Development of Tetra-primer ARMS PCR Based Genotyping Method for A Clinically Important SNP CYP2C19*3 (rs4986893)

By:

Wasifa Ar Rahman

ID: 16136006

A Dissertation Submitted to BRAC University in Partial Fulfillment of the Requirements for The Degree of Bachelor of Science in Biotechnology

Department of Mathematics and Natural Sciences

BRAC University

August 2020

Declaration

It is hereby declared that

1. The thesis submitted is my original work while completing degree at Brac University.

2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.

3. The thesis does not contain material that has been accepted, or submitted, for any other degree or diploma at a university or other institution.

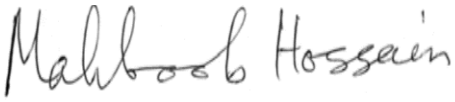
4. I have acknowledged all the main sources of help.

Student's Full Name & Signature:

Wasifa Ar Rahman

ID 16136006

Approval

<table>
<tr><td colspan="2">The thesis/project titled “Development of Tetra-primer ARMS PCR Based Genotyping Method for A Clinically Important SNP CYP2C19*3 (rs4986893)” submitted by
Wasifa Ar Rahman (ID 16136006)
of Spring, 2016 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of Bachelor of Science in Biotechnology on 26th of August 2020.</td></tr>
<tr><td colspan="2">Examining Committee:</td></tr>
<tr><td>Supervisor:
(Member)</td><td> 
_____
M. Mahboob Hossain, PhD
Professor, Department of Mathematics and Natural Sciences
Brac University</td></tr>
<tr><td>Co-Supervisor:
(Member)</td><td>_____
MD. Mahmudul Hasan Akash,
Lecturer, Department of Genetic Engineering and Biotechnology
Dhaka University</td></tr>
<tr><td>Program Coordinator:
(Member)</td><td>_____
Iftekhar Bin Naser, PhD
Assistant Professor, Department of Mathematics and Natural Sciences
Brac University</td></tr>
<tr><td>Departmental Head:
(Chair)</td><td>_____
A F M Yusuf Haider, PhD
Professor and Chairperson, Department of Mathematics and Natural Sciences
Brac University</td></tr>
</table>

Acknowledgement

First and foremost, I would like to express my thanks to almighty Allah because He has given me the opportunity, and strength to finish this research. I am also thankful for His blessings to my daily life, good health and healthy mind.

I acknowledge my esteem to Professor **A F M Yusuf Haider**, Chairperson of MNS Department, and Assistant Professor **Iftekhar Bin Naser**, Coordinator of the Biotechnology Program of MNS Department of BRAC University for allowing me and encouraging me to complete my undergraduate thesis.

My regards, gratitude, indebtedness, and appreciation goes to my respected supervisors **MD. Mahmudul Hasan Akash**, Lecturer of Department of Genetic Engineering and Biotechnology, Dhaka University and Professor **Dr. M Mahboob Hossain**, Coordinator of Microbiology Program of MNS Department, BRAC University for their constant supervision, constructive criticism, expert guidance, enthusiastic encouragement to pursue new ideas and never ending inspiration throughout the entire period of my research work. I would like to thank and express my deepest gratitude for guiding me in my report writing and providing time to time suggestions regarding the setting of my experimental designs, interpretation of results, and subsequence directions for the whole work without being a bit of impatient. It would have been impossible to submit my report without their cordial help.

I would like to extend my appreciation to the respective Lab officers Asma Binte Afzal and Md. Nazrul Islam for their suggestions and moral support during my work.

I also appreciate my thesis partners Afia Adiba Rupoma, Anika Tahsin, Rifaquat Ahmed for their kind cooperation and active support throughout my work.

Finally, I would like to extend my gratitude to the members of my family and friends for their prayerful concerns and supports.

Wasifa Ar Rahman

August 2020

Abstract

Treatment of a particular disease with satisfactory outcome has become quite a rare and challenging issue nowadays. This issue is mostly responsible for the polymorphism in the metabolizing gene which alters its metabolized drug function. To solve this problem, SNP genotyping is a great concept. The SNP used in this study rs4986893 present in the gene CYP2C19 causes loss of function to the associated enzyme. As a result, an individual carrying a single copy of the mutant allele becomes a poor metabolizer for a very important anti-coagulant drug, Clopidogrel. The study is conducted by the tetra primer ARMS PCR technique which is an allele specific PCR. The PCR utilize four sets of primer by differentiating homozygous and heterozygous product. Here, in the 3' end of the primordial the single base mismatch is added. This mutation causes the primordial to amplify a single allele. The ARMS PCR is a cost effective and easy handling procedure. To do the PCR, buccal cells were collected from various range of healthy individual by and their DNA is extracted in phenol-chloroform-isoamylalcohol extraction protocol. The extracted DNA was tested by ARMS PCR method to identify if the individual was carrying the SNP variant or not by resolving the bands in agarose gel electrophoresis. To complete the study with the population perspective of people carrying the SNP variant CYP2C19*3 more thermal cycle optimization step was needed for accomplishing further works.

Table of Contents

List of Figures

List of Tables

CHAPTER 1

INTRODUCTION

1.1 Pharmacology:

Pharmacology is the study of drug action. The systems in our bodies involve the interaction of chemical substances along with it to identify ways in which our biological systems affect drugs.

The umbrella of pharmacology is a wide one. It encloses uses of various segments such as the examination of the different classes of drugs, their therapeutic uses, their societal roles, and their mechanisms of action, among other things. In an instance of this how the drug molecules and drug receptors are interacting as well the consequences of these interactions are studied under the umbrella of pharmacology. (What is Pharmacology? n.d.)

These consist the use of pharmaceutical drugs in medicine, drugs used in scientific research, the study of drug actions within the overall health sciences, as well as the regulation and development of various pharmaceuticals.

Pharmacology sums up its entire study into two divisions. They are:

1. Pharmacokinetics
2. Pharmacodynamics

1.1.1 Pharmacokinetics:

Pharmacokinetics is the study of how individuals are differently affected by the same drug according to their genetic information. Pharmacokinetics, based on the Greek terms pharmakon (drug) and kinetikos (movement), is used to describe a compound's absorption, delivery, metabolism, and excretion. It refers to the passage of drugs through and out the body the timing of their absorption, bioavailability, delivery, metabolism, and excretion. An individual's type of reaction to a given drug depends on the drug's intrinsic pharmacological properties at its site of action (Turfus et al., 2017). Nonetheless, onset speed, intensity and response length typically depend on parameters such as:

- The rate and degree to which the drug is taken from its place of administration
- The rate and degree of the delivery of the drug to various tissues, including the site of action
- The rate at which the drug is withdrawn from the body (Derek et al., 2018).

1.1.2 Pharmacodynamics:

Pharmacodynamics is the study of how drugs affect the body. The most common mechanism is the drug's association with tissue receptors either found in the cell membranes or the intracellular fluid. The degree of receptor activation, and the subsequent biological reaction, is linked to the triggering drug concentration (the 'agonist'). The relation is defined by the dosage response curve measuring the dosage (or concentration) of the drug against its effect. (Lins et al., 1999).

1.1.3 Pharmacogenetics:

Pharmacogenetics is the study of understanding how genetic variation affects the effects of the drug treatment outcome. Although the words pharmacogenetics and pharmacogenomics are sometimes used interchangeably, pharmacogenetics usually refers to the effects of a single genetic marker whereas pharmacogenomics is broader in nature, referring to the collective impact of genome-wide heterogeneity for modulating the drug response profile of a person.

Pharmacogenetics can affect both the pharmacokinetics and drug pharmacodynamics. This variation has implications for dosage, therapeutic response, probability of side effects, and the possibility of reactions to hypersensitivity.

The single nucleotide polymorphism, or SNP, is the most common cause for genetic variation, and thus the cause for a pharmacogenetic approach to drug therapy.

1.1.4 Application of Pharmacogenetics:

- Pharmacogenetics is the study of changes in the sequence / characteristics of DNA and RNA as relating to drug reaction and toxicity.
- All proteins are products of the genes and many (maybe most) exhibit genetic polymorphism. The structure, composition, and / or concentration of proteins is determined by single-nucleotide polymorphisms (SNPs), gene deletions, and gene amplifications.
- The effects of genetic polymorphisms may result in either increased or decreased clinical effectiveness or increased or decreased toxic effects.
- Personalized treatment requires appropriate targeting (e.g., limiting the use of trastuzumab (Herceptin)) based on cancer phenotype (i.e., only those whose tumors over-express HER2).

- In medication reactions and toxicity, epigenetic factors bring additional complexity to personalized medicine.
- Medications that cause serious adverse reactions in conjunction with other products or certain conditions (e.g., pregnancy) are used today with adequate warnings and legal liability to the prescriber of the drugs. A strong responsibility for the correct use of pharmacogenetic knowledge should be applied in future.
- For an ever-expanding list of medications, the U.S. Food and Drug Administration (FDA) has released labeling regulations and publishes updated pharmacogenetic information.
- Knowledgeable, individualized medication or medication dose selection based on patient genetic information should replace the one-drug model and/or one-dose model fits all.
- Active medications previously discarded due to toxicity in certain patients can be useful when aimed at patients with a genetic profile suitable. It would minimize drug production costs, which eventually will reduce drug costs.
- The Consortium for Implementing Clinical Pharmacogenetics (CPIC) formulates and distributes practice guidelines to facilitate the conversion of pharmacogenetic knowledge into actionable prescription decisions. (Stan et al., 2011)

1.2 The drug; Clopidogrel:

Plavix, the brand name of the drug “Clopidogrel” is an antiplatelet agent. It functions by reducing the risk of myocardial infarction (MI) and stroke in patients with acute coronary syndrome (ACS) and in patients with atherosclerotic vascular disease (indicated by a recent MI or stroke, or established peripheral arterial disease. (Morgantown et al., 2017).

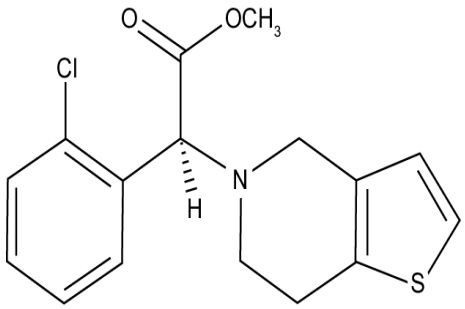


Figure 1.1: Chemical Structure of Clopidogrel

To treat or to prevent further occurrences of arterial thrombosis, Clopidogrel is given, which occurs due to a blood clot (thrombus) formation inside an artery. A transient ischemic attack (TIA) or ischemic stroke can be caused through thrombosis in the cerebral arteries in the brain. In the peripheral vessels and in the heart, thrombosis can cause respectively peripheral artery disease and in the coronary arteries can cause ACS. Clopidogrel (Platelet inhibitors) interrupt the formation of the thrombus, which involves the rapid recruitment and activation of platelets. (Dean et al., 2018).

1.2.1 Clopidogrel is a Prodrug:

After oral administration, clopidogrel is rapidly absorbed, clopidogrel is not detected in human plasma following to its extensive metabolism. Clopidogrel is a prodrug which is absorbed in the intestine and gets activated in the liver. (Savi et al., 1992).

The conversion of clopidogrel to its active metabolite requires two sequential oxidative steps:

1. The formation of 2-oxo-clopidogrel
2. The conversion of 2-oxoclopidogrel to the active metabolite R-130964 (clop-AM).

Enzymes (cytochrome P450) involved in the metabolism of clopidogrel are CYP1A2, CYP2B6, CYP2C9, CYP2C19 and CYP3A4/5. (Sangkuhl et al., 2010).

1.2.2 Mechanism:

- Clopidogrel converts into 2-oxo-clopidogrel by cytochrome P450 monooxygenase-dependent metabolism (in vitro)
- Following that, hydrolysis of 2-oxo-clopidogrel generates the active metabolite.

The study by (Kazui et al., 2010) concluded that CYP2C19 contributes substantially to both oxidative steps.

