskal-Wallis test were conducted for comparison of clinical parameters between three groups for normally and non-normally distributed variables, respectively.

The frequencies of HLA alleles in NAFLD patients and healthy controls were calculated and summarised. The positivity of HLA-DQA1 and -DQB1 alleles was calculated by direct count where the alleles were represented in the tables as yes/no fashion regardless of its homozygous or heterozygous state to reduce overestimation. Since an individual carries two alleles and the alleles are not mutually exclusive, the sum of percentages in the disease association table may not be one hundred. The phenotype frequencies of HLA-DQA1 and HLA-DQB1 alleles were compared among the study groups using the Chi square (χ^2) test or Fisher's exact test when required. Bonferroni's correction was applied for multiple comparisons and the corrected *P* values (P_c) were calculated by multiplying the *P* value with the number of tests performed. A value of $P_c < 0.05$ was considered statistically significant. Odds ratio (OR) and 95% confidence interval (CI) were estimated whenever applicable. Association of histological ordinal variables and genotypes were determined using univariate analysis via the Jonckheere–Terpstra test. Multivariate analysis of histological ordinal variables was performed using ordinal regression. Statistical analyses were performed using SPSS 16.0 (Chicago IL) with a two-sided $P < 0.05$ considered as being statistically significant.

7.4 Results

The demographic and clinical data of the NAFLD and control groups are presented in Table 7.2. A total of 379 subjects were included in the study, 191 with NAFLD and 188 healthy controls. There is no significant difference in age, gender and ethnicity distribution between the two groups. Upon stratification of NAFLD patients according to histology, 20 were identified as NAFL, 94 were NASH without significant fibrosis (fibrosis score <2) and 77 were NASH with significant fibrosis (fibrosis score ≥2). Table 7.3 shows the demographic and clinical data of the NAFLD patients. There is no significant difference in age, gender and ethnicity differences between the three groups of patients. In terms of the clinical parameters, patients with NASH with significant fibrosis have a significantly higher BMI (p =0.004), HbA1c (p <0.001), AST (p =0.017) and ALT (p = 0.025). On the other hand, patients with NAFL have a significantly higher GGT (p = 0.028).

Table 0.2: Demographic and clinical data of the subjects

Characteristics	***n* (%) or Mean ± SD**		
	Control (*n* = 188)	**NAFLD (*n* = 191)**	***P* value**
Race			0.217
Malay	78 (41.5)	92 (48.2)	
Chinese	73 (38.8)	58 (30.4)	
Indian	37 (19.7)	41 (21.5)	
Gender			0.054
Males	74 (39.4)	95 (49.7)	
Females	114 (60.6)	96 (50.3)	
Age (years)	48.7 ± 13.3	50.8 ± 11.2	0.264
BMI (kg/m^2)	21.8 ± 2.3	29.2 ± 5.1	<0.0001
HbA1c (%)	5.6 ± 0.8	6.5 ± 1.4	<0.0001
HDL cholesterol (mg/dl)	56.8 ± 47.5	46.1 ± 10.3	<0.0001
LDL cholesterol (mg/dl)	88.5 ± 17.8	118.2 ± 41.6	<0.0001
Total cholesterol (mg/dl)	178.3 ± 24.2	196.5 ± 45.2	<0.0001
Triglycerides (mg/dl)	106.5 ± 35.6	159.6 ± 73.3	<0.0001
AST (IU/L)	20.9 ± 5.2	50.0 ± 29.0	<0.0001
ALT (IU/L)	26.6 ± 8.6	85.7 ± 48.6	<0.0001
GGT (IU/L)	37.9 ± 18.2	105.1 ± 92.2	<0.0001

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 0.3: Demographic and clinical data of NAFLD patients

Characteristics	*n* (%) or Mean ± SD			*P* value
	NAFL (*n* = 20)	NASH (F<2) (*n* = 94)	NASH(F≥2) (*n* = 77)	
Race				0.212
Malay	12 (60.0)	46 (48.9)	34 (44.2)	
Chinese	3 (15.0)	33 (35.1)	22 (28.6)	
Indian	5 (25.0)	15 (16.0)	21 (27.3)	
Gender				0.551
Males	12 (60)	44 (47.6)	24 (51.1)	
Females	8 (40)	50 (52.4)	23 (48.9)	
Age (years)	49.4 ± 11.5	49.9 ± 11.7	52.4 ± 10.5	0.399
BMI (kg/m^2)	26.5 ± 3.2	29.2 ± 4.8	29.9 ± 5.6	0.004
HbA1c (%)	5.8 ± 0.9	6.2 ± 1.1	6.9 ± 1.6	<0.001
HDL cholesterol (mg/dl)	45.8 ± 12.6	45.6 ± 9.4	46.9 ± 10.8	0.665
LDL cholesterol (mg/dl)	116.0 ± 44.1	120.5 ± 42.6	115.8 ± 40.0	0.745
Total cholesterol (mg/dl)	206.5 ± 35.8	197.8 ± 45.0	192.4 ± 47.7	0.436
Triglycerides (mg/dl)	186.6 ± 101.5	157.0 ± 73.4	155.9 ± 63.8	0.233
AST (IU/L)	33.9 ± 11.1	49.6 ± 28.9	54.7 ± 30.8	0.017
ALT (IU/L)	60.7 ± 30.4	85.0 ± 48.1	93.1 ± 51.1	0.025
GGT (IU/L)	124.8 ± 161.5	90.5 ± 76.1	117.9 ± 84.5	0.028

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.

The distribution of HLA-DQA1 and HLA–DQB1 allele groups in the subjects is demonstrated in Table 7.4. Five HLA-DQA1 and five HLA-DQB1 allele groups were characterised and the results indicated no significant difference in the allele distribution between the two groups in all ethnicities (Table 7.4).Overall, the allele frequency of DQA1*01 and DQB1*03 is the highest in both the healthy controls and NAFLD patients, having an overall frequency of 70% and 62% respectively. When the subjects were stratified according to three ethnic groups, there is a higher frequency of DQA1*06 in Malay healthy control as compared to Malay NAFLD patients (P = 0.036) but it does not remain statistically different after Bonferroni correction for multiple testing. In Chinese, the allele frequency of DQB1*02 is higher in healthy controls compared to NAFLD patients but the difference is no longer significant after Bonferroni correction.

In a second comparison, the NAFLD patients were stratified into NAFL, NASH without significant fibrosis and NASH with significant fibrosis. Table 7.5 shows that the HLA-DQB1*06 frequency was significantly the lowest in NASH with significant fibrosis (18.2%) followed by NASH without significant fibrosis (42.6%) and NAFL (45%) (P = 0.002). The observed difference remained significant after multiple correction (P_c = 0.02), suggesting a protective effect of this allele against NASH with significant fibrosis. On the other hand, there was no significant difference in the HLA-DQA1 allele group frequency in the comparisons. In Table 7.6 where patients were stratified to NAFL (n = 20), NASH without advanced fibrosis (F<3) (n = 124) and NASH with advanced fibrosis (F≥3) (n = 47), the frequency of DQB1*06 remained significantly lowest in patient with NASH with advanced fibrosis (12.8%), followed by NASH without advanced fibrosis (38.7%) and highest in patients with NAFL (45%) (P = 0.003). It remained statistically significant after correction for multiple testing (P_c = 0.03).HLA-DQB1*06 allele group was significantly associated with reduced risk of NASH with advanced fibrosis compared to NASH without advanced fibrosis (OR 0.272, 95% CI 0.14-0.55, P < 0.001).

On multivariate analysis, HLA-DQB1*06 was significantly associated with lower risk of lobular inflammation (OR 0.84, 95% CI 0.39-0.97, P = 0.016) (Table 7.7). The HLA-DQB1*06 was also found to confer a lower risk of hepatic fibrosis in NAFLD patients (OR 0.30, 95% CI 0.14-0.67, P< 0.001 and Bonferroni corrected P_c = < 0.01). This association remained strong even after adjustment for other histological features (P< 0.001) suggesting the role of HLA-DQB1*06 as a protective marker against the development of hepatic fibrosis.

Table 0.4: Distribution of HLA-DQA1 and HLA-DQB1 alleles in subjects

Alleles	All ethnicities *n* (%)			Malay *n* (%)			Chinese *n* (%)			Indian *n* (%)		
	Control (*n*= 188)	NAFLD (*n*=191)	*P* value (P_c value)	Control (*n*= 78)	NAFLD (*n*=92)	*P* value (P_c value)	Control (*n*= 73)	NAFLD (*n*=58)	*P* value (P_c value)	Control (*n*= 37)	NAFLD (*n*=41)	*P* value (P_c value)
DQA1*01	132 (70.2)	135 (70.7)	0.920	58 (74.4)	68 (73.9)	0.947	47 (64.4)	39 (67.2)	0.732	27 (73.0)	28 (68.3)	0.651
DQA1*02	32 (17.0)	43 (22.5)	0.180	13(16.7)	20 (21.7)	0.405	8 (11.0)	5 (8.6)	0.657	11 (29.7)	18 (43.9)	0.196
DQA1*03	59 (31.4)	45 (23.6)	0.088	18 (23.1)	20 (21.7)	0.835	30 (41.1)	17 (29.3)	0.162	11 (29.7)	8 (19.5)	0.294
DQA1*05	44 (23.4)	37 (19.4)	0.338	10 (12.8)	14 (15.2)	0.655	27 (37.0)	14 (24.1)	0.115	7 (18.9)	9 (22.0)	0.741
DQA1*06	60 (31.9)	51 (26.7)	0.265	35 (44.9)	27 (29.3)	0.036 (0.360)	19 (26.0)	15 (25.9)	0.983	6 (16.2)	9 (22.0)	0.521
DQB1*02	48 (25.5)	46 (24.1)	0.744	19 (24.4)	26 (28.3)	0.566	19 (26.0)	6 (10.3)	0.023 (0.230)	10 (27.0)	14 (34.1)	0.496
DQB1*03	122 (64.9)	112 (58.6)	0.210	47 (60.3)	48 (52.2)	0.290	51 (69.9)	39 (67.2)	0.748	24 (64.9)	25 (61.0)	0.723
DQB1*04	14 (7.4)	11 (5.8)	0.508	6 (7.7)	6 (6.5)	0.767	7 (9.6)	3 (5.2)	0.344	1 (2.7)	2 (4.9)	0.618
DQB1*05	84 (44.7)	99 (51.8)	0.164	42 (53.8)	59 (64.1)	0.174	29 (39.7)	23 (39.7)	0.993	13 (35.1)	17 (41.5)	0.566
DQB1*06	62 (33.0)	63 (33.0)	0.999	20 (25.6)	24 (26.1)	0.947	24 (32.9)	24 (41.4)	0.316	18 (48.6)	15 (36.6)	0.282

NAFLD non-alcoholic fatty liver disease, *Pc* corrected P value

Table 0.5: Distribution of HLA-DQA1 and HLA-DQB1 alleles in NAFLD patients

Alleles	*n* (%)			*P* value (P_c value)
	NAFL (*n* = 20)	NASH (F<2) (*n* = 94)	NASH(F≥2) (*n* = 77)	
DQA1*01	15 (75.0)	69 (73.4)	51 (66.2)	0.535
DQA1*02	5 (25.0)	20 (21.3)	18 (23.4)	0.911
DQA1*03	5 (25.0)	22 (23.4)	18 (23.4)	0.987
DQA1*05	3 (15.0)	20 (21.3)	14 (18.2)	0.766
DQA1*06	6 (30.0)	23 (24.5)	22 (28.6)	0.783
DQB1*02	7 (35.0)	21 (22.3)	18 (23.4)	0.477
DQB1*03	11 (55.0)	54 (57.4)	47 (61)	0.841
DQB1*04	1 (5.0)	5 (5.3)	5 (6.5)	0.936
DQB1*05	9 (45.0)	45 (47.9)	45 (58.4)	0.315
DQB1*06	9 (45.0)	40 (42.6)	14 (18.2)	0.002 (0.02)

NASH non-alcoholic steatohepatitis

Table 0.6: Distribution of HLA-DQA1 and HLA-DQB1 alleles in NAFLD patients

Alleles	*n* (%)			*P* value (P_c value)
	NAFL (*n* = 20)	NASH (F<3) (*n* = 124)	NASH(F≥3) (*n* = 47)	
DQA1*01	15 (75.0)	89 (71.8)	31 (66.0)	0.685
DQA1*02	5 (25.0)	28 (22.6)	10 (21.3)	0.945
DQA1*03	5 (25.0)	28 (22.6)	12 (25.5)	0.909
DQA1*05	3 (15.0)	25 (20.2)	9 (19.1)	0.863
DQA1*06	6 (30.0)	31 (25.0)	14 (29.8)	0.770
DQB1*02	7 (35.0)	28 (22.6)	11 (23.4)	0.480
DQB1*03	11 (55.0)	70 (56.5)	31 (66.0)	0.499
DQB1*04	1 (5.0)	8 (6.5)	2 (4.3)	0.849
DQB1*05	9 (45.0)	61 (49.2)	29 (61.7)	0.279
DQB1*06	9 (45.0)	48 (38.7)	6 (12.8)	0.003 (0.03)

NASH non-alcoholic steatohepatitis

Table 7.7 Association of DQB1*06 allele group with histological features in NAFLD patients

Histology	Univariate *P* value[a]	Multivariate *P* value[b]	OR (95% CI)
Steatosis			
> 33% vs. < 33%	0. 594	0.043	0.78 (0.37-1.63)[c]
Lobular inflammation			
≥ 2 foci vs. < 2 foci	0. 040	0.016	0.84 (0.39-0.97)[c]
Hepatocellular ballooning			
≥ 1 vs. < 1	0. 855	0.996	2.71 (1.00-7.34)[c]
Fibrosis			
≥ 2 vs. < 2	<0. 001	<0.001	0.30 (0.14-0.67)[c]

[a] Jonckheere-Terpstra test
[b] Ordinal regression
[c] Multivariate logistic regression
OR odds ratio, *CI* confidence interval

7.5 Discussions

This study explored the association of HLA-DQA1 and HLA-DQB1 with susceptibility to NAFLD, in which data were previously limited. The HLA typing was only performed on HLA class II DQ region as SNPs located in the HLA-DQ locus were found to be associated with various liver disease, including chronic hepatitis B, spontaneous resolution of HCV infection, hepatitis B induced HCC and primary biliary cirrhosis (Duggal et al., 2013; Hirschfield et al., 2009; Jiang et al., 2013; Mbarek et al., 2011).

Despite no significant difference in the allele frequency of HLA-DQA1 and HLA-DQB1 between NAFLD patients and healthy controls, it is found that HLA-DQB1*06 is associated with a lower risk of fibrosis development when the histological severity among NAFLD patients were compared.

Among patients with NAFLD, the HLA-DQB1*06 allele appeared to be the least presented in NASH with significant fibrosis compared to patients without significant fibrosis. Interestingly, the allele was also found to confer a lower risk of lobular inflammation and hepatic fibrosis in patients with NAFLD. Patients with significant fibrosis (F≥2) are at higher risk of progression to decompensated cirrhosis, portal hypertension and liver cancer, hence having higher rate of mortality compared to patients with NAFL who often remain stable for a number of years (Hagstrom et al., 2017; Sanyal et al., 2006). The ability to predict those who are likely to have significant fibrosis and/or advanced fibrosis could optimise the management of NAFLD as patients with significant fibrosis require more intensive lifestyle modification, while patients with advanced fibrosis require close surveillance for the development of oesophageal varices and enrolment into clinical trials (Dyson et al., 2014).

Since this is a preliminary study, it remains unclear how the HLA alleles may influence the NAFLD spectrum. The propensity of developing hepatocellular injury could be accounted from various immune cells. In a state of liver injury, there is a rapid recruitment of immune cells such as neutrophils, leukocytes, monocytes, inflammatory macrophages, natural killer T (NKT) cells, B cells, and T cells to the liver (Racanelli & Rehermann, 2006). The HLA-DQ cell surface receptor proteins are expressed on the antigen presenting cells such as dendritic cells, Kupffer cells and macrophages (Racanelli & Rehermann, 2006). Macrophages for example, promote apoptosis of hepatic stellate cells (HSCs) (Friedman, 2005) and matrix degradation (Hironaka et al., 2000) in addition to being an important source of the pro-inflammatory cytokines (Zhan & An, 2010). The release of proinflammatory cytokines, such as tumor necrosis factor α (TNFα), interleukin-6 (IL-6), and interleukin-1β (IL-1β) (Bilzer, Roggel, & Gerbes, 2006; Kubes & Mehal, 2012) results in T cell activation, induction of apoptosis and activation of HSCs, where the induced HSCs trans-differentiate into activated proliferating, fibrogenic, and contractile cell, resulting in hepatic fibrosis (Carpino et al., 2004). This then initiates the progression to the more severe complications such as cirrhosis and hepatocellular carcinoma (Ganz & Szabo, 2013; Peverill, Powell, & Skoien, 2014).

Besides antigen presenting cells, aberrant MHC class II expression have also been found in hepatocytes in stimulating the inflammatory CD4 T cells during inflammatory immune response (Herkel et al., 2003). NAFLD, being the hepatic manifestation of metabolic syndrome, is often associated with a high number of obese patients, and it has been reported that adipocytes can directly activate T cells via the MHC class II pathway which is exacerbated in obesity (Deng et al., 2013).

