

https://dspace.org/

Gene variant screening for biomarkers significant for effective ARV therapy with a focus on tenofovir nephrotoxicity

Tshabalala, Elizabeth Sibongile

2018
https://hdl.handle.net/10539/26975

Downloaded from a DSpace repository. https://dspace.org/

Gene variant screening for biomarkers significant for effective ARV therapy with a focus on tenofovir nephrotoxicity

Elizabeth Sibongile Tshabalala

Supervisors: Prof Michèle Ramsay and Dr. Dalu Mancama

A dissertation submitted to the Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Master of Science in Medicine (Human Genetics)

Johannesburg, 2017

DECLARATION

I, Elizabeth Sibongile Tshabalala, declare that this dissertation is my own, unaided work. It is being submitted for the degree of Master of Science in Medicine (Human Genetics) at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.

……..

Elizabeth Sibongile Tshabalala

…… day of ……... 20…… in ……

PRESENTATIONS ARISING FROM THIS STUDY

1. Southern African Society for Human Genetic (SASHG), Pretoria, September 2015. Presented a poster entitled 'Next generation sequencing to identify new variants and to establish the frequency of known variants in genes implicated in Tenofovir efficacy and nephrotoxicity.'
2. Young researcher's forum (affiliated with SASHG), Pretoria, September 2015. Presented a poster entitled 'Next generation sequencing to identify new variants and to establish the frequency of known variants in genes implicated in Tenofovir efficacy and nephrotoxicity.'
3. Molecular Biosciences Research Thrust Research Day (MBRT), December 2015). Presented a poster entitled 'Next generation sequencing to identify new variants and to establish the frequency of known variants in genes implicated in Tenofovir efficacy and nephrotoxicity.'
4. Pharmacogenetics and Precision Medicine Conference, University of Cape Town, South Africa, 4 – 9 April 2016. Presented a poster entitled 'Next generation sequencing to identify new variants and to establish the frequency of known variants in genes implicated in Tenofovir efficacy and nephrotoxicity.'
5. Wits Health Sciences Research Day & Postgraduate Expo, September 2016. Presented a poster entitled 'Next generation sequencing to identify variants and to establish the frequency of known variants in genes implicated in Tenofovir efficacy and nephrotoxicity'.
6. South African Society for Bioinformatics (SASBi) and the South African Genetics Society (SAGS) Conference, Nelson Mandela School of Medicine, South Africa, 20 – 23 September 2016. Presented a poster title 'Targeted ultra-deep sequencing of a Southern African cohort to identify and characterise common and novel variants in genes implicated in Tenofovir efficacy and nephrotoxicity.'
7. Soweto Matlosana Collaborating Centre for HIV/AIDS and TB (SoMCHAT) Emerging Researcher's Mini Conference, Chris Hani Baragwanath Academic Hospital (PHRU), 18 November 2016. Oral presentation entitled 'Genetic screening for biomarkers significant for Tenofovir efficacy and nephrotoxicity.'
8. Molecular Biosciences Research Thrust (MBRT), December 2016. Presented a poster entitled 'Next generation sequencing to identify variants and to establish the frequency of known variants in genes implicated in Tenofovir efficacy and nephrotoxicity'.

ABSTRACT

Recent South African studies report an increase in hospital admissions due to adverse drug reactions (ADRs), including to tenofovir disoproxil fumarate (TDF). Pharmacogenetic studies conducted in African populations have reported notable differences compared to other populations in variants associated with ADRs, including TDF-induced nephrotoxicity (clinically manifesting as acute kidney injury (AKI)), warranting their further investigation in local populations.

The aim of this study was to identify variants in eight TDF implicated ADME genes and to select a subset of SNPs for a genetic association study for TDF-associated nephrotoxicity.

Using next generation sequencing (NGS) to screen a cohort of 40 HIV-infected patients of Bantu-speaking origin, 275 variants were identified, including three loss of function and 40 novel variants. Based on allele frequency information (MAF>0.1) and prior association with ADRs, nine polymorphisms within five genes were prioritised for genotyping. This was done in cases presenting with AKI following TDF treatment (n=53) and in controls of similar ethnic origin not presenting with AKI (n=84). The cohort was genotyped using TaqMan® assays. A case-control association study was performed with genotypes, alleles and haplotypes using χ^2 tests. Suggestive associations were detected for the *ABCC2* 1249A allele (p=0.02) and *ABCC2* haplotypes AAC and AAT (p=0.02 and 0.03 respectively) and TDF induced AKI. The *ABCC2* GTT haplotype (p=0.02) appeared to be protective of TDF induced AKI. However, the observed associations were not deemed to be significant following corrections for multiple testing.

A limitation of the study was its relatively low power because of the small sample size. Therefore, findings from this work will need to be replicated in more suitably powered cohorts. Despite this, we observed evidence that suggests the potential influence of a subset of *ABCC2* variants on TDF-induced AKI.

ACKNOWLEDGEMENTS

I would like to thank the following people, without whom the study would not have been possible.

- The nursing staff Ms Ntabiseng Masindi, Ms Katlego Motlhaoleng, Ms Modiehi Rakgokong and Ms Motshidisi Khumalo for patient recruiting, sample collection, patient data collecting and management.
- Mr Pattamukkil Abraham for managing the data collected for this cohort.
- Mr Turflos Netshilindi for extracting the DNA from the stavudine (d4T) cohort samples used for NGS sequencing.
- Thanks to Dr Natasha Beeton-Kempen for NGS training and data processing and Dr Phiyani Lebea of the CSIR, South Africa, for the use of the Ion PGM sequencer.
- I am thankful to Dr Inger Jonasson, Ms Susana Haggqvist and Dr Adam Ameur (National Genomics Infrastructure, SciLifeLab Department of Immunology, Genetics and Pathology, Uppsala University, Sweden) for sequencing performed on the Ion S5.
- I would like to thank Dr Ananyo Choudhury for giving me access to and information on the Bantu-speaking cohort, and assistance with the NGS data analysis.
- I am very grateful to all the patients who participated in the study without whom the study would not be possible.
- Funders
    - DST support grant (# V6YET50)
    - CSIR parliamentary grant (# V1YBT96)
- Many thanks to members of teams at the CSIR and SBIMB (especially Dr. Liesl Henry) for the advice, guidance, support and ongoing encouragement.
- To my loving family and friends, thanks for the continuous support, patience and encouragement.

Sincere thanks to my supervisors, Prof Ramsay and Dr. Mancama, for the guidance, encouragement, patience and support. I appreciate the opportunity given to work with experienced scientists of high calibre.

LIST OF FIGURES

Figure1.1. Global estimates of people living with HIV in 2015. 1

Figure 1.2. Genetic variations in pathways of drug absorption, disposition, metabolism and excretion (ADME) 11

Figure 1.3. Visual representation of the proximal tubules of the kidneys and the main transporter proteins involved in the absorption and excretion of TDF. 13

Figure 1.4: *APOL1* variants implicated in HIV associated nephropathy (HIVAN). 19

Figure 2.1. Summary of the Next Generation Sequencing (NGS) workflow that was implemented in this study. 28

Figure 2.2. Ion PGM chip sequencing capacities and performance. 33

Figure 2.3. Summary of the NGS data filtering strategy that was employed for sequencing data analysis. 34

Figure 2.4. Loading density plots for the Ion 318™ chips used to produce the NGS data 40

Figure 2.5. Next generation sequencing coverage 42

Figure 2.6. Loading density plots for the Ion 530™ chips used to produce the NGS data……...43

Figure 2.7. Venn diagram showing the total number of variants produced by Ion Torrent sequencers PGM™ and S5™ 44

Figure 2.8. Variant classification for data generated using Ion Torrent PGM™ sequencing. 46

Figure 2.9. Variant classification for data generated via Ion Torrent S5™ sequencing. 48

Figure 2.10: Summary of the filtering strategy that was employed for prioritising variants identified in the IonS5™ data. 49

Figure 2.11. Gene panel and filtering of high quality variants identified by the IonS5™. 50

Figure 2.13. Novel variants classification 55

Figure 2.14. Characterisation of 11 prioritised variants 58

Figure 3.1. Summary of patients enrolled in the study. 74

LIST OF TABLES

Table 1.1. Adverse events of ARVs.....5

Table 1.2. A summary of studies that have observed association between genetic variants and clinical indicators of TDF-associated kidney tubular dysfunction 15

Table 1.3: Pharmacogenes associated with ARV adverse drug reactions and disease progression in HIV positive patients 17

Table 1.4. Renal histology in HIV-infected patients in South Africa.....18

Table 2.1. Program run to amplify target DNA regions 29

Table 2.2. Program run to partially digest the primers 29

Table 2.3. Program run to perform the ligation reaction 29

Table 2.4. Primers sequences and their expected amplicon size. 36

Table 2.5. The PCR reaction set up used to amplify sequences with the NEB kit 37

Table 2.6. The PCR reaction set up used to amplify sequences using the iProof^{TM} kit 37

Table 2.7. Sequencing reaction set up for the BigDye® terminator v3.1 cycle sequencing kit 38

Table 2.8. Sequencing cycling conditions 38

Table 2.9. Classification of Loss of function variants 51

Table 2.11. Nine variants previously reported to be associated with TDF associated nephropathy .57

Table 3.1. List of the nine TaqMan® SNP genotyping assays used in this study 71

Table 3.2. TaqMan® SNP genotyping assay reaction setup.....71

Table 3.3. Primers sequences and their expected amplicon size 72

Table 3.4. Descriptive statistic of the cases vs the controls.....75

Table 3.5. TaqMan® SNP genotyping assay robustness and validation 76

Table 3.6. Genotype distribution and statistics for association analysis of the variants assessed in this study in relation to acute kidney inury in the Bantu-speaking cohort.....77

Table 3.7. Allele distribution and statistics for association analysis of the variants assessed in this study in relation to acute kidney injury in the Bantu-speaking cohort 78

Table 3.8. Haplotype frequencies and associations with TDF induced acute kidney injury in a Bantu-speaking cohort 80

TABLE OF CONTENTS

DECLARATION ..... **II**

PRESENTATIONS ARISING FROM THIS STUDY ..... **III**

ABSTRACT ..... **IV**

ACKNOWLEDGEMENTS ..... **V**

LIST OF FIGURES ..... **VI**

LIST OF TABLES ..... **VII**

TABLE OF CONTENTS ..... **IX**

ACRONYM AND SYMBOL LIST ..... **XI**

CHAPTER 1: INTRODUCTION **1**

1.1 HIV/AIDS PREVALENCE AND TREATMENT **1**

1.1.1. HIV/AIDS PREVALENCE AND MAGNITUDE 1
1.1.2. ARVS AND HAART IN SOUTH AFRICA 2

1.2 NTRI AND TDF EFFICACY **3**

1.3 TREATMENT FAILURE **4**

1.3.1 ADVERSE EVENTS AND NON-COMPLIANCE 4
a) Mitochondrial toxicity *5*
b) Lactic acidosis *6*
c) Renal toxicity *6*
- TDF-associated nephropathy: acute kidney injury 7
- Other factors of renal failure 9

1.4 THE ROLE OF PHARMACOGENETICS IN DRUG TREATMENT OUTCOME **10**

1.4.1 THE PHARMACOGENETICS OF ARV TREATMENT 11
a) The role of drug metabolising enzyme variants *12*
b) The role of drug transporter gene variants *13*
c) The role of other important gene variants *15*
1.4.2 *APOL1* VARIANTS AND KIDNEY DISEASE 17