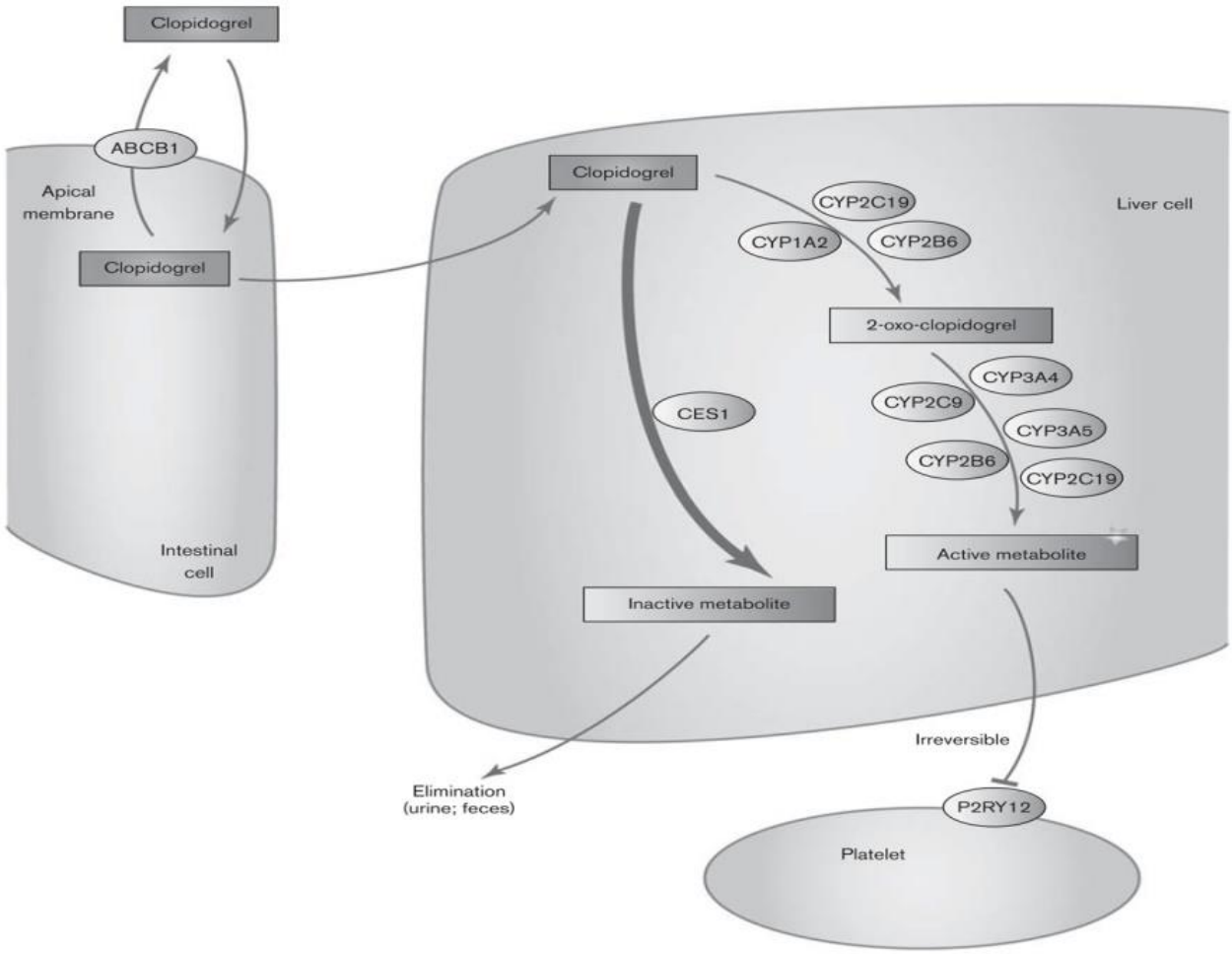


Figure 1. 2: Clopidogrel's activation from prodrug. The drug clopidogrel from the intestinal cell converts into 2-oxo-clopidogrel by cytochrome P450 monooxygenase-dependent metabolism CYP2C19 (in vitro) continuation to that, hydrolysis of 2-oxo-clopidogrel generates the active metabolite in the liver cell.

1.3 Metabolizing Gene, CYP2C19:

The major system for metabolizing lipids, hormones, toxins, and drugs are formed by the cytochrome P450 superfamily (CYP) which is a large and diverse group of enzymes. By various polymorphic feature of the CYP genes can result in reduced, absent, or increased drug metabolism.

A range of clinically important drugs, such as antidepressants, benzodiazepines, voriconazole (Moriyama et al., 2016) some proton pump inhibitors, and the antiplatelet agent, clopidogrel are metabolized by the contribution of the CYP2C19 enzyme.

The variability of clopidogrel metabolism and treatment outcomes between individuals is partly determined by variant alleles of the *CYP2C19* gene.

The CYP2C19 gene is highly polymorphic—35 variant star (*) alleles are catalogued at the Pharmacogenetic Variation (PharmVar) Consortium.

1.4 Mechanism of action:

Clopidogrel belongs to the second generation of thienopyridine antiplatelet agents. After the activation of clopidogrel from prodrug into an active metabolite in the liver it starts its mechanism of action.

After forming the active metabolite then it reaches the blood stream. In the blood stream it irreversibly binds to the P2Y purinoreceptor 12 on the platelet cell surface which is the clopidogrel's exhibition of its pharmacodynamic effect. It prevents ADP from binding to and activating it. P2Y12 receptor activation causes a complex cascade of intracellular events resulting in reduced protein kinase (PKA) phosphorylation of vasodilator-stimulated phosphoprotein (VASP) and subsequent glycoprotein (GP) IIb / IIIa receptor activation, granule release, platelet aggregation amplification, and platelet aggregate stabilization. At the same time, a significant portion of the ingested clopidogrel (at least 85 – 90 percent) undergoes first-pass metabolism in the liver where it is hydrolyzed to the inactive carboxylic acid metabolite SR26334 by carboxylesterase 1 (CES1). It prevents the activation of the P2Y12 receptor-associated Gi protein from inactivating adenylate cyclase in the platelet, leading to cAMP build-ups. As a result, only about 2 percent of the clopidogrel dose administered is converted to clop-AM and enters the systemic circulation with the advance to pharmacokinetic effect. At this point it should be remembered that CES1 also hydrolyzes the 2-oxo-clopidogrel and clop-AM. This formed up cAMP then stimulates the calcium efflux pumps and prevents calcium accumulation in the platelet which would cause activation and aggregation later on. Consequently, irreversible clop-AM binding to the P2Y12 receptor results in the inactivation of the GP IIb / IIIa receptor and in the destabilization of the thrombus for the platelet lifespan. (Collet et al., 2011).

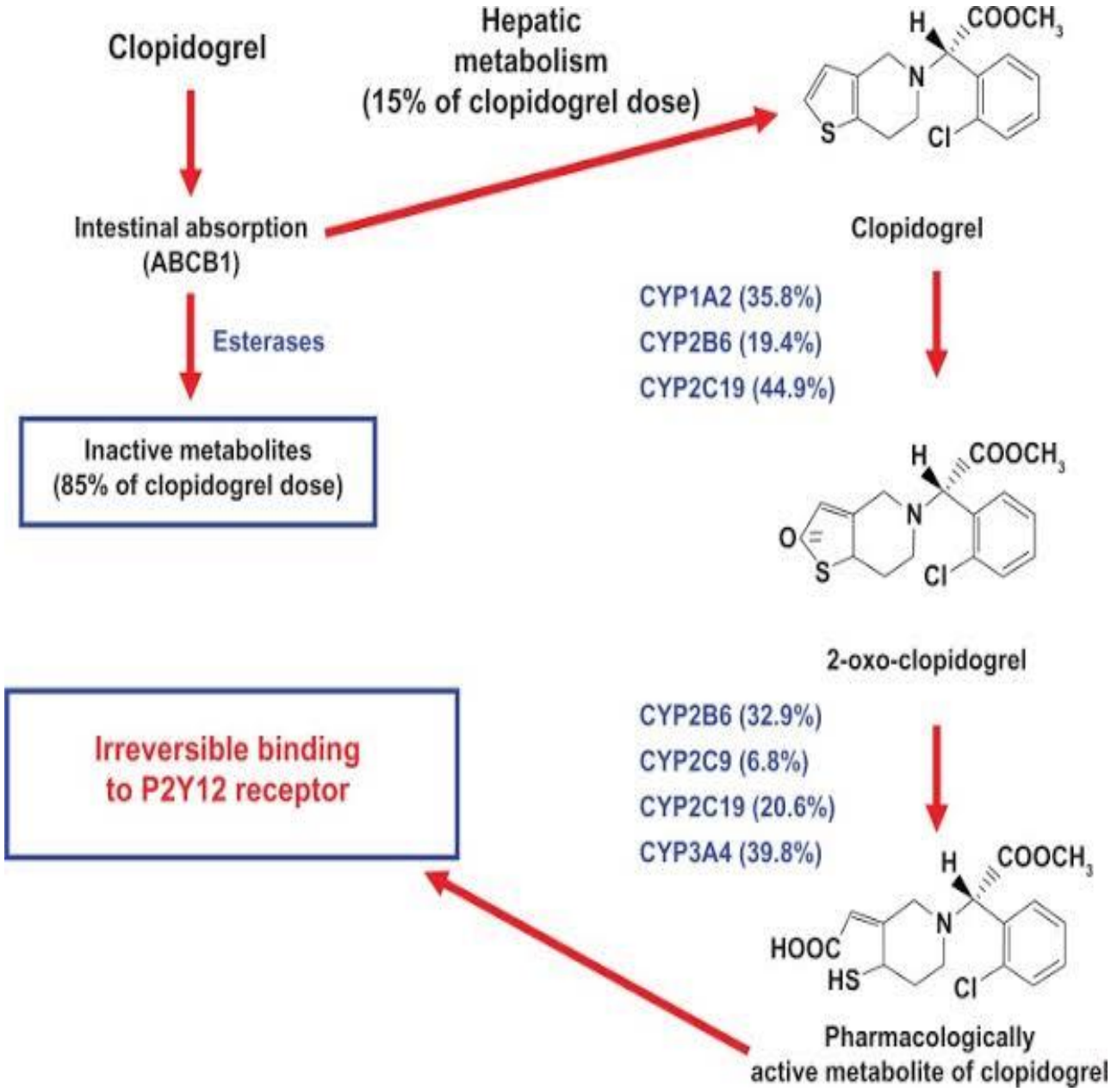


Figure 1.3: Pharmacokinetics of clopidogrel. At same time of platelet aggregation, a significant portion of the ingested clopidogrel (80-90) % undergoes first pass metabolism and prevents the activation of the P2Y12 receptor-associated Gi protein. As a result, only about 2 percent of the clopidogrel dose administered is converted to clop-AM and enters the systemic circulation with the advance to pharmacokinetic effect.

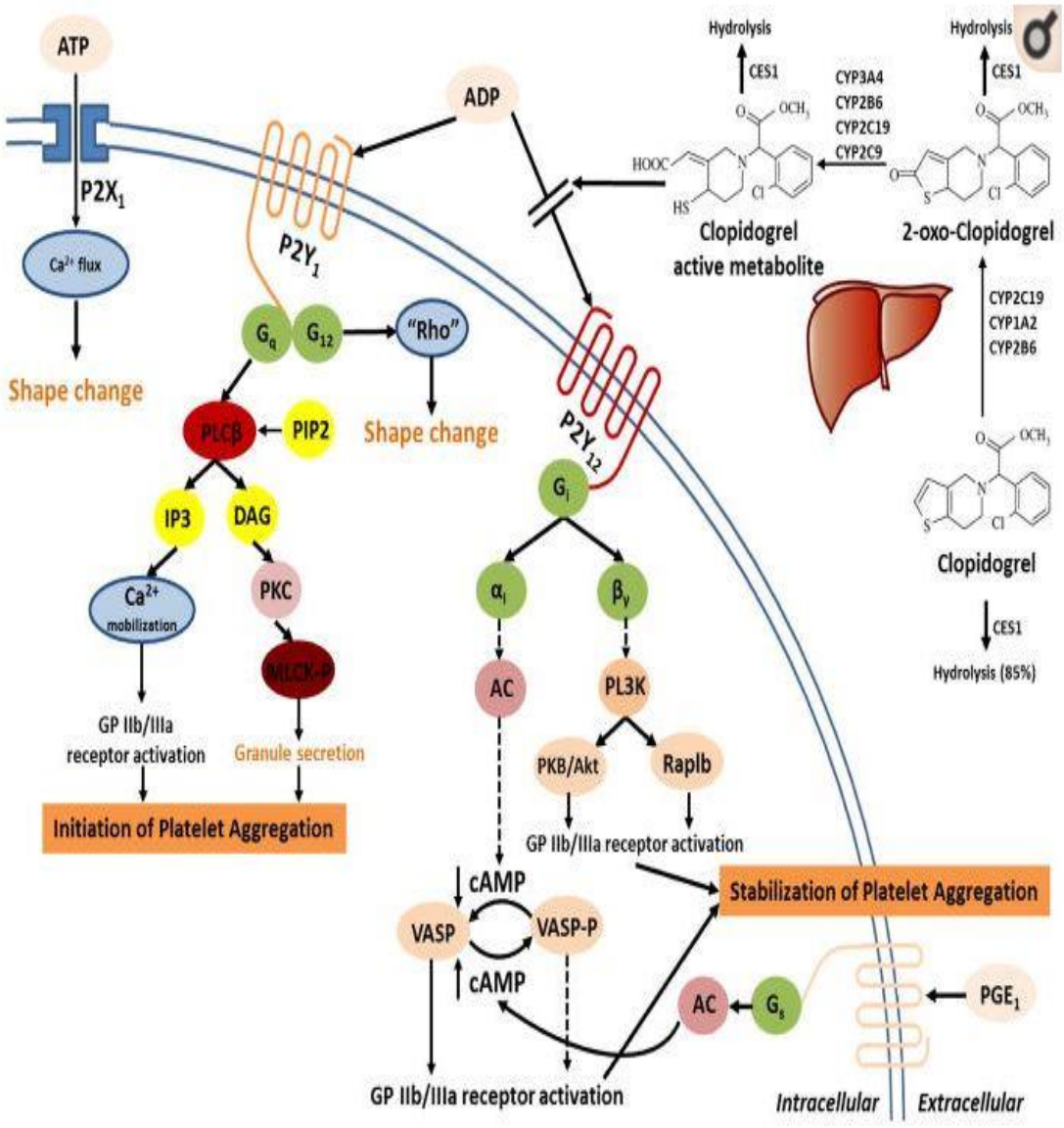


Figure 1. 4: Pharmacodynamics of clopidogrel. The active metabolite of the drug binding in the P2Y purinoreceptor 12 on the platelet cell surface exhibits its pharmacodynamic effect.

