ANALYSIS OF THE ASSOCIATION BETWEEN SLCO1B1 GENETIC POLYMORPHISM (rs4149056) AND LIPID PROFILE AMONG STATIN USERS OF HUSM, KELANTAN

NUR FARAH IZZATI BINTI YUSRI

UNIVERSITI SAINS MALAYSIA

2022

ANALYSIS OF THE ASSOCIATION BETWEEN SLCO1B1 GENETIC POLYMORPHISM (rs4149056) AND LIPID PROFILE AMONG STATIN USERS OF HUSM, KELANTAN

by

NUR FARAH IZZATI BINTI YUSRI

Dissertation submitted in partial fulfillment of
the requirements of the degree of
Master of Science (Biomedicine) Mixed Mode

SEPTEMBER 2022

ACKNOWLEDGEMENT

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

The highest gratitude goes to Allah S.W.T for giving me this opportunity to finish this thesis. During one semester of research, I have worked with a great number of people for research study and the making of the thesis. It is pleasure to convey my thanks and gratitude to all of them.

I would like to express my gratitude to my supervisor, Dr Nur Salwani Bakar, lecturer from School of Health Sciences, USM, for her supervision and guidance in finishing this research study.

I would like to express my appreciation to my Co-supervisors Assoc. Prof. Dr. Sarina Sulung and Assoc. Prof. Dr. Edinur Hisham Atan for their valuable assistance in providing facilities and support for completing this project.

It is pleasure to extent my sincere thanks Mr. Mohamad Ros Sidek for his guidance in the completion of the studies in Laboratory of Human Genome Center. My endless gratitude to my friends for their support and encouragements in many ways.

Finally, I would like to express my gratitude to my parents Mr Yusri Mat salleh and Mrs Zaikiki Abu Bakar for their support and encouragements throughout study period.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS, ABBREVIATIONS AND ACRONYMS

%	Percentage
>	More than
<	Less than
^{0}C	Degree celcius
µL	Microliter
ASCVD	Atherosclerotic cardiovascular disease
CNV	Copy number polymorphism
CVD	Cardiovascular diseases
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
EDTA	Ethylenediaminetetraacetic acid
HDL-c	High density lipoprotein cholesterol
HMG-CoA	3-hydroxy-3-methylglutaryl coenzyme A
HWE	Hardy- weimberg equilibrium
LDL-c	Low density lipoprotein cholesterol
MAF	Minor allele frequencies
PCSK9	Proprotein convertase subtilisin/kexin type 9
RNA	Ribonucleic acids
SD	Standard deviation
SLCO1B1	Solute carrier organic anion transporter family member 1B1
SNP	Single nucleotide polymorphism
STR	Short tandem repeat polymorphism

TAE	Tris-acetate EDTA
TC	Total cholestrol
TG	Triglycerides
UV	Ultraviolet
VLDL	Very low density lipoprotein

ANALISIS PERKAITAN ANTARA POLIMORFIK DALAM GEN *SLCO1B1* (rs4149056) DAN PFOFAIL LIPID DALAM KALANGAN PENGGUNA STATIN DARI HUSM, KELANTAN

ABSTRAK

Statin adalah ubat yang telah terbukti dapat merawat penyakit hiperlipidemia. Walaupun begitu, beberapa kejadian mengenai kurang keberkesanan ubat tersebut telah dilaporkan. Polimorfisme nukleotida tunggal (SNP) dalam *SLCO1B1* rs41419056 telah dilaporkan boleh mengubah keberkesanan statin. Sehingga kini, tiada kajian genetik kaitan hubungan di Malaysia mengenai pengaruh SNPs tersebut terhadap keberkesanan statin. Oleh itu, tujuan kajian ini dilakukan adalah untuk mengkaji perkaitan antara *SLCO1B1* rs4149056 dan pengguna statin di HUSM, Kelantan. Terdapat 72 orang pengguna statin telah terlibat dalam kajian ini. Semua maklumat berkenaan demografi dan klinikal pesakit telah diperolehi melalui pemeriksaan rekod perubatan pesakit. Semua sampel DNA yang digunakan di dalam kajian ini telah diperolehi daripada penyelidik sebelum ini. Sampel DNA pengguna statin telah melalui proses genotip melalui kaedah ARMS-PCR diikuti dengan elektroforesis gel. Keputusan menunjukkan tiada perbezaan bagi demografi dan faktor klinikal yang dianalisa di antara kumpulan 1 (mereka yg mencapai tahap LDL-c <2.60 mmol/L) dan kumpulan 2 (mereka yang mencapai tahap LDL-c >2.60 mmol/L). Selain itu, nilai kekerapan alel minor (MAF) yang diperolehi adalah 0.23. Namun demikian, tiada statistik yang signifikan di antara pemboleh ubah (umur, kaum, jantina, jenis statin, ubat bersamaan, suplemen, genotip, aktiviti fizikal,diet dan status merokok) dan pencapaian tahap LDL-c <2.60 mmol/L. Kesimpulannya, *SLCO1B1* rs4149056

polimorfisme tidak memberi kesan ke atas perubahan profil lipid pada 72 pengguna statin hiperlipidemik.

