Joint Biotechnology Master Program

Palestine Polytechnic University

Deanship of Higher Studies and Scientific Research

Bethlehem University

Faculty of Science

Design and Optimization of a CYP2C19*2 Genotyping Assay Based on the SuperSelective Primer Method

By

Salah Alddin Mohamed Al-Jubeh

In Partial Fulfillment of the Requirements for the Degree

Master of Science

September 2021

The undersigned hereby certify that they have read and recommend to the Faculty of Scientific Research and Higher Studies at the Palestine Polytechnic University and the Faculty of Science at Bethlehem University for acceptance a thesis entitled:

Design and Optimization of a CYP2C19*2 Genotyping Assay Based on the SuperSelective Primer Method

By

Salah Alddin Mohamed Al-Jubeh

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in biotechnology

Graduate Advisory Committee:

Committee Member (Student's Supervisor)	Date
Dr. Yaqoub Ashhab, Palestine Polytechnic University	
Committee Member (Internal Examiner)	Date
Dr. Suhair Lolas Hamameh	
Committee Member (External Examine)	Date
Dr. Muneef Ayyash	

Approved for the Faculties

Dean of Graduate Studies and Scientific Research	Dean of Faculty of Science
Palestine Polytechnic University	Bethlehem University
Date	Date

I. Abstract

Introduction: Percutaneous coronary intervention (PCI) is a nonsurgical technique carried out frequently to treat patients suffering obstructive coronary artery disease. These patients are treated with a combination of aspirin and an antiplatelet drug (clopedogril) for one year following the intervention. CYP2C19 enzyme is considered the major hepatic enzyme that contributes in the conversion of clopedogril from inactive prodrug to active drug. Several CYP2C19 genetic variants are associated with a significant decreased enzymatic activity. Accordingly, CYP2C19 genotyping is recommended to test patients for the most frequent poor metabolizer genotype, particularly the CYP2C19*2 variant. The aim of this study is to develop and optimize a new technique for CYP2C19*2 genotyping using the recently invented SuperSelective primer method.

Material and method: BLAST analysis, PCR assay, and Sanger sequencing were used to establish and validate the reference control of the three possible genotypes, namely, the wild-type *1/*1, the heterozygous *1/*2, and the mutant *2/*2. Real-time PCR assay was utilized in the test to optimize the parameters of a SuperSelective primer "CYP2C19*2 24-14/13-6:1:1". The newly developed assay was used to genotype 81 genomic DNA samples and Sanger sequencing was used for validating accuracy of the new assay.

Result and discussion: The newly developed Real-time PCR assay utilizing "CYP2C19*2 24-14/13-6:1:1" SuperSelective primer showed a 100% accuracy in genotyping and distinguishing the heterozygous from the wild-type samples with a minimum discriminatory cutoff of (Ct =7). The assay demonstrated a high genotyping efficiency of all the tested clinical samples across a broad range of their genomic DNA concentrations.

Conclusion: The SuperSelective based CYP2C19*2 genotyping assay, developed in this study, is an accurate and relatively fast genotyping assay that can be used in labs with resource-limited settings for genotype-guided antiplatelet therapy.

II. ملخص

المقدمة : يتم إجراء رأب الوعاء التاجي بشكل متكرر للأشخاص الذين يعانون من متلازمة الشريان التاجي الحادة في جميع أنحاء العالم. يتضمن العلاج إعطاء المرضى زوج من الأدوية المانعة لتخثر الصفائح وهم الأسبرين والكلوبيدوجريل لمدة عام واحد بعد إجراء رأب الوعاء التاجي. يعتبر الإنزيم CYP2C19 هو الإنزيم الكبدي الرئيسي المسؤول عن آلية تحويل دواء الكلوبيدوجريل، وهذا الإنزيم له عدة متغيرات جينية مرتبطة بانخفاض النشاط الأنزيمي. يوصى بعمل التنميط الجيني لإنزيم CYP2C19 للمتغيرات المسؤولة عن انخفاض مستوى النشاط وخاصة CYP2C19*2. الهدف من هذه الدراسة تصميم وتكييف طريقة جديدة للتنميط الجيني باستخدام برايمر خاص يسمى SuperSelective التي تم اختراعها مؤخرا.

المادة والطريقة: تم استخدام تحليل BLAST ومقايسة جهاز البلمرة المتعدد المتسلسل PCR وتسلسل Sanger لإنشاء عينات مرجعية تحتوي على التغير الجيني وعينات لا تحتوي على التغير الجيني. تم استخدام تفاعل البوليميريز المتسلسل بالزمن الحقيقي بتوظيف البرايمر "CYP2C19*2 24-14/13-6:1:1" لمعايرة وتكييف التقنية. التقنية المصممة استخدمت للتنميط الجيني ل 81 عينة حمض نووي.

النتائج والمناقشة: استخدام تفاعل البوليميراز المتسلسل بالزمن الحقيقي بتوظيف البرايمر "CYP2C19*2 24-14/13-6:1:1" لديه القدرة على معرفة التركيب الجيني والتمييز بين الاليل الحامل للطفرة من الاليل الغير حامل للطفرة بناءا على قيمة ال Ct (7>). إضافة لذلك، الفحص المطور لديه القدرة على التنميط الجيني للعينات السريرية عبر مجموعة واسعه من تركيزات الحمض النووي الجينومي.

الخلاصة: الفحص الذي تم تطويره في هذه الدراسة باستخدام البرايمر SuperSelective قادر على تمييز التنميط الجيني للإنزيم CYP2C19 بطريقة سريعة، ممكنة، دقيقة، وبأرخص تكلفة.

III. Declaration

I declare that this thesis with title "*Design and Optimization of a CYP2C19*2 Genotyping Assay Based on the SuperSelective Primer Method*" is an original report of my research, has been written by me and has not been submitted for any previous degree. The experimental work is almost entirely my own work; the collaborative contributions have been indicated clearly and acknowledged. Due references have been provided on all supporting literatures and resources.

Name and signature: Salah Alddin Al-Jubeh

Date: 28/9/2021

IV. Dedication

I dedicate my dissertation work to my family and many friends. A special feeling of gratitude to my loving parents, Mohamed Jameel and Dalal Al-Khatib whose words of encouragement and push for tenacity ring in my ears. I also dedicate this dissertation to the memory of my grandfather and grandmother whose passed away during the Covid-19 pandemic.

I also dedicate this work and give special thanks to my brothers and sisters, who have never stop support, love and pray. And finally to my friends and my colleagues in the medical laboratory field.

ACKNOWLEDGMENT

First and most important of all, I direct all my thanks to God (الله, ALLAH) who overwhelmed me with His mercy, care and love, where his majesty never disappointed me, despite that I don't deserve all that he gave me, and I will never will.

I would like to express my deepest thanks, appreciation and gratefulness to my thesis supervisor: Dr. Yaqoub Ashhab. He always stood by me in good as well as bad times in helping me to complete this work despite difficulties arising here and there.

I direct my special thanks to Asma' Tamimi for her encouragement and technical support in my thesis.

I would also like to express my appreciation to the faculty staff and lecturers in the Biotechnology Master program in Palestine Polytechnic University: Dr. Yaqoub Ashhab, Dr. Robin Abu Ghazaleh, Dr. Rami Arafeh, Mr. Hasan Taradeh, Mr. Zayd Tarade, and Mr. Ayman Eideh for their tremendous help and assistance.

I also like to thank my lovely friend Mr Nassim Al-Bakri for his assistance and support during completion my thesis.

V. List of abbreviations

Abbreviation	Meaning	Page
%	Percent	2
°C	Celsius degree	15
µL	Micro liter	21
µM	Micro mole	20
ABCB1	ATP Binding Cassette Subfamily B Member 1	3
ACS	Acute Coronary Syndrome	3
ADP	Adenosine diphosphate	3
ARMS	Amplification Refractory Mutation System	11
ATP	Adenosine triphosphate	6
Bp	Base pair	11
BRAF	B-rapidly Accelerated Fibrosarcoma	43
CES-1	Carboxy-Lesterase 1	3
Ct	Cycle threshold	21
CYP	Cytochrome P450	2
CYP2C19	Cytochrome P450 Family 2 Subfamily C Member 19	2
CYP2C9	Cytochrome P450 family 2 subfamily C member 9	2
dNTP	Deoxyribose nucleoside triphosphate	20
DPO	Dual Primer Oligonucleotide	12
ds-DNA	Double-strand DNA	41
EDTA	Ethylenediaminetetraacetic acid	18
EGRF	Epidermal Growth Factor	43
FDA	Food and Drug Administration	2
gDNA	Genomic DNA	18
He	Heterozygous	23
Ho	Homozygous	23
HRM	High Resolution Melting	14
INR	International Normalized Ratio	2
LFA	Lateral Flow Assay	13
m-RNA	Messenger- RNA	6

Ng	Nano gram	21
nm	Nano mole	20
NTC	no template control	26
PCI	Percutaneous Coronary Intervention	3
PCR	Polymerase Chain Reaction	10
PGMNs	Poly-acrylic Acid Modified Gold Magnetic Nanoparticles	13
P-gp	P-glycoprotein	6
Poly-I	Polydeoxyinosine	12
PM	Poor metabolizer	14
PPA	Poly-acrylic acid	13
qPCR	Quantitative Polymerase Chain Reaction	14
RFLP	Restriction Fragment Length Polymorphism	11
RFU	Relative Fluorescent Unit	32
SNP	Single Nucleotide Polymorphism	1
Std	Standard	22
SYBR	Sybase Replication Server	21
TE	Tris-EDTA	21
UM	Ultra-rapid Metabolizer	14
Wt	Wild type	22

VI. Table of figures

VII. Table of tables

Table of content

CHAPTER 1

1. Introduction

1.1 Pharmacogenetics

It is a well-known fact that patients' responses to medication vary from person to person. According to their responses to a given drug, patients can be classified into three general categories: good responders, poor responders, and non-responders. This variation in response depends on genetic and non-genetic factors. Of the non-genetic factors is the route of administration, which includes buccal and sublingual oral administration, dry powder and liquid spray inhalation, nasal drops and sprays, otic, ocular, or rectal administration, and finally, parenteral injection. Other non-genetic variables include age (people of advanced age may be less responsive than those of a younger age), gender (women are more likely to experience adverse drug interactions), clinical status (e.g., kidney function, liver function, and protein status), drug-drug interaction (inhibition of the metabolism of one drug by the other), lifestyle (e.g., food-drug interaction, smoking, and alcohol consumption), disease, infection, and pregnancy, in addition to other variables (Abul-Husn, Obeng, Sanderson, Gottesman, & Scott, 2014).

One of the most important factors affecting variation in drug response is the genetic factor, which is known as *pharmacogenetics*. Pharmacogenetics has been described as a tool that "can predict and/or explain how individuals respond to drugs" and as a "prominent component of personalized, precision medicine"(Rifai, Horvath, Wittwer, & Park, 2018). Pharmacogenetics predicts gene variability at the chromosome and nucleotide level, and based on that, gene variability, especially single nucleotide polymorphisms (SNPs), lead us to understand how individuals respond to medication, and can guide treatment either by adjusting drug dose or alternating to another drug. Additionally, related work associating drug response with many genes, and ultimately the whole genome, is known as *pharmacogenomics*, and the terms are commonly used interchangeably.

Variations in response to medication, either due to genetic or non-genetic factors, should be taken into consideration when making medical decisions in order to achieve the best benefit from the medication. Therefore, drug therapy decisions are based on several clinical measurements. For example, blood pressure should be observed for patients being administered antihypertensive drugs

(Jorga, Holt, & Johnston, 2004), and dosage of the anticoagulant drug Coumadin is adjusted depending on prothrombin time (PT) and the international normalized ratio (INR) derived from it (Pruthi, 2013).

Drug therapy decisions based on monitoring a patient's clinical measurements are not sufficient to determine the drug response before administration. Therefore, physicians' decisions should be based on genetic variant analysis, which aims to improve drug efficacy and prevent up to 60 % of adverse drug reactions. For these reasons, pharmacogenetic testing is designed to predict the best choice of drug as well as the optimal route of administration (Yang, Tavassolian, Haddad, Bailey, & Gholami, 2019).

There are many drugs for which the response in the human body is affected by genetic polymorphism. Examples of these drugs are listed in a table created by the Food and Drug administration (FDA), which includes oncology drugs (Afatinib and Alectinib), neurology drugs (Amifampridine Phosphate and Carbamazepine), cardiology drugs (clopidogrel and Metoprolol), psychiatric drugs (Amoxapine and amphetamine), hematology drugs (warfarin and voxeloter), and others. In addition, the table includes tens of pharmacogenomic biomarkers that affect drug response.

One of the commonly used drugs in the FDA list is a Coumadine (crystalline warfarin sodium). Coudamine is a racemic mixture of the R- and S-enantiomers which act as an anticoagulant by inhibiting vitamin K-dependent coagulation factors (factors II, VII, IX and X, and the proteins C and S). In 2011, warfarin was acknowledged to be the leading cause of hospitalization among American adults, with the percentage of patients experiencing adverse effects reaching 33.3 % (Budnitz, Lovegrove, Shehab, & Richards, 2011). During warfarin metabolism, the CYP2C9 enzyme plays a major role in warfarin metabolism, and allelic variants of CYP2C9 affect enzyme activity and therefore warfarin metabolism. Thus, a pharmacogenetics test for CYP2C9 was designed to test for the CYP2C9*2 and CYP2C9*3 allelic variants. Patients who have these variants are expected to have decreased CYP2C9 enzyme activity, by ~30–40% and ~80–90% respectively, compared to patients with the homozygous CYP2C9*1 genotype. Therefore, patients with the CYP2C9*2 and CYP2C9*3 allelic variants are at high risk for bleeding during therapy and require lower doses of warfarin to achieve appropriate pharmacological targets (Johnson, Gong, Whirl-Carrillo, Gage, Scott, Stein, Anderson, Kimmel, Lee, & Pirmohamed, 2011).

1.2 Pharmacokinetics and Pharmacodynamics Mechanisms

During the twenty-first century, pharmacology has studied two important mechanisms which summarize the drug response for each drug. These mechanisms are pharmacokinetics and

pharmacodynamics. Pharmacokinetics describes the kinetic (movement) pathways when the drug enters the body and acts on it, which involves absorption, distribution, metabolism, and elimination of a drug. On the other hand, pharmacodynamics explains how the body responds to the drug by studying biochemical, physiological, and molecular effects, which involve receptor binding (to proteins, enzymes, or ion channels), post-receptor effects, and chemical interactions (drug-drug interactions) as described in Figure (1.1) (Yang et al., 2019). Genetic polymorphisms can affect pharmacokinetic and pharmacodynamics elements by affecting any protein structure involved in drug metabolism. Interestingly, the United States Food and Drug Administration (FDA) cannot approve any drug before performing pharmacokinetic and pharmacodynamics analyses.

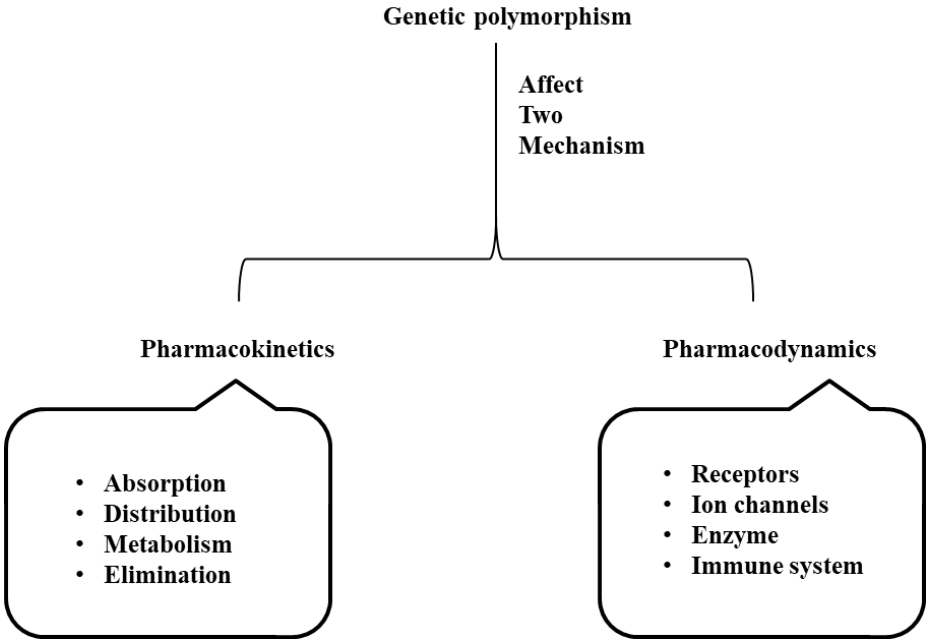

Figure 1.1: Genetic polymorphisms in gene coding for protein can affect drug response by two major mechanisms called pharmacokinetics and pharmacodynamics.

1.3 Clopidogrel (Plavix) and cytochrome P450 family 2 subfamily C member 19 (CYP2C19)

Clopidogrel (Plavix) is an antiplatelet prodrug that binds to the P2Y12 platelet receptor causing reduced platelet aggregation in patients suffering from acute coronary syndromes (ACSs) undergoing percutaneous coronary intervention (PCI) who have stable ischemic heart disease (Ma, Chen, Zhang, Zhang, & Huang, 2019). Clopidogrel is orally administrated as an inactive prodrug at a dosage of 75 mg daily. Only 15% of the drug is converted to active form, and the remaining 85% is eliminated by carboxy-lesterase 1 (CES-1) (Hasan, Basri, Hin, & Stanslas, 2013). The plasma membrane P-glycoprotein (ATP Binding Cassette Subfamily B Member 1: ABCB1) in the gastrointestinal tract

adsorbs the drug into blood vessels where it then circulates to the liver, where it is metabolized in two oxidation steps carried out by a group of cytochrome P450 enzymes. The first oxidation step converts clopidogrel into 2-oxo-clopidogrel, followed by the second oxidation step carried out by esterases to convert approximately 85% of the 2-oxo-clopidogrel into the inactive form, which is then eliminated in urine and stool. Afterwards, active metabolites target the P2RY12 platelet receptors and bind to them irreversibly, leading to a reduction in platelet aggregation, which require approximately 10 days for restoration of function (Hasan et al., 2013; Johnson & Cavallari, 2013; Maegdefessel, Azuma, Tsao, & management, 2010). Figure (1.2) describes the mechanism of action of clopidogrel.