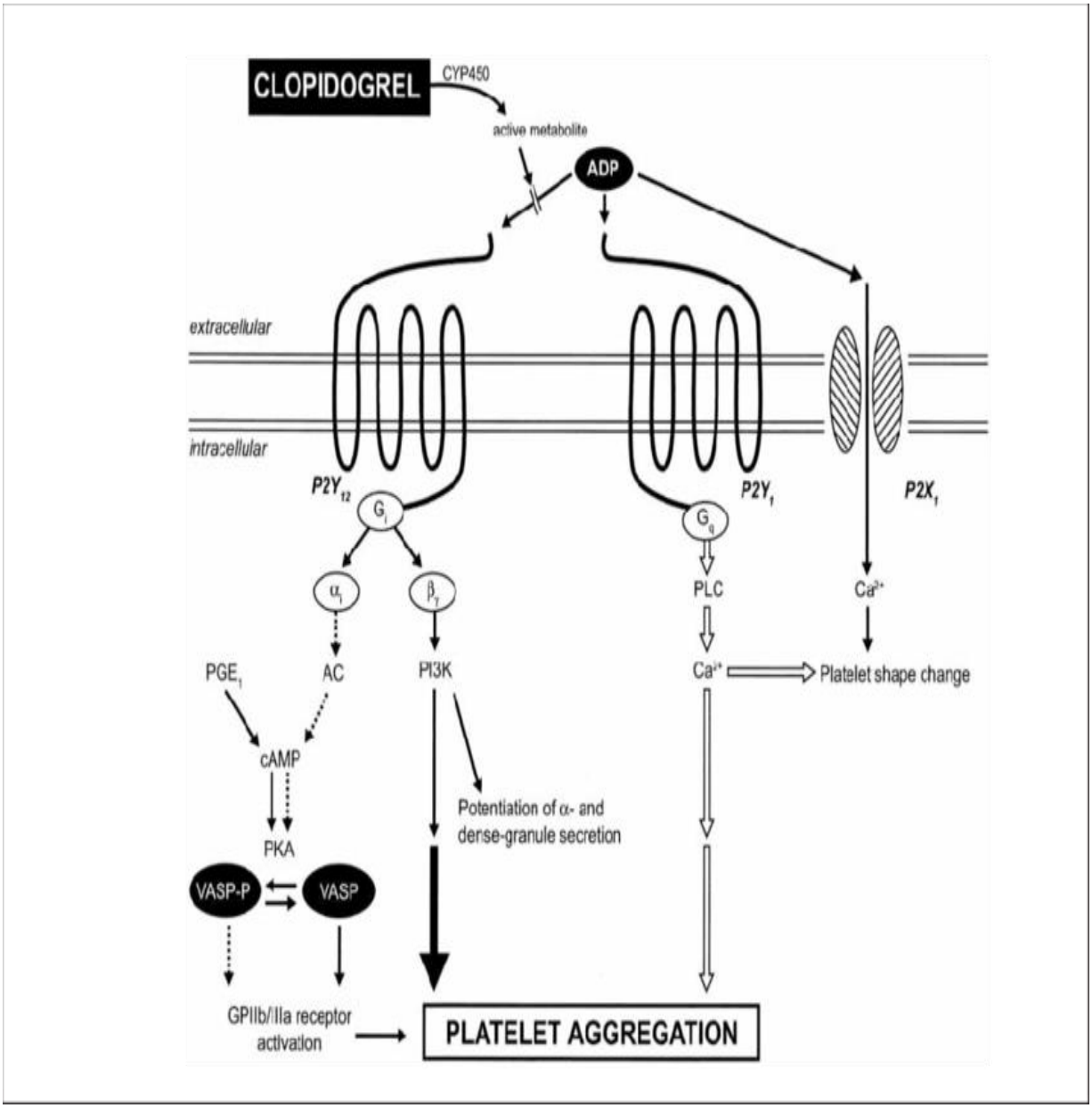


Figure 1. 5: Clopidogrel's mechanism of action. After formation of the active metabolite it reaches the blood steam and binds to the P2Y purinoreceptor 12 on the platelet cell surface. It prevents ADP from binding to and activating it. P2Y12 receptor activation causes a complex cascade of intracellular events resulting in reduced protein kinase (PKA) phosphorylation of vasodilator-stimulated phosphoprotein (VASP) and subsequent glycoprotein (GP) IIb / IIIa receptor activation, granule release, platelet aggregation amplification, and platelet aggregate stabilization.

1.5 Single Nucleotide Polymorphism (SNP):

Single nucleotide polymorphisms, frequently called SNPs (pronounced "snips"), are the most common type of genetic variation among people. A difference of a nucleotide means a difference in a single DNA building block represents each SNP. For example, in a certain stretch of DNA a SNP may replace the nucleotide cytosine (C) with the nucleotide thymine (T).

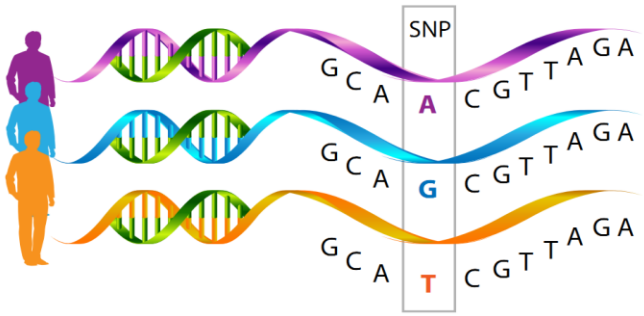


Figure 1. 6: A difference of single nucleotide from A>G in individual whether it should be A>T, represents a SNP.

1.5.1 Importance of SNP's:

In the study of human health some of the genetic differences have proven to be very important.

SNPs can-

- Help predict an individual's response to certain drugs,
- Identify susceptibility to environmental factors such as toxins
- Determine risk of developing particular diseases
- Used to track the inheritance of disease genes within families (What are single nucleotide polymorphisms? n.d.)

1.5.2 Allelic Variant CYP2C19*3: (SNP variant of CYP2C19 in the allele 3)

Clopidogrel gets effective by its conversion to an active metabolite by CY2C19. Individuals who carry 2 non-functional copies of the CYP2C19 gene are classified as CYP2C19 poor metabolizers.

Which contains a c.636G>A variant in exon 4 that causes a premature stop codon. The drug will not affect on the individuals because they have no enzyme activity and cannot activate clopidogrel via the CYP2C19 pathway.

Table 1. 1: CYP2C19 metabolizers.

Phenotype	Genotype	Examples
CYP2C19 poor metabolizer (A;A)	An individual carrying 2 non function alleles	*3/*3 (homozygous mutant)
CYP2C19 slow metabolizer (A;G)	Carrier of a CYP2C19*3 allele	*3 (heterozygous mutant)
CYP2C19 normal metabolizer (G;G)	Normal	

1.6 Genotyping:

Genotyping is used to classify variation in any biological species, from microorganisms to humans, at different locations in the DNA sequence. That variability, the genotype, exists naturally, and can be known as an individual's genetic fingerprint.

Compared with a reference sequence derived from the general population or a given subgroup, it is known as distinct, and can vary from the reference sequence in many ways.

Forms of mutation include single base variations (commonly referred to as single nucleotide variants and polymorphisms, or SNVs and SNPs), insertions and deletions (indels), and even gene copies (copy number variation or CNV). (Understanding Genotyping n.d.).

1.6.1 SNP Genotyping:

There are several ways to detect novel SNPs and known ones. (What is Genotyping? n.d.) These include:

- DNA sequencing
- mass spectrometry
- molecular beacons
- microarrays with SNP
- methods based on PCR

1.6.1.1 DNA Sequencing:

Here is a typical workflow for SNP detection and genotype calling from massively parallel sequencing data, starting from unmapped reads (in fasta format) (Li et al., 2013).

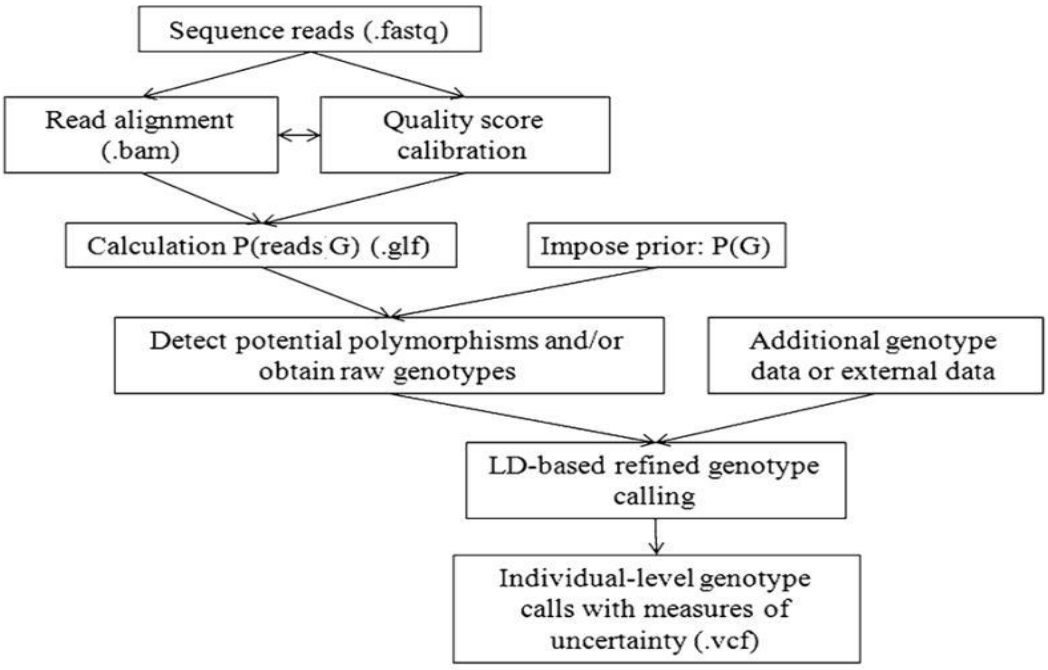


Figure 1. 7: A typical workflow for SNP detection and genotype

1.6.1.2 Targeted Next Generation Sequencing (NGS):

Targeted next-generation sequencing uses deep sequencing to identify known and novel variants within the interest field. It allows for speedier and more cost-effective sequencing of different

genome areas for in-depth research than whole genome sequencing (WGS). This approach needs less sample input and generates a smaller volume of data than WGS, making analysis more manageable. (Targeted Next Generation Sequencing n.d.) The most popular methods of NGS are-

Hybridization Capture: It is performed in a solution using biotinylated oligos (i.e., baits or probes) to capture complementary sequences from the sample library.

Amplicon Sequencing: It uses PCR primers to amplify the sequences of interest.

1.6.1.3 Mass Spectrometry:

Multiplexing, which allows simultaneous evaluation of multiple SNPs, is an effective, rapid and economical way to increase the medium-throughput genotyping performance and isformed using matrix-assisted time-of-flight mass spectrometry (MALDI-TOF MS) laser desorption ionization. Mass spectrometers allow standard microtitre plate formats, usually 384 wells, to be automatically detected and used. The MALDI spectra SNP analysis can be performed automatically through appropriate programs such as GenoTools, which enables high throughput calling without user intervention (Pusch et al., 2001).