ANALYSIS OF THE ASSOCIATION BETWEEN SLCO1B1 GENETIC POLYMORPHISM (rs4149056) AND LIPID PROFILE AMONG STATIN USERS HUSM, KELANTAN

ABSTRACT

Statin is a medication that was proven to treat hyperlipidaemia effectively. Despite its well-known efficacy in statin users, the cases of its effectiveness have been reported. Single nucleotide polymorphism of rs4149056 in *SLCO1B1* gene has been reported to alter the efficacy of statin. Until now there is no genetic association study among Malaysian population has been reported pertaining the influence of SNP on the efficacy of statin. Thus, the aim of this study is to investigate the association between *SLCO1B1* rs4149056, patient's demographic profiles and other clinical features among statin's users from HUSM, Kelantan. Information regarding the demographic and clinical features of statin's users obtained through examination of patient's medical record. The extracted DNAs for this study were obtained from previous researcher. The genotyping of selected DNA sample of statin's users was performed by using ARMS-PCR technique followed by gel electrophoresis. The results indicated that, there were no different in demographic profiles and clinical features between group 1 (those who achieved LDL-c level <2.60 mmol/L) and group 2 (those who achieved LDL-c level >2.60 mmol/L) except for concomitant drugs. The minor allele frequency (MAF) of the SNP value obtained was 0.23. However, there is no variables (age, race, gender, type of statin, concomitant drugs, supplement, genotype, physical activity, diet and smoking status) that are statistically significant associated with the achievement of LDL-c goal of <2.60 mmol/L. In conclusion, the *SLCO1B1*

rs41419056 polymorphism had no effect on lipid level changes in 72 hyperlipidaemic statin users.

CHAPTER 1

INTRODUCTION

1.1 Introduction

Hyperlipidaemia is a condition where the lipids level in the bloodstream is elevated and these lipids including fats, fatty acids, cholesterol, cholesterol esters, phospholipids, and triglycerides (Jain *et al.*, 2007). Furthermore, the elevation of lipids level can increase the risk of cardiovascular diseases (Xie *et al*., 2012). According to the Yen *et al*. (2017), the prevalence of hyperlipidaemia increases from 26.9% in 2006 to 35.1% in 2011 and there are 6.2 million Malaysian adults living with hyperlipidaemia.

To overcome the elevation level of lipids, statin medication can be applied as it has cholesterol-lowering effect by inhibiting the conversion of 3-hydroxy-3-methylglutaryl coenzyme A to mevalonate for the synthesis of cholesterol (Adams *et al.*, 2015). There are several types of statins including lovastatin, pravastatin, simvastatin, atorvastatin, cerivastatin, fluvastatin, pravastatin, pitavastatin and rosuvastatin (Schachter, 2005). The effectiveness of statin in the prevention of acute cardiovascular diseases and ischemic heart diseases have been proven by several studies as statins can decrease low-density lipoprotein (LDL) cholesterol, decrease triglycerides level and increase high-density lipoprotein (HDL) cholesterol level (Beltowski *et al*., 2009).

Statin have been the best-selling drug class not only in the US but also in the world as it has been proven to have a favourable safety profile and beneficial in the treatment of cardiovascular diseases (Golomb & Evans, 2008). However, there is large variability in lipid lowering response to statins. This can be due to environmental,

dietary factors and single- nucleotide polymorphisms (SNPs), or other DNA sequence variations, in genes encode for drug metabolizing enzymes or transporters (Romaine *et al.*, 2010).

The present of certain SNPs could greatly affect the density and activity of the transporters that will lead to changes in the rate the handling the drugs by hepatocytes with subsequent effect on the pharmacokinetic parameters of medication (Saber-ayad *et al.*, 2018). According to the study conducted by Zhang *et al.* (2007), SNPs of 521T > C, Val174Ala, rs4149056, influence the total cholesterol-lowering efficacy of pravastatin treatment in Chinese coronary heart diseases patients. Besides that, the presence of the rs4149056 SNP was related to the significantly less of LDL cholesterol lowering response to pravastatin in elderly (Akao *et al.*, 2012).

1.2 The Problem Statements and Significance of the Study

The leading cause of mortality globally is cardiovascular disease where 17.7 million of people died annually due to the cardiovascular diseases such as stroke and coronary artery disease (Sazlina *et al*., 2020). High cholesterol level is associated with the high risk of cardiovascular diseases. In order to reduce the cholesterol level in blood stream, statin medication has been prescribed and widely used globally. However, there are several cases where the statin medication does not produce desire effect and cause adverse side effects. This is due to the variability of response toward the statin where genetic variation can influence efficacy of drugs when being prescribed.

In this study, the association between *SLCO1B1* genetic polymorphism rs4149056 and lipid profiles among outpatient statin's users will be investigated. This study should be conducted in order to identify the prevalence and the factors that can affect statin

effectiveness in statin users. Furthermore, duet to the lack of data pertaining the SNP in Malaysia, this study showed to be beneficial for the future studies especially in Malaysia. Genetic polymorphism in *SLCO1B1* can be the potential genetic marker that is important to measure the effectiveness of statin therapy. This study can provide the evidence regarding the genetic variation that can influence the statin efficacy and might contribute in the developing of more individualized targeted therapies, also called personalized medicine, to ensure the effectiveness and safety of statin treatment.

1.3 Research Questions

1. What factors that significantly contribute to the lower efficacy of statin?
2. What is the genotype frequencies of the *SLCO1B1* rs4149056 among the statins' users in HUSM?
3. Is there any association between *SLCO1B1* rs4149056 and the change in lipid profile?

1.4 Hypothesis

There is association between *SLCO1B1* genetic polymorphism rs4149056 and lipid profiles among outpatient statin's user in HUSM, Kelantan.

1.5 Research Objectives

1.5.1 General objectives

To study the association between *SLCO1B1* genetic polymorphism rs4149056 and lipid profiles among outpatient statin user in HUSM, Kelantan.