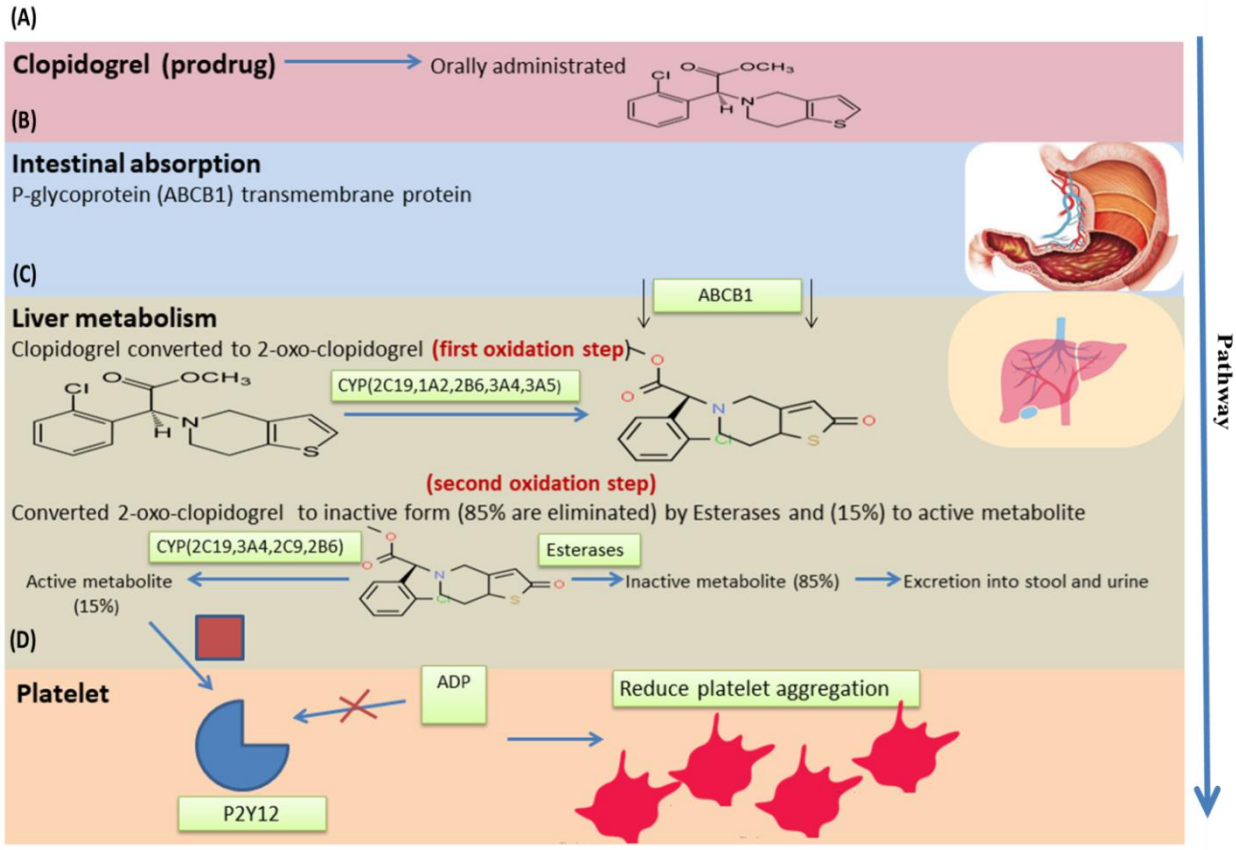

Figure 1.2: Mechanism of action of clopidogrel. (A) Oral intake of clopidogrel as a prodrug. (B) Absorption of clopidogrel by an intestinal transmembrane protein (ABCB1). (C) Two oxidation steps in the liver to convert clopidogrel into its active form. (D) Active metabolites target platelets to reduce platelet aggregation.

1.4 Genetic Factors Affecting Response to Clopidogrel

The response to clopidogrel depends on several pharmacokinetics and pharmacodynamics factors which are summarized in Table (1.1) (Small et al., 2010; Taubert et al., 2006; Zhu et al., 2013). The most important of these factors is the genetic variation of the CYP2C19 enzyme produced in the liver.

The major action of clopidogerl is reducing the aggregation and life span of platelets. Accordingly, the effectiveness of clopidogrel (the pharmacodynamics effect) is measured by its ability to reduce platelet aggregation in a patient.

Table 1.1: Pharmacokinetic and pharmacodynamic factors of clopidogrel.

Clopidogrel pharmacokinetics	
Absorption	Uptake of clopidogrel is determined by a transmembrane P-glycoprotein (P-gp, ABCB1). P-gp is the major determinant of clopidogrel intestinal absorption and oral bioavailability.
Distribution	Only about 2% of the administered clopidogrel dose is converted to active metabolite and enter the systemic circulation.
Metabolism	Number of cytochrome P450 (CYP) enzymes, including CYP2C19, CYP1A2, CYP2B6, CYP2C9 and CYP3A4, play an important role in the bioactivation and metabolism of clopidogrel via a two-step process in the liver. CYP2C19 is the major hepatic enzyme which contributes to the formation of approximately 50% of the overall clopidogrel.
Clopidogrel Pharmacodynamics	
Clopidogrel exhibits its pharmacodynamic effect by specifically and irreversibly binding to P2Y12, a subtype of the adenosine diphosphate (ADP) receptor, on the surface of platelets.	

1.4.1 The CYP2C19 Gene and its Variants

The CYP2C19 gene (Ensembl ID: ENSG00000165841) is a protein coding gene that is a member of the cytochrome P450 gene family. The protein product is found mainly in the liver and 25 other tissues. The gene is located on chromosome 10q24 within a cluster of cytochrome P450 genes, and its mRNA transcript contains nine exons (Scott et al., 2012). Since the nineties of the last century up until now many significant variants of the CYP2C19 gene have been identified, with more than 25 variant alleles being discovered. Figure (1.3) illustrates the CYP2C19 gene and its most significant variants that affect enzyme activity and function. Additionally, a genome-wide association study conducted on 429 healthy

Amish individuals was carried out to identify gene variants that influence clopidogrel response for 7 days of clopidogrel administration by measuring ex vivo platelet aggregation. A cluster of cytochrome enzymes (CYP2C18–CYP2C19–CYP2C9–CYP2C8) showed a high degree of statistical significance for correlation to clopidogrel response. Due to a high linkage disequilibrium with the CYP2C19*2 variant, patients with this variant treated with clopidogrel who were undergoing coronary intervention were evaluated for any complications. Ultimately, patients carrying the CYP2C19*2 genotype were found to be at greater risk for cardiovascular ischemic events or death at 1 year follow-up (Shuldiner et al., 2009).

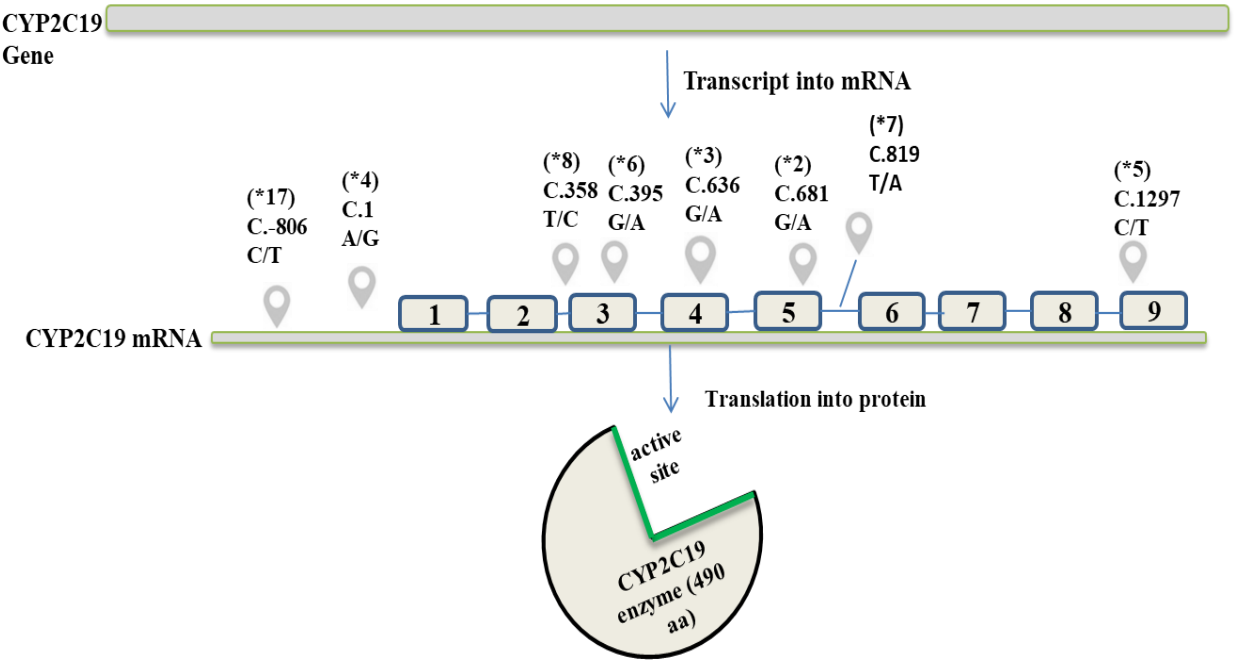

Figure 1.3: CYP2C19 gene variants. Loss of function (*2-*8) and gain of function (*17) alleles. Exons are represented by numbered rectangles. CYP2C19*4-*8 are less frequent (rare) (<1%).

The presence of multiple variants of the CYP2C19 gene, as depicted in Figure (1.3), affects the enzyme's function in drug metabolism, with different variants resulting in enzymes with either decreased or increased activity. The most well-known variant associated with an increase in enzyme activity is the CYP2C19*17 variant, whereas a decrease in enzyme activity is associated with the CYP2C19*2, *3, *4, *5, *6, *7 and *8 variants. Moreover, the most frequent loss of function allele variants are CYP2C19*2 and CYP2C19*3. These two loss of function variants each contain a single nucleotide polymorphism SNP that completely inactivates the CYP2C19 enzyme as explained in Figure (.14). The CYP2C19*2 variant has a G to A transition in exon 5 which creates an aberrant splice site that alters the mRNA reading frame and results in a nonfunctional truncated protein. This protein is

characterized by a deletion of amino acids 215-227, with a shift in reading frame, producing a premature stop codon 20 amino acids downstream (De Morais, Wilkinson, Blaisdell, Nakamura, et al., 1994). Consequently, the resulting protein of the CYP2C19*2 variant lacks the binding site for the heme, and is therefore catalytically inactive, as shown in Figure (1.4.b).

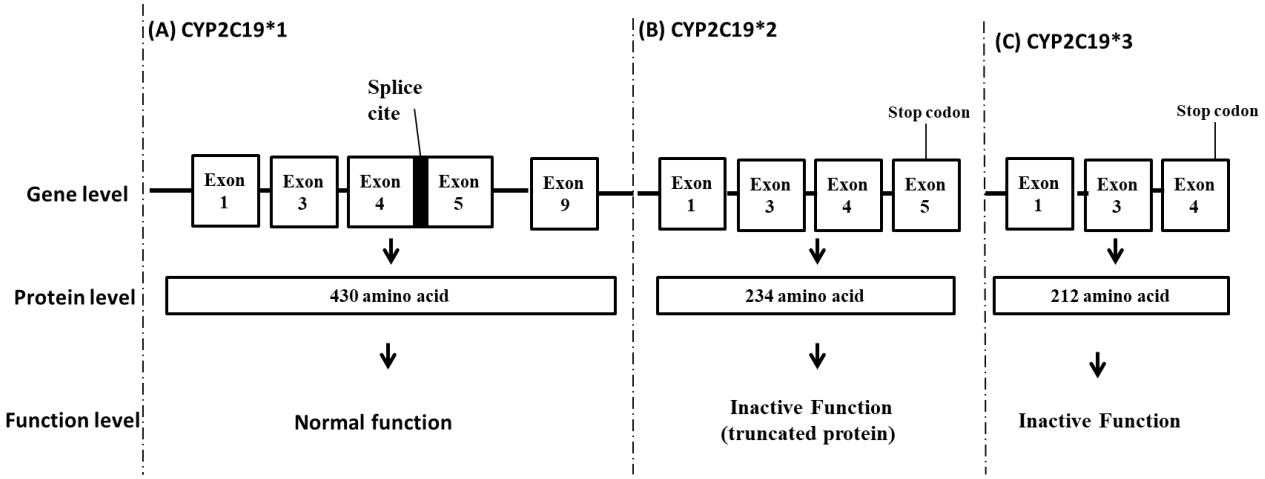

Figure 1.4: CYP2C19 translated process. (A) describe normal translation of CYP2C19. (B) describe abnormal protein translation of CYP2C19*2. (C) describe abnormal protein translation of CYP2C19*3.

Another important loss of function allele is the CYP2C19*3 variant, with a G to A transition in exon 4 leading to the replacement of the amino acid tryptophan with a stop codon (UAG) at amino acid position 212, as described in Figure (1.4.c). This SNP produces a CYP2C19 protein of only 212-amino acids, which lacks the heme-binding site, and is therefore inactive (De Morais, Wilkinson, Blaisdell, Meyer, et al., 1994). Accordingly, patients who are carriers of either the *2 or *3 SNP variants are considered either intermediate or poor metabolizers of clopidogrel as well as other drugs related to metabolic pathways that involve CYP2C19. The FDA in 2010 added a red warning box for the use of the antiplatelet drug Plavix in patients who have a poor or intermediate metabolizer genotype with the *2 or *3 variants (Food & Administration, 2010).

Patients with other less frequent non-functional variants of CYP2C19 are also classified as having a poor metabolizer phenotype, however, these variants are considered very rare with a frequency of less than 1%. These include loss of function variants such as CYP2C19*4, CYP2C19*5, CYP2C19*6, and CYP2C19*7. For example, the CYP2C19*4 variant has an A to G transition in the initiation codon (ATG), producing a GTG sequence in exon 1, which consequently stops the protein translation process (Ferguson et al., 1998). The CYP2C19*5 variants has a C to T transition that leads to an exchange of the amino acid arginine for tryptophan in the heme-binding region, producing a non-

functional protein (Xiao et al., 1997). The CYP2C19*6 variant was first discovered in Caucasians displaying a poor metabolizer of mephenytoin phenotype, which is caused by a conversion from G to A in exon 3, which changes the amino acid from arginine to glutamine in position 132, and thus diminishes the catalytic activity of the CYP2C19 enzyme towards S-mephenytoin (Ibeanu et al., 1998).

1.5 CYP2C19 genetic testing guideline

Considerations must be made for the identification of patients carrying any loss of function allelic variants of CYP2C19, especially in cases involving the administration of clopidogrel. Therefore, patients suffering from acute coronary syndrome (ACS) and undergoing percutaneous coronary intervention (PCI) are treated with anti-platelet aggregation drugs, which can include a dual therapy of clopidogrel (75 mg/daily) and aspirin. Aspirin acts as a blocker for thromboxane A2 synthesis, thereby reducing platelet aggregation and reducing the need for any secondary prevention for an extended duration. However, aspirin alone is not sufficient to achieve the target of reducing platelet aggregation due to its side effects and effectiveness, and particularly due to aspirin allergy, as well as others factors (Kumar & Cannon, 2009). Based on several studies, clopidogrel is considered more effective than aspirin in patients suffering from heart disease, and is known to improve health status in patients undergoing PCI (Sabatine et al., 2008). In recent years, dual antiplatelet therapy, including a P2Y12-receptor antagonist and aspirin, resulted in a reduction in recurrent ischemic events, such as stent thrombosis, compared to cases where aspirin was administered alone (Pareek & Bhatt, 2018).

As shown above, clopidogrel drug responsiveness and its effectiveness is contingent on the presence of certain SNP variants in the CYP2C19 gene. Therefore, patients are classified into four categories based on CYP2C19 status: poor metabolizers, who are homozygous for two loss of function alleles (ex: *2/*2); extensive metabolizers, who are homozygous for two functional alleles (ex: *1/*1); intermediate metabolizers, who are heterozygous for one functional allele with one loss of function allele or increased activity allele (ex: *1/*2 or *1/*17); or ultra-rapid metabolizers, who are homozygous for increased allele activity (ex: *17/*17) or heterozygous with one normal function allele and increased allele activity (ex: *1/*17) (Johnson, Gong, Whirl-Carrillo, Gage, Scott, Stein, Anderson, Kimmel, Lee, Pirmohamed, et al., 2011). Table (1.2) describes CYP2C19 genotyping and patient classification based on CYP2C19 status, and in addition contains FDA recommendations for each genotype.

Table 1.2: Antiplatelet therapy recommendations based on CYP2C19 status for administration of clopidogrel in ACS/PCI patients.

Phenotype	Therapeutic recommendations	Example (genotype)
Ultrarapid-metabolizer (UM) and extensive metabolizer (EM)	Strong recommendation[A1]	UM (*1/*17, *17/*17), EM (*1/*1)
Intermediate metabolizer	Moderate recommendation[A2]	(*1/*2, *1/*3, *2/*17)
Poor metabolizer (PM)	Strong recommendation[A2]	(*2/*2, *2/*3, *3/*3)

A1 2- Strong recommendation of 75 mg/daily of clopidogrel, A2- Moderate recommendation for alternative antiplatelet therapy (Prasugrel or Ticagrelor), A3- Strong recommendation for alternative antiplatelet therapy (Prasugrel or Ticagrelor).

1.6 CYP2C19 Variants Across the World

Across different populations the allele frequency of genetic variations in the CYP2C19 gene are diverse. Hundreds of studies have been carried out to assess allele frequency for variants of the CYP2C19 gene, especially for the CYP2C19*2, *3, and *17 variants. In Norwegian populations the allele frequencies of the CYP2C19*2, *3, and *17 variants were found to be 18.1%, 0.6% and 22.0%, respectively (Rudberg et al., 2008). A similar frequency was found in the Swedish population with 14.4% for CYP2C19*2, 0.7% for CYP2C19*3, and 20.0% for CYP2C19*17 (Ramsjö et al., 2010; Yamada et al., 1998). Allele frequencies in the Korean population were found to be 20.9% for CYP2C19*2, 11.6 % for CYP2C19*3, and 1.3 % for the extensive metabolizer CYP2C19*17 (Herrlin et al., 1998; Kim, Song, Kim, Park, & therapeutics, 2010). In the Middle East and Arabian countries, many studies have assessed the allele frequency of CYP2C19 variants. In Saudi Arabians, allele frequency for alleles resulting in the poor metabolizer phenotype was 8.2%, with only one case each for CYP2C19*2 and CYP2C19*3, while 26.9% displayed the extensive metabolizer phenotype with the CYP2C19*17 allele (Al-Jenoobi et al., 2013). In the Jordanian population, a high allele frequency of intermediate metabolizer phenotype variants was found, reaching 33.0%, when researchers assessed most of the poor metabolizer phenotype variants (CYP2C19*2,*3, 4*,5* and *6) (Al-Jenoobi et al., 2013).

1.7 Molecular Techniques Used for CYP2C19 Genotyping

Multiple methods and techniques have been used for SNP detection and genotyping over the last century. These methods are able to detect a change in a single nucleotide in the human genome as well as the genomes of other organisms. Since Kary Mullis invented PCR, the technique has become a cornerstone for most methods used for genotyping and mutation detection. DNA amplification and detection usually including the following steps: (I) DNA is extracted from a biological sample; (II) specific primers are design for a certain target DNA segment; (III) specific thermal cycling protocols are establish; and (IV) visualization of the amplification product by any visualization system, such as agarose gel. However, genotyping methods differ from each other depending on multiple parameters, including the number of identifiable SNPs, cost per DNA sample, time, technical expertise, available equipment, and method accuracy.