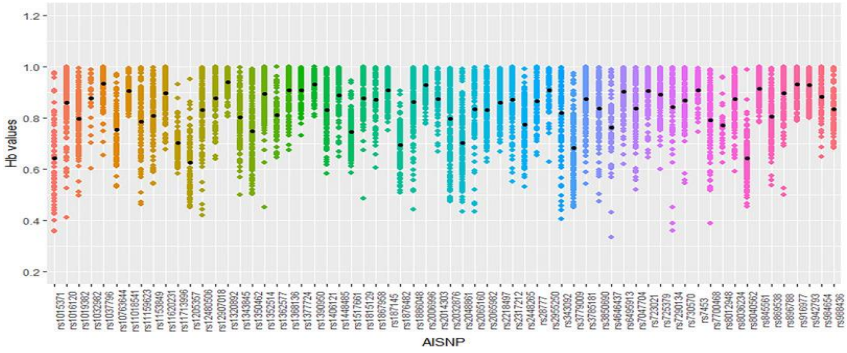


Figure 1.8: Mass Spectrometry

1.6.1.4 Molecular Beacon:

- SNP detection
- Allele discrimination
- Pathogen detection
- Multiplexing
- Viral load quantification

- Gene expression analysis
- Gene copy determination
- Endpoint genotyping
- in vitro quantification or detection

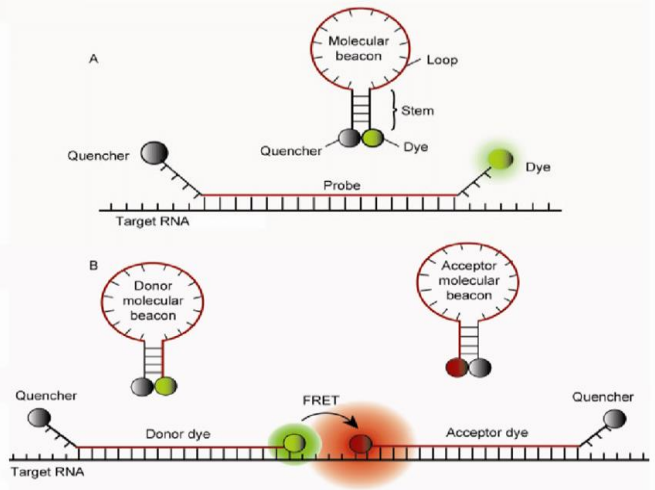


Figure 1. 9: Molecular Beacon

1.6.1.5 Microarrays with SNP:

SNP arrays, originally designed to genotype a patient's GWAS DNA and cosegregation studies to establish the link between a disease locus and a chromosome region, can also be used to detect CNVs. For example, deletions (or uniparental disomic) covering several of the neighboring SNPs included in the array may display loss of heterozygosity (LOH), while gains in copy numbers (e.g. duplication) may be identified by automated detection of an increased number of different genotypes. (Negrini 2016).

In single nucleotide mutation screening, these arrays took over from the frequently tedious method of using polymorphism (SSCP) single-strand conformation analysis, conformation sensitive gel electrophoresis (CSGE) or high-performance liquid chromatography (DHPLC) denaturing. SNP arrays are still the expense when compared to solutions focused on NGS. Generate a virtual karyotype to determine each SNP's copy number on the array and then align the SNPs in chromosomal order with advanced software (Wan Lam et al., 2010).

1.6.1.6 Methods Based on PCR:

Highly precise and sensitive methods should be used to identify / detect SNPs or mutations. Polymerase chain reaction (PCR) has provided the analytical output needed for many molecular analyzes in this regard.

PCR-based SNP / mutation detection methods are generally categorized into two specific types-

1. Polymorphic or mutant allele-directed analyzes using primers matched to replaced nucleotides or using oligonucleotides to block or clamp the non-directed template
2. Melting curve analysis combined with real-time PCR techniques using hydrolysis samples, hybridization samples (Matsuda 2017).

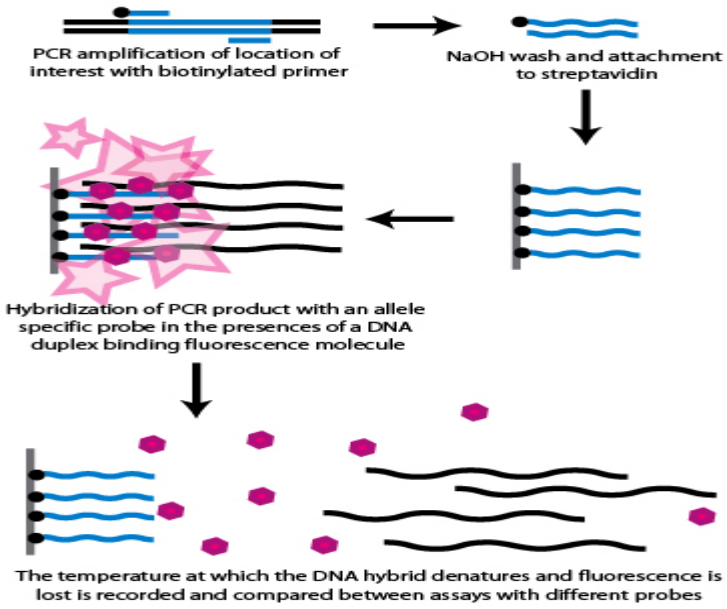


Figure 1.10: PCR method in SNP detection

1.6.1.7 Tetra Primer ARMS PCR:

The reaction of the tetra-primer amplification refractory mutation system – polymerase chain (ARMS – PCR) is a simple and economical process for genotype polymorphisms (SNPs) with single nucleotides. In a single PCR it uses four primers and is accompanied only by gel electrophoresis.

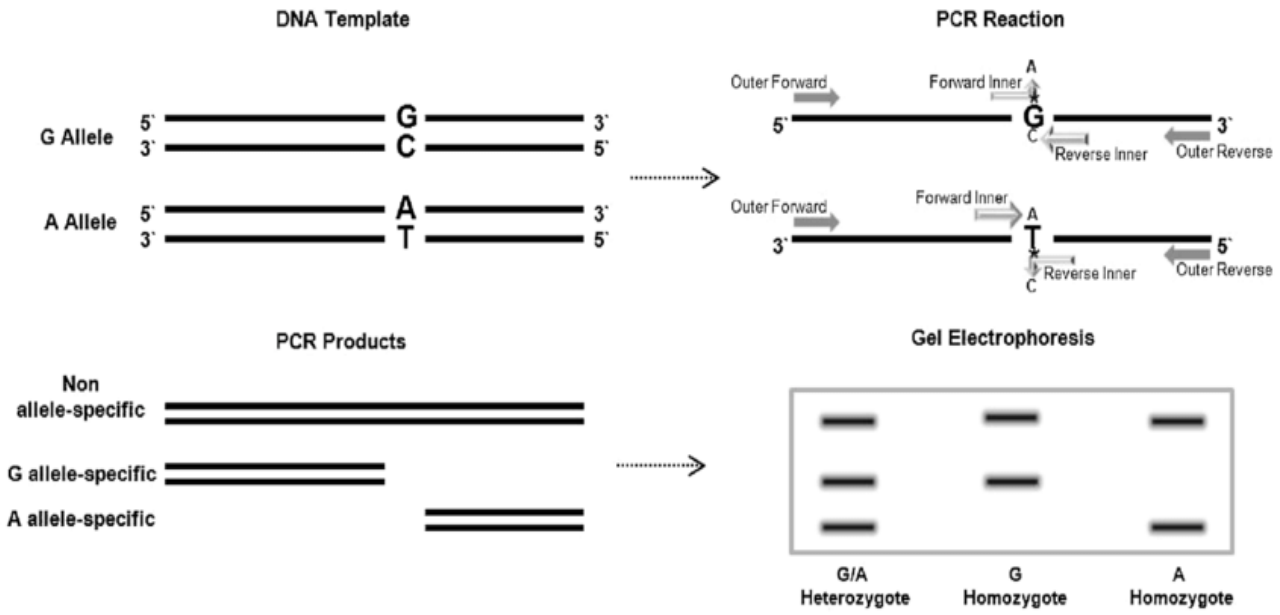


Figure 1. 11: Method of tetra primer ARMS PCR. Here, in this method G and A allele specific PCR products produces homozygote band and because of the mismatched inner primer at the SNP locus it produces heterozygote band.

Two SNPs is generally chosen which supported different amplification conditions. Methods of extraction of DNA, temperature annealing, protocols of the PCR cycles, reagents, and concentration of primers are also studied. The use of tetra-primer ARMS – PCR may be hampered for SNPs in cytosine and guanine-rich DNA areas, and unpurified DNA samples.

Using tetra-primer ARMS – PCR meets the demands of modern genomics research and enables SNPs to be studied in a fast, reliable, and low-cost way (Medrano et al., 2014).

1.6.1.8 Agarose Gel Electrophoresis:

The beauty of ARMS PCR is that no hybridization steps are needed. This will make this technique more suitable and more diagnostic-reliable.

To get the results, the amplified fragments are placed directly on the 2 percent agarose gel. The samples are loaded into the agarose gel sequentially, and run for 60 minutes.

1.7 Therapeutic Implications:

The U.S. In March 2010, A new label for clopidogrel was approved by the Food and Drug Administration (FDA) with the addition of a boxed warning about pharmacogenetics, noting decreased therapy efficacy in poor metabolizers (defined as having two CYP2C19 alleles loss-of-function). The boxed warning further notes that "tests are available to determine the CYP2C19 genotype of a patient and can be used as an aid in evaluating clinical protection and suggest alternative diagnosis or treatment approaches in patients classified as CYP2C19 poor metabolizers." (Janice et al., 2013).

1.7.1 Percutaneous Coronary Intervention:

The genetic variability decreases clopidogrel activation. This raises the risk of severe cardiovascular problems (percutaneous coronary intervention) in patients undergoing balloon angioplasty or stent placement. There have been no reports of any adverse health effects in other patients.

Table 1.2: DPWG (2017) Recommendations for Clopidogrel and CYP2C19 Phenotype:

Phenotype	**Recommendation**
Poor Metabolizer	Coronary percutaneous intervention: ▪ Prasugrel and ticagrelor alternatives are not metabolized with CYP2C19 (or to a lesser extent) Other indications: ▪ Determine the extent of platelet aggregation inhibition using clopidogrel ▪ Consider an alternative in poor responders CYP2C19 (or to a lesser extent) is not metabolized by Prasugrel and ticagrelor

1.8 Importance of the Study:

Medications are necessary to treat any health issues. But it would not be effective if the given medications cannot function adequately. Some medicated drugs can over function or less function in an individual's body. This difference holds the genotype to phenotype changes which leads to the importance of Pharmacogenetic analysis into individuals. This pharmacogenetic approach signifies to particular SNP in populations which causes variation of drug dose.

CYP2C19*3 is such kind of an allelic variant which functions as a poor metabolizer in clopidogrel therapy causing loss of function of drug. For this reason, the pharmacogenetic study of this particular SNP reveals of it reduce function in population given Clopidogrel therapy.

To solve this issue, the genotyping method can easily identify the responsible SNP in individual and can treat them with the right dose of medication.

1.9 Objectives of the Study:

The objectives of the study were as following:

1. To develop a simple, inexpensive and easy-to-perform genotyping method for the study SNP.

2. To determine the allele frequency of the study SNP in Bangladeshi population.

So, the study possesses the idea of further Pharmacogenetic approaches in this field of genotyping particular SNP which leads to the valuable research for personalized drug to have the proper medicated result in population.

CHAPTER 2

MATERIALS AND METHOD

2.1 Study Design:

The study is designed into three major categories of *In silico* experiment, In vitro experiment and data analysis. The primers were designed for the SNP CYP2C19*3 of CYP2C19 *Homo sapiens* gene. There are four sets of primers. Here, the 3' end of the primers is changed such that the regular allele can be amplified by one set of the primers and the mutant allele can be amplified by others. The single base mismatch is added at the 3' end of the primordial. Such mutation causes the primordial to amplify a single allele. The primers were designed according to their best parameters to function in correct compatibility.

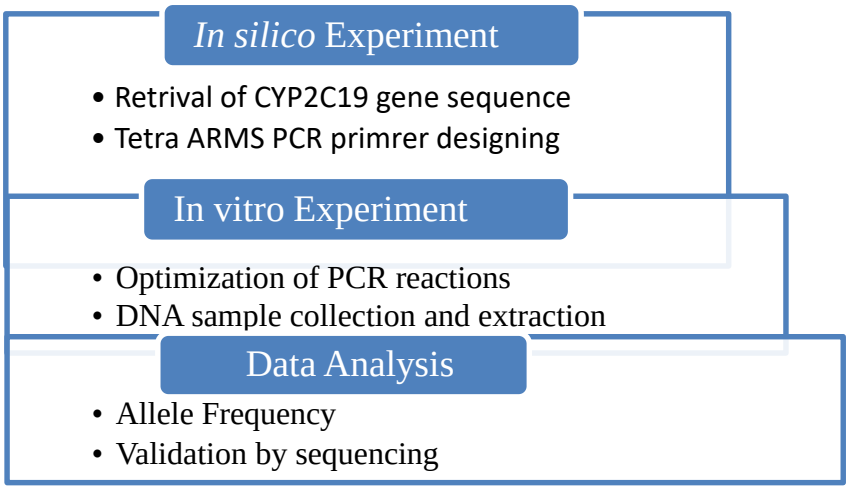


Figure 2.1: A workflow involved in this study design. There are 3 major parts of 1) In silico experiment, 2) In vitro experiment, 3) Data analysis. Each of the parts series of steps mentioned in the above figure.

2.1.1 *In silico* Experiment:

This section of the experiment comprises *In silico* approaches to retrieve CYP2C19 gene sequence and from the obtained gene sequence the tetra ARMS allele specific primers were designed to further optimize for the PCR reactions. Finally, the sequences were synthesized in Batch Primer 3.

