DEVELOPMENT OF GENOTYPING SYSTEMS FOR
PHARMACOGENOMICS PROFILING

Fatima A. Eshumani

A thesis submitted in partial fulfilment of the requirements for the degree of
Magister Scientiae in the Department of Biotechnology,
University of the Western Cape

Supervisor: Professor Mongi Benjeddou
Co-supervisor: Dr. Bronwyn Kirby

November 2016
KEYWORDS

Pharmacogenetics

Single Nucleotide polymorphism

Polymerase chain reaction

Anticancer drugs

Cholesterol lowering drugs

Efficacy and toxicity

High resolution melt

Genotyping

Minor allele frequency

Haplotype
ABSTRACT

Genetic variability in genes encoding drug metabolizing enzymes, transporters and targets are known to be the main factors of inter-individual differences in therapeutic outcome. Genetic factors are estimated to be responsible for about 15-30% of inter-individual variation in drug disposition and response. Single-nucleotide polymorphisms (SNPs) are the most prevalent class of genetic variation that could explain the variability in drug efficacy and undesired side effects for patients.

The aims of this study were to develop and evaluate the performance of robust and high throughput techniques for genotyping ten polymorphisms related to anticancer drugs and ten polymorphisms related to cholesterol lowering drugs. SNaPshot minisequencing and high resolution melt analysis (HRM) genotyping panels were developed, optimized, and their performances were evaluated and compared. SNaPshot minisequencing systems were developed and successfully optimized for the genotyping of ten SNPs associated with anticancer drug therapy, and ten SNPs associated with cholesterol lowering drugs. These systems were used to genotype the selected SNPs in 130 healthy Cape Admixed participants residing in Cape Town, South Africa. Population genetics data obtained for the studied SNPs were analysed using several statistical analysis software tools. Important population genetic parameters were calculated. Among others, allelic and genotypic frequencies were determined and compared with other populations in the world.

High resolution melt analysis (HRM) genotyping panels were developed, optimized and their performance were evaluated and compared to the SNaPshot assays. HRM was explored as an alternative inexpensive and rapid methodology to genotype five SNPs related to anticancer therapy and five SNPs related to cholesterol lowering therapy (statins). Unlike the SNaPshot
assays, rigorous optimization was required for the detection heterozygous genotypes via HRM.

Both assays were validated using direct sequencing and compared to each other. The HRM system is a closed tube, cheap and (theoretically) rapid method for identifying genetic variations. HRM was however found to be more time consuming, needed further optimization, primer redesigning and more evaluation.

The developed genotyping systems could be further validated using clinical samples from patients. This could help in optimizing drug therapy for cancer and cholesterol treatment.
Declaration

I declare that ‘Development of genotyping systems for Pharmacogenomics profiling’ is my own work that has not been submitted for any degree or examination in any other university and that all the sources I have used have been indicated and acknowledged by complete references.

Full Name: Fatma A. Eshumani

Signature:

Date: November 2016
ACKNOWLEDGEMENTS

First and foremost all praises and gratitude to Allah Almighty for sustaining me through my study, only through your strength and guidance was I able to persevere and complete my Masters.

I would like to express my deep gratitude to my supervisor Prof. Mongi Benjeddou and my co-supervisor Dr. Bronwyn Kirby, for giving me the opportunity to do master degree and for their guidance, encouragement and advice.

I am sincerely grateful to my husband and my brother Jamal for their love, invaluable support and endless patience.

I owe a huge amount of gratitude to my beloved parents, the rest of my brothers, sisters and my husband’s family for their kind words, support, prayers and unwavering confidence in me.

A special thank you to my colleagues in pharmacogenetics lab as well as all staff in Biotechnology department for all the assistance you have given me during the duration of my study.

Finally, I would like to thank the education ministry of Libya and the Libyan embassy in South Africa for financial assistance.
DEDICATION

I dedicate this thesis to

My parents

My husband

And

My little daughter, Asmaa

The Prophet Muhammad (Peace Be Upon Him) said, regarding knowledge:

“Acquire knowledge and impart it to the people...One who treads a path in search of knowledge, has his path to Paradise made easy by God” (Sunan Tirmidhi, Hadith 107; Riyadh us-Saleheen, 245).
# LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>BLAST</td>
<td>Basic Alignment Tool</td>
</tr>
<tr>
<td>Bp</td>
<td>Base pair</td>
</tr>
<tr>
<td>dbSNP</td>
<td>Database of Single Nucleotide Polymorphisms</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>EXO I</td>
<td>Exonuclease I</td>
</tr>
<tr>
<td>FDA</td>
<td>US Food and Drug Administration</td>
</tr>
<tr>
<td>HRM</td>
<td>High-Resolution Melt</td>
</tr>
<tr>
<td>HWE</td>
<td>Hardy-Weinberg equilibrium</td>
</tr>
<tr>
<td>MAF</td>
<td>Minor Allele Frequency</td>
</tr>
<tr>
<td>MgCl₂</td>
<td>Magnesium Chloride</td>
</tr>
<tr>
<td>NCBI</td>
<td>National Centre for Biotechnology Institute</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>SAP</td>
<td>Shrimp Alkaline Phosphatase</td>
</tr>
<tr>
<td>SNP</td>
<td>Single nucleotide polymorphism</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

| Figure 1.1 | Chemical structure of Cholesterol. | 17 |
| Figure 1.2 A | Structure of lipoproteins. | |
| Figure 1.2 B | Relationship of lipoproteins components to diameter and density. | 18 |
| Figure 1.3 | Structure of HMG-CoA and different types of statins. | 21 |
| Figure 2.1 | Electropherogram profile of the SNPs related to anticancer drugs. Green peaks indicate an A nucleotide, blue peaks indicate a G nucleotide, black peaks indicate a C nucleotide, and red peaks indicate T nucleotide. | 38 |
| Figure 2.2 | Allele frequencies of the selected SNPs related to anticancer drugs in the Cape Admixed population compared to other ethnic groups. | 43 |
| Figure 3.1 | Electropherogram profile of the SNPs related to cholesterol lowering drugs. | 57 |
| Figure 3.2 | Allele frequencies of the selected SNPs related to cholesterol lowering drugs in the Cape Admixed population compared to other ethnic groups. | 61 |
| Figure 4.1 | Gradient PCR of the SNP (rs60140950) on 2% agarose gel. | 73 |
| Figure 4.2 | MgCl₂ optimization of the SNP (rs60140950) on 2% agarose gel. | 74 |
Figure 4.3  HRM mutation scan of SNP rs1801133.  

Figure 4.4A  HRM mutation scan of SNP rs4149036.  

Figure 4.4 B  HRM mutation scan of SNP rs2306283.
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 2.1</td>
<td>A list of genes and selected SNPs associated with anticancer drugs.</td>
<td>31</td>
</tr>
<tr>
<td>Table 2.2</td>
<td>Multiplex PCR primers for the selected SNPs related to anticancer drugs used in SNaPshot™ genotyping.</td>
<td>33</td>
</tr>
<tr>
<td>Table 2.3</td>
<td>SNaPshot minisequencing reaction primers of the selected SNPs related to anticancer drugs.</td>
<td>36</td>
</tr>
<tr>
<td>Table 2.4</td>
<td>Genotype and allele frequencies of selected SNPs related to anticancer drugs in 130 healthy Cape Admixed individuals.</td>
<td>41</td>
</tr>
<tr>
<td>Table 2.5</td>
<td>Comparison of MAF of the selected SNPs related to anticancer drugs in the Cape Admixed population to other ethnic groups.</td>
<td>42</td>
</tr>
<tr>
<td>Table 3.1</td>
<td>Multiplex PCR primers for SNPs related to Cholesterol lowering drugs used in SNaPshot™ genotyping.</td>
<td>53</td>
</tr>
<tr>
<td>Table 3.2</td>
<td>SNaPshot minisequencing reaction primers of the selected SNPs related to cholesterol lowering drugs.</td>
<td>55</td>
</tr>
<tr>
<td>Table 3.3</td>
<td>Genotype and allele frequencies of the selected SNPs related to cholesterol lowering drugs in 130 healthy Cape Admixed individuals.</td>
<td>59</td>
</tr>
<tr>
<td>Table 3.4</td>
<td>Comparison of MAF of the selected SNPs related to cholesterol lowering drugs in the Cape Admixed population to other ethnic groups.</td>
<td>60</td>
</tr>
<tr>
<td>Table 3.5 A</td>
<td>Haplotype structure defined by 5 SNPs in the SLC01B1 gene in the Cape Admixed population.</td>
<td>63</td>
</tr>
<tr>
<td>Table 3.5 B</td>
<td>Haplotype structure defined by 5 SNPs in the SLC01B3 gene in the Cape Admixed population.</td>
<td>64</td>
</tr>
</tbody>
</table>
Admixed population.  

Table 4.1  A list of selected SNPs primers used in HRM analysis  

Table 4.2  Confirmation of SNaPshot genotyping using direct sequencing.  

Table 4.3  Performance of HRM in comparison to SNaPshot minisequencing.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title Page</td>
<td>i</td>
</tr>
<tr>
<td>Keywords</td>
<td>ii</td>
</tr>
<tr>
<td>Abstract</td>
<td>iii</td>
</tr>
<tr>
<td>Declaration</td>
<td>v</td>
</tr>
<tr>
<td>Acknowledgements</td>
<td>vi</td>
</tr>
<tr>
<td>Dedication</td>
<td>vii</td>
</tr>
<tr>
<td>List of Abbreviations</td>
<td>viii</td>
</tr>
<tr>
<td>List of Figures</td>
<td>ix</td>
</tr>
<tr>
<td>List of Tables</td>
<td>xi</td>
</tr>
<tr>
<td>Chapter one</td>
<td></td>
</tr>
<tr>
<td>Literature review</td>
<td></td>
</tr>
<tr>
<td>1.1. Introduction.</td>
<td>1</td>
</tr>
<tr>
<td>1.2. Human genetic variation</td>
<td>2</td>
</tr>
<tr>
<td>1.2.1. Insertion and deletion variation</td>
<td>3</td>
</tr>
<tr>
<td>1.2.2. Short Tandem Repeats.</td>
<td>3</td>
</tr>
<tr>
<td>1.2.3. Single nucleotide polymorphisms.</td>
<td>4</td>
</tr>
<tr>
<td>1.3. Precision medicine and Pharmacogenetics / Pharmacogenomics</td>
<td>4</td>
</tr>
<tr>
<td>1.4. Pharmacogenetics and ethnic background.</td>
<td>6</td>
</tr>
<tr>
<td>1.5. Genetic variation and drug response.</td>
<td>7</td>
</tr>
<tr>
<td>1.6. Drug transporters.</td>
<td>7</td>
</tr>
</tbody>
</table>
1.7. Drug metabolizing enzymes.
1.8. Drug targets.
1.9. Pharmacogenetics profiling systems.
   1.9.1. Allele specific PCR.
   1.9.2. SNaPshot minisequencing (single base extension reaction).
   1.9.3. High resolution melt analysis.
1.10. Pharmacogenetics of anticancer drugs.
   1.10.1. Genetics of Cancer.
   1.10.2. Cancer treatment and side effects.
1.11. Pharmacogenetics of cholesterol lowering drugs.
   1.11.1. Cholesterol.
   1.11.2. Lipoproteins.
   1.11.3. Familial cholesterolemia.
   1.11.4. Cholesterol lowering drugs.
1.12. Genetic polymorphisms of drugs metabolizing enzymes, transporters and targets.
   1.12.1. Genetic variations associated with FDA approved anticancer drugs.
      1.12.1.1. Catechol O-methyltransferase enzyme.
      1.12.1.2. Dihydropyrimidine dehydrogenase enzyme.
      1.12.1.3. Thiopurine S-methyltransferase enzyme.
      1.12.1.4. Transmembrane Protein 43.
      1.12.1.5. Methylene tetrahydrofolate reductase enzyme.
1.12.1.6. Glutathione S-transferase P enzyme. 24

1.12.2. Genetic variations associated with FDA approved cholesterol lowering drugs. 24

1.12.2.1. Solute carrier organic anion trasporter1B1 gene. 24

1.12.2.2. Solute carrier organic anion trasporter1B3 gene. 25

1.13. Statistical analyses. 25

1.13.1. GenAlEx. 25

1.13.2. The Hardy-Weinberg equilibrium. 25

1.14. The aim of the study. 26

Chapter Two Development of a pharmacogenomic profiling panel for anticancer drugs.

2.1. Introduction. 27

2.2. Materials and Methods 29

2.3. Results 37

2.4. Discussion 44

2.5. Conclusion 49

Chapter Three Development of a Pharmacogenomic profiling panel of Cholesterol lowering drugs

3.1. Introduction 50

3.2. Materials and Methods 51

3.3. Results 56
Chapter One

Literature review

1.1. Introduction

The improvement of new methods for high-throughput single nucleotide polymorphism (SNP) analysis is one of the most exciting areas in genetic studies (Quintáns et al. 2004). Recently, countless methodologies for high-throughput SNP analysis have been developed such as: FRET analysis by real time PCR, DNA microarrays, Pyrosequencing, MALDI-TOF spectrometry and TaqMan probes or Molecular Beacons (Carracedo et al. 1998; Syvänen 2001). The application of these methodologies has revolutionized many of the biomedical sciences. However, it has had a major influence when applied to the molecular study of human genetic variations (Carracedo et al. 1998).

There is increasing interest in SNP typing due to their application in pharmacological response and genetic disease diagnosis (Evans and Relling 1999). It is well known that individuals’ DNA differs which makes some people more susceptible to specific diseases than others (Pennisi 2007). It has long been recognized that genetic variations in drug metabolizing enzymes, transporters and targets underlie the inter-individual variability in drug response. These differences that influence drug response might be due to age, sex, body weight, nutrition, organ function, infections, concomitant medication, environmental factors, and genetic makeup (Sadée and Dai 2005; Jacobs 2014). Hence, the goal of both Pharmacogenetics and Pharmacogenomics (PGXs) is to carefully study these variations to improve drug efficacy (Lee et al. 2005) with the ultimate aim to individualize therapy by
choosing the right drug for each patient by classifying them into genetically definable groups that have similar drug responses (Kalow 2002). The advantage of studying human genetics is the discovery and description of the genetic contribution to many human diseases. This is a motivation for increasing studies to define the contribution of several genes to the development of diseases such as cancer, heart disease, and diabetes (Pennisi 2007).

Africa has a high level of genetic diversity, but in spite of this diversity, population genetic studies are limited (Henry et al. 2008). South Africa is the home of native and immigrant population groups (Benjeddou 2010). The mix of these different ethnicities led to establishment of a unique population called Cape Coloured (Abrahams et al. 2011) which are designated here as Cape Admixed. The genomic intermixture in the Cape Admixed can form an essential foundation for subsequent pharmacogenetics studies in order to predict individualised therapy specially in Africa (Ikediobi et al. 2011). Note, that in this study the term “precision medicine” will be used interchangeably with “individualised” or “personalised therapy”.

1.2. Human genetic variations

Human genetic variation is the differences in the occurrence of mutations between the genomes of individuals and populations. The genomes of any two individuals differs at approximately one in every 1,000 base pairs (Venter et al. 2001) which results in an identity of up to 99.9% (Kruglyak and Nickerson 2001). Variation can appear in different ways, it can be in a single base pair, insertions/deletions of many nucleotide-long fragments of deoxyribonucleic acid (DNA) or loss of entire genes (Pelak et al. 2010). The rate of human genetic variation due to differences between populations is modest, and individuals from different populations can be genetically more similar than individuals from the same
population (Witherspoon et al. 2007). Major variations can be found within populations which can provide beneficial information about the histories and backgrounds of humans (Cavalli-Sforza et al. 1994). Human genetic variation has a direct influence on a wide range of biological and medical disciplines. Thus, the study of human genetic diversity is relevant to several research areas including human and population genetics, molecular biology, evolutionary biology, biological anthropology, health sciences and clinical medicine (Benjeddou 2010).

1.2.1. Insertion and deletion variation
Insertion and deletion (INDEL) variation refers to a type of genetic variation in which a specific nucleotide base pair is added or removed at some location of the DNA molecule during replication (Mills et al. 2006). INDELs are widely spread in the genome as they encompass a total of 3 million of the 15 million known genetic variants in humans (Mills et al. 2006). There are five types of INDELs, including insertions and deletions of single-base pairs, monomeric base pair expansions, multi-base pair expansions of 2–15 bp repeat units, transposon insertions, and INDELs containing random DNA sequences (Mills et al. 2006).

1.2.2. Short Tandem Repeats variation
Short Tandem Repeats variation (STRs) are regions of DNA composed of short nucleotide sequence (2 to 6 base pairs long) repeated many times (Willems et al. 2014). STRs are also considered to be the most polymorphic regions in the human genome, and play a role in genetic diseases and have been used in forensics, in determining genetic profiles and population genetics. Studies have shown that STRs sites are prone to mutations during DNA replication (Willems et al. 2014).

1.2.3. Single nucleotide polymorphisms
Single-nucleotide polymorphisms (SNPs) are the most prevalent class of genetic variation amongst people and are a valuable resource for mapping complex genetic traits (Marth et al. 1999; Reumers et al. 2008). SNPs are introduced by the substitution of one nucleotide with another at a particular locus within the DNA sequence (Simko et al. 2006). On average SNPs occur at approximately one in every 1000 bp (Komar 2009). SNPs could appear as bi-, tri-, or tetra-allelic polymorphisms. However, as tri-allelic and tetra-allelic SNPs are rare or almost do not exist in humans, SNPs are often described as bi-allelic markers (Brookes 1999).

SNPs in the coding region (cSNPs) can be either synonymous or nonsynonymous. Synonymous SNPs do not affect the protein sequence, while nonsynonymous SNPs change the amino acid sequence of a protein and may subsequently change its structure and function (Liao and Lee 2010). SNPs in regulatory regions of a gene can result in a disease phenotype because they may affect gene splicing, transcription factor binding sites, and other functional sites at the transcriptional level (Reumers et al. 2008). The majority of SNPs occur in the non-coding region of the human genome where they do not have direct effect on a phenotype (Liao and Lee 2010). Studies have shown that not all SNPs are useful genetic markers. However, SNPs found in the promoter and expressed regions have a direct impact on gene function and are considered to be worthwhile molecular markers (Paris et al. 2003). SNPs act as biological markers that assist scientists to distinguish genes that are linked to diseases (Cargill et al. 1999).