1.5 PHARMACOGENETICS IN AFRICA **20**

1.6 PROBLEM STATEMENT **23**

1.7 JUSTIFICATION **24**

1.8 AIMS AND OBJECTIVES **25**

CHAPTER 2: NEXT GENERATION SEQUENCING OF ADME GENES AND VARIANT PRIORITIZATION **26**

2.1 METHODOLOGY **26**

2.1.2 STUDY PARTICIPANTS 26
3.1.2 TARGETED NEXT GENERATION SEQUENCING OF 40 HIV-POSITIVE BLACK SOUTH AFRICANS 26
2.1.3 NEXT GENERATION SEQUENCING 27
a) Ion AmpliSeq™ DNA Library preparation *28*
b) Template preparation *30*
c) Enrichment and quality control *31*
d) Data processing and filtering *33*

2.1.4 NGS SEQUENCING DATA VALIDATION 35
a) Primer design *36*
b) Polymerase chain reaction and amplicon purification *36*
c) Sequencing reaction set up *37*
2.1.5 CHARACTERIZATION OF IDENTIFIED VARIANTS 38

2.2: RESULTS 39

2.2.1 SEQUENCING CHIP PERFORMANCE 39
2.2.2 SEQUENCING DATA ANALYSIS 43
a) Ion PGM™ vs Ion S5™ *43*
2.2.3 VARIANTS FILTERING AND CHARACTERIZATION 48
a) Gene panel filtering *48*
b) Bioinformatic functional analysis *51*
c) Novel variant identification in four genes and their classification *54*
2.2.4 PANEL OF POLYMORPHISMS SELECTED FOR GENOTYPING THE TDF COHORT 55

2.3: DISCUSSION 61

2.3.1 NEXT GENERATION SEQUENCING OF A BANTU SPEAKING COHORT 61
a) Ion PGM™ vs Ion S5™ *62*
b) Variant discovery and characterization *63*
c) Prioritized panel of SNPs *64*
d) Study limitations *66*

CHAPTER 3: GENE VARIANT SCREENING FOR VARIANTS IMPLICATED IN TDF INDUCED AKI USING A CASE-CONTROL STUDY DESIGN 68

3.1 METHODOLOGY 68

3.1.1 STUDY PARTICIPANTS 68
3.1.2 GDNA ISOLATION AND GENOTYPING 69
a) Sample collection and genomic DNA isolation *69*
b) Genotyping TDF cohort samples *70*
- Real Time PCR and TaqMan® genotyping 70
- Genotyping validation: Sanger sequencing 71

3.1.3 STATISTICAL ANALYSIS 72

3.2 RESULTS 73

3.2.1 CHARACTERISTICS OF THE STUDY POPULATION 73
3.2.2 GENOTYPING RESULTS AND ASSOCIATION TESTING 75
a) Chi-square tests for genotype association *76*
b) Chi-square tests for allele association *78*
c) Chi-square tests for haplotype association *79*

3.3: DISCUSSION 80

3.3.1 ROBUSTNESS OF GENOTYPING ASSAYS 80
3.3.2 TDF COHORT DESCRIPTIVE STATISTICS 81
3.3.3 CHI-SQUARED ASSOCIATION TESTING 82
a) Genotype and allele association testing *82*
b) Haplotype association testing *84*
c) ABCC2 and TDF associated nephrotoxicity *85*
d) Study limitations *87*

CHAPTER 4: FUTURE PERSPECTIVES 88

CHAPTER 5: CONCLUSION 89

CHAPTER 6: REFERENCES **91**

APPENDIX A: ETHICS CLEARANCE **112**

APPENDIX B: LIST OF 65 GENES IN NEXT GENERATION SEQUENCING (NGS) PANEL USED FOR THE AMPLISEQ™ PRIMER PANEL DESIGN **113**

APPENDIX C: CLINICAL RESPONSE FORM (CRF) **114**

APPENDIX D: ARTICLE **117**

APPENDIX E: NGS SEQUENCING VALIDATION OF D4T COHORT **118**

APPENDIX F: GDNA EXTRACTION & QUALITY CONTROL OF D4T COHORT SAMPLES: **122**

APPENDIX G: EVALUATING THE SALTING OUT METHOD FOR GDNA EXTRACTION **123**

APPENDIX H: ASSAY OPTIMIZATION **125**

APPENDIX I: TAQMAN® SNP GENOTYPING ASSAY ROBUSTNESS AND VALIDATION **126**

1. *ABCB1* RS2032582 126
2. *ABCC10* RS2125739 128
3. *ABCC2* RS2273697 129
4. *ABCC2* RS17222723 130
5. *ABCC2* RS3740066 131
6. *ABCC4* RS3742106 132
7. *ABCC4* RS1751034 133
8. *APOL1* RS73885319 134
9. *APOL1* RS60910145 135

APPENDIX J: TDF GENOTYPING DATABSE **136**

APPENDIX K: ASSESSMENT OF THE PERFORMANCE OF THE PRIMERS USED FOR PCR PRIOR TO SANGER SEQUENCING TO VALIDATE THE GENOTYPING CALLS **140**

APPENDIX L: ASSESSMENT AND VALIDATION OF TAQMAN SNP GENOTYPING ASSAY **142**

APPENDIX M: CONFIRMATION OF AMBIGUOUS GENOTYPIC CALLS **143**

APPENDIX N: TENOFOVIR/ ADEFOVIR PATHWAY, PHARMACOKINETICS **146**

ACRONYM AND SYMBOL LIST

°C: Degrees Celsius

mg: milligram

µg: microgram

ml: milliliter

µl: microliter

µM: micromolar

ng: nanogram

3TC: Lamivudine

ABCB1: ATP-binding cassette subfamily B member 1

ABC: Abacavir

ABCC: ATP-binding cassette subfamily C

ADRs: Adverse Drug Reactions

AIDS: Acquired Immune Deficiency Syndrome

AKI: Acute kidney injury

APOL1: Apolipoprotein L1

ARV: Antiretroviral

ATP: Deoxyadenosine Triphosphate

AZT: Azidothymidine, also called Zidovudine,

CD4: Cluster of Differentiation 4

DNA: Deoxyribonucleic Acid

DLV: Delavirdine

d4T: Stavudine

DDC: Zalcitabine

DDI: Didanosine

dNTPs: Deoxynucleotide triphosphates

EFV: Efavirenz

ExAC: Exome Aggregation Consortium

FTC: Emtricitabine

ETR: Etravirine

HIV: Human immunodeficiency virus

HIVAN: Human immunodeficiency virus associated nephropathy

KGP: The 1000 Genomes Project

LPV/r: Lopinavir/ritonavir

NGS: Next Generation Sequencing

NVP: Nevirapine

PCR: Polymerase Chain Reaction

SA: South Africa

SCL22A: Solute carrier family 22 member

SNP: Single Nucleotide Polymorphism

TDF: Tenofovir Disproxil Fumarate

UNAIDS: Joint United Nations Programme on HIV/AIDS

UNICEF: United Nations Children's Fund

VEP: Variant Effect Predictor

WHO: World Health Organisation

CHAPTER 1: INTRODUCTION

1.1 HIV/AIDS prevalence and treatment

1.1.1. HIV/AIDS prevalence and magnitude

Since the discovery of the human immunodeficiency virus (HIV) three decades ago (Barré-Sinoussi *et al.* 1983, Montagnier *et al.* 1984, Gallo 1988), the scale of the HIV/AIDS (acquired immunodeficiency syndrome) epidemic continues to be a global health issue, with an estimated 36.7 million people living with HIV in 2015 (UNAIDS 2016a). The epidemic continues to be most severe in Southern Africa (Figure 1.1), with South Africa having the fourth highest prevalence rate (19.2% of adults) and an estimated 7 million people living with HIV (UNAIDS 2016b).

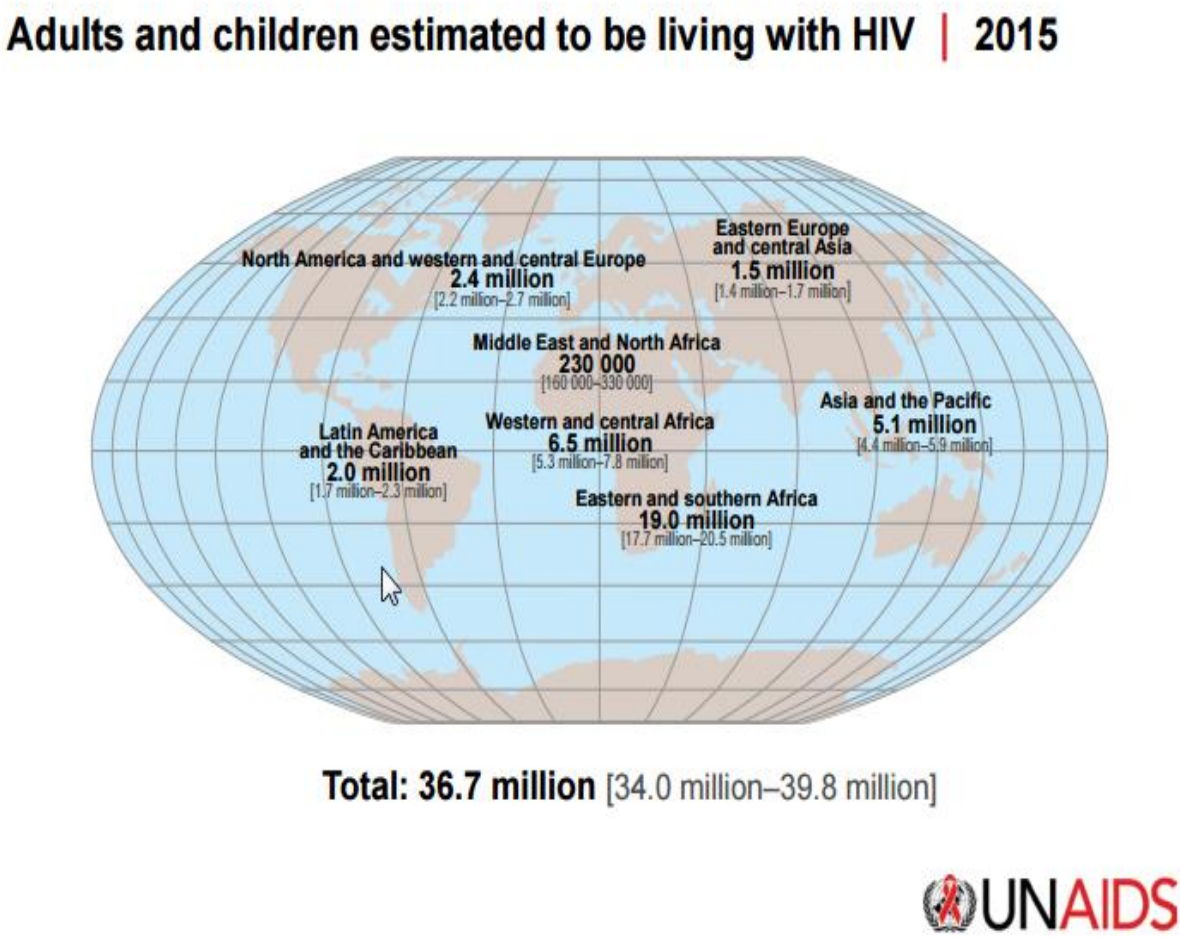


Figure1.1. Global estimates of people living with HIV in 2015. The estimates are listed in black and the ranges for the estimates are listed in grey. Source: (UNAIDS 2016a).