1.7.1 Restriction Fragment Length Polymorphism (RFLP)

The restriction fragment length polymorphism (RFLP) method was first used for CYP2C19*2 genotyping by simply amplifying the region of interest by PCR followed by digestion of the PCR product and visualization on a gel based material such as polyacrylamide (De Morais, Wilkinson, Blaisdell, Meyer, et al., 1994). RFLP consists of three basic steps, starting with DNA extraction from a biological sample (venous blood, saliva, tissue, etc.), followed by PCR amplification of the specific region of interest (which requires primer design and master mix preparation), and then restriction digest. For example, one protocol calls for the amplification of two regions, a 167 bp region for CYP2C19*2 and a 233 bp region for CY92C19*3. The PCR products are then analyzed by using a specific concentration of agarose gel, such as 2%. After successful PCR amplification, the PCR product is incubated with a restriction endonuclease, such as SmaI for CY92C19*2 and BamHI for CYP2C19*3. The PCR product, which may contain digested and undigested products, is run on 4% gel agarose. If two bands appear on the gel (117 bp and 50 bp for CYP2C19*2, and 137 bp and 96 bp for CYP2C19*3), then the product contains digested products and the genotype is wild type. If only one band appears, then the products are undigested and the genotype is homozygous. In contrast, digested and undigested bands reveal a heterozygous genotype. The RFLP method is considered simple, rapid and easy to use (Ohkubo et al., 2006).

1.7.2 The Amplification Refractory Mutation System (ARMS)

Another simple method used for CYP2C19 genotyping is the amplification refractory mutation system (ARMS), which is designed for SNP or point mutation detection. The ARMS method uses two

complementary reactions, first, a PCR reaction containing specific primers for the region of interest, and a second PCR reaction containing internal specific primers which have the ability to amplify only the mutant or wild type allele, depending on primer design. Both PCR amplicons are loaded into a gel and subject to electrophoresis, and based on the fragment length one is able to distinguish the genotype of the target gene. To illustrate, an ARMS protocol was developed for CYP2C19*17 genotyping. Four primers were designed to fit the requirement: outer forward, outer reverse, internal forward, and internal reverse primers for PCR amplification, which are all placed in the same tube. The outer forward and reverse primer create an amplicon with a product size of 462 bp. The internal forward primer for the T allele with the outer reverse primer creates an amplicon of 227 bp, and the reverse inner primer for the C allele with the outer forward primer creates a product of 292 bp. In either case, after PCR amplification and loading into the gel for electrophoresis, all human DNA samples should contain a band of 462 bp, which is considered a control for successful PCR amplification. Three bands with the various lengths are visualized when the genotype is heterozygous (CT), while only two bands appear (292 bp and 462 bp) when the genotype is wild type (CC), or homozygous (TT) with product lengths of 227 bp and 462 bp (Jin et al., 2020). The ARMS method is considered the most favorable method based on cost-benefit ratio compared with other SNP detection methods, and it is also reliable, simple, and fast. However designing four compatible primers is considered challenging, and may increase the probability of amplifying unwanted products, or the production of primer dimers (Duta-Cornescu et al., 2009). Moreover the ARMS method does not include a restriction endonuclease digestion step

On the other hand, RFLP is more technical, requiring several optimization steps, which increases the cost-effectiveness of the former technique. However, Costa et al. compared the two methods of CYP2C19*2 genotyping, and they concluded that RFLP-PCR is more efficient than the ARMS system. RFLP can detect at least one polymorphic allele in 74.5% of 200 samples, while the ARMS system did not detect any polymorphic allele, making RFLP-PCR highly sensitive compared to ARMS-PCR. The reason for this, as they mentioned, is due to several factors, including hairpin loop formation, primer dimer formation, and a dramatic difference in the annealing temperature for each primer (Costa et al., 2019).

1.7.3 The Dual Priming Oligonucleotide (DPO) System

Serious attention has been given to SNP genotyping in the molecular and genetic sciences in order to overcome the disadvantages of routinely used methods. Jong-Yoon Chun et al. invented a dual

priming oligonucleotide (DPO) system for multiplex detection of SNPs. This method involves special primers designed for more efficient allelic discrimination. The DPO system primers differ structurally from conventional primers. These special primers are composed of two separate segments connected by a polydeoxyinosine [poly(I)] linker which form a bubble like structure. This bubble structure improves the annealing temperature between primer segments and creates stable annealing which blocks the formation of any nonspecific priming and the formation of secondary structures. The DPO system was used for CYP2C19 genotyping in which nine samples differed in genotype (3 wild type, 3 heterozygous, and 3 homozygous), and the system was compared to genotyping using conventional primers. The multiplex PCR-DPO system more clearly distinguishes all the CYP2C19 genotypes compared to conventional multiplex PCR. In addition, the DPO system has several advantages over other methods, including a decrease in false positive results, prevention of primer-dimer formation, and the ability to give reliable, fast, and cost-effective results in one-PCR step (Chun et al., 2007).

1.7.4 Lateral Flow Assay Based on GoldMag Nanoparticles (GMNS)

A novel method was invented by Hui and Zhang et al. in 2016 for DNA point mutation detection. This method was built by a combination of nanoparticle based DNA detection and lateral flow assay (LFA) for nucleic acid analysis, especially for SNP and point mutation detection. Lateral flow assay based on GoldMag nanoparticles systems are composed of multiple elements, including poly-acrylic acid (PPA) modified gold magnetic nanoparticles (PGMNs), anti-digoxin antibodies, streptavidin, gout anti-mouse antibodies, and lateral flow assay strips. According to the manufacturer's instructions, DNA should be extracted from biological samples to obtain a sufficient quantity of DNA for later use. Based on the ARMS-PCR principle, two tubes with the same DNA template are run separately, with each tube containing a discriminatory primer, and the DNA is amplified by PCR. PGMN's conjugated with anti-digoxin antibody are inoculated onto a conjugate pad on the LFA strip. This strip is already pre-immobilized with a defined test line (the T-line) and a control line (the C-line) on a porous nitrocellulose membrane by streptavidin and gout anti-mouse antibody. Consequently, for genotyping, PCR products from the separate tubes are inoculated onto separate LFA strips, and DNA is captured by the PGMN's-anti digoxin antibodies and migrates across the nitrocellulose membrane and binds on the T-line (streptavidin and Gout anti-mouse antibody) and a red band is visualized. If the genotype is homozygous, then a red band appears on both the T and C lines only when inoculated with DNA from the mutant tube, while the T-line does not change color when DNA from the wild type tube is inoculated. Otherwise, the genotype is determined to be wild type (Hui et al., 2016).

Recently in 2019, a PCR lateral flow assay was design for CYP2C19 genotyping for detection of the poor metabolizer phenotype caused by the *2 and *3 variants. Xuhong and Sinong et al. succeeded in designing a lateral flow assay for CYP2C19 genotyping. This assay is characterized by its ability to carry out genotyping in 5 minutes after a PCR product is ready. Assay specificity reached 99.7%, with a limit of detection of 5 ng of genomic product. However, the cost of this test is still a limitation for it to be used in routine laboratories (Xuhong et al., 2019).

1.7.5 High Resolution Melting (HRM) Analysis in Real-Time PCR

Reducing the number of steps to gather results as quickly as possible is considered a challenge in the field of molecular diagnostics. Real time PCR or quantitative PCR (qPCR) are important tools in molecular diagnostics which monitor DNA amplification in real time by measuring dye intensity during the reaction. Of the multiple uses of real-time PCR, it plays an essential role in SNP genotyping. Additionally, introducing additional steps after PCR amplification called high resolution melting (HRM) analysis allows genotypes to be distinguished by the difference in DNA melting products. HRM analysis is a simple warming step done by gradually increasing the temperature and monitoring the color intensity against temperature. Based on that, a melt curve is produce, indicating the nature of the nucleic acid components (Wittwer, Reed, Gundry, Vandersteen, & Pryor, 2003). HRM analysis has been developed for CY92C19*17 genotyping. If the mutation is present there will be a corresponding change in the melting temperature and melting curve compared with those of the non-mutant template. In SNP genotyping by HRM analysis, a homozygous allelic variant is characterized by temperature shifting in the HRM melt curve, while a heterozygous variant is characterized by a change in the melt curve shape due to a heteroduplex component (Ghasemi et al., 2016). A novel multiplex HRM assay was invented to detect clopidogrel resistance. In this assay both the poor metabolizer (PM) and ultra-rapid metabolizer (UM) genotypes are detected for CYP2C19, and the assay is a highly sensitive and highly specific technique compared to the digital fluorescence molecular hybridization (DFMH) assay based on the molecular beacon (Zhang, Ma, You, Zhang, & Fu, 2017). Overall, the HRM assay is distinguished by its ability to be convenient, fast, closed-tubed, cost efficient, and consistently accurate.

1.7.6 Taqman Allelic Discrimination assay

TaqMan probes are hydrolysis probes that are designed to increase the specificity of quantitative PCR. Also known as the 5'-nuclease assay, the assay exploit the 5'-exonuclease activity of certain thermostable polymerases, such as Taq or Tth. The principle of TaqMan assay is that the probes

contains a florescent reporter at the 5' end and a quencher at the 3' end, when hybridize to its sequence specific target and during the annealing the Taq polymerase cleaved the reporter. As a result, the reporter is separated from the quencher, and the resulting fluorescence signal is proportional to the amount of amplified product in the sample (Livak, 1999). TaqMan assay was designed for CYP2C19 genotyping, the assay included both the specific primers and fluorescently labeled TaqMan such as FAM and VIC to amplify and detect the presence or absence of the allelic variants of CYP2C19 (*2,*3 and *17). Based on this assay, The presence of wild-type and variant alleles was defined by comparing the relative end-point fluorescence created by the degradation of each fluorescently labeled TaqMan probe (FAM and VIC) (Cervinski et al., 2013).

TaqMan assay have some advantages over the other allelic discrimination assays, which is considered high specific, a high signal-to-noise ration and ability to perform multiplex format. While the disadvantages are mainly concern about the initial cost of the probes and some difficulties in assay design.

1.8 Superselective Primers

A primer is well-known as a "short nucleic acid sequence that provides a starting point for DNA synthesis." and should be as sensitive, specific, and selective as possible for the target sequence in order to carry out its role. S. Marras et al. developed a highly selective nucleic acid amplification primer called the "SuperSelective" primer. A SuperSelective primer is a multi-part primer composed of 3 parts, which are all nucleic acid based: a 5'-segment "anchor sequence" component that is 15-40 bp in length and complimentary to the intended target, a "bridge sequence" of 15-20 bp that does not match the target sequence, and a "foot sequence" of 5-8 bp that is complimentary to the target sequence. Figure (1.5) describes the multi-part structure of a SuperSelective primer. Mismatching of the bridge component with intervening sequences creates a bubble-like structure that separates the anchor from the foot. During DNA amplification, the anchor sequence binds its intended sequence, while the bridge forms a bubble with the intervening sequence and the foot binds to its own intended sequence only if it is completely complementary to it. Subsequently, DNA polymerase starts copying even if a small or rare amount of target template is present in a large amount of unspecific DNA (Marras, Vargas-gold, Tyagi, & Kramer, 2018).

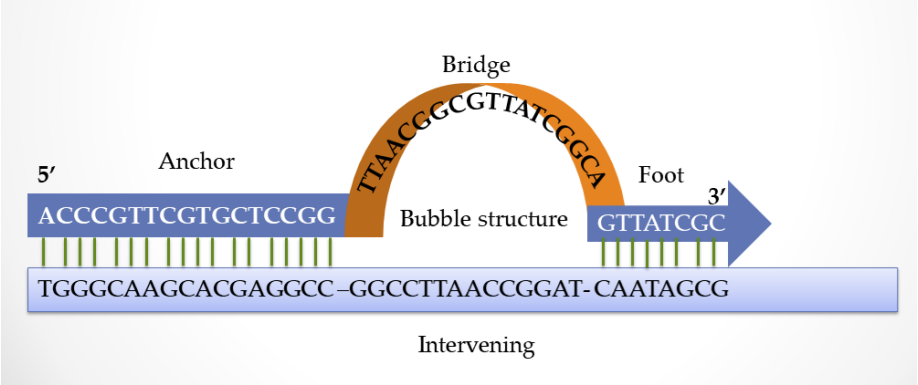

Figure 1.5 : The multi-part SuperSelective primer. 5' Anchor sequence connected to the bridge sequence and fsequence. The SuperSelective sequence is: 5'ACCCGTTCGTGCTCCGGTTAACGGCGTTATCCGGCAGTTATCGC-3'

The lengths of the three major parts of the SuperSelective primer have been evaluated for achieving optimal performance (Vargas, Kramer, Tyagi, & Marras, 2016). The most preferred parameters for an anchor sequence should fit the following: length of about 20-30 bp, melting temperature of above 50 °C, and few nucleotide mismatches from the intended sequence that do not affect performance. The most preferable parameters for the foot sequence are as follows: length between 6-8 bp, mismatch nucleotide position preferably toward the 3' end (particularly at the site before the last nucleotide, called "3' penultimate nucleotide").

The bridge is the middle part of the SuperSelective primer that is not bound to the template sequence and forms a bubble with the intervening sequence. The bubble forms a circular structure called "circumference", which is the sum of the bridge length and intervening sequence length plus four nucleotides. This circumference was subject to multiple experiments to determine the most preferable size of the bubble structure (Vargas et al., 2016). Studies show that it is not necessary for the bridge to be the same length of the intervening sequence, and the length of the intervening sequence can be zero, but, it is more preferable for the bridge length to be at least six nucleotides, and the intervening sequence to be at least ten nucleotides in length.

CHAPTER 2

2 Problem statement and objectives

Percutaneous coronary intervention (PCI) is frequently carried out for patients in Palestinian hospitals. To illustrate, 250-300 PCI interventions are carried out monthly at Al-Ahli Hospital alone in Hebron (Personal communication with the hospital chief medical officer based on their internal system data). These patients are routinely treated with the anti-platelet protocol compromised of 75 mg clopidogrel and 100 mg aspirin for one year after PCI. Unfortunately, a fraction of these patients do not respond to clopidogrel, and typically show a higher risk for stent thrombosis. The Food and Drug Administration (FDA) listed many warnings and precautions for clopidogrel administration for patients with loss-of-function alleles for the CYP2C19 gene. A warning box recommends testing patients for CYP2C19 poor metabolizer genotypes. The tests available for CYP2C19 genotyping are expensive, time consuming, and laborious. Here, a new method was developed to improve genotyping of CYP2C19. The proposed method is cheaper and time-saving compared to available methods as it includes only one step for genotyping.

The newly developed method aims to facilitate the genotyping of the CYP2C19*2 SNP by a newly designed SuperSelective primer. Accordingly, this study aims to establish and prove that our genotyping algorithm is able to differentiate the CYP2C19*2 SNP carrier allele from non-carrier allele in a fast and efficient way to classify patients as either poor metabolizers or normal metabolizers of clopidogrel.

The specific objectives of the study are:

- To design a SuperSelective primer set and optimize assay conditions.
- To discriminate between CYP2C19*2 genotypes using real-time PCR.
- To study the frequency of CYP2C19 poor metabolizer genotypes in collected samples.

CHAPTER 3

3 Materials and methods

3.1 Sample Collection

Blood samples were collected from eighty blood donor volunteers who visited the blood bank at the Patient's Friends Society Hospital (Al-Ahli Hospital) during Jan/2020 to Feb/2021. The 81 blood donors were 15 females and 66 males and all of them were adults (19-60 years). The blood samples were collected after obtaining written informed consent (see appendix 7.2).

3.2 Ethical Approval

The present study was ethically approved by Palestine Polytechnic University Ethical Review Committee (No. EH 01/2019) (see appendix 7.1).

3.3 DNA Isolation and Purification

From every participant, 3 ml blood was withdrawn and collected on ethylenediaminetetraacetic acid (EDTA) tubes and directly transported to the Palestine-Korea Biotechnology Center for DNA extraction. Genomic DNA (gDNA) was extracted from 200 ul of buffy coat using a commercially available kit (QIAamp® DNA Mini Kit Cat. No. 51304). The quality and quantity of gDNA were measured and evaluated by Implen NanoPhotometer®. The gDNA samples were either used directly or stored at -20°C for further analysis.

3.4 Primer Design

Primers were designed to amplify DNA containing the region of interest in CYP2C19 (NG_008384.3: 22474-24454). Forward primer (FP), reverse primer (RP), as well as forward sequencing (FS) and reverse sequencing (RS) primers were designed for amplification and sequencing.

Figure 3.1 is a schematic drawing showing the positions of the two SNPs associated with clopidogrel resistance, namely *rs4244285* and *rs4986893*. The positions and sequences of the FP and RP are also shown. Table 3.1 summarizes the primer characteristics Of all primers designed and used in the study.

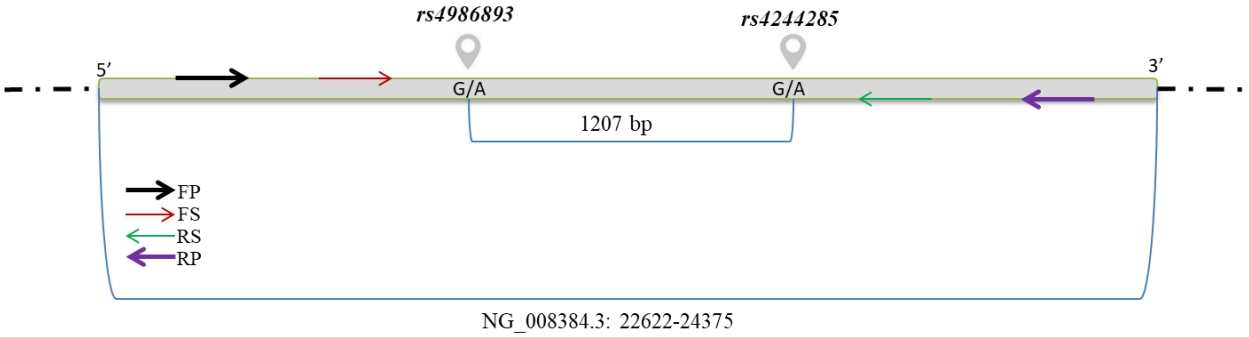

Figure 3.1: A schematic drawing of the DNA segment of the CYP2C19 gene that contains the two poor metabolizer SNPs (rs4244285 (CYP2C19*2), rs4986893 (CYP2C19*3). FP: Forward primer, RP: Reverse primer, FS: Forward sequencing, RS: Reverse sequencing.

Table 3.1 Primer sequences and their physicochemical characteristics.