1.3. Precision medicine and Pharmacogenetics / Pharmacogenomics

Currently, the terms pharmacogenetics and pharmacogenomics which have so much in common are used interchangeably to refer to studies investigating the contribution of inheritance to variation in the drug response phenotype (Weinshilboum and Wang 2006). More accurately, pharmacogenetics specifically refers to approaches that are concerned with
studying the genetic influence of individual difference to drug responses, identifying variations that are caused by hereditary factors (Issa 2002), and uses a patient’s genetic information in order to improve and personalise drug treatment (Holm 2008). Conversely, pharmacogenomics is defined as all genes that influence drug responses, and uses genome wide analysis to identify such genes and search for novel drug targets (Srinivasan et al. 2009; Wu and Fuhlbrigge 2008). Generally, genetic factors are responsible for about 15 to 30% of differences in drug metabolism and rarely can account for up to 95% for certain drugs. These genetic differences can affect pharmacokinetics and pharmacodynamics (Choi and Song 2008; Nebert et al. 2008; Benjeddou 2010). Pharmacokinetics describes the process of drug transport in the body in terms of its absorption, distribution, metabolism, and elimination (Shell 1982; McLeod and He 2012), while pharmacodynamics is defined as the study of drug concentration and the pharmacological effects on the body (Craig 1993; Sinnollareddy et al. 2012). However, the variability in drug response can be due to non-genetic factors, such as age, concomitant diseases, food, environment factors and organ functioning (Choi and Song 2008; Nebert et al. 2008; Benjeddou 2010).

The focus of pharmacogenetics studies is pharmacokinetics, while pharmacogenomics focuses on how genetic variation is linked to an individual’s pharmacodynamics response (Kalow 2006; Weinshilboum and Wang 2006). The promise of both branches of study is to deliver personalised medicine which is based on clinical, genomic, genetic and environmental information of each patient (Limdi and Veenstra 2010). Such developments in pharmacogenetics and pharmacogenomics studies can be used to define diseases and biomarkers. This will lead to a patient-tailored therapeutic drug that hopefully would be more effective and will result in fewer adverse drug reactions (Kalow 2002; Oscarson 2003; Hedgcoe 2004). Ultimately, it is hoped that by knowing the important genetic variations that
exist between racial and ethnic groups, it will be possible to classify patients into low, intermediate and high dose groups (Shastry 2006).

1.4. Pharmacogenetics and ethnic background

Race, ancestry and ethnicity have been identified as major factors in pharmacogenetic discovery and improving our understanding of population differences in drug response (Urban 2010). This means that the variant that affects a response to a type of drug in one ethnic group may not have the same impact in another group, and this is due to gene-gene or gene-environment interaction differences (Tate and Goldstein 2004). For example, warfarin treatment shows a wide variance among patients of different racial groups, and this variance might be due to polymorphisms in the gene encoding for vitamin K epoxide reductase complex 1. One example of this is that heparin and warfarin dosages for Chinese patients are lower than those commonly prescribed for Caucasian patients (Daar and Singer 2005). Another example is that the cardiovascular drug BiDil, is recommended for African people because as it has low side effects compared with Caucasians (Taylor et al. 2004). Understanding ethnic genomic differences in populations is required to improve drug treatment and to establish public health policies (Shastry 2006; Suarez-Kurtz 2008b). Africa is a widely diverse landscape. Sub-Saharan Africans have the highest human genomic diversity in the world, however this diversity is understudied (Tishkoff et al. 2009; Benjeddo 2010). South Africa has a wide population diversity as a result of combination of different groups from Europe, Asia and Africa (Adhikari 2005; Benjeddou 2010). The unique South African Cape admixed population which is distinguished by wide ethnic intermixture diversity comprising about 9 - 11% Asian, 21 - 28% white, 32 - 43% Khoisan and 20 - 36% black populations (de Wit et al. 2010; Erasmus et al. 2012). The mixed ancestries in this population may help set the foundation for future pharmacogenetics traits studies, in order to
predict personalized therapy in diverse populations particularly in Africa (Ikediobi et al. 2011). This wide ethnic diversity supplies a valuable opportunity to study genomic patterns (Conrad et al. 2006). In addition, the complexity in the Admix populations creates an ideal situation for exploring the role of genetic ancestry of population, as well as clinical and environmental factors on human health (Suarez-Kurtz 2008a; Via et al. 2009).

1.5. Genetic variation and drug response

Inter-individual differences in drug disposition are substantial reasons for adverse drug reactions and deficiency in drug response (Sim et al. 2013). Despite that, genetic variation in the human genome plays a main role in variable response to drugs and other xenobiotics (Brockmöller and Tzvetkov 2008). Polymorphisms in drug metabolising enzymes, drug transporters, drug receptors, and ion channels genes can lead to adverse drug reactions or can change the efficacy of therapeutics (Meyer 2000). Many of the newly identified SNPs could play a significant role in the expression level and activity of the corresponding protein (Venter et al. 2001). Such polymorphisms once they occur in genes encoding drug transporters or drug metabolizing enzymes might change the disposition of the drug, and subsequently, its efficacy and toxicity may be affected (Robert et al. 2005).

1.6. Drug transporters

Transport proteins play an important role in absorption, distribution, and excretion of many medications. They play a significant role in drug response, serving as drug targets and determining the most effective dose of each drug (Leabman et al. 2003).

On the basis of function, membrane transport proteins are divided to major superfamilies which are the ABC (ATP-binding cassette) transporters, and the SLC (solute carrier) transporters. While the human ABC transporter family comprises eight sub-families with 48
members, the SLC family has 47 subfamilies with more than 365 members (Leabman et al. 2003). Genetic polymorphisms in drug transporters are considered to be the main factors affecting the variation of patients’ responses to medication (Kalliokoski and Niemi 2009).

Most of the SLC family members act as influx transporters for nutrients and substances essential for the cell, whereas ABC members function as efflux transporters for wasted metabolites and toxins, including many anticancer drugs (Nakanishi and Tamai 2011).

In humans, SLC transporters primarily are expressed in the liver, kidneys, lungs and intestines, but can be expressed in most tissues where they are either localized at the basolateral or apical plasma membranes of polarized cells (Wojtal et al. 2009). Furthermore, SLCs have two gene superfamilies which contain the major organic cation and anion transporters. The SLCO superfamily (previously called SLC21) consists of the organic anion transporting polypeptides (OATPs), while SLC22A that makes up the organic cation transporters (OCTs) and the organic anion transporters (OATs) (Roth et al. 2012).

SLCs membrane-associated transporters mediate the main physiological functions by influx and efflux of endogenous substrates including amino acids, lipids and bile acids, thyroid hormones and xenobiotics (Hagenbuch and Meier 2004; Shitara and Sugiyama 2006; Meier et al. 2007). These transporters are the main focus of many pharmacokinetics studies as they play a significant role in drug absorption and are therefore important determinants of drug distribution (Meier et al. 2007; Le Vée et al. 2015).

1.7. Drug metabolizing enzymes

Most xenobiotic compounds including drugs, environmental pollutants and endogenous compounds such as steroids and prostaglandin are metabolized by a diverse group of proteins called drug metabolizing enzymes (DMEs) (Sheweita 2000). Conceptually, DMEs can be
divided into two categories, the oxidative or phase I reactions, in which lipophilic substances are catalysed into more water soluble forms which are more easily excreted. These reactions can either lead to activation or inactivation of the compound. Phase I enzymes includes the cytochrome P450s (CYP), epoxide hydrolases (EH) and flavin mono-oxygenases (FMO) (Evans and Relling 1999; Pinto and Dolan 2011), while the conjugative or Phase II reactions increase the solubility of compounds, which is typically achieved by conjugation to a highly polar agent such as glucuronic acid. Phase II enzymes include thiopurine S-methyltransferases (TPMT), uridine glucuronyl transferases (UGT), N-acetyltransferases (NAT), glutathione S-transferases (GST) and sulphonyl transferases (SULT) (Sheweita 2000; Clouthier and Pelletier 2012). Most DMEs are polymorphic, appearing as small insertions and deletions, gene amplification and deletions, or as SNPs (Sim et al. 2013). Their presence can affect drug efficacy and toxicity (Serpe et al. 2014).

**1.8. Drug targets**

The total number of drug targets is currently limited for many reasons, including the inability to understand the existence of splice variants or interactions between the encoded proteins from gene sequences and the unknown function of most DNA in the genome (Imming et al. 2006). Drug targets have been classified into enzymes, substrates, metabolites and proteins, receptors, ion channels, transport proteins, DNA / RNA and ribosome, and the targets of monoclonal antibodies. In addition, some drugs act through physicochemical mechanisms, or have unrecognized mechanisms of action (Imming et al. 2006; Landry and Gies 2008). Genetic variability in drug targets can have an unequivocal impact on drug efficacy, and that can be exemplified by sequence variations in the gene for the β2-adrenoreceptor which influences its response to β2-agonists and the reno-protective action of angiotensin-
converting enzyme inhibitors are impacted by SNPs in the gene for angiotensin-converting enzymes (Wood et al. 2003).

1.9. Pharmacogenetics profiling systems

There are several techniques that have been developed for genotyping SNPs in recent years including: DNA Microarrays, MALDI-TOF spectrometry, FRET analysis, SNPllex, Pyrosequencing, TaqMan, High resolution melt technique (HRM) and Allele Specific Amplification (ASA) (De Monaco et al. 2014).

1.9.1. Allele specific PCR

Allele specific PCR is also called PCR allele specific amplification (PASA), is a method used for the identification of single nucleotide polymorphisms by using allele specific primers (Papp et al. 2003). A recently developed technique called PCR amplification of multiple specific alleles (PAMSA) involves using at least two allele specific primers in the same reaction and allows for the detection of all SNPs present in one sample (Ishiguro et al. 2005; Hansson and Kawabe 2005). Many PAMSA methods allow for the discrimination of alleles by gel electrophoresis. Liu et al. (1997) described a variant of PAMSA called bidirectional-PASA, in which one allele is amplified in one direction whereas the other allele is amplified in the opposite direction. However, for this method four primers are needed and accurate optimization of the target fragment can be time consuming and problematic (Liu et al. 1997; Sasvari-Szekely et al. 2000; Waterfall and Cobb 2001, 2002). In PAMSA, a primer length difference of between 2 and 5 nucleotides is used, and alleles are resolved on acrylamide gels. This method may require the molar ratio of allele-specific primers to be optimized (Okimoto and Dodgson 1996). Gaudet et al. (2007) described that three unlabelled primers were necessary to amplify specific alleles. The alleles-specific primers that were used had a
destabilizing mismatch within five bases of the 3’ end and a 5’ tail for the amplification of different length PCR products, which were detected on an agarose gel (Gaudet et al. 2007).

1.9.2 SNaPshot minisequencing (Single Base Extension reaction)

There is an increasing focus on SNPs in the last few years. Traditionally, SNP analysis was performed with costly and time consuming techniques based on conventional PCR followed by restriction fragment analysis and gel electrophoresis (Quintáns et al. 2004). Recently, many high-throughput methods for studying SNPs have been developed including next generation sequencing (NGS), MALDI-TOF spectrometry, DNA Microarrays, FRET analysis and TaqMan assays or Molecular beacons (Quintáns et al. 2004). Despite the growing number of techniques, one method which is growing in popularity is the SNaPshot minisequencing assay which relies on a single base extension with a labelled ddNTP (Brandstätter et al. 2003). It is also less time-consuming and relatively rapid, while remaining robust and accurate compared to the traditional SNP analysis techniques like gel visualisation (Syvänen 1999). The commercialized multiplex SNaPshot system has been used increasingly in various genotyping studies, because it is highly accurate that can effectively detect many polymorphisms in a single assay (Paneto et al. 2011; Li et al. 2012). Technically, this method uses only fluorescently labelled dideoxynucleotide triphosphates (ddNTPs) to identify the allelic variant during the elongation of a primer located beside the SNP of interest (Huang et al. 2011; Li et al. 2012). The chemical structure of ddNTPs (lacking a 2’ and 3’ hydroxyl group) prohibits any additional elongation, resulting in a product one nucleotide longer than the primer itself. This additional nucleotide can be visualized by a capillary automated sequencer. Each ddNTP is labelled with a specific fluorescent colour (fluorophore) for the different bases G,C,A,Ts which allows them to be identified (Rogers et al. 2012) and a fifth colour is used to label the internal size marker (Quintáns et al. 2004).
1.9.3. High resolution melt analysis

High resolution melt (HRM) analysis is a closed-tube assay (i.e. less chance of cross-contamination than other techniques), commonly used to detect mutations and polymorphisms in blindly screened DNA samples. The DNA is amplified by PCR with increasing temperature, in the presence of a saturating dye which only fluoresce in the presence of double stranded DNA followed by observing the progressive change in fluorescence during DNA denaturation in order to provide information about target DNA quantity (Ramón-Laca et al. 2014; Lim et al. 2015). The fluorescence data can be instantly analysed by using suitable software supplied by the manufacturer (Prajantasen et al. 2015). Basically, the HRM technique depends on the melting temperature of double stranded DNA, GC composition, sequence length and heterozygosity of the target (Reed et al. 2007; Baniecki et al. 2015). HRM analysis is rapidly becoming the favoured choice to detect pathogenic variants because it is simple, easy to use, cheap, flexible, non-destructive, sensitive and specific (Vossen et al. 2009; Vondráčková et al. 2015). Theoretically HRM is used to generate target amplicons less than 500bp, however due to its sensitivity and specificity, amplicons up to 1000 bp have been successfully amplified (Botezatu et al. 2015; Vondráčková et al. 2015). HRM has been successfully used to screen for mutations and genotyping of humans, animals, plants and microbes, as well as epigenetics and general population studies (Vondráčková et al. 2015). It is also a good method for individualized medicine to predict drug response (Reed et al. 2007).

1.10. Pharmacogenetics of anticancer drugs

Variability in response and toxicity of anticancer drugs are consistently observed among patients which is a real problem in clinical practice that may lead to therapy failure or adverse drug reactions (Bosch et al. 2006; Huang and Ratain 2009). However, genetic variations in
drug metabolizing enzymes, transporters and targets genes have major effects on the observed differences in chemotherapeutic outcome. The aim of studying pharmacogenetics is to identify the genetic base for these differences and use it to predict optimal personalized drug regimes and dosage for the patient (Lee et al. 2005; Huang and Ratain 2009). There are two approaches used to evaluate how genetic variation contributes to differences in drug response and toxicity in humans (Huang and Ratain 2009). The first approach focuses on candidate genes which are involved in human pathophysiology, pharmacology, and cancer biology, based on the important role that genetic variations play in the pharmacokinetics or pharmacodynamics of a drug and the likelihood of affecting its efficacy and/or toxicity (Huang and Ratain 2009). The second is genome-wide approach that refers to the study of genetic variations within the human genome for their impact on drug therapy, which any genetic variant in the genome can contribute to variation in drug action (Huang and Ratain 2009).

Pharmacogenomic approaches have been applied to several anticancer drugs in an attempt to determine related inherited variations that are able to predict a patient’s response to cancer drugs (Lee et al. 2005).

1.10.1. The Genetics of Cancer

Cancer is a common genetic disease that originally develops from normal cells that are able to proliferate and become abnormal. These cells keep dividing to form more cells without normal cell cycle controls, forming a solid mass of excess tissue called a tumor which can be either malignant (cancerous) or benign (not cancerous) (Weinberg 2013). Tumor suppressor genes and oncogenes are two essential classes of genes that play a role in the carcinogenic process. The epigenetic changes could be inherited from parents or somatic (acquired) changes that are associated with environment or lifestyle (Hemminki and Hemminki 2005).
About 5 to 10% of all cancers are caused by inherited mutations, while 90 to 95% are as a result of other environmental factors (Mgbakor et al. 2014). There is a noteworthy increase in cancer worldwide, specifically in less developed countries (Torre et al. 2015). It was estimated by the GLOBOCAN project (http://www.globocan.iarc.fr) that in 2008, 12.7 million new cancer cases and 7.6 million cancer deaths occurred worldwide, with 56% of the new cancer cases and 64% of the cancer deaths occurring in the less developed countries around the world (Ferlay et al. 2010; Jemal et al. 2011). Similarly, Torre and co-workers estimated that of the 14.1 million new cancer cases and 8.2 million deaths which occurred in 2012, 57% of the new cases and 65% of deaths occurred in the less developed regions of the world (Torre et al. 2015). This increase can be attributed to population growth, senility and harmful lifestyle (Jemal et al. 2011).

In Africa, cancer is increasing at an alarming rate. In 2008, about 715,000 new cancer cases and 542,000 cancer deaths occurred on the continent (Cancer), whereas 847,000 new cancer cases (6% of the whole world) and 591,000 deaths (7.2% of the whole world) were projected for 2012 (Parkin et al. 2014). As Africa has a wealth of extraordinary genetic diversity, it is thought to be an appropriate place to investigate cancer aetiology due to the impact of environmental factors on different genetic variations (Mgbakor et al. 2014).

1.10.2. Cancer treatment and side effects

Cancer treatment depends on the nature of the cancer and its progression. There are many types of cancer treatment including surgery, photodynamic therapy, chemotherapy, radiotherapy, immunotherapy, targeted therapy, hyperthermia, stem cell transplant, lasers, and blood product donation and transfusion (Thurston 2006).
Chemotherapy refers to the use of medicines to treat cancer. Chemotherapeutic agents are classed according to the cellular phase in which they are active (S phase, M phase, G1 phase, G2 phase dependent) (Pazdur 2005). Nonspecific agents that do not target a specific cell cycle phase such as alkylating agents, show a linear dose-response curve, so further drug dosages increases cell killing, whereas cell cycle specific drugs have a plateau with regards to cell killing ability, so increases in drug dosage will not result in increased cells being killed (Page and Takimoto 2002).