The global incidence and transmission rates have, however, declined; the UNAIDS report of 2014 reported a 35% decline in the number of new HIV infections since 2000, a 42% decline in AIDS-related deaths since the 2004 peak and, most importantly, a 58% decline in the number of new HIV infections among children since 2000 (UNAIDS 2016a). This progress is largely due to effective treatment and awareness campaigns (WHO *et al.* 2013) where globally there has been an 84% increase in access to anti-retroviral (ARVs) since 2010 (UNAIDS 2016a).

1.1.2. ARVs and HAART in South Africa

Effects of highly active antiretroviral therapy (HAART) include reduced viral load, increased CD4 and CD8 T-cell count and subsequent prevention of new opportunistic infections. In addition, improved clinical symptoms such as weight gain and general well-being, increased survival, improved quality of life, and a reduction of HIV transmission are seen (WHO *et al.* 2013). Although HAART has significantly improved patient prognosis, some drugs are not available in developing countries.

South Africa (SA) has the largest ARV program in the world with 3.5 million people receiving treatment in 2016 (Reid 2016). To limit the emergence of drug-resistant strains, an effective ARV treatment program is key. In SA, the initial regimen was predominantly stavudine (2′,3′-didehydro-2′,3′-dideoxythymidine (d4T)) based but because of the emergence of adverse events such as neuropsychiatric side-effects, lactic acidosis, central nervous system (CNS) side effects and lipodystrophy (Carr *et al.* 1998, Tsiodras *et al.* 2000, Andrew 2003, Carr 2003, Geddes *et al.* 2006, Bolhaar and Karstaedt 2007, Andrew *et al.* 2016), the South African guidelines have recently been updated (DoHSA 2013) and changed based on documented literature and reports about patient responses and side-effects. With this paradigm shift in HIV infection treatment, from d4T to TDF, the implication of this is the possible emergence of various combinations of strains that are prevalent in the South African population. This is due to the possible emergence and subsequent transmission of mutant and or drug resistant viruses. Most reported cases or transmitted resistant viruses involve nucleoside RT inhibitors (NRTIs) and non-nucleoside RT inhibitors (nNRTIs) (Michaud *et al.* 2012). With the evolution of viral quasi-species, there is a continuous need for new effective drugs targeting drug-resistant HIV strains. Switching from a d4T based regimen to TDF was a cost-effective strategy as demonstrated by Ronsen and colleagues (Rosen *et al.*

2008). After a cost-effectiveness analysis, they inferred that due to the cost of complications associated with d4T toxicities, it would cost less to switch to TDF.

1.2 NTRI and TDF efficacy

Effective treatment has significantly improved the prognosis of people with HIV and transmission rates have significantly reduced in pregnant women and has subsequently led to a decline in vertical transmission (Baroncelli *et al.* 2015).

NRTIs and NtRTIs form the backbone of ARV regimens, including in SA. NRTIs mimic the structure of natural deoxynucleotides (dNTPs) and their mechanism of action involves them competing with the natural dNTPs. They are effective as reverse transcriptase (RT) inhibitors as they have a higher affinity to the RT enzyme which integrates them into the growing strand during DNA replication, resulting in chain termination. NtRTIs' mechanism of action is similar to that of NRTIs as they compete with the host's natural deoxyadenosine triphosphate (dATP), thus successfully preventing viral DNA replication (De Clercq 1998, Maga *et al.* 2010). The intracellular persistence and long intracellular half-life of NRTIs and NtRTIs has contributed to their long-term clinical efficacy (Arribas *et al.* 2008, Cihlar and Ray 2010). TDF is an NtRTI and is widely used because of its efficiency in viral load suppression, subsequent reconstitution of the immune system (Gallant *et al.* 2006) especially in treatment-naive patients (Arribas *et al.* 2008), and renal safety after long term usage (144 weeks) (Izzedine *et al.* 2005a). To reduce vertical transmission, guidelines for prevention of mother to child transmission (PMTCT) recommended a triple drug ARV regimen (TDF, 3TC/FTC and EFV) to be initiated early during pregnancy (Townsend *et al.* 2008, Mofenson 2010). A South African study reported that a triple drug ARV regimen containing TDF is preferred for effective and rapid administration of ARV initiation during pregnancy, as renal dysfunction is significantly less common in pregnant women than in other adults eligible for treatment (Townsend *et al.* 2008, Myer *et al.* 2013, WHO *et al.* 2013). TDF/FTC combination is preferred because of its potent and durable clinical efficacy, long-term safety and tolerance demonstrated in long-term clinical studies (Gallant *et al.* 2004, Cassetti *et al.* 2007, Arribas *et al.* 2008, Smith *et al.* 2009).

TDF fixed-dose combination therapy is also preferred in breastfeeding women and in children between 2 and 10 years old by the European Medicines Agency and US food and drug administration (FDA) (Giacomet *et al.* 2005, WHO *et al.* 2013, Lim *et al.* 2015). TDF

has been an effective drug for salvage therapy in paediatric patients (Hazra *et al.* 2005, Viganò *et al.* 2011). TDF and TDF-FDC based pre-exposure prophylaxis (PrEP) regime have been shown to be effective in reducing HIV transmission and preventing HIV infection (Baeten *et al.* 2012), even in high-risk individuals (Okwundu *et al.* 2012). A study conducted in SA evaluated and confirmed the safety and effectiveness of a TDF-based microbicide gel (Karim *et al.* 2011), especially in women who adhered to treatment (Karim *et al.* 2010). However some studies reported that adherence is low among African women (Marrazzo *et al.* 2015) and some women experienced side effects (Van Damme *et al.* 2012). These findings highlight the potency of TDF and benefit in HIV/AIDS treatment and prevention in SA, specifically in antiretroviral pre-exposure prophylaxis and PMTCT.

1.3 Treatment failure

Treatment failure typically occurs due to the emergence of drug resistance, non-compliance, or impaired drug metabolism (Knobler *et al.* 2003). Drug resistant strains are common and occur as a result of mutations associated with the viral genome. There are a number of factors influencing non-compliance or adherence to ARV treatment. A significant factor is adverse drug reactions that are discussed below. Some factors are dependent on the patients' commitment. These issues include social problems, poverty and food insecurity, substance abuse, denial, inadequate treatment literacy, work-related issues, pill burden, depression and mental illness, dosage requirements, and inexperience and conflict of opinion between health care providers and alternative care providers. At the same time, pharmacokinetic and pharmacodynamic changes due to drug interactions, and the use of fertility interventions (such as ceasing treatment to start a family) for example, rank highly amongst other concerns (Conradie 2012). Impaired drug metabolism is influenced by a number of factors including genetic variations, discussed in section 1.4.

1.3.1 Adverse events and non-compliance

One major factor that influences patient adherence to ARV treatment is the occurrence of adverse events such as drug induced toxicity, contraindications and metabolic changes. Like most drugs, there are side-effects associated with taking ARVs. The mechanism of action of each drug and its pharmacokinetics influences the kind of adverse reaction that is associated with each drug. The pharmacodynamics of each drug also influences drug interactions,

especially for HAART (Rathbun and Liedtke 2011, Maarseveen *et al.* 2012, Margolis *et al.* 2013). Common side-effects associated with the use of ARVs include nausea, insomnia, and irritable bowel, with some being so severe that they jeopardise a patient's health and wellbeing or end up being fatal (Monforte *et al.* 1998, McIlleron *et al.* 2007). Complications include: renal toxicity, mitochondrial toxicity, lipoatrophy, hyperlactatemia and lactic acidosis, metabolic disorders, myopathy, neuropathy, neuropsychiatric disorders, hepatic stenosis, raised cholesterol, cardiovascular events, gastrointestinal intolerance, pancreatitis, insulin resistance, allergic reactions and hypersensitivity reaction syndrome (HRS) (Beach 1998, Tozzi *et al.* 2008, Tozzi 2010). Key adverse events commonly associated with the use of ARVs are described in more detail below and are summarised in Table 1.1.

Table 1.1. Adverse events of ARVs

CLASS	DRUG	ADVERSE EFFECTS
NRTIs	Zidovudine	Anemia, nausea, rash
	Stavudine and didanosine	Nausea, lipoatrophy, dyslipidemia, pancreatitis, lactic acidosis, hepatic steatosis, heart disease, distal sensory peripheral neuropathy (DSPN)
	Abacavir	Hypersensitivity reaction syndrome (HSR), hepatotoxicity, heart disease
	Tenofovir	Renal insufficiency, bone loss
nNRTIs	Efavirenz	CNS adverse effects, rash, hepatotoxicity, lipoatrophy, teratogenicity, hypertriglyceridemia
	Nevirapine	Rash, HSR, hepatotoxicity
	Etravirine	Rash, hepatotoxicity
PIs	All PIs	Nausea, diarrhoea, rash, dyslipidemia, insulin resistance, hepatotoxicity
	Atazanavir	Jaundice, scleral icterus,
	Indinavir	Jaundice, scleral icterus, nephrolithiasis
	Lopinavir fosamprenavir	Heart disease
Entry inhibitors	Enfuvirtide	Injection site reactions, pneumonia, HSR
	Maraviroc	Cough, fever, respiratory tract
Integrase inhibitors	Raltegravir	Headache, insomnia, dizziness
Source: Adapted from Hawkins et al, 2010		

a) Mitochondrial toxicity

Mitochondrial toxicity is associated with NRTIs. This is as a result of NTRIs inhibiting mitochondrial DNA polymerase (POLδ) (Martin *et al.* 1994, Lim *et al.* 2003) responsible for the replication of mitochondrial DNA (mtDNA). Inhibiting mtDNA may result in the production of structurally and functionally different proteins leading to adverse side-effects.

Mitochondrial toxicities are regarded as the most significant and manifest in different ways which include myopathy, neuropathy, lipoatrophy and lactic acidosis (reviewed in (Margolis *et al.* 2013).

b) Lactic acidosis

Lactic acidosis occurs as a result of accumulation of lactate due to mitochondrial toxicity. Mitochondrial toxicity results in insufficient oxidative phosphorylation and subsequently reduced oxidation of lactate. The severity of accumulation of lactate ranges from hyperlactaemia without acidosis to severe lactic acidemia (Margolis *et al.* 2013). Lactic acidosis is a serious adverse event associated with NRTIs, such as d4T (Coghlan *et al.* 2001, Falcó *et al.* 2002, Falcó *et al.* 2003) and the incidence rates of lactic acidosis in South Africa has increased (Geddes *et al.* 2006, Bolhaar and Karstaedt 2007). Risk factors for lactic acidosis include: being female, pregnancy, obesity, older age group, and lower CD4 count (AIDSinfo 2016). Gynaecomastia (enlargement of male breasts) and lipomastia (enlargement of adipose tissue causing enlargement of breast) has been reported to occur in both boys and girls especially during adolescence (DoHSA, 2014). Several reports of differences in the sex prevalence in NRTI-induced lactic acidosis have been documented, even in children (Bolhaar and Karstaedt 2007, Currier 2007). It is associated with obesity and the aetiology is postulated to be dose-dependent, sex-linked (mitochondrial genetic variation) and more prevalent in obese women (Currier 2007).

c) Renal toxicity

Mitochondrial toxicity is postulated to be the mechanism by which TDF-associated renal toxicity occurs (Hall. *et al.* 2011). Renal toxicity has been reported to be a significant problem in patients on TDF (including PIs such as indinavir and atazanavir) (Andiman *et al.* 2009, Bickel *et al.* 2013, Yombi *et al.* 2014) and in clinical cases it presents as acute kidney disease, chronic kidney disease (end-stage kidney disease (ESKD)) and Fanconi syndrome (FS) (proximal renal tubular dysfunction) (Kinai and Hanabusa 2005, Han *et al.* 2006, Kinai and Hanabusa 2009, Brennan *et al.* 2011, Scherzer *et al.* 2012). Renal function has also generally been shown to decline over time in children and adolescents on antiretroviral therapy (ART) (Pontrelli *et al.* 2012). TDF is also associated with oesteomalacia, which is caused by phosphate wasting and calcitriol (bioactive vitamin D) deficiency (Woodward *et*

al. 2009, Lucey *et al.* 2013). In addition, the risk of bone toxicity is higher in children (Gafni *et al.* 2006). TDF is, however, still used in the current regimen because its effectiveness in the management of HIV (viral load and transmission) outweighs the potential impact it has after long-term exposure. Some studies have suggested that TDF can be safely administered without routine monitoring of renal function (Nelson *et al.* 2007, Stöhr *et al.* 2011).