2.1.1.1 Retrieval of the *Homo sapiens* CYP2C19 gene:

To retrieve the Homo sapiens CYP2C19 gene the steps were followed are given below-

Firstly, the NCBI gateway web page was browsed by using the URL: https://www.ncbi.nlm.nih.gov/Then in the search box "Homo sapiens CYP2C19" was typed and in the drop-down menu database "Gene" was selected. To start the query search button was clicked. After that, various results were shown as per the entry and the first result with the ID: 1557 CYP2C19 was selected. That particular entry revealed information regarding the gene. From there, in the section of "Genomic regions, transcripts, and products" to the "Go to nucleotide" "Genbank" option was selected to visit the GenBank entry. Then in the GenBank entry page to the right corner "Change region shown" menu was selected in the region from 94762681 to 94855547 to get the complete sequence of the gene. Beside that menu "send to" option was selected into "Complete Record" chosen the destination "File" in the format of "FASTA" and finally the "Create file" button was clicked to download the complete gene sequence of CYP2C19. From this FASTA output result primers were designed in the further progression of this experiment.

2.1.1.2 Designing of the Tetra-primer ARMS-PCR primers:

Designing of the tetra ARMS PCR primers were done in Batch Primer3 v1.0 which is an automated server and the gateway website is https://wheat.pw.usda.gov/demos/BatchPrimer3/. To develop different types of primers in a high-throughput manner BatchPrimer3 is a comprehensive web primer design program.

BatchPrimer3 is developed based on Primer3. It functions by adopting the Primer3 core program as a major primer design engine to choose the best primer pairs. To pick position-restricted primers a new score-based primer picking module is incorporated into BatchPrimer3 v1.0 which implements several types of primer designs including generic primers, SSR primers together with SSR detection, and SNP genotyping primers (including single-base extension primers, allele-specific primers, and tetra-primers for tetra-primer ARMS PCR), as well as DNA sequencing primers (You et al., 2008).

Table 2.1: The parameters finalized for the ARMS-Primer sets from this website are given below:

Parameters	min	opt	max
Primer size (bp)	20	28	30
Primer Tm (℃)	50	65	80
Max Tm difference (℃)	5		
Primer GC%	20	80	
Inner product size (bp)	100	250	400
Relative number between inner product sizes	0.5	1.5	
Max #N's	0		
Salt concentration (µg)	50		
DNA concentration (µg)	50		
Max self-complementarity	8		
Max 3' self-complementarity	3		

2.1.2 In Vitro Experiments:

This section focuses on the affectivity of the designed primers in PCR reaction. By the optimizing procedure of primers in PCR reaction and later individual DNA samples genotyping by their allele specific primers are the overall phases of in vitro experiments.

2.1.2.1 In vitro optimization of PCR conditions:

The first phase of in vitro experiment for optimizing PCR is to detect whether our designed primers work for the genotyping of our extracted DNA samples. The DNA extraction was done by collecting samples from 50 various aged individuals then their extracted DNA samples were amplified in PCR. The individual's extracted DNA samples were used as DNA templates for each PCR reaction.

2.1.2.2 First attempt of Polymerase Chain Reaction:

The steps of discovering whether our designed primers work or not are given below,

4 nuclease free PCR tubes were labeled. There were 2 pairs of allele specific primers. So, 2 pairs of primer × 2 reactions each = 4 tubes. The other tube was for non-template negative control. In the first few attempts each of them ran at 60°C, 59°C and 61°C. These temperatures were selected according to the optimal temperature of the designed primers. Annealing temperature should be 5°C below the Tm of our designed primers. The PCR master mix was prepared according to the table. All the components were added sequentially in the same order stated below in the table. The PCR thermal cycler instrumentation was configured with the cycling condition shown in the figure.

Table 2.2: Recipe for the first attempt of PCR reactions.

Component	Volume (µL)
10x Reaction Buffer	2.5
dNTP (10 mM)	0.5
Outer Forward Primer	0.5
Outer Reverse Primer	0.5
Inner Forward Primer	0.5
Inner Reverse Primer	0.5
Taq DNA Pol	0.2
PCR Water	16.8
Template DNA	3.0
	Total= 25 µL

After the preparation of the master mix the 4 nuclease free tubes were run for polymerase chain reaction each tube containing individual template DNA samples and 2 pairs of allele specific primer. The other tube was for non-template negative control to get ensured of the primer functioning. The PCR vials were placed in the region of the instrument where it supposed to produce 60°C, 59°C and 61°C. The PCR reaction was started. After the PCR reaction was completed the PCR products were resolved in a 1.5% agarose gel for the PCR reaction which ran at 60°C annealing temperature and for other two annealing temperatures 59°C and 61°C, the PCR products were resolved in a 2.0% agarose gel. Each of the PCR products resolving in the gel was run at 80 voltage current for 60 minutes. The DNA bands in the agarose gel were visualized under UV rays and captured the resulted bands.

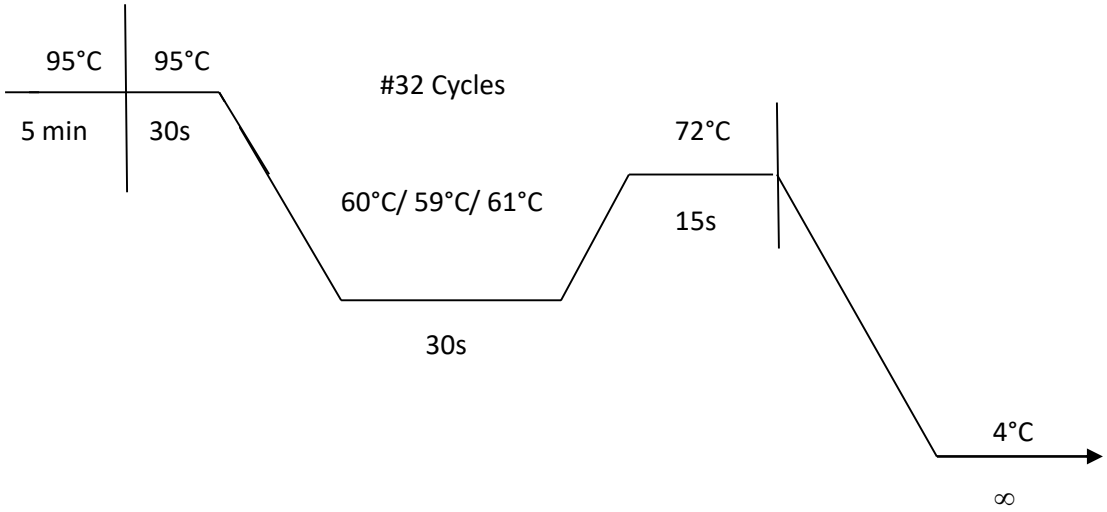


Figure 2.2: Thermal Cycling condition for the first attempt of a polymerase chain reaction: the final elongation was for 15 seconds which was later increased in the fixation step

2.1.2.3 Fixing the PCR Conditions:

The following steps were followed to yield good PCR Product.

4 nuclease free PCR tubes were labeled. There were 2 pairs of allele specific primers. So, 2 pairs of primer × 2 reactions each = 4 tubes. The other tube was for non-template negative control. In the fixing of PCR conditions each of them ran at 58°C. These temperatures were selected according to the optimal temperature of the designed primers. Annealing temperature should be 5°C below the Tm of our designed primers. The PCR master mix was prepared according to the table. All the components were added sequentially in the same order stated below in the table. The PCR thermal cycler instrumentation was configured with the cycling condition shown in the figure.

Table 2. 3: Recipe of the PCR reaction while fixing the conditions for primer pairs:

Component	Volume (µL)
10x Reaction Buffer	2.5
dNTP (10mM)	0.5
Outer Forward Primer	0.5
Outer Reverse Primer	0.5
Inner Forward Primer	0.5
Inner Reverse Primer	0.5
Taq DNA Pol	0.2
PCR Water	16.8
Template DNA	3.0
	Total= 25 µL

After the preparation of the master mix the 4 nuclease free tubes were run for polymerase chain reaction each tube containing individual template DNA samples and 2 pairs of allele specific primer. The other tube was for non-template control to get ensured of the primer functioning. The PCR vials were placed in the region of the instrument where it supposed to produce 58°C. The PCR reaction was started. After the PCR reaction PCR product was in the gel at 80 voltage current for 60 minutes. The DNA bands in the agarose gel were visualized under UV rays and captured the resulted bands.

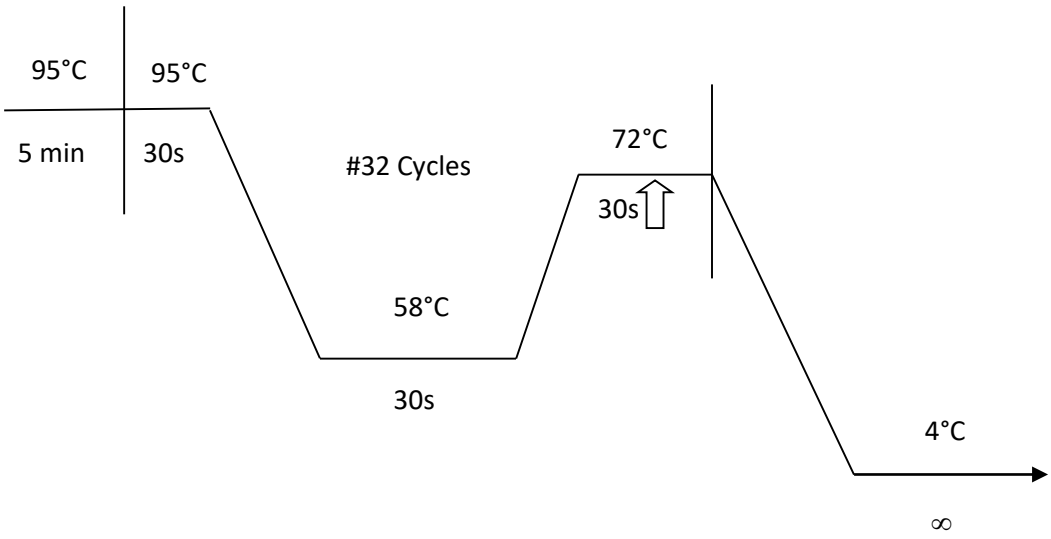


Figure 2. 3: Thermal cycling condition for fixing the PCR condition: The elongation time was increased in this fixing step. It increased to 30 seconds instead of 15 seconds.

2.1.2.4 Confirmation of the PCR reaction:

The thermal cycling condition optimized in the previous experiments needed a final confirmation. To get the desired sharp band of the PCR products several more optimization steps were needed to confirm the thermal cycle.

Due to the unavoidable reasons of COVID-19 and the severity of the pandemic, our university was shut down along with its all physical presence activities. The wet lab experiment had to leave in the mid-way with no further continuation to get the proper and final result of this project.

2.1.2.5 Performing PCR reaction on individual DNA samples:

Every PCR reaction was performed with individual DNA samples to confer their genotyping results. To do that individual samples were collected and their DNA was extracted. So far 50 samples were collected and got good extracted DNA from 35 samples. The confirmation of extracted DNA was done by gel electrophoresis in 0.8% agarose gel for 50 minutes.

2.1.2.6 Sample Collection:

The buccal samples were collected in the gargling method from different age of healthy individuals. The participants were selected without any bias of gender, religion, disease condition etc. The gargling samples were collected in 0.9% NaCl in dH2O (distilled water).

Healthy participants were asked for their consent with the required information for the paperwork. Participants had saline and gargled with it for about 1 minute by motion with the liquid and moving it to the both cheeks. In this sample collecting section the participants were strictly monitored with the time of their gargling for at least of 1 minute and doing the movement to the cheeks properly. After that, to collect the sample they were provided with an autoclaved sterilized falcon tube to pour their gargle fluid into it.

The collected falcon tubes with samples in it were labeled with the same identification number which was used in their consent form for further work.

2.1.2.7 DNA Extraction from the Collected Samples:

The samples were transferred to eppendorf tube of 1500 µl sample from the falcon tube. The transferred samples were centrifuged at 4500 rpm for 5 minutes. This step was repeated 3 /4 times until getting a good. Then the pellet was suspended in 500 µl lysis buffer [10 mM Tris (pH 8.0), 10 mM EDTA, and 2.0% SDS] and 50 µl 10% SDS, followed by 5–10 µl 20 mg/ml proteinase K was added. The samples were incubated 2h at 65°C. The DNA was then extracted from each sample with an equal volume of phenol:chloroform:isoamylalcohol solution (25:24:1) and mixed gently by inverting the tubes for 3 min. The samples were then centrifuged for10 min with 10,000 g (4°C), and the upper aqueous layer was transferred to a fresh, sterilized microcentrifuge tube. RNase A (10 µl of 10 mg/ml) was added and the solution was incubated at 37°C for 30 min. Equal volumes of chloroform:isoamylalcohol solution were added and centrifuged again with 10,000 g (4°C) for 10 min. The upper aqueous layer was transferred to a sterilized microcentrifuge tube, and double the volume of chilled isopropanol was added, along with one-tenth volume of 3 M sodium acetate, and chilled at -20°C for 1 h for precipitation. After 1 h, the samples were centrifuged at 10,000 g (4°C) for 10 min. After decanting the supernatant, 250µl 70% ethanol was added and the pellet was dissolved; the mixture was centrifuged at 10,000g for 10 min and the supernatant was decanted gently. The pellet was air dried under laminar air flow and the dried pellet was resuspended in 40 µl nuclease-free water and frozen at -20°C for storage.