Primer Name	DNA Sequence (5′-3′)	Annealing Temperature (°C)	GC(%)	Length (bp)	Reference
FP	AGGCTGTAATTGTTAATTCGA	50.0	33.3	21	This study
RP	GAGGGTTGTTGATGTCCATC	50.0	50.0	20	This study
FS	GCTCCATTATTTTCCAGAAACG	50.0	40.9	21	This study
RS	GTCCCGAGGGTTGTTGATGT	50.0	55.0	20	This study
FSS	CTTAGATATGCAATAATTTTCCCACCAACCA CCAGCCTTTCCCAG	60.0	49.9	45	This study
RP2	GTCAATGAATCACAAATACG	60.0	35.0	20	This study

3.5 Bioinformatics Analysis

The DNA segment of the CYP2C19 gene (NG_008384.3: 22474-24454) was used as a query to search the GenBank database for similar DNA sequences (paralogous genes) that might bind to the primers designed for CYP2C19. The search was carried out using the BLAST program with the following parameters: Homo sapiens (taxid:9606) for organism, using the somewhat similar sequences algorithm, and 100 maximum target sequences with 7 nucleotides as the word size. BLASTn alignment results were visualized by the SnapGene Viewer (GSL Biotech; available at snapgene.com). In order to avoid any mis-priming, all the primers used in this study were aligned with the CYP2C19 gene sequence alongside the other paralogous genes of the cytochrome P450 family.

3.6 SuperSelective Primer Design

SuperSelective primers were adapted from the original idea of Vargas et al (Vargas et al., 2016). These primers have a unique design that comprises three basic parts: a short 3' sequence (5-8 bp)

known as the “foot”, a unique sequence (15-20 bp) in the middle part known as the “bridge”, which is not complementary to the DNA target, and a sequence (15-40 bp) called the “anchor” at the 5' part of the primer that helps to stabilize primer binding (Marras et al., 2018; Vargas et al., 2016). A SuperSelective primer for *rs4244285* (CYP2C19*2) was designed with the following parameters: a 24 bp anchor sequence, a 14 bp bridge, a 13bp intervening sequence, and an 8 bp foot. The foot sequence has: 7 bp complementary to both the mutant [*rs4244285* carrier] and wild type [*rs4244285* non-carrier] targets, one interrogating nucleotide that is complimentary to the corresponding nucleotide in the mutant allele. Such a design is expressed in the following format:“24-14/13-6:1:1”. The physicochemical characteristics of the primer were determined only for the anchor sequence. The designed SuperSelective and reverse primers are listed above in Table 3.1.

The SuperSelctive primer sequences for (CYP2C19*2) were examined with the aid of the Mfold web server (Zuker, 2003) and the OligoAnalyzer computer program (Integrated DNA Technologies, Coralville, IA) to check for any internal hairpin structures and self-dimer or heterodimer formation with the reverse primer.

3.7 PCR Optimization

PCR amplification was carried out using 0.1 ng gDNA, 0.4 µm of each primer, 200 µM dNTPs mixture (Promega/Cat.U1330), 2 mM $MgSO_4$, 0.6 units of Taq DNA polymerase, 10X Taq reaction buffer, and PCR grade water in a 25 µl reaction volume according to the Taq polymerase kit (Hylabes/Cat. HTD0078) in a BioRad T100™ Thermal Cycler. The thermal cycling protocol was as follows: initial denaturation at 95°C for 5 min, 40 cycles of denaturing at 95°C for 45 sec, annealing at 50°C for 45 sec, and elongation at 72°C for 1:40 min. The final extension cycle was 72°C for 10 min. PCR products were then mixed with 3 µl of DNA loading dye (0.25% w/v bromophenol blue, 0.25% w/v xylene cyanol FF, and 30% v/v glycerol) and were resolved using a 1.5% agarose gel. The electrophoresis was carried out at 100 Volts (V) for approximately 30 minutes and the gel was visualized and documented using the Gel Doc™ XR+ molecular imager.

3.8 DNA Sequencing

PCR products with both CYP2C19 SNPs (NG_008384.3: 22474-24454) were sequenced as follows: PCR products were cleaned using a Sigma spin post reaction clean up column (Sigma Aldrich/Cat.S5059-70EA). Then, sequencing reactions were carried out using BigDye terminator v1.1 5x and were purified using BigDye Xterminator™ solution and SAM™ solution. All steps of the

sequencing protocol followed the kit manufacturer's instructions. Afterwards, purified samples were analyzed using the SeqStudio™ Genetic Analyzer System.

3.9 Real Time PCR Optimization

Reference gDNA samples confirmed by sequencing for *rs4244285* (CYP2C19*2) were used for real time optimization. The reference samples include heterozygous, homozygous, and wild-type genotypes. Real-time PCR amplification was carried out using 5 µl iTaq Universal SYBR Green supermix (Cat.1725120), 1 µl gDNA, 0.1 µl FSS, 0.5 µl RP2 primers, and 3.4 µl nuclease-free H_2O in a 10 µl reaction volume. The thermal cycling protocol consisted of a temperature hold at 95°C for 3 min followed by thermal cycling consisting of 50 cycles as follows: initial denaturation at 95°C for 10 sec, annealing at 60°C for 15 second, and elongation at 60°C for 30 sec using the CFX connect™ Real-Time system. SYBR green fluorescence intensity was measured at 530 nm at each cycle and plotted against temperature. Finally, the cycle threshold (Ct) was determined for each reaction.

After real-time PCR amplification, melt curve analysis was generated by increasing the temperature from 65°C to 95°C at 0.3°C increments. In addition, the amplification reaction mixtures were loaded on 1.5% agarose gel in order to exclude the possibility of interference by any nonspecific products in the mixtures that could either increase or decrease the Ct value.

3.10 Serial Dilution for Reference Samples and Copy Number Calculations

Serial dilution of reference samples (heterozygous (1*/2*), homozygous (2*/2*), and wild-type (1*/1*) for *rs4244285* ((CYP2C19*2) were carried out using TE buffer. DNA copy number per µl (copies/µL) was calculated for human gDNA using the following formula:

$$\boldsymbol{number\ of\ copies} = \frac{\boldsymbol{amount\ of\ dsDNA\ (ng) * Avagadro\ number\ (6.022 * 10^{23})}}{\boldsymbol{DNA\ tempate\ length\ (bp) * conversion\ factor\ of\ ng\ to\ g\ (1 * 10^{9}) * average\ molecular\ weight\ for\ dsDNA\ (660)}}$$

Stocks solutions of the reference samples were prepared with a genomic concentration equal to 33ng/µl (Std1). According to the copy number equation, 1 ng of gDNA contains approximately 330 copies, and based on that, a 5-fold serial dilution using TE buffer was prepared as explained in Table 3.2 and stored at -20°C for later use as standards in real-time PCR. Accordingly, real-time PCR amplification was carried out for each standard.

Table 3.2: Stock solutions from 33 ng/µl of Wild-type and heterozygous samples were determined by using Implen Nanophotometer and then 5-fold serially diluted by using TE buffer.

concentration	Std1	Std2	Std3	Std4	Std5	Std6
Copies/µl	10^4	10^3	10^2	10	1.0	0.1
ng/µl	33	3.3	0.33	-	-	-

3.11 Data Analysis and Genotypic Discrimination Algorithm

A genotypic discrimination algorithm was built for sorting the samples based on (CYP2C19*2) SNP carrier status. The algorithm was built based on discriminating the SNP carrier by using real-time PCR and the SuperSelective primer. To perform the discrimination, cycle threshold (Ct) for each sample was determined and compared with standards in the same real-time PCR amplification reaction. The real-time PCR amplification reaction contained both standards (wild-type and heterozygous or homozygous) for (CYP2C19*2) in each reaction mixture.

The genotyping discrimination algorithm was sequenced in several steps to finally classify the patients as suspected good responders or non-responders to clopidogrel. After real-time PCR amplifications were complete, the Ct of the standards were determined and the difference in Ct between the wild-type and heterozygous standards was calculated at the same gDNA concentration as follows:

$$\Delta Ct = |HeStd1\ Ct - WtStd1\ Ct|$$

The difference between the heterozygous standard and the wild type standard must be at least 7 cycles. If ΔCts = 7cycles, the amplifications were accepted and considered efficient and discriminatory. Otherwise, the amplifications were rejected and repeated.

Ct values of unknown samples were calculated and compared with the standards. If any sample had a Ct difference from the heterozygous standard of at least 7 cycles, the sample was suspected to be wild type and considered a non-carrier for *rs4244285* (CYP2C19*2). If the sample had a difference in Ct from the wild type standard of at least 7 cycles, the sample was suspected to be heterozygous for (CYP2C19*2) and considered a carrier for *rs4244285*. This genotypic discrimination algorithm is described in Figure 3.2.

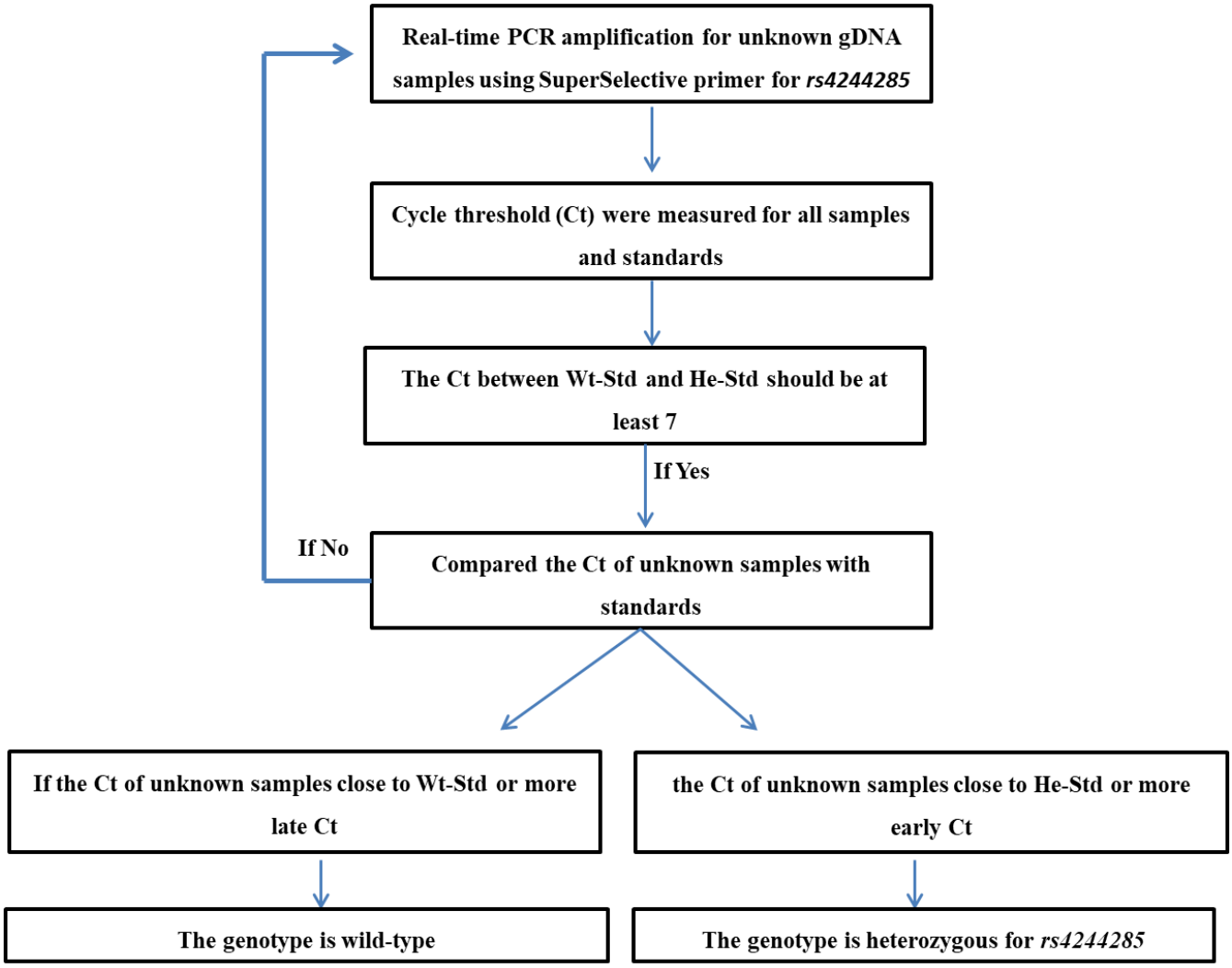

Figure 3.2: Schematic for genotypic discrimination algorithm.

CHAPTER 4

4 Results and discussions

4.1 Reference Samples Establishment

4.1.1 Primer Design

In order to exclude any non-specific amplification, all the primers were designed considering the possible interference from paralogous genes of the cytochrome p450 family. The FP and RP primers were designed to amplify the segment of interest of the CYP2C19, which contains the most frequent poor metabolizer SNPs, namely *rs4244285 and rs4986893.* Both primers were designed to target regions with significant mismatch with other paralogous genes of the cytochrome p450 family, including CYP2C9, CYP2C18, CYP2C8, and CYP2E1. Figure 4.1a shows a schematic diagram for the region of interest in the CYP2C19 gene. The schematic figure includes both SNPs and primers. In order to avoid poor sequencing quality at the beginning of each sequence, the FP was located 332 bp upstream the *rs4986893* SNP, and the RP is 177 bp downstream the (CYP2C19*2) SNP.

For a selective specific amplification The region of interest of CYP2C19 was aligned against the most similar paralogous genes of the cytochrome p450 family as shown in Figure 4.1b. The schematic alignment shows a high degree of homology observed around the (CYP2C19*2) SNP. Therefore, the FP was designed in a region of mismatch with all paralogous genes except CYP2C9, which is mismatched at only 6 nucleotides as shown in the bottom left of Figure 4.1. Due to the position of the (CYP2C19*2) SNP, we were obliged to locate the RP in a region that has high sequence similarity with 4 paralogous genes as shown in the bottom right of Figure 4.1. The CYP2C9 sequence corresponding to the RP primer has 3 mismatched nucleotides, while other genes have at least 6 mismatched nucleotides.

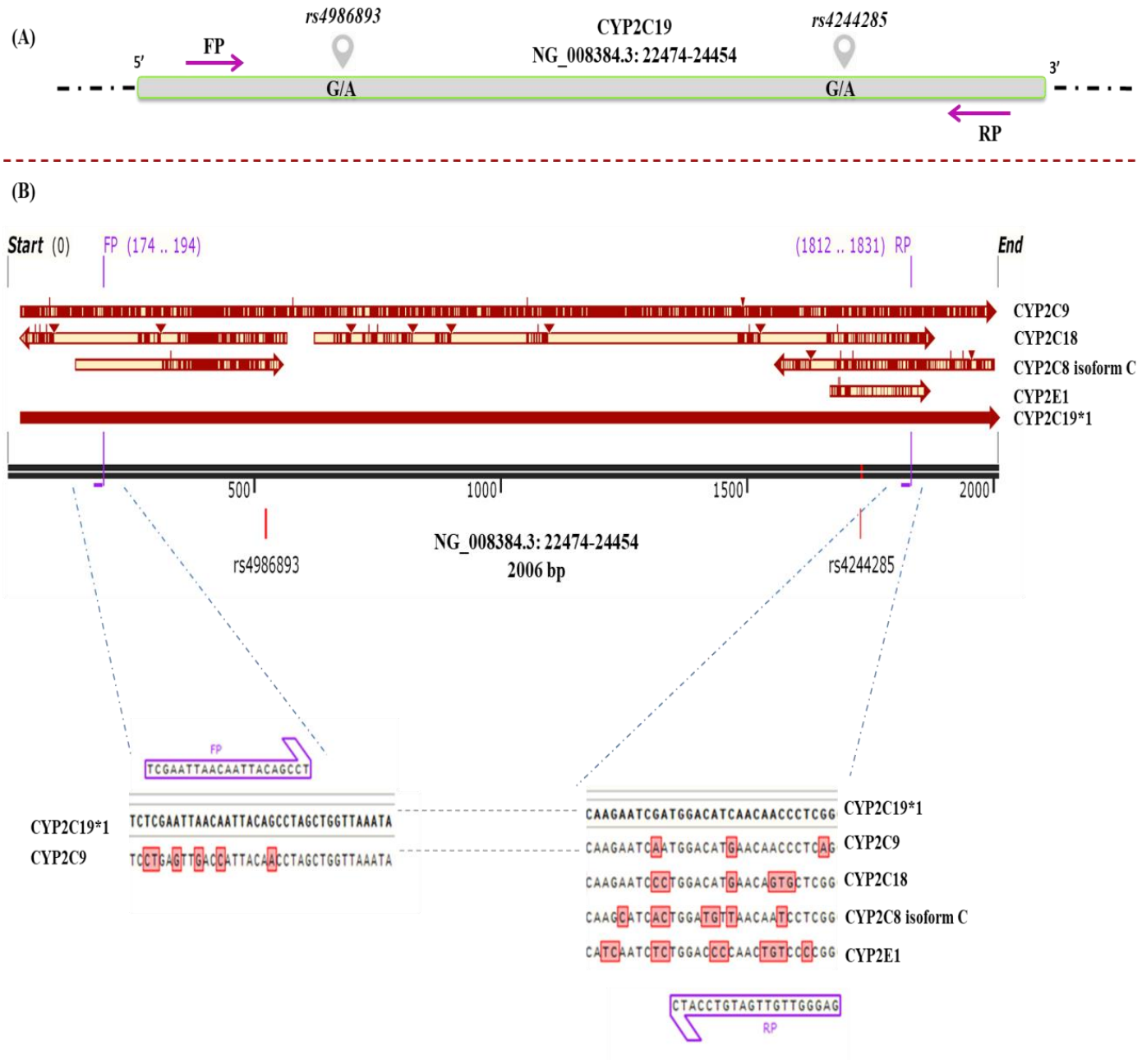

Figure 4.1: Schematic diagram and BLAST aligment sequences for the region of interset in CYP2C19. (4.1a) Schematic representaion of the region of interest in CYP2C19. The schematic includes the location of the SNPs (gray marks) and both primers (purple arrow). (4.1b) BLAST alignment of a portion of CYP2C19 showing the placement of the designed primers, *rs4244285* (CYP2C19*2) , *rs4986893* (CYP2C19*3), and the sequences of the paralogous genes.

4.1.2 Amplification and Sequencing

The target sequence fragment of the CYP2C19 gene was amplified by PCR and resolved by agarose gel electrophoresis (Figure 4.2). The amplified PCR product was purified from the agarose gel

and sequenced using sequencing primers. Sanger sequencing for *rs4244285* is shown in the electropherogram in Figure 4.3. Figure 4.3(A, B and C) shows the normal genotype (G/G), heterozygous (G/A), and homozygous (A/A) genotypes for *rs4244285* respectively.

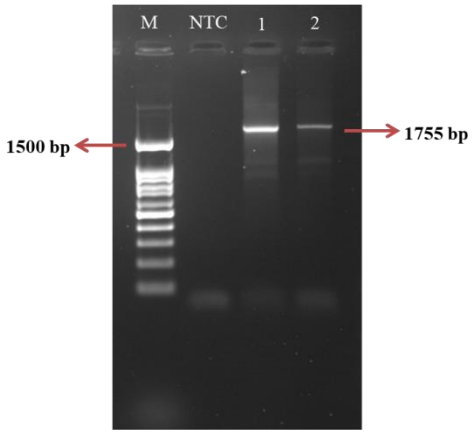

Figure 4.2: Gel electrophoresis analysis of CYP2C19 segment of interest. Lane 1-2 are the amplified products of CYP2C19 in 2 different samples .Product size is1755 pb. Lane NTC is a negative control template and lane M is the synthesized DNA reference marker.