- TDF-associated nephropathy: acute kidney injury

The main target of TDF nephrotoxicity is the proximal tubules. In severe cases, TDF leads to a breakdown of solute transport in this nephron segment. This event of tubular dysfunction is classified as renal FS or acute kidney injury (AKI) (Verhelst *et al.* 2002). However, TDF-induced AKI occurs with or without proximal tubulopathy (Zaidan *et al.* 2013). TDF is also associated with other renal dysfunctions such as TDF-related FS with nephrogenic diabetes insipidus and ESKD (Verhelst *et al.* 2002, Karras *et al.* 2003, Rollot *et al.* 2003, Schaaf *et al.* 2003). Several studies have identified lesions in kidney biopsies from patients taking TDF (Créput *et al.* 2003, Bickel *et al.* 2013). FS is characterised by urine wasting of solutes normally exclusively reabsorbed in the proximal tubules (PT) (i.e., phosphate, bicarbonate, glucose, low molecular weight proteins, and amino acids) (Verhelst *et al.* 2002). Phosphate loss is clinically important as osteomalacia, bone pain and fractures can result. This is classified as TDF-induced bone toxicity (Fux *et al.* 2008).

AKI is reversible and common, with a global incidence rate of 21.6% in adults (Susantitaphong *et al.* 2013). Studies have shown that AKI associated with TDF can be reversed once treatment has been switched (Patel *et al.* 2013), therefore monitoring of patients on TDF is essential. The Kidney Disease Improving Clinical Outcomes (KDIGO) criteria are the gold standard method for monitoring renal impairment or injury. KDIGO is a global organisation that develops and implements kidney disease guidelines through evidence-based clinical practice and research. They define acute kidney injury as any of the following: an increase in serum creatinine by more than 1.5 times the baseline, creatinine increases >26.4umol/l above the upper limit of normal, urine output <0.5ml/kg/hr for 6 hrs, and reductions in eGFR, which is known or presumed to have occurred within the 7 days prior. Their definition of chronic kidney disease (CKD) is defined as "abnormalities of kidney structure or function, present for >3 months, with implications for health". Since TDF-associated nephrotoxicity occurs within 3 months (Roe *et al.* 2008) it can be classified

as AKI. Severe renal dysfunction has been reported by some studies in certain populations in Africa. In one such study, significant increases in eGFR were observed after starting ART initiation, where 74% of the patients were on TDF combination therapy (Reid *et al.* 2008).

A 2005 South African study indicated that the HIV infected patients were two times as likely to develop adverse drug reactions (ADRs), where 6.3% of hospital admissions were due to ADR (and an additional 6.3% of inpatients developed ADR during their stay) where ART was implicated in hospital admissions. Drugs used for opportunistic infections and antibiotics were implicated in 67% of hospital-acquired ADRs (Mehta *et al.* 2008). Since the introduction of TDF-based therapy in 2010, and the subsequent increase in people on ARVs, the incidence of AKI among HIV patients is notably high in South Africa. A follow-up study from the same group indicated 2.9% of hospital deaths were associated with TDF, rifampicin or co-trimoxazole (an antibiotic) (Mouton *et al.* 2015). The National HIV & Tuberculosis Health Care Worker (HCW) Hotline ADRs report indicated that kidney injury was among the three most common ADRs (including rashes and drug induced liver injury (DILI)), with EFV (rashes), anti-tuberculosis drugs (DILI) and TDF (kidney injury) being implicated (Njuguna *et al.* 2016). Similar findings were described by a study in four South African hospitals where 8.33% of admissions were ADR related and renal impairment, hypoglycaemia, liver injury, and haemorrhage were most common with TDF, insulin, rifampicin, and warfarin being implicated, respectively (Mouton *et al.* 2016). TDF also has an added effect on AKI patients causing a more rapid and worsening renal function, higher proportion of proteinuria and acidosis and delayed renal recovery (Seedat *et al.* 2016). These findings reflect our nation's burden of HIV/AIDS and tuberculosis (TB) co-infection and show an increase in the incidence of ADR. The South African guidelines recently shifted to initiating therapy immediately after diagnosis (Reid 2016). In their efforts to adopt the WHO guidelines to reduce HIV transmission, the incidence of ADR is expected to increase proportionately.

TDF is associated with multiple drug interactions, which may lead to increased risk of renal failure (Zimmermann *et al.* 2006). The role of drugs such as ritonavir, abacavir (ABC), acyclovir**,** amprenavir, didanosine (ddl), lopinavir- ritonavir, indinavir and atazanavir (ATV) are also associated with tubulopathies in particular, has been documented to play a significant role in the pathogenesis of TDF-induced kidney damage (Izzedine *et al.* 2005c, Hussain *et al.* 2006, Piketty *et al.* 2006, Hawkins and Ball 2007, Goicoechea *et al.* 2008, Kiser *et al.*

2008b). Therefore, careful planning and monitoring of patients on such regimens has to be performed regularly as TDF-associated toxicities are likely to occur. Switching from TDF to another ARV in patients with TDF-associated renal failure has been shown to reverse the symptoms and improve renal function within two to three months (Giacomet *et al.* 2013). Weight is an important criterion to consider before administering TDF, especially in children (Lim *et al.* 2015) and pregnant women weighing more than 90kg (Best *et al.* 2015). Recent studies suggested that the incidence of AKI is notably high in South Africa, although this is preventable to some degree (Mouton *et al.* 2015).

- Other factors of renal failure

The prevalence of kidney disease in patients with HIV is 6% in South Africa (Fabian and Naicker 2009), with 10% of the patients expected to experience renal toxicity or renal failure during the course of the disease (Bihl 2003, Franceschini et al. 2005). Several studies have documented HIV-associated nephropathy (HIVAN) (Ray *et al.* 1998, Schwartz and Klotman 1998, Han *et al.* 2006), which is a disease that is directly caused by HIV infection (Klotman and Klotman 1998). This is the most common source of renal failure associated with HIV infection and is prevalent in kidney biopsies of HIV-infected patients of African ancestry (Kimmel *et al.* 1998, Klotman and Klotman 1998, Kimmel 2003, Kimmel *et al.* 2003, Fabian *et al.* 2013), suggesting a genetic predisposition to acquiring HIVAN. This has been confirmed by the APOL1 studies by Kasembeli and colleagues (Kasembeli *et al.* 2015a). ESKD is 18 times more likely to occur in African-Americans than in Caucasian Americans and HIVAN is the third leading cause of ESKD in African-Americans (Lucas *et al.* 2008). Other factors that may influence renal failure include diabetes mellitus (diabetic nephropathy), hypertension and glomerular disease which are common in Africa. Alternative aetiologies of renal failure include hypertension, infections, dysimmune disorders, malignancies and nephropathies of undetermined origin (Naicker and Fabian 2010, Naicker 2013).

a) Other factors that influence ARV treatment outcome

Factors such as race, sex, and ethnicity often play a key role in dosage, ARV susceptibility, tolerance and effective HIV management (Kimmel *et al.* 1998, Klotman

and Klotman 1998, Currier 2007). Data furthermore infers the influence of genetic factors, as described in detail below, on increased risk of ARV-associated adverse events. South Africa's ARV regimen is predominantly reverse transcriptase inhibitor-based (i.e. previously d4T based and currently TDF-based) thus adverse events such as renal toxicity, lipodystrophy, hepatotoxicity and HRS (most especially in children under five) are of concern. With a country that has relatively high incidences of diseases such as HBV, TB and cancer, some associated with the high prevalence of HIV, co-infection becomes harder to manage. This is specifically when ARV toxicity occurs in patients susceptible to these diseases (i.e. not prescribing nNRTIs in patients with HBV infection or EFV and other TB treatment drugs) (McIlleron *et al.* 2007). Another factor to consider when starting ARV treatment and management is access to consistent ARV medication, as it affects therapeutic outcome (Fatti *et al.* 2010). Pre-existing renal failure conditions may be exacerbated by TDF administration (Brennan *et al.* 2011).

1.4 The role of pharmacogenetics in drug treatment outcome

Personalised medicine or precision medicine includes genomic medicine, where genomic information is used to make more informed decisions. Pharmacogenetics is the study of genetic differences or variations in the metabolism of drugs, which can affect individual responses to drug (Roden *et al.* 2002, Kalra *et al.* 2009). Pharmacogenetics involves analysing the genetic basis for the inter-individual variation in the body's drug disposition and patient clinical outcome to drug therapy such as drug efficiency and toxicity (Mini and Nobili 2009). Disposition is defined as the fate of a drug or xenobiotic in relation to its absorption, distribution, metabolism, and excretion (Caldwell *et al.* 1995, Ma and Lu 2011). These genetic variations potentially predispose a particular individual to having a specific therapeutic effect or adverse effect/s to a drug. Therefore, it is the understanding of how the genetic profile affects an individual's response to drugs (Ma and Lu 2011). The inter-individual responses to drugs are a complex, multifactorial process which often makes it difficult to predict the functional impact of genetic variation. Single-nucleotide polymorphisms (SNPs) are defined as single nucleotide changes or variations that occur in human DNA occurring at an allele frequency greater than 1% (Hoehe *et al.* 2003). Nucleotide changes occurring with a frequency lower than this are referred to as rare and novel variants. These polymorphisms ultimately drive the evolution of a particular disease and the subsequent response to treatment.

Understanding pharmacokinetics and pharmacodynamics will assist in understanding the influences of these polymorphisms, as drugs interact with numerous targets such as transport proteins, carrier proteins, receptors and metabolising enzymes. The purpose of pharmacogenomics or pharmacogenetics is to ascertain the role of sequence variation in the human genome that influences drug disposition, metabolism, efficacy and toxicity. Most pharmacogenetic studies have focused on genes coding for proteins involved in drug metabolism, immune modulation factors, drug transportation proteins, lipid binding proteins, cytokines, chemokine receptors and proteins involved in glucuronidation (reviewed in (Rodríguez-Nóvoa *et al.* 2005, Rodríguez-Nóvoa *et al.* 2006, Tozzi 2010, Rathbun and Liedtke 2011).

1.4.1 The pharmacogenetics of ARV treatment

Genetic variations in pathways of drug absorption, disposition, metabolism and excretion (ADME) are the main targets of HIV pharmacogenetic studies, which seek to explain the inter-patient variability observed in patients treated with drugs such as ARVs (Figure 1.2).