1.5.2 Specific objectives

1. To genotype *SLCO1B1* gene of rs4149056 in 72 outpatient statin users in Hospital USM.
2. To determine the genotype frequencies of *SLCO1B1* gene rs4149056 and its association with changes in lipids profile among outpatient statin users.
3. To determine the extent by which other factors (genetic, type of statin, lipid profiles, concomitant drugs, and patient demographic) contribute to the independent factor(s) associated with the achievement of LDL-c goal of < 2.6 mmol/L.

CHAPTER 2

LITERATURE REVIEW

2.1 Hyperlipidaemia

Globally, hyperlipidaemia become the main cause of cardiovascular related morbidity and mortality (Weinreich & Frishman, 2014). Hyperlipidaemia can be defined as disorders caused by the elevation of lipids in the blood vessels and these lipids including triglycerides, phospholipids, cholesterol, and cholesterol esters (Nouh *et al*., 2019). In Malaysia, the prevalence of hyperlipidaemia increases from 26.9% in 2006 to 35.1% in 2011 and there are 6.2 million Malaysian adults living with hyperlipidaemia (Yen *et al*., 2017). Hyperlipidaemia is the main cause of atherosclerosis and related strongly to ischemic heart diseases (Shattat, 2014).

Hyperlipidaemia can be divided into two category which are familial hyperlipidaemia and acquired hyperlipidaemia. Individual that categorized in the familial hyperlipidaemia inherit diseases through birth while for acquired hyperlipidaemia, the diseases obtain from the unhealthy lifestyle (Hill & Bordoni, 2022).

2.1.1 Pathogenesis of hyperlipidaemia

Cholesterol synthesis start with the addition of two-carbon atom to acetyl-CoA and it occur in the cell cytoplasm while the fatty acid oxidation occur in the mitochondria. Liver and adipose tissue are the main sites of cholesterol synthesis. Hyperlipidaemias occur once adipose lipolysis and hepatic VLDL synthesis exceeds the rate of clearance of plasma VLDL. Unhealthy lifestyle and diets in environmentally and genetically predisposed individuals increase the risk of hyperlipidaemia. Impaired adiposeness of

lipids during positive caloric balance might increase lipid deposited in non-adipose tissue organ that resulting in lipotoxicity (Nouh *et al*., 2019).

2.1.2 Risk factor of hyperlipidaemia

Risk factors of hyperlipidaemia can be categorized into two group which are modifiable risk factors and nonmodifiable risks factors in which modifiable risk factors including medication, nutrition and physical inactivity while nonmodifiable risks factors included age, gender, genetic and chronic diseases (Nouh *et al*., 2019).

As mentioned above, age is a risk factor of hyperlipidaemia in a group of nonmodifiable risk factors. According to the study by Rosada *et al*. (2020), the prevalence of hyperlipidaemia increases with age as the levels of total cholesterol in elder group (64%) higher than the young group (23.7%) resulting from loss of hepatic LDL receptor associated with age, higher body mass index, larger waist circumference and decrease in levels of sec hormone. Not only an old age individual has high risk of hyperlipidaemia, children also has high tendency of hyperlipidaemia due to the genetic disorders resulting from the mutation of single or multiple genes that cause overproduction and defective clearance of the cholesterol, TG and LDL (Nouh et al., 2019). Female has high risk of hyperlipidaemia due to the drastic escalation of biological aging and menopause-related endocrines change. In addition, age is associated with the increasing of hyperlipidaemia in women from 14.9% in third decade to 56.4% by the age of 60 (Cho *et al*., 2020).

In modifiable risk factors, physical inactivity, medication and nutrition are included. Physical inactivity and unhealthy diet can lead to the weight gain and obesity that can contribute to the high level of cholesterol in bloodstream. In addition, consumes food that contain high calories resulting in the stored of excess calories in a body as a fat.

Medication such as thiazides, retinoids, estrogens and glucocorticoids can also increase the risk of hyperlipidaemia (Nouh *et al*., 2019).

2.1.3 Diagnosis and Treatment

In order to decrease the level of lipids in bloodstream, The National Cholesterol Education Program has established guidelines for diagnosis and treatment of hyperlipidaemia based on individuals lo-density lipoprotein (LDL) concentration and their cardiac risk factors as epidemiological studies shows the relationship between cholesterol concentration and risk of CVD (Weinreich & Frishman, 2014).

The American Heart Association urge its citizen to evaluate their cholesterol level especially for individuals who has family history of stroke, high blood pressure or heart diseases. Lipid profile is the blood test that usually being conducted in order to screen the lipids level during routine health care evaluation (Nelson, 2013).

The treatment of hyperlipidaemia can be divided into two ways which are medication and non-medication treatment. Statin is a medication that usually being prescribed as it can lower the LDL-c level in bloodstream and reduce the morbidity and mortality of CVD (Nelson, 2013). Besides that, Niacin medication also can be used as it can increase the level of HDL-c and Fibrates can decrease the level of triglyceride in bloodstream (Nouh *et al*., 2013). In non-medication treatment, patients should improve their lifestyle by consumes less food that contain high calories, active physically, stop smoking and reduce weight (Nelson, 2013).

LDL-c level for individuals with high risk oh hyperlipidaemia should be lower than 70 mg/dL after treatment. While for diabetes mellitus patients, LDL-c goal level should be lower than 100 mg/dL and patient with high risks of CVD, LDL-c goal level should be below 70mg/dL. Besides that, increase the HDL-c level minimally 40 mg/dL also

essential for the treatment of hyperlipidaemia. Triglyceride goal level should be below 150 mg/dL (Nouh *et al*., 2019).