The novel finding of this study is that the HLA-DQB1*06 is significantly associated with lower risk of hepatic fibrosis. To our knowledge, this is the first study to demonstrate the association of HLA alleles with the histological findings of NAFLD. Being the major myofibroblastic cells in the liver, hepatic stellate cells promotes fibrogenesis through the expression of cytokines and chemokines (Peverill et al., 2014). Features of antigen-presenting cells in hepatic stellate cells were demonstrated through the high expression of HLA class II molecules during fibrogenesis (Vinas et al., 2003). In addition, MHC-class-II-dependent high immunoglobulinemia and hepatic deposition of immunoglobulins were discovered in porcine-serum-induced hepatic fibrosis in rats (Baba & Doi, 2004; Baba, Saeki, Onodera, & Doi, 2005) implying the involvement of MHC class II pathway in the development of hepatic fibrosis.

Unlike the previous studies, our study did not find any association of HLA-DQA1 and HLA–DQB1 alleles with susceptibility to NAFLD. In the earlier reports by Celikbilek *et al.* (2011) and Amzolini *et al*. (2015), risk of NAFLD is associated with a higher frequency of HLA-DQ5, a HLA serotype subgroups containing a β-chain encoded by HLA-DQB1*05 allele. The discrepancy in the results between our cohort and the previous two cohorts may be due to population differences in the distribution of HLA alleles. Despite the high frequency of HLA-DQB1*05 in the NAFLD patients in our study (51.8%), the frequency is also high in the healthy controls (44.7%). This frequency is in accordance with another study carried out in a healthy Malaysian cohort (45%) (Azira, Zeehaida, & Nurul Khaiza, 2013), but higher than that in the Turkish healthy control subjects (16.1%) (Celikbilek et al., 2011). The discrepancy may also be due to the recruitment of NAFLD patients in the Turkish and Romanian studies, which are not based on biopsy-proven classification of the liver condition nor with clear indication of histopathological severity. The small number of

patients (n=66 and n=46 respectively) in the studies mentioned may also contribute to the discrepancy (Amzolini et al., 2015; Celikbilek et al., 2011).

One of the limitation of this study is that liver biopsy was not carried out in the healthy controls due to ethical consideration in exposing healthy controls to a relatively invasive procedure. Possible disease misclassification in the controls was minimised by the strict selection criteria imposed including normal BMI, fasting plasma glucose, lipid profile and liver enzymes. Another limitation is the lack of high resolution characterisation using sequence-based typing due to its requirement of high cost and facilities that are not routinely available. Despite low resolution, HLA typing by PCR-SSO is more reliable and reproducible compared to serological methods and it is also more feasible to be used as clinical assay involving large-scale screening, as compared to SSP method (Kulcsarova, Kralovicova, Parnicka, Ferencik, & Buc, 2000). This is also the first report of such a study in the Asian population. It is a preliminary study which proposes the utility of HLA alleles in predicting the risk of NASH with significant fibrosis. The HLA typing data for HLA-DQA1 and HLA-DQB1 in the Malaysian healthy control will also be available in the repository for future reference by studies involving liver disease or metabolic syndrome. The strength of this study is the recruitment of biopsy-proven NAFLD patients with clear indication of histopathological severity that allows complete investigation of the HLA-DQA1 and HLA-DQB1 alleles within a wide spectrum of NAFLD (Chalasani et al., 2012). Risk stratification of NAFLD according to the severity is crucial as NASH with significant fibrosis signify greater risk of progression to cirrhosis and end stage liver disease.

7.6 Conclusion

In conclusion, this study provides the first evidence of the protective effect of HLA-DQB1*06 against the development of significant fibrosis and higher grade of lobular inflammation in non-alcoholic fatty liver disease. Future studies involving larger sample size may consider performing HLA typing on HLA-DQB1*06 to confirm the findings.

CHAPTER 8: CONCLUSION

Given that genetics and epigenetics play a crucial role in the development and progression of diseases, the present study discovered the association of various genetic determinants in conferring higher risk of NAFLD and unravelled the potential of several serum biomolecules to be used as diagnostic tools. The genetic determinants investigated include *GCKR* SNPs, CNV and HLA. The *GCKR* rs780094 and rs1260326 and the CNV gains at locus 13q12.11 were associated with NASH with significant fibrosis. On the other hand, the HLA-DQB1*06 was found to be associated with lower risk of NASH with significant fibrosis. The results suggest the potential relationship of genomic variability underlying the molecular pathways that lead to the individual progression of disease and the development of histologic features.

In the aspects of diagnostic biomarkers, miR-122-5p and miR-193-3p were found to be potential biomarkers in distinguishing the NASH from NAFLD patients. The field of miRNA application as non-invasive diagnostic biomarkers is appealing due to their relative stability and not subjected to inter-observer variability. It awaits standardisation of biological assays protocol to be implemented in the clinical setting. However, in the diagnosis of advanced fibrosis, the miRNAs did not substantially improve the diagnostic performance of the existing techniques. Meanwhile, the non-invasive scoring system including FIB-4 score and NAFLD fibrosis score outperformed miRNAs in the detection of advanced fibrosis in NAFLD patients. In clinical practice, application of these non-invasive biomarkers could help in the stratification of patients with higher risk of disease progression and at the same time reduce the need for liver biopsy.

As a whole, future studies could consider deriving an algorithm or combining a panel of genetic determinants and molecular biomarkers in the diagnosis and prediction of the disease progression. Based on the ability of respective markers in reflecting different pathophysiological components involved in NAFLD, they could serve as invaluable pieces to the puzzle in defining the course and outcome of the disease in individual.

REFERENCES

Adams, L. A., Lymp, J. F., St Sauver, J., Sanderson, S. O., Lindor, K. D., Feldstein, A., & Angulo, P. (2005). The natural history of nonalcoholic fatty liver disease: a population-based cohort study. *Gastroenterology, 129*(1), 113-121.

Ahmed, M. (2015). Non-alcoholic fatty liver disease in 2015. *World J Hepatol, 7*(11), 1450-1459. doi:10.4254/wjh.v7.i11.1450

Akhavan Rezayat, A., Dadgar Moghadam, M., Ghasemi Nour, M., Shirazinia, M., Ghodsi, H., Rouhbakhsh Zahmatkesh, M. R., . . . Akhavan Rezayat, K. (2018). Association between smoking and non-alcoholic fatty liver disease: A systematic review and meta-analysis. *SAGE Open Med, 6*, 2050312117745223. doi:10.1177/2050312117745223

Akuta, N., Kawamura, Y., Suzuki, F., Saitoh, S., Arase, Y., Kunimoto, H., . . . Kumada, H. (2016). Impact of circulating miR-122 for histological features and hepatocellular carcinoma of nonalcoholic fatty liver disease in Japan. *Hepatol Int, 10*(4), 647-656. doi:10.1007/s12072-016-9729-2

Alexander, M., Loomis, A. K., Fairburn-Beech, J., van der Lei, J., Duarte-Salles, T., Prieto-Alhambra, D., . . . Alazawi, W. (2018). Real-world data reveal a diagnostic gap in non-alcoholic fatty liver disease. *BMC Med, 16*(1), 130. doi:10.1186/s12916-018-1103-x

Alkhouri, N., Carter-Kent, C., Lopez, R., Rosenberg, W. M., Pinzani, M., Bedogni, G., . . . Nobili, V. (2011). A combination of the pediatric NAFLD fibrosis index and enhanced liver fibrosis test identifies children with fibrosis. *Clin Gastroenterol Hepatol, 9*(2), 150-155. doi:10.1016/j.cgh.2010.09.015

Alkhouri, N., & McCullough, A. J. (2012). Noninvasive Diagnosis of NASH and Liver Fibrosis Within the Spectrum of NAFLD. *Gastroenterol Hepatol (N Y), 8*(10), 661-668.

Ameres, S. L., Martinez, J., & Schroeder, R. (2007). Molecular basis for target RNA recognition and cleavage by human RISC. *Cell, 130*(1), 101-112. doi:10.1016/j.cell.2007.04.037

Amzolini, A.-M., Fortofoiu, M., Elena, S., Popescu, M., Burada, F., Vere, C. C., & Ciurea, T. (2015). Genetic Factors Involved in the Development and Progression of Nonalcoholic Fatty Liver Disease. *Current Health Sciences Journal, 41*(4).

Andersen, C. L., Jensen, J. L., & Orntoft, T. F. (2004). Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Res, 64*(15), 5245-5250. doi:10.1158/0008-5472.CAN-04-0496

Andreasen, D., Fog, J. U., Biggs, W., Salomon, J., Dahslveen, I. K., Baker, A., & Mouritzen, P. (2010). Improved microRNA quantification in total RNA from clinical samples. *Methods, 50*(4), S6-9. doi:10.1016/j.ymeth.2010.01.006

Angulo, P. (2010). Long-term mortality in nonalcoholic fatty liver disease: is liver histology of any prognostic significance? *Hepatology, 51*(2), 373-375. doi:10.1002/hep.23521

Angulo, P., Bugianesi, E., Bjornsson, E. S., Charatcharoenwitthaya, P., Mills, P. R., Barrera, F., . . . George, J. (2013). Simple noninvasive systems predict long-term outcomes of patients with nonalcoholic fatty liver disease. *Gastroenterology, 145*(4), 782-789 e784. doi:10.1053/j.gastro.2013.06.057

Angulo, P., Hui, J. M., Marchesini, G., Bugianesi, E., George, J., Farrell, G. C., . . . Day, C. P. (2007). The NAFLD fibrosis score: a noninvasive system that identifies liver fibrosis in patients with NAFLD. *Hepatology, 45*(4), 846-854. doi:10.1002/hep.21496

Angulo, P., Kleiner, D. E., Dam-Larsen, S., Adams, L. A., Bjornsson, E. S., Charatcharoenwitthaya, P., . . . Bendtsen, F. (2015). Liver Fibrosis, but No Other Histologic Features, Is Associated With Long-term Outcomes of Patients With Nonalcoholic Fatty Liver Disease. *Gastroenterology, 149*(2), 389-397 e310. doi:10.1053/j.gastro.2015.04.043

Anstee, Q. M., & Day, C. P. (2013). The genetics of NAFLD. *Nat Rev Gastroenterol Hepatol, 10*(11), 645-655. doi:10.1038/nrgastro.2013.182

Anstee, Q. M., Seth, D., & Day, C. P. (2016). Genetic Factors That Affect Risk of Alcoholic and Nonalcoholic Fatty Liver Disease. *Gastroenterology, 150*(8), 1728-1744 e1727. doi:10.1053/j.gastro.2016.01.037

Anstee, Q. M., Targher, G., & Day, C. P. (2013). Progression of NAFLD to diabetes mellitus, cardiovascular disease or cirrhosis. *Nat Rev Gastroenterol Hepatol, 10*(6), 330-344. doi:10.1038/nrgastro.2013.41

Ardito, F., Giuliani, M., Perrone, D., Troiano, G., & Lo Muzio, L. (2017). The crucial role of protein phosphorylation in cell signaling and its use as targeted therapy (Review). *Int J Mol Med, 40*(2), 271-280. doi:10.3892/ijmm.2017.3036

Arroyo, J. D., Chevillet, J. R., Kroh, E. M., Ruf, I. K., Pritchard, C. C., Gibson, D. F., . . . Tewari, M. (2011). Argonaute2 complexes carry a population of circulating microRNAs independent of vesicles in human plasma. *Proc Natl Acad Sci U S A, 108*(12), 5003-5008. doi:10.1073/pnas.1019055108

Ascha, M. S., Hanouneh, I. A., Lopez, R., Tamimi, T. A., Feldstein, A. F., & Zein, N. N. (2010). The incidence and risk factors of hepatocellular carcinoma in patients with nonalcoholic steatohepatitis. *Hepatology, 51*(6), 1972-1978. doi:10.1002/hep.23527

Asselah, T., Marcellin, P., & Bedossa, P. (2014). Improving performance of liver biopsy in fibrosis assessment. *J Hepatol, 61*(2), 193-195. doi:10.1016/j.jhep.2014.03.006

Auguet, T., Aragones, G., Berlanga, A., Guiu-Jurado, E., Marti, A., Martinez, S., . . . Richart, C. (2016). miR33a/miR33b* and miR122 as Possible Contributors to Hepatic Lipid Metabolism in Obese Women with Nonalcoholic Fatty Liver Disease. *Int J Mol Sci, 17*(10). doi:10.3390/ijms17101620

Ayed, K., Jendoubi, S. A., Makhlouf, M., Sfar, I., & Abdallah, T. B. (2003). Comparison of HLA Class I and II Molecular and Serological Typing within Clinical Laboratory. *Saudi J Kidney Dis Transpl, 14*(1), 39-42.

Aykut, U. E., Akyuz, U., Yesil, A., Eren, F., Gerin, F., Ergelen, R., . . . Yilmaz, Y. (2014). A comparison of FibroMeter NAFLD Score, NAFLD fibrosis score, and transient elastography as noninvasive diagnostic tools for hepatic fibrosis in patients with biopsy-proven non-alcoholic fatty liver disease. *Scand J Gastroenterol, 49*(11), 1343-1348. doi:10.3109/00365521.2014.958099

Azira, N. M., Zeehaida, M., & Nurul Khaiza, Y. (2013). HLA DR/DQ type in a Malay population in Kelantan, Malaysia. *Malays J Pathol, 35*(1), 65-69.

Baba, Y., & Doi, K. (2004). MHC class II-related genes expression in porcine-serum-induced rat hepatic fibrosis. *Exp Mol Pathol, 77*(3), 214-221. doi:10.1016/j.yexmp.2004.08.001

Baba, Y., Saeki, K., Onodera, T., & Doi, K. (2005). Serological and immunohistochemical studies on porcine-serum-induced hepatic fibrosis in rats. *Exp Mol Pathol, 79*(3), 229-235. doi:10.1016/j.yexmp.2005.08.007

Bala, S., Petrasek, J., Mundkur, S., Catalano, D., Levin, I., Ward, J., . . . Szabo, G. (2012). Circulating microRNAs in exosomes indicate hepatocyte injury and inflammation in alcoholic, drug-induced, and inflammatory liver diseases. *Hepatology, 56*(5), 1946-1957. doi:10.1002/hep.25873

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. doi:10.1101/gr.173301

Barros, B. S. V., Santos, D. C., Pizarro, M. H., del Melo, L. G. N., & Gomes, M. B. (2017). Type 1 Diabetes and Non-Alcoholic Fatty Liver Disease: When Should We Be Concerned? A Nationwide Study in Brazil. *Nutrients, 9*(8). doi:10.3390/nu9080878

Barros, F., Setubal, S., Martinho, J. M., Ferraz, L., & Gaudencio, A. (2016). Correlation of Non-Alcoholic Fatty Liver Disease and Features of Metabolic Syndrome in Morbidly Obese Patients in the Preoperative Assessment for Bariatric Surgery. *Arq Bras Cir Dig, 29*(4), 260-263. doi:10.1590/0102-6720201600040011

Bartel, D. P. (2004). MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell, 116*(2), 281-297.

Becker, P. P., Rau, M., Schmitt, J., Malsch, C., Hammer, C., Bantel, H., . . . Geier, A. (2015). Performance of Serum microRNAs -122, -192 and -21 as Biomarkers in Patients with Non-Alcoholic Steatohepatitis. *PLoS One, 10*(11), e0142661. doi:10.1371/journal.pone.0142661

Beckmann, J. S., Estivill, X., & Antonarakis, S. E. (2007). Copy number variants and genetic traits: closer to the resolution of phenotypic to genotypic variability. *Nat Rev Genet, 8*, 639-646.

Bedogni, G., Miglioli, L., Masutti, F., Tiribelli, C., Marchesini, G., & Bellentani, S. (2005). Prevalence of and risk factors for nonalcoholic fatty liver disease: the Dionysos nutrition and liver study. *Hepatology, 42*(1), 44-52. doi:10.1002/hep.20734

Bedossa, P., & Patel, K. (2016). Biopsy and Noninvasive Methods to Assess Progression of Nonalcoholic Fatty Liver Disease. *Gastroenterology, 150*(8), 1811-1822 e1814. doi:10.1053/j.gastro.2016.03.008

Bedossa, P., Poitou, C., Veyrie, N., Bouillot, J. L., Basdevant, A., Paradis, V., . . . Clement, K. (2012). Histopathological algorithm and scoring system for evaluation of liver lesions in morbidly obese patients. *Hepatology, 56*(5), 1751-1759. doi:10.1002/hep.25889

Bedossa, P., & Poynard, T. (1996). An algorithm for the grading of activity in chronic hepatitis C. The METAVIR Cooperative Study Group. *Hepatology, 24*(2), 289-293. doi:10.1002/hep.510240201

Beer, N. L., Tribble, N. D., McCulloch, L. J., Roos, C., Johnson, P. R., Orho-Melander, M., & Gloyn, A. L. (2009). The P446L variant in GCKR associated with fasting plasma glucose and triglyceride levels exerts its effect through increased glucokinase activity in liver. *Hum Mol Genet, 18*(21), 4081-4088. doi:10.1093/hmg/ddp357

Bellentani, S., Bedogni, G., Miglioli, L., & Tiribelli, C. (2004). The epidemiology of fatty liver. *Eur J Gastroenterol Hepatol, 16*(11), 1087-1093.