2.1.3 Data Analysis

2.1.3.1 Allele Frequency and Validation by Sequencing:

To get the overall allele frequency among the target population SNP genotyping is very essential to perform. To do this, enough good PCR results were needed from an individual's DNA sample to calculate the allele frequency. It would help to retrieve the idea of percentage population carrying the SNP variant. Along with it, to validate the SNP genotyping sequencing is much needed. Sequencing can easily bring out the genetic differences and accurate mapping of the gene in a cost-effective way.

This step was much needed to meet the objective of the study but because of the pandemic situation mentioned before it was not possible to continue to complete the last steps of this study (Bhatia et al., 2013).

CHAPTER 3

RESULTS

3.1 DNA Sequence of *Homo sapiens* CYP2C19 Gene:

The NCBI gateway website leads to the "GenBank" of *Homo sapiens* CYP2C19 Gene which featured some important characteristics of the gene. It is quiet of a large DNA sequence of 94,762,681 to 94,855,547 base pairs. In the molecular location it is present on the chromosome 10 and considering the cytogenetic location is found to be on the long arm of chromosome 10 at position 23.33.

Table 3. 1: The different sequence attributes of the *Homo sapiens* CYP2C19 gene

Homo sapiens CYP2C19	
Official Symbol	CYP2C19 (provided by HGNC)
Official Full Name	Cytochrome p450 family 2 subfamily C
Accession code	NC_000010
Gene ID	1557
Gene Type	Protein Coding
EC Number	1.14.14.51
Refseq Status	Reviewed
Organism	*Homo sapiens*
Location	10q23.33
Length of the gene sequence	92867 bp
Region	Complement (746..11948)
Total Exon Count	9
Exon Position	(1….193,12378..12540,12710..12859,17819..17979,19141..19317,57816..57957,80157..80344,87237..87378,90053..92867)
Total transcript variant	3
Transcript ID	NM_000769.4
RefSeq Protein ID	NO_000760.1

3.2 Designing the Allele Specific Primers for Tetra-Primer ARMS PCR:

The search result in dbSNP from the NCBI gateway shows the 5' flanking position of the CYP2C19*3 SNP 5'-AGCACCCCCTG-3'.

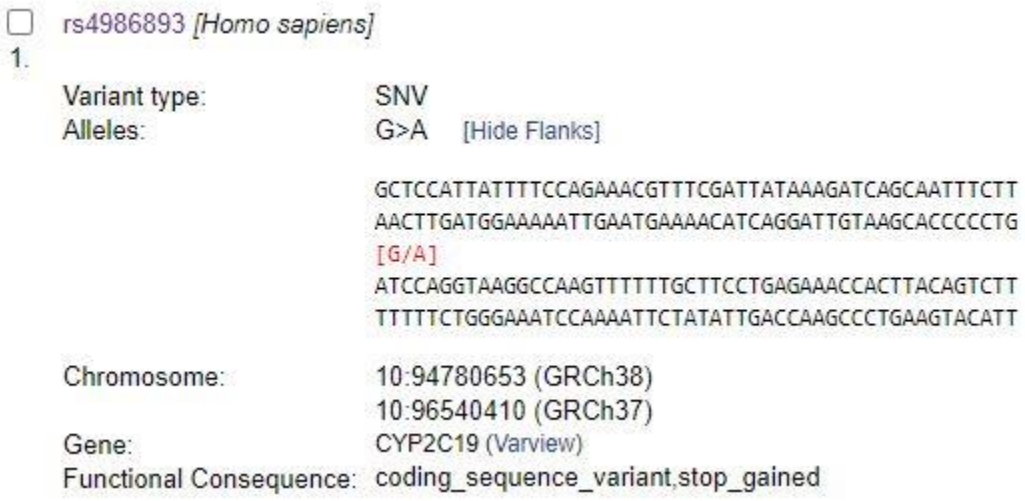


Figure 3. 1: search result in dbSNP from NCBI gateway indicating the flanking sequence of CYP2C19*3 SNP.

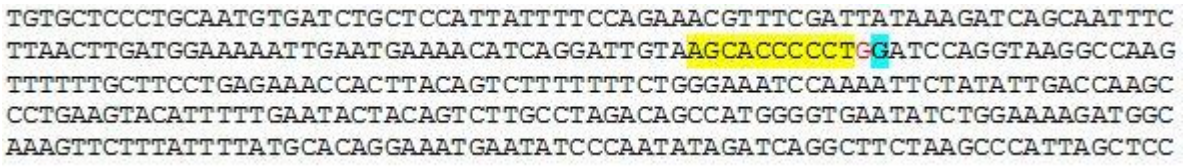


Figure 3.2: From the above figure it shows the word document of CYP2C19 gene sequence where the flanking sequence of the SNP CYP2C19*3 was searched and the important portion, positions were highlighted as 5' flanking sequence in yellow, penultimate position in red and the SNP locus in blue.

In the table (3.2) given below it display's the designed primers using the parameters described previously from the gateway website of BatchPrimer3 (v1.0). The website produced the best melting temperature, GC content and other features to function the allele specific primer sets in efficient way. The ARMS PCR process is based on the adjustment of the primers for different alleles. Here, the 3' end of the primers is changed such that the regular allele can be amplified by one set of the primers and the mutant allele can be amplified by others. The single base mismatch is added at the 3' end of the primordial. Such a mutation causes the primordial to amplify a single allele. The outer forward primer, outer reverse primer and the inner reverse primer are 28bp in length and the inner forward primer is 24bp in length.

The designed primers have melting temperature of between 70°C where the lowest melting temperature (69.97) counts for the outer reverse primer and the highest temperature (70.36) counts for the inner forward primer. It has GC content of 42%. The inner reverse primer (mutant) has the lowest GC content (42.43) and the inner forward primer has the highest GC content (54.17).

Table 3.2: Display's the designed primer from the website with all its important parameters like Tm, GC content.

Primer Type	Primer Sequence	Product Size	Tm	GC content
Outer Primer Forward Wildtype	TCCCTGCAATGTGATCTGCTCCATTATT	278	69.99	42.86
Outer Primer Reverse	TTTGCCATCTTTTCCAGATATTCACCCC		69.97	42.86
Inner Primer Forward	ATCAGGATTGTAAGCACCCCCGGA	184	70.36	54.17
Inner Primer Reverse Mutant	AAGCAAAAAACTTGGCCTTACCTGGCTC	145	70.25	46.13

Graphical Representation helps to provide a visual clue to the pattern of distribution of bases along DNA sequences. It helps into see the base distribution in DNA or RNA sequence following the variations in family of gene. From the designed primers showing the graphical representation of CYP2C19*3 variant with designed mutant primer in the figure below:

Figure 3.3: The figure displaying the CYP2C19*3 SNP locus in blue at forward inner primer, the penultimate position in red and the mutated inner reverse primer.

3.3 PCR Optimization Result:

In the first few attempts of PCR optimization 60°C annealing temperature gave wild type homozygous band (faint) for 2 samples in the first two wells. Other two samples did not give any bands in the next two wells.

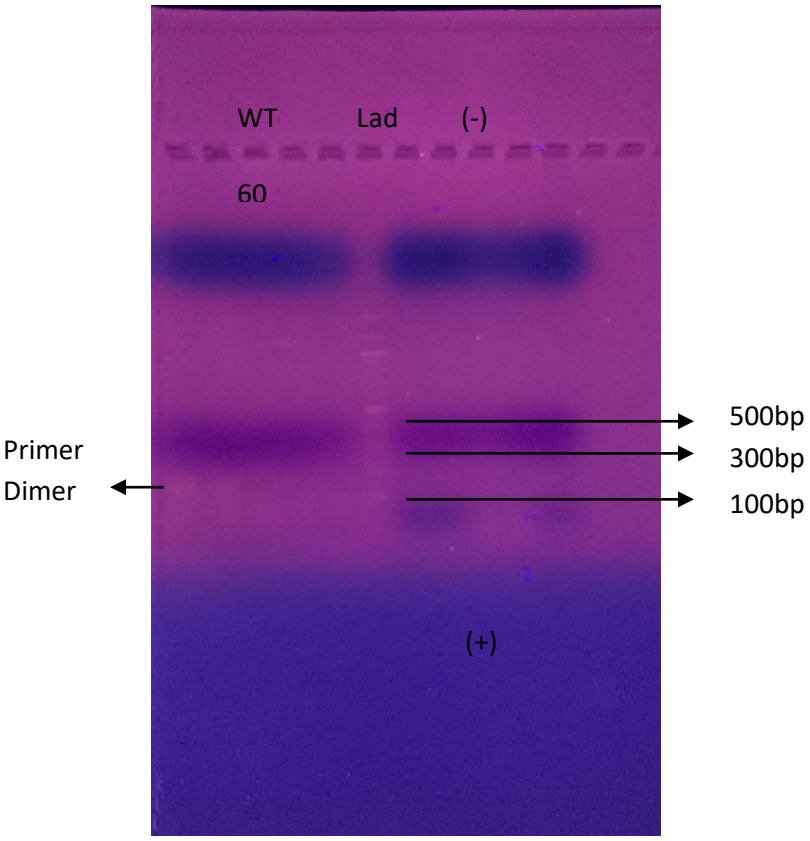


Figure 3.4: In this attempt of PCR optimization the PCR products were resolved in a 1.5% agarose gel in 1x TAE buffer ran for 60 minutes at 80 voltages.

There were total of 4 samples ran together in the gel electrophoresis but it gave result for only 2 bands in the first two wells other 2 samples did not give any band.

The bands resulted in the gel electrophoresis were faint bands and they were wild type homozygous bands.

100 bp ladder was used. The bands were not quiet in their actual product sizes. Therefore, the products were nonspecific or formed primer dimer.

In the next attempt, the PCR products were run at 59°C annealing temperature. It gave result for 3 samples out of 4 samples. Two of them gave wild type homozygous (nonspecific bands) and another product gave smear result in the gel electrophoresis.

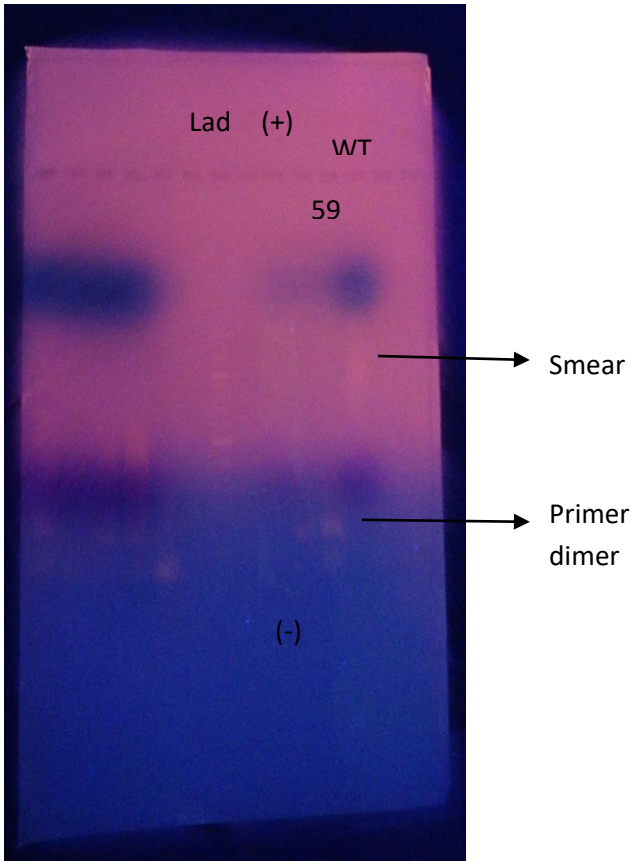


Figure 3.5: In this attempt of PCR optimization the PCR products were resolved in a 2.0% agarose gel ran in 1x TAE buffer for 60 minutes at 80 voltages.

There were total of 4 samples ran together in the gel electrophoresis but it gave result for only 3 samples. Two of them gave visible wild type homozygous band another product gave smear result.

100 bp ladder was used. The bands could not be accounted with the ladder for its product size. Therefore, the products were nonspecific or formed primer dimer.

Another attempt for PCR optimization was held at 61°C annealing temperature. Total of 4 samples ran together but it did not produce any sort of band in the gel electrophoresis.