2.2 Lipid Profile

According to the Department of Statistics Malaysia (https://www.dosm.gov.my/v1_/), ischaemic heart diseases become the main cause of medically certified deaths in 2020 about 17% followed by pneumonia (11.4%), and cerebrovascular disease (16.3%). Cardiovascular diseases (CVD) become the leading cause of death globally and it is associated with unhealthy diet, physical inactivity and obesity that resulting in hyperlipidaemia and hypertension (Sazlina *et al*., 2020).

Nigam (2011) state that, lipid profile is a test that are conducted for cardiovascular risk prediction and now become the routine test. This test consists of four parameters which are total cholesterol, low-density lipoprotein (LDL) cholesterol, high-density lipoprotein (HDL) cholesterol and triglycerides. Elevation of LDL and triglycerides, and low of HDL are related with obesity (Sane *et al*., 2019). Obesity can be defined as the unbalanced fat storage in the body that might affect health (Omer, 2020). Table below shows the normal range of lipid profile.

Table 2.1 Normal range of lipid profile

Variable assessed	**Normal range (mg/dl)**
Cholesterol	<200
Triglyceride	<150
HDL	>60
LDL	<130
VLDL	<30

2.2.1 LDL-c

Low-density lipoprotein cholesterol (LDL-c) delivers fat molecules around the body through blood circulation. LDL is produced by the metabolization of very low-density lipoprotein (VLDL) to IDL by lipoprotein lipase (LPL). Then, IDL is converted to LDL by hepatic triglyceride lipase (HTGL). LDL-c is removed from the blood vessels by LDL receptors into the hepatocyte. Since LDL receptors are necessary for the uptake of LDL-c in blood, the genetic decreases in LDL receptors can cause the incapability of hepatocytes in absorbing LDL-c, thus can lead to the increase of LDL-c in blood. In heterozygous mutation, there are some LDL receptor present, thus the level of LDL-c present in the blood usually 300 mg/dL. For homozygous mutation, resulting in absence of LDL receptor that can lead to the increase of LDL-c level about 1000 mg/dL (Pirahanci *et al*., 2021).

2.2.2 HDL-c

High-density lipoprotein cholesterol (HDL-c) plays a crucial role in reverse cholesterol transport (RCT), in which the excess cholesterol is removed from the peripheral vessels and transported into the liver for disposal. HDL-c consists of apolipoproteins which are ApoA-I (70%) and ApoA-II (20%) that are important for normal HDL-c biosynthesis. ApoA-I synthesis in both liver and intestines while ApoA-II synthesis in liver. There is a clear reverse relationship between HDL-c levels and risks of coronary heart disease (CHD). In addition, it has been estimated that an increase of 1 mg/dL of HDL-c, can reduce the risk of CHD by about 3% in women and 2% in men (Kosmas *et al*., 2018).

2.2.3 Triglycerides (TG)

Triglycerides are the main elements of vegetables and body fat in humans and animals. TG present in the blood for the bidirectional transference of adipose fat and blood glucose from the liver. Besides that, they are also the main components of human skin oils. Triglycerides are composed of three fatty acids esterified with glycerol backbone. Besides LDL-c, TG is also associated with the risk of ASCVD. Mendelian disorders such as familial chylomicronemia syndrome caused by several LPL-associated genes, including the LPL gene itself, and type 3 dyslipidaemia caused by apolipoprotein E (APOE) gene mutations elevate triglycerides levels and increase the risk of ASCVD (Tada *et al*., 2020).

2.2.4 Cholesterol

Cholesterol consists of a hydrocarbon tail, a central sterol nucleus made of four hydrocarbon rings, and a hydroxyl group. There are several biological functions of cholesterol including assisting in the synthesis of steroids and vitamin D, acts as precursor to bile acids, play a major role in maintaining rigidity and fluidity of the cellular membrane, and crucial for successful of cellular homeostasis. Even though, cholesterol is important for the functioning of the cell, the elevation of cholesterol levels can cause serious problem as well as genetic disorder implications including familial hypercholesterolemia, and atherosclerosis (Craig *et al*., 2018).

2.3 Statin

Hyperlipidaemia is a risk factor of cardiovascular disease (CVD) thus inhibition cholesterol therapy is important in managing and preventing of CVD (Shuhaili *et al*., 2017). The effectiveness of lipid-lowering drug can treat hyperlipidaemia and reduce the incidence and mortality of CVD. There are various lipid-lowering drugs including statins, proprotein convertase subtilisin/kexin type 9 (PCSK9) inhibitor, cholesterol absorption inhibitors, the bile acid sequestrant, nicotinic acid, and fibrates (Wang, 2020).

Mevalonate pathways involve in the conversion of mevalonate into cholesterol, thus in order to reduce the formation of cholesterol statin medication is prescribed as it can inhibit the formation of cholesterol (Buhaescu & Izzedine, 2007). Statin is widely used globally in treating hyperlipidaemia and there are several types of statins including Atorvastatin (Lipitor and Torvast), simvastatin (Zocor and Lipex), lovastatin (Mevacor, Altocor and Altoprev), pitavastatin (Livalo and Pitava), rosuvastatin (Crestor), fluvastatin (Lescol), and pravastatin (Pravachol, Lipostat and Selektine) (Shuhaili *et al*., 2017). The effectiveness of Statin is proven based on studies by Feingold (2021) state that statin can lower LDL-c level about 60% and reduce triglycerides level. Figure 2.1 below shows the structure of several types of statins.