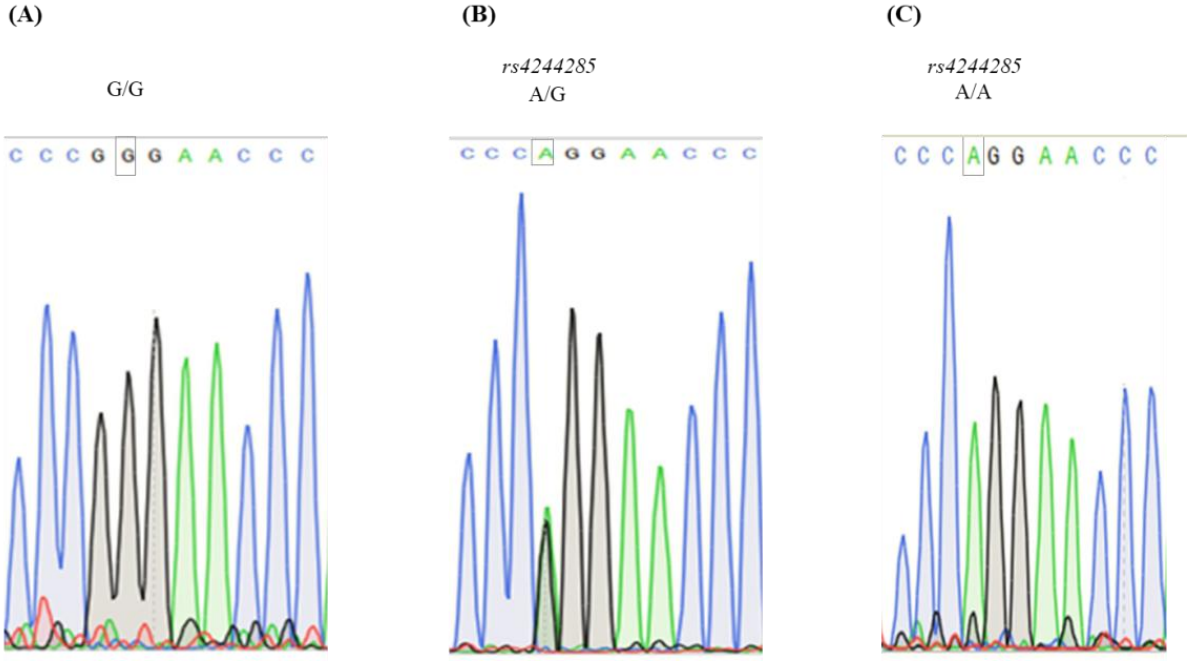

Figure 4.3: Sanger sequencing for *rs4244285.* (4.3a) Wild-type for *rs4244285* (normal genotype). (4.3b) Heterozygous for *rs4244285* (G>A). (4.3c) Homozygous for *rs4244285* (A/A).

Sanger sequencing was performed for both CYP2C19*2 and CYP2C19*3 for 20 gDNA samples. Among the 20 samples, all the samples were showed a wild-type genotype for CYP2C19*3 (data not

shown). So, SuperSelective primer design and optimization for CYP2C19*3 didn't carried out due to the missing of reference samples for CYP2C19*3 and the rarely frequent SNP among society.

4.2 SuperSelective Primer Design and Optimization

4.2.1 Primers Design

A SuperSelective primer is a multi-part primer consisting of only a DNA nucleotide sequence to be copied by DNA polymerase. A SuperSelective primer for (CYP2C19*2) was designed to be selective, able to amplify only the target sequence and rejects any non-target sequence even if only a small quantity of target sequence is present in the mixture. The primer selectivity for (CYP2C19*2) was achieved by the three parts of the primer that were designed with the most preferable specifications s mentioned by Vargas, Diana Y., et al (Vargas et al., 2016).

The first part of a SuperSelective primer is the anchor sequence. For the, it was designed to be significantly mismatched with other paralogous genes, as the anchor is similar to and functions like a conventional primer. In addition, the anchor sequence is the first part to hybridize to the intended sequence, so, the annealing temperature of the reaction is determined based on this hybridization. The annealing temperature for the anchor sequence is preferably 3 ℃ to 10℃ above the calculated Tm in order to enhance specific hybridization corresponding to the annealing temperature of the anchor sequence for (CYP2C19*2). Anchor sequences consisting of 24 bp did not anneal with other primer parts at the same annealing temperature.

The second part of a SuperSelective primer is the bridge sequence. It was designed to be completely mismatched with the intervening sequence and was designed to comply as much as possible with the most preferred specifications. The bridge consists of 14 bp and did not hybridize to any sequence in the human genome under the calculated annealing temperature. In addition, the bridge did not form any secondary structures with the reverse primer or any other SuperSelective primer parts which adds to the selectivity of that primer.

The third part of a SuperSelective primer is the foot sequence. The foot sequence for (CYP2C19*2) was designed to completely match with the intended sequence and has the most preferred specifications for foot length, ranging between 6-8 nucleotides with the mismatch nucleotide position to the unintended target at a penultimate position toward the 3' end. The foot sequence for (CYP2C19*2) consists of 8 nucleotides with a mismatch nucleotide at a penultimate position (-2).

Accordingly, when the primer is hybridized to a template target, the bridge sequence in the primer and the intervening sequence in the template form a single stranded “bubble’. The bubble functionally separates the efficient formation of the anchor hybrid from the formation of the foot hybrid. Circumference of the bubble is the sum of the bridge and intervening sequence plus 4 nucleotides (2 nucleotides hybridize from the anchor and 2 nucleotides hybridize from the foot) and it is equal to 31 bp (14 nucleotides of the bridge and 13 nucleotides from the intervening sequence plus 4 nucleotides).

The SuperSelective primer for CYP2C19*2 was designed for the exact location of the SNP, so, BLASTn was used for CYP2C19 analysis with other paralogous genes for sequence similarity. Figure 4.4 shows the schematic representation and BLASTn analysis for the region of interest of CYP2C19. Figure 4.4a is a schematic diagram showing the amplicon (240bp) amplified by the SuperSelective primer and reverse primer. The region of interest of CYP2C19 was aligned against the most similar paralogous genes as shown in Figure 4.4b. Three similar segments show high similarity with the CYP2C19 segment. These segments are portions of the CYP2C9, CYP2C8, and CYP2C18 genes. The forward SuperSelective (FSS) primer was designed to be mismatched with other similar genes as described in the bottom left of Figure 4.4b. RP2 was designed in a region totally mismatched with the CYP2C8 segment and partially mismatched with other paralogous genes as described in the bottom right of Figure 4.4b. RP2 was designed based on the primer design guidelines for real-time PCR (Rodríguez, Rodríguez, Córdoba, & Andrade, 2015) to perform efficient amplification, so, the RP2 position is the nearest suitable place for the FSS primer.

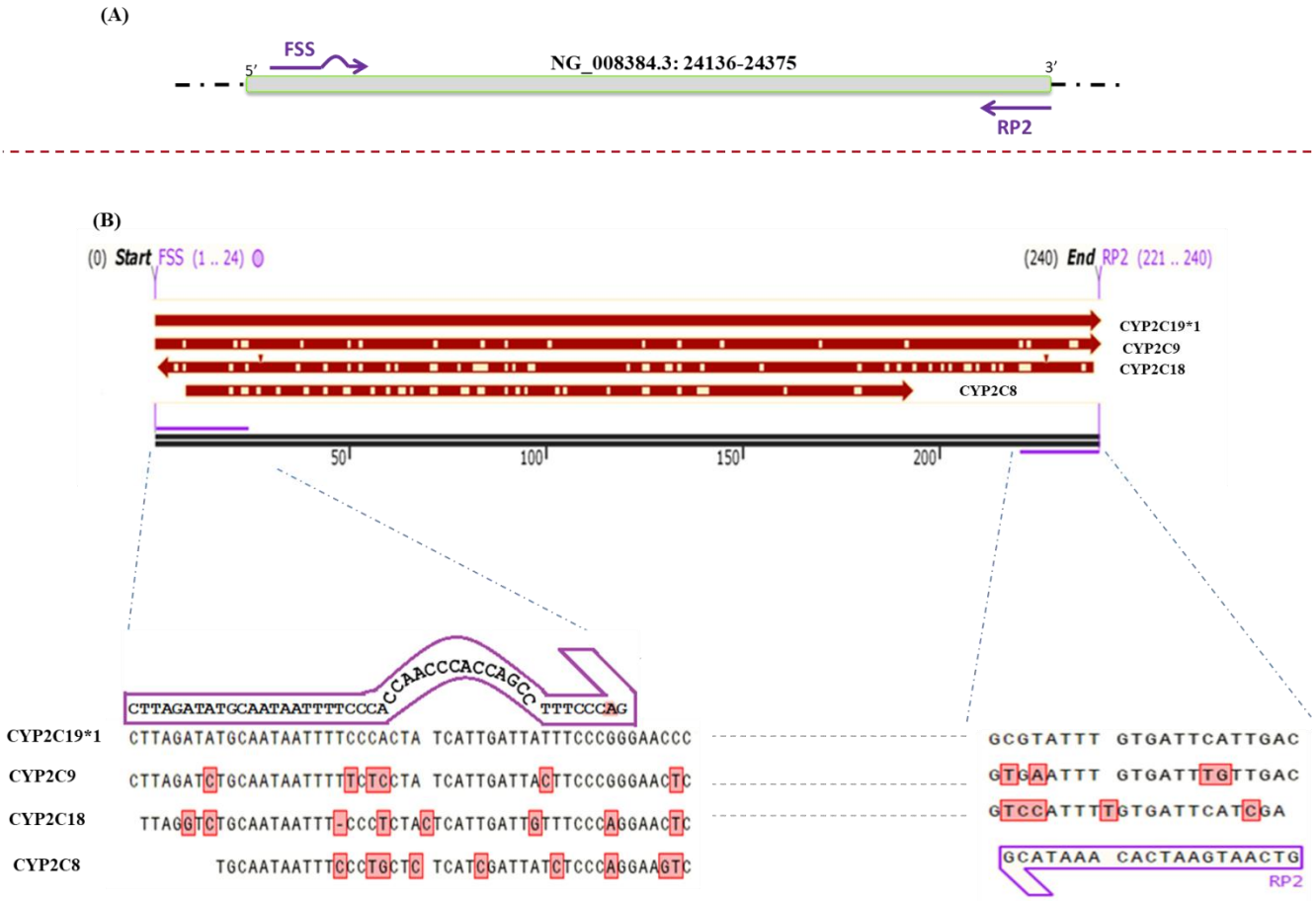

Figure 4.4: Schematic diagram and BLASTN aligment for CYP2C19 region of interest around *rs4244285*. Figure 4.4.a schematic diagram of the region of interest of CYP2C19 around *rs4244285*. The schematic includes the accession number of the region and the FSS and RP2 primers (purple arrow). Figure 4.4b shows the BLAST alignment of a portion of CYP2C19 showing the placement of the FSS and the RP2 primers and multiple sequence alignment against the paralogous genes CYP2C9, CYP2C18, and CYP2C8.

4.2.2 SuperSelctive Optimization

The SuperSelective primer was designed to selectively amplify DNA fragments containing the *rs4244285* SNP. The SuperSelective primer is referred to as CYP2C19*2 24-14/13-6:1:1, which indicates that the length of the anchor sequence is 24 bp, and the bridge length is 14 bp, that is across from an intervening sequence that is 13 bp in length, and the foot length is 8 bp with the interrogating nucleotide located at the penultimate position from the primer's 3'end as described in Figure 4.5a. Initial optimization of the Real-time PCR assay was carried out using the designed SuperSelective primer for CYP2C19*2 using confirmed reference samples that represent the wild-type *1/*1, the heterozygous *1/*2, and the mutant *2/*2 genotypes. The reference samples were designated wild type standard "Wt-Std", heterozygous standard "He-Std" and mutant standard "Ho-Std". Figure 4.5b shows the real-time PCR amplifications of He-Std1 and Wt-Std1, and Figure 4.5c shows the real-time PCR

amplifications of He-Std2 and Wt-Std2. The fluorescence intensity reflected by the SYBER green DNA dye used in the reactions, (y-axis) was measured during the elongation phase of each cycle (x-axis) which reflects the number of amplicons synthesized.

The cycle threshold was measured for each reaction that contained heterozygous and wild-type templates. He-Std1 shows early amplification in the reaction, with a Ct value equal to 21, whereas Wt-Std1 shows a later amplification with a Ct value equal to 30. The difference in Ct, between the two standards is 9 cycles, and enough to be considered discriminatory. He-Std2 and Wt-Std2 also showed the same exact discriminatory measure.

Regardless of the gDNA concentration, the Ct of the heterozygous standards shows early amplification when compared to the Ct of the wild-type standards. Accordingly, the SuperSelective primer for CYP2C19*2 is specifically selective to initiate the amplification of the heterozygous standards before the wild-type standards, due to the hybridization of the foot to the carrier allele *rs4244285* (CYP2C19*2) which increases the chances of encountering DNA polymerase compared to a non-allele carrier (CYP2C19*1) for *rs4244285*. Heterozygous standards for (CYP2C19*2) contain two alleles, one of them carries the SNP, which means that half of the DNA fragments are considered specific targets for the SuperSelective primer, so hybridization is initiated early. On the other hand, in the wild-type standards, there are no alleles for (CYP2C19*2), and hybridization of the SuperSelective primer is weak, which is demonstrated in the late Ct. The capability of the developed assay to discriminate wild-type genotype from heterozygous genotypes of the standard samples prompted us to examine the accuracy of the assay using a group of DNA samples with unknown genotypes for validation purpose. After performing the genotyping with the developed SuperSelective primer technique, the results were also confirmed and validated by Sanger sequencing method.

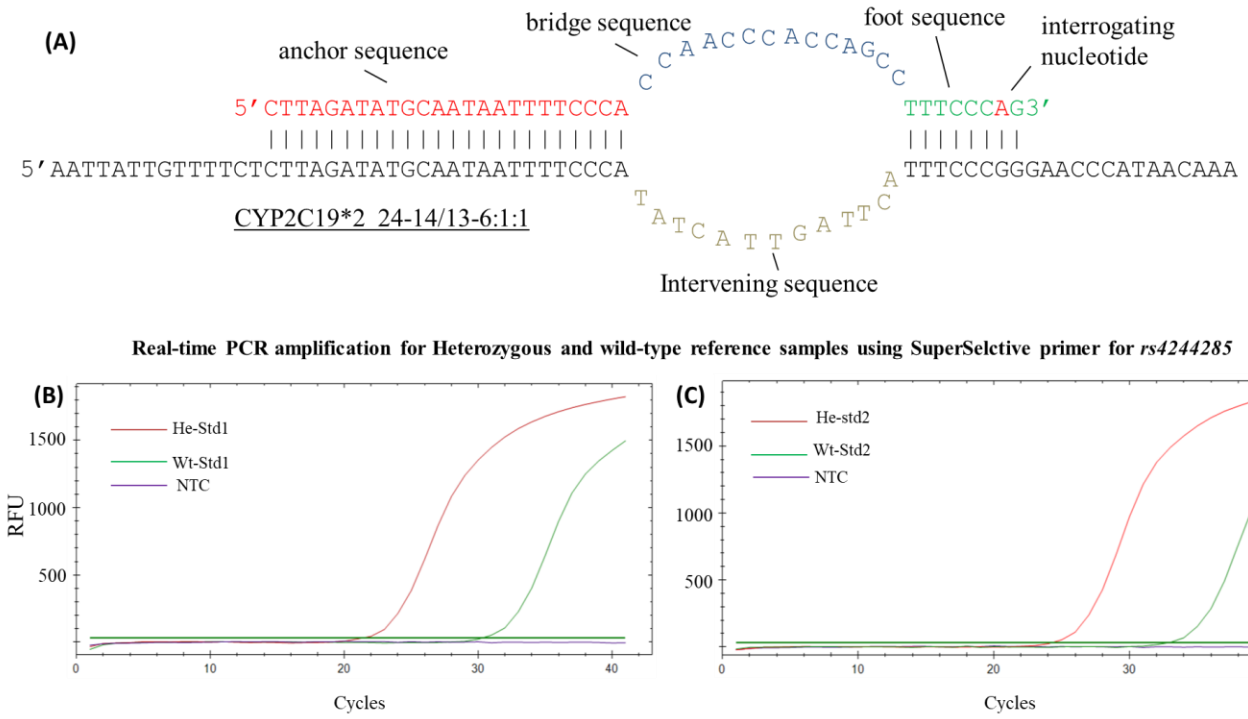

Figure 4.5: Structure of the SuperSelective primer for the detection of CYP2C19*2 in the presence of a wild-type sequence and a demonstration of real-time PCR assay. (4.5a) The SuperSelective primer CYP2C19*2 24-14/13-6:1:1 contains a 5'-anchor sequence, 3'foot sequence, and a bridge sequence that links the anchor to the foot sequence. (4.5b) and (4.5c) Real-time PCR assays employing the SuperSelective primer for the amplification of the wild-type and heterozygous standards for CYP2C19*2. The red line represents the heterozygous standard, the green line represents the wild-type standard, and the purple line represents the negative control template. The number of cycles is represent by the x-axis and relative flourescence intensity (RFU) is represent by the y-axis.

4.2.3 Effect of DNA concentration on SuperSelective performance Samples

A serial dilution of reference samples was prepared in order to evaluate the performance of the SuperSelective primer for *rs4244285* with different concentrations of gDNA template. To carry this out, real-time PCR assays employing the SuperSelective primer CYP2C19*2 24-14/13-6:1:1 were performed. Six reactions were initiated with different quantities of wild-type, heterozygous, and homozygous templates. The reactions contained 10^4, 10^3, 10^2,10^1, 1.0, and 0.1 copies of each standard prepared as described in Section 3.10 in the materials and methods. Real-time PCR amplifications were carried out for each standard, and the Ct on the x-axis was plotted against fluorescence intensity (RFU) on the y-axis as shown in Figure 4.6.

Wild-type standards were tested using the SuperSelective primer. Six reactions were initiated for all wild-type template copy numbers and were plotted on a green line (Figure 4.7). Wt-Std5 and Wt-Std6 did not produce any fluorescence until 50 cycles, while Wt-Std1, Wt-Std2, Wt-Std3 and Wt-Std4

showed amplification, but with a late Ct (> 34). Figure 4.7 shows real-time florescence results obtained with the CYP2C19*2 24-14/13-6:1:1 primer for reactions containing heterozygous standards that are plotted on red lines, except He-Std6 (which did not shows any amplification) and all homozygous standards, plotted on blue lines.