Bellentani, S., Saccoccio, G., Masutti, F., Croce, L. S., Brandi, G., Sasso, F., . . . Tiribelli, C. (2000). Prevalence of and risk factors for hepatic steatosis in Northern Italy. *Ann Intern Med, 132*(2), 112-117.

Benedict, M., & Zhang, X. (2017). Non-alcoholic fatty liver disease: An expanded review. *World J Hepatol, 9*(16), 715-732. doi:10.4254/wjh.v9.i16.715

Benjamini, Y., & Hochberg, Y. (1995a). Controlling the false discovery rate - a practical and powerful approach to multiple testing. *J R Stat Soc Ser B Methodol, 57*, 12.

Benjamini, Y., & Hochberg, Y. (1995b). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Statist. Soc. B, 57*(1).

Benz, F., Roderburg, C., Vargas Cardenas, D., Vucur, M., Gautheron, J., Koch, A., . . . Luedde, T. (2013). U6 is unsuitable for normalization of serum miRNA levels in patients with sepsis or liver fibrosis. *Exp Mol Med, 45*, e42. doi:10.1038/emm.2013.81

Bi, M., Kao, W. H., Boerwinkle, E., Hoogeveen, R. C., Rasmussen-Torvik, L. J., Astor, B. C., . . . Kottgen, A. (2010). Association of rs780094 in GCKR with metabolic traits and incident diabetes and cardiovascular disease: the ARIC Study. *PLoS One, 5*(7), e11690. doi:10.1371/journal.pone.0011690

Bilzer, M., Roggel, F., & Gerbes, A. L. (2006). Role of Kupffer cells in host defense and liver disease. *Liver Int, 26*(10), 1175-1186. doi:10.1111/j.1478-3231.2006.01342.x

Bissels, U., Wild, S., Tomiuk, S., Holste, A., Hafner, M., Tuschl, T., & Bosio, A. (2009). Absolute quantification of microRNAs by using a universal reference. *RNA, 15*(12), 2375-2384. doi:10.1261/rna.1754109

Blondal, T., Jensby Nielsen, S., Baker, A., Andreasen, D., Mouritzen, P., Wrang Teilum, M., & Dahlsveen, I. K. (2013). Assessing sample and miRNA profile quality in serum and plasma or other biofluids. *Methods, 59*(1), S1-6. doi:10.1016/j.ymeth.2012.09.015

Blum, J. S., Wearsch, P. A., & Cresswell, P. (2013). Pathways of antigen processing. *Annu Rev Immunol, 31*, 443-473. doi:10.1146/annurev-immunol-032712-095910

Bosca-Watts, M. M., Minguez, M., Planelles, D., Navarro, S., Rodriguez, A., Santiago, J., . . . Mora, F. (2018). HLA-DQ: Celiac disease vs inflammatory bowel disease. *World J Gastroenterol, 24*(1), 96-103. doi:10.3748/wjg.v24.i1.96

Brea, A., & Puzo, J. (2013). Non-alcoholic fatty liver disease and cardiovascular risk. *Int J Cardiol, 167*(4), 1109-1117. doi:10.1016/j.ijcard.2012.09.085

Brenner, C., Galluzzi, L., Kepp, O., & Kroemer, G. (2013). Decoding cell death signals in liver inflammation. *J Hepatol, 59*(3), 583-594. doi:10.1016/j.jhep.2013.03.033

Bril, F., & Cusi, K. (2017). Management of Nonalcoholic Fatty Liver Disease in Patients With Type 2 Diabetes: A Call to Action. *Diabetes Care, 40*(3), 419-430. doi:10.2337/dc16-1787

Brunt, E. M., Janney, C. G., Di Bisceglie, A. M., Neuschwander-Tetri, B. A., & Bacon, B. R. (1999). Nonalcoholic steatohepatitis: a proposal for grading and staging the histological lesions. *Am J Gastroenterol, 94*(9), 2467-2474. doi:10.1111/j.1572-0241.1999.01377.x

Brunt, E. M., Kleiner, D. E., Wilson, L. A., Belt, P., Neuschwander-Tetri, B. A., & Network, N. C. R. (2011). Nonalcoholic fatty liver disease (NAFLD) activity score and the histopathologic diagnosis in NAFLD: distinct clinicopathologic meanings. *Hepatology, 53*(3), 810-820. doi:10.1002/hep.24127

Bruschi, F. V., Tardelli, M., Claudel, T., & Trauner, M. (2017). PNPLA3 expression and its impact on the liver: current perspectives. *Hepat Med, 9*, 55-66. doi:10.2147/HMER.S125718

Bugianesi, E., Leone, N., Vanni, E., Marchesini, G., Brunello, F., Carucci, P., . . . Rizzetto, M. (2002). Expanding the natural history of nonalcoholic steatohepatitis: from cryptogenic cirrhosis to hepatocellular carcinoma. *Gastroenterology, 123*(1), 134-140.

Bugianesi, E., Moscatiello, S., Ciaravella, M. F., & Marchesini, G. (2010). Insulin resistance in nonalcoholic fatty liver disease. *Curr Pharm Des, 16*(17), 1941-1951.

Busk, T. M., Bendtsen, F., Nielsen, H. J., Jensen, V., Brunner, N., & Moller, S. (2014). TIMP-1 in patients with cirrhosis: relation to liver dysfunction, portal hypertension, and hemodynamic changes. *Scand J Gastroenterol, 49*(9), 1103-1110. doi:10.3109/00365521.2014.934910

Buzzetti, E., Lombardi, R., De Luca, L., & Tsochatzis, E. A. (2015). Noninvasive Assessment of Fibrosis in Patients with Nonalcoholic Fatty Liver Disease. *Int J Endocrinol, 2015*, 343828. doi:10.1155/2015/343828

Buzzetti, E., Pinzani, M., & Tsochatzis, E. A. (2016). The multiple-hit pathogenesis of non-alcoholic fatty liver disease (NAFLD). *Metabolism, 65*(8), 1038-1048. doi:10.1016/j.metabol.2015.12.012

Cai, W., Weng, D. H., Yan, P., Lin, Y. T., Dong, Z. H., Mailamuguli, & Yao, H. (2019). Genetic polymorphisms associated with nonalcoholic fatty liver disease in Uyghur population: a case-control study and meta-analysis. *Lipids Health Dis, 18*(1), 14. doi:10.1186/s12944-018-0877-3

Cales, P., Laine, F., Boursier, J., Deugnier, Y., Moal, V., Oberti, F., . . . Lunel, F. (2009). Comparison of blood tests for liver fibrosis specific or not to NAFLD. *J Hepatol, 50*(1), 165-173. doi:10.1016/j.jhep.2008.07.035

Calzadilla Bertot, L., & Adams, L. A. (2016). The Natural Course of Non-Alcoholic Fatty Liver Disease. *Int J Mol Sci, 17*(5). doi:10.3390/ijms17050774

Carpino, G., Franchitto, A., Morini, S., Corradini, S. G., Merli, M., & Gaudio, E. (2004). Activated hepatic stellate cells in liver cirrhosis. A morphologic and morphometrical study. *Ital J Anat Embryol, 109*(4), 225-238.

Celikbilek, M., Baskol, M., Taheri, S., Deniz, K., Dogan, S., Zararsiz, G., . . . Yucesoy, M. (2014). Circulating microRNAs in patients with non-alcoholic fatty liver disease. *World J Hepatol, 6*(8), 613-620. doi:10.4254/wjh.v6.i8.613

Celikbilek, M., Selcuk, H., & Yilmaz, U. (2011). A new risk factor for the development of non-alcoholic fatty liver disease: HLA complex genes. *Turk J Gastroenterol, 22*(4), 395-399.

Cermelli, S., Ruggieri, A., Marrero, J. A., Ioannou, G. N., & Beretta, L. (2011). Circulating microRNAs in patients with chronic hepatitis C and non-alcoholic fatty liver disease. *PLoS One, 6*(8), e23937. doi:10.1371/journal.pone.0023937

Ceylan, A. C., Sahin, I., Erdem, H. B., Kayhan, G., Simsek-Kiper, P. O., Utine, G. E., . . . Alikasifoglu, M. (2019). An eight-case 1q21 region series: novel aberrations and clinical variability with new features. *J Intellect Disabil Res*. doi:10.1111/jir.12592

Chalasani, N., Guo, X., Loomba, R., Goodarzi, M. O., Haritunians, T., Kwon, S., . . . Nonalcoholic Steatohepatitis Clinical Research, N. (2010). Genome-wide association study identifies variants associated with histologic features of nonalcoholic Fatty liver disease. *Gastroenterology, 139*(5), 1567-1576, 1576 e1561-1566. doi:10.1053/j.gastro.2010.07.057

Chalasani, N., Younossi, Z., Lavine, J. E., Charlton, M., Cusi, K., Rinella, M., . . . Sanyal, A. J. (2018). The diagnosis and management of nonalcoholic fatty liver disease: Practice guidance from the American Association for the Study of Liver Diseases. *Hepatology, 67*(1), 328-357. doi:10.1002/hep.29367

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 Association for the Study of Liver Diseases, American College of Gastroenterology, and the American Gastroenterological Association. *Hepatology, 55*(6), 2005-2023. doi:10.1002/hep.25762

Chan, W. K., Sthaneshwar, P., Nik Mustapha, N. R., & Mahadeva, S. (2014). Limited utility of plasma M30 in discriminating non-alcoholic steatohepatitis from steatosis--a comparison with routine biochemical markers. *PLoS One, 9*(9), e105903. doi:10.1371/journal.pone.0105903

Chan, W. K., Tan, A. T., Vethakkan, S. R., Tah, P. C., Vijayananthan, A., & Goh, K. L. (2013). Non-alcoholic fatty liver disease in diabetics--prevalence and predictive factors in a multiracial hospital clinic population in Malaysia. *J Gastroenterol Hepatol, 28*(8), 1375-1383. doi:10.1111/jgh.12204

Chang, J., Nicolas, E., Marks, D., Sander, C., Lerro, A., Buendia, M. A., . . . Taylor, J. M. (2004). miR-122, a mammalian liver-specific microRNA, is processed from hcr mRNA and may downregulate the high affinity cationic amino acid transporter CAT-1. *RNA Biol, 1*(2), 106-113.

Chang, P. E., Goh, G. B., Ngu, J. H., Tan, H. K., & Tan, C. K. (2016). Clinical applications, limitations and future role of transient elastography in the management of liver disease. *World J Gastrointest Pharmacol Ther, 7*(1), 91-106. doi:10.4292/wjgpt.v7.i1.91

Chang, Y., Jung, H. S., Cho, J., Zhang, Y., Yun, K. E., Lazo, M., . . . Ryu, S. (2016). Metabolically Healthy Obesity and the Development of Nonalcoholic Fatty Liver Disease. *Am J Gastroenterol, 111*(8), 1133-1140. doi:10.1038/ajg.2016.178

Chavez-Tapia, N. C., Tellez-Avila, F. I., Barrientos-Gutierrez, T., Mendez-Sanchez, N., Lizardi-Cervera, J., & Uribe, M. (2010). Bariatric surgery for non-alcoholic steatohepatitis in obese patients. *Cochrane Database Syst Rev*(1), CD007340. doi:10.1002/14651858.CD007340.pub2

Cheah, M. C., McCullough, A. J., & Goh, G. B. (2017). Current Modalities of Fibrosis Assessment in Non-alcoholic Fatty Liver Disease. *J Clin Transl Hepatol, 5*(3), 261-271. doi:10.14218/JCTH.2017.00009

Chen, C. F., Yeh, S. H., Chen, D. S., Chen, P. J., & Jou, Y. S. (2005). Molecular genetic evidence supporting a novel human hepatocellular carcinoma tumor suppressor locus at 13q12.11. *Genes Chromosomes Cancer, 44*(3), 320-328. doi:10.1002/gcc.20247

Chen, C. H., Huang, M. H., Yang, J. C., Nien, C. K., Yang, C. C., Yeh, Y. H., & Yueh, S. K. (2006). Prevalence and risk factors of nonalcoholic fatty liver disease in an

adult population of taiwan: metabolic significance of nonalcoholic fatty liver disease in nonobese adults. *J Clin Gastroenterol, 40*(8), 745-752.

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. doi:10.1016/j.ajhg.2009.10.016

Chen, X., Ba, Y., Ma, L., Cai, X., Yin, Y., Wang, K., . . . Zhang, C. Y. (2008). Characterization of microRNAs in serum: a novel class of biomarkers for diagnosis of cancer and other diseases. *Cell Res, 18*(10), 997-1006. doi:10.1038/cr.2008.282

Chin, L., & Gray, J. W. (2008). Translating insights from the cancer genome into clinical practice. *Nature, 452*(7187), 553-563. doi:10.1038/nature06914

Choo, S. Y. (2007). The HLA system: genetics, immunology, clinical testing, and clinical implications. *Yonsei Med J, 48*(1), 11-23. doi:10.3349/ymj.2007.48.1.11

Chugh, P., & Dittmer, D. P. (2012). Potential pitfalls in microRNA profiling. *Wiley Interdiscip Rev RNA, 3*(5), 601-616. doi:10.1002/wrna.1120

Chwist, A., Hartleb, M., Lekstan, A., Kukla, M., Gutkowski, K., & Kajor, M. (2014). A composite model including visfatin, tissue polypeptide-specific antigen, hyaluronic acid, and hematological variables for the diagnosis of moderate-to-severe fibrosis in nonalcoholic fatty liver disease: a preliminary study. *Pol Arch Med Wewn, 124*(12), 704-712.

Cichoz-Lach, H., Celinski, K., Prozorow-Krol, B., Swatek, J., Slomka, M., & Lach, T. (2012). The BARD score and the NAFLD fibrosis score in the assessment of advanced liver fibrosis in nonalcoholic fatty liver disease. *Med Sci Monit, 18*(12), CR735-740.

Clement, P. M., Johansson, H. E., Wolff, E. C., & Park, M. H. (2006). Differential expression of eIF5A-1 and eIF5A-2 in human cancer cells. *FEBS J, 273*(6), 1102-1114. doi:10.1111/j.1742-4658.2006.05135.x

Coenen-Stass, A. M. L., Magen, I., Brooks, T., Ben-Dov, I. Z., Greensmith, L., Hornstein, E., & Fratta, P. (2018). Evaluation of methodologies for microRNA biomarker detection by next generation sequencing. *RNA Biol, 15*(8), 1133-1145. doi:10.1080/15476286.2018.1514236

Dai, D., Wen, F., Zhou, S., Su, Z., Liu, G., Wang, M., . . . He, F. (2017). Association of MTTP gene variants with pediatric NAFLD: A candidate-gene-based analysis of single nucleotide variations in obese children. *PLoS One, 12*(9), e0185396. doi:10.1371/journal.pone.0185396

Dai, W., Ye, L., Liu, A., Wen, S. W., Deng, J., Wu, X., & Lai, Z. (2017). Prevalence of nonalcoholic fatty liver disease in patients with type 2 diabetes mellitus: A meta-analysis. *Medicine (Baltimore), 96*(39), e8179. doi:10.1097/MD.0000000000008179

Dalva, K., & Beksac, M. (2014). HLA typing with sequence-specific oligonucleotide primed PCR (PCR-SSO) and use of the Luminex technology. *Methods Mol Biol, 1109*, 87-99. doi:10.1007/978-1-4614-9437-9_6

Davoren, P. A., McNeill, R. E., Lowery, A. J., Kerin, M. J., & Miller, N. (2008). Identification of suitable endogenous control genes for microRNA gene expression analysis in human breast cancer. *BMC Mol Biol, 9*, 76. doi:10.1186/1471-2199-9-76

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

De Silva, S., Li, W., Kemos, P., Brindley, J. H., Mecci, J., Samsuddin, S., . . . Alazawi, W. (2018). Non-invasive markers of liver fibrosis in fatty liver disease are unreliable in people of South Asian descent. *Frontline Gastroenterol, 9*(2), 115-121. doi:10.1136/flgastro-2017-100865

Dedeoglu, B. G. (2014). High-throughput approaches for microRNA expression analysis. *Methods Mol Biol, 1107*, 91-103. doi:10.1007/978-1-62703-748-8_6

Demir, M., Lang, S., Nierhoff, D., Drebber, U., Hardt, A., Wedemeyer, I., . . . Steffen, H. M. (2013). Stepwise combination of simple noninvasive fibrosis scoring systems increases diagnostic accuracy in nonalcoholic fatty liver disease. *J Clin Gastroenterol, 47*(8), 719-726. doi:10.1097/MCG.0b013e3182819a89

Deng, T., Lyon, C. J., Minze, L. J., Lin, J., Zou, J., Liu, J. Z., . . . Hsueh, W. A. (2013). Class II major histocompatibility complex plays an essential role in obesity-induced adipose inflammation. *Cell Metab, 17*(3), 411-422. doi:10.1016/j.cmet.2013.02.009

Di Costanzo, A., Belardinilli, F., Bailetti, D., Sponziello, M., D'Erasmo, L., Polimeni, L., . . . Arca, M. (2018). Evaluation of Polygenic Determinants of Non-Alcoholic Fatty Liver Disease (NAFLD) By a Candidate Genes Resequencing Strategy. *Sci Rep, 8*(1), 3702. doi:10.1038/s41598-018-21939-0

Diabetes Genetics Initiative of Broad Institute of, H., Mit, L. U., Novartis Institutes of BioMedical, R., Saxena, R., Voight, B. F., Lyssenko, V., . . . Purcell, S. (2007). Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science, 316*(5829), 1331-1336. doi:10.1126/science.1142358