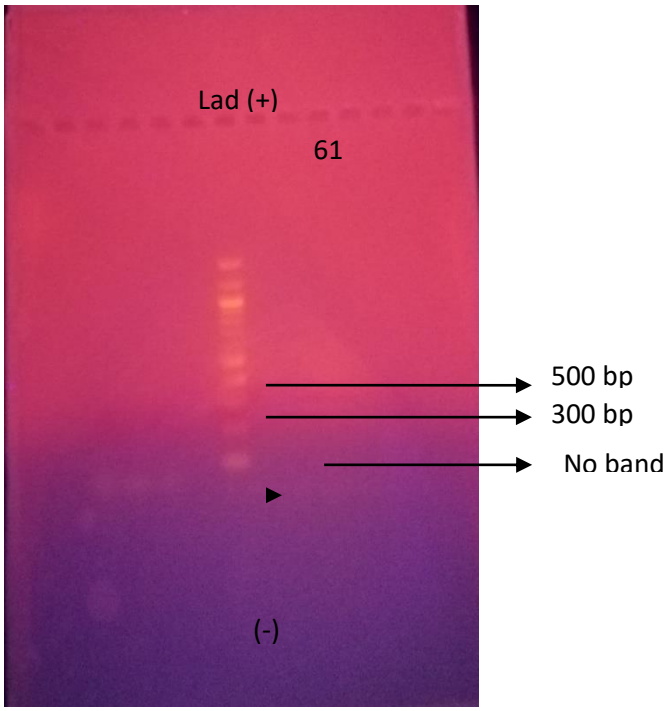


Figure 3.6: In this attempt of PCR optimization the PCR products were resolved in a 2.0% agarose gel ran in 1x TAE buffer for 60 minutes at 80 voltages.

There were total of 4 samples ran together but did not show any band in the gel electrophoresis.

In the fixing step of PCR optimization occurred at 58°C annealing temperature by increasing the elongation to 30 seconds instead of previous 15 seconds. Total of 4 samples ran together and got wild type homozygous (faint) band for only two samples in the first two wells.

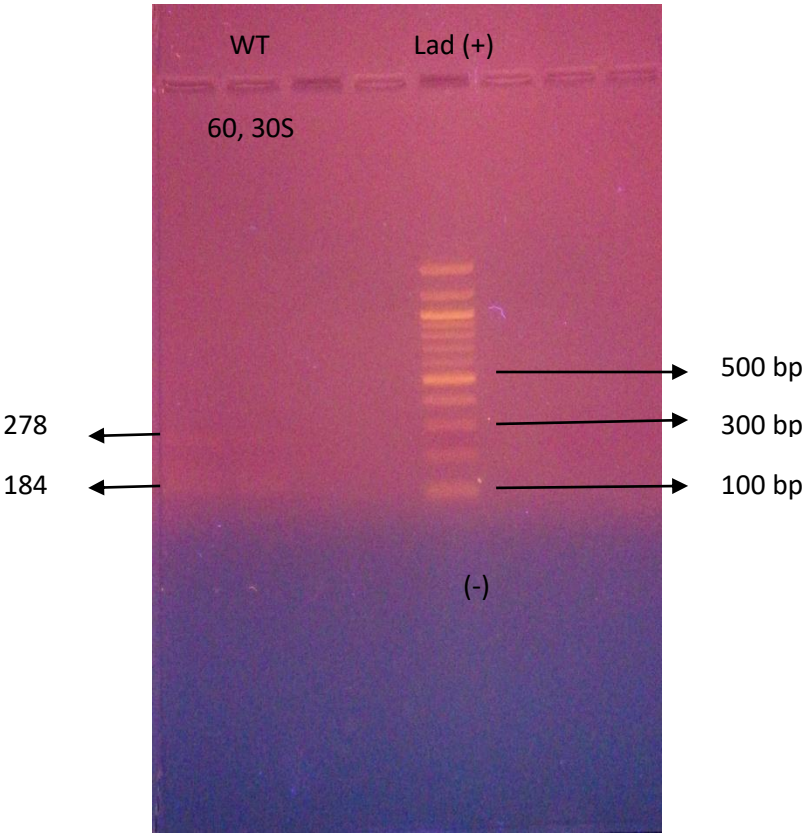


Figure 3.7: In this attempt of PCR optimization the PCR products were resolved in a 2.0% agarose gel ran in 1x TAE buffer for 60 minutes at 80 voltages.

There were total of 4 samples ran together in the gel electrophoresis but it gave result for only 2 bands in the first two wells other 2 samples did not give any band.

The bands resulted in the gel electrophoresis were faint bands and they were wild type homozygous bands.

100 bp ladder was used. The bands comparing to the ladder were quite specific in their product sizes according to the designed primers.

CHAPTER 4

DISCUSSION

4.1 General Discussion:

Medications are necessary for treating various health issues in individuals. Thus, medications can be over functioned or less functioned in different individuals relating to their genetic makeup. Carriage of cytochrome P450 (CYP) 2C19 loss-of-function (LoF) polymorphism associated with high platelet reactivity (HPR) on treatment and an increased risk of ischemic events in patients during clopidogrel therapy.

SNP is the usual type of genetic variation among individuals. The rate of occurring SNPs normally throughout a person's DNA is almost once in every 1,000 nucleotides on average, which means in a person's genome there are roughly 4 to 5 million SNPs. SNPs are usually are classified as SNVs that occur in the population at > 1 per cent. Depending on the number of SNPs cataloged in Build 149 of the SNP database, dbSNP maintained by the National Center for Biotechnology Information (NCBI) and with a genome size of 3.4 x 109 bp, the human genome will produce an SNP approximately once every 22 bases. Other-model systems exhibit a similarly high SNP frequency (Prediger et al., 2017). The features of the variations may be unique or occur in many individuals. Most commonly, these variations are found in the DNA between genes. They can act as biological markers to locate genes that are associated with disease. A more direct role can be played by SNPs occurring within a gene or in a regulatory region near a gene by affecting the gene's function (NCBI dbSNP build 2017; Gregory 2005).

Such change in the CYP2C19 gene, the allelic variant CYP2C19*3 effects the bioactivation of its metabolizing drug by causing poor metabolism. The allele frequencies of this SNP are about 2 percent of Caucasians, 4 percent of African Americans, and 14 percent of Chinese. One widely tested non-functional variant for CYP2C19 is CYP2C19*3 which contains a c.636G>A variant in exon 4 which causes premature stop codon. As noted above, ACS / PCI patients that are poor metabolizers with CYP2C19 who are treated with clopidogrel have elevated risks for significant cardiovascular vents like stent thrombosis relative to patients with equivalent care without a non-functional allele. For Asian populations, the CYP2C19*3 allele levels are ~2–9 per cent (Morgantown et al., 2017).

The above clinical significance was the target to work on this SNP considering its loss of function activity to the metabolizing drug clopidogrel which is vastly used (75mg/300mg) to treat cardiovascular issues (Yusuf et al., 2001).

The further approach of this work is to genotype the SNP to get the overall allelic frequency rate of CYP2C19*3 variant in different ages of individuals among the Bangladeshi population. Allele frequency measures in a population of genetic locus of the relative frequency of an allele. It is usually expressed as a proportion or percentage. If the allele frequency of CYP2C19*3 variant can be measured it can show the genetic diversity of a species population or equivalently the richness of its gene pool. The different "forces" that might lead to changes in the distribution and frequencies of alleles can be done by the population genetics study. To relate genotype frequency to allele frequency Hardy Weinberg principle model can be applied. Hardy Weinberg principle can determine how much variation exists in natural populations throughout the gene pool by a calculative mathematical formula (Genet 2003; Khan Academy, 2020).

As the genotyping was another main approach regarding this study, the effect of this variant into individuals which lower its associated drug's functioning and as no particular report were submitted regarding this CYP2C19*3 variant among Bangladeshi population with its genotype frequency was the motivation to work on this study.

The main intention of this study is to introduce a cost-effective method to detect this particular SNP to get benefitted in an individual's treatment with their personal genotyped result. For doing these various processes are developed to do the genotyping. As mentioned before some genotyping methods are mostly expensive, requires sophisticated laboratory platform and more over for routine diagnosis requires high throughput experiments which are not regularly available procedures to work in any diagnostic laboratory for a country like Bangladesh. For mass spectrometry it requires exploration in order to understand what the different analytical methods can reveal about experiment. For micro array procedure it includes the high cost of a single experiment, to design probe based on sequences of low specificity it requires the large number of designed probe and as the commonly used microarray platforms utilize only one set of probes designed by the manufacturer because of the lack of control over the pool of analyzed transcripts. Other disadvantages of microarrays are their relatively low accuracy, precision and specificity along with it to get the variations in hybridization by the high sensitivity of experimental set up.

Considering the genetic material in micro array with its amplification process together with other factors may effect on the estimation of gene expression (Mao et al., 2020).

Considering all these aspects the study was conducted with Tetra-Primer ARMS PCR. The PCR unique to alleles is an effective tool for single-gene disorders with the SNPs. In recent days, the ARMS PCR is one of the essential methods for the diagnosis of genetic disorder. The method of restriction digestion is not 100 per cent effective and not all sequences have the site of restriction. It is also a form of the gold standard for Sickle cell anemia and thalassemia, including genetic condition. This is also important in the identification of JAK2 and HIV mutations. Commercial kits for the various disorders such as HbS and JAK2 gene SNPs detection kits are now available. S.Kwok and colleagues used ARMS-PCR in module studies of type 1 HIV. This method was used by the pioneers C.R.Newton and coworkers for the SNP study of the ATT gene deficiency. The important intention to use this PCR technique was to work in an easy method which would be cost effective and can be done in any molecular biology laboratory or in any diagnostic laboratory in a country like Bangladesh. The PCR is run for about an hour followed by agarose gel electrophoresis which is a less time consumable process (What is PCR or Allele Specific PCR? 2019).

Allele specific ARMS PCR process is based on the adjustment of the primers for different alleles. Here, the 3' end of the primers is changed such that the regular allele can be amplified by one set of the primers and the mutant allele can be amplified by others. The single base mismatch is added at the 3' end of the primordial. Such mutation causes the primordial to amplify a single allele. The mismatch between the primer and the DNA template plays a crucial role in achieving the amplification here. The addition of a mismatch at the 3’ends of the priming changes the temperature of the annealing for that particular allele. If the mismatch is low the amplification chances are increasing. By adding a strong mismatch near the 3 'end of the primer (ideally at -2 position) it can't bind to the DNA sequences in the non-complementary allele and thus terminates amplification. C:T, G:A and A:G are the strong base pairs that decrease the amplification cycle by up to 100 fold. In general, the reverse primer or another primer (which is not modified) remains the same. In addition, the primer must satisfy all the desired priming requirements. By this feature allele specific ARMS PCR is excel with low cost efficiency within less time.

The four primer sets were designed by the gateway website BatchPrimer3. From there, with the best parameters for functioning the allele specific PCR primers were supposed to detect wild type

homozygous and mutant heterozygous bands to identify the SNP variant among individuals. The PCR bands got from this study were wild type homozygous bands in every optimization step of PCR. They were all faint bands. From the fainted band result it cannot be said that the designed primers are working efficiently to get the best result. For this some more optimization steps were needed to confirm the thermal cycle. The wild type homozygous bands indicated the amplification of normal gene type in individuals though it should be done with much more proper results of PCR to do the SNP genotyping.

The further approach of this study was to validate by sequencing. It allows for speedier and more cost-effective sequencing of different genome areas for in-depth research. It provides accuracy to customize fine mapping in a cost-effective way. For better characterization of variants, reducing false positives and maximizing capture of population structural differences are the implications of sequencing. The validation of genotyping gets more accuracy by sequencing.

4.2 In silico Analysis:

The designed primers were in the ideal range of their parameters to function properly. For adequate specificity primer length should be long enough and short enough for primers to bind easily to the template at the annealing temperature. The ideal primer length is between 18-24 bp. All the primers used in this study were 24bp only the inner reverse primer was 28 bp because it would give the allele specificity by having a deliberate mismatch at 3' end. GC content of primer is a percentage of G's and C's of the total bases which should be between 40%-60%. The primer used in this study had GC content of 42% on average which satisfy the ideal condition. Primer melting temperature (Tm) is the temperature at which one half of the DNA duplex will dissociate to become single stranded and indicates the duplex stability. Generally, the primers with melting temperature are between 52-58°C producing the best result. The primers used in this study had a melting temperature of 70°C which is a bit higher than the ideal range. But this should be kept in consideration that parameters differ from the ideal range can also produce best results with enough optimization steps (Primer Designing Guide for PCR n,d.).

The inner forward primer (WT) and the inner reverse primer (MT) respectively had the product size of 278 and 145. These amplicon sizes produce in the PCR reaction can be easily resolved in the agarose gel because they differ in their product sizes of more than 50 bp. The inner forward primer had a product size of 184 bp which also differ in the size of more than 50 bp to resolve in

an agarose gel electrophoresis. Overall, this set of primers were found to be the best functioning primers theoretically to detect the CYP2C19*3 SNP.