Cycle thresholds for Wt-Stds, He-Stds, and Ho-Stds were plotted on the y-axis against the genomic concentration (copy/µl) of each standard on the x-axis as shown in Figure 4.7. All standards showed an inverse relationship between the Ct and the concentration of the standards. In addition, the Ct distribution of wild type standards ranged from 34.62 to 42.0 (34.62 for Wt-Std1, 39.60 for Wt-Std2, 41.24 for Wt-Std3, and 42.0 for Wt-Std4). Ct distribution of He-Stds ranged from 25.02 to 36.71 (25.02 for He-Std1, 25.81 for He-Std2, 31.14 for He-Std3, 35.04 for He-Std4, and 36.71 for He-Std5) and the Ct distribution of Ho-Stds ranged from 24.07 to 37.37 (24.07 for Ho-Std1, 26.47 for Ho-Std2, 31.57 for Ho-Std3, 35.04 for Ho-Std4, 33.7 for Ho-Std5, and 37.37 for Ho-Std6). In view of that, real-time assays using the CYP2C19*2 24-14/13-6:1:1 primer show a clear differentiation between the wild-type and carrier genotypes of *rs4244285* (homozygous or heterozygous) in a series of different gDNA concentrations, while the difference in Ct between Ho-Stds and He-Stds show a somewhat similar distribution, especially in the three routine clinical concentrations of the DNA obtained in extraction (10^4 into 10^2 copies/µl), as shown in Figure 4.7.

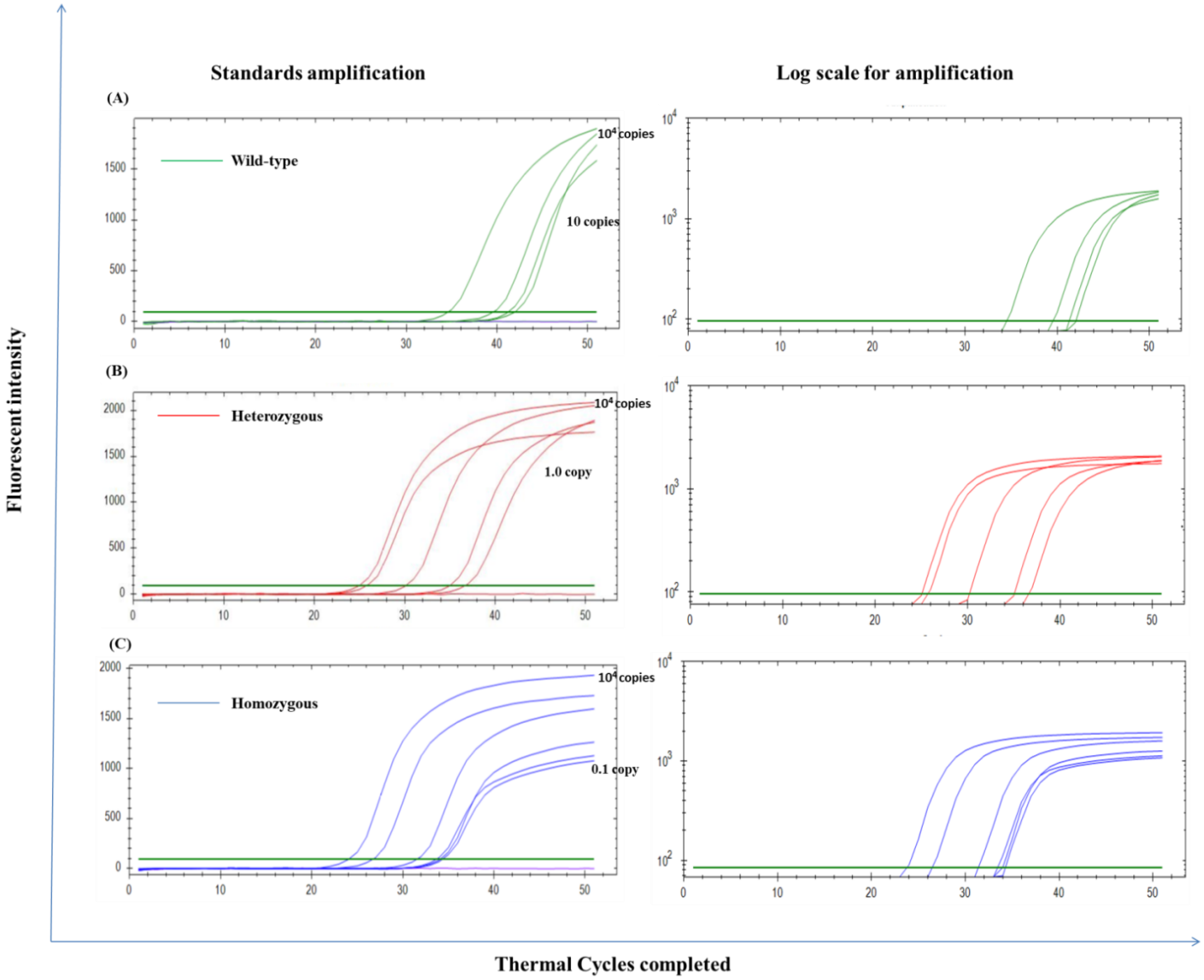

Figure 4.6: Real-time PCR amplification for CYP2C19*2 standards using the CYP2C19*2 24-14/13-6:1:1 primer. (4.6a) Real-time fluorescence results obtained for wild-type standards (green line). (4.6b) Real-time fluorecence results obtained for heterozygous standards (red line). (4.6c) Real-time fluorecence results obtained for homozygous standards (blue line). NTC in each reaction is represented by the purple line, and the reactions were carried out in triplicates.

Figure 4.8 shows the combined results for the cycle thresholds for all standards used to evaluate the employment of the CYP2C19*2 24-14/13-6:1:1 primer. Evidently, the difference in Ct between heterozygous standards and wild-type standards represents the observed cut-off value on which the distinction is based. However, the differences in Ct between the standards at the same genomic concentration were calculated and listed in Table 4.4. The lowest Ct observed out of the heterozygous and wild-type standards is 7.87 cycles. The smallest difference in Ct is considered the borderline on which the discrimination of this assay is based. Here, CYP2C19 genotyping for CYP2C19*2 was based

on the difference in Ct between the standards, and if the unknowns samples were analyzed using the CYP2C19*2 24-14/13-6:1:1 primer, the samples could be classified as either heterozygous or wild-type.

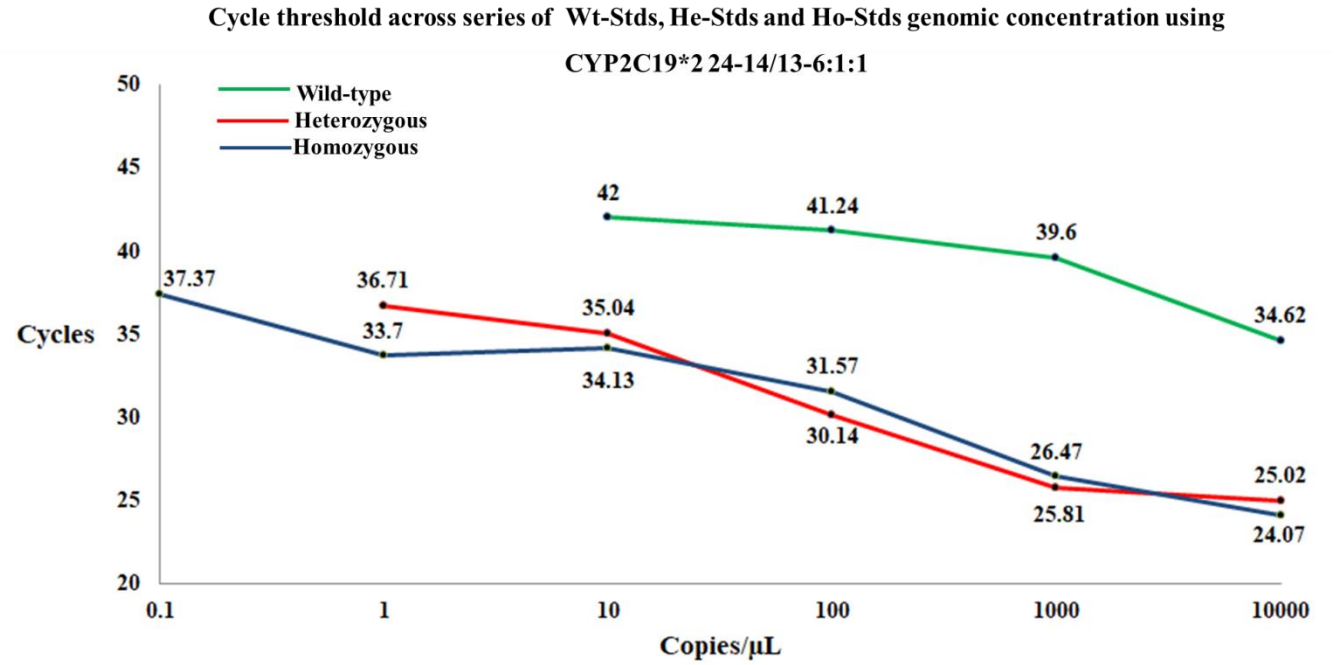

Figure 4.7: Cycle threshold values across a series of different genomic DNA concentrations using the CYP2C19*2 24-14/13-6:1:1 primer for wild-type standards, heterozygous standards, and homozygous standards. Ct values for each standard were plotted in black dots and a green line connects the wild-type standards, a red line connects the heterozygous standards, and a blue line connects the homozygous standards.

Table 4.1: The difference in Ct between wild-type and heterozygous standards at the same genomic concentration.

ΔCt = |HeStdn Ct – WtStdn Ct|				
	|HeStd1 Ct – WtStd1 Ct|	|HeStd2 Ct – WtStd2 Ct|	|HeStd3 Ct – WtStd3 Ct|	|HeStd4 Ct – WtStd4 Ct|
ΔCt	9.6	13.79	11.1	7.87

A weak discrimination was observed by employing CYP2C19*2 24-14/13-6:1:1 primer for the amplification of He-Std5 and Ho-Std5. The Ct for each standard are 36.71 and 33.7 respectively. The Ct difference between both standards (HoStd5-HeStd5) is a 3 thermal cycle. Three thermal cycle are not strong enough to classify the homozygous from the heterozygous, and thus, the genotyping was limited for sorting the samples with heterozygous or mutant genotypes as one class from the samples with wild-type genotypes as another class.

4.2.4 Validation of SuperSelective CYP2C19*2 Primer

Real-time assays were carried out using the CYP2C19*2 24-14/13-6:1:1 primer for samples from 81 healthy individuals in order to genotype individuals as heterozygous or wild-type in a blind experiment. Figure 4.8 shows the Ct for a group of genomic samples "with unknown genotype" in addition to the heterozygous and wild-type standards. Human genomic DNA concentrations of a concentration higher than 30ng/µl were manipulated to be within the same concentration as the standards, which ranged from 30 to 35 ng/µl, and were then amplified as shown in Figure 4.8a. Human genomic DNA concentrations of less than 30ng/µl were manipulated to be within the same concentration of the standards, which ranged from 3.0 to 9.0ng/µl, and were then amplified as shown in Figure 4.9b. Real-time PCR assays were carried out for all genomic DNA samples and the Ct for each sample was plotted in a scatter plot as shown in Figures 4.8c and 4.8d.

According to the genotypic discrimination algorithm, the difference in Ct between He-Std1 and Wt-Std1 is equal to 9.0 cycles, which following the rules of the algorithm, was considered discriminatory enough to classify the unknown samples as either wild-type or heterozygous. Fifteen samples, with genomic concentrations ranging between 30-35 ng/µl, were plotted in Figure 4.8c. Two of them showed early amplification, with a Ct near that of the He-Std1, and the rest of the samples had a late Ct, closer to that of Wt-Std1, with the exception of one sample, which had a Ct equal to 27, far away from He-Std1 by an average of 6 cycles. Four samples, with genomic concentrations ranging between 3.0 to 9.0ng/µl, were plotted in Figure 4.8d. One of them had a Ct approximately equal to that of He-Std2, and the rest of the samples had Ct values shifting toward that of Wt-Std2.

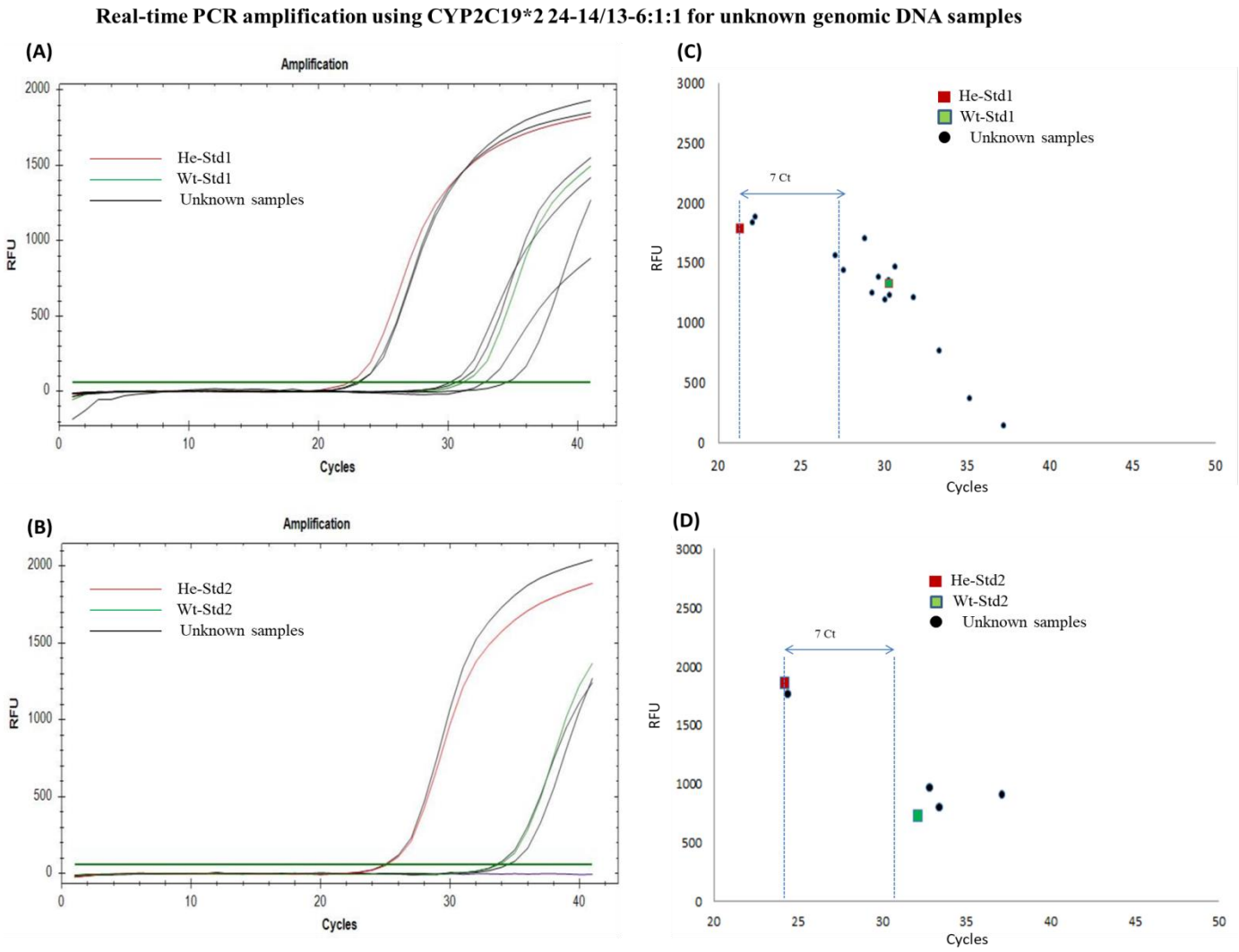

Figure 4.8: Real-time PCR amplifications using the CYP2C19*2 24-14/13-6:1:1 primer for genotyping unknown genomic DNA samples. (4.8a) Real-time PCR amplifications for a group of unknown samples (black curves) with genomic concentrations ranging between 30-35 ng/µl in the presence of the amplification of He-Std1 (red curve) and Wt-Std1 (green curve). (4.8b) Real-time PCR amplifications for a group of unknown samples (black curves) with genomic concentrations ranging between 3.0-9.0 ng/µl in the presence of the amplification of He-Std2 (red curve) and Wt-Std2 (green curve). (4.8c) Cycle threshold (x-axis) for the unknown samples (black dots), He-Std1 (red square,) and Wt-Std1 (green sequre), which were plotted against fluorescence intensity (RFU) (y-axis). (4.8d) Cycle threshold (x-axis) for the unknown samples in (A) (black dots), He-Std2 (red square) in (B), and Wt-Std2 (green square) were plotted againts fluorescence intensity (RFU) (y-axis).

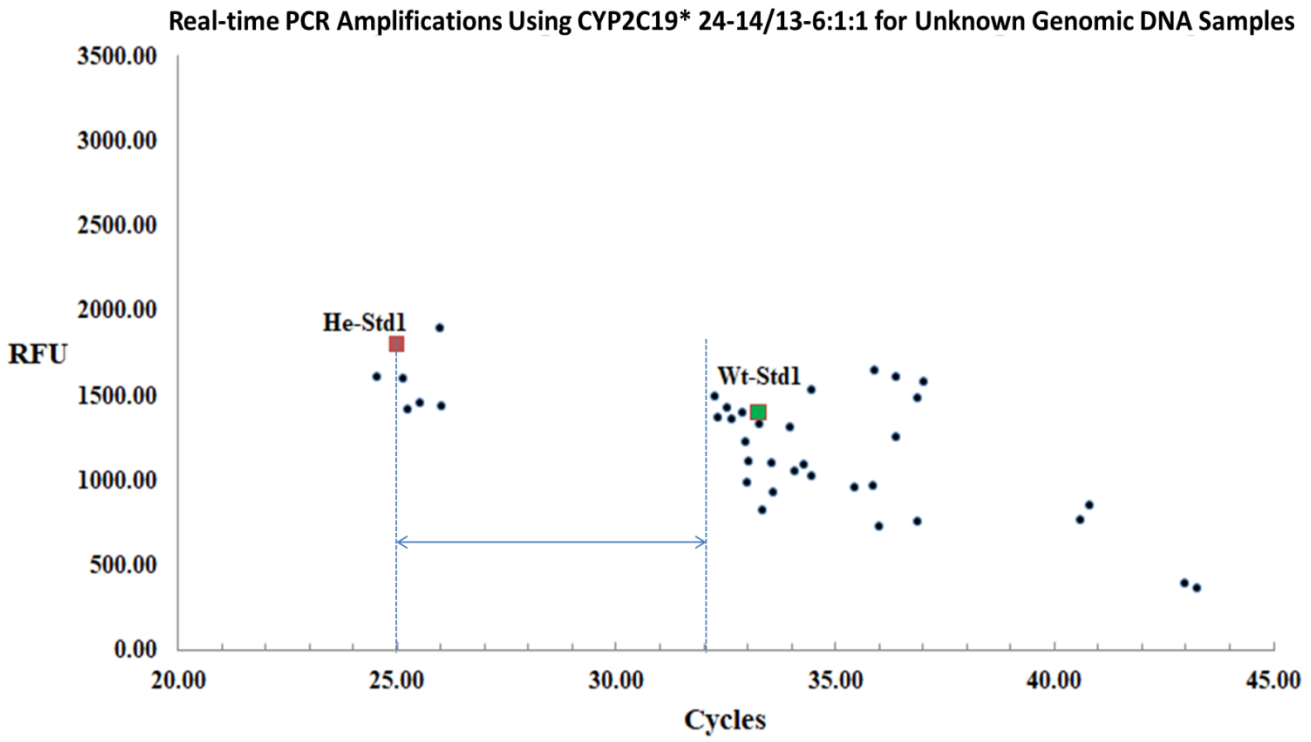

Figure 4.9: Real-time PCR amplifications using the CYP2C19*2 24-14/13-6:1:1 primer for genotyping unknown genomic DNA samples with different initial concentrations (3.0-155 ng/ul). Cycle threshold (x-axis) for the unknown samples (black dots), He-Std (red square), and Wt-Std (green square) were plotted againts fluorescence intensity (RFU) (y-axis).