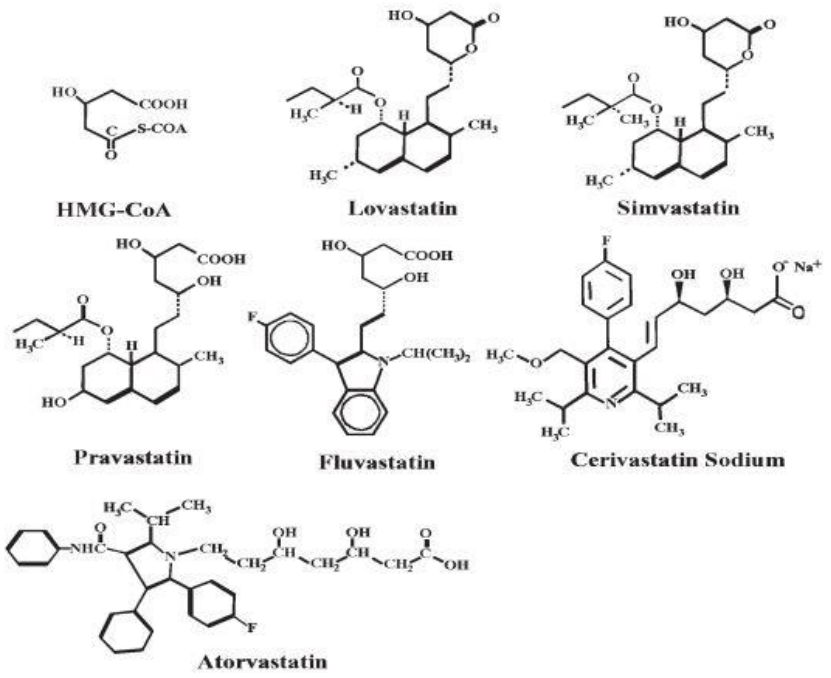

Figure 2.1 Types of statins (Shuhaili *et al*., 2017)

2.3.1 Mechanism of action

Statin works by competitively inhibit HMG-CoA reductase by blocking the conversion of HMG-CoA to mevalonic acid, HMG-CoA reductase is the first enzyme of the mevalonate pathway. Then, statin activate the production of microsomal HMG-CoA reductase and cell surface low-density lipoprotein (LDL) receptors as cholesterol synthesis in the liver is being interrupted resulting in the increase of LDL clearance form blood vessel and LDL-c level in blood may decrease about 20% to 55% (Buhaescu & Izzedine, 2007).

Statin blocking the conversion of HMG-CoA to mevalonate by binds to the active site within the reductase and prevent the binding of HMG-CoA. Since statin has higher affinity compared to HMG-CoA, it can prevent the binding of HMG-CoA (Davies *et al*., 2016). The mevalonate pathway can be summarized in figure 2.2.

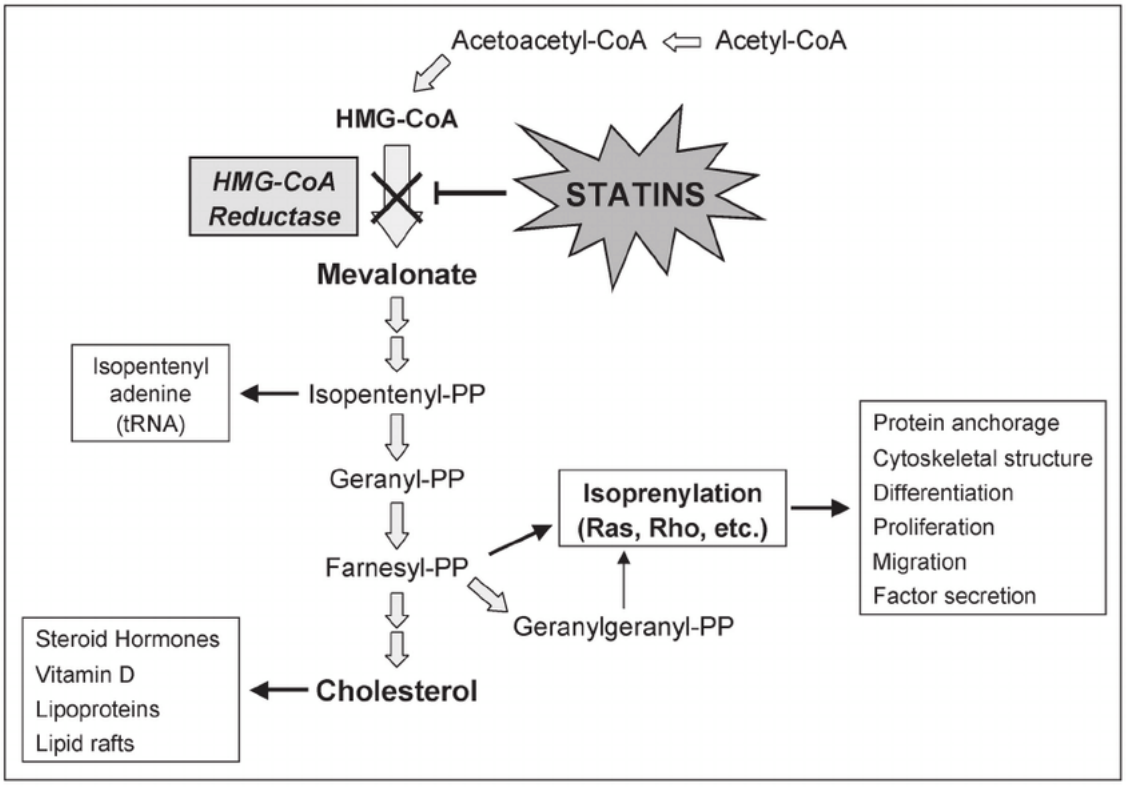

Figure 2.2 The mevalonate pathway (Hedwich, 2007).