Chemotherapy drugs can be divided into many classes which are: Alkylating agents such as cyclophosphamide which is activated in the liver (Lind 2011); platinum compounds such as cisplatin, carboplatin and oxaliplatin; and Antimetabolites which are subclassified into Antifolates (e.g. Methotrexate), Antipyrimidines (e.g. fluorouracil) and Antipurines (e.g. mercaptopurine and thioguanine) (Lind 2011). Although anticancer therapeutics are effective, current anticancer treatments have several drawbacks which ultimately affect their efficiency including that they are non-specific, and kill both normal and cancerous cells (Kawabe 2004). Some patients reported that they have experienced on average of 20 symptoms during the treatment, of which 7 were psychosocial and 13 were physical (Griffin et al. 1996). The most common side effects reported were nausea, vomiting, sore mouth, tiredness, hair loss, difficulty sleeping, loss of taste, loss of appetite, passing more urine, dry skin and effects on their social lives (Griffin et al. 1996; Carelle et al. 2002). For most patients, vomiting is not as common or as acute as before due to improvements in antiemetic drugs. Nevertheless, vomiting is still a frequent symptom and in particular nausea remains a wide issue (Bloechl-Daum et al. 2006). Today, fatigue is other common side effect of anticancer drugs that affects about 80% of patients receiving chemotherapy (Henry et al. 2008).


1.11. Pharmacogenomics of cholesterol lowering drugs

The main factors that could impact the response of cholesterol lowering drugs (statins) include diseases, correlated with medications and adherence to treatment, biologic and physiologic conditions, and genetic background (Hutz and Fiegenbaum 2008). It has been suggested that genetic variations in influx and efflux transporters in the liver may modify the disposition of cholesterol lowering drugs. There are a considerable number of transporters that might have a role in statins transport, however relatively few have been analysed (Romaine et al. 2010). Few mutations have been reported that are medically related to SLCO genes. Several studies have focused on the impact of SLCO variants on drug disposition, taking in account pharmacokinetics of drugs (Hagenbuch and Stieger 2013).

The Genetic diversity in the solute carrier organic anion transporter 1B1 (SLCO1B1) and solute carrier organic anion transporter 1B3 (SLCO1B3) genes varies greatly among continental populations which can help in pharmacogenetic studies conducted in mixed populations (Sortica et al. 2012). Such variations in these genes are responsible for the uptake of structurally different drugs and endogenous compounds from blood to the liver (Kalliokoski and Niemi 2009; Hagenbuch and Stieger 2013). A large number of variations have been discovered in SLCO1B1 which affect transport function (Niemi et al. 2011). The influence of SLCO1B1 polymorphisms on transport function leads to its substrate specificity (Tirona et al. 2001). Although clinical statins have been identified that are substrates of organic anion transporting polypeptide 1B1 (OATP1B1) (Furihata et al. 2009; Niemi 2010), many of statins are substrates of other hepatic organic anion transporting polypeptides (OATPs), such as pitavastatin which is a substrate of organic anion transporting polypeptide 1B3 (OATP1B3) (Fujino et al. 2005). Recent studies have identified the altered pharmacokinetics of pravastatin and pitavastatin associated with polymorphisms in SLCO1B1 (Smith et al. 2007).
1.11.1. Cholesterol

Cholesterol is a waxy, fatty substance that is an essential constituent in every cell in the body (Figure 1.1). Cholesterol has important biological functions, and is mainly produced in the liver and CNS. Cholesterol is naturally absorbed through the intestinal tract (Charlton-Menys and Durrington 2008). It is the major sterol component of the plasma membrane and organelle membranes, except for the mitochondrial membranes (Lange et al. 2004). The quantity of cholesterol in cells depends on the type of organelle and compartment. The endoplasmic reticulum (ER), where cholesterol is formed contains 0.1-2% of cholesterol in mole%, depending on cell type (Lange et al. 2004). Cholesterol is very important for many biological processes; including the formation of lipids in cell membranes, as well as being a precursor for the synthesis of vitamin D, steroid hormones and bile acids (Brookes et al. 2009). Changes in cholesterol levels can have deep influences on membrane transport (Charlton-Menys and Durrington 2008). Synthesized cholesterol is carried by two types of lipoproteins that are produced by the liver and intestine. High density lipoproteins (HDL) carry cholesterol from cells to the liver, where it is broken down as a waste product. Conversely, low density lipoproteins (LDL) carry cholesterol from the liver to body cells. Excess cholesterol can accumulate in the walls of arteries and lead to atherosclerosis (Gylling 2004). Most cholesterol is synthesized in the body, while cholesterol gained from food is a minor component (Ikonen 2006).

![Chemical structure of cholesterol](image)

**Figure 1.1 -** Chemical structure of cholesterol
1.11.2. Lipoproteins

Lipids are insoluble in plasma and therefore need to be transported as complexes with proteins, these complexes are called lipoproteins (Lusis and Pajukanta 2008). Lipoproteins are spherical particles (Figure 1.2 A), whose major constituents can include non-polar lipids of cholesterol esters and triglycerides, as well as polar components including free cholesterol, phospholipids and proteins (Shen et al. 1977).

Figure 1.2 A- Structure of lipoproteins

Figure 1.2 B- Relationship of lipoproteins components to diameter and density (Ridker 2014).

Lipoproteins are categorized into classes with different ratios of fat and protein which lead to differences in density and size of the complexes (Figure 1.2 B). The classes are: very low density lipoproteins (VLDL), low density lipoproteins(LDL), intermediate density lipoproteins(IDL) and high density lipoproteins(HDL), as well as chylomicrons (Zhang et al. 2011).

1.11.3. Familial hypercholesterolemia
Familial hypercholesterolemia (FH) is an inherited autosomal (not related to the sex chromosomes) co-dominant genetic disorder. FH is mainly caused (85–90% of the cases) by mutations in the LDL receptor gene (LDLR), while 5–10% of the cases are linked to mutations in apolipoprotein B (APOB). Approximately 1% of the cases are the result of mutations in proprotein convertase subtilisin/kexin type 9 gene (PCSK9) (Davidson et al. 2011; Gouni-Berthold and Berthold 2015). FH is characterized by very high LDL cholesterol, tendon xanthomas and premature cardiovascular disease (Chiou and Charng 2012).

FH has been classified into two forms, heterozygous (heFH) form in which the patient inherits FH gene from one parent (Kwiterovich et al. 1974). HeFH is a relatively common disease with an estimated prevalence of 1/500 in most populations (Durrington 2001; Chiou and Charng 2012). However, the rare homozygous (hoFH) form occurs when the patient inherits the FH gene from both parents (Raal and Santos 2012). Globally the prevalence of hoFH is estimated to be 1 case in 1 million people (Raal and Santos 2012).

HeFH is more frequent in specific populations, which is shown 1/100 among Afrikaners, 1/170 among Christian Lebanese, 1/270 among French Canadians (Izar et al. 2010) and 1/208 in the Hokuriku district of Japan (Mabuchi et al. 2011; Raal and Santos 2012). However, prevalence hoFH in these groups is estimated to be 1/30000, 1/100000, 1/275000 and 1/171167, respectively (Raal and Santos 2012).

1.1.4. Cholesterol lowering drugs

Hydroxy methylglutaryl coenzyme A (HMG-CoA) Reductase Inhibitors (also called statins) are a type of drugs that reduce low-density lipoprotein cholesterol levels (Cho et al. 2015). Structurally statins are analogues of 3-hydroxy-3-methylglutaryl coenzyme A reductase (HMG-CoA) (Figure 1.3) and strongly inhibit hepatic conversion of HMG-CoA reductase, thereby limiting cholesterol biosynthesis (Endres 2005; Pahan 2006).
There are five types of statins available as a medication that approved by US Food and Drug Administration (FDA), including lovastatin, simvastatin, pravastatin, fluvastatin and atorvastatin (Manzoni and Rollini 2002). Current studies show that due to their multiple functions, statins could be used for pharmaceutical treatment of several disorders including cancer, diabetes, inflammation, neurodegeneration and demyelination (Pahan 2006). In spite of the effectiveness of statins in lowering cholesterol levels in the blood, they have some side effects that cannot be disregarded, such as the fact that they cause myopathy, severe liver problems (Brookes et al. 2009), and an increased risk of hemorrhagic strokes (Goldstein et al. 2008). Statins play a role in the development of heart disease and atherosclerosis because cholesterol plays in a main role in synthesising lipids and hormones in the body, therefore statins do not decrease cholesterol production only, but might affect other products synthesis (Brookes et al. 2009).
1.12. Genetic polymorphisms of drug metabolizing enzymes, transporters and targets

It has well known that genetic variability in genes encoding drug metabolizing enzymes, transporters and targets comprises the main causes of inter-individual differences in therapeutic outcome (Evans 2004; Serpe et al. 2014).

The following enzymes, transporters and targets harbouring well determined polymorphisms with clinical effect on FDA anticancer efficacy, and specific transporters with clinical effect on FDA cholesterol lowering drugs efficacy, were chosen for this study.

**Figure 1.3**- Structure of HMG-CoA and different types of statins (Shitara and Sugiyama 2006).
1.12. 1. Genetic variations associated with FDA approved anticancer drugs

1.12.1.1. Catechol O-methyltransferase enzyme:
Catechol O-methyltransferase enzyme (COMT) is encoded by the COMT gene which is an important gene involved in cisplatin cancer drug metabolism and is linked with ototoxicity in children (Fung et al. 2011). Ototoxicity is a serious problem that affects about 60% of patients receiving cisplatin (Ross et al. 2009). The genetic variations rs9332377 and rs4646316 in the COMT gene have been identified and are associated with hearing damage in children who receiving cisplatin (Ross et al. 2009).

1.12.1.2. Dihydropyrimidine dehydrogenase enzyme:
Dihydropyrimidine dehydrogenase enzyme (DPYD) is the first and rate-limiting factor in pyrimidines catabolism (Seck et al. 2005; Largillier et al. 2006; Saif 2013). It has been proposed that DPYD deficiency is associated with toxicities to the cancer drug fluorouracil (Li et al. 2014), and genetic polymorphisms in DPYD gene were responsible for DPYD deficiency (Luo et al. 2015). The SNPs rs2297595 and rs3918290 in the DPYD gene are linked to a decreased efficacy of fluorouracil and also cause severe toxicity (Gentile et al. 2015). Patients with DPYD mutations may have severe drug-adverse effects following fluoropyrimidines (capecitabine and fluorouracil) therapy (Gross et al. 2008). The variant rs3918290 is associated with severe toxicity to capecitabine (Deenen et al. 2011), and the SNP 2297595 was found to be associated with toxicity, particularly in patients with breast and gastroesophageal cancer (Gross et al. 2008).

1.12.1.3. Thiopurine S-methyltransferase enzyme:
Thiopurine S-methyltransferase (TPMT) is an enzyme that catalyses the S-methylation of thiopurines such as azathioprine, thioguanine, and mercaptopurine (Spire-Vayron de la Moureyre et al. 1998; Zeglam et al. 2015). Patients with inherited deficiency in TPMT activity treated with standard doses of thiopurine cancer drugs are at an elevated risk of
thiopurine toxicity (Weinshilboum 2001). The three SNPs in the *TPMT* gene, rs1142345, rs1800460 and rs1800462 have been reported as the most common mutant alleles linked with the loss of TPMT activity in many populations (Murugesan *et al.* 2010) and with the incidence of adverse effects to thiopurines (Gazouli *et al.* 2010).

TPMT also plays a significant role in cisplatin metabolism and correlated with ototoxicity in paediatric patients (Fung *et al.* 2011). The SNPs rs1142345 and rs1800460 are associated with hearing loss in children treated with cisplatin (Ross *et al.* 2009).

1.12.1.4. Transmembrane Protein 43:
Transmembrane Protein 43 (TMEM43) is a 43 kDa putative membrane protein whose structure and function are undetermined (Siragam *et al.* 2014). This transporter is encoded by the *TMEM43* gene (van der Zwaag *et al.* 2009). The SNP rs2228001 in the *TMEM43* gene might decrease the activity of the transporter, and patients with the rare allele who receive cisplatin may be at higher risk of ototoxicity (Caronia *et al.* 2009).

1.12.1.5. Methylene-tetrahydrofolate reductase enzyme:
Methylene-tetrahydrofolatereductase (MTHFR) is an important enzyme that regulates folate metabolism (Shen *et al.* 2001) and is encoded by the *MTHFR* gene (Kałużna *et al.* 2015). Polymorphisms in *MTHFR* and the variant genotypes lead to decreased MTHFR enzyme activity and lower plasma folate level (Shen *et al.* 2001). Studies have shown that the common SNP rs1801133 is associated with a decreased activity of the enzyme MTHFR (Cortese and Motti 2001). In addition, this variant significantly increased the toxicity of methotrexate when used with carboplatin as a treatment for ovarian cancer (Toffoli *et al.* 2003; Kim 2009). Severe acute toxicity could occur in patients with the AA genotype after the first cycle of adjuvant CMF (Cyclophosphamide, Methotrexate, Fluorouracil) chemotherapy (Toffoli *et al.* 2000b).
1.12.1.6. Glutathione S-transferase P enzyme:
For the purpose of cell protection, GSTs enzymes catalyze the conjugation of glutathione to xenobiotics to form glutathione disulfide (Hayes and McLellan 1999; Geng et al. 2014). The genetic variant of GSTP1 gene (rs1695) has been indicated to be relevant to the response of anthracycline-based chemotherapy (e.g. epirubicin) in breast cancer (Tulsyan et al. 2013). The variant G is correlated with increased risk of toxicity in colorectal cancer patients (Braun et al. 2009) and increased risk of neutropenia in lupus erythematosus patients treated with cyclophosphamide (Zhong et al. 2006). However, in women with breast cancer, this SNP was not associated with adverse drug reactions after cyclophosphamide containing chemotherapy (Ekhart et al. 2008; Low et al. 2009).

1.12.2. Genetic polymorphisms associated with FDA approved cholesterol lowering drugs:
1.12.2.1. Solute carrier organic anion trasporter1B1 (SLCO1B1) gene:
It was identified that all types of statins that are used clinically are substrates of the transporter OATP1B1 (gene name SLCO1B1) (Niemi et al. 2011). Numerous variants have been found within the SLCO1B1 gene (Romaine et al. 2010). The common SNPs rs4149056 and rs2306283 are associated with decreased the hepatic transporter OATP1B1 activity (Niemi et al. 2011).

The SNP rs11045819 genotype CC has been shown to exhibit significantly less LDL reduction and higher post-treatment LDL levels (Couvert et al. 2008). It has been identified that the SNPs rs4149015 and rs4149056 were linked with reduced plasma concentrations of pravastatin on paediatric patients with HeFH and cardiac transplant recipients treating with immunosuppressive drugs (Hedman et al. 2006). The SNP rs4149036 is associated with
triglyceride lowering, but only among those homozygous for the rare allele (Thompson et al. 2005).

1.12.2.2. Solute carrier organic anion trasporter1B3 (SLCO1B3) gene

Statins are among the drugs that are substrates of OATP1B3 (gene name SLCO1B3) (Kindla et al. 2011). OATP1B3 is an important hepatic transporter that shares several substrates with OATP1B1, such as rosvastatin and pravastatin (Ho et al. 2006; Seithel et al. 2007; Niemi et al. 2011). It was identified that the SNP rs72559743 reduces protein expression and activity of the protein, contrasted with SNP rs4149117 that increases the activity in vitro (Letschert et al. 2004). The SNP 60140950 has a harmful effect on the function of the protein (Geraldine et al.). It was reported that the SNP rs12299012 reduces the uptake of rosvastatin (Schwarz et al. 2011). To date there is no clinical data available for SNP rs57585902.

1.13. Statistical Analyses

1.13.1 GenAlEx

Genetic Analysis in Excel (GenAlEx) is an in-MS-Excel application used for population genetic analysis. GeneAlEx can be used to calculate allele frequency, the Hard-Weinberg Equilibrium and the Shannon and Fixation indices (Peakall and Smouse 2012).

1.13.2. The Hardy-Weinberg Equilibrium

The Hardy-Weinberg equilibrium (HWE) hypothesises that allele and genotype frequencies in a population remain constant or they are in equilibrium from generation to generation, unless other external influences are presented (Hardy 2003). These influences include: mutation, selection, population size, mating, gene migration and random genetic drift. However, this means static allele frequencies of HW which assumes no mutation, no
migration, no selection, random mating (no inbreeding) and a large population (no genetic
drift) do not exist in nature (Butler 2005). However, the HWE is still an ideal state used to
determine genotype frequencies.

1.14. The aims of the study:

This study aims to develop and evaluate the performance of robust and high throughput
techniques for genotyping ten polymorphisms related to anticancer drugs and ten
polymorphisms related to cholesterol lowering drugs. These assays will be developed using a
SNaPshot minisequencing panel which is considered to be an accurate, less time-consuming
and relatively rapid method for detection many polymorphisms in a single assay, as well as
high resolution melt analysis as an alternative, fast and cheap typing method. In addition,
HRM assays will be evaluated and the results validated by comparing them to those obtained
by SNaPshot assays (the gold standard assay for genotyping). Lastly, both assays will be
validated using direct sequencing.

In addition, the minor allele frequencies of ten previously identified SNPs
pharmacogenetically associated with FDA approved anticancer drugs, and the minor allele
frequencies of ten previously identified SNPs pharmacogenetically associated with FDA
cholesterol lowering drugs will be determined. Allele and genotype frequencies for the
investigated SNPs will be compared to different population groups from around the world.
Chapter Two

Development of a pharmacogenomic profiling panel for anticancer drugs

2.1. Introduction

Cancer in Africa is a serious health issue compounded by communicable diseases (e.g. HIV), changes in life style and age (Parkin et al. 2008). It has been noticed that the differences in cancer drug responses and toxicity is a serious problem amongst patients which might lead to treatment failure or adverse drug reactions (Bosch et al. 2006; Huang and Ratain 2009). However, it has been shown that the genetic polymorphisms in drug metabolizing enzymes, transporters and targets genes may explain the differences of chemotherapeutic outcome. There are several chemotherapeutics that are used for cancer treatment. Despite their effectiveness, many exhibit serious side effects (Kawabe 2004). The importance of studying genetic variations has been introduced in the case of chemotherapeutics such as Azathioprine and mercaptopurine with variant alleles in the TPMT gene has been shown to put patients at risk for toxicity. This has helped physicians develop personalised dose regimes (Wilke et al. 2007). Cisplatin is a widely used and highly effective anticancer agent. Nevertheless, ototoxicity can affect about 60% of patients who are receiving this treatment (Ross et al. 2009). It was reported that the SNPs rs9332377 and rs4646316 in the COMT gene and the SNPs rs1142345 and rs1800460 in the TPMT gene, are linked to ototoxicity in children receiving cisplatin (Ross et al. 2009). Hence, pharmacogenetics studies aim to identify these genetic variations and use them as a baseline to predict the optimal chemotherapy dose for each patient (Lee et al. 2005; Huang and Ratain 2009).