Real-time PCR amplifications were carried out using the CYP2C19*2 24-14/13-6:1:1 primer for genotyping a group of human genomic DNA samples of concentration approximately equal to the concentration of the standards. Here, the same real-time PCR amplifications were carried out without modifying the sample concentration, and the genomic concentrations ranged from 3.0 to 155ng/µl, as shown in Figure 4.9. Forty samples were amplified, and the Ct for the samples was plotted as black dots against the fluorescence intensity (RFU). He-Std1 was plotted as a red square with a Ct equal to 24.02, and Wt-Std1 was plotted as a green square, with a Ct of approximately 33.5. Six samples demonstrated early amplification and were genotyped as heterozygous by using the SuperSelective primer, while the rest of the samples demonstrated slower amplification and were considered homozygous wild-type for CYP2C19*2.

Figure (4.10) shows the genomic DNA concentrations of 40 samples with unknown genotypes. According to the developed assay, six samples their genotypes were heterozygous. The samples with the highest genomic concentrations showed late Ct values and were wildtypes using our SuperSelective

primer. This experiment showed that genomic concentrations did not seem to affect the reactions performance in the displayed genomic concentration. To demonstrate this weak effect of the DNA concentration of the tested samples, the performance of CYP2C19*2 24-14/13-6:1:1 was evaluated across different genomic concentrations ranging from 3ng/µl to 160ng/µl. As can be seen in Fig 4. 10, the developed CYP2C19*2 genotyping assay showed a good specificity and discriminatory power across a relatively wide range of DNA concretions.

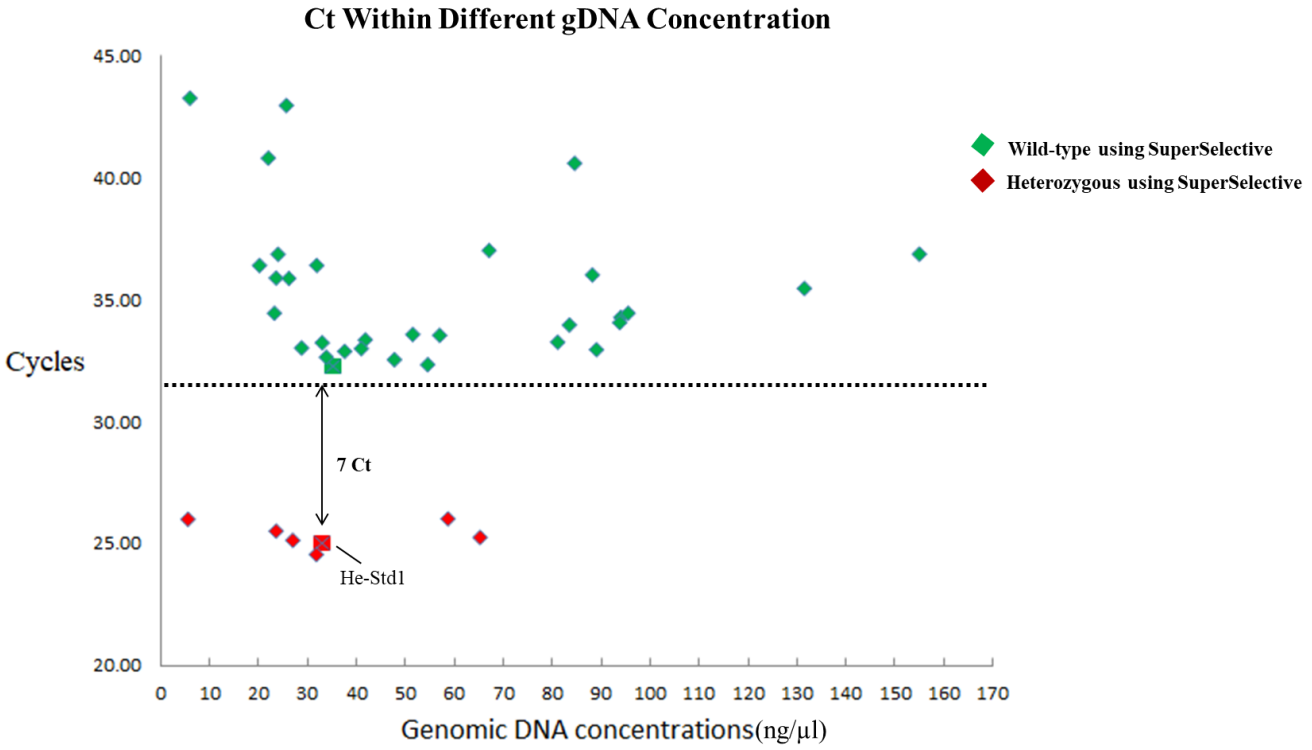

Figure 4.10: Genomic DNA samples with different concentrations and their Ct value. Red squares represent the heterozygous genotype and green squares represent the wild-type genotype determined using the SuperSelective primer.

All the samples in this study that were classified based on the genotypic discrimination algorithm as heterozygous were confirmed using Sanger sequencing, and eight randomly chosen homozygous wild-type samples were also confirmed by sequencing. In total, 20 samples (24.7%) from 81 samples were confirmed by sequencing, and their real-time PCR and Sanger sequencing results matched. The genotypic frequencies of the unknown samples were classified as follows: two samples (2.46 %) were homozygous for CYP2C19*2, ten samples (12.34%) were heterozygous for CYP2C19*2, and 69 samples (85.1%) were homozygous wild-type.

The CYP2C19*2 allele is associated with decreased enzyme activity, and its frequency varies across populations. In the studied group of samples, the CYP2C19*2 allele frequencies were found to be 2.46% poor metabolizers for clopedogril (CYP2C19*2/ *2), 12.34% intermediate metabolizers (CYP2C19*1/ *2), and finally 85% extensive metabolizers (CYP2C19*1/ *1). This CYP2C19*2 allele frequency can give a preliminary estimate of the presence of the CYP2C19*2 allele in the Palestinian population. However, further studies with more extensive number of samples are needed to address this issue to evaluate the actual allele frequency in the population.

The CYP2C19*2 allele frequency demonstrates interethnic variability. This allele is most frequently found in European and African populations (Kurose, Sugiyama, Saito, & pharmacokinetics, 2012), and in Middle Eastern populations CYP2C19*2 is considered the most frequent poor metabolizer allele with a frequency reaching up to 13.46% (Fricke-Galindo et al., 2016). CYP2C19 genotyping for the poor metabolizer genotype is expected to be useful to improve the quality of health care given during treatment with drugs that rely in their metabolism activity on CYP2C19 specifically clopedogril.

CYP2C19*3 allele frequency also showed interethnic variability. The studies showed that the more frequent of CYP2C19*3 present in Native Oceanians (14.42 %) and East Asians (6.89 %) across the world and was rarely found in the rest of the ethnic group (Fricke-Galindo et al., 2016). None of the poor metabolizer genotype were detected in a group of Saudi Arabians except the CYP2C19*2 (Al-Jenoobi et al., 2013).

4.3 The Specificity of SuperSelective CYP2C19*2 Primer

As SYBR green is a non-specific dye that has the ability to bind to double-stranded DNA, it is important to rule out any interference from non-specific amplicons. In order to ensure that the generated SYBR green fluorescence signal is produced by the specific PCR amplicon and not due to any PCR artifacts, a melting curve analysis was carried out after completing the SuperSelective primer based assay. A melting curve was generated by increasing the temperature in small increments and monitoring the fluorescent signal at each step. The produced fluorescent signal at each step is plotted as a function of temperature. A melting-curve analysis was generated after each reaction carried out using the CYP2C19*2 24-14/13-6:1:1 primer in order to evaluate the specificity of the primer for the target amplicon and to exclude any non-specific products, such as primer-dimers, in the reaction.

Figure 4.11 shows a real-time PCR amplification for heterozygous standard, homozygous standard, and wild-type standard, and the melt-curve analysis generated after the completion of each reaction. The melt-curve analyses showed that only one product was amplified without the presence of non-specific amplification or the formation of primer-dimers. In addition, the amplification products were loaded for agarose gel electrophoresis to determine the size of the amplification products as shown in Figure 4.11c. The target amplicon resulting from real-time PCR employing the CYP2C19*2 24-14/13-6:1:1 primer is 240 bp in length, as observed in the agarose gel.

As demonstrated by the melt curve analysis and agarose gel electrophoresis, the SuperSelective primer is specific for the amplification of the gene target, and only the target amplicon was amplified, without the presence of non-specific products or primer-dimer formation.

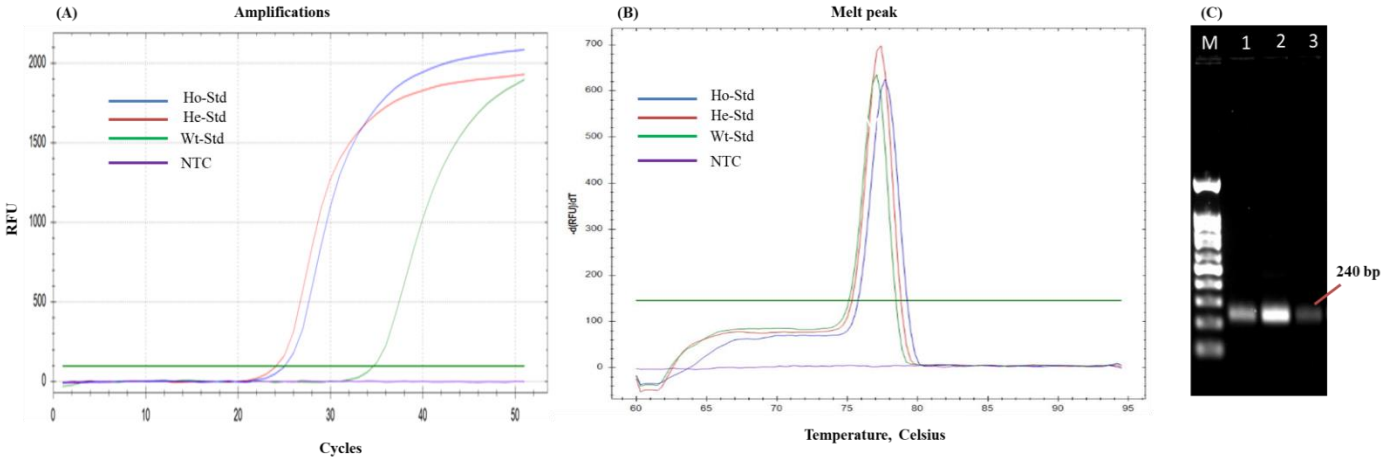

Figure 4.11: Validation and specificity of real-time PCR amplification using the CYP2C19*2 24-14/23-6:1:1 primer. (4.11a) Real-time PCR amplification for heterozygous standard (red curve), homozygous standard (blue curve), and wild-type standard (green curve). (4.11b) Melt curve analysis for heterozygous standard (red peak), homozygous standard (blue peak), and wild-type standard (green peak). (4.11c) Agarose gel electrophoresis for the heterozygous standard (lane 1), homozygous standard (lane 2), and wild-type standard (lane 3), and lane M is the product of the synthesized DNA reference marker.

Multiple molecular approaches were used for CYP2C19 genotyping, including RFLP (Ohkubo et al., 2006), ARMS (Kumar & Cannon, 2009), and DPO (Chun et al., 2007). Primers used in each assay are non-compatible with real-time PCR. The only method used for CYP2C19 genotyping and compatible with real-time PCR is High resolution melting curve analysis (HRM). HRM (Zhang et al., 2017) is a closed tube method able to identify the genotype and is considered a rapid, high-throughput, and convenient assay. HRM genotyping analysis can be influenced by several factors, including the

quality of the DNA, as well as quantity, amplicon size, salt concentration used in the reaction, in addition to the need for melt software analysis for genotyping analysis.

The primer length between 18-24 bases tend to be sequence specific if the annealing of the PCR reaction is set approximately a few degrees below the primer Tm. In ARMS PCR method, two inner primers design to discriminate the wild-type allele from the mutant allele, for SNP detection by ARMS-PCR the primer design for the amplification of mutant allele is completely matches with the mutant allele and one bp mismatches with the wild type allele. In the ARMS-PCR amplification scenario, there is a good chance for the primer binding on the both templets (wild-type and mutant) and that decrease the power of allelic discrimination (Jin et al., 2020). While in SuperSelective primer, the 5' anchor part is complimentary for the mutant and wild-type sequence, the bridge part is mismatches with the target sequence and the 3' short foot completely matches with mutant target and one bp mismatches with wild-type. The foot length for CYP2C19*2 SuperSelective primer is equal to 8 bp and one bp that mismatches with wild-type reduces significantly the hybrid probability for wild-type templet. The one bp mismatches with the wild-type allele and its penultimate position in the foot made the discrimination is based on it, here in, the allelic discrimination power based on the interrogating nucleotides in the short length foot.

Real-time PCR results can either be quantitative or qualitative. In this study, a real-time assay was designed to evaluate the results as qualitative measurements. Additionally, the results can be evaluated in a closed tube system without gel electrophoresis. The main advantage of our CYP2C19*2 genotyping assay over the previously TaqMan real-time PCR assays is the use of the SYBR green, which is a ds-DNA specific dye. The use of this dye provides a low cost method to detect the amplicon during real-time PCR with the ability to generate a melt-curve analysis after the amplification.

Veragas et.al invented the SuperSelective primer for the quantitative detection of rare mutant targets (Vargas et al., 2016). They employed the SuperSelective primer BRAF V600E 24-14/14-5:1:1 for the detection of the mutant BRAF gene that substitutes the amino acid valine (V) by glutamic acid (E) at position 600 at the protein level. This mutation is associated with certain diagnoses such as melanoma (Ascierto et al., 2012) and hairy cell leukemia (Tiacci et al., 2011). Accordingly, they employed their SuperSelective primer for the quantitative detection of BRAF mutant targets in a plasmid containing DNA fragments from mutant and wild type templates. The BRAF V600E 24-14/14-5:1:1 SuperSelective primer was used to carry out amplification of 10^6 mutant templates in the presence

of 10^6 wild-type templates, and the results showed the difference in Ct between the wild-type and mutant template to be approximately 20 cycles.

The SuperSelective primer CYP2C19*2 24-14/13-6:1:1 in this study was employed for genomic DNA fragments obtained from clinical samples. Genomic DNA fragments in the clinical samples are highly variable from person to person, and from time to time in a given person, and depending on the approach used for DNA extraction (Chacon Cortes & Griffiths, 2014). Vergas et.al employed the SuperSelective primer EGFR L858R 24-14/14-5:1:1 for the quantitative detection of EGFR mutants in the digested human genomic DNA encoding the EGFR L858R mutation in the presence of digested human genomic DNA from wild-type human cells, and the Ct of the reaction was measured using a real-time assay. Accordingly, the designed assay for use of the SuperSelective primer with human genomic DNA from clinical samples is somewhat less selective than those employed for plasmids containing the mutant and wild-type DNA. In addition, the high Ct value still represented for wild-type template either for plasmid or clinical samples, and the difference in Ct between the wild-type and mutant templates, is considered discriminatory at constant genomic DNA concentration, which is approximately 11 cycles. Overall, the bifunctionality of the SuperSelective primer is specific and unaffected by the presence of DNA fragments from the rest of the human genome, and that is compatible with the results obtained from our study.

Another SuperSelective primer was designed and applied for allelic-specific detection of chronic lymphocytic leukemia. The assay distinguished the samples containing the wild-type allele from the mutant allele in the samples based on the difference in Ct (>7 cycles between the two alleles) (Wang et al., 2017). However, their study does not explain their reliance on the difference in Ct between the two alleles, and instead just mentions it.

SuperSelective primer designed in this experiment shows the ability to genotyping CYP2C19*2 in a singleplex assay. In contrast to TaqMan probe assays, the developed technique relies on using SYBR green dye. Using this nonspecific DNA dye reduces the cost of genotyping. The disadvantage of such nonspecific dye is the possibility to obtain false positive signals due non-target amplification. However, the design of the SuperSelective primer is to avoid such mispriming. One of the potential limitations of the SuperSelective-primer-based genotyping using SYBR green is difficulty to use it in a multiplex format. This limitation can be addressed using the High Resolution Melting Analysis approach.

The proposed assay in this study can be carried out in virtually any of the many spectrofluorometric thermal cyclers available in hospitals and laboratories around the world. The test is considered a rapid genotyping assay to distinguish PCI patients who are either poor or intermediate metabolizers, who have low or no benefit from clopidogrel treatment, from patient with normal CYP219 enzyme activity who can benefit from the drug. This affordable phramacogenetic test can be of a great value for hospitals with limited resources.

SuperSelective primer designed in this study was limited for genotyping the samples as wild-type and allele carrier for CYP2C19*2, and another molecular approaches could be applied to determine the sample as heterozygous or homozygous such as DNA sequencing. Whatever the patients have a homozygous or heterozygous genotype, the FDA recommends changing the clopidogrel into another anti-platelet therapy such as Prasugrel or Ticagrelor which made the test applicable for anti-platelet guided therapy for patients undergoing PCI(Johnson & Cavallari, 2013).

5 Conclusion

This research aimed to develop and adapt a new method for CYP2C19*2 genotyping. Based on a quantitative and qualitative analysis of our designed SuperSelective primer and its real-time PCR optimization, this SuperSelective primer can distinguish between samples carrying a wild-type allele from samples carrying a mutant allele in a fast, feasible, specific, precise, and more cost-effective assay. The results indicate the possibility to apply the assay on clinical samples and improve health care provided in the medical field.

To achieve the goals of this study, the procedure has been designed and implemented meticulously in each step. A great deal of bioinformatics analysis for the gene of interest was carried out for primer design, especially for the SuperSelective primer. The assay was then optimized using the prepared reference samples and subsequently the assay was applied for genotyping unknown samples. The assay optimization steps were performed in order to make it as sensitive, specific, and precise as possible, to obtain a well-designed genotypic discrimination assay.