Diehl, A. M. (2005). Hepatic complications of obesity. *Gastroenterol Clin North Am, 34*(1), 45-61. doi:10.1016/j.gtc.2004.12.012

Diehl, P., Fricke, A., Sander, L., Stamm, J., Bassler, N., Htun, N., . . . Peter, K. (2012). Microparticles: major transport vehicles for distinct microRNAs in circulation. *Cardiovasc Res, 93*(4), 633-644. doi:10.1093/cvr/cvs007

Doganay, L., Katrinli, S., Colak, Y., Senates, E., Zemheri, E., Ozturk, O., . . . Doganay, G. D. (2014). HLA DQB1 alleles are related with nonalcoholic fatty liver disease. *Mol Biol Rep, 41*(12), 7937-7943. doi:10.1007/s11033-014-3688-2

Donnelly, K. L., Smith, C. I., Schwarzenberg, S. J., Jessurun, J., Boldt, M. D., & Parks, E. J. (2005). Sources of fatty acids stored in liver and secreted via lipoproteins in patients with nonalcoholic fatty liver disease. *J Clin Invest, 115*(5), 1343-1351. doi:10.1172/JCI23621

Duggal, P., Thio, C. L., Wojcik, G. L., Goedert, J. J., Mangia, A., Latanich, R., . . . Thomas, D. L. (2013). Genome-wide association study of spontaneous resolution of hepatitis C virus infection: data from multiple cohorts. *Ann Intern Med, 158*(4), 235-245. doi:10.7326/0003-4819-158-4-201302190-00003

Dulai, P. S., Singh, S., Patel, J., Soni, M., Prokop, L. J., Younossi, Z., . . . Loomba, R. (2017). Increased risk of mortality by fibrosis stage in nonalcoholic fatty liver disease: Systematic review and meta-analysis. *Hepatology, 65*(5), 1557-1565. doi:10.1002/hep.29085

Dulai, P. S., Sirlin, C. B., & Loomba, R. (2016). MRI and MRE for non-invasive quantitative assessment of hepatic steatosis and fibrosis in NAFLD and NASH: Clinical trials to clinical practice. *J Hepatol, 65*(5), 1006-1016. doi:10.1016/j.jhep.2016.06.005

Dunckley, H. (2012). HLA typing by SSO and SSP methods. *Methods Mol Biol, 882*, 9-25. doi:10.1007/978-1-61779-842-9_2

Duvnjak, M., Barsic, N., Tomasic, V., & Lerotic, I. (2009). Genetic polymorphisms in non-alcoholic fatty liver disease: clues to pathogenesis and disease progression. *World J Gastroenterol, 15*(48), 6023-6027.

Dyson, J. K., Anstee, Q. M., & McPherson, S. (2014). Non-alcoholic fatty liver disease: a practical approach to diagnosis and staging. *Frontline Gastroenterol, 5*(3), 211-218. doi:10.1136/flgastro-2013-100403

Ekstedt, M., Hagstrom, H., Nasr, P., Fredrikson, M., Stal, P., Kechagias, S., & Hultcrantz, R. (2015). Fibrosis stage is the strongest predictor for disease-specific mortality in NAFLD after up to 33 years of follow-up. *Hepatology, 61*(5), 1547-1554. doi:10.1002/hep.27368

Ellis, E. L., & Mann, D. A. (2012). Clinical evidence for the regression of liver fibrosis. *J Hepatol, 56*(5), 1171-1180. doi:10.1016/j.jhep.2011.09.024

Enache, L. S., Enache, E. L., Ramiere, C., Diaz, O., Bancu, L., Sin, A., & Andre, P. (2014). Circulating RNA molecules as biomarkers in liver disease. *Int J Mol Sci, 15*(10), 17644-17666. doi:10.3390/ijms151017644

Engelhard, V. H. (1994). Structure of peptides associated with MHC class I molecules. *Curr Opin Immunol, 6*(1), 13-23.

Erlich, H. (2012). HLA DNA typing: past, present, and future. *Tissue Antigens, 80*(1), 1-11. doi:10.1111/j.1399-0039.2012.01881.x

Estep, M., Armistead, D., Hossain, N., Elarainy, H., Goodman, Z., Baranova, A., . . . Younossi, Z. M. (2010). Differential expression of miRNAs in the visceral

adipose tissue of patients with non-alcoholic fatty liver disease. *Aliment Pharmacol Ther, 32*(3), 487-497. doi:10.1111/j.1365-2036.2010.04366.x

European Association for the Study of the, L., European Association for the Study of, D., & European Association for the Study of, O. (2016). EASL-EASD-EASO Clinical Practice Guidelines for the management of non-alcoholic fatty liver disease. *Diabetologia, 59*(6), 1121-1140. doi:10.1007/s00125-016-3902-y

Fabbrini, E., Sullivan, S., & Klein, S. (2010). Obesity and nonalcoholic fatty liver disease: biochemical, metabolic, and clinical implications. *Hepatology, 51*(2), 679-689. doi:10.1002/hep.23280

Falk, K., Rotzschke, O., Stevanovic, S., Jung, G., & Rammensee, H. G. (1991). Allele-specific motifs revealed by sequencing of self-peptides eluted from MHC molecules. *Nature, 351*(6324), 290-296. doi:10.1038/351290a0

Fang, Y. L., Chen, H., Wang, C. L., & Liang, L. (2018). Pathogenesis of non-alcoholic fatty liver disease in children and adolescence: From "two hit theory" to "multiple hit model". *World J Gastroenterol, 24*(27), 2974-2983. doi:10.3748/wjg.v24.i27.2974

Fassio, E., Alvarez, E., Dominguez, N., Landeira, G., & Longo, C. (2004). Natural history of nonalcoholic steatohepatitis: a longitudinal study of repeat liver biopsies. *Hepatology, 40*(4), 820-826. doi:10.1002/hep.20410

Filipowicz, W., Bhattacharyya, S. N., & Sonenberg, N. (2008). Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? *Nat Rev Genet, 9*(2), 102-114. doi:10.1038/nrg2290

Fitzpatrick, E., & Dhawan, A. (2014). Noninvasive biomarkers in non-alcoholic fatty liver disease: current status and a glimpse of the future. *World J Gastroenterol, 20*(31), 10851-10863. doi:10.3748/wjg.v20.i31.10851

Friedman, S. L. (2005). Mac the knife? Macrophages- the double-edged sword of hepatic fibrosis. *J Clin Invest, 115*(1), 29-32. doi:10.1172/JCI23928

Fujii, H., Enomoto, M., Fukushima, W., Tamori, A., Sakaguchi, H., & Kawada, N. (2009). Applicability of BARD score to Japanese patients with NAFLD. *Gut, 58*(11), 1566-1567; author reply 1567. doi:10.1136/gut.2009.182758

Gallo, A., Tandon, M., Alevizos, I., & Illei, G. G. (2012). The majority of microRNAs detectable in serum and saliva is concentrated in exosomes. *PLoS One, 7*(3), e30679. doi:10.1371/journal.pone.0030679

Ganz, M., & Szabo, G. (2013). Immune and inflammatory pathways in NASH. *Hepatol Int, 7 Suppl 2*, 771-781. doi:10.1007/s12072-013-9468-6

Gerhard, G. S., Chu, X., Wood, G. C., Gerhard, G. M., Benotti, P., Petrick, A. T., . . . Argyropoulos, G. (2013). Next-generation sequence analysis of genes associated with obesity and nonalcoholic fatty liver disease-related cirrhosis in extreme obesity. *Hum Hered, 75*(2-4), 144-151. doi:10.1159/000351719

Gilad, S., Meiri, E., Yogev, Y., Benjamin, S., Lebanony, D., Yerushalmi, N., . . . Chajut, A. (2008). Serum microRNAs are promising novel biomarkers. *PLoS One, 3*(9), e3148. doi:10.1371/journal.pone.0003148

Git, A., Dvinge, H., Salmon-Divon, M., Osborne, M., Kutter, C., Hadfield, J., . . . Caldas, C. (2010). Systematic comparison of microarray profiling, real-time PCR, and next-generation sequencing technologies for measuring differential microRNA expression. *RNA, 16*(5), 991-1006. doi:10.1261/rna.1947110

Glen, J., Floros, L., Day, C., Pryke, R., & Guideline Development, G. (2016). Non-alcoholic fatty liver disease (NAFLD): summary of NICE guidance. *BMJ, 354*, i4428. doi:10.1136/bmj.i4428

Goffredo, M., Caprio, S., Feldstein, A. E., D'Adamo, E., Shaw, M. M., Pierpont, B., . . . Santoro, N. (2016). Role of TM6SF2 rs58542926 in the pathogenesis of nonalcoholic pediatric fatty liver disease: A multiethnic study. *Hepatology, 63*(1), 117-125. doi:10.1002/hep.28283

Goh, G. B., Pagadala, M. R., Dasarathy, J., Unalp-Arida, A., Pai, R. K., Yerian, L., . . . McCullough, A. J. (2015). Age impacts ability of aspartate-alanine aminotransferase ratio to predict advanced fibrosis in nonalcoholic Fatty liver disease. *Dig Dis Sci, 60*(6), 1825-1831. doi:10.1007/s10620-015-3529-8

Goh, S. C., Ho, E. L., & Goh, K. L. (2013). Prevalence and risk factors of non-alcoholic fatty liver disease in a multiracial suburban Asian population in Malaysia. *Hepatol Int, 7*(2), 548-554. doi:10.1007/s12072-012-9359-2

Gorden, A., Yang, R., Yerges-Armstrong, L. M., Ryan, K. A., Speliotes, E., Borecki, I. B., . . . Consortium, G. (2013). Genetic variation at NCAN locus is associated with inflammation and fibrosis in non-alcoholic fatty liver disease in morbid obesity. *Hum Hered, 75*(1), 34-43. doi:10.1159/000346195

Gorker, I., Gurkan, H., Ulusal, S., Atli, E., Ayaz, G., Ceylan, C., . . . Berberoglu, K. K. (2018). Investigation of Copy Number Variation by arrayCGH in Turkish Children and Adolescents Diagnosed with Autism Spectrum Disorders. *Noro Psikiyatr Ars, 55*(3), 215-219. doi:10.5152/npa.2017.21611

Gough, S. C., & Simmonds, M. J. (2007). The HLA Region and Autoimmune Disease: Associations and Mechanisms of Action. *Curr Genomics, 8*(7), 453-465. doi:10.2174/138920207783591690

Grimsby, J., Coffey, J. W., Dvorozniak, M. T., Magram, J., Li, G., Matschinsky, F. M., . . . Grippo, J. F. (2000). Characterization of glucokinase regulatory protein-deficient mice. *J Biol Chem, 275*(11), 7826-7831.

Guerrero, R., Vega, G. L., Grundy, S. M., & Browning, J. D. (2009). Ethnic differences in hepatic steatosis: an insulin resistance paradox? *Hepatology, 49*(3), 791-801. doi:10.1002/hep.22726

Guha, I. N., Parkes, J., Roderick, P., Chattopadhyay, D., Cross, R., Harris, S., . . . Rosenberg, W. M. (2008). Noninvasive markers of fibrosis in nonalcoholic fatty

liver disease: Validating the European Liver Fibrosis Panel and exploring simple markers. *Hepatology, 47*(2), 455-460. doi:10.1002/hep.21984

Hagstrom, H., Elfwen, O., Hultcrantz, R., & Stal, P. (2018). Steatohepatitis Is Not Associated with an Increased Risk for Fibrosis Progression in Nonalcoholic Fatty Liver Disease. *Gastroenterol Res Pract, 2018*, 1942648. doi:10.1155/2018/1942648

Hagstrom, H., Nasr, P., Ekstedt, M., Hammar, U., Stal, P., Hultcrantz, R., & Kechagias, S. (2017). Fibrosis stage but not NASH predicts mortality and time to development of severe liver disease in biopsy-proven NAFLD. *J Hepatol, 67*(6), 1265-1273. doi:10.1016/j.jhep.2017.07.027

Hamaguchi, M., Kojima, T., Takeda, N., Nakagawa, T., Taniguchi, H., Fujii, K., . . . Ida, K. (2005). The metabolic syndrome as a predictor of nonalcoholic fatty liver disease. *Ann Intern Med, 143*(10), 722-728.

Harrison, S. A., Fecht, W., Brunt, E. M., & Neuschwander-Tetri, B. A. (2009). Orlistat for overweight subjects with nonalcoholic steatohepatitis: A randomized, prospective trial. *Hepatology, 49*(1), 80-86. doi:10.1002/hep.22575

Harrison, S. A., Oliver, D., Arnold, H. L., Gogia, S., & Neuschwander-Tetri, B. A. (2008). Development and validation of a simple NAFLD clinical scoring system for identifying patients without advanced disease. *Gut, 57*(10), 1441-1447. doi:10.1136/gut.2007.146019

Harrison, S. A., Torgerson, S., & Hayashi, P. H. (2003). The natural history of nonalcoholic fatty liver disease: a clinical histopathological study. *Am J Gastroenterol, 98*(9), 2042-2047. doi:10.1111/j.1572-0241.2003.07659.x

Hashimoto, E., Yatsuji, S., Tobari, M., Taniai, M., Torii, N., Tokushige, K., & Shiratori, K. (2009). Hepatocellular carcinoma in patients with nonalcoholic steatohepatitis. *J Gastroenterol, 44 Suppl 19*, 89-95. doi:10.1007/s00535-008-2262-x

Hayward, B. E., Dunlop, N., Intody, S., Leek, J. P., Markham, A. F., Warner, J. P., & Bonthron, D. T. (1998). Organization of the human glucokinase regulator gene GCKR. *Genomics, 49*(1), 137-142. doi:10.1006/geno.1997.5195

Hegele, R. A. (2007). Copy-number variations and human disease. *Am J Hum Genet, 81*(2), 414-415; author reply 415. doi:10.1086/519220

Henrichsen, C. N., Vinckenbosch, N., Zollner, S., Chaignat, E., Pradervand, S., Schutz, F., . . . Reymond, A. (2009). Segmental copy number variation shapes tissue transcriptomes. *Nat Genet, 41*(4), 424-429. doi:10.1038/ng.345

Herkel, J., Jagemann, B., Wiegard, C., Lazaro, J. F., Lueth, S., Kanzler, S., . . . Lohse, A. W. (2003). MHC class II-expressing hepatocytes function as antigen-presenting cells and activate specific CD4 T lymphocyutes. *Hepatology, 37*(5), 1079-1085. doi:10.1053/jhep.2003.50191

Hernaez, R., McLean, J., Lazo, M., Brancati, F. L., Hirschhorn, J. N., Borecki, I. B., . . . Speliotes, E. K. (2013). Association between variants in or near PNPLA3, GCKR,

and PPP1R3B with ultrasound-defined steatosis based on data from the third National Health and Nutrition Examination Survey. *Clin Gastroenterol Hepatol, 11*(9), 1183-1190 e1182. doi:10.1016/j.cgh.2013.02.011

Hironaka, K., Sakaida, I., Matsumura, Y., Kaino, S., Miyamoto, K., & Okita, K. (2000). Enhanced interstitial collagenase (matrix metalloproteinase-13) production of Kupffer cell by gadolinium chloride prevents pig serum-induced rat liver fibrosis. *Biochem Biophys Res Commun, 267*(1), 290-295. doi:10.1006/bbrc.1999.1910

Hirschfield, G. M., Liu, X., Xu, C., Lu, Y., Xie, G., Lu, Y., . . . Siminovitch, K. A. (2009). Primary biliary cirrhosis associated with HLA, IL12A, and IL12RB2 variants. *N Engl J Med, 360*(24), 2544-2555. doi:10.1056/NEJMoa0810440

Holling, T. M., Schooten, E., & van Den Elsen, P. J. (2004). Function and regulation of MHC class II molecules in T-lymphocytes: of mice and men. *Hum Immunol, 65*(4), 282-290. doi:10.1016/j.humimm.2004.01.005

Holoshitz, J. (2013). The quest for better understanding of HLA-disease association: scenes from a road less travelled by. *Discov Med, 16*(87), 93-101.

Hou, C., Feng, W., Wei, S., Wang, Y., Xu, X., Wei, J., . . . Zheng, K. (2018). Bioinformatics Analysis of Key Differentially Expressed Genes in Nonalcoholic Fatty Liver Disease Mice Models. *Gene Expr, 19*(1), 25-35. doi:10.3727/105221618X15341831737687

Huntzinger, E., & Izaurralde, E. (2011). Gene silencing by microRNAs: contributions of translational repression and mRNA decay. *Nat Rev Genet, 12*(2), 99-110. doi:10.1038/nrg2936

Husby, S., Koletzko, S., Korponay-Szabo, I. R., Mearin, M. L., Phillips, A., Shamir, R., . . . Nutrition. (2012). European Society for Pediatric Gastroenterology, Hepatology, and Nutrition guidelines for the diagnosis of coeliac disease. *J Pediatr Gastroenterol Nutr, 54*(1), 136-160. doi:10.1097/MPG.0b013e31821a23d0

Ibrahim, S. H., Kohli, R., & Gores, G. J. (2011). Mechanisms of lipotoxicity in NAFLD and clinical implications. *J Pediatr Gastroenterol Nutr, 53*(2), 131-140. doi:10.1097/MPG.0b013e31822578db

International HapMap, C., Altshuler, D. M., Gibbs, R. A., Peltonen, L., Altshuler, D. M., Gibbs, R. A., . . . McEwen, J. E. (2010). Integrating common and rare genetic variation in diverse human populations. *Nature, 467*(7311), 52-58. doi:10.1038/nature09298

Ionita-Laza, I., Rogers, A. J., Lange, C., Raby, B. A., & Lee, C. (2009). Genetic association analysis of copy-number variation (CNV) in human disease pathogenesis. *Genomics, 93*(1), 22-26. doi:10.1016/j.ygeno.2008.08.012

Ishak, K., Baptista, A., Bianchi, L., Callea, F., De Groote, J., Gudat, F., . . . et al. (1995). Histological grading and staging of chronic hepatitis. *J Hepatol, 22*(6), 696-699.