Africa has a wealth of genetic diversity, so it is considered to be a good place to study how environmental factors influence the different genetic variations in relation to cancer treatment
(Mgbakor et al. 2014). South Africa is a multi-ethnic country and susceptibility to different types of cancer vary between ethnic groups (Bezwoda et al. 1997). It is home to different groups that include Khoisan, Xhosa, Zulu, Venda, and Sotho Pedi, the Afrikaners and the Cape Admixed the latter being a uniquely admixed population of immigrant Europeans, Asians and indigenous populations (Hardy et al. 2008; Du Plessis et al. 2015). The complexity in Admix populations in particular creates advantages for studying the clinical, environmental, and genetic ancestry of population differences in health outcomes (Suarez-Kurtz 2008a;Via et al. 2009). Therefore, the objectives of this part of the project was to develop a SNaPshot minisequencing system to investigate the genotypic and allelic profile of ten previously reported single nucleotide polymorphisms relevant to anticancer drugs in 130 South African Cape Admixed individuals living in Cape Town. The SNaPshot methodology is high throughput and fast, as it is based on single base extension with a labelled ddNTP (Hu et al. 2016). The minor allele frequencies (MAF) of the Cape admixed population were compared to several other populations. Ethnic populations that were included in this study were: British in England and Scotland (GBR) which represent European Caucasian, Colombian in Medellin, Colombia (CLM) and Mexican ancestry in Los Angeles, California (MXL) both represent Latino admixed population, Gujarati Indian in Houston, Texas (GIH) represent Asian, African ancestry in South-western US represent African American (ASW), Luhya in Webuya, Kenya (LWK) and Yoruba in Ibadan, Nigeria (YRI) both represent sub-Saharan African populations. The SNPs were selected from various population groups as per Ensembl(http://www.Ensembl.org/Homo_sapiens/Variation/Population?db=core;r=22:19967669-19968669:v=rs9332377;vdb=variation;vf=4986831).
2.2. Materials and Methods

2.2.1. Sample Collection

Biological samples were obtained in the form of buccal swabs from 130 apparently healthy, unrelated volunteers from the Cape Admixed population living in Cape Town, South Africa. Ethnicity of volunteers was determined by self-report. To confirm the volunteers’ mixed ancestry, the questionnaires included a short genealogy of two generations for maternal and paternal family members, as well as place of birth, home language and religious affiliation. Ethical approval for this project was obtained from the Senate Research Committee of the University of the Western Cape with a registration no: 10/9/40.

2.2.2. DNA Extraction

Genomic DNA was extracted using a standard salt lysis method. The cotton tip of each swab was cut off with a clean scalpel surgical blade. The excised pieces of the swab were added to a sterile 1.5ml Eppendorf tube containing 600 µl of salt lysis buffer (0.4 M NaCl, 0.01 M Tris-HCl (pH 8), 2 mM EDTA, 1% SDS) and 3µl of 20mg.ml⁻¹ Proteinase K. The tubes were shaken vigorously and incubated shaking overnight at 60°C at 11 5 rpm in a G24 Environmental Incubator shaker (New Brunswick Scientific Co., Inc., USA). After which the entire volume was transferred to a sterile tube. The lysis solution containing the biological material in the swab was recovered by using spin columns. The end of a 0.5 ml tube was perforated with a sterile surgical needle and placed inside a clean 1.5 ml Eppendorf tube and centrifuged at 14000 rpm for 1 min in an Eppendorf centrifuge 5417 (Eppendorf, Germany). The collected volume was added to the separated lysis material. To precipitate the DNA, 200 µl of 6M NaCl was added and the tubes were shaken vigorously for 15 sec and centrifuged at 5000 rpm for 15min. The supernatant containing the DNA was transferred to another sterile tube. A volume of 600 µl isopropanol was added and the tubes were incubated
at -80°C for 30 min. The DNA was collected by centrifugation at 14000 rpm for 30 min then washed with 100 µl of 70% Ethanol and centrifuged at 14 000 rpm for 10 min to remove excess salts. The pellets were air dried and re-suspended in a final volume of 50 µl of 1x TE buffer (1M Tris and 0.5M EDTA, pH 7). Samples were stored at -20°C. The concentration of DNA was measured by using a Nanodrop ND 1000 UV-Vis Spectrophotometer (Thermo Fisher Scientific, Inc., USA).

2.2.3. SNP Selection

A total of ten SNPs, seven non-synonymous and three intronic, were targeted for this part of this project. SNP selection was based on the available evidence on clinical relevance of genetic variants in anticancer therapy as gathered from the Pharmacogenomics Knowledgebase (http://www.pharmgkb.org), US Food and Drug Administration website (FDA) (http://www.FDA.gov) and Ensembl database (http://www.Ensembl.org). Genetic variants were initially selected based on the FDA's table of pharmacogenomic biomarkers in drug labelling for cancer therapy, and narrowed down to ten SNPs using additional clinical relevance data available in the Pharmacogenomics Knowledgebase (http://www.pharmgkb.org) (Table 2.1).
Table 2.1 - A list of genes and selected SNPs associated with anticancer therapy:

<table>
<thead>
<tr>
<th>SNP ID</th>
<th>Gene</th>
<th>Type of gene</th>
<th>Drug</th>
<th>Clinical observation</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs9332377</td>
<td>COMT</td>
<td>Drug metabolizing enzyme</td>
<td>Cisplatin</td>
<td>Ototoxicity</td>
</tr>
<tr>
<td>rs3918290</td>
<td>DPYD</td>
<td>Drug metabolizing enzyme</td>
<td>Capecitabine, Fluorouracil and Tegafur</td>
<td>Drug toxicity</td>
</tr>
<tr>
<td>rs2297595</td>
<td>DPYD</td>
<td>Drug metabolizing enzyme</td>
<td>Capecitabine and Fluorouracil</td>
<td>Drug toxicity</td>
</tr>
<tr>
<td>rs2228001</td>
<td>TMEM43</td>
<td>Drug transporter</td>
<td>Cisplatin</td>
<td>Ototoxicity and toxicity</td>
</tr>
<tr>
<td>rs1142345</td>
<td>TPMT</td>
<td>Drug metabolizing enzyme</td>
<td>Cisplatin Thioguanine</td>
<td>Ototoxicity</td>
</tr>
<tr>
<td>rs1800460</td>
<td>TPMT</td>
<td>Drug metabolizing enzyme</td>
<td>Cisplatin Thioguanine</td>
<td>Drug toxicity</td>
</tr>
<tr>
<td>rs4646316</td>
<td>COMT</td>
<td>Drug metabolizing enzyme</td>
<td>Cisplatin</td>
<td>Ototoxicity</td>
</tr>
<tr>
<td>rs1801133</td>
<td>MTHFR</td>
<td>Drug target</td>
<td>Cyclophosphamide, Methotrexate and Carboplatin</td>
<td>Drug toxicity</td>
</tr>
<tr>
<td>rs1800462</td>
<td>TPMT</td>
<td>Drug metabolizing enzyme</td>
<td>Azathioprine , Thioguanine and purine analogous</td>
<td>Drug toxicity</td>
</tr>
<tr>
<td>rs1695</td>
<td>GSTP</td>
<td>Drug metabolizing enzyme</td>
<td>Cyclophosphamide, Epirubicin and Platinum compounds</td>
<td>Drug toxicity</td>
</tr>
</tbody>
</table>

2.2.4. Primer Design

Primers were designed to genotype the selected SNPs in one multiplex PCR. The primers, which are listed in Table 2.2, were designed to have an annealing temperature between 55°C and 60°C using Primer3plus software (http://primer3plus.com/cgi-bin/dev/primer3plus.cgi) and Ensembl database (http://www.Ensembl.org). Primers were aligned on Ensembl using BLAST (http://www.ensembl.org/Multi/Tools/Blast) to test for nonspecific amplification. Multiplex SNaPshot minisequencing primers listed in Table 2.3, in which poly (gcat) tails were added at the 5’ end of the specific hybridizing part of the primer to generate a range of different sized primers. The final sizes ranged between 30 to 75 bp. Both PCR and SNaPshot minisequencing primers were synthesised by Inqaba Biotec, South Africa.
Table 2.2 - Multiplex PCR primers for the selected SNPs related to anticancer drugs used in SNaPshot™ genotyping.

<table>
<thead>
<tr>
<th>Gene</th>
<th>NCBI (dbSNP)</th>
<th>Forward primer (5’ to 3’)</th>
<th>Reverse Primer (5’ to 3’)</th>
<th>Amplicon length (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>COMT</td>
<td>rs9332377</td>
<td>GTCTCCAGGGACCACCATACCA</td>
<td>GCTGGGTGAGTGGAACACAT</td>
<td>107</td>
</tr>
<tr>
<td>DPYD</td>
<td>rs3918290</td>
<td>TGGACAAAGCTCTTTTCTGA</td>
<td>TCACCAACTTATGCAATTCTC</td>
<td>149</td>
</tr>
<tr>
<td>DPYD</td>
<td>rs2297595</td>
<td>CCAGCAACAAAAAGAGCAAT</td>
<td>AAAACAAGAATTCTTTGAAACAT</td>
<td>194</td>
</tr>
<tr>
<td>TMEM43</td>
<td>rs2228001</td>
<td>GCCTCAAAACCGAGAAGATG</td>
<td>CTGCCCTAGTTTGCTTCTC</td>
<td>237</td>
</tr>
<tr>
<td>TPMT</td>
<td>rs1142345</td>
<td>GGGGAATTGACTGCTTTTTGA</td>
<td>TTGCAATCTGCAAGACACAT</td>
<td>316</td>
</tr>
<tr>
<td>TPMT</td>
<td>rs1800460</td>
<td>CCCTGATACCTGAGCCAGAG</td>
<td>TTACCATTGCGATCACCTG</td>
<td>375</td>
</tr>
<tr>
<td>COMT</td>
<td>rs4646316</td>
<td>ACGGTTTCTTTGAGGAGTGA</td>
<td>TCTGCAAGAGACACATGCTT</td>
<td>431</td>
</tr>
<tr>
<td>MTHFR</td>
<td>rs1801133</td>
<td>AGGACAGTTGGGAGTTGGG</td>
<td>CTCACCTGGAGGAAAGAT</td>
<td>470</td>
</tr>
<tr>
<td>TPMT</td>
<td>rs1800462</td>
<td>TCTGCTTTCTTGCAATTGCTT</td>
<td>CAGGAATTTGCGTATTGTTG</td>
<td>265</td>
</tr>
<tr>
<td>GSTP</td>
<td>rs1695</td>
<td>GCTGGGGCTCACAGAAGCAG</td>
<td>GTGCAAGTTTGCTTGTCC</td>
<td>381</td>
</tr>
</tbody>
</table>
2.2.5. PCR Amplification

PCR amplification was performed using the KAPA2G Fast Multiplex PCR mix (KAPA BIOSYSTEMS, South Africa) in a final volume of 25 µl according to manufacturer’s instructions. Each reaction contained a minimum of 20 ng genomic DNA. PCR amplifications were performed in a GeneAmp 2700/ 2720 (Applied Biosystems, USA) thermal cycler. PCR cycling conditions were carried out as follows: initial denaturation at 95°C for 3 min, denaturation at 95°C for 15 sec, annealing at 60°C for 15 sec, elongation at 72°C for 15 sec, for a total of 30 cycles, followed by final extension at 72°C for 1 min and a hold step at 4°C. PCR products were confirmed by agarose gel electrophoresis on 2% gels and visualised under UV light.

2.2.6. Post PCR Purification

The PCR products were purified to remove excess primers and unincorporated ddNTPs using a Thermosensitive Alkaline Phosphatase (FastAP) (Thermoscientific, U.S.A) and Exonuclease I (Exo1) (Thermoscientific, U.S.A). PCR products were incubated with 0.5 µl of Exo1 and 1 µl of FastAP at 37°C for 15 min followed by thermal deactivation of the enzymes at 75°C for 15 min in a GeneAmp 2700/ 2720 (Applied Biosystems, USA) thermal cycler.

2.2.7. SNaPshot minisequencing (Single Base Extension)

The reaction was performed using the SNaPshot® kit according to the manufacturer’s specifications (Applied Biosystems, U.S.A). The final 10 µl reaction volume contained 3 µl of the purified PCR products, 0.2 µM of primers (final concentration), 2 µl of SNaPshot® ready mix and 3µl water. Negative and positive controls for the mini-sequencing reactions were prepared according to manufacturer’s instructions. A GeneAmp 2700 thermal cycler (Applied Biosystems, USA) was used for the minisequencing reactions. Sequencing cycling
conditions consisted of 25 cycles of denaturation at 96°C for 10 sec, annealing at 50°C for 5 sec, and extension at 60°C for 30 sec, followed by a holding step at 4°C.
Table 2.3- SNaPshot minisequencing reaction primers of the selected SNPs related to anticancer drugs.

<table>
<thead>
<tr>
<th>NCBI (dbSNP)</th>
<th>Amino Acid Change</th>
<th>Nucleotide Change</th>
<th>Minisequencing Reaction Primers</th>
<th>Primer Direction</th>
<th>Position Accession number (NC_000006.12)</th>
<th>Size bp</th>
<th>Poly GCAT tail</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs9332377</td>
<td>Intron</td>
<td>C&gt;T</td>
<td>AGTATCCGGACTCAAGGACCGTGACCACA</td>
<td>Reverse</td>
<td>19968169</td>
<td>30</td>
<td>0</td>
</tr>
<tr>
<td>rs3918290</td>
<td>Intron</td>
<td>C&gt;T</td>
<td>ACTGAACCTAAGGCTGACTTTCCAGACAAC</td>
<td>Reverse</td>
<td>97450058</td>
<td>35</td>
<td>5</td>
</tr>
<tr>
<td>rs2297595</td>
<td>M166V</td>
<td>T&gt;C</td>
<td>ACGAAACTTATGGATGCCCTTTAGTGCC</td>
<td>Reverse</td>
<td>97699535</td>
<td>40</td>
<td>10</td>
</tr>
<tr>
<td>rs2228001</td>
<td>G939L</td>
<td>G&gt;T</td>
<td>GCAGCGCTTCCCACCTGTTCCATTGGAG</td>
<td>Reverse</td>
<td>14145949</td>
<td>45</td>
<td>15</td>
</tr>
<tr>
<td>rs1142345</td>
<td>T240C</td>
<td>T&gt;C</td>
<td>AGATGAATGTCTTTTCATTCTCTGTATG</td>
<td>Forward</td>
<td>18130687</td>
<td>50</td>
<td>20</td>
</tr>
<tr>
<td>rs1800460</td>
<td>A154T</td>
<td>C&gt;T</td>
<td>GGCAAAATTTGACATTTGGTGGATAGAGGA</td>
<td>Reverse</td>
<td>18138997</td>
<td>55</td>
<td>25</td>
</tr>
<tr>
<td>rs4646316</td>
<td>Intron</td>
<td>C&gt;T</td>
<td>CACACCCAGACGACAGCAGCGAGAAA</td>
<td>Forward</td>
<td>19964609</td>
<td>60</td>
<td>30</td>
</tr>
<tr>
<td>rs1801133</td>
<td>A222V</td>
<td>G&gt;A</td>
<td>AGCACTTGAAGGAGGGTGCTTGCCGAGAG</td>
<td>Reverse</td>
<td>11796321</td>
<td>65</td>
<td>35</td>
</tr>
<tr>
<td>rs1800462</td>
<td>A80P</td>
<td>C&gt;G</td>
<td>AAGTGTAATGTATGATTTATGCCAGTTT</td>
<td>Reverse</td>
<td>18143724</td>
<td>70</td>
<td>40</td>
</tr>
<tr>
<td>rs1695</td>
<td>L105V</td>
<td>A&gt;G</td>
<td>GACGGCGTGAGGACCTCCTCGCAATAC</td>
<td>Forward</td>
<td>67585218</td>
<td>75</td>
<td>45</td>
</tr>
</tbody>
</table>
2.2.8. Post Extension Purification

Post Extension products were purified by adding 1 U of FastAP to the 10 µl reaction volume and incubated at 37°C for 30 minutes followed by 15 minutes at 75°C to inactivate the enzyme in a GeneAmp 2700 thermal cycler (Applied Biosystems, USA).

2.2.9. Electrophoresis of the minisequencing products

Reactions were set up in a 96-well plate in POP4 polymer (Applied Biosystems, USA). To each well, 8.7 µl HiDi Formamide (Applied Biosystems, USA) and 0.3 µl GeneScan Liz 120 size standard (Applied Biosystems, USA) were mixed with 1µl of the purified minisequencing product. The plate was briefly centrifuged for 1 min at 1000 rpm and incubated in a thermal cycler for DNA denaturation at 95°C for 5 min. Subsequently, the fluorescently labelled fragments were separated on 36 cm-long capillary containing POP4 polymer on ABI Prism 3500 DNA sequencer (Applied Biosystems, USA). Data was analysed using GeneMapper® IDX Software Version 1.4 (Applied Biosystems, USA).

2.2.10. Statistical Analysis

The statistical analysis of Genotype and allele frequencies, as well as the deviation from the Hardy-Weinberg Equilibrium were calculated using GenAlEx 6.5 software (Peakall and Smouse 2012). Allele and genotype frequencies are given with binomial proportion 95% confidence intervals (CI). The Chi-square test was used to determine if individual variants were in Hardy-Weinberg Equilibrium (HWE) at each locus.