2.3.2 Efficacy of statin

There are several studies that prove the efficacy of statin in lipid-lowering effect and how the medication itself reduces the risk of CVD. The studies conduct by Adhyaru & Jaconson (2013), state that statin has been proven in the risk reduction of atherosclerotic cardiovascular diseases (ASCVD) in the primary and secondary prevention in which each 1mmol/l reduction in LDL-cholesterol (LDL-C) levels in the plasma can reduces the risk of ASCVD events by 22% after 5 years.

Besides that, studies performed by Garcia-gil *et al*. (n.d.) also shows the positive results on the efficacy of statin as statin was effective in preventing ASCVD across the range of coronary risk in patients with high adherence to therapy (MPR>70%). In addition, statin drugs decreased ASCVD risk by 30% and 26% in individuals with

cardiovascular risk of ≥7.5-9.9% and ≥10%-19.9%, respectively and decreased all-cause mortality by 27% in those at highest risk.

2.3.3 Development of side effects

Even though statin medication has been proven in the lowering of cholesterol level, however there are some adverse effects that has been reported occur to the statin user. The most common adverse effect that has been reported is muscle adverse effect including muscle pain, fatigue and weakness as well as rhabdomyolysis (Golomb % Evans, 2008). Many muscular complaints, especially exercise-induced, may occur in the 10% of treated patients. (Beltowski *et al*., 2009).

Besides that, Myositis also one of adverse effect of statin medication in which the inflammation of the muscles occurs. The risk of myositis increases when certain other medications are taken with statins such as fibrate. Fibrates is another drug that reduce cholesterol. The risk of myositis increases greatly compared to someone who takes a statin alone. Besides muscle adverse effect, there are several side effects that commonly occur which is less chronic that muscle adverse effect. These side effects including headache, difficulty in sleeping, flushing of the skin, drowsiness, dizziness and nausea or vomiting (Bruce, 2020).

2.4 Genetic Polymorphism

Genetic polymorphism can be described as the heterozygous DNA variation that present in more than 1% of population (Stankiewicz and Lupski, 2012). Genetic polymorphism can happen in any coding and non-coding region of DNA (Singh & Kulathinal, 2013). Polymorphism of DNA may cause by chance processes or external

agents including viruses or radiation and it can be inherited from parent to child (Ismail & Essawi, 2012).

According to Ismail and Essawi (2012), there are several types of DNA polymorphism including tandem repeat polymorphism, short tandem repeat polymorphism (STR), copy-number polymorphism (CNV) and single nucleotide polymorphism (SNP).

2.4.1 Single nucleotide polymorphism (SNP)

As reported by Stankiewicz and Lupski (2012), SNPs has been discovered in the initial phase of DNA sequencing analysis and showed that human genome is vary by single nucleotide changes. In simple way, SNP can be defined as variation of single nucleotide that occur at specific genomic position (Zou *et al*., 2020). SNPs are the most common type of variant in human genome as it occurs every 100-300 base pairs with an allele frequency more than 1% and at least 10 million of SNPs present in the genome (Robert & Pelletier, 2018). Furthermore, SNPs may result in the changes of genomic sequence, either in the coding, intergenic, or noncoding region (Vallejos-Vidal *et al*., 2020).

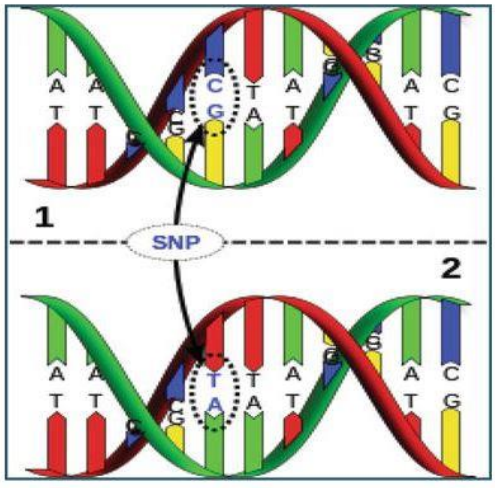

Figure 2.3 DNA sequences contained SNP (Sukhumsirichart, 2018)

Based on the figure 2.3, Individual 1 has C nucleotide in his DNA sequence which is similar with the majority group of population while the DNA sequence in individual 2 has T nucleotide which is similar with minority group of population. It can be concluded that there is present of SNP at the specific position between allele C and T (Sukhumsirichart, 2018).

SNP may occur in coding region, non-coding region or intergenic region and may affect the gene products. In coding region, there are two type of SNPs which are synonymous and nonsynonymous. The synonymous SNP will not alter the amino acid sequence and effect of the protein while nonsynonymous may alter the amino acid sequence thus effect the protein. The nonsynonymous can either be missense or nonsense. Both of SNP may lead to the non-function protein product. Besides that, SNP that occur in non-coding region or in the intergenic region may resulting in gene splicing, transcription factor binding, messenger RNA degradation or sequence of non-coding RNA (Sukhumsirichat, 2018).

SNPs are very beneficial especially in human studies as it can be considered important tool and useful biomarkers in disease-association studies due to their common frequency, ease of analysis, and low genotyping costs (Srinivasan *et al*., 2016). For example, in pathogenic studies as SNP may influence the diseases progression or immune respond toward the pathogen (Vallejos-Vidal *et al*., 2020).