Jakobsen, M. U., Berentzen, T., Sorensen, T. I., & Overvad, K. (2007). Abdominal obesity and fatty liver. *Epidemiol Rev, 29*, 77-87. doi:10.1093/epirev/mxm002

Jaksik, R., Iwanaszko, M., Rzeszowska-Wolny, J., & Kimmel, M. (2015). Microarray experiments and factors which affect their reliability. *Biol Direct, 10*, 46. doi:10.1186/s13062-015-0077-2

Jia, D., Wei, L., Guo, W., Zha, R., Bao, M., Chen, Z., . . . He, X. (2011). Genome-wide copy number analyses identified novel cancer genes in hepatocellular carcinoma. *Hepatology, 54*(4), 1227-1236. doi:10.1002/hep.24495

Jiang, D. K., Sun, J., Cao, G., Liu, Y., Lin, D., Gao, Y. Z., . . . Yu, L. (2013). Genetic variants in STAT4 and HLA-DQ genes confer risk of hepatitis B virus-related hepatocellular carcinoma. *Nat Genet, 45*(1), 72-75. doi:10.1038/ng.2483

Jin, P., & Wang, E. (2003). Polymorphism in clinical immunology - From HLA typing to immunogenetic profiling. *J Transl Med, 1*(1), 8. doi:10.1186/1479-5876-1-8

Jones, D. E., & Donaldson, P. T. (2003). Genetic factors in the pathogenesis of primary biliary cirrhosis. *Clin Liver Dis, 7*(4), 841-864.

Jorde, L. B., & Wooding, S. P. (2004). Genetic variation, classification and 'race'. *Nat Genet, 36*(11 Suppl), S28-33. doi:10.1038/ng1435

Kagansky, N., Levy, S., Keter, D., Rimon, E., Taiba, Z., Fridman, Z., . . . Malnick, S. (2004). Non-alcoholic fatty liver disease--a common and benign finding in octogenarian patients. *Liver Int, 24*(6), 588-594. doi:10.1111/j.1478-3231.2004.0969.x

Kalia, H. S., & Gaglio, P. J. (2016). The Prevalence and Pathobiology of Nonalcoholic Fatty Liver Disease in Patients of Different Races or Ethnicities. *Clin Liver Dis, 20*(2), 215-224. doi:10.1016/j.cld.2015.10.005

Kamatani, Y., Wattanapokayakit, S., Ochi, H., Kawaguchi, T., Takahashi, A., Hosono, N., . . . Matsuda, K. (2009). A genome-wide association study identifies variants in the HLA-DP locus associated with chronic hepatitis B in Asians. *Nat Genet, 41*(5), 591-595. doi:10.1038/ng.348

Karlsen, T. H., Franke, A., Melum, E., Kaser, A., Hov, J. R., Balschun, T., . . . Schreiber, S. (2010). Genome-wide association analysis in primary sclerosing cholangitis. *Gastroenterology, 138*(3), 1102-1111. doi:10.1053/j.gastro.2009.11.046

Kaswala, D. H., Lai, M., & Afdhal, N. H. (2016). Fibrosis Assessment in Nonalcoholic Fatty Liver Disease (NAFLD) in 2016. *Dig Dis Sci, 61*(5), 1356-1364. doi:10.1007/s10620-016-4079-4

Khimji, A. K., Shao, R., & Rockey, D. C. (2008). Divergent transforming growth factor-beta signaling in hepatic stellate cells after liver injury: functional effects on ECE-1 regulation. *Am J Pathol, 173*(3), 716-727. doi:10.2353/ajpath.2008.071121

Khov, N., Sharma, A., & Riley, T. R. (2014). Bedside ultrasound in the diagnosis of nonalcoholic fatty liver disease. *World J Gastroenterol, 20*(22), 6821-6825. doi:10.3748/wjg.v20.i22.6821

Kim, D., Kim, W. R., Kim, H. J., & Therneau, T. M. (2013). Association between noninvasive fibrosis markers and mortality among adults with nonalcoholic fatty liver disease in the United States. *Hepatology, 57*(4), 1357-1365. doi:10.1002/hep.26156

Kim, D., Kim, W. R., Talwalkar, J. A., Kim, H. J., & Ehman, R. L. (2013). Advanced fibrosis in nonalcoholic fatty liver disease: noninvasive assessment with MR elastography. *Radiology, 268*(2), 411-419. doi:10.1148/radiol.13121193

Kirschner, M. B., Edelman, J. J., Kao, S. C., Vallely, M. P., van Zandwijk, N., & Reid, G. (2013). The Impact of Hemolysis on Cell-Free microRNA Biomarkers. *Front Genet, 4*, 94. doi:10.3389/fgene.2013.00094

Kirschner, M. B., Kao, S. C., Edelman, J. J., Armstrong, N. J., Vallely, M. P., van Zandwijk, N., & Reid, G. (2011). Haemolysis during sample preparation alters microRNA content of plasma. *PLoS One, 6*(9), e24145. doi:10.1371/journal.pone.0024145

Kitamoto, A., Kitamoto, T., Nakamura, T., Ogawa, Y., Yoneda, M., Hyogo, H., . . . Hotta, K. (2014). Association of polymorphisms in GCKR and TRIB1 with nonalcoholic fatty liver disease and metabolic syndrome traits. *Endocr J, 61*(7), 683-689.

Klein, J., & Sato, A. (2000). The HLA system. First of two parts. *N Engl J Med, 343*(10), 702-709. doi:10.1056/NEJM200009073431006

Kleiner, D. E., Brunt, E. M., Van Natta, M., Behling, C., Contos, M. J., Cummings, O. W., . . . Nonalcoholic Steatohepatitis Clinical Research, N. (2005). Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology, 41*(6), 1313-1321. doi:10.1002/hep.20701

Kolios, G., Valatas, V., & Kouroumalis, E. (2006). Role of Kupffer cells in the pathogenesis of liver disease. *World J Gastroenterol, 12*(46), 7413-7420.

Komatsu, S., Ichikawa, D., Takeshita, H., Tsujiura, M., Morimura, R., Nagata, H., . . . Otsuji, E. (2011). Circulating microRNAs in plasma of patients with oesophageal squamous cell carcinoma. *Br J Cancer, 105*(1), 104-111. doi:10.1038/bjc.2011.198

Kozlitina, J., Smagris, E., Stender, S., Nordestgaard, B. G., Zhou, H. H., Tybjaerg-Hansen, A., . . . Cohen, J. C. (2014). Exome-wide association study identifies a TM6SF2 variant that confers susceptibility to nonalcoholic fatty liver disease. *Nat Genet, 46*(4), 352-356. doi:10.1038/ng.2901

Kroh, E. M., Parkin, R. K., Mitchell, P. S., & Tewari, M. (2010). Analysis of circulating microRNA biomarkers in plasma and serum using quantitative reverse transcription-PCR (qRT-PCR). *Methods, 50*(4), 298-301. doi:10.1016/j.ymeth.2010.01.032

Kubes, P., & Mehal, W. Z. (2012). Sterile inflammation in the liver. *Gastroenterology, 143*(5), 1158-1172. doi:10.1053/j.gastro.2012.09.008

Kulcsarova, E., Kralovicova, J., Parnicka, Z., Ferencik, S., & Buc, M. (2000). [Comparison of the results of HLA typing using serologic and molecular genetics methods]. *Bratisl Lek Listy, 101*(3), 134-137.

Kumar, V., Kato, N., Urabe, Y., Takahashi, A., Muroyama, R., Hosono, N., . . . Matsuda, K. (2011). Genome-wide association study identifies a susceptibility locus for HCV-induced hepatocellular carcinoma. *Nat Genet, 43*(5), 455-458. doi:10.1038/ng.809

Larter, C. Z., Chitturi, S., Heydet, D., & Farrell, G. C. (2010). A fresh look at NASH pathogenesis. Part 1: the metabolic movers. *J Gastroenterol Hepatol, 25*(4), 672-690. doi:10.1111/j.1440-1746.2010.06253.x

Lawrie, C. H., Gal, S., Dunlop, H. M., Pushkaran, B., Liggins, A. P., Pulford, K., . . . Harris, A. L. (2008). Detection of elevated levels of tumour-associated microRNAs in serum of patients with diffuse large B-cell lymphoma. *Br J Haematol, 141*(5), 672-675. doi:10.1111/j.1365-2141.2008.07077.x

Lee, I., Ajay, S. S., Chen, H., Maruyama, A., Wang, N., McInnis, M. G., & Athey, B. D. (2008). Discriminating single-base difference miRNA expressions using microarray Probe Design Guru (ProDeG). *Nucleic Acids Res, 36*(5), e27. doi:10.1093/nar/gkm1165

Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., . . . Kim, V. N. (2003). The nuclear RNase III Drosha initiates microRNA processing. *Nature, 425*(6956), 415-419. doi:10.1038/nature01957

Li, L., Liu, H., Hu, X., Huang, Y., Wang, Y., He, Y., & Lei, Q. (2018). Identification of key genes in nonalcoholic fatty liver disease progression based on bioinformatics analysis. *Mol Med Rep, 17*(6), 7708-7720. doi:10.3892/mmr.2018.8852

Liang, L., Wong, C. M., Ying, Q., Fan, D. N., Huang, S., Ding, J., . . . He, X. (2010). MicroRNA-125b suppressesed human liver cancer cell proliferation and metastasis by directly targeting oncogene LIN28B2. *Hepatology, 52*(5), 1731-1740. doi:10.1002/hep.23904

Liang, X. T., Pan, K., Chen, M. S., Li, J. J., Wang, H., Zhao, J. J., . . . Xia, J. C. (2011). Decreased expression of XPO4 is associated with poor prognosis in hepatocellular carcinoma. *J Gastroenterol Hepatol, 26*(3), 544-549. doi:10.1111/j.1440-1746.2010.06434.x

Lim, L. P., Lau, N. C., Garrett-Engele, P., Grimson, A., Schelter, J. M., Castle, J., . . . Johnson, J. M. (2005). Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. *Nature, 433*(7027), 769-773. doi:10.1038/nature03315

Lipowsky, G., Bischoff, F. R., Schwarzmaier, P., Kraft, R., Kostka, S., Hartmann, E., . . . Gorlich, D. (2000). Exportin 4: a mediator of a novel nuclear export pathway in higher eukaryotes. *EMBO J, 19*(16), 4362-4371. doi:10.1093/emboj/19.16.4362

Liu, X. L., Pan, Q., Zhang, R. N., Shen, F., Yan, S. Y., Sun, C., . . . Fan, J. G. (2016). Disease-specific miR-34a as diagnostic marker of non-alcoholic steatohepatitis in a Chinese population. *World J Gastroenterol, 22*(44), 9844-9852. doi:10.3748/wjg.v22.i44.9844

Loiseau, P., al-Daccak, R., David, F., & Colombani, J. (1991). Molecular HLA-class II typing: advantages and clinical application. *Nouv Rev Fr Hematol, 33*(6), 461-463.

Lonjou, C., Borot, N., Sekula, P., Ledger, N., Thomas, L., Halevy, S., . . . Regi, S. s. g. (2008). A European study of HLA-B in Stevens-Johnson syndrome and toxic epidermal necrolysis related to five high-risk drugs. *Pharmacogenet Genomics, 18*(2), 99-107. doi:10.1097/FPC.0b013e3282f3ef9c

Lopez-Riera, M., Conde, I., Quintas, G., Pedrola, L., Zaragoza, A., Perez-Rojas, J., . . . Jover, R. (2018). Non-invasive prediction of NAFLD severity: a comprehensive, independent validation of previously postulated serum microRNA biomarkers. *Sci Rep, 8*(1), 10606. doi:10.1038/s41598-018-28854-4

Lopez, I. C., Aroca, F. G., Bernal, M. D. F., Mompean, J. A. L., Bernal, A. B., Martinez, A. M. H., . . . Paricio, P. P. (2017). Utility of the ELF Test for Detecting Steatohepatitis in Morbid Obese Patients with Suspicion of Nonalcoholic Fatty Liver Disease. *Obes Surg, 27*(9), 2347-2353. doi:10.1007/s11695-017-2606-9

Losekann, A., Weston, A. C., de Mattos, A. A., Tovo, C. V., de Carli, L. A., Espindola, M. B., . . . Coral, G. P. (2015). Non-Alcoholic Steatohepatitis (NASH): Risk Factors in Morbidly Obese Patients. *Int J Mol Sci, 16*(10), 25552-25559. doi:10.3390/ijms161025552

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. doi:10.1086/518312

Lund, E., Guttinger, S., Calado, A., Dahlberg, J. E., & Kutay, U. (2004). Nuclear export of microRNA precursors. *Science, 303*(5654), 95-98. doi:10.1126/science.1090599

Macaluso, F. S., Maida, M., & Petta, S. (2015). Genetic background in nonalcoholic fatty liver disease: A comprehensive review. *World J Gastroenterol, 21*(39), 11088-11111. doi:10.3748/wjg.v21.i39.11088

MacDonald, J. R., Ziman, R., Yuen, R. K., Feuk, L., & Scherer, S. W. (2014). The Database of Genomic Variants: a curated collection of structural variation in the human genome. *Nucleic Acids Res, 42*(Database issue), D986-992. doi:10.1093/nar/gkt958

Mace, A., Kutalik, Z., & Valsesia, A. (2018). Copy Number Variation. *Methods Mol Biol, 1793*, 231-258. doi:10.1007/978-1-4939-7868-7_14

Mahdi, B. M. (2013). A glow of HLA typing in organ transplantation. *Clin Transl Med, 2*(1), 6. doi:10.1186/2001-1326-2-6

Makkonen, J., Pietilainen, K. H., Rissanen, A., Kaprio, J., & Yki-Jarvinen, H. (2009). Genetic factors contribute to variation in serum alanine aminotransferase activity independent of obesity and alcohol: a study in monozygotic and dizygotic twins. *J Hepatol, 50*(5), 1035-1042. doi:10.1016/j.jhep.2008.12.025

Mangalam, A. K., Taneja, V., & David, C. S. (2013). HLA class II molecules influence susceptibility versus protection in inflammatory diseases by determining the cytokine profile. *J Immunol, 190*(2), 513-518. doi:10.4049/jimmunol.1201891

Marabita, F., de Candia, P., Torri, A., Tegner, J., Abrignani, S., & Rossi, R. L. (2015). Normalization of circulating microRNA expression data obtained by quantitative real-time RT-PCR. *Brief Bioinform*. doi:10.1093/bib/bbv056

Marabita, F., de Candia, P., Torri, A., Tegner, J., Abrignani, S., & Rossi, R. L. (2016). Normalization of circulating microRNA expression data obtained by quantitative real-time RT-PCR. *Brief Bioinform, 17*(2), 204-212. doi:10.1093/bib/bbv056

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.

Mause, S. F., & Weber, C. (2010). Microparticles: protagonists of a novel communication network for intercellular information exchange. *Circ Res, 107*(9), 1047-1057. doi:10.1161/CIRCRESAHA.110.226456

Mayeux, R. (2004). Biomarkers: potential uses and limitations. *NeuroRx, 1*(2), 182-188. doi:10.1602/neurorx.1.2.182

Mbarek, H., Ochi, H., Urabe, Y., Kumar, V., Kubo, M., Hosono, N., . . . Matsuda, K. (2011). A genome-wide association study of chronic hepatitis B identified novel risk locus in a Japanese population. *Hum Mol Genet, 20*(19), 3884-3892. doi:10.1093/hmg/ddr301

McCarty, T. R., Echouffo-Tcheugui, J. B., Lange, A., Haque, L., & Njei, B. (2018). Impact of bariatric surgery on outcomes of patients with nonalcoholic fatty liver disease: a nationwide inpatient sample analysis, 2004-2012. *Surg Obes Relat Dis, 14*(1), 74-80. doi:10.1016/j.soard.2017.09.511

McDonald, J. S., Milosevic, D., Reddi, H. V., Grebe, S. K., & Algeciras-Schimnich, A. (2011). Analysis of circulating microRNA: preanalytical and analytical challenges. *Clin Chem, 57*(6), 833-840. doi:10.1373/clinchem.2010.157198

McPherson, S., Anstee, Q. M., Henderson, E., Day, C. P., & Burt, A. D. (2013). Are simple noninvasive scoring systems for fibrosis reliable in patients with NAFLD and normal ALT levels? *Eur J Gastroenterol Hepatol, 25*(6), 652-658. doi:10.1097/MEG.0b013e32835d72cf

McPherson, S., Hardy, T., Henderson, E., Burt, A. D., Day, C. P., & Anstee, Q. M. (2015). Evidence of NAFLD progression from steatosis to fibrosing-steatohepatitis using paired biopsies: implications for prognosis and clinical management. *J Hepatol, 62*(5), 1148-1155. doi:10.1016/j.jhep.2014.11.034

McPherson, S., Stewart, S. F., Henderson, E., Burt, A. D., & Day, C. P. (2010). Simple non-invasive fibrosis scoring systems can reliably exclude advanced fibrosis in patients with non-alcoholic fatty liver disease. *Gut, 59*(9), 1265-1269. doi:10.1136/gut.2010.216077

Mefford, J., & Witte, J. S. (2012). The Covariate's Dilemma. *P Lo S Genet, 8*, e1003096.