2.3. Results:

The sample group enrolled in this study consisted of 130 healthy individuals from the Cape Admixed Population; 12 (9%) females and 118 (91%) males between the ages of 14 and 68 years. Minor optimization was needed for the primer sets of the two SNPs (rs1801133 and
Due to poor amplification, the concentration of these two primers pairs was increased to 0.4 µM in the PCR amplification. Also minor optimization was required for these two SNPs in SNAPSHOT minisequencing reaction, in which the concentration was also increased to 0.4 µM. This resulted in good profiles for each. In certain cases the fragments migrated further than its indicated size which is likely due to minisequencing chemistry (Quintáns et al. 2004). Genotypes were confirmed with direct sequencing as described in Chapter 4. Electropherogram profile of the SNAPSHOT minisequencing genotyping system is displayed in Figure 2.1.

**Figure 2.1**- Electropherogram profile of the SNPs related to anticancer drugs. Green peaks indicate an A nucleotide, blue peaks indicate a G nucleotide, black peaks indicate a C nucleotide and red peaks indicate a T nucleotide.

The allelic frequency of each SNP was determined to be in HWE (P>0.05), with the exception of SNPs rs1801133 and rs3918290. The expected and observed genotype
frequencies in the study are determined under the Hardy-Weinberg Equilibrium. The genotype and allele frequencies of the ten selected SNPs relevant to anticancer drugs observed in Cape Admixed Population are summarized in Table 2.4. The minor allele frequency for the ten variants was compared globally to seven ethnic groups, which were: BGR (British in England and Scotland), CLM (Colombian in Medellin, Colombia), MXL (Mexican ancestry in Los Angeles in California), GIH (Gujarati Indian in Houston), ASW (African ancestry in Southwest US), LWK (Luhya in Webuye, Kenya) and YRI (Yoruba in Ibadan, Nigeria). The comparison was summarized in Table 2.5 and depicted in Figure 2.2. In our study, rs3918290 was the only investigated SNP found to be monomorphomic. The heterozygosity was observed in two intronic and seven nonsynonomous SNPs. The intronic variant rs9332377 genotype frequencies for wild-type (CC), heterozygote (CT) and homozygote (TT) were 74.6%, 24.6% and 0.8%, respectively. Furthermore, the intronic variant rs4646316 genotype frequencies showed homozygote 57.7% for the wild-type (CC), 38.5% for heterozygote (CT) and 3.8% for homozygote (TT). The M166V (rs2297595) genotype frequencies for wild-type (TT), heterozygote (TC), and homozygote (CC) were 91.5%, 8.5% and 0.0%, respectively. The G939L (rs2228001) genotype frequencies for wild-type (TT), heterozygote (TG), and homozygote (GG) were 46.9%, 47.7% and 5.4%, respectively. T240C (rs1142345) genotype frequencies for wild-type (TT), heterozygote (TC), and homozygote (CC) were 94.6%, 5.4% and 0.0%, respectively. The A154T (rs1800460) genotype frequencies for wild-type (CC), heterozygote (CT), and homozygote (TT) were 97.7%, 2.3% and 0.0%, respectively. The A222V (rs1801133) genotype frequencies for wild-type (GG), heterozygote (GA), and homozygote (AA) were 48.5%, 13% and 38.5%, respectively. A80P (rs1800462) genotype frequencies for wild-type (CC), heterozygote (CG), and homozygote (GG) were 99.2%, 0.8% and 0.0%, respectively. L105V
(rs1695) genotype frequencies for wild-type (GG), heterozygote (GA), and homozygote (AA) were 12.3%, 50% and 37.7%, respectively.
### Table 2.4 - Genotype and allele frequencies of selected SNPs related to anticancer drugs in 130 healthy Cape Admixed individuals.

<table>
<thead>
<tr>
<th>Amino Acid Substitution</th>
<th>Gene</th>
<th>dbsNP ID</th>
<th>Observed Genotype Frequency</th>
<th>Allele Frequency</th>
<th>HWE (P)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Genotype %</td>
<td>95% CI</td>
<td>Allele %</td>
</tr>
<tr>
<td>Intron</td>
<td>COMT</td>
<td>rs9332377</td>
<td>CC 74.6</td>
<td>67.12 – 82.08</td>
<td>C 86.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CT 24.6</td>
<td>17.2 – 32</td>
<td>T 13.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TT 0.8</td>
<td>0.0 – 2.33</td>
<td></td>
</tr>
<tr>
<td>Intron</td>
<td>DPYD</td>
<td>rs3918290</td>
<td>CC 100.0</td>
<td>100-100</td>
<td>C 100.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CT 0.0</td>
<td>0.0</td>
<td>T 0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TT 0.0</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>M166V</td>
<td>DPYD</td>
<td>rs2297595</td>
<td>TT 91.5</td>
<td>86.71-96.29</td>
<td>T 95.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TC 8.5</td>
<td>3.71-13.29</td>
<td>C 4.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CC 0.0</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>G939L</td>
<td>TMEM43</td>
<td>rs2228001</td>
<td>TT 46.9</td>
<td>38.32-55.48</td>
<td>T 70.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TG 47.7</td>
<td>39.11-56.29</td>
<td>G 29.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GG 5.4</td>
<td>1.51-9.29</td>
<td></td>
</tr>
<tr>
<td>T240C</td>
<td>TPMT</td>
<td>rs1142345</td>
<td>TT 94.6</td>
<td>90.71-98.49</td>
<td>T 97.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TC 5.4</td>
<td>1.51-9.29</td>
<td>C 2.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CC 0.0</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>A154T</td>
<td>TPMT</td>
<td>rs1800460</td>
<td>CC 97.7</td>
<td>95.12-100</td>
<td>C 98.8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CT 2.3</td>
<td>0.0-4.88</td>
<td>T 1.2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TT 0.0</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>Intron</td>
<td>COMT</td>
<td>rs4646316</td>
<td>CC 57.7</td>
<td>49.21-66.19</td>
<td>C 76.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CT 38.5</td>
<td>30.14-46.86</td>
<td>T 23.1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TT 3.8</td>
<td>0.51-7.09</td>
<td></td>
</tr>
<tr>
<td>A222V</td>
<td>MTHFR</td>
<td>rs1801133</td>
<td>GG 48.5</td>
<td>39.91-57.09</td>
<td>G 55</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GA 13</td>
<td>7.22-18.78</td>
<td>A 45</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>AA 38.5</td>
<td>30.14-46.86</td>
<td></td>
</tr>
<tr>
<td>A80P</td>
<td>TPMT</td>
<td>rs1800462</td>
<td>CC 99.2</td>
<td>97.67-100</td>
<td>C 99.6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CG 0.8</td>
<td>0.0-2.33</td>
<td>G 0.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GG 0.0</td>
<td>0.0</td>
<td></td>
</tr>
<tr>
<td>L105V</td>
<td>GSTP</td>
<td>rs1695</td>
<td>GG 12.3</td>
<td>6.65-17.95</td>
<td>G 37.3</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GA 50</td>
<td>41.41-58.59</td>
<td>A 62.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>AA 37.7</td>
<td>29.37-46.03</td>
<td></td>
</tr>
</tbody>
</table>

P-Value will be NA for monomorphic SNPs (MAF =0).
Table 2.5 - Comparison of MAF of the selected SNPs related to anticancer drugs in the Cape Admixed population to other ethnic groups.

<table>
<thead>
<tr>
<th>rsSNP ID</th>
<th>Amino acid change</th>
<th>Minor Allele</th>
<th>Cape Admixed</th>
<th>British</th>
<th>Colombian</th>
<th>Mexican</th>
<th>Indian</th>
<th>African American</th>
<th>Luhya</th>
<th>Yuroba</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs9332377</td>
<td>Intron</td>
<td>T</td>
<td>13.1</td>
<td>15.9</td>
<td>14.9</td>
<td>8.6</td>
<td>21.4</td>
<td>33.6</td>
<td>33.3</td>
<td>35.2</td>
</tr>
<tr>
<td>rs3918290</td>
<td>Intron</td>
<td>T</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>1.5</td>
<td>0.8</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>rs2297595</td>
<td>M166V</td>
<td>C</td>
<td>4.2</td>
<td>8.2</td>
<td>6.4</td>
<td>7.0</td>
<td>3.4</td>
<td>4.9</td>
<td>10.1</td>
<td>0.9</td>
</tr>
<tr>
<td>rs2228001</td>
<td>G939L</td>
<td>G</td>
<td>29.2</td>
<td>38.5</td>
<td>29.3</td>
<td>25.8</td>
<td>34</td>
<td>33.6</td>
<td>27.3</td>
<td>26.9</td>
</tr>
<tr>
<td>rs1142345</td>
<td>T240C</td>
<td>C</td>
<td>2.7</td>
<td>3.3</td>
<td>2.1</td>
<td>4.7</td>
<td>2.4</td>
<td>9.8</td>
<td>11.6</td>
<td>6.0</td>
</tr>
<tr>
<td>rs1800460</td>
<td>A154T</td>
<td>T</td>
<td>1.2</td>
<td>2.7</td>
<td>1.1</td>
<td>3.9</td>
<td>0.0</td>
<td>2.5</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>rs4646316</td>
<td>Intron</td>
<td>T</td>
<td>23.1</td>
<td>20.3</td>
<td>21.8</td>
<td>16.4</td>
<td>9.2</td>
<td>23.8</td>
<td>16.7</td>
<td>16.2</td>
</tr>
<tr>
<td>rs1801133</td>
<td>A222V</td>
<td>A</td>
<td>45.0</td>
<td>32.4</td>
<td>54.3</td>
<td>46.9</td>
<td>15</td>
<td>13.9</td>
<td>7.1</td>
<td>10.6</td>
</tr>
<tr>
<td>rs1800462</td>
<td>A80P</td>
<td>G</td>
<td>0.4</td>
<td>1.1</td>
<td>1.1</td>
<td>0.0</td>
<td>0.0</td>
<td>0.8</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>rs1695</td>
<td>L105V</td>
<td>A</td>
<td>62.7</td>
<td>68.1</td>
<td>64.4</td>
<td>43.8</td>
<td>68.9</td>
<td>54.1</td>
<td>49.0</td>
<td>60.2</td>
</tr>
</tbody>
</table>

a. This study, b. Data from Ensembl database (http://www.Ensembl.org).
Figure 2.2 - Allele frequencies of the selected SNPs related to anticancer drugs in the Cape Admixed population compared to other ethnic groups.

Cape Admixed (This study); British, Colombian, Mexican, Indian, African American, Luhya and Yuroba (Ensembl database).
2.4. Discussion:

The SNaPshot minisequencing assay is one of the most efficient, rapid and accurate ways for high throughput genotyping that requires little optimization steps (Quintáns et al. 2004). In this part of the project, SNaPshot minisequencing method was successfully used and optimized to investigate 10 SNPs relevant to anticancer drugs within 130 unrelated healthy individuals of the Cape Admixed population in South Africa. Cancer is increasing globally, especially in less developed countries (Torre et al. 2015). It is estimated that approximately 5-10% of all cancers are caused by inherited mutations (Mgbakor et al. 2014). It is well known that chemotherapeutics effectiveness is associated with ethnic differences which have profound effects on the response and toxicity to these agents (Chen et al. 2010). Differences in patients drug metabolizing enzymes, transporters and targets genes are considered to be the major source on the variability in treatment outcome (Evans 2004; Serpe et al. 2014).

The Cape Admixed population has high genetic diversity compared to the indigenous populations. This unique population has contributed to inter-individual genetic variability, which may account for the observed physiological differences, especially with respect to drug uptake and metabolism. From the results presented it can be seen that the Cape Admixed population shares many genetic variations with other ethnic groups.

The intronic variants rs9332377 and rs4646316 in the COMT gene which encodes for the enzyme catechol-O-methyltransferase, have previously been shown to be linked to an increased risk of hearing damage in a study on 162 paediatric cancer patients receiving cisplatin (Ross et al. 2009). However, a recent study by Yang et al. (2013) on 213 children with medulloblastoma and 41 with solid tumors showed no relation
between hearing damage and these variants. In our population, the MAF observed for the intronic SNPs in *COMT* gene was 13.1% for rs9332377, which was a bit lower than Colombian (14.9%) and higher than Mexican populations (8.6%), and 23.1% for rs4646316 which was slightly higher than Colombian (21.8%), Indian (21.4%) and British (20.3%) populations.

The intronic variant rs3918290 and M166V (rs2297595), were detected in the *DPYD* gene which encodes the enzyme dihydropyrimidine dehydrogenase. These variants were reported by Gentile *et al.* (2015) to be responsible for decreasing efficacy of fluorouracil and cause acute toxicity. However, the incidence of severe toxicity was observed in SNP rs3918290 in patients who underwent treatment with fluoropyrimidines, such as fluorouracil and capecetabine, but no clear association with the M166V (rs2297595) variant and severe toxicity was found (Toffoli *et al.* 2015). Deenen *et al.* (2011) enrolled advanced colorectal cancer Dutch patients in a study which described that the DPYD variants rs3918290 and rs2297595 lead to severe toxicity with capecitabine-based chemotherapy. Furthermore, rs3918290 was also confirmed as a cause of severe toxicity to tegafur as fluorouracil prodrugs (Terrazzino *et al.* 2013). It was suggested by some studies that patient’s gender could affect the toxicity of fluorouracil. Many studies have showed women are more prone to fluorouracil toxicity than men (Schwab *et al.* 2008). However, it was described by another study that the correlation of the variant rs3918290 with fluorouracil toxicity was significantly higher in men than in women (Schwab *et al.* 2008; Caudle *et al.* 2013). The intronic *DPYD* gene SNP rs3918290 was not observed in the Cape admixed, British, Hispanic (Mexican and Colombian) and sub-Saharan populations. However, it was observed in African Americans with a MAF of 0.8% and Indians with a MAF of 1.5%. The substitution M166V (rs2297595) was observed with a MAF
of 4.2% which was a bit lower than the observed frequency for African Americans (MAF of 4.9%) and slightly higher than Indians that reported a MAF of 3.4%.

In the **TMEM43** gene which encodes for the transmembrane protein 43 (van der Zwaag *et al.* 2009), the SNP rs2228001 with the amino acid substitution G939L was previously shown to increase the risk of cisplatin-induced hearing loss (Caronia *et al.* 2009). Caronia *et al.* investigated 91 osteosarcoma patients receiving cisplatin, and found that only 32 patients had ototoxicity in which patients with GG genotype were more prone to have ototoxicity than patients with the GA and AA genotypes. Moreover, this variant has been predicted to cause severe toxicity in a study consisting of 101 patients with bladder cancer treated with platinum-based chemotherapy (Sakano *et al.* 2010). In our population, this variant showed a MAF with 29.2% that was similar to Colombian population (29.3%). However, the MAF was higher than those of Mexican and sub-Saharan African populations (Table 2.5).

The variants T240C (rs1142345) and A154T (rs1800460) in the **TPMT** gene, which encodes for thiopurine methyltransferase, were found to increase the risk of hearing loss in child cancer patients treated with cisplatin (Ross *et al.* 2009). Conversely, a study by Yang *et al.* observed no relationship between these variants and hearing loss in children taking cisplatin. Moreover, he found no functional differences in TPMT knockout against wild-type mice after treatment with cisplatin (Yang *et al.* 2013). It has been identified that patients with inherited low levels of TPMT enzyme activity are at a higher risk for toxicity when treated with thiopurines, such as thioguanine, azathioprine, and mercaptopurine (Murugesan *et al.* 2010). Moreover, it was determined that the **TPMT** polymorphisms (T240C, A154T and A80P), were the most common mutant alleles correlated with loss of TPMT enzyme activity in many populations (Murugesan *et al.* 2010). Furthermore, in previous studies on children...
with inflammatory bowel disease (IBD) treated with thiopurines, these polymorphisms did not show a significant correlation with adverse effects to these drugs (Stocco et al. 2005; Ridder et al. 2006; Gazouli et al. 2010). The MAF for the variant T240C was 2.7% in our population, which was quite similar to Indians (2.4%). However, it is slightly higher than Colombians (2.1%) and lower than the British population (3.3%). While, the MAF for the variant A154T was 1.2%, which is similar to Colombian (1.1%) and lower than British, Mexican and African Americans (Table 2.5). Further, the MAF for the variant A80P was 0.4% which was lower than British (1.1%), Colombian (1.1%) and African American (0.8%). However, it was significantly high compared to Mexican, Indian and sub-Saharan populations which all had a MAF of 0.0%.

In the MTHFR gene encodes for methylene-tetrahydrofolate reductase (Kałużna et al. 2015), the substitution A222V (ra1801133) has been shown to reduce enzyme activity and affect chemosensitivity of cancer cells (Martin et al. 2006). Ulrich and co-workers investigated 220 chronic myelogenous leukemia patients, and they showed that in patients with the AA genotype, MTHFR activity appeared to be decreased and they are at a higher risk of methotrexate toxicity (Ulrich et al. 2001). In another study on Caucasian ovarian cancer patients treated with methotrexate only or combined with carboplatin, showed that AA genotype was associated with methotrexate-induced hyperhomocysteinaemia and patients were at a higher risk of treatment related toxicity (Toffoli et al. 2003). Moreover, this SNP has also been observed to reduce chemosensitivity to methotrexate in breast cancer cell lines (Sohn et al. 2004). Toffoli and co-workers also found a higher incidence of acute toxicity in Caucasian breast cancer patients carrying the AA genotype during CMF (cyclophosphamide, methotrexate, and fluorouracil) treatment (Toffoli et al. 2000a). Furthermore, in a
study consisting of 248 women with breast cancer, African American and Caucasian women with the GA and AA genotypes for A222V (rs1801133) were shown to have an increased cancer survival rate (Martin et al. 2006). However, in another study, no relationship between this variation and disease-free survival was found in breast cancer patients treated with FEC (cyclophosphamide, epirubicin, and fluorouracil) or CMF (Pare et al. 2007). The variant A222V (rs1801133) was detected in our studied population with a MAF of 45% which was lower than the Mexican population (46.9%), however, it was clearly higher than African Americans, Indians, Luhya and Yoruba that showed a MAF of 13.9%, 15%, 7.1% and 10.6%, respectively.