2.4.2 Worldwide prevalence of *SLCO1B*1 rs4149056

The SNP rs4149056 located in the chromosomes 12 with the substitution of Thymine to Cytosine (T>C). Based on Ensemble genome browser (https://asia.ensembl.org/index.html), the average frequency of mutant C allele in worldwide about 9%. Europe population shows the highest frequency at 16% followed

by American and East Asian about 13% and 12% respectively. However, African population shows the lowest frequency of mutant allele about 1% (Figure 2.4).

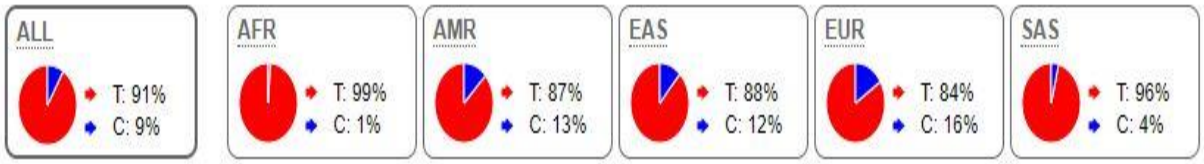

Figure 2.4 The genotype frequency of *SLCO1B1* rs4149056 (https://asia.ensembl.org/index.html)

SLCO1B1 (rs4149056) polymorphism has been genotyped in various population, however it has not been genotyped in Malaysia. Thus, there is no evidence of study in any source. The table below shows the frequency of genotype in different populations based on ensemble genome browser (https://asia.ensembl.org/index.html).

Table 2.2: The frequency genotype of *SLCO1B1* rs4149056 in different population.

Genotype	TT (%)	TC (%)	CC (%)
African			
Barbados (African Caribbean)	95.8	4.2	0
Southwest US (African Ancestry)	88.5	9.8	1.6
Nigeria (Esan)	100	0	0
Western Division (Gambian)	100	0	0
Webuye, Kenya (Luhya)	96	4	0
Sierra Leone (Mende)	100	0	0
Ibadan, Nigeria (Yoruba)	98.1	1.9	0
American			
Medellin, Colombia (Colombian)	67	29.8	3.2
Los Angeles, California (Mexican Ancestry)	84.4	15.6	0
Lima, Peru (Peruvian)	74.1	23.5	2.4
Puerto Rico (Puerto Rican)	76.9	22.1	1
European			

Utah Residents (Northern and Western European ancestry)	71.7	27.3	1
Finland (Finnish)	66.7	30.3	3
England and Scotland (British)	72.5	26.4	1.1
Spain (Iberian)	77.6	21.5	0.9
Italy (Toscani)	60.7	35.5	3.7
East Asian			
Xishuangbanna, China (Chinese Dai)	73.1	25.8	1.1
Beijing, China (Han Chinese)	74.8	23.3	1.9
China (Southern Han Chinese)	79	18.1	2.9
Tokyo, Japan (Japanese)	78.8	18.3	2.9
Ho Chi Minh City, Vietnam (Kinh)	80.8	18.2	1
South Asian			
Bangladesh (Bengali)	89.5	10.5	0
Gujarati Indian (residue in Houston, TX)	96.1	3.9	0
Indian Telugu (residue in UK)	88.2	10.8	1
Lahore, Pakistan (Punjabi)	93.8	5.2	1
Sri Lankan Tamil (residue in UK)	91.2	8.8	0

2.5 ARMS-PCR

Polymerase chain reaction (PCR) is the most common procedure for SNP genotyping. Nowadays, there are many PCR procedure that are available and it is important to choose the most suitable one for the successful of the study. One of the PCR procedures that are simple and economical for the SNP genotyping is ARMS- PCR, as it involves single PCR reaction and followed by gel electrophoresis (Medrano & De Oliveira, 2014). The amplification refractory-mutation system (ARMS) or allele specific polymerase chain reaction is an efficient method to detect any mutation that involve single base change or small deletions. ARMS use specific sequence of PCR primer to amplify the target DNA if only the target allele contained in the DNA sample,

then the PCR product will be analysed for the presence or absence of target allele (Haines *et al*., 2001).

ARMS-PCR used four primers to determine genotype by amplification of two non-allele specific primer (outer Primer) in the region that contain SNP. The outer primer fragment is produced. The outer primer fragment serves as a template to the two allele specific fragments (inner primer) to produce allele specific fragments (Medrano & De Oliveira, 2014).

2.5.1 Principle of ARMS-PCR

The principle of this procedure based on mismatch in 3' terminus of allele-specific primer resulting in the specification of primer into only one allele of SNP and refractory to another allele. DNA polymerase extends the primer when its 3' end complementary to the template, thus amplicon is produced. To increase the specificity of the reaction, the mismatch not only occur in 3' terminus but also in the position-2 from 3' terminus of the same allele specific primer. This mismatch destabilized the base pairs of the primer and non-target templates (Medrano & De Oliveira, 2014). The schematic illustration of SRMS-PCR for SNP genotyping can be summarized in figure 2.5.