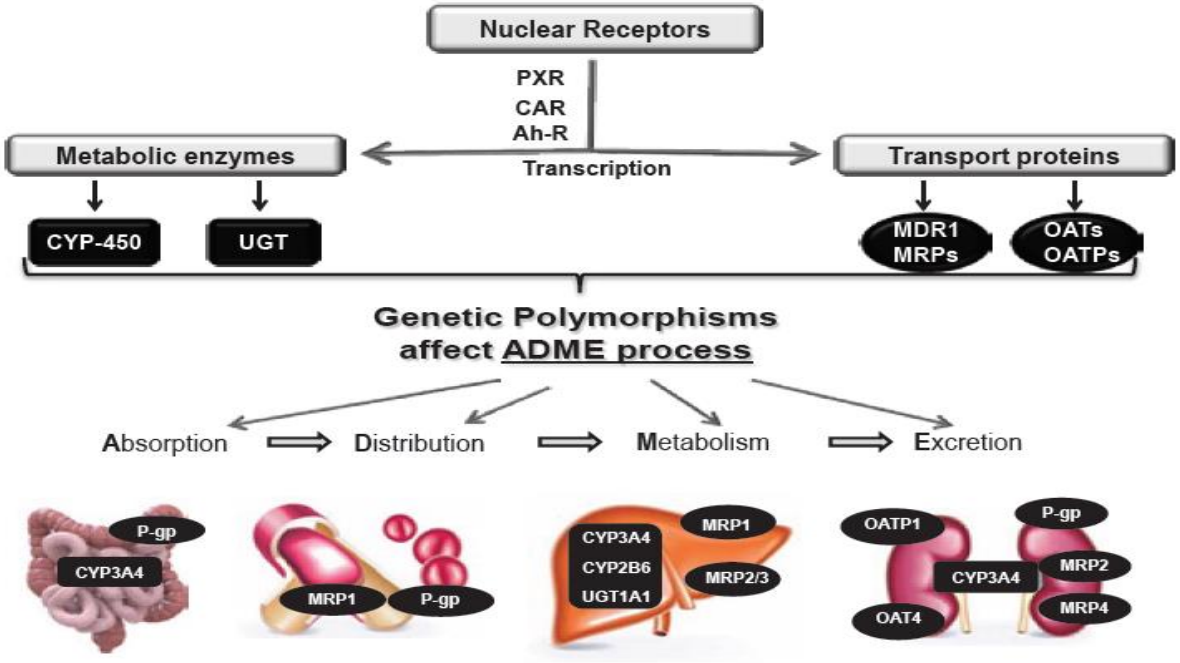


Figure 1.2. Genetic variations in pathways of drug absorption, disposition, metabolism and excretion (ADME) are main targets of HIV pharmacogenetic studies used to explain the inter-patient variability observed in patients treated with drugs such ARVs. Nuclear Receptors (NR) regulate metabolic enzymes and transport proteins and polymorphisms in these genes may influence patient tolerance or susceptibility to ARVs. Abbreviations: Ah-R: Aryl hydrocarbon receptor; PXR: pregnane X receptor; CAR: constitutive androstane receptor; CYP-450: cytochrome-450, UGT: uridine diphosphate glucuronosyl transferase; P-gp: P-glycoprotein; MRPs: multidrug resistance proteins, OATs: organic anion transporters; OATPs: Organic anion transporter polypeptides. Source: (Barco and Nóvoa 2013).

a) The role of drug metabolising enzyme variants

Several known genes have been identified to be largely responsible for variances in drug metabolism and response. The most common are the cytochrome P450 (*CYP*) genes, which are a superfamily of heme-binding proteins. They are responsible for the synthesis and metabolism of various chemicals and molecules (Nebert and Russell 2002). Mutations or variations in the genetic structure of CYP genes can affect the function of these enzymes; the effects are documented or seen in the metabolism of drugs.

CYP enzymes are primarily found in liver cells but are also located in cells throughout the body. The metabolism of ARVs commonly involves CYP3A4 and CYP2B6 and their re-absorption and distribution occurs through P-glycoprotein (P-pg) transport (Fellay *et al.* 2002, Marzolini *et al.* 2004a, Marzolini *et al.* 2004b). However, NtRTIs such as TDF are renally excreted (Rodríguez-Nóvoa *et al.* 2005, Rodríguez-Nóvoa *et al.* 2006). nNRTIs such as nevirapine (NVP), efavirenz (EFV) and delavirdine (DLV) are metabolized mainly by CYP3A4, CYP2D6, CYP2B6 and glucuronidation (Erickson *et al.* 1999, Dickinson *et al.* 2010). Protease inhibitors such as nelfinavir are metabolized by CYP2C19, CYP3A4, and CYP2D6 (Hirani *et al.* 2004, Dickinson *et al.* 2010). Other drugs such as maraviroc and raltegravir are metabolised by CYP3A4 and glucuronidation (Kassahun *et al.* 2007, Abel *et al.* 2008, Dickinson *et al.* 2010).

There is a strong association between *CYP2B6* genetic variations and adverse events evident in EFV-treated patients (Rodriguez-Novoa *et al.* 2005, Rotger *et al.* 2005, Wang *et al.* 2006, Saitoh *et al.* 2007a, Haas *et al.* 2009, Aurpibul *et al.* 2012, Masebe *et al.* 2012). *CYP2B6* 516 T/T genotype is correlated with both greater EFV plasma exposure and increased CNS toxicity (Haas *et al.* 2004); whereas the *CYP2B6* 516 G/T genotype and the *CYP2B6* 516 T/T genotype is associated with increased clearance and thus decreased exposure, with the G/G genotype having the least clearance (Rotger *et al.* 2005). This polymorphism has also been reported in the South African black population (MAF=0.23) and is correlated to the emergence of EFV-associated neurotoxicity where individuals who are homozygous for the T/T genotype showed severe neuropsychiatric adverse events (Gounden *et al.* 2010). This polymorphism is also associated with increased exposure to NVP (Rotger *et al.* 2007). Some additional *CYP2B6* variations associated with EFV-associated CNS toxicity include the 983 T/C, 415 A/G and the 785 A/G genotypes (Nolan *et al.* 2006, Wang *et al.* 2006, Swart *et al.* 2013). *CYP2B6* polymorphisms are also associated with NVP- associated hepatotoxicity such as altered hepatic function and altered plasma concentrations of NVP (Spence *et al.* 1995,

Rotger *et al.* 2005, Haas *et al.* 2006, Penzak *et al.* 2007, Saitoh *et al.* 2007b, Frasco *et al.* 2012).

b) The role of drug transporter gene variants

ARVs are substrates for transport proteins, with the most relevant and the most significant to ARVs transport including p-glycoprotein (p-gp1), multi-drug resistance proteins (MRPs), and the organic anionic transporter (OAT). These transporters carry molecules across the cell membrane. The OAT protein family is involved in the active transport of drugs in the proximal tubules of the kidneys during elimination. They are expressed in the basolateral membrane of the proximal tubules, while the MRPs are apically expressed. The drugs are actively taken up by OAT proteins and secreted into the lumen of the tubule via apically located MRP proteins (Miller 2001, Cihlar et al. 2001, Motohashi et al. 2002, Ray et al. 2006, Uwai et al. 2007) (Figure1.3). MRP proteins, which are members of the ATP-binding cassette superfamily, are scientifically proven to regulate intracellular NRTI concentrations (Eilers *et al.* 2008).

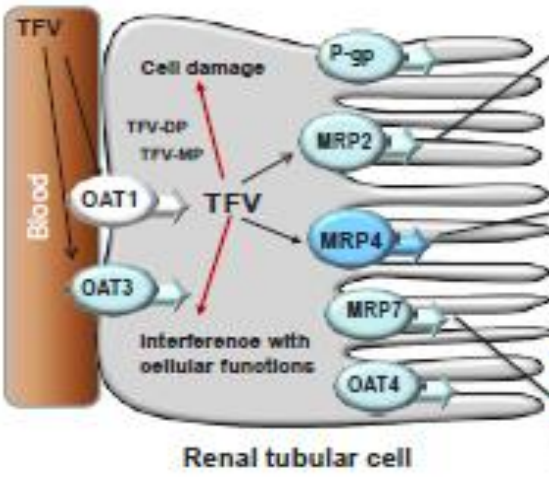


Figure 1.3. Visual representation of the proximal tubules of the kidneys and the main transporter proteins involved in the absorption and excretion of TDF. Abbreviations: MRP2, multidrug-resistance protein 2; MRP4, multidrug-resistance protein 4; OAT1, organic anion transporter 1; OAT4, organic anion transporter 4; P-gp, P-glycoprotein. Source: (Barco 2013, Barco and Nóvoa 2013).

The most significant transporter of drugs is the MRP family of proteins which are associated with adverse events linked to protease inhibitors and NTRIs such as TDF and EFV (Fellay *et al.* 2002). P-gp proteins facilitate the transportation of a variety of molecules including sugars, fats, amino acids and drugs (Marzolini *et al.* 2004a). P-gps are widely expressed in somatic cells including hepatocytes, renal proximal tubular cells, small intestinal cells, lymphoid cells and the capillary endothelial cells. Together with MRP2 and MPR4, they play important roles in regulating intracellular concentrations of NRTIs.

Because MRPs are energy-dependent pumps that efflux their substrates into the glomerular filtrate and nucleotide analogues are known to cause mitochondrial toxicity, concentration-dependent toxicity is postulated to be the main cause of renal dysfunction. With regard to TDF-associated renal toxicity, SNPs often evaluated include those in the ATP-binding cassette subfamily C members 2, 4 and 10 *(ABCC2, ABCC4* and *ABCC10*), ATP-binding cassette subfamily B member 1 (*ABCB1*), solute carrier family 22 members 6 and 11 (*SLC22A6* and *SLC22A11*) genes (Izzedine *et al.* 2006, Rodríguez-Nóvoa *et al.* 2009, Nishijima *et al.* 2012, Giacomet *et al.* 2013). The variants are associated with varying clinical outcomes such as increased TDF concentration, and moderate to severe risk of renal dysfunctions. Table 1.2 summarises variants described previously that are associated with TDF-associated nephropathy.

Table 1.2. A summary of studies that have observed association between genetic variants and clinical indicators of TDF-associated kidney tubular dysfunction