The SNP rs1695 in the GSTP gene, that encodes for glutathione S-transferase, results in the amino acid change L105V, was observed in 94 women with breast cancer receiving anthracycline based CEF (cyclophosphamide, epirubicin, and fluorouracil) treatment. The study found that GG genotypes were associated with homological toxicity (Zárate et al. 2007). However, a study on Chinese patients with breast cancer, has shown that patients with the G allele have increased survival following cyclophosphamide-based chemotherapy than those with the A allele (Ge et al. 2013). In another study, colorectal cancer patients were found to be at an increased risk of toxicity when treated with cyclophosphamide (Braun et al. 2009). Moreover, patients with AA genotype had increased oxaliplatin-induced neuropathy (Lecomte et al. 2006). The MAF for L105V variant was 62.7% in our population which was lower than the Colombians (64.4%), Indians (68.9%) and British population (68.1), but higher than Mexican, African American and sub-Saharan populations (Table2.5).
2.5. Conclusion

Pharmacogenetics gives a unique way for studying and investigating the variability in response of chemotherapeutic drugs which have a narrow curative effect and show inter-individual differences. However, understanding the impact of genetic variations in drug metabolizing enzymes, transporters and targets could provide a solution to precision cancer treatment based on a patient’s genetic profile, ensuring they receive the drug with the highest efficacy and lowest toxicity. This study investigated the allele and genotype frequencies of ten SNPs in six pharmacogeneticaly cancer relevant genes in the South African Cape Admixed population. SNaPshot method was used successfully for high throughput genotyping. The genotypic and allelic frequencies obtained were compared globally to different ethnic groups. Ultimately, the genotype and allele frequency data of this study could assist in providing effective cancer treatments in the correct dosages for the South African Cape Admixed population.
Chapter Three

Development of a Pharmacogenomic profiling panel of cholesterol lowering drugs

3.1. Introduction

Many transporters play an essential role in the membrane transport of clinical drugs, as they are involved in absorption, distribution and excretion of drugs (Tamai 2012). In addition, the ability of transporters to change the tissue concentrations of drugs such as statins, affects their pharmacological and/or toxicological properties (Giacomini et al. 2010; Shitara et al. 2013). However, it has been recognized that genetic variations play a role in determining a patient’s drug response and susceptibility to toxicity to statins (Needham and Mastaglia 2014).

The genetic structure in the organic anion-transporting polypeptide 1B1 (OATP1B1) and the organic anion-transporting polypeptide 1B3 (OATP1B3) varies from one population to another around the world which creates a great chance for pharmacogenetic studies (Sortica et al. 2012). The solute carrier organic anion transporter family member 1B1 (SLCO1B1) gene and the solute carrier organic anion transporter family member 1B3 (SLCO1B3) gene that code for OATP1B1 and OATP1B3, respectively are thought to mediate the influx and metabolism of statins in the liver (Schwarz et al. 2011). Some studies have previously shown connections between genetic variations in SLCO genes and the pharmacokinetics of substrate uptake. For example, the common variation V174A (rs4149056) in the SLCO1B1 reduces the activity of the OATP1B1 transporter which leads to increase plasma concentrations of several statins (Niemi et al. 2011).

In this part of this study, we developed a robust genotyping system based on the SNaPshot minisequencing reaction to investigate the genotypic and allelic distributions of five SNPs in
the solute carrier organic anion transporter 1B1 (SLCO1B1) gene and five SNPs in the solute carrier organic anion transporter 1B3 (SLCO1B3) gene related to cholesterol lowering drugs. In addition, we generated important pharmacogenomic data for FDA approved cholesterol lowering drugs. The data was compared to different ethnic groups worldwide. The haplotype structure of these genes was designated and possible implications of the genetic variations on cholesterol lowering drugs in the South African Cape Admixed population are discussed.

3.2. Materials and Methods

3.2.1. Sample Collection

Samples collection was as described in Chapter Two.

3.2.2. DNA Extraction

Genomic DNA was extracted from buccal swab samples using a standard salt-lysis method and stored at -20°C as described in Chapter Two.

3.2.3. SNP Selection

Ten single nucleotide polymorphisms (SNPs) were investigated in this part of the study. SNPs were selected based on the available evidence on clinical relevant genetic variants in cholesterol lowering therapy as gathered from the Pharmacogenomics Knowledgebase (http://www.pharmgkb.org), US Food and Drug Administration website (FDA) (http://www.FDA.gov), the UCSF-PMT (http://www.pharmacogenetics.ucsf.edu/) database and Ensembl database (http://www.Ensembl.org). Genetic variants were initially selected based on the FDA's table of pharmacogenomic biomarkers in drug labelling for cholesterol lowering drugs, and narrowed down to ten SNPs using additional clinical relevance available in the Pharmacogenomics Knowledgebase (http://www.pharmgkb.org).
3.2.4. Primer Design

All primers were designed as described in Chapter two. PCR primers were designed to have an annealing temperature between 52°C and 55°C. Table 3.1 shows a list of PCR primers and Table 3.2 shows minisequencing primers.
**Table 3.1** - Multiplex PCR primers for the selected SNPs related to Cholesterol lowering drugs used in SNaPshot™ genotyping.

<table>
<thead>
<tr>
<th>Gene</th>
<th>NCBI (dbSNP)</th>
<th>Forward primer (5’ to 3’)</th>
<th>Reverse Primer (5’ to 3’)</th>
<th>Amplicon length (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SLCO1B1</td>
<td>rs4149056</td>
<td>CATGGCTTTGCTCTTCTCTTCC</td>
<td>TGTAAGAAAGGCCCAATGTT</td>
<td>404</td>
</tr>
<tr>
<td>SLCO1B1</td>
<td>rs2306283</td>
<td>GGGGAAGATAATGGTGCAAAA</td>
<td>GCTGCTGTGTGTTCTCAAA</td>
<td>561</td>
</tr>
<tr>
<td>SLCO1B1</td>
<td>rs4149015</td>
<td>TGGCCTTGGGTCTACATTTTC</td>
<td>CGTGGGTATGTATGGGAGACTGGA</td>
<td>661</td>
</tr>
<tr>
<td>SLCO1B1</td>
<td>rs11045819</td>
<td>CAACATCGACCTATCCACTTGG</td>
<td>TGTTAATGGCGAACTGTT</td>
<td>249</td>
</tr>
<tr>
<td>SLCO1B1</td>
<td>rs4149036</td>
<td>CCTCTGTGCCACTATCAGTACC</td>
<td>GCTTCAGTGAAATGATGGGAAC</td>
<td>348</td>
</tr>
<tr>
<td>SLCO1B3</td>
<td>rs4149117</td>
<td>CATTGGGGCATTCAGTGTCT</td>
<td>TCTCACAAGGTAACCTGCCACT</td>
<td>307</td>
</tr>
<tr>
<td>SLCO1B3</td>
<td>rs57585902</td>
<td>ACCCACCTTTGTCTGCTGCTGTT</td>
<td>GGTGGTTGGTTTCTGCTTTCT</td>
<td>594</td>
</tr>
<tr>
<td>SLCO1B3</td>
<td>rs60140950</td>
<td>GCACACAAGATCACAGCAGAATG</td>
<td>CTTGGTTGCTAAATAGC</td>
<td>465</td>
</tr>
<tr>
<td>SLCO1B3</td>
<td>rs12299012</td>
<td>GGTGGAATGCCCAAGAGATA</td>
<td>CCCCCTACATCCCCCTATT</td>
<td>888</td>
</tr>
<tr>
<td>SLCO1B3</td>
<td>rs72559743</td>
<td>ACACACTGTGCTGATCTCTCT</td>
<td>TGTTGACCTCTGTTGCAG</td>
<td>764</td>
</tr>
</tbody>
</table>
3.2.5. **PCR Amplification:**

The PCR reactions were performed as described in Chapter two. Cycling parameters were carried out as follows: initial denaturation at 95°C for 3 min, a total of 30 cycles of denaturation at 95°C for 15 s, annealing at 55°C for 15 sec, extension at 72°C for 15 sec, followed by final extension at 72°C for 1 min and a hold step at 4°C.

3.2.6. **Post PCR purification:**

PCR products were purified as described in Chapter Two.

3.2.7. **SNaPshot minisequencing reaction (Single base extension):**

Minisequencing reaction was performed in a 10 µl reaction volume using 3 µl of the purified PCR products, 0.4 µM of primers, and 2 µl of SNaPshot™ ready reaction mix (Applied Biosystems, U.S.A). Sequence cycling was performed according to the manufacturer’s instructions.
Table 3.2-SNaPshot minisequencing reaction primers of the selected SNPs related to cholesterol lowering drugs.

<table>
<thead>
<tr>
<th>NCBI (dbSNP)</th>
<th>Amino Acid Change</th>
<th>Nucleotide Change</th>
<th>Minisequencing Reaction Primers</th>
<th>Primer Direction</th>
<th>Position Accession number (NC_000006.12)</th>
<th>Size bp</th>
<th>Poly GCAT tail</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs4149056</td>
<td>V174A</td>
<td>T&gt;C</td>
<td>GGAATCTGGGTCATACATGTTGGATATATG</td>
<td>Forward</td>
<td>21178615</td>
<td>30</td>
<td>0</td>
</tr>
<tr>
<td>rs2306283</td>
<td>N130D</td>
<td>G&gt;A</td>
<td>AGTTACAGGTATTCTAAAGAAAATTAATTC</td>
<td>Forward</td>
<td>21176804</td>
<td>35</td>
<td>5</td>
</tr>
<tr>
<td>rs4149015</td>
<td>Intron</td>
<td>G&gt;A</td>
<td>CATTTCACACATATACATGTACATA</td>
<td>Reverse</td>
<td>21130388</td>
<td>40</td>
<td>10</td>
</tr>
<tr>
<td>rs11045819</td>
<td>P155T</td>
<td>C&gt;A</td>
<td>GACTCTATACCCCTTTCATTCTAAATTA</td>
<td>Reverse</td>
<td>21176879</td>
<td>45</td>
<td>15</td>
</tr>
<tr>
<td>rs4149036</td>
<td>Intron</td>
<td>C&gt;A</td>
<td>TTATTATTATCTCTTTAATGGCAATGC</td>
<td>Forward</td>
<td>21174806</td>
<td>50</td>
<td>20</td>
</tr>
<tr>
<td>rs4149117</td>
<td>S112A</td>
<td>G&gt;T</td>
<td>CCTCTTATGGGAATCTTATTTGGAC</td>
<td>Forward</td>
<td>20858546</td>
<td>55</td>
<td>25</td>
</tr>
<tr>
<td>rs57585902</td>
<td>T147A</td>
<td>A&gt;G</td>
<td>GGAATAGTAAGTTACCTTGATGGAGCTA</td>
<td>Reverse</td>
<td>20861096</td>
<td>60</td>
<td>30</td>
</tr>
<tr>
<td>rs60140950</td>
<td>G256A</td>
<td>G&gt;C</td>
<td>GAATAACTCTAAGGACTTGTGGTTG</td>
<td>Forward</td>
<td>20875274</td>
<td>65</td>
<td>35</td>
</tr>
<tr>
<td>rs12299012</td>
<td>V560A</td>
<td>T&gt;C</td>
<td>CAGGAGGTACCACATTTATCTTTTGACTG</td>
<td>Forward</td>
<td>20883599</td>
<td>70</td>
<td>40</td>
</tr>
<tr>
<td>rs72559743</td>
<td>G522C</td>
<td>G&gt;T</td>
<td>CACCTACGGGTCTCTATTATGAATGTTT</td>
<td>Reverse</td>
<td>20883484</td>
<td>75</td>
<td>45</td>
</tr>
</tbody>
</table>
3.2.8. Post Extension Purification

Post-extension products were purified as described in Chapter Two.

3.2.9. Electrophoresis of the minisequencing products

The fluorescently labelled fragments were separated on an ABI Prism 3500 Genetic Analyzer (Applied Biosystems, USA) as described in Chapter Two. Data was analyzed using GeneMapper® IDX Software Version 1.4 (Applied Biosystems, USA).

3.2.10. Statistical Analysis

The statistical analysis of the data was done using the freely available software, GenAlEx as described in Chapter Two, as well as the SHEsis analysis platform was used to deduce the haplotype frequencies (Yong and Lin 2005; Li et al. 2009).

3.3. Results:

3.3.1. SNaPshot minisequencing Genotyping:

Our studied population composed of a 130 healthy, unrelated Cape Admixed participants residing in Cape Town, South Africa. There were 12 (9%) females and 118 (91%) males between the ages of 14 and 68 years.

In this chapter, a SNaPshot multiplex assay was developed for genotyping five SNPs in \textit{SLCO1B1} and five SNPs in \textit{SLCO1B3} transporters genes in intronic and coding regions that are associated with FDA cholesterol lowering drugs (statins). The original primer concentration for the amplification reactions was 0.2 µM. However, minor optimization of the SNaPshot minisequencing reaction was required and the primer concentration was increased to 0.4µM for all the SNPs to get better results for each. As described in Chapter Two a
notable number of shifts for fragments during capillary electrophoresis were detected. Genotypes were confirmed with direct sequencing is discussed in detail in Chapter Four. A typical electropherogram profile of the SNaPshot minisequencing genotyping system of the ten selected SNPs relevant to cholesterol lowering drugs is displayed in Figure 3.1.

![Electropherogram profile of the SNPs related to cholesterol lowering drugs.](image)

**Figure 3.1**- Electropherogram profile of the SNPs related to cholesterol lowering drugs. Green peaks indicate an A nucleotide, blue peaks indicate a G nucleotide, black peaks indicate a C nucleotide and red peaks indicate a T nucleotide.

SNaPshot genotyping results in the Cape Admixed are reported in Table 3.3. The minor allele frequency (MAF) for the ten variants was compared to seven different ethnic groups which were: BGR (British in England and Scotland), CLM (Colombian in Medellin, Colombia), MXL (Mexican ancestry in Los Angeles in California), GIH (Gujarati Indian in Houston), ASW (African ancestry in Southwest US), LWK (Luhya in Webuye, Kenya) and YRI (Yoruba in Ibadan, Nigeria), summarized in Table 3.4 and depicted in Figure 3.2.

Only two SNPs (rs4149056 and rs2306283) were determined to be in HWE (p>0.05). Four out of the ten studied SNPs were monomorphic in the Cape Admixed population. None of the participants were homozygous for the variant allele for N130D (rs2306283), P155T
(rs11045819), G256A (rs60140950), S112A (rs4149117), V174A (rs4149056) and the intronic (rs4149036). The N130D variant genotype frequencies for wild-type homozygote (GG), heterozygote (GA) and mutant homozygote (AA) were 50.8%, 36.2% and 13%, respectively (Table 3.3). The P155T variant genotype frequencies for wild-type homozygote (CC), heterozygote (CA) and mutant homozygote (AA) were 61.5%, 38.5% and 0.0%, respectively. The genotype frequencies for G256A variant wild-type homozygote (GG), heterozygote (GC) and mutant homozygote (CC) were 96.1%, 3.1% and 0.8%, respectively. Interestingly, while this variant was not observed in African American and Yoruba; it was observed in the Cape Admixed, British, Colombian and Mexican populations (Table 3.4). The S112A variant genotype frequencies for wild-type homozygote (GG), heterozygote (GT) and mutant homozygote (TT) were 63.1%, 36.9% and 0.0%, respectively. The intronic variant rs4149036 genotype frequencies for the wild-type homozygote (CC), heterozygote (CA) and mutant homozygote (AA) were 48.5%, 50% and 1.5%, respectively. The variant V174A genotype frequency for wild-type homozygote (CC) was 90.8%, while heterozygote (CT) and mutant homozygote (TT) were 8.4% and 0.8%, respectively. The intronic variant rs4149015 was not observed in the Cape Admixed and Yoruba. However, it was observed in the other populations. Two variants, T147A (rs57585902) and V560A (rs12299012) were not observed in the Cape Admixed, British, Colombian and Mexican; however, they were observed in the African American and the sub-Saharan populations. No population data was found for the variant G522C (rs72559743) on the database, which was not observed in the Cape Admixed population.
Table 3.3 - Genotype and allele frequencies of the selected SNPs related to cholesterol lowering drugs in 130 healthy Cape Admixed individuals.

<table>
<thead>
<tr>
<th>Amino Acid Substitution</th>
<th>Gene</th>
<th>dbSNP ID</th>
<th>Observed Genotype Frequency</th>
<th>Allele Frequency</th>
<th>HWE (P)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Genotype</td>
<td>%</td>
<td>95% CI</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V174A</td>
<td>SLCO1B1</td>
<td>rs4149056</td>
<td>TT</td>
<td>90.8</td>
<td>85.83-95.77</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TC</td>
<td>8.4</td>
<td>3.63-13.17</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CC</td>
<td>0.8</td>
<td>0.0-2.33</td>
</tr>
<tr>
<td>N130D</td>
<td>SLCO1B1</td>
<td>rs2306283</td>
<td>GG</td>
<td>50.8</td>
<td>42.21-59.39</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GA</td>
<td>36.2</td>
<td>27.94-44.46</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>AA</td>
<td>13</td>
<td>7.22-18.78</td>
</tr>
<tr>
<td>Intron SLCO1B1</td>
<td>rs4149015</td>
<td></td>
<td>GG</td>
<td>100.0</td>
<td>100-100</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GA</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>AA</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>P155T</td>
<td>SLCO1B1</td>
<td>rs11045819</td>
<td>CC</td>
<td>61.5</td>
<td>53.14-69.86</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CA</td>
<td>38.5</td>
<td>30.14-46.86</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>AA</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Intron SLCO1B1</td>
<td>rs4149036</td>
<td></td>
<td>CC</td>
<td>48.5</td>
<td>39.91-57.09</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CA</td>
<td>50</td>
<td>41.41-58.59</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>AA</td>
<td>1.5</td>
<td>0.0-3.59</td>
</tr>
<tr>
<td>S112A</td>
<td>SLCO1B3</td>
<td>rs4149117</td>
<td>GG</td>
<td>63.1</td>
<td>54.81-71.39</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GT</td>
<td>36.9</td>
<td>28.61-45.19</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TT</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>T147A</td>
<td>SLCO1B3</td>
<td>rs57585902</td>
<td>AA</td>
<td>100.0</td>
<td>100-100</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>AG</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GG</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>G256A</td>
<td>SLCO1B3</td>
<td>rs60140950</td>
<td>GG</td>
<td>96.1</td>
<td>92.77-99.43</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GC</td>
<td>3.1</td>
<td>0.12-6.08</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CC</td>
<td>0.8</td>
<td>0.0-2.33</td>
</tr>
<tr>
<td>V560A</td>
<td>SLCO1B3</td>
<td>rs12299012</td>
<td>TT</td>
<td>100.0</td>
<td>100-100</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TC</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CC</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>G522C</td>
<td>SLCO1B3</td>
<td>rs72559743</td>
<td>GG</td>
<td>100.0</td>
<td>100-100</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GT</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TT</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

P-Value will be NA for monomorphic SNPs (MAF =0).
Table 3.4 - Comparison of MAF of the SNPs related to cholesterol lowering drugs of the Cape Admixed population to other ethnic groups.