4.3 In vitro Analysis:

In the optimization step of PCR reactions, the annealing temperature were set in the first few attempts were 60°C, 59°C, 61°C. These annealing temperatures were selected considering the other parameters of primers GC content, Tm etc.

In the first attempt at 60°C annealing temperature the PCR products were really fainted. For the detection of the CYP2C19*3 variant total of 4 samples were run for PCR reaction but it gave result for only 2 samples. The possible reason can be the DNA templates for the other two samples were less in amount. The faint bands were wild type homozygous bands. They were amplified result of normal gene. The PCR bands were below compared to the ladder. For this reason, product size could not be optimized and it ensures that the bands were nonspecific or formed primer dimer.

In the next attempt of PCR reaction, thermal cycle's annealing temperature was set at 59°C. The annealing temperature was lowered because to bind the primer to the template region. Sometimes too high annealing temperature unable the primer to bind in the template region. In this step the bands were also fainted but quiet more visible than the previous attempted result. There were a total of 4 samples given in the PCR reaction but it gave result for only 3 samples. Two of them gave wild type homozygous band and another one gave smear result. The causes for getting smear result in PCR can be the template concentration was too high or not diluted properly. So, templates overload probably a reason for getting smear result. Another reason can be the improperly prepared gel. The gel will not polymerize or solidify evenly if the gel is not poured correctly. If the wells are filled too much excess sample can cause smear. To determine the product size 100 bp ladder was used. In this attempt the ladder got too dissolved into the gel. The PCR products were below compared to the ladder which clearly indicates that the bands were nonspecific or formed primer dimer.

In the third attempt of PCR to avoid nonspecific band formation and to get a more specified result the annealing temperature was increased to 61°C. But it certainly did not give any band in PCR product resolving in the agarose gel. The probable reason can be the annealing temperature was

too high to bind the band to the template region. This could be the reason for not getting any result from this thermal cycle.

In the fourth attempt, the PCR cycle produced wild type homozygous band for two samples. This time the bands were fainted but quite specific in their product size. From the 100 bp ladder comparison the product sizes were 278 and 184 which were amplification of normal gene. Here in this step another modification was done in the elongation step. In the previous attempts the elongation was for 15 seconds but in this step, it was increased to 30 seconds. This can elongate the primers in more time to complete one cycle.

So, from the PCR optimization step it can give the idea that the individuals who were detected for CYP2C19*3 variant were carriers of normal gene. Though the results could be more specified and accurate to get the overall genotype frequency among target population.

The individuals who were tested for this particular variant are aged between 23-24 years and have no other major health issues like diabetes, breathing problem or any adverse habit like smoking. Some of them have family history of cardiovascular disease, diabetes but it has not developed into them. So, the participants were healthy individual who had normal type of gene regarding this SNP.

In the sample collection section, it was a gargle sample collection method which is described earlier. Here, in the gargle collection from participants were strictly monitored for 1 minute because this is the enough time to get good amount of inner cheek cells. Time less than that might not able to get into good amount of inner cheek cells.

In DNA extraction, the very first step was to centrifuge 1500µL samples at 4500 rpm for 5 minutes. But the step was repeated for 4/5 times until getting a good pellet. Because a good amount of pellet increases the cell concentration. Another modification was done in extraction procedure. The water bath step mentioned in the protocol was for about 1 to 3 hours. For this study, water bath (65°C) was set for 2 hours because hot water bath softens the cell walls, and membranes for releasing the DNA. Along with this it further deactivates the enzymes in the mixture that can degrade DNA. Above all this longer heating can denature the DNA. That is why 2 hours was good enough to work with.

Before collecting samples in gargle sample method there was some trial and error going on. At first, it was thought that the samples should be collected by buccal swab method and the extraction should be done by using a kit. But in case of DNA extraction from buccal swab method there were some difficulties which changed the way for collecting samples and to extract the DNA in manual procedure that is described earlier in the study.

The problem faced in buccal swab sample's DNA extraction was, after collecting the samples it was placed in shaker incubator for 2 to 3 hours by placing the swab in the PBS solution. This step was done to bring out the cells from the swab to the solution. After this step the whole procedure was done according to the kit's given protocol but it did not result in any sort of extracted DNA when it was resolved in agarose gel. The probable cause for not getting any result from this procedure can be that not enough cells were out form the swab to the solution to get desired DNA. For fixing this issue the swabs were kept in the shaker incubator by increasing the time to overnight but the result was the same as before. That is why the collection and DNA extraction method was shifted to gargle sample collection and manual extraction procedure.

SNP genotyping, Allele frequency, validation by sequencing could finish the study in an accomplished manner. But due to the unavoidable reasons of COVID-19 and the severity of the pandemic, our university was shut down along with its all physical presence activities. The wet lab experiment had to stopped in the mid-way with no further continuation to get the proper and final result of this project. To fulfill the motivation of this study more thermal cycle optimization and confirmation steps were needed. If this could be accomplished it could find out the allele frequency and validate the study through sequencing. All this work could have been completed the study by bringing an effective outcome.

CHAPTER 5

CONCLUSION

Having medication to get improve from a particular disease is necessary. The medication is working efficiently or not has become an important issue nowadays. Sometimes the drug given to treat a particular disease get altered metabolism by its associated gene polymorphism which will not give any fruitful outcome from the medication provided. For this, SNP genotyping in individuals are holding great importance nowadays to have the personalized medicational approach. The test can bring the clear idea of percentage population carrying polymorphism which alters drug response and can benefit them in their treatments given.

The test is conducted with regular diagnostic applications like DNA extraction, gel electrophoresis, polymerase chain reaction which are mostly available in a country like Bangladesh. Rather than approaching to other expensive methods like mass spectrometry, micro arrays which requires costly materials and platforms working with tetra primer ARMS PCR is way cheaper and cost-effective method.

ARMS PCR is done by the basis of amplifying all wild type and mutant alleles together inside a single tube PCR reaction. The device uses four primers to amplify a non-allele-specific, larger fragment containing the mutation site and allele-specific amplicons for each of the two forms of allelic type. Two allele specific (inner) primers are built in opposite orientation and can simultaneously amplify both the wild-type and the mutant amplicons in conjunction with the typical primers. As primers can be constructed to amplify fragments of different sizes for each allele, a standard agarose gel electrophoresis can be used to isolate the amplicons easily and interpret the result. This specificity of this allele specific PCR technique is efficient and easy to conduct.

The availability of CYP2C19 genetic testing, however, may promote personalized antiplatelet therapy with early detection of individuals with impaired CYP2C19 activity.

Recent studies have shown that CYP2C19-genotype controlled antiplatelet therapy results in a greater likelihood of achieving a therapeutic level of platelet reactivity on treatment which can also be cost-effective for ACS patients undergoing Chemotherapy.

Bibliography:

1. Angiolillo DJ, Fernandez-Ortiz A, Bernardo E, Alfonso F, Macaya C, Bass TA, et al. (2007). Variability in individual responsiveness to clopidogrel: clinical implications, management, and future perspectives. National Library of Medicine. J Am Coll Cardiol. 10;49(14):1505–16.
2. Bhatia, Dharminder & Wing, Rod & Singh, Kuldeep. (2013). Genotyping by sequencing, its implications and benefits. Crop Improvement. 40, 101-111.
3. Ching-Wan Lam, Sui-Fan Tong. (2010). Microarrays for Personalized Genomic Medicine. Advances in Advances in Clinical Chemistry. 52, 1-18.
4. Collet JP, Hulot JS, Anzaha G, Pena A, Chastre T, Caron C, et al. (2011). High doses of clopidogrel to overcome genetic resistance: the randomized crossover CLOVIS-2 (Clopidogrel and Response Variability Investigation Study 2). JACC Cardiovasc Interv. 4(4): 392–402.
5. Derek G. Waller, DM, FRCP, Anthony P. Sampson. (2018). Pharmacokinetics. Medical Pharmacology and Therapeutics (Fifth Edition), 33-62.
6. Frank M You, Naxin Huo, Yong Qiang Gu, Ming-cheng Luo, Yaqin Ma, Dave Hane, Gerard R Lazo, Jan Dvorak and Olin D Anderson. (2008). BatchPrimer3: a high throughput web application for PCR and sequencing primer designing. BMC Bioinformatics. 9:253. doi:10.1186/1471-2105-9-253.
7. J Hum Genet. (2003)48, 654-658. doi,org/10.1007/s10038-003-0096-1.
8. Janice Y. Chyou, ... Marc S. Sabatine. (2013). A Companion to Braunwald's Heart Disease. Cardiovascular Therapeutics: (Fourth Edition).
9. Katrin Sangkuhl, Teri E. Klein, and Russ B. (2010). Altmana, Pharmacogenet Genomics. 20(7): 463–465.
10. Kazui M, Nishiya Y, Ishizuka T, Hagihara K, Farid NA, Okazaki O, et al. (2010). Identification of the human cytochrome P450 enzymes involved in the two oxidative steps in the bioactivation of clopidogrel to its pharmacologically active metabolite. Drug Metab Dispos. 38:92–99.
11. Kazuyuki Matsuda. (2017). PCR-Based Detection Methods for Single-Nucleotide Polymorphism or Mutation: Real-Time PCR and Its Substantial Contribution Toward Technological Refinement. ;80:45-72.

12. Khan Academy, (2020). Allele frequency & the gene pool.
13. Laura Dean, MD. (2018). Clopidogrel Therapy and CYP2C19 Genotype. Medical Genetics Summaries - NCBI Bookshelf.
14. Lins R, Broekhuysen J, Necciari J, Deroubaix X, Semin Thromb Hemost. (1999). Pharmacokinetic profile of 14C-labeled clopidogrel. 25 Suppl 2:29-33.
15. Massimo Negrini. (2016). SNP Arrays. 5(4): 27.
16. Mao, Young & Lu. (2020). The application if single nucleotide polymorphism microarrays in cancer research. Curr Genomics. 8(4): 219–228. doi: 10.2174/138920207781386924
17. Morgantown, WV. (2017). CLOPIDOGREL BISULFATE- clopidogrel tablet, film coated [package insert].
18. Moriyama B., Obeng A.O., Barbarino J., Penzak S.R., et al. (2017). Clinical Pharmacogenetics Implementation Consortium (CPIC) Guidelines for CYP2C19 and Voriconazole Therapy. Clin Pharmacol Ther. 102(1):45-51. doi: 10.1002/cpt.583. Epub 2017 Apr 18.
19. Prediger E. (2017). Integrated DNA Technologies. Coralville, [Online].
20. Pusch W, Kraeuter KO, Froehlich T, Stalgies Y, Kostrzewa M, Biotechniques. (2001). Genotools SNP manager: a new software for automated high-throughput MALDI-TOF mass spectrometry SNP genotyping. 30(1):210-5.
21. Ruan Felipe Vieira Medrano & Camila Andréa de Oliveira. (2014). Guidelines for the Tetra-Primer ARMS–PCR Technique Development. Molecular Biotechnology 56: 599–608.
22. Savi P, Herbert JM, Pflieger AM, Dol F, Delebassee D, Combalbert J, Defreyn G, Maffrand JP, Biochem Pharmacol. (1992). Importance of hepatic metabolism in the antiaggregating activity of the thienopyridine clopidogrel. 44(3):527-32.
23. S.C.Turfus, R.Delgoda, D.Picking, B.J.Gurley, 2017, Pharmacokinetics, Pharmacognosy, Fundamentals Applications and Strategies, 495-512.
24. Small DS, Farid NA, Payne CD, Konkoy CS, Jakubowski JA, Winters KJ, et al. (2010). Effect of intrinsic and extrinsic factors on the clinical pharmacokinetics and pharmacodynamics of prasugrel. Clinical pharmacokinetics. 49(12):777–98.
25. Stan K. Bardal, Douglas S. Martin. (2011). Applied Pharmacology.

26. T. Ryan Gregory, James A. Nicol, Heidi Tamm, Bellis Kullman, Kaur Kullman, Ilia J. Leitch, Brian G. Murray, Donald F. Kapraun, Johann Greilhuber, Michael D. Bennett. (2007). Eukaryotic genome size databases. Nucleic Acids Research. 35: D332–D338.
27. Targeted Next Generation Sequencing, (n.d.)
28. Understanding Genotyping, (n.d.)
29. What is Pharmacology? (n.d.)
30. What Are Single Nucleotide Polymorphisms? (n.d.)
31. What is Genotyping? (n.d.)
32. What is PCR or Allele Specific PCR? (2019).
33. Yun Li, corresponding author Wei Chen, Eric Yi Liu, and Yi-Hui Zhou. (2013). Single Nucleotide Polymorphism (SNP) Detection and Genotype Calling from Massively Parallel Sequencing (MPS) Data. 5(1): 3–25.
34. Yusuf S., Zhao F., Mehta S.R., Chrolavicius S., et al. (2001). Effects of clopidogrel in addition to aspirin in patients with acute coronary syndromes without ST-segment elevation. N Engl J Med. 16;345(7):494–502.