Gene (protein)	Polymorphism SNP ID	Nucleotide change	Risk genotype	Change in primary protein sequence	Clinical outcomes	Reference
ABCC2 **(MPR2)**	rs717620	c. −24C>T	CC (WT)		Increased risk of AKI (Urine phosphate wasting, decreased eGFR, increased TDF concentration and β2-microglobulinuria)	(Rodríguez-Nóvoa *et al.* 2009, Pushpakom *et al.* 2011, Manosuthi *et al.* 2014)
			CT		Moderate risk of AKI	(Rodríguez-Nóvoa *et al.* 2009)
			TT		Lower risk of AKI	(Rodríguez-Nóvoa *et al.* 2009)
	rs2273697	c.1249G>A	GA/AA	p.Val417Ile	Aminoaciduria	(Rodríguez-Nóvoa *et al.* 2009, Pushpakom *et al.* 2011)
	rs17222723	c.3563T>A	AA	p.Val1188Glu	More frequent in the control group (protective*)	(Izzedine *et al.* 2006)
	rs3740066	c.3972C>T	CC (WT)		Associated with TDF-AKI (β2- microglobulinuria)	(Izzedine *et al.* 2006, Rodríguez-Nóvoa *et al.* 2009)
	rs717620 & rs2273697	c.−24C>T & c.1249G>A	CC and AA	p.Val417Ile	Increased risk of kidney tubular dysfunction (additive)	(Nishijima *et al.* 2012)
ABCC4 **(MRP4)**	rs899494	c.669C>T	CC(WT)		Risk for AKI (phosphorous wasting)	(Rodríguez-Nóvoa *et al.* 2009)
	rs3742106	c.4131T>G	TG/GG		Altered uric acid excretion	(Rodríguez-Nóvoa *et al.* 2009)
	rs1751034	c.3463A>G	GG		Higher intracellular TFV-DP (35%)	(Kiser *et al.* 2008a)
ABCC10 **(MRP7)**	rs9349256	c526G>A	GG		Associated with TDF-AKI, Urine phosphate wasting and (β2- microglobulinuria	(Pushpakom *et al.* 2011)
	rs2125739	c.2759T>C	CC		Marginally associated with TDF-AKI	(Pushpakom *et al.* 2011)
ABCB1 **(PgP-1)**	rs2032582	c.2677T>A/G	AA	p.Ser893Thr (A), p.Ser893Ala (G)	Significantly associated with TDF	(Nishijima *et al.* 2012)
SLC22A11 **(OAT4)**	rs11231809	T>A	TT(WT)		Altered uric acid excretion	(Rodríguez-Nóvoa *et al.* 2009)

c) The role of other important gene variants

The majority of pharmacogenetic variations that influence treatment and HIV development and that have been documented so far have been identified in Caucasian cohorts given the higher number of studies conducted in such populations (Rodríguez-Nóvoa *et al.* 2005). Examples of such clinically relevant variants include the Human Leukocyte Antigen (*HLA*) – B*5701 genetic variant linked to ABC drug associated hypersensitivity reaction syndrome (HRS) (Mallal et al. 2002, Rodríguez-Nóvoa et al. 2007). The prevalence of the *HLA-B*5701*

is 5-8% in Caucasian patients, 2% in Asians, 1% in Hispanics and 0.5% in African Americans (Martin *et al.* 2004, Li *et al.* 2007, Maiers *et al.* 2007, Mallal *et al.* 2008, Saag *et al.* 2008). Because of the high linkage disequilibrium of *HLA-B*5701* to the polymorphism 355 T>G (rs2395029) in HCP5 (HLA complex P5 gene), genotyping this SNP offers a cheaper alternative method of screening for ABC tolerance (Rodríguez-Nóvoa *et al.* 2010). Screening for other *HLA* variants associated with HRS to NVP has been evaluated and also indicates the importance of population-specific genetic HIV biomarkers. *HLA-DRB1*0101* has been identified as a potential biomarker and *HLA-DRB1*0102 and HLA-B*5801* were identified in a South African population (Martin *et al.* 2005, Phillips *et al.* 2013). Variants in the MHC region (*HLA* variants) have been linked to different responses to treatment, disease progression (and non-progression), CD4 count and viral load (Limou *et al.* 2009, Tozzi 2010).

Other genetic variants are not found in pharmacogenes per se, but influence HIV infection or disease progression (Table 1.3), such a clinically relevant gene variant is the CCR5Δ32 genotype that has been found to be more prevalent in Caucasian populations. CCR5Δ32 is associated with immunity against R5 tropic viral infection. People with a homozygous CCR5Δ32 profile lack the CCR5 chemokine co-receptors on the surface of their T cells and are resistant to virus infections that require this entry mechanism but are susceptible to X4 virus infections (which depends on the CCX4 co-receptor for entry) (Liu *et al.* 1996, Eugen-Olsen *et al.* 1997). People with a heterozygous CCR5Δ32 genotype have fewer CCR5 molecules on the T cell surfaces (reviewed in (McDermott 1998, Novembre *et al.* 2005). As a result, individuals with the CCR5Δ32 deletion have innate resistance to CCR5 tropic viruses. Another example of an important genetic variation associated with HIV and its disease progression is rs8177832, which causes a H186R change in the *APOBEC3G* gene. It is frequently identified in individuals of African ancestry and may be associated with more rapid disease progression (An *et al.* 2004).

Table 1.3: Pharmacogenes associated with ARV adverse drug reactions and disease progression in HIV positive patients

Gene	**Risk allele**	**Clinical significance**	**Population with highest frequency**	**References**
Human leukocyte antigen (*HLA*)	*B*5701*	Hypersensitivity to abacavir	Caucasian	(Mallal *et al.* 2002, Mallal *et al.* 2008)
	*DRB1*0101*	Hypersensitivity to abacavir	Caucasian	(Liu et al. 1996, Eugen-Olsen et al. 1997)
HLA complex P5 (*HCP5*)	*rs2395029*	Hypersensitivity to abacavir	Caucasian	(Rodríguez-Nóvoa *et al.* 2010).
Chemokine co-receptor R5 (*CCR5*)	CCR5Δ32 deletion	Resistance to CCR5 tropic HIV virus	Caucasian	(McDermott 1998, Novembre *et al.* 2005)
APOBEC3G	rs8177832	Rapid disease progression to AIDS in HIV positive patients	African	(An *et al.* 2004)

1.4.2 *APOL1* variants and kidney disease

Genetic variants in the apolipoprotein L1 gene (*APOL1*) have been found to be associated with nephropathy. End-stage renal disease (ESRD) presents as focal segmental glomerulosclerosis (FSGS) (Genovese *et al.* 2010). *APOL1* variants are commonly associated with some form of structural kidney dysfunction. These variants are associated with hypertension-associated ESRD, idiopathic FSGS, and HIV-associated collapsing glomerulosclerosis (Genovese *et al.* 2010, Tzur *et al.* 2010, Kopp *et al.* 2011, Ulasi *et al.* 2013, Kasembeli *et al.* 2015a, Purswani *et al.* 2016). The prevalence of chronic kidney disease is approximately 6.0%-48.5% in Africa (Fabian and Naicker 2009) and HIVAN accounts for 24%-83% of cases in South Africa (Table 1.4).

Table 1.4. Renal histology in HIV-infected patients in South Africa

Histology	Durban[21]	JHB[20]	Cape Town[22]	JHB[1]
Biopsy numbers	**30**	**99**	**192**	**364**
Classic HIVAN (%)	83	27	24.4	32.7
FSGS		3	32.8	11.3
HIV Immune Complex Disease (%) (mostly with hepatitis B or C co-infection)		21	30.2	11.8
Mesangial proliferative		6		
Membranoproliferative (type I and III) (%)	7			2.7
Lupus-like (%)				4.4
IgA				
Membranous (%)	13.3	13	5.2	7.7
Exudative-proliferative				
HIV TTP/HUS (thrombotic microangiopathy)				
Various glomerulonephropathies (%) (heterogenous group with different aetiologies)	7	41	24	29.4
Minimal change (%)		2		3.3
Immunotactoid				
Amyloidosis				

Renal histology in HIV-infected patients in South Africa. Copied from (Kasembeli *et al.* 2015b)

Variants in the *APOL1* gene have stronger association with renal failure than the *MYH9* (Myosin, Heavy Chain 9) gene and these variants occur more frequently in African Americans and African populations (Freedman *et al.* 2010, Genovese *et al.* 2010, Tzur *et al.* 2010). In a study conducted by Lipkowitz and colleagues, African Americans with hypertension possessed particular risk variants that were associated with an individual's significantly higher odds of developing nephropathies such as end-stage renal disease (Lipkowitz *et al.* 2013). The *APOL1* variants rs73885319 (a serine to glycine substitution at amino acid position 342 (p.S342G)) and rs71785313 (in-frame deletions which result in asparagine and tyrosine deletions at amino acid positions 388 and 389 respectively (p.N388del: Y389del)) are commonly associated with some form of structural kidney dysfunction (Genovese *et al.* 2010, Tzur *et al.* 2010). These variants are also associated with hypertension-attributed ESRD, idiopathic FSGS, and HIV-associated collapsing glomerulosclerosis (Reeves-Daniel *et al.* 2011). There have been several studies emerging from African populations studying *APOL1* and HIVAN/CKD (Behar *et al.* 2011, Kopp *et al.* 2011, Tzur *et al.* 2012, Ulasi *et al.* 2013) where *APOL1* variants were associated with HIVAN/CKD. These authors studied the risk variants referred to as the G1, G2 and G3 alleles (Figure 1.4). Ko and colleagues were the first to describe the G3 haplotype. They showed that *APOL1* genotypes are geographically restricted where the prevalence of the G2 variants was 3-8%. The Yoruba population (Nigeria) had the highest prevalence of G1 variants with 39% and the *APOL1* G3 haplotype was present in all populations, but the Yoruban population (Ko *et al.* 2013). Another study by Behar and colleagues found the

association with lack of HIVAN among people of Ethiopian ancestry, where the prevalence was lower than the African Americans, HIV-positive African Americans and Western African populations (Behar *et al.* 2013).

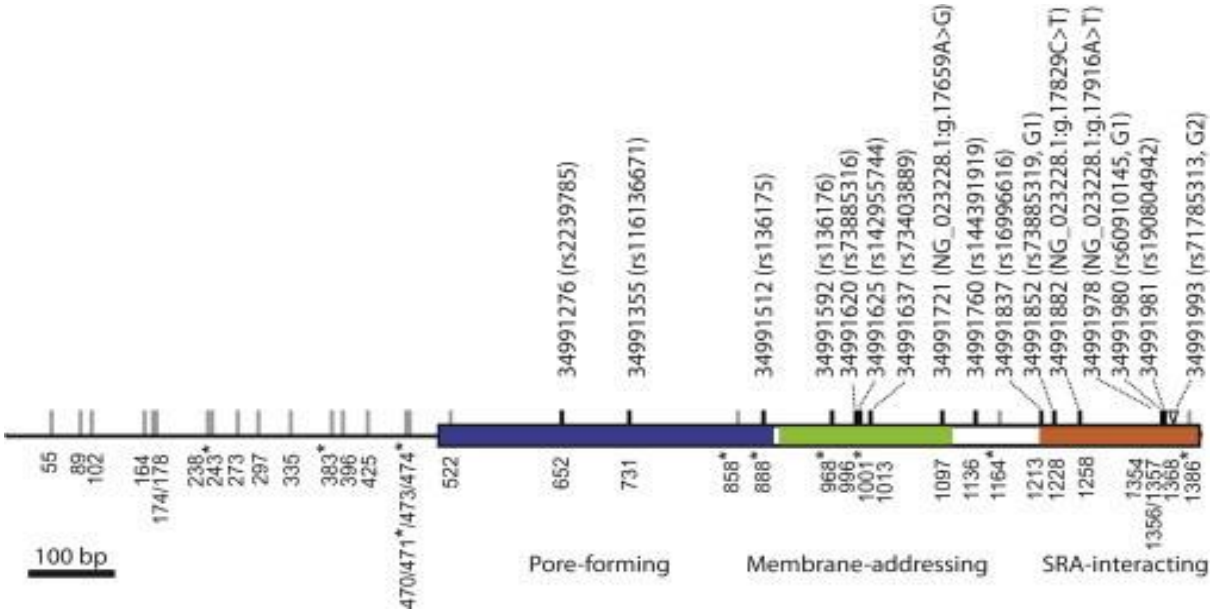


Figure 1.4: *APOL1* variants implicated in HIV associated nephropathy (HIVAN). The *APOL1* variants rs73885319 (a serine to glycine substitution at amino acid position 342 (p. S342G)), rs71785313 (in-frame deletions which result in asparagine and tyrosine deletions at amino acid positions 388 and 389 respectively (p. N388del: Y389del)) and rs60910145 (an isoleucine to a methionine substitution at position 384 (p. I384M)) are implicated in HIVAN. These alleles are referred to as G1 and G2 alleles where the G1 allele refers to rs73885319 and rs60910145 and the G2 allele refers to. rs71785313. G3 haplotype contains eight SNPs positions across the 1.4 kb region. They include two synonymous variants (positions 858 and 1 164), two non-synonymous variants (positions 888 (rs136175) and 968 (rs136176) located at the pore-forming domain and membrane-addressing domain, respectively) and four intronic variants (positions 243, 383, 471, and 474). Blue indicates the region of which codes for the pore-forming domains, green and orange codes for the membrane-addressing and SRA-interacting domains, respectively. The adjacent intron contains variants in the G3 allele. Copied form (Ko *et al.* 2013).