<table>
<thead>
<tr>
<th>dbSNP ID</th>
<th>Amino acid change</th>
<th>Minor Allele</th>
<th>Cape admixed</th>
<th>British</th>
<th>Colombian</th>
<th>Mexican</th>
<th>Indian</th>
<th>African-American</th>
<th>Luhya</th>
<th>Yoruba</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs4149056</td>
<td>V174A</td>
<td>C</td>
<td>5.0</td>
<td>14.3</td>
<td>18.1</td>
<td>7.8</td>
<td>1.9</td>
<td>6.6</td>
<td>2.0</td>
<td>0.9</td>
</tr>
<tr>
<td>rs2306283</td>
<td>N130D</td>
<td>A</td>
<td>31.2</td>
<td>64.3</td>
<td>52.1</td>
<td>62.5</td>
<td>44.7</td>
<td>25.4</td>
<td>15.7</td>
<td>18.5</td>
</tr>
<tr>
<td>rs4149015</td>
<td>Intron</td>
<td>A</td>
<td>0.0</td>
<td>2.2</td>
<td>6.9</td>
<td>2.3</td>
<td>5.8</td>
<td>0.8</td>
<td>3.0</td>
<td>0.0</td>
</tr>
<tr>
<td>rs11045819</td>
<td>P155T</td>
<td>A</td>
<td>19.2</td>
<td>15.4</td>
<td>14.9</td>
<td>5.5</td>
<td>2.9</td>
<td>6.6</td>
<td>2.5</td>
<td>5.1</td>
</tr>
<tr>
<td>rs4149036</td>
<td>Intron</td>
<td>A</td>
<td>26.5</td>
<td>18.1</td>
<td>25.0</td>
<td>12.5</td>
<td>10.7</td>
<td>54.9</td>
<td>56.6</td>
<td>59.7</td>
</tr>
<tr>
<td>rs4149117</td>
<td>S112A</td>
<td>T</td>
<td>18.5</td>
<td>13.2</td>
<td>11.2</td>
<td>13.3</td>
<td>6.8</td>
<td>51.6</td>
<td>72.2</td>
<td>65.3</td>
</tr>
<tr>
<td>rs57585902</td>
<td>T147A</td>
<td>G</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>4.1</td>
<td>2.5</td>
<td>5.6</td>
</tr>
<tr>
<td>rs60140950</td>
<td>G256A</td>
<td>C</td>
<td>2.3</td>
<td>15.9</td>
<td>14.9</td>
<td>4.7</td>
<td>9.2</td>
<td>0.0</td>
<td>0.5</td>
<td>0.0</td>
</tr>
<tr>
<td>rs12299012</td>
<td>V560A</td>
<td>C</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.8</td>
<td>2.0</td>
<td>2.0</td>
<td>2.8</td>
</tr>
<tr>
<td>rs72559743</td>
<td>G522C</td>
<td>T</td>
<td>0.0</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
</tbody>
</table>

a. This study, b. Data from Ensembl database (http://www.Ensembl.org)
Figure 3.2- Allele frequencies of the selected SNPs related to cholesterol lowering drugs compared to other ethnic groups. Cape Admixed (This study); British, Colombian, Mexican, Indian, African American, Luhya and Yuroba (Ensembl database).
3.3.2. Haplotype analysis:

Haplotypes were calculated using the SHesis online platform. Twelve haplotypes were distinguished from the five SNPs in the *SLCO1B1* gene, and four haplotypes were identified from the five SNPs in the *SLCO1B3* gene (Table 3.5A and Table 3.5 B). The most frequently observed haplotypes in the *SLCO1B1* gene were T G G C C (33.2%), T A G C C (28.3%), T G G C A (14.7), T G G A A (9.4%) and T G G A C (7.8%), while the most frequently observed haplotypes in *SLCO1B3* gene were G A G T G (79.2%) and T A G T G (18.5%).
**Table 3.5 A** - Haplotype structure defined by the 5 SNPs in the *SLCO1B1* gene in the Cape Admixed population.

<table>
<thead>
<tr>
<th>Haplotype No.</th>
<th>Haplotypes (1)</th>
<th>Frequency %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Haplotype *1</td>
<td>T G G C C</td>
<td>33.2</td>
</tr>
<tr>
<td>Haplotype *2</td>
<td>T A G C C</td>
<td>28.3</td>
</tr>
<tr>
<td>Haplotype *3</td>
<td>T G G C A</td>
<td>14.7</td>
</tr>
<tr>
<td>Haplotype *4</td>
<td>T G G A A</td>
<td>9.4</td>
</tr>
<tr>
<td>Haplotype *5</td>
<td>T G G A C</td>
<td>7.8</td>
</tr>
<tr>
<td>Haplotype *6</td>
<td>C G G C A</td>
<td>1.7</td>
</tr>
<tr>
<td>Haplotype *7</td>
<td>T A G A C</td>
<td>1.6</td>
</tr>
<tr>
<td>Haplotype *8</td>
<td>C G G C C</td>
<td>1.5</td>
</tr>
<tr>
<td>Haplotype *9</td>
<td>C A G C A</td>
<td>0.7</td>
</tr>
<tr>
<td>Haplotype *10</td>
<td>C A G C C</td>
<td>0.6</td>
</tr>
<tr>
<td>Haplotype *11</td>
<td>C G G A C</td>
<td>0.5</td>
</tr>
<tr>
<td>Haplotype *12</td>
<td>T A G C A</td>
<td>0.0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>100</strong></td>
</tr>
</tbody>
</table>

(1) - Haplotype sequences are based on the position of SNPs on chromosome 12.

**Table 3.5 B** - Haplotype structure defined by the 5 SNPs in the *SLCO1B3* gene in the Cape Admixed population.

<table>
<thead>
<tr>
<th>Haplotype No.</th>
<th>Haplotypes (1)</th>
<th>Frequency %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Haplotype*1</td>
<td>G A G T G</td>
<td>79.2</td>
</tr>
<tr>
<td>Haplotype*2</td>
<td>T A G T G</td>
<td>18.5</td>
</tr>
<tr>
<td>Haplotype*3</td>
<td>G A C T G</td>
<td>2.3</td>
</tr>
<tr>
<td>Haplotype*4</td>
<td>T A C T G</td>
<td>0.0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>100</strong></td>
</tr>
</tbody>
</table>
3.4. Discussion:

The SNaPshot minisequencing reaction was selected as it is an easy, accurate, and high throughput method for SNP analysis. Minimal optimization of increasing the primer concentration to 0.4 µM for each SNP was required to get good profiles. As a part of the project, ten SNPs in *SLCO1B1* and *SLCO1B3* genes were genotyped in 130 unrelated healthy individuals of the Cape Admixed population in South Africa.

Drug transporters are important proteins involved in drug absorption, tissue accumulation and elimination from the body. Thus, recent interest has focussed on the contribution of inter-individual variability that affect drug transporter function in drug response (DeGorter 2012). OATPs transporters encoded by the *SLCO* genes are uptake transporters with wide substrate specificity although statins are considered to be their commonly specified substrates. Furthermore, hepatocytes are the major site of statins action where transporter proteins from OATP1B subfamily are expressed (DeGorter 2012; Romaine *et al*. 2010).

The common variant V174A (rs4149056) was found to be associated with altering the transporter activity of pravastatin *in vitro* (Tirona *et al*. 2001). Moreover, different studies have confirmed that this substitution is linked to decreased hepatic transporter activity of rosvastatin, atorvastatin and pravastatin (Tirona *et al*. 2003; Katz *et al*. 2006; Oswald *et al*. 2008). Thompson *et al*. studied the variant V174A (rs4149056), and noticed a significant effect on the ability of fluvastatin to affect HDL cholesterol (Thompson *et al*. 2005). However, they did not notice any significant associations for LDL cholesterol lowering.

Furthermore, a study enrolled elderly patients with vascular risk or at risk of vascular disease showed that the V174A variant showed less pravastatin induced low LDL-Cholesterol levels (Akao *et al*. 2012). A previous study identified that the V174A was correlated with reduced plasma concentrations of pravastatin in paediatric heterozygous familial
hypercholesterolemia (HeFH) (Hedman et al. 2006). The MAF for V174A recorded by Rajput et al. in study conducted on Pakistani population was observed at 23.9% (Rajput et al. 2014). However, in the present study the variant was observed to have a MAF of 5% in the Admixed population which was lower than Mexican and African American, but higher than sub-Saharan and Indian populations (Table 3.4).

It was reported that the variant N130D (rs2306283) was responsible for OATP expression in the liver cells (Nies et al. 2013). In a study, it has been found that N130D associated with a significant increase in atorvastatin response, which was an evaluated reduction in LDL-cholesterol (Rodrigues et al. 2011). Also, Kadam and co-workers identified in their recent study of 177 hypercholesterolaemic Indian patients undergoing atorvastatin treatment, that the variant N130D displayed a notable reduction of LDL-cholesterol in response to atorvastatin (Kadam et al. 2016). This variation was observed in our population with a MAF of 31.2% which was significantly lower than the Hispanic, Indian and British populations.

Hedman et al. (2006) reported that the intronic SNP rs4149015 was associated with reduced plasma concentrations of pravastatin in children with heterozygous familial hypercholesterolemia (HeFH). This variant was not observed in Cape Admixed or in the Yoruba. However, it was recorded in British, Hispanic, Indian and African American populations (Table 3.4).

In a study of 420 French patients with hypercholesterolemia used fluvastatin, Couvert and co-workers reported that the SLCO1B1 variant, P155T (rs11045819) reduced LDL in the patients with the CC genotype more than in CA and AA genotypes (Couvert et al. 2008). In the Cape Admixed, this variant demonstrated a MAF at 19.2% which was higher than Colombian (14.9%) and British (15.4%) populations.
The intronic SNP rs4149036 was identified to be linked with triglyceride lowering in atorvastatin treated individuals with the uncommon homozygous AA (Thompson et al. 2005). This study found that the MAF for this variant was 26.5% which was a bit higher than that in Colombians (25%).

The \textit{SLCO1B3} S112A (rs41429117), has been shown to increase transporter activity \textit{in vitro} (Letschert et al. 2004). The T allele of S112A in Italians and Hungarians was observed at 29.4% and 47.8%, respectively (Nagy et al. 2015). It was detected in Chinese (35%), Malay (20%) and Indians (8%) in a study by Chew et al. (2011) and 15% in Caucasians (Baker et al. 2009). However, in the present study this variant was observed with a MAF of 18.5%.

In a study of HeLa cell lines, V560A (rs12299012) in \textit{SLCO1B3} showed reduced uptake activity for rosuvastatin compared to the wild-type (Schwarz et al. 2011). To our knowledge, clinical data is currently lacking for the variant T147A (rs57585902). However, this SNP and V560A (rs12299012) were not observed in Chinese, Malays and Indians (Chew et al. 2011) which corresponds with our population, British, Hispanic and Gujarati Indian populations. Both variants were observed in Caucasians with a MAF of 0.5% and 1.6%, respectively (Baker et al. 2009).

The variant G256A (rs60140950) associated with the expression of \textit{SLCO1B3} gene was reported to have a negative effect on protein function in Caucasian Europeans (Geraldine et al.). Chew and colleagues also investigated the variant G256A (rs60140950) and it was found in Malays and Indians with MAF of 4% and 5%, respectively, but not in Chinese (Chew et al. 2011). Similarly, a previous study (Baker et al. 2009) detected this variant in Caucasians with a MAF of 1.9% which roughly corresponded with our population result that showed a MAF at 2.3%.
The rare SLCO1B3 variant, G522C (rs72559743) is reported to decrease protein expression and reduce the transporter activity which may be because of its location in the key regions of the transporter OATP1B3 (Letschert et al. 2004; Schwarz et al. 2011). It is notable that there is no population data on this variant on the database. However, Letschert et al. (2004) detected this SNP in Caucasians with a MAF of 2%. Chew et al. (2011) did not observe it in Chinese, Malays and Indians populations. This allele was also not detected in our study of the Cape Admixed.

The incorporation of haplotypes in pharmacogenetic studies is believed to provide a more complete picture of loci that are applicable in genetic medicine to determine individual or population response to relevant drugs (Crawford and Nickerson 2005). In our study represented here, the haplotype structure defined for the ten SNPs was determined for the Cape Admixed population. Twelve haplotypes were detected from the five SNPs in SLCO1B1 and four haplotypes were inferred from the five SNPs in SLCO1B3. This data of haplotype structure could provide the basis of more clinical and in vitro studies to determine a suitable dosage for individuals with high efficacy and low toxicity.

3.5. Conclusion:

To conclude, in this chapter, we have focused on particular genes (SLCO1B1 and SLCO1B3) which encode for transporters located on the basolateral membrane of hepatocytes. These transporters have a significant role in the transport of statins. SNaPshot assay was developed to genotype five SNPs in SLCO1B1 and five SNPs in SLCO1B3. The obtained results were compared to seven different ethnic groups. In addition, haplotypes were inferred from these genes. Studying solute carrier transporters (SLCs) genetic variations and their pharmacogenetics implications in African populations will fill the gap in the missing data of
pharmacogenetics. Pharmacogenetics can be a robust tool to investigate the effect of genetic differences in these genes and their impact on drug response.

In the next chapter, we are going to discuss high resolution melt technique as an alternative genotyping system.
Chapter Four

Development and performance evaluation of High Resolution Melt Analysis Genotyping System

4.1. Introduction

SNPs are commonly investigated as markers of genetic variation in pharmacogenetics studies of cancer and other complex diseases (Martino et al. 2010). During the past decade, many high throughput genotyping techniques have been improved and developed (Martino et al. 2010). However, in the recent years it has been suggested that high resolution melt technique (HRM) is the most widely used technique in genome analysis (Martino et al. 2010). HRM is a cheap, fast and easy method to discriminate single base variations and small insertions and deletions (Druml and Cichna-Markl 2014; Prajantasen et al. 2015). HRM analysis was first introduced by Ririe et al. (1997). It is a real-time PCR-based technique for mutation and genotyping studies (Ezgu et al. 2014; Mastoraki et al. 2015). It is based on the principle that different PCR products vary in their melting temperature depending on the percentage of GC, length and sequence (Druml and Cichna-Markl 2014; Prajantasen et al. 2015). HRM is ideal for SNP genotyping, however, it is not recommended for screening genes with highly variable SNPs (Vondráčková et al. 2015). Given that a HRM approach offers several, the aim of this part of the study was to develop an HRM-based method for genotyping SNPs associated with FDA approved anticancer and lowering cholesterol drugs in the Cape Admixed population. HRM performance was evaluated and compared with SNaPshot as it is the standard genotyping system which its results were confirmed by direct sequencing.
4.2. Materials and Methods

4.2.1. Sample Collection

Samples were collected as described in Chapter 2.

4.2.2. DNA Extraction

Genomic DNA was extracted as described in Chapter 2.

4.2.3. SNP Selection

SNPs were selected as described in Chapters 2 and 3.

4.2.4. Primer Design

All primers were designed by using Primer3plus tool and synthesized by Inqaba Biotec, South Africa (Table 4.1). The amplicons were made with the size ranging from 154-280 bp then aligned at Ensembl using BLAST (http://www.ensembl.org/Multi/Tools/Blast). Primers were diluted to a working stock of 2 μM.
Table 4.1 - A list of selected SNPs primers used in HRM analysis.

<table>
<thead>
<tr>
<th>Gene</th>
<th>dpSNP ID</th>
<th>Nucleotide Change</th>
<th>Forward primer</th>
<th>Reverse primer</th>
<th>Amplicon length (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>DPYD</td>
<td>rs3918390</td>
<td>C/T</td>
<td>TGCCAATTCTCTTGTTTAAGATG</td>
<td>GACAAATGTTCCCCCAGAA</td>
<td>195</td>
</tr>
<tr>
<td>DPYD</td>
<td>rs2297595</td>
<td>T/C</td>
<td>GCGAAGGATTTCTGATCTG</td>
<td>AATGTCACCAGTGAGTTT</td>
<td>278</td>
</tr>
<tr>
<td>TPMT</td>
<td>rs1800460</td>
<td>C/T</td>
<td>TTACCATTGCGATCACC</td>
<td>TAAAACCATGAGGGATGGA</td>
<td>265</td>
</tr>
<tr>
<td>MTHFR</td>
<td>rs1801133</td>
<td>G/A</td>
<td>GAAAGCTCGTGATGATGA</td>
<td>TGTGGAAAGTGCAAGATCA</td>
<td>231</td>
</tr>
<tr>
<td>TPMT</td>
<td>rs1142345</td>
<td>T/C</td>
<td>ATTTATCTATGTCTATTACTTACTTGT</td>
<td>CATGTACTCTTTTCTGTTTGATGTTAACAT</td>
<td>154</td>
</tr>
<tr>
<td>SLCO1B1</td>
<td>rs4149056</td>
<td>T/C</td>
<td>TCTACATAGGGTTTAAAGGAATCTGG</td>
<td>AAAGTACAAAGGGAAAGTGATCATAC</td>
<td>208</td>
</tr>
<tr>
<td>SLCO1B3</td>
<td>rs60140950</td>
<td>G/C</td>
<td>CACTATCAGAAATACTCTAAGGACTC</td>
<td>CTGACTCTAGATGATTTGAGTATGGTTT</td>
<td>343</td>
</tr>
<tr>
<td>SLCO1B3</td>
<td>rs57585902</td>
<td>A/G</td>
<td>CATCAGAAATTTCAACATCAAGTTATC</td>
<td>GGAATAATTGTTCCTCTATTTGTTCTAAA</td>
<td>209</td>
</tr>
<tr>
<td>SLCO1B1</td>
<td>rs11045819</td>
<td>C/A</td>
<td>ATTAATCAAATTTATCATCAATAGAGCA</td>
<td>GGGCGAACTGTGTATATTAAACACTATA</td>
<td>216</td>
</tr>
<tr>
<td>SLCO1B1</td>
<td>rs2306283</td>
<td>G/A</td>
<td>ATTCAGTGATGTTCTTACAGTACAGGGTAT</td>
<td>AATTATGTCTGTAAAGTCAATGTTTTTC</td>
<td>280</td>
</tr>
</tbody>
</table>
4.2.5. HRM Analysis

HRM was done using the KAPA HRM FAST PCR kit. The final reaction volume of 20 µl contained 15 ng of DNA and was set up according to manufacturer’s instructions. HRM analysis was performed on a 36-well Rotor Gene Q (Qiagen) real-time PCR thermocycler. Thermocycling conditions were: initial hold at 95°C for 3 min, followed by a total of 40 cycles of denaturation at 95°C for 5 sec, annealing at 63°C for 20 sec and extension at 72°C for 30 sec with fluorescence acquisition, then HRM was performed over a melting temperature ranging from 70°C-95°C at the rate of 0.1 °C per sec. Melt profiles were then analysed using the Rotor Gene data collection software (Qiagen).