Meister, G. (2013). Argonaute proteins: functional insights and emerging roles. *Nat Rev Genet, 14*(7), 447-459. doi:10.1038/nrg3462

Mestdagh, P., Van Vlierberghe, P., De Weer, A., Muth, D., Westermann, F., Speleman, F., & Vandesompele, J. (2009). A novel and universal method for microRNA RT-qPCR data normalization. *Genome Biol, 10*(6), R64. doi:10.1186/gb-2009-10-6-r64

Miele, L., Dall'armi, V., Cefalo, C., Nedovic, B., Arzani, D., Amore, R., . . . Boccia, S. (2014). A case-control study on the effect of metabolic gene polymorphisms, nutrition, and their interaction on the risk of non-alcoholic fatty liver disease. *Genes Nutr, 9*(2), 383. doi:10.1007/s12263-013-0383-1

Miele, L., De Michele, T., Marrone, G., Antonietta Isgro, M., Basile, U., Cefalo, C., . . . Grieco, A. (2017). Enhanced liver fibrosis test as a reliable tool for assessing fibrosis in nonalcoholic fatty liver disease in a clinical setting. *Int J Biol Markers, 32*(4), e397-e402. doi:10.5301/ijbm.5000292

Mitchell, P. S., Parkin, R. K., Kroh, E. M., Fritz, B. R., Wyman, S. K., Pogosova-Agadjanyan, E. L., . . . Tewari, M. (2008). Circulating microRNAs as stable blood-based markers for cancer detection. *Proc Natl Acad Sci U S A, 105*(30), 10513-10518. doi:10.1073/pnas.0804549105

Mitchison, N. A. (2004). T-cell-B-cell cooperation. *Nat Rev Immunol, 4*(4), 308-312. doi:10.1038/nri1334

Miyaaki, H., Ichikawa, T., Kamo, Y., Taura, N., Honda, T., Shibata, H., . . . Nakao, K. (2014). Significance of serum and hepatic microRNA-122 levels in patients with non-alcoholic fatty liver disease. *Liver Int, 34*(7), e302-307. doi:10.1111/liv.12429

Mohamed, R. A., Nabih, M. L., ElShobaky, M. B., & Khattab, H. M. (2014). The value of noninvasive scoring systems for the diagnosis of advanced fibrosis in Egyptian patients with nonalcoholic fatty liver disease. *Egypt J Intern Med, 26*(4), 162-169.

Mohanty, S. R., Troy, T. N., Huo, D., O'Brien, B. L., Jensen, D. M., & Hart, J. (2009). Influence of ethnicity on histological differences in non-alcoholic fatty liver disease. *J Hepatol, 50*(4), 797-804. doi:10.1016/j.jhep.2008.11.017

Murakami, Y., Toyoda, H., Tanaka, M., Kuroda, M., Harada, Y., Matsuda, F., . . . Shimotohno, K. (2011). The progression of liver fibrosis is related with overexpression of the miR-199 and 200 families. *PLoS One, 6*(1), e16081. doi:10.1371/journal.pone.0016081

Musso, G., Gambino, R., Cassader, M., & Pagano, G. (2011). Meta-analysis: natural history of non-alcoholic fatty liver disease (NAFLD) and diagnostic accuracy of non-invasive tests for liver disease severity. *Ann Med, 43*(8), 617-649. doi:10.3109/07853890.2010.518623

Nalbantoglu, I. L., & Brunt, E. M. (2014). Role of liver biopsy in nonalcoholic fatty liver disease. *World J Gastroenterol, 20*(27), 9026-9037. doi:10.3748/wjg.v20.i27.9026

Nielsen, R., Paul, J. S., Albrechtsen, A., & Song, Y. S. (2011). Genotype and SNP calling from next-generation sequencing data. *Nat Rev Genet, 12*, 443-451.

Niu, L., Cui, X., Qi, Y., Xie, D., Wu, Q., Chen, X., . . . Liu, Z. (2016). Involvement of TGF-beta1/Smad3 Signaling in Carbon Tetrachloride-Induced Acute Liver Injury in Mice. *PLoS One, 11*(5), e0156090. doi:10.1371/journal.pone.0156090

Niwa, R., & Slack, F. J. (2007). The evolution of animal microRNA function. *Curr Opin Genet Dev, 17*(2), 145-150. doi:10.1016/j.gde.2007.02.004

Nobili, V., Donati, B., Panera, N., Vongsakulyanon, A., Alisi, A., Dallapiccola, B., & Valenti, L. (2014). A 4-polymorphism risk score predicts steatohepatitis in children with nonalcoholic fatty liver disease. *J Pediatr Gastroenterol Nutr, 58*(5), 632-636. doi:10.1097/MPG.0000000000000279

Nobili, V., Parkes, J., Bottazzo, G., Marcellini, M., Cross, R., Newman, D., . . . Rosenberg, W. M. (2009). Performance of ELF serum markers in predicting fibrosis stage in pediatric non-alcoholic fatty liver disease. *Gastroenterology, 136*(1), 160-167. doi:10.1053/j.gastro.2008.09.013

Noble, J. A., & Valdes, A. M. (2011). Genetics of the HLA region in the prediction of type 1 diabetes. *Curr Diab Rep, 11*(6), 533-542. doi:10.1007/s11892-011-0223-x

Nowakowska, B. (2017). Clinical interpretation of copy number variants in the human genome. *J Appl Genet, 58*(4), 449-457. doi:10.1007/s13353-017-0407-4

Obika, M., & Noguchi, H. (2012). Diagnosis and evaluation of nonalcoholic fatty liver disease. *Exp Diabetes Res, 2012*, 145754. doi:10.1155/2012/145754

Okamoto, M., Miyake, T., Kitai, K., Furukawa, S., Yamamoto, S., Senba, H., . . . Hiasa, Y. (2018). Cigarette smoking is a risk factor for the onset of fatty liver disease in nondrinkers: A longitudinal cohort study. *PLoS One, 13*(4), e0195147. doi:10.1371/journal.pone.0195147

Ollila, H. M., Fernandez-Vina, M., & Mignot, E. (2015). HLA-DQ allele competition in narcolepsy: a comment on Tafti et al. DQB1 locus alone explains most of the risk and protection in narcolepsy with cataplexy in Europe. *Sleep, 38*(1), 147-151. doi:10.5665/sleep.4342

Orho-Melander, M., Melander, O., Guiducci, C., Perez-Martinez, P., Corella, D., Roos, C., . . . Kathiresan, S. (2008). Common missense variant in the glucokinase regulatory protein gene is associated with increased plasma triglyceride and

C-reactive protein but lower fasting glucose concentrations. *Diabetes, 57*(11), 3112-3121. doi:10.2337/db08-0516

Pacifici, M., Delbue, S., Kadri, F., & Peruzzi, F. (2014). Cerebrospinal fluid MicroRNA profiling using quantitative real time PCR. *J Vis Exp*(83), e51172. doi:10.3791/51172

Pais, R., Charlotte, F., Fedchuk, L., Bedossa, P., Lebray, P., Poynard, T., . . . Group, L. S. (2013). A systematic review of follow-up biopsies reveals disease progression in patients with non-alcoholic fatty liver. *J Hepatol, 59*(3), 550-556. doi:10.1016/j.jhep.2013.04.027

Palmer, N. D., Musani, S. K., Yerges-Armstrong, L. M., Feitosa, M. F., Bielak, L. F., Hernaez, R., . . . Speliotes, E. K. (2013). Characterization of European ancestry nonalcoholic fatty liver disease-associated variants in individuals of African and Hispanic descent. *Hepatology, 58*(3), 966-975. doi:10.1002/hep.26440

Paunic, V., Gragert, L., Madbouly, A., Freeman, J., & Maiers, M. (2012). Measuring ambiguity in HLA typing methods. *PLoS One, 7*(8), e43585. doi:10.1371/journal.pone.0043585

Peltier, H. J., & Latham, G. J. (2008). Normalization of microRNA expression levels in quantitative RT-PCR assays: identification of suitable reference RNA targets in normal and cancerous human solid tissues. *RNA, 14*(5), 844-852. doi:10.1261/rna.939908

Perez-Gutierrez, O. Z., Hernandez-Rocha, C., Candia-Balboa, R. A., Arrese, M. A., Benitez, C., Brizuela-Alcantara, D. C., . . . Chavez-Tapia, N. C. (2013). Validation study of systems for noninvasive diagnosis of fibrosis in nonalcoholic fatty liver disease in Latin population. *Ann Hepatol, 12*(3), 416-424.

Periasamy, M. (2007). Indian Migration into Malaya and Singapore During the British Period. *Biblioasia. Singapore: National Library Board Singapore*.

Peter, A., Stefan, N., Cegan, A., Walenta, M., Wagner, S., Konigsrainer, A., . . . Schleicher, E. (2011). Hepatic glucokinase expression is associated with lipogenesis and fatty liver in humans. *J Clin Endocrinol Metab, 96*(7), E1126-1130. doi:10.1210/jc.2010-2017

Petit, J. M., Masson, D., Guiu, B., Rollot, F., Duvillard, L., Bouillet, B., . . . Verges, B. (2016). GCKR polymorphism influences liver fat content in patients with type 2 diabetes. *Acta Diabetol, 53*(2), 237-242. doi:10.1007/s00592-015-0766-4

Petta, S., Miele, L., Bugianesi, E., Camma, C., Rosso, C., Boccia, S., . . . Craxi, A. (2014). Glucokinase regulatory protein gene polymorphism affects liver fibrosis in non-alcoholic fatty liver disease. *PLoS One, 9*(2), e87523. doi:10.1371/journal.pone.0087523

Peverill, W., Powell, L. W., & Skoien, R. (2014). Evolving concepts in the pathogenesis of NASH: beyond steatosis and inflammation. *Int J Mol Sci, 15*(5), 8591-8638. doi:10.3390/ijms15058591

Piao, X., Zhang, X., Wu, L., & Belasco, J. G. (2010). CCR4-NOT deadenylates mRNA associated with RNA-induced silencing complexes in human cells. *Mol Cell Biol, 30*(6), 1486-1494. doi:10.1128/MCB.01481-09

Pillai, R. S., Bhattacharyya, S. N., & Filipowicz, W. (2007). Repression of protein synthesis by miRNAs: how many mechanisms? *Trends Cell Biol, 17*(3), 118-126. doi:10.1016/j.tcb.2006.12.007

Pinto, D., Marshall, C., Feuk, L., & Scherer, S. W. (2007). Copy-number variation in control population cohorts. *Hum Mol Genet, 16 Spec No. 2*, R168-173. doi:10.1093/hmg/ddm241

Pirazzi, C., Adiels, M., Burza, M. A., Mancina, R. M., Levin, M., Stahlman, M., . . . Romeo, S. (2012). Patatin-like phospholipase domain-containing 3 (PNPLA3) I148M (rs738409) affects hepatic VLDL secretion in humans and in vitro. *J Hepatol, 57*(6), 1276-1282. doi:10.1016/j.jhep.2012.07.030

Pirola, C. J., Fernandez Gianotti, T., Castano, G. O., Mallardi, P., San Martino, J., Mora Gonzalez Lopez Ledesma, M., . . . Sookoian, S. (2015). Circulating microRNA signature in non-alcoholic fatty liver disease: from serum non-coding RNAs to liver histology and disease pathogenesis. *Gut, 64*(5), 800-812. doi:10.1136/gutjnl-2014-306996

Prinja, S., Bahuguna, P., Duseja, A., Kaur, M., & Chawla, Y. K. (2018). Cost of Intensive Care Treatment for Liver Disorders at Tertiary Care Level in India. *Pharmacoecon Open, 2*(2), 179-190. doi:10.1007/s41669-017-0041-4

Pritchard, C. C., Cheng, H. H., & Tewari, M. (2012). MicroRNA profiling: approaches and considerations. *Nat Rev Genet, 13*(5), 358-369. doi:10.1038/nrg3198

Pritchard, C. C., Kroh, E., Wood, B., Arroyo, J. D., Dougherty, K. J., Miyaji, M. M., . . . Tewari, M. (2012). Blood cell origin of circulating microRNAs: a cautionary note for cancer biomarker studies. *Cancer Prev Res (Phila), 5*(3), 492-497. doi:10.1158/1940-6207.CAPR-11-0370

Promrat, K., Kleiner, D. E., Niemeier, H. M., Jackvony, E., Kearns, M., Wands, J. R., . . . Wing, R. R. (2010). Randomized controlled trial testing the effects of weight loss on nonalcoholic steatohepatitis. *Hepatology, 51*(1), 121-129. doi:10.1002/hep.23276

Qureshi, K., Clements, R. H., & Abrams, G. A. (2008). The utility of the "NAFLD fibrosis score" in morbidly obese subjects with NAFLD. *Obes Surg, 18*(3), 264-270. doi:10.1007/s11695-007-9295-8

Racanelli, V., & Rehermann, B. (2006). The liver as an immunological organ. *Hepatology, 43*(2 Suppl 1), S54-62. doi:10.1002/hep.21060

Raitoharju, E., Seppala, I., Lyytikainen, L. P., Viikari, J., Ala-Korpela, M., Soininen, P., . . . Lehtimaki, T. (2016). Blood hsa-miR-122-5p and hsa-miR-885-5p levels associate with fatty liver and related lipoprotein metabolism-The Young Finns Study. *Sci Rep, 6*, 38262. doi:10.1038/srep38262

Raszeja-Wyszomirska, J., Szymanik, B., Lawniczak, M., Kajor, M., Chwist, A., Milkiewicz, P., & Hartleb, M. (2010). Validation of the BARD scoring system in Polish patients with nonalcoholic fatty liver disease (NAFLD). *BMC Gastroenterol, 10*, 67. doi:10.1186/1471-230X-10-67

Ratziu, V., Bellentani, S., Cortez-Pinto, H., Day, C., & Marchesini, G. (2010). A position statement on NAFLD/NASH based on the EASL 2009 special conference. *J Hepatol, 53*(2), 372-384. doi:10.1016/j.jhep.2010.04.008

Rees, M. G., Wincovitch, S., Schultz, J., Waterstradt, R., Beer, N. L., Baltrusch, S., . . . Gloyn, A. L. (2012). Cellular characterisation of the GCKR P446L variant associated with type 2 diabetes risk. *Diabetologia, 55*(1), 114-122. doi:10.1007/s00125-011-2348-5

Regnell, S. E., & Lernmark, A. (2011). Hepatic steatosis in type 1 diabetes. *Rev Diabet Stud, 8*(4), 454-467. doi:10.1900/RDS.2011.8.454

Roberts, T. C., Coenen-Stass, A. M., & Wood, M. J. (2014). Assessment of RT-qPCR normalization strategies for accurate quantification of extracellular microRNAs in murine serum. *PLoS One, 9*(2), e89237. doi:10.1371/journal.pone.0089237

Rockey, D. C., Caldwell, S. H., Goodman, Z. D., Nelson, R. C., Smith, A. D., & American Association for the Study of Liver, D. (2009). Liver biopsy. *Hepatology, 49*(3), 1017-1044. doi:10.1002/hep.22742

Roessler, S., Lin, G., Forgues, M., Budhu, A., Hoover, S., Simpson, R. M., . . . Wang, X. W. (2015). Integrative genomic and transcriptomic characterization of matched primary and metastatic liver and colorectal carcinoma. *Int J Biol Sci, 11*(1), 88-98. doi:10.7150/ijbs.10583

Romeo, S., Cohen, J. C., & Hobbs, H. H. (2006). No association between polymorphism in PEMT (V175M) and hepatic triglyceride content in the Dallas Heart Study. *FASEB J, 20*(12), 2180; author reply 2181-2182. doi:10.1096/fj.06-1004ufm

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. doi:10.1038/ng.257

Rosenberg, W. M., Voelker, M., Thiel, R., Becka, M., Burt, A., Schuppan, D., . . . European Liver Fibrosis, G. (2004). Serum markers detect the presence of liver fibrosis: a cohort study. *Gastroenterology, 127*(6), 1704-1713.

Royo, F., Zabala, A., Paz, N., Acquadro, F., Echevarria, J. J., Zabalza, I., . . . Parada, L. A. (2013). Genome-wide analysis of DNA copy number changes in liver steatosis. *Br J Med Med Res, 3*(4), 1773-1785.

Ruffillo, G., Fassio, E., Alvarez, E., Landeira, G., Longo, C., Dominguez, N., & Gualano, G. (2011). Comparison of NAFLD fibrosis score and BARD score in predicting fibrosis in nonalcoholic fatty liver disease. *J Hepatol, 54*(1), 160-163. doi:10.1016/j.jhep.2010.06.028

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.