A more recent study by Kasembeli and colleagues looked at the distribution of *APOL1* haplotypes of various allele combinations in a South African population and observed an association of HIVAN with *APOL1* variants. The study cohort included HIVAN patients presenting with collapsing FSGS, HIV positive (HIV+) patients presenting with and without CKD, HIV positive controls with no kidney disease, and population controls who were HIV- and healthy. Approximately 79% of the HIVAN patients carried two copies of the risk alleles

compared to 2% of the population controls. Most notably, the effect size of the variants was higher in the South African black population compared to the African American populations (Kasembeli *et al.* 2015a). Histology studies assessing renal biopsies of HIV-infected patients presenting with renal failure indicated that a significant proportion of the patients presented with classic HIVAN, illustrating the need for a predictive biomarker for renal failure.

A study by Reeves-Daniel and colleagues found that transplant patients receiving kidneys from donors who are carriers of these two variants had significantly shorter graft survival, whereas receptors carrying the two risk variant alleles do not increase the risk of five year allograft loss (Reeves-Daniel *et al.* 2011). It has therefore been proposed that *APOL1* genotyping has the potential to improve donor selection and the benefit of maximising long-term renal allograft survival (Reeves-Daniel *et al.* 2011). It is postulated that these variants occur through natural selection at high frequencies in Africa because of their protective effect against sleeping sickness (Genovese *et al.* 2010).

1.5 Pharmacogenetics in Africa

With the ongoing challenges facing Africa, including increasing prevalence of non-communicable diseases and infectious disease burden, pharmacogenetics offers the prospect of predicting drug efficacy and safety, thus alleviating or minimising the cost of treatment failure (i.e. hospitalisation due to adverse drug reactions) and improving dosage optimisation. The last decade has witnessed increased interest in the genetic diversity of African populations and the contribution that this diversity confers on human evolution and phenotypic traits such as disease susceptibility (Schuster *et al.* 2010, Ramsay *et al.* 2011, May *et al.* 2013, Gurdasani *et al.* 2015, Busby *et al.* 2016).

It is well established that human genetic variation can influence drug outcome. Most clinical drug trials and genetic and pharmacogenetic studies are based on Caucasian, American and Asian populations. Subsequently, most dosing recommendations are based on clinical trials conducted in Caucasian or Asian populations (Chirac 2006, Masimirembwa and Hasler 2013). These studies have led to the development of pharmacogenetic-based interventions. However, African populations are yet to benefit as pharmacogenetic data is limited in Africa (Ramsay *et al.* 2011, Masimirembwa and Hasler 2013, Pillai *et al.* 2013). Where African studies have been performed, they have included mostly African-American subjects, at the

same time studies on native African populations often being limited to relatively small groups of participants and selected genes. Thus far, very few variants have been sufficiently validated in these populations to justify their use in guiding clinical decisions around drug choice or dosing.

Recent studies have shown that African populations reflect greater genetic diversity as compared to other populations. The 1000 genomes project consortium, aimed at deep characterization of the human genome sequence variation, noted that populations of African ancestry contributed the largest number of variants (66%) and also identified the highest fraction of novel variants. This indicated that African populations reflect greater genetic diversity as compared to other populations (Abecasis 2010). Africa being the origin of modern populations is the reason for the high genetic diversity. The Bantu population is postulated to have originated from West Africa, having spread through Sub-Saharan Africa as determined through analysis of paternal lineage based on markers on the Y chromosome (Berniell-Lee *et al.* 2009). Subsequently, African American, Yoruba and Luhya populations have been used as representative populations for African ancestry studies and data are easily accessible through the HapMap and 1000 Genomes Project studies (Abecasis 2010, Altshuler *et al.* 2010, Abecasis 2012, Auton 2015). However, recent studies have highlighted key differences in allele frequencies between these populations and other study populations, especially from Sub-Saharan Africa (Jallow *et al.* 2009, Li *et al.* 2012, Gurdasani *et al.* 2015).

Previous African studies report variation in allele frequencies in variants of clinical significance and some differed significantly compared to other populations. Variants of significance are implicated in the efficacy and outcome of ARVs, anti-mycobacterial and anti-cancer drugs (Subbaraman *et al.* 2007, Aminkeng *et al.* 2013). A pharmacogenetic study comparing allele frequencies of populations of African (Southern - South Africa, Eastern - Tanzania and Kenya, Saharan- Sudan) and European ancestries (Canadian) revealed that variants associated with ADRs were more frequent in African populations (Aminkeng *et al.* 2013). Most significant differences were seen in CYP3A family of genes and they are also associated with poor responses to drugs (such as anti-TB, anti-cancer and ARVs) and reduced survival rates. When analysing the differences among African populations, researchers noted that a risk variant for anthracycline–induced cardiotoxicity was the highest in the Saharan population, a risk variant for cisplatin-ADR was the highest in the Eastern population and a risk variant for DILI (due to nevirapine use) was the most common in the Southern African

population but rare in the Saharan population. The authors inferred that the higher presence of risk alleles was comparable with increased frequency of ADR among African populations and may explain the differences in treatment outcome and survival rates, as African populations were more susceptible to ADR, some of which were extreme (Aminkeng *et al.* 2013). At the same time, it has become clear that genetic diversity significantly influences how populations across the continent respond to drug-based therapy. This diversity provides an opportunity to discover entirely novel polymorphisms that affect therapeutic outcome amongst patients of African ancestry.

Reflecting the continent's high HIV burden, a predominant number of pharmacogenetic studies have focused on investigating variants that affect ARV efficacy and outcome. Variations in *CYP2B6* result in the slow metabolism of several drugs including EFV and the prevalence of *CYP2B6**6 is higher in African populations, making patients of African ancestry more susceptible to EFV-associated adverse events (Nyakutira *et al.* 2008, Gounden *et al.* 2010, Masebe *et al.* 2012, Ngaimisi *et al.* 2013, Swart *et al.* 2013). Likewise, evidence suggests that *CYP2D6* variants associated with the ultra-fast-metabolising phenotype are present at a higher prevalence in patients from Sub-Saharan Africa, leading to reduced efficacy of common ARVs such as EFV and Lopinavir/Ritonavir therapy in such individuals due to sub-therapeutic dosing (Gorny *et al.* 2010, Mukonzo *et al.* 2014). Variants in the CYPA4 and CYP2B6 enzymes and an *ABCB1* gene may also affect ART success by influencing CD4+ recovery rates amongst local patients (Parathyras et al. 2009, Swart et al. 2012). In a South African study, an *ABCB1* variant, rs3842, was significantly associated with low plasma EFV concentrations while rs1128503 was associated with high plasma EFV concentrations in HIV/AIDS patients (Swart *et al.* 2012).

In South Africa, the most common AIDS-related opportunistic infection is TB and co-administration of ARVs and anti-TB drugs is common practice. ARVs and anti-TB drugs are among the most common drugs implicated in ADRs (Mouton *et al.* 2015, Mouton *et al.* 2016, Njuguna *et al.* 2016). Variants in the gene encoding *SLC01B1*, for example, have been shown to be common in South Africa and are associated with increased rifampicin clearance (i.e. a higher dose is recommended) (Chigutsa *et al.* 2011). Likewise, in a randomised controlled trial investigating the potential merit of pharmacogenetics-based therapy, it was demonstrated that a *NAT*-2 genotype–guided regimen potentially reduced isoniazid-induced liver injury and early treatment failure in TB patients (Azuma *et al.* 2013). Yimer and colleagues conducted a

study in Ethiopian individuals and identified variants for the *NAT2* slow-acetylator genotype (rs1801280 or rs1799930 or rs1799931) and *ABCB1* 3435 TT genotype (rs1045642) as biomarkers for predisposition to drug induced liver injury in TB-HIV co-infected patients (Yimer *et al.* 2011).

In addition to the role of genetic factors, multiple other factors that affect the variability of individuals to ARV response need to be taken into account when investigating the potential role of host genetic factors on the outcome of these drugs. These include ethnicity, sex, age, and body weight, contra interactions with other drugs or food, binding to plasma proteins, hepatic impairment, disease status, and pregnancy. With individuals of African descent having been shown to possess a large number of variants and genetic diversity (Tishkoff *et al.* 2009, Bains 2013, Gurdasani *et al.* 2015), studies based on such populations can contribute significantly to elucidating inter-patient variability to treatment and disease susceptibility in Africa, and to understanding drug resistance, adverse reactions and drug metabolism.

1.6 Problem statement

Currently, there is no effective vaccine or cure for HIV infection and HIV/AIDS still remains a chronic illness, while we rely heavily on ARV treatment to reduce the symptoms of HIV/AIDS. With treatment failure, often due to patient non-compliance, adverse drug reactions and the emergence of drug-resistant strains of the HIV virion, being an on-going issue in South Africa, prescribing effective treatment with maximum effect becomes a challenge. One approach to this has been the paradigm shift from a trial-and-error approach to prescribing medication, to the use of more modern pharmacogenetic approaches to improve treatment outcome. HIV-ART pharmacogenetic studies help elucidate the inter-patient variability observed in patients treated with ARVs, and in instances, diagnostic applications based on this principle have significantly improved the success rates achieved with ARV treatment (Barco 2013).

Pharmacogenetic studies conducted in African populations have uncovered notable differences compared to other populations in the contribution of genetic variants to the metabolism and efficacy of drugs, further highlighting the need to comprehensively map the genetic profile of such populations in order to develop more tailored recommendations for drug-based treatment (Kerb *et al.* 2009, Gounden *et al.* 2010, Chigutsa *et al.* 2011, Azuma *et*

al. 2013). Given the significant differences observed across populations to date in variants associated with drug therapeutic outcome, there exists a need to comprehensively map and better understand the implications of such changes in populations of African ancestry. Within Sub-Saharan Africa a greater understanding of population genetic structure is still required before such knowledge can be fully exploited to develop similar strategies for drug-based therapy. Thus, very few variants have been sufficiently validated in these populations to justify their use in guiding clinical decisions around drug choice or dosing, revealing the presence of multiple variants that influence drug efficacy and outcome in a number of ailments including HIV/AIDS, TB and malaria, which continue to dominate the continent's disease burden, with nephrotoxicity being of particular concern.

1.7 Justification

The dual role of pharmacogenetics in HIV treatment is to optimise treatment and to minimise adverse events (Michaud *et al.* 2012). Pharmacogenetic studies have led to the development of relatively inexpensive genetic-based tests to improve the chances of good therapeutic outcome. For abacavir (ABC) for example, screening for a HLA – B*5701 genetic variant linked to drug hypersensitivity and fatality, significantly improves treatment success in certain populations (Kauf *et al.* 2010). There is also a cost-benefit saving. The financial and social costs associated with eventually identifying optimal treatment also increases exponentially following failed ART, where hospitalisation, mortality and morbidity due to ARV-induced adverse events are increased (Núñez *et al.* 2006, Mehta *et al.* 2008, Mok and Minson 2008, Mouton *et al.* 2015, Christine Njuguna 2016, Mouton *et al.* 2016). In South Africa where ARVs are widely used and poor treatment outcome is observed often enough to warrant clinical intervention, a case exists for exploring the potential development of genetic tests to guide ARV prescription. This may particularly be the case for TDF, where higher than expected incidences of side-effects such as nephropathies and a poorer clinical outcome are observed compared to other regimens (Mouton *et al.* 2015, Mouton *et al.* 2016, Njuguna *et al.* 2016, Seedat *et al.* 2016).