4.2.6. HRM sensitivity:

HRM sensitivity was calculated by the number of genotypes obtained over the total reactions (Martino et al. 2010), while error rate was calculated as the number of genotypes different between HRM and SNaPshot divided by the number of genotypes obtained by HRM (Cui et al. 2013).

4.2.7. Sequencing analysis:

Direct sequencing was used to validate SNaPshot genotyping results, which the later was used to evaluate and validate the HRM performance as a standard for genotyping.
4.3. Results:

4.3.1. PCR and HRM optimization:

4.3.1.1. Optimizing annealing temperature:

Firstly, all primers sets needed to be optimized to determine an appropriate annealing temperature (T\textsubscript{a}). Therefore, gradient PCR was performed for each primer set and any resultant amplicons were analyzed by 2% agarose gel electrophoresis. Figure 4.1 showed the result of optimizing annealing temperature for the SNP rs60140950 HRM primer which the bands had quite similar brightness. However, the annealing temperature at 63 gave the best result.

Figure 4.1- Annealing temperature optimization of the SNP (rs60140950) by gradient PCR on 2% agarose gel. 1: at 58°C, 2: at 59°C, 3: at 60°C, 4: at 62°C, 5: at 63°C.
4.3.1.2. Optimizing MgCl$_2$ concentration:

MgCl$_2$ concentration was the main factor that affected our results. Five concentrations of MgCl$_2$ (2.5mM, 3.5mM, 5mM, 7.5mM and 10mM) were tested. However, the best result that could discriminate heterozygote samples was 10mM MgCl$_2$ (Figure 4.2).

![Image of 2% agarose gel](image)

**Figure 4.2**- MgCl$_2$ optimization of the SNP (rs60140950) on 2% agarose gel. 1: 2.5mM, 2.5mM, 3: 7.5mM, 4: 10 mM.

4.3.2. Genotyping:

The homozygotes samples of the investigated SNPs rs3918290, rs1800460, and rs57585902 were detected easily. However, no heterozygotes samples for the SNPs rs1801133, rs1142345, rs11045819 and rs60140950 were able to be discriminated. They were detected only for the SNPs rs4149056, rs2306283 and rs2297595 (Figure 4.3).

As shown in Figure 4.3, for the SNP rs1801133, the wild type G and the mutant A were detected on the HRM mutation scan, however, the melting temperature ($T_m$) of the mutant A was slightly shifted to the left compared to the wild type G, and the reason of that is AT base pair had a lower melting temperature compared to a GC base pair (Cui *et al.* 2013).
**Figure 4.3** - HRM mutation scan for SNP rs1801133. The wild type G and mutant A can clearly be distinguished.

In Figure 4.4 A and B, heterozygosity in the SNPs rs4149056 and rs2306283 is clearly detected as their peaks were much lower and wider than the peaks for the homozygote samples.

**Figure 4.4 A** - HRM mutation scan of SNP rs4149056. The homozygotes and heterozygote peaks can be clearly distinguished based on peak height.
4.3.3. Evaluation of SNaPshot and HRM systems:

We compared sensitivity of HRM with the SNaPshot assay which was confirmed by direct sequencing (Table 4.2). HRM sensitivity, defined as the number of genotypes obtained over the total number of reactions (Martino et al. 2010), while error rate was calculated by taking the number of genotypes different between HRM and SNaPshot divided by the number of genotypes obtained by HRM (Cui et al. 2013).

The data represented here showed a mean level of HRM sensitivity of 89.2%, while the mean error rate was 20.8 % (Table 4.3). The difference in HRM sensitivity which was demonstrated in this study could be attributed to several factors. However, some previous studies showed that the sensitivity of HRM was affected by the instrument type (De Leeneer et al. 2008; Herrmann et al. 2007; Cui et al. 2013). The differences detected in this study could be due to performing the reactions on a capillary system which is less sensitive than newer micro-titre plate systems.
Table 4.2- Confirmation of SNaPshot genotyping using direct sequencing

<table>
<thead>
<tr>
<th>SNPs</th>
<th>Number of samples</th>
<th>Correct SNaPshot genotypes</th>
<th>*Error rate %</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs2297595</td>
<td>130</td>
<td>126</td>
<td>3.2</td>
</tr>
<tr>
<td>rs1801133</td>
<td>130</td>
<td>127</td>
<td>2.4</td>
</tr>
<tr>
<td>rs1142345</td>
<td>130</td>
<td>130</td>
<td>0</td>
</tr>
<tr>
<td>rs1800460</td>
<td>130</td>
<td>130</td>
<td>0</td>
</tr>
<tr>
<td>rs3918390</td>
<td>130</td>
<td>130</td>
<td>0</td>
</tr>
<tr>
<td>rs4149056</td>
<td>130</td>
<td>121</td>
<td>7.4</td>
</tr>
<tr>
<td>rs60140950</td>
<td>130</td>
<td>126</td>
<td>3.2</td>
</tr>
<tr>
<td>rs57585902</td>
<td>130</td>
<td>130</td>
<td>0</td>
</tr>
<tr>
<td>rs11045819</td>
<td>130</td>
<td>130</td>
<td>0</td>
</tr>
<tr>
<td>rs2306283</td>
<td>130</td>
<td>130</td>
<td>0</td>
</tr>
</tbody>
</table>

*Error rate = number of genotypes different between SNaPshot and direct sequencing/number of correct genotypes obtained by SNaPshot.
Table 4.3- Performance of HRM in comparison to SNaPshot.

<table>
<thead>
<tr>
<th>dp SNP ID</th>
<th>Type of drug</th>
<th>Total samples</th>
<th>Correct profiles</th>
<th>HRM</th>
<th>SNaPshot</th>
<th>Err/AA</th>
<th>Err/AB</th>
<th>Err/BB</th>
<th>Error rate %</th>
</tr>
</thead>
<tbody>
<tr>
<td>rs2297595</td>
<td>Anticancer</td>
<td>50</td>
<td>49</td>
<td>98</td>
<td>96.9</td>
<td>0/47</td>
<td>1/3</td>
<td>0</td>
<td>2.1</td>
</tr>
<tr>
<td>rs1801133</td>
<td>Anticancer</td>
<td>50</td>
<td>38</td>
<td>76</td>
<td>97.7</td>
<td>5/25</td>
<td>7/7</td>
<td>0/18</td>
<td>31.6</td>
</tr>
<tr>
<td>rs1142345</td>
<td>Anticancer</td>
<td>50</td>
<td>48</td>
<td>96</td>
<td>100</td>
<td>0/48</td>
<td>2/2</td>
<td>0</td>
<td>4.2</td>
</tr>
<tr>
<td>rs1800460</td>
<td>Anticancer</td>
<td>50</td>
<td>50</td>
<td>100</td>
<td>100</td>
<td>0/50</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>rs3918390</td>
<td>Anticancer</td>
<td>50</td>
<td>50</td>
<td>100</td>
<td>100</td>
<td>0/50</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>rs4149056</td>
<td>Cholesterol lowering</td>
<td>50</td>
<td>44</td>
<td>88</td>
<td>93.1</td>
<td>0/44</td>
<td>5/6</td>
<td>0</td>
<td>11.1</td>
</tr>
<tr>
<td>rs60140950</td>
<td>Cholesterol lowering</td>
<td>50</td>
<td>49</td>
<td>98</td>
<td>96.9</td>
<td>0/49</td>
<td>1/1</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>rs2306283</td>
<td>Cholesterol lowering</td>
<td>50</td>
<td>32</td>
<td>64</td>
<td>100</td>
<td>0/21</td>
<td>18/22</td>
<td>0/7</td>
<td>56.3</td>
</tr>
<tr>
<td>rs11045819</td>
<td>Cholesterol lowering</td>
<td>50</td>
<td>36</td>
<td>72</td>
<td>100</td>
<td>0/36</td>
<td>14/14</td>
<td>0</td>
<td>38.9</td>
</tr>
<tr>
<td>rs57585902</td>
<td>Cholesterol lowering</td>
<td>50</td>
<td>50</td>
<td>100</td>
<td>100</td>
<td>0/50</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

(a) Sensitivity = number of genotypes obtained/total reactions
(b) Error rate = number of genotypes different between HRM and SNaPshot /number of correct genotypes obtained by HRM.

AA = Homozygous for the ancestral allele, AB= Heterozygous, BB= Homozygous for the minor allele.
4.4. Discussion:

Numerous techniques are developed for SNP genotyping; however, HRM assay are less time consuming, easy and cheap method (Reed and Wittwer 2004; Erali et al. 2008). In addition, it does not require any post PCR steps, as it is done in one tube or plate (Wittwer et al. 2003; Norambuena et al. 2009). In comparison with SNaPshot that needs minor optimization, HRM has been described as an uncomplicated method which can detect several SNPs in one multiplex assay (Pati et al. 2004). In addition, it is reliable, rapid and accurate (Syvänen 1999). Despite that, it is more expensive than other methods (Pati et al. 2004).

The aim of this present study was to develop and explore an inexpensive, fast and alternative method to genotype different SNPs associated with anticancer and cholesterol lowering drugs, and to evaluate and validate the data obtained from HRM by comparing it with the SNaPshot system which is considered the gold standard for genotyping. In this study, five SNPs related to anticancer drugs and five SNPs related to cholesterol lowering drugs were selected to be amplified and analysed individually and a total of 50 samples were genotyped for each SNP. Several factors such as amplicon size, MgCl₂ concentration and DNA quality were taken into consideration in order to successfully performed HRM assay on the selected SNPs.

It has been shown that HRM is more sensitive when small amplicons are used because the differences in the melting temperature (Tm) are greater compared to larger amplicons (Gundry et al. 2008). However, genotyping SNPs by HRM with larger products (amplicon size of 600 to 1000 bp) has been reported (Reed and Wittwer 2004). Thus, in this study, primers were designed to generate the shortest possible amplicons flanking the mutation of interest with high efficiency, low template secondary structure, and low complementarity.
After determining the optimal conditions, HRM was successfully performed. Genotype for each SNP was determined after the melting peaks were normalized. Heterozygotes are easily distinguished by a change in shape and width of melting curve (Graham et al. 2005; Cui et al. 2013; Gundry et al. 2008), while homozygotes could be identified by a shift in melting temperature (Cui et al. 2013), which is in our study differed by approximately 0.5°C. The homozygous SNPs detected in this study were class 1 (G/A, C/T), class 2 (C/A) and class 3 (C/G). In a study by Liew et al., reported that about 4% of homozygous human SNPs of class 3 and class 4 (T/A) cannot be detected by HRM due to the small difference in melting temperature (Tm) generated by homozygous T/A and C/G base pairs (Liew et al. 2004).

4.5. Conclusion:

In this chapter, HRM was used as an alternative, inexpensive and rapid methodology to genotype five SNPs related to anticancer drug therapy and five SNPs related to cholesterol lowering drug therapy (statins). A rigorous optimization was required for the detection heterozygous genotypes. The obtained SNaPshot results were confirmed by direct sequencing to use them as a standard for evaluating HRM performance.

Ultimately, this study found that the SNaPshot assay was the more appropriate method as it needed minor optimization compared to HRM.
Chapter five

Conclusion

SNPs are the most prolific type of variations in the genome, and have been used as molecular markers in a wide range of studies. A number of SNP genotyping technologies have evolved in the last few years. In genetic research, development of new techniques for high-throughput SNP analysis is one of the most stimulating areas with regards to using SNPs, particularly in studies of normal and pathogenic human variations. Despite the remarkable technological advancements, further improvements are still necessary which must focus on the development of cost effective and less time consuming methods. Therefore, progress in genotyping methods is essential for the development of precision medicine as these technologies will advance the healthcare industry and contribute to the evolution of medical science.

African populations have the most genomic diversity in the globe. However, in spite of being the origin of all the modern human beings, genetics studies on these ethnically diverse populations is still limited. In particular, South Africa is home to different ethnic groups from Europe, Asia and Africa. The mix of these groups led to the establishment of a unique population called South African coloured population, which are distinguished in this study as the Cape Admixed. Intensive studies on the genomic diversity in Africa and particularly in South Africa should be made, which could compensate for the lack of pharmacogenetics data that might have significant medical implications and serve as a solution for the wide health burden in this continent.
In this study, we have developed a SNaPshot and a HRM assay for pharmacogenomics profiling. The performance of both genotyping systems was evaluated using direct sequencing.

Ten SNPs associated with anticancer therapy were genotyped in 130 individuals within the Cape Admixed population. The results obtained were compared globally to seven different ethnic groups which were: British in England and Scotland (GBR) who served as representative European Caucasian; Colombian in Medellin, Colombia (CLM) and Mexican ancestry in Los Angeles, California (MXL) who both represented admixed populations; Gujarati Indian in Houston, Texas (GIH) represented Asians, while African ancestry in Southwest USA represented African Americans (ASW); and Luhya in Webuya, Kenya (LWK) and Yoruba in Ibadan, Nigeria (YRI) both served as representatives of sub-Saharan African populations.

In addition to anticancer drugs profiling, five SNPs of *SLCO1B1* and five SNPs of *SLCO1B3* associated with cholesterol lowering drugs were investigated and the data obtained was also compared to the seven different ethnic groups. The haplotype structures were inferred for the *SLCO1B1* and *SLCO1B3* genes. From the obtained results, as expected the Cape Admixed population shares genetic characteristics with several other global populations.

A SNaPshot minisequencing system was successfully designed, developed and used to offer a quick, high throughput and accurate technique that required minor optimization. HRM method was designed and used to genotype five SNPs related to anticancer drugs and five SNPs related to cholesterol lowering drugs. The performance of HRM was evaluated by comparing with the SNaPshot assay results as it is the gold standard genotyping technique. Both assays were validated using direct sequencing. Despite its advantages - closed tube,
cheap and rapid method for identifying genetic variations, HRM is more time consuming to optimize and may require primer redesigning as well as optimization of reaction conditions.

In future research, we suggest exploring the use of the MassARRAY® System (Agena Bioscience) for pharmacogenomics SNPs genotyping. This system combines mass spectrometry, a sensitive and robust chemistry, and advanced data analysis software to meet the assay design, validation, and performance needs of genomic laboratories. In addition, the developed genotyping systems could be further validated using clinical samples from patients. This could help in optimizing drug therapy for cancer and cholesterol.
References


Erasmus RT, Soita DJ, Hassan MS, Blanco-Blanco E, Vergotine Z, Kengne AP, Matsha TE (2012) High prevalence of diabetes mellitus and metabolic syndrome in a South


Gundry CN, Dobrowolski SF, Martin YR, Robbins TC, Nay LM, Boyd N, Coyne T, Wall MD, Wittwer CT, Teng DH-F (2008) Base-pair neutral homozygotes can be discriminated by calibrated high-resolution melting of small amplicons. Nucleic acids research 36 (10):3401-3408


Hayes JD, McLellan LI (1999) Glutathione and glutathione-dependent enzymes represent a co-ordinately regulated defence against oxidative stress. Free radical research 31 (4):273-300


Jacobs CW (2014) Pharmacogenomics of solute carrier transporter genes in the Xhosa population.


Letschert K, Keppler D, König J (2004) Mutations in the SLCO1B3 gene affecting the substrate specificity of the hepatocellular uptake transporter OATP1B3 (OATP8). Pharmacogenetics and genomics 14 (7):441-452


10.1038/ng0208-129


Manzoni M, Rollini M (2002) Biosynthesis and biotechnological production of statins by filamentous fungi and application of these cholesterol-lowering drugs. Applied Microbiology and Biotechnology 58 (5):555-564


Rajput TA, Naveed AK, Khan S, Farooqi Z-U-R (2014) Frequencies of two functionally significant SNPs and their haplotypes of organic anion transporting polypeptide 1B1 SLCO1B1 gene in six ethnic groups of Pakistani population. Iranian journal of basic medical sciences 17 (6):441


Reed GH, Kent JO, Wittwer CT (2007) High-resolution DNA melting analysis for simple and efficient molecular diagnostics.


treatment-related severe toxicity: a prospective clinical trial by the German 5-FU Toxicity Study Group. Journal of Clinical Oncology 26 (13):2131-2138


distribution and association with paclitaxel pharmacokinetics. Clinical Pharmacology & Therapeutics 81 (1):76-82


receptor activation. Journal of Pharmacology and Experimental Therapeutics 304 (1):223-228


Via M, Ziv E, Burchard EG (2009) Recent advances of genetic ancestry testing in biomedical research and direct to consumer testing. Clinical genetics 76 (3):225-235