This study clearly distinguishes between the carriers of a mutant allele from the carriers of a wild type allele based on the difference in Ct in real-time PCR. The study was limited to genotyping the samples as heterozygous or wild-type. Further research is needed to build an assay able to distinguish the heterozygous from homozygous genotype using real-time PCR utilizing a SuperSelective primer.

6 References

Abul-Husn, N. S., Obeng, A. O., Sanderson, S. C., Gottesman, O., & Scott, S. A. (2014). Implementation and utilization of genetic testing in personalized medicine. *Pharmacogenomics and personalized medicine, 7*, 227.

Al-Jenoobi, F. I., Alkharfy, K. M., Alghamdi, A. M., Bagulb, K. M., Al-Mohizea, A. M., Al-Muhsen, S., . . . toxicology. (2013). CYP2C19 genetic polymorphism in Saudi Arabians. *112*(1), 50-54.

Ascierto, P. A., Kirkwood, J. M., Grob, J.-J., Simeone, E., Grimaldi, A. M., Maio, M., . . . Mozzillo, N. J. J. o. t. m. (2012). The role of BRAF V600 mutation in melanoma. *10*(1), 1-9.

Budnitz, D. S., Lovegrove, M. C., Shehab, N., & Richards, C. L. (2011). Emergency hospitalizations for adverse drug events in older Americans. *New England Journal of Medicine, 365*(21), 2002-2012.

Cervinski, M. A., Schwab, M. C., Lefferts, J. A., Lewis, L. D., Lebel, K. A., Tyropolis, A. M., . . . Tsongalis, G. J. J. A. j. o. c. p. (2013). Establishment of a CYP2C19 genotyping assay for clinical use. *139*(2), 202-207.

Chacon Cortes, D. F., & Griffiths, L. J. J. o. B. S. f. A. M. (2014). Methods for extracting genomic DNA from whole blood samples: current perspectives. *2014*(2), 1-9.

Chun, J.-Y., Kim, K.-J., Hwang, I.-T., Kim, Y.-J., Lee, D.-H., Lee, I.-K., & Kim, J.-K. J. N. a. r. (2007). Dual priming oligonucleotide system for the multiplex detection of respiratory viruses and SNP genotyping of CYP2C19 gene. *35*(6), e40.

Costa, I., Santos, T., Bento, I., Siqueira, B., Barbosa, A., Silva, K., . . . Morais, M. J. G. M. R. (2019). RFLP-PCR is more efficient than ARMS-PCR for identifying CYP2C19* 2 polymorphism in atherosclerotic patients. *18*(3), 1-5.

De Morais, S., Wilkinson, G. R., Blaisdell, J., Meyer, U. A., Nakamura, K., & Goldstein, J. A. J. M. p. (1994). Identification of a new genetic defect responsible for the polymorphism of (S)-mephenytoin metabolism in Japanese. *46*(4), 594-598.

De Morais, S., Wilkinson, G. R., Blaisdell, J., Nakamura, K., Meyer, U. A., & Goldstein, J. A. J. J. o. B. C. (1994). The major genetic defect responsible for the polymorphism of S-mephenytoin metabolism in humans. *269*(22), 15419-15422.

Duta-Cornescu, G., Simon-Gruita, A., Constantin, N., Stanciu, F., Dobre, M., Banica, D., . . . Stoian, V. J. B. (2009). A comparative study of ARMS-PCR and RFLP-PCR as methods for rapid SNP identification. *14*(6), 4845-4850.

Ferguson, R. J., De Morais, S. M., Benhamou, S., Bouchardy, C., Blaisdell, J., Ibeanu, G., . . . Therapeutics, E. (1998). A new genetic defect in human CYP2C19: mutation of the initiation codon is responsible for poor metabolism of S-mephenytoin. *284*(1), 356-361.

Food, & Administration, D. (2010). FDA drug safety communication: reduced effectiveness of Plavix (clopidogrel) in patients who are poor metabolizers of the drug [safety announcement]. March 12, 2010. In.

Fricke-Galindo, I., Céspedes-Garro, C., Rodrigues-Soares, F., Naranjo, M., Delgado, A., De Andrés, F., . . . LLerena, A. J. T. p. j. (2016). Interethnic variation of CYP2C19 alleles,'predicted'phenotypes and 'measured'metabolic phenotypes across world populations. *16*(2), 113-123.

Ghasemi, Z., Hashemi, M., Ejabati, M., Ebrahimi, S. M., Manjili, H. K., Sharafi, A., & Ramazani, A. J. A. j. o. m. b. (2016). Development of a high-resolution melting analysis method for CYP2C19* 17 genotyping in healthy volunteers. *8*(4), 193.

Hasan, M. S., Basri, H. B., Hin, L. P., & Stanslas, J. J. I. J. o. N. (2013). Genetic polymorphisms and drug interactions leading to clopidogrel resistance: why the Asian population requires special attention. *123*(3), 143-154.

Herrlin, K., Massele, A. Y., Jande, M., Alm, C., Tybring, G., Abdi, Y. A., . . . Therapeutics. (1998). Bantu Tanzanians have a decreased capacity to metabolize omeprazole and mephenytoin in relation to their CYP2C19 genotype. *64*(4), 391-401.

Hui, W., Zhang, S., Zhang, C., Wan, Y., Zhu, J., Zhao, G., . . . Li, N. J. N. (2016). A novel lateral flow assay based on GoldMag nanoparticles and its clinical applications for genotyping of MTHFR C677T polymorphisms. *8*(6), 3579-3587.

Ibeanu, G. C., Goldstein, J. A., Meyer, U., Benhamou, S., Bouchardy, C., Dayer, P., . . . Therapeutics, E. (1998). Identification of new human CYP2C19 alleles (CYP2C19* 6 and CYP2C19* 2B) in a Caucasian poor metabolizer of mephenytoin. *286*(3), 1490-1495.

Jin, C., Li, Z., Zheng, X., Shen, K., Chao, J., Dong, Y., . . . Zhu, W. J. J. o. C. L. A. (2020). Development and validation of T-ARMS-PCR to detect CYP2C19* 17 allele. *34*(1), e23005.

Johnson, J. A., & Cavallari, L. H. J. P. r. (2013). Pharmacogenetics and cardiovascular disease—implications for personalized medicine. *65*(3), 987-1009.

Johnson, J. A., Gong, L., Whirl-Carrillo, M., Gage, B. F., Scott, S. A., Stein, C., . . . Pirmohamed, M. (2011). Clinical Pharmacogenetics Implementation Consortium Guidelines for CYP2C9 and VKORC1 genotypes and warfarin dosing. *Clinical Pharmacology & Therapeutics, 90*(4), 625-629.

Johnson, J. A., Gong, L., Whirl-Carrillo, M., Gage, B. F., Scott, S. A., Stein, C., . . . Therapeutics. (2011). Clinical Pharmacogenetics Implementation Consortium Guidelines for CYP2C9 and VKORC1 genotypes and warfarin dosing. *90*(4), 625-629.

Jorga, A., Holt, D., & Johnston, A. (2004). *Therapeutic drug monitoring of cyclosporine.* Paper presented at the Transplantation proceedings.

Kim, K. A., Song, W. K., Kim, K. R., Park, J. Y. J. J. o. c. p., & therapeutics. (2010). Assessment of CYP2C19 genetic polymorphisms in a Korean population using a simultaneous multiplex pyrosequencing method to simultaneously detect the CYP2C19* 2, CYP2C19* 3, and CYP2C19* 17 alleles. *35*(6), 697-703.

Kumar, A., & Cannon, C. P. (2009). *Acute coronary syndromes: diagnosis and management, part I.* Paper presented at the Mayo Clinic Proceedings.

Kurose, K., Sugiyama, E., Saito, Y. J. D. m., & pharmacokinetics. (2012). Population differences in major functional polymorphisms of pharmacokinetics/pharmacodynamics-related genes in Eastern Asians and Europeans: implications in the clinical trials for novel drug development. *27*(1), 9-54.

Livak, K. J. J. G. a. b. e. (1999). Allelic discrimination using fluorogenic probes and the 5' nuclease assay. *14*(5-6), 143-149.

Ma, Q., Chen, G.-Z., Zhang, Y.-H., Zhang, L., & Huang, L.-A. J. C. m. j. (2019). Clinical outcomes and predictive model of platelet reactivity to clopidogrel after acute ischemic vascular events. *132*(9), 1053.

Maegdefessel, L., Azuma, J., Tsao, P. S. J. V. h., & management, r. (2010). Modern role for clopidogrel in management of atrial fibrillation and stroke reduction. *6*, 95.

Marras, S., Vargas-gold, D., Tyagi, S., & Kramer, F. R. (2018). Highly selective nucleic acid amplification primers. In: Google Patents.

Ohkubo, Y., Ueta, A., Ando, N., Ito, T., Yamaguchi, S., Mizuno, K., . . . Kurono, Y. J. J. o. h. g. (2006). Novel mutations in the cytochrome P450 2C19 gene: A pitfall of the PCR-RFLP method for identifying a common mutation. *51*(2), 118-123.

Pareek, M., & Bhatt, D. L. J. E. H. J. S. (2018). Dual antiplatelet therapy in patients with an acute coronary syndrome: up to 12 months and beyond. *20*(suppl_B), B21-B28.

Pruthi, R. K. (2013). *Review of the American College of Chest Physicians 2012 guidelines for anticoagulation therapy and prevention of thrombosis.* Paper presented at the Seminars in hematology.

Ramsjö, M., Aklillu, E., Bohman, L., Ingelman-Sundberg, M., Roh, H.-K., & Bertilsson, L. J. E. j. o. c. p. (2010). CYP2C19 activity comparison between Swedes and Koreans: effect of genotype, sex, oral contraceptive use, and smoking. *66*(9), 871-877.

Rifai, N., Horvath, A. R., Wittwer, C. T., & Park, J. (2018). *Principles and applications of molecular diagnostics*: Elsevier.

Rodríguez, A., Rodríguez, M., Córdoba, J. J., & Andrade, M. J. (2015). Design of primers and probes for quantitative real-time PCR methods. In *PCR Primer Design* (pp. 31-56): Springer.

Rudberg, I., Mohebi, B., Hermann, M., Refsum, H., Molden, E. J. C. P., & Therapeutics. (2008). Impact of the ultrarapid CYP2C19* 17 allele on serum concentration of escitalopram in psychiatric patients. *83*(2), 322-327.

Sabatine, M. S., Hamdalla, H. N., Mehta, S. R., Fox, K. A., Topol, E. J., Steinhubl, S. R., & Cannon, C. P. J. A. h. j. (2008). Efficacy and safety of clopidogrel pretreatment before percutaneous coronary intervention with and without glycoprotein IIb/IIIa inhibitor use. *155*(5), 910-917.

Scott, S. A., Sangkuhl, K., Shuldiner, A. R., Hulot, J.-S., Thorn, C. F., Altman, R. B., . . . genomics. (2012). PharmGKB summary: very important pharmacogene information for cytochrome P450, family 2, subfamily C, polypeptide 19. *22*(2), 159.

Shuldiner, A. R., O'Connell, J. R., Bliden, K. P., Gandhi, A., Ryan, K., Horenstein, R. B., . . . Gibson, Q. J. J. (2009). Association of cytochrome P450 2C19 genotype with the antiplatelet effect and clinical efficacy of clopidogrel therapy. *302*(8), 849-857.

Small, D. S., Farid, N. A., Payne, C. D., Konkoy, C. S., Jakubowski, J. A., Winters, K. J., & Salazar, D. E. J. C. p. (2010). Effect of intrinsic and extrinsic factors on the clinical pharmacokinetics and pharmacodynamics of prasugrel. *49*(12), 777-798.

Taubert, D., von Beckerath, N., Grimberg, G., Lazar, A., Jung, N., Goeser, T., . . . therapeutics. (2006). Impact of P-glycoprotein on clopidogrel absorption. *80*(5), 486-501.

Tiacci, E., Trifonov, V., Schiavoni, G., Holmes, A., Kern, W., Martelli, M. P., . . . Wells, V. A. J. N. E. J. o. M. (2011). BRAF mutations in hairy-cell leukemia. *364*(24), 2305-2315.

Vargas, D. Y., Kramer, F. R., Tyagi, S., & Marras, S. A. J. P. o. (2016). Multiplex real-time PCR assays that measure the abundance of extremely rare mutations associated with cancer. *11*(5), e0156546.

Wang, L., Fan, J., Francis, J. M., Georghiou, G., Hergert, S., Li, S., . . . Xiao, S. J. G. r. (2017). Integrated single-cell genetic and transcriptional analysis suggests novel drivers of chronic lymphocytic leukemia. *27*(8), 1300-1311.

Wittwer, C. T., Reed, G. H., Gundry, C. N., Vandersteen, J. G., & Pryor, R. J. J. C. c. (2003). High-resolution genotyping by amplicon melting analysis using LCGreen. *49*(6), 853-860.

Xiao, Z.-S., Goldstein, J. A., Xie, H.-G., Blaisdell, J., Wang, W., Jiang, C.-H., . . . Therapeutics, E. (1997). Differences in the Incidence of the CYP2C19Polymorphism Affecting the S-Mephenytoin Phenotype in Chinese Han and Bai Populations and Identification of a New RareCYP2C19 Mutant Allele. *281*(1), 604-609.

Xuhong, Y., Sinong, Z., Jianping, L., Yu, C., Juanli, Z., Chao, Z., . . . biotechnology. (2019). A PCR-lateral flow assay system based on gold magnetic nanoparticles for CYP2C19 genotyping and its clinical applications. *47*(1), 636-643.

Yamada, H., Dahl, M.-L., Lannfelt, L., Viitanen, M., Winblad, B., & Sjöqvist, F. J. E. j. o. c. p. (1998). CYP2D6 and CYP2C19 genotypes in an elderly Swedish population. *54*(6), 479-481.

Yang, C., Tavassolian, N., Haddad, W. M., Bailey, J. M., & Gholami, B. (2019). A Fast Parameter Identification Framework for Personalized Pharmacokinetics. *Scientific reports, 9*(1), 1-10.

Zhang, L., Ma, X., You, G., Zhang, X., & Fu, Q. J. S. r. (2017). A novel multiplex HRM assay to detect Clopidogrel resistance. *7*(1), 1-8.

Zhu, H.-J., Wang, X., Gawronski, B. E., Brinda, B. J., Angiolillo, D. J., Markowitz, J. S. J. J. o. P., & Therapeutics, E. (2013). Carboxylesterase 1 as a determinant of clopidogrel metabolism and activation. *344*(3), 665-672.

Zuker, M. J. N. a. r. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *31*(13), 3406-3415.

7 Appendix

7.1 Appendix A-Ethical approval

University Graduates Union
Palestine Polytechnic University (PPU)

رابطة الجامعيين/ محافظة الخليل
جامعـــة بوليتكنـك فلسـطين

DECISION OF PALESTINE POLYTECHNIC UNIVERSITY ETHICAL REVIEW COMMITTEE

(For work involving human subjects and data)

Ethical Application Reference:	EH 01 / 2019

Primary Researcher

Name:	Dr. Yaqoub Ashhab
Position and duty station:	Researcher / Palestine Korea Biotechnology Center
Email:	yashhab@ppu.edu

Title of research: Study the diagnostic value of a group of lung cancer genetic and epigenetic biomarkers within the Palestinian population

Decision of the Ethical Review Committee

Decision: Approved ☒ / Rejected ☐

Date: **25th November, 2018: Reviewed 10th of November 2019**

Notes/conditions:

This work complies with the WMA DECLARATION OF HELSINKI – ETHICAL PRINCIPLES FOR MEDICAL RESEARCH INVOLVING HUMAN SUBJECTS.

The Palestine Polytechnic University Research Council is solely authorized to consider requests for ethical approval for research conducted by its staff and/or in the name of the university that is related to the use of human subjects and data.

Chairman of the committee: (name and signature)

Murad Abusulaih

18/11/2019

فلسطين – الخليل – ص.ب: ١٩٨
مباني وادي الهرية: تلفاكس: ٠٢-٢٢٣٣٠٥٠ ٠٢-٢٢٣٠٠٦٨
مباني أبو رمان: تلفاكس: ٠٢-٢٢٣١٩٢١

P.O.Box: 198 , Hebron , Palestine
Wadi Al Hareih Campus: Telefax: 00970-2-2233050 , 2230068
Abu Roman Campus: Telefax: 00970-2-2231921
www. ppu.edu Email: info@ppu.edu

7.2 Appendix B-Informed consent

إقرار بالموافقة على المشاركة في تجربة سريرية

المركز الفلسطيني الكوري للتكنولوجيا الحيوية – جامعة بوليتكنك فلسطين

هدف الدراسة

تهدف هذه الدراسة للكشف عن وجود التغير الجيني في الإنزيم الخاص بتحويل دواء مانع لتجمع الصفائح الدموية عند مجموعة من مرضى القلب والأوعية الدموية. وجود هذا التغير الجيني يجعل المرضى لا يستجيبون لهذا الدواء مما يعرضهم لخطر الإصابة بجلطة دموية أخرى.

الباحث الرئيسي : الدكتور يعقوب الأشهب

الباحث المساعد : أ. صلاح الدين الجعبة

اسم المشارك	رقم الهوية	رقم العينة	العمر	الجنس	رقم الهاتف	مكان السكن

هناك عدة أمور يجب على المشترك بالبحث معرفتها قبل الموافقة على الانضمام للبحث وهي كالتالي :

1- **السماح للباحث بسحب عينة دم بمقدار 5 مل والسماح له بتسجيل البيانات الطبية اللازمة بالدراسة وذلك يشمل جميع الفحوصات المخبرية والاشارات السريرية**
2- **لا يترتب على المشترك أي مخاطر صحية بحيث يتم سحب العينة عن طريق فنيين مهرة**
3- **يستفيد المشترك من هذا البحث بمعرفة إن كان هناك أي تغيير جيني قد يؤثر على استجابة الدواء في حال وجود الطفرة**
4- **لا يتقاضى المشترك أي فائدة مالية من الانضمام إلى الدراسة**
5- **يتم الاحتفاظ بسرية المعلومات الخاصة بالمشتركين وذلك يشمل كل من المعلومات الشخصية والمعلومات الطبية الخاصة بالمشترك**
6- **لا يوجد أي شخص مخول بالوصول لمعلومات المشترك إلا الباحث الرئيسي والباحثين المساعدين**

موافقة أو توقيع المشارك في البحث :

حصلت على شرح مفصل عن الدراسة وأهدافها وإجراءاتها، ومنافعها، والمخاطر المحتملة وعن الحرية الكاملة للمشاركة. أفهم كل المعلومات التي قُدمت ووصلتني إجابة على كل أسئلتي. أوافق على أن أشارك في هذه الدراسة بطوعية وبدون أي نوع من الاجبار أو الضغوط.

الاسم	
التوقيع	
التاريخ	

في حال عدم قدرة المشارك على التوقيع بنفسه، نطلب من حضرتكم (المسؤول أو الوصي) التوقيع أدناه

الاسم (الوصي أو المسؤول)	
التوقيع	
التاريخ	

توقيع الباحث الرئيسي	
التاريخ	

انتهاء نموذج الموافقة