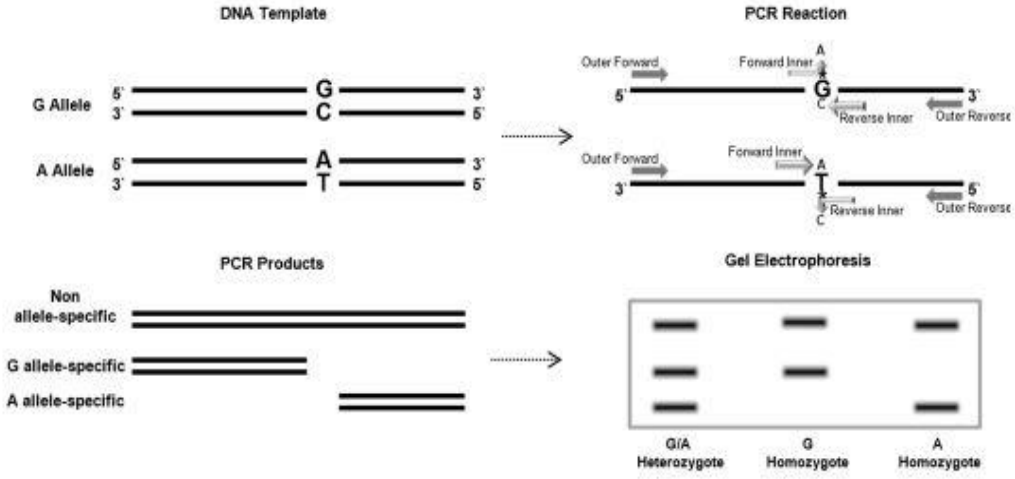

Figure 2.5 ARMS-PCR for SNP genotyping (Medrano & De Oliveria, 2014).

CHAPTER 3

METHODOLOGY

3.1 Study Design

This is a cross-sectional retrospective study to determine the prevalence of the *SLCO1B1* rs4149056 among the outpatient statin users and to evaluate its association with the patients' lipid profiles. The information such as demographic information including age, race and gender, type of statin being consumed, the daily dose, period of treatment, concurrent diseases and concomitant drug, and the laboratory results of lipids profile (total cholesterol, LDL-c, HDL-c, and triglycerides), will be obtained from archived records from the previous researcher.

The blood sample of participants were obtained from the sample storage located in Human Genome centre Laboratory.

3.1.1 Inclusion criteria

Patients who were hyperlipidaemic and be treated with statin at the aged in between 18 to 75 at the time of recruitment and received statin for at least 6 weeks were selected.

3.1.2 Exclusion criteria

Patients that have thyroidal diseases, acute or chronic kidney diseases, acute or chronic liver diseases and taking any drug that interact with statin disposition were excluded

3.1.3 Sample Size

Original study indicated that the total number of patient cohort for the statin study was based on a comparison of proportions between two groups using other SNP of interest (ii.e.,CETP rs708272) and not the SLCO1B1 SNP. Referring to a study in Thailand, the indicated SNP has resulted in reduced LDL-c levels from statin treatment by 35.0 % in the GG + GA group and 23.0 % in the AA group (Wanmasae *et al*., 2017). The power (1-β) was set to 80 %, and the significance level (α) was set to 0.05. The total number of subjects for the original study was around 249 patients, expecting a difference of 12 % LDL-c decrease between the two groups and assuming a 10 % drop-out rate because some patients may not have complete lipid profiles or be on other lipid-lowering drugs.

However, total of 72 statin users have been chosen since only 72 sample managed to be genotyped due to the limited time for conducting the study. In addition, inclusion and exclusion factors still included in determining the sample size.

3.1.4 Selection of SNP

SLCO1B1 gene (Gene ID:10599) encoded for the protein of organic anion polypeptide 1B1 (OATP1B1) (Santos *et al.*, 2011) which is expressed on the sinusoidal membrane of human hepatocytes (Huang *et al*., 2012). This transmembrane protein responsible in transporting anionic drugs such as statins, irinotecan, rifampin, repaglinide and methotrexate (Gurusamy & Shewade, 2014). This gene locus occupied 109 kb and located on chromosomes 12 (Chr 12p12.2) (Wilke *et al.*, 2012).

SLCO1B1 rs4149056 is a single nucleotide polymorphism that can be found in exon 5 (Santos *et al*., 2011) resulting in the substitution of an amino acids from valine to alanine (Turkmen *et al*., 2022). This SNP has been chosen in this study as it may

influence the effectiveness of the statin therapy resulting in the lower efficacy in reducing LDL-c (Turkmen *et al*., 2022).

3.1.5 Primer design

The primers for *SLCO1B1* rs4149056 was designed by using programmed PRIMER1 (http://primer1.soton.ac.uk/primer1.html). The primers were subjected to BLAT Search Genome (https://genome.ucsc.edu/cgi-bin/hgBlat) to ensure the specificity. The primers were reconstituted to 100 µM by adding ddH_2O, followed by dilution to 10 µM as a working stock for the PCR amplification. The list of primers as shown in Table 3.1

Table 3.1: Primer sequences for *SLCO1B1* rs4149056

Gene	Primer
SLCO1B1	Forward outer 5’ CATATTGTCAAAGTTTGCAAAGTGA 3’
	Reverse outer 5’ TTCAAAAGTAGACAAAGGGAAAGTGATCA 3’
	Forward inner 5’ ATCTGGGTCATACATGTGGATATAGGT 3’
	Reverse inner 5’ ATTCCACGAAGCATATTACCCATGACCG 3’

3.2 DNA Extraction

Extraction of DNA is the initiator in genomic studies thus it can be considered essential procedure. DNA extraction is a technique to purify the DNA by separating DNA from cell membranes, protein or other cellular components by using physical or chemical method (Gupta, 2019). The first DNA extraction was performed by Friedrich Miescher in 1869 in which he accidentally purified DNA from nucleus while analysed protein from leukocytes and discovered the difference of DNA compare to the protein, thus the term ‘*nuclein*’ was invented (Dairawan & Shetty, 2020).