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. doi:10.1038/35057149

Salameh, H., Hanayneh, M. A., Masadeh, M., Naseemuddin, M., Matin, T., Erwin, A., & Singal, A. K. (2016). PNPLA3 as a Genetic Determinant of Risk for and Severity of Non-alcoholic Fatty Liver Disease Spectrum. *J Clin Transl Hepatol, 4*(3), 175-191. doi:10.14218/JCTH.2016.00009

Samuel, V. T., Petersen, K. F., & Shulman, G. I. (2010). Lipid-induced insulin resistance: unravelling the mechanism. *Lancet, 375*(9733), 2267-2277. doi:10.1016/S0140-6736(10)60408-4

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. doi:10.1002/hep.24806

Santos, V. N., Leite-Mor, M. M., Kondo, M., Martins, J. R., Nader, H., Lanzoni, V. P., & Parise, E. R. (2005). Serum laminin, type IV collagen and hyaluronan as fibrosis markers in non-alcoholic fatty liver disease. *Braz J Med Biol Res, 38*(5), 747-753. doi:/S0100-879X2005000500012

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

Sanyal, A. J., Banas, C., Sargeant, C., Luketic, V. A., Sterling, R. K., Stravitz, R. T., . . . Mills, A. S. (2006). Similarities and differences in outcomes of cirrhosis due to nonalcoholic steatohepatitis and hepatitis C. *Hepatology, 43*(4), 682-689. doi:10.1002/hep.21103

Sanyal, A. J., Brunt, E. M., Kleiner, D. E., Kowdley, K. V., Chalasani, N., Lavine, J. E., . . . McCullough, A. (2011). Endpoints and clinical trial design for nonalcoholic steatohepatitis. *Hepatology, 54*(1), 344-353. doi:10.1002/hep.24376

Satapathy, S. K., & Sanyal, A. J. (2015). Epidemiology and Natural History of Nonalcoholic Fatty Liver Disease. *Semin Liver Dis, 35*(3), 221-235. doi:10.1055/s-0035-1562943

Sattar, N., Forrest, E., & Preiss, D. (2014). Non-alcoholic fatty liver disease. *BMJ, 349*, g4596. doi:10.1136/bmj.g4596

Saw, S. H. (2007). *The Population of Peninsular Malaysia*: Institute of Southeast Asian Studies.

Saxena, R., Voight, B. F., Lyssenko, V., Burtt, N. P., de Bakker, P. I., Chen, H., . . . Purcell, S. (2007). Genome-wide association analysis identifies loci for type 2

diabetes and triglyceride levels. *Science, 316*(5829), 1331-1336. doi:10.1126/science.1142358

Sayed, D., & Abdellatif, M. (2011). MicroRNAs in development and disease. *Physiol Rev, 91*(3), 827-887. doi:10.1152/physrev.00006.2010

Sayiner, M., Koenig, A., Henry, L., & Younossi, Z. M. (2016). Epidemiology of Nonalcoholic Fatty Liver Disease and Nonalcoholic Steatohepatitis in the United States and the Rest of the World. *Clin Liver Dis, 20*(2), 205-214. doi:10.1016/j.cld.2015.10.001

Schaffer, M., & Olerup, O. (2001). HLA-AB typing by polymerase-chain reaction with sequence-specific primers: more accurate, less errors, and increased resolution compared to serological typing. *Tissue Antigens, 58*(5), 299-307.

Schmittgen, T. D., & Livak, K. J. (2008). Analyzing real-time PCR data by the comparative C(T) method. *Nat Protoc, 3*(6), 1101-1108.

Schnabl, B., Kweon, Y. O., Frederick, J. P., Wang, X. F., Rippe, R. A., & Brenner, D. A. (2001). The role of Smad3 in mediating mouse hepatic stellate cell activation. *Hepatology, 34*(1), 89-100. doi:10.1053/jhep.2001.25349

Schwarzenbach, H., da Silva, A. M., Calin, G., & Pantel, K. (2015). Data Normalization Strategies for MicroRNA Quantification. *Clin Chem, 61*(11), 1333-1342. doi:10.1373/clinchem.2015.239459

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. doi:10.1053/j.gastro.2009.01.050

Shah, A. G., Lydecker, A., Murray, K., Tetri, B. N., Contos, M. J., Sanyal, A. J., & Nash Clinical Research, N. (2009). Comparison of noninvasive markers of fibrosis in patients with nonalcoholic fatty liver disease. *Clin Gastroenterol Hepatol, 7*(10), 1104-1112. doi:10.1016/j.cgh.2009.05.033

Shahba, S., Jafari Shakib, R., Jamshidi, A., Vojdanian, M., Akhtari, M., Aslani, S., . . . Mahmoudi, M. (2018). Association study of copy number variation in BMP8A gene with the risk of ankylosing spondylitis in Iranian population. *Journal of Cellular Biochemistry, 0*(0). doi:doi:10.1002/jcb.28120

Shi, Y., & Massague, J. (2003). Mechanisms of TGF-beta signaling from cell membrane to the nucleus. *Cell, 113*, 685-700.

Singer, J. B., Lewitzky, S., Leroy, E., Yang, F., Zhao, X., Klickstein, L., . . . Paulding, C. A. (2010). A genome-wide study identifies HLA alleles associated with lumiracoxib-related liver injury. *Nat Genet, 42*(8), 711-714. doi:10.1038/ng.632

Singh, A., Le, P., Lopez, R., & Alkhouri, N. (2018). The utility of noninvasive scores in assessing the prevalence of nonalcoholic fatty liver disease and advanced fibrosis in type 1 diabetic patients. *Hepatol Int, 12*(1), 37-43. doi:10.1007/s12072-017-9840-z

Singh, S., Allen, A. M., Wang, Z., Prokop, L. J., Murad, M. H., & Loomba, R. (2015). Fibrosis progression in nonalcoholic fatty liver vs nonalcoholic steatohepatitis: a systematic review and meta-analysis of paired-biopsy studies. *Clin Gastroenterol Hepatol, 13*(4), 643-654 e641-649; quiz e639-640. doi:10.1016/j.cgh.2014.04.014

Sookoian, S., & Pirola, C. J. (2011). Meta-analysis of the influence of I148M variant of patatin-like phospholipase domain containing 3 gene (PNPLA3) on the susceptibility and histological severity of nonalcoholic fatty liver disease. *Hepatology, 53*(6), 1883-1894. doi:10.1002/hep.24283

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.

Sparso, T., Andersen, G., Nielsen, T., Burgdorf, K. S., Gjesing, A. P., Nielsen, A. L., . . . Pedersen, O. (2008). The GCKR rs780094 polymorphism is associated with elevated fasting serum triacylglycerol, reduced fasting and OGTT-related insulinaemia, and reduced risk of type 2 diabetes. *Diabetologia, 51*(1), 70-75. doi:10.1007/s00125-007-0865-z

Speliotes, E. K., Yerges-Armstrong, L. M., Wu, J., Hernaez, R., Kim, L. J., Palmer, C. D., . . . Consortium, G. (2011). Genome-wide association analysis identifies variants associated with nonalcoholic fatty liver disease that have distinct effects on metabolic traits. *PLoS Genet, 7*(3), e1001324. doi:10.1371/journal.pgen.1001324

Spengler, E. K., & Loomba, R. (2015). Recommendations for Diagnosis, Referral for Liver Biopsy, and Treatment of Nonalcoholic Fatty Liver Disease and Nonalcoholic Steatohepatitis. *Mayo Clin Proc, 90*(9), 1233-1246. doi:10.1016/j.mayocp.2015.06.013

Stal, P. (2015). Liver fibrosis in non-alcoholic fatty liver disease - diagnostic challenge with prognostic significance. *World J Gastroenterol, 21*(39), 11077-11087. doi:10.3748/wjg.v21.i39.11077

Starkey Lewis, P. J., Dear, J., Platt, V., Simpson, K. J., Craig, D. G., Antoine, D. J., . . . Park, B. K. (2011). Circulating microRNAs as potential markers of human drug-induced liver injury. *Hepatology, 54*(5), 1767-1776. doi:10.1002/hep.24538

Sterling, R. K., Lissen, E., Clumeck, N., Sola, R., Correa, M. C., Montaner, J., . . . Investigators, A. C. (2006). Development of a simple noninvasive index to predict significant fibrosis in patients with HIV/HCV coinfection. *Hepatology, 43*(6), 1317-1325. doi:10.1002/hep.21178

Stranger, B. E., Forrest, M. S., Dunning, M., Ingle, C. E., Beazley, C., Thorne, N., . . . Dermitzakis, E. T. (2007). Relative impact of nucleotide and copy number variation on gene expression phenotypes. *Science, 315*(5813), 848-853. doi:10.1126/science.1136678

Sumida, Y., Nakajima, A., & Itoh, Y. (2014). Limitations of liver biopsy and non-invasive diagnostic tests for the diagnosis of nonalcoholic fatty liver

disease/nonalcoholic steatohepatitis. *World J Gastroenterol, 20*(2), 475-485. doi:10.3748/wjg.v20.i2.475

Sumida, Y., & Yoneda, M. (2018). Current and future pharmacological therapies for NAFLD/NASH. *J Gastroenterol, 53*(3), 362-376. doi:10.1007/s00535-017-1415-1

Sumida, Y., Yoneda, M., Hyogo, H., Itoh, Y., Ono, M., Fujii, H., . . . Japan Study Group of Nonalcoholic Fatty Liver, D. (2012). Validation of the FIB4 index in a Japanese nonalcoholic fatty liver disease population. *BMC Gastroenterol, 12*, 2. doi:10.1186/1471-230X-12-2

Sun, W., Cui, H., Li, N., Wei, Y., Lai, S., Yang, Y., . . . Chen, D. F. (2016). Comparison of FIB-4 index, NAFLD fibrosis score and BARD score for prediction of advanced fibrosis in adult patients with non-alcoholic fatty liver disease: A meta-analysis study. *Hepatol Res, 46*(9), 862-870. doi:10.1111/hepr.12647

Szabo, G., & Bala, S. (2013). MicroRNAs in liver disease. *Nat Rev Gastroenterol Hepatol, 10*(9), 542-552. doi:10.1038/nrgastro.2013.87

Tai, C. M., Huang, C. K., Tu, H. P., Hwang, J. C., Yeh, M. L., Huang, C. F., . . . Yu, M. L. (2016). Interactions of a PPARGC1A Variant and a PNPLA3 Variant Affect Nonalcoholic Steatohepatitis in Severely Obese Taiwanese Patients. *Medicine (Baltimore), 95*(12), e3120. doi:10.1097/MD.0000000000003120

Takaki, A., Kawai, D., & Yamamoto, K. (2013). Multiple hits, including oxidative stress, as pathogenesis and treatment target in non-alcoholic steatohepatitis (NASH). *Int J Mol Sci, 14*(10), 20704-20728. doi:10.3390/ijms141020704

Tan, C. K., Chong, H. C., Tan, E. H., & Tan, N. S. (2012). Getting 'Smad' about obesity and diabetes. *Nutr Diabetes, 2*, e29.

Tan, C. K., Leuenberger, N., Tan, M. J., Yan, Y. W., Chen, Y., Kambadur, R., . . . Tan, N. S. (2011). Smad3 deficiency in mice protects against insulin resistance and obesity induced by a high-fat diet. *Diabetes, 60*(2), 464-476. doi:10.2337/db10-0801

Tan, G. W., & Tan, L. P. (2017). High-Throughput RT-qPCR for the Analysis of Circulating MicroRNAs. *Methods Mol Biol, 1580*, 7-19. doi:10.1007/978-1-4939-6866-4_2

Tan, Y., Ge, G., Pan, T., Wen, D., & Gan, J. (2014). A pilot study of serum microRNAs panel as potential biomarkers for diagnosis of nonalcoholic fatty liver disease. *PLoS One, 9*(8), e105192. doi:10.1371/journal.pone.0105192

Tang, Y. C., & Amon, A. (2013). Gene copy-number alterations: a cost-benefit analysis. *Cell, 152*(3), 394-405. doi:10.1016/j.cell.2012.11.043

Tanwar, S., Trembling, P. M., Guha, I. N., Parkes, J., Kaye, P., Burt, A. D., . . . Rosenberg, W. M. (2013). Validation of terminal peptide of procollagen III for the detection and assessment of nonalcoholic steatohepatitis in patients with nonalcoholic fatty liver disease. *Hepatology, 57*(1), 103-111. doi:10.1002/hep.26030

Teli, M. R., James, O. F., Burt, A. D., Bennett, M. K., & Day, C. P. (1995). The natural history of nonalcoholic fatty liver: a follow-up study. *Hepatology, 22*(6), 1714-1719.

Thiele, N. D., Wirth, J. W., Steins, D., Koop, A. C., Ittrich, H., Lohse, A. W., & Kluwe, J. (2017). TIMP-1 is upregulated, but not essential in hepatic fibrogenesis and carcinogenesis in mice. *Sci Rep, 7*(1), 714. doi:10.1038/s41598-017-00671-1

Thomas, P. A. (2017). Fibroscan to Detect Significant Fibrosis in NAFLD. *Journal of Clinical and Experimental Hepatology, 7*, S40. doi:10.1016/j.jceh.2017.05.077

Tilg, H., & Moschen, A. R. (2010). Evolution of inflammation in nonalcoholic fatty liver disease: the multiple parallel hits hypothesis. *Hepatology, 52*(5), 1836-1846. doi:10.1002/hep.24001

Treeprasertsuk, S., Piyachaturawat, P., Soontornmanokul, T., Wisedopas-Klaikaew, N., Komolmit, P., & Tangkijavanich, P. (2017). Accuracy of noninvasive scoring systems to assess advanced liver fibrosis in Thai patients with nonalcoholic fatty liver disease. *Asian Biomedicine, 10*(s1), s49-s55. doi:https://doi.org/10.5372/1905-7415.1000.521

Trepo, E., Romeo, S., Zucman-Rossi, J., & Nahon, P. (2016). PNPLA3 gene in liver diseases. *J Hepatol, 65*(2), 399-412. doi:10.1016/j.jhep.2016.03.011

Turchinovich, A., Baranova, A., Drapkina, O., & Tonevitsky, A. (2018). Cell-Free Circulating Nucleic Acids as Early Biomarkers for NAFLD and NAFLD-Associated Disorders. *Front Physiol, 9*, 1256. doi:10.3389/fphys.2018.01256

Vader, W., Stepniak, D., Kooy, Y., Mearin, L., Thompson, A., van Rood, J. J., . . . Koning, F. (2003). The HLA-DQ2 gene dose effect in celiac disease is directly related to the magnitude and breadth of gluten-specific T cell responses. *Proc Natl Acad Sci U S A, 100*(21), 12390-12395. doi:10.1073/pnas.2135229100

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, 51*(4), 1209-1217. doi:10.1002/hep.23622

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, 52*(4), 1274-1280. doi:10.1002/hep.23823

Valenti, L., Alisi, A., & Nobili, V. (2012). I148M PNPLA3 variant and progressive liver disease: a new paradigm in hepatology. *Hepatology, 56*(2), 1883-1889.

Vallet-Pichard, A., Mallet, V., Nalpas, B., Verkarre, V., Nalpas, A., Dhalluin-Venier, V., . . . Pol, S. (2007). FIB-4: an inexpensive and accurate marker of fibrosis in HCV infection. comparison with liver biopsy and fibrotest. *Hepatology, 46*(1), 32-36. doi:10.1002/hep.21669

van Belle, T. L., Coppieters, K. T., & von Herrath, M. G. (2011). Type 1 diabetes: etiology, immunology, and therapeutic strategies. *Physiol Rev, 91*(1), 79-118. doi:10.1152/physrev.00003.2010

Varkonyi-Gasic, E., Wu, R., Wood, M., Walton, E. F., & Hellens, R. P. (2007). Protocol: a highly sensitive RT-PCR method for detection and quantification of microRNAs. *Plant Methods, 3*, 12. doi:10.1186/1746-4811-3-12

Vaxillaire, M., Cavalcanti-Proenca, C., Dechaume, A., Tichet, J., Marre, M., Balkau, B., . . . Group, D. S. (2008). The common P446L polymorphism in GCKR inversely modulates fasting glucose and triglyceride levels and reduces type 2 diabetes risk in the DESIR prospective general French population. *Diabetes, 57*(8), 2253-2257. doi:10.2337/db07-1807

Verrijken, A., Beckers, S., Francque, S., Hilden, H., Caron, S., Zegers, D., . . . Van Gaal, L. (2013). A gene variant of PNPLA3, but not of APOC3, is associated with histological parameters of NAFLD in an obese population. *Obesity (Silver Spring), 21*(10), 2138-2145. doi:10.1002/oby.20366

Vespasiani-Gentilucci, U., Dell'Unto, C., De Vincentis, A., Baiocchini, A., Delle Monache, M., Cecere, R., . . . Picardi, A. (2018). Combining Genetic Variants to Improve Risk Prediction for NAFLD and Its Progression to Cirrhosis: A Proof of Concept Study. *Can J Gastroenterol Hepatol, 2018*, 7564835. doi:10.1155/2018/7564835

Vickers, K. C., Palmisano, B. T., Shoucri, B. M., Shamburek, R. D., & Remaley, A. T. (2011). MicroRNAs are transported in plasma and delivered to recipient cells by high-density lipoproteins. *Nat Cell Biol, 13*(4), 423-433. doi:10.1038/ncb2210

Vilar-Gomez, E., Calzadilla-Bertot, L., Wai-Sun Wong, V., Castellanos, M., Aller-de la Fuente, R., Metwally, M., . . . Romero-Gomez, M. (2018). Fibrosis Severity as a Determinant of Cause-Specific Mortality in Patients With Advanced Nonalcoholic Fatty Liver Disease: A Multi-National Cohort Study. *Gastroenterology, 155*(2), 443-457 e417. doi:10.1053/j.gastro.2018.04.034

Vinas, O., Bataller, R., Sancho-Bru, P., Gines, P., Berenguer, C., Enrich, C., . . . Rodes, J. (2003). Human hepatic stellate cells show features of antigen-presenting cells and stimulate lymphocyte proliferation. *Hepatology, 38*(4), 919-929. doi:10.1053/jhep.2003.50392

von Roenn, N. (2018). Spotlight on Impactful Research: Increased Risk of Mortality by Fibrosis Stage in Nonalcoholic Fatty Liver Disease: Systemic Review and Meta-Analysis. *Clinical Liver Disease, 12*(2), 35-38. doi:doi:10.1002/cld.726

Wahl, K., Rosenberg, W., Vaske, B., Manns, M. P., Schulze-Osthoff, K., Bahr, M. J., & Bantel, H. (2012). Biopsy-controlled liver fibrosis staging using the enhanced liver fibrosis (ELF) score compared to transient elastography. *PLoS One, 7*(12), e51906. doi:10.1371/journal.pone.0051906

Wai, C. T., Greenson, J. K., Fontana, R. J., Kalbfleisch, J. D., Marrero, J. A., Conjeevaram, H. S., & Lok, A. S. (2003). A simple noninvasive index can predict

both significant fibrosis and cirrhosis in patients with chronic hepatitis C. *Hepatology, 38*(2), 518-526. doi:10.1053/jhep.2003.50346

Wang, J. Z., Cao, H. X., Chen, J. N., & Pan, Q. (2018). PNPLA3 rs738409 underlies treatment response in nonalcoholic fatty liver disease. *World J Clin Cases, 6*(8), 167-175. doi:10.12998/wjcc.v6.i8.167

Wang, K., Zhang, S., Weber, J., Baxter, D., & Galas, D. J. (2010). Export of microRNAs and microRNA-protective protein by mammalian cells. *Nucleic Acids Res, 38*(20), 7248-7259. doi:10.1093/nar/gkq601

Watson, N. F., Ton, T. G., Koepsell, T. D., Gersuk, V. H., & Longstreth, W. T., Jr. (2010). Does narcolepsy symptom severity vary according to HLA-DQB1*0602 allele status? *Sleep, 33*(1), 29-35.

Weber, J. A., Baxter, D. H., Zhang, S., Huang, D. Y., Huang, K. H., Lee, M. J., . . . Wang, K. (2010). The microRNA spectrum in 12 body fluids. *Clin Chem, 56*(11), 1733-1741. doi:10.1373/clinchem.2010.147405

Weyand, C. M., Hicok, K. C., Conn, D. L., & Goronzy, J. J. (1992). The influence of HLA-DRB1 genes on disease severity in rheumatoid arthritis. *Ann Intern Med, 117*(10), 801-806.

Whalley, S., Puvanachandra, P., Desai, A., & Kennedy, H. (2007). Hepatology outpatient service provision in secondary care: a study of liver disease incidence and resource costs. *Clin Med (Lond), 7*(2), 119-124.

Williams, C. D., Stengel, J., Asike, M. I., Torres, D. M., Shaw, J., Contreras, M., . . . Harrison, S. A. (2011). Prevalence of nonalcoholic fatty liver disease and nonalcoholic steatohepatitis among a largely middle-aged population utilizing ultrasound and liver biopsy: a prospective study. *Gastroenterology, 140*(1), 124-131. doi:10.1053/j.gastro.2010.09.038

Williams, T. M. (2001). Human leukocyte antigen gene polymorphism and the histocompatibility laboratory. *J Mol Diagn, 3*(3), 98-104. doi:10.1016/S1525-1578(10)60658-7

Witwer, K. W. (2015). Circulating microRNA biomarker studies: pitfalls and potential solutions. *Clin Chem, 61*(1), 56-63. doi:10.1373/clinchem.2014.221341

Wong, V. W., Vergniol, J., Wong, G. L., Foucher, J., Chan, H. L., Le Bail, B., . . . de Ledinghen, V. (2010). Diagnosis of fibrosis and cirrhosis using liver stiffness measurement in nonalcoholic fatty liver disease. *Hepatology, 51*(2), 454-462. doi:10.1002/hep.23312

Wong, V. W., Wong, G. L., Chim, A. M., Tse, A. M., Tsang, S. W., Hui, A. Y., . . . Chan, H. L. (2008). Validation of the NAFLD fibrosis score in a Chinese population with low prevalence of advanced fibrosis. *Am J Gastroenterol, 103*(7), 1682-1688. doi:10.1111/j.1572-0241.2008.01933.x

Wong, V. W., Wong, G. L., Choi, P. C., Chan, A. W., Li, M. K., Chan, H. Y., . . . Chan, H. L. (2010). Disease progression of non-alcoholic fatty liver disease: a prospective

study with paired liver biopsies at 3 years. *Gut, 59*(7), 969-974. doi:10.1136/gut.2009.205088

Xu, C., Liu, S., Fu, H., Li, S., Tie, Y., Zhu, J., . . . Zheng, X. (2010). MicroRNA-193b regulates proliferation, migration and invasion in human hepatocellular carcinoma cells. *Eur J Cancer, 46*(15), 2828-2836. doi:10.1016/j.ejca.2010.06.127

Xu, J., Wu, C., Che, X., Wang, L., Yu, D., Zhang, T., . . . Lin, D. (2011). Circulating microRNAs, miR-21, miR-122, and miR-223, in patients with hepatocellular carcinoma or chronic hepatitis. *Mol Carcinog, 50*(2), 136-142. doi:10.1002/mc.20712

Xu, R., Tao, A., Zhang, S., Deng, Y., & Chen, G. (2015). Association between patatin-like phospholipase domain containing 3 gene (PNPLA3) polymorphisms and nonalcoholic fatty liver disease: a HuGE review and meta-analysis. *Sci Rep, 5*, 9284. doi:10.1038/srep09284

Xun, Y. H., Fan, J. G., Zang, G. Q., Liu, H., Jiang, Y. M., Xiang, J., . . . Shi, J. P. (2012). Suboptimal performance of simple noninvasive tests for advanced fibrosis in Chinese patients with nonalcoholic fatty liver disease. *J Dig Dis, 13*(11), 588-595. doi:10.1111/j.1751-2980.2012.00631.x

Xun, Y. H., Guo, J. C., Lou, G. Q., Jiang, Y. M., Zhuang, Z. J., Zhu, M. F., . . . Shi, J. P. (2014). Non-alcoholic fatty liver disease (NAFLD) fibrosis score predicts 6.6-year overall mortality of Chinese patients with NAFLD. *Clin Exp Pharmacol Physiol, 41*(9), 643-649. doi:10.1111/1440-1681.12260

Yadav, H., Quijano, C., Kamaraju, A. K., Gavrilova, O., Malek, R., Chen, W., . . . Rane, S. G. (2011). Protection from obesity and diabetes by blockade of TGF-beta/Smad3 signaling. *Cell Metab, 14*(1), 67-79. doi:10.1016/j.cmet.2011.04.013

Yamada, A., Cox, M. A., Gaffney, K. A., Moreland, A., Boland, C. R., & Goel, A. (2014). Technical factors involved in the measurement of circulating microRNA biomarkers for the detection of colorectal neoplasia. *PLoS One, 9*(11), e112481. doi:10.1371/journal.pone.0112481

Yamada, H., Suzuki, K., Ichino, N., Ando, Y., Sawada, A., Osakabe, K., . . . Hashimoto, S. (2013). Associations between circulating microRNAs (miR-21, miR-34a, miR-122 and miR-451) and non-alcoholic fatty liver. *Clin Chim Acta, 424*, 99-103. doi:10.1016/j.cca.2013.05.021

Yang, H., Chen, G., Song, C., Li, D., Ma, Q., Chen, G., & Li, X. (2018). A novel index including SNPs for the screening of nonalcoholic fatty liver disease among elder Chinese: A population-based study. *Medicine (Baltimore), 97*(13), e0272. doi:10.1097/MD.0000000000010272

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. doi:10.1007/s11033-010-0212-1

You, K., Li, S. Y., Gong, J., Fang, J. H., Zhang, C., Zhang, M., . . . Zhuang, S. M. (2018). MicroRNA-125b Promotes Hepatic Stellate Cell Activation and Liver Fibrosis by Activating RhoA Signaling. *Mol Ther Nucleic Acids, 12*, 57-66. doi:10.1016/j.omtn.2018.04.016

Younossi, Z. M., Anstee, Q. M., Marietti, M., Hardy, T., Henry, L., Eslam, M., . . . Bugianesi, E. (2018). Global burden of NAFLD and NASH: trends, predictions, risk factors and prevention. *Nat Rev Gastroenterol Hepatol, 15*(1), 11-20. doi:10.1038/nrgastro.2017.109

Younossi, Z. M., Koenig, A. B., Abdelatif, D., Fazel, Y., Henry, L., & Wymer, M. (2016). Global epidemiology of nonalcoholic fatty liver disease-Meta-analytic assessment of prevalence, incidence, and outcomes. *Hepatology, 64*(1), 73-84. doi:10.1002/hep.28431

Younossi, Z. M., Stepanova, M., Afendy, M., Fang, Y., Younossi, Y., Mir, H., & Srishord, M. (2011). Changes in the prevalence of the most common causes of chronic liver diseases in the United States from 1988 to 2008. *Clin Gastroenterol Hepatol, 9*(6), 524-530 e521; quiz e560. doi:10.1016/j.cgh.2011.03.020

Zain, S. M. (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. *P Lo S One, 8*, e58538.

Zain, S. M., Mohamed, R., Cooper, D. N., Razali, R., Rampal, S., Mahadeva, S., . . . Mohamed, Z. (2014). Genome-wide analysis of copy number variation identifies candidate gene loci associated with the progression of non-alcoholic fatty liver disease. *PLoS One, 9*(4), e95604. doi:10.1371/journal.pone.0095604

Zain, S. M., Mohamed, R., Mahadeva, S., Cheah, P. L., Rampal, S., Basu, R. C., & Mohamed, Z. (2012). 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. doi:10.1007/s00439-012-1141-y

Zain, S. M., Mohamed, Z., Mahadeva, S., Cheah, P. L., Rampal, S., Chin, K. F., . . . Mohamed, R. (2013). Impact of leptin receptor gene variants on risk of non-alcoholic fatty liver disease and its interaction with adiponutrin gene. *J Gastroenterol Hepatol, 28*(5), 873-879. doi:10.1111/jgh.12104

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, 8*(3), e58538. doi:10.1371/journal.pone.0058538

Zarrei, M., MacDonald, J. R., Merico, D., & Scherer, S. W. (2015). A copy number variation map of the human genome. *Nat Rev Genet, 16*(3), 172-183. doi:10.1038/nrg3871

Zender, L., Villanueva, A., Tovar, V., Sia, D., Chiang, D. Y., & Llovet, J. M. (2010). Cancer gene discovery in hepatocellular carcinoma. *J Hepatol, 52*(6), 921-929. doi:10.1016/j.jhep.2009.12.034

Zender, L., Xue, W., Zuber, J., Semighini, C. P., Krasnitz, A., Ma, B., . . . Lowe, S. W. (2008). An oncogenomics-based in vivo RNAi screen identifies tumor suppressors in liver cancer. *Cell, 135*(5), 852-864. doi:10.1016/j.cell.2008.09.061

Zetterquist, H., Bengtsson, M., Backstrom, G., Egle-Jansson, I., Ekdahl, A. M., Grunnet, N., . . . Olerup, O. (1997). Report from the HLA class II typing by PCR-SSP Multicentre Study. *Eur J Immunogenet, 24*(3), 191-199. doi:10.1111/j.1365-2370.1997.00261.x

Zhan, Y. T., & An, W. (2010). Roles of liver innate immune cells in nonalcoholic fatty liver disease. *World J Gastroenterol, 16*(37), 4652-4660.

Zhang, F., Fan, Y. C., Mu, N. N., Zhao, J., Sun, F. K., Zhao, Z. H., . . . Wang, K. (2014). Exportin 4 gene expression and DNA promoter methylation status in chronic hepatitis B virus infection. *J Viral Hepat, 21*(4), 241-250. doi:10.1111/jvh.12136

Zhang, H., Li, Q. Y., Guo, Z. Z., Guan, Y., Du, J., Lu, Y. Y., . . . Su, S. B. (2012). Serum levels of microRNAs can specifically predict liver injury of chronic hepatitis B. *World J Gastroenterol, 18*(37), 5188-5196. doi:10.3748/wjg.v18.i37.5188

Zhang, H., Wei, S., Ning, S., Jie, Y., Ru, Y., & Gu, Y. (2013). Evaluation of TGFbeta, XPO4, elF5A2 and ANGPTL4 as biomarkers in HCC. *Exp Ther Med, 5*(1), 119-127. doi:10.3892/etm.2012.750

Zhang, Q. Q., & Lu, L. G. (2015). Nonalcoholic Fatty Liver Disease: Dyslipidemia, Risk for Cardiovascular Complications, and Treatment Strategy. *J Clin Transl Hepatol, 3*(1), 78-84. doi:10.14218/JCTH.2014.00037

Zhang, Y., Jia, Y., Zheng, R., Guo, Y., Wang, Y., Guo, H., . . . Sun, S. (2010). Plasma microRNA-122 as a biomarker for viral-, alcohol-, and chemical-related hepatic diseases. *Clin Chem, 56*(12), 1830-1838. doi:10.1373/clinchem.2010.147850

Zhou, H. T., Shi, Z. Z., Zhou, Z. X., Jiang, Y. Y., Hao, J. J., Zhang, T. T., . . . Zhang, Y. (2013). Genomic changes in rectal adenocarcinoma associated with liver metastasis. *Cancer Biomark, 13*(4), 281-288. doi:10.3233/CBM-130351

Zhou, J., Lemos, B., Dopman, E. B., & Hartl, D. L. (2011a). Copy-number variation: the balance between gene dosage and expression in Drosophila melanogaster. *Genome Biol Evol, 3*, 1014-1024.

LIST OF PUBLICATIONS AND PAPERS PRESENTED

Publications

Tan, H. L., Zain, S. M., Mohamed, R., Rampal, S., Chin, K. F., Basu, R. C., Cheah, P. L., Mahadeva, S., & Mohamed, Z. (2014). Association of glucokinase regulatory gene polymorphisms with risk and severity of non-alcoholic fatty liver disease: an interaction study with adiponutrin gene. *J Gastroenterol, 49*(6), 1056-1064.

Impact factor (2017): 3.3

Zain, S. M., Mohamed, Z., Pirmohamed, M., Tan, H. L., Alshawsh, M. A., Mahadeva, S., Chan, W. K., Mustapha, N. R. N., Mohamed, R. (2015). Copy number variation in exportin-4 (*XPO4*) gene and its association with histological severity of non-alcoholic fatty liver disease. *Sci Rep*. *5*, 13306.

Impact factor (2017): 4.122

Tan, H. L., Mohamed, R., Mohamed, Z., & Zain, S. M. (2016). Phosphatidylethanolamine N-methyltransferase gene rs7946 polymorphism plays a role in risk of nonalcoholic fatty liver disease: evidence from meta-analysis. *Pharmacogenet Genomics, 26*(2), 88-95.

Impact factor (2017): 2.25

Zain, S. M., Mohamed, Z., Mahadeva, S., Cheah, P. L., Rampal, S., Chin, K. F., et al. (2013). Impact of leptin receptor gene variants on risk of non-alcoholic fatty liver disease and its interaction with adiponutrin gene. *J Gastroenterol Hepatol, 28*(5), 873-879.

Impact factor (2017):3.483

Conference

Hwa-Li Tan, Zahurin Mohamed, Shamsul Mohd Zain, Sanjiv Mahadeva, Nik Raihan Nik Mustapha, Wah-Kheong Chan, Rosmawati Mohamed. (2014, 22nd -24th August). Mirnome analysis reveals differential expression of microRNA in nonalcoholic steatohepatitis (NASH). Physiology International Scientific Congress 2014, Putra World Trade Centre, Kuala Lumpur, Malaysia. (Poster Presentation)

Hwa-Li Tan, Zahurin Mohamed, Shamsul Mohd Zain, Sanjiv Mahadeva, Nik Raihan Nik Mustapha, Wah-Kheong Chan, Rosmawati Mohamed. (2015, 11th – 13th March). Diagnostic Value of Serum microRNAs in Non-Alcoholic Steatohepatitis. Golden Helix Symposium 2015, Nexus Bangsar South, Kuala Lumpur, Malaysia. (Poster Presentation)

*Abstract of poster presentation was published in : 2015 golden helix symposium-next generation pharmacogenomics. March 11-13, 2015, kuala lumpur, malaysia: abstracts. (2015). *Public Health Genomics, 18 Suppl 1*, 1-51.

Impact factor: 1.5

Manuscript in preparation

Serum miRNAs in non-alcoholic fatty liver disease: Distinguishing steatohepatitis and significant fibrosis

Association of HLA-DQB1*06 in conferring a lower risk of significant fibrosis

Non-invasive scoring system in diagnosis of advanced fibrosis in non-alcoholic fatty liver disease (NAFLD)