1.8 Aims and objectives

The aim of this study was to comprehensively identify all variants in eight TDF associated ADME genes and to select a subset of SNPs for a genetic association study for TDF-associated nephrotoxicity. To achieve this, the objectives of this work will include:

1: To identify, map, characterise and prioritise variants in eight ADME genes implicated in TDF efficacy and nephrotoxicity in 40 HIV-infected patients of Bantu-speaking origin.

2: To perform an association study with the prioritized panel of variants and to evaluate their potential association with TDF treatment outcome in relation to AKI.

CHAPTER 2: NEXT GENERATION SEQUENCING OF ADME GENES AND VARIANT PRIORITIZATION

Recent pharmacogenetic studies conducted in African populations reported notable differences compared to other populations in variants associated with ADRs (Jallow *et al.* 2009, Li *et al.* 2012, Gurdasani *et al.* 2015), warranting their further investigation in local populations. Given the potential for the discovery of novel varianrs and the expected differences in variant allele frequency in our study population compared to populations from publicly available databases, targeted sequencing of ADME genes in an existing HIV positive South African cohort (referred to as the stavudine, d4T, cohort) was undertaken in the first part of this study. Individuals from the d4T cohort were selected for next generation sequencing (NGS) to comprehensively identify and determine the prevalence of variants in eight ADME genes implicated in TDF associated nephropathy and TDF efficacy. Following an assessment of the variants identified and those described in the literature, we selected a subset of the SNPS for a case-control association analysis described in Chapter 3, the second part of the study.

2.1 METHODOLOGY

2.1.2 Study participants

The clinical cohort consisted of 40 HIV-positive black South African patients of Bantu origin, recruited from the Soweto catchment area at the Perinatal HIV Research Unit (PHRU), Chris Hani Baragwanath Hospital. The patients were undergoing either d4T or EFV based ARV therapy at the time of recruitment and blood collection. Blood samples were collected with consent and ethics approval for the study granted by the CSIR Research Ethics Committee (REF: 58/2013) (Appendix A).

3.1.2 Targeted next generation sequencing of 40 HIV-positive black South Africans

In order to select and identify a suitable panel of genes from which known or potentially novel polymorphisms could be identified to screen local patients undergoing TDF therapy, current literature was reviewed to first identify relevant genes associated with drug

metabolism, absorption, elimination and disposition. Particular attention was paid to specifically include genes that contain polymorphisms commonly associated with TDF-induced nephrotoxicity, those involved in kidney function including drug transport, and those that have been implicated as early biomarkers of renal failure. The final gene panel designated the name CSIR PGx Panel 2.1b, included 65 protein coding genes. A list of these genes is summarised in Appendix B. This list was submitted online to the Ion AmpliSeq™ Designer portal (https://www.ampliseq.com/), from which a panel of custom primers was designed to enable simultaneous amplification of all the target genes of interest prior to sequencing. The predicted amplification size of the panel was 242.58 kb which represents a total of 861 amplicons; the amplicon range was between 125 and 345 bps and in total these were predicted to cover approximately 98.57% of the coding (exon) sequences contained within the 65 genes. High-throughput Massively Parallel DNA Sequencing (MPS) of the target genes was achieved using NGS on an Ion Personal Genome Machine (Life Technologies, California, USA) platform as described in detail below. The kits used in the sequencing workflow are all from Applied Biosystems (Life Technologies, California, USA) unless stated otherwise.

2.1.3 Next Generation Sequencing

Next generation sequencing was performed using Ion Torrent technology (Life Technologies, California, USA). Ion Torrent technology is a semiconductor technology based on the principle of "sequencing by synthesis" where a complementary strand is built based on the template strand. During this synthesis, hydrogen ions are released and the detector detects the pH change. The software directly translates chemically encoded information (A, C, G, T) into digital information (0, 1) on a semiconductor chip. The sequencing reaction occurs in a micro-well on the chip, containing a template attached to Ion Sphere™ particles. Below each micro-well is an ion-sensitive layer which measures the change in pH and converts it into voltage. The wells are flooded with a single species of deoxyribonucleotide triphosphate (dNTP) at a time and if the nucleotide is incorporated, the released hydrogen ion is detected. Multiple nucleotides may be incorporated per a single flow due to homopolymer repeats, which limit the technology especially if the signal is higher than the detection range. This process occurs simultaneously in millions of wells, regardless of the chip capacity, and takes up to several hours to complete (Thermo Fisher Scientific, 2016). The sequencing workflow used in this study consisted of the following steps: a) Ion Ampliseq™ panel design b) library

preparation, c) template preparation, d) enrichment of positive Ion sphere™ particles (ISPs), e) quality control f) sequencing and data processing and variants filtering. The sequencing workflow is summarised in Figure 2.1.

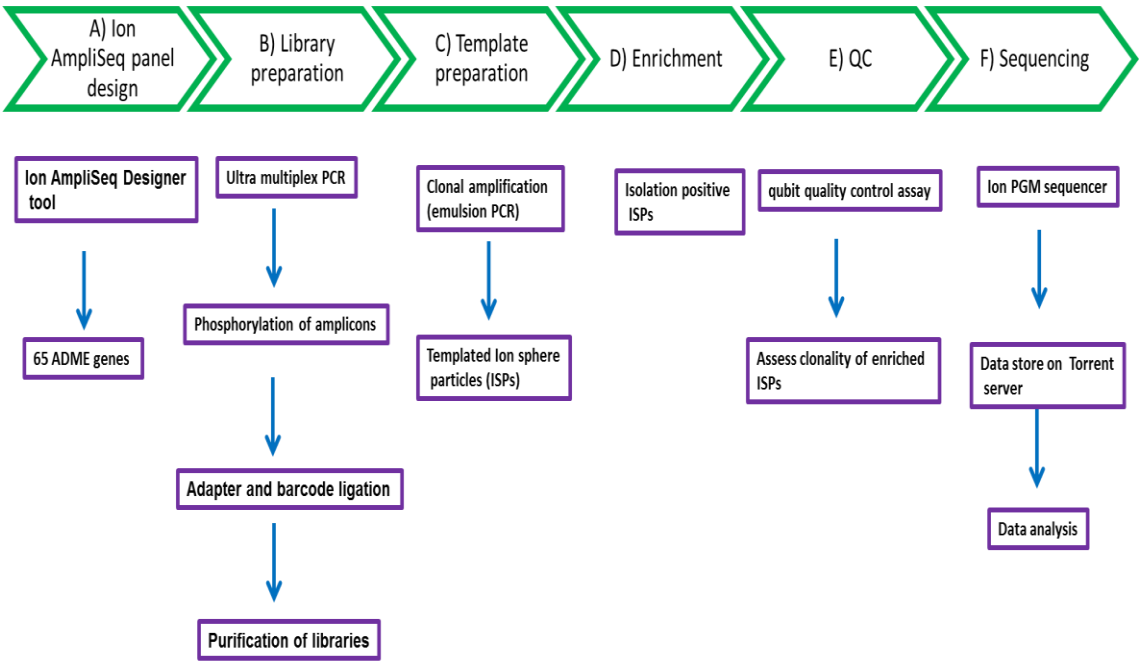


Figure 2.1. Summary of the Next Generation Sequencing (NGS) workflow that was implemented in this study. The sequencing workflow consists of the following steps: A) Ampliseq™ panel design B) library preparation, C) template preparation, D) enrichment of positive ISPs E) quality control F) sequencing and data processing and variants filtering.

a) Ion AmpliSeq™ DNA Library preparation

An amplicon library of the targeted regions within the 65 genes described previously was prepared for each individual patient using an Ion AmpliSeq™ Custom Panel of primers designed with Ion AmpliSeq Designer software (https://www.ampliseq.com/) (Appendix C). For the purposes of this study, data from a subset of eight transporter genes (*ABCB1*, *ABCC2*, *ABCC4*, *ABCC10*, *SLC22A1*, *SLC22A6*, *SLC22A8* and *SLC22A11*) were prioritised due to the direct relevance of these genes to TDF treatment and AKI. The remainder of the genes were utilised for other studies outside the scope of this work. I am first author of a publication that is under revision and development for which the abstract is shown in Appendix D that describes the variants identified from the study of all 65 genes.

Library preparation for each individual was performed using an Ion AmpliSeq™ Library Kit 2.0 and Ion Xpress Barcode Adapter kits 1–16 and 17-32 (catalogue numbers 4471250 and

4474009) according to the manufacturer's instructions. The use of barcodes allowed for multiplex parallel sequencing of up to 20 individuals and easy identification during data analysis of each individual sample. The library preparation workflow consisted of four steps. 1) PCR amplification of genomic DNA targets 2), digestion of excess primers and phosphorylation of fragments, 3) ligation of barcoded adapters and 4) purification of each library. Using an Ion AmpliSeq™ Library Kit 2.0, PCR amplification of genomic DNA targets was achieved by starting with 10ng of each patient sample and amplifying the target region based on the recommended program (Table 2.1).

Table 2.1. Program run to amplify target DNA regions

Stage	**Step**	**Temperature (°C)**	**Time**
Hold	Activate enzyme	99	2 min
Cycle (16X) *	Denature	99	15 sec
	Activate and extend	60	4 min^
Hold	-	10	Hold

*Recommended no. of amplification cycles
^ Recommended annealing and extension time for reactions with ≤1536 primer pairs per pool

The next step involved partially digesting primer sequences using 2 µl of Fafu™ reagent, provided in the kit, and digesting based on the program in Table .2.2. A mixture of Ion P1 Adapter and Ion Xpress™ Barcode X at a final dilution of 1:4 for each adapter was prepared and 2 µl was used. The barcoded adapters were then ligated to the amplicons using 2µL of DNA ligase, switch solution and run using the program in Table 2.3. The barcoded adapters allow for multiplex parallel sequencing and easy identification of each sample and their unique design allow for detection during template.

Table 2.2. Program run to partially digest the primers

Temperature (°C)	**Time (mins)**
50	10
55	10
60	20

Table 2.3. Program run to perform the ligation reaction

Temperature (°C)	**Time (mins)**
22	30 min
72	10 min
10	Hold (for up to 1 hour)

Library purification was performed using Agencourt Ampure XP beads (Beckman Coulter Inc, California, USA), which allow concentration dependent size selection. The Ampure beads are made of made polystyrene surrounded by a layer of magnet and coated with carboxyl molecules. In the presence of polyethylene glycol (PEG) and salt 20% (w/v) PEG, 2.5M NaCl), DNA can irreversibly bind to the beads. Binding of the DNA is PEG and salt-dependent, making the volumetric ratio of the beads to DNA critical. A low concentration is used for DNA clean-up, i.e. they bind high molecular weight molecule because of their strong electrostatic interaction. Higher concentrations bind a range of DNA fragments, with larger fragments displacing smaller fragments. If the concentration is high enough (precise), smaller fragments can bind. In this study both low and high concentrations were used to remove high molecular weight DNA (gDNA) and the primers, respectively.

In line with the recommendations of the protocol, 200 pM libraries were prepared and the Qubit® fluorometer (Life Technologies, California, USA) was used for quantification. The libraries were combined by adding equal amounts (3.6 - 5 µl) of each equalised 200 pM library into a single pool. The 40 patient samples were prepared in batches of 14-20 patient samples to reduce human error and were sequenced in batches to improve the coverage for each patient sequenced. Some of the patient samples that did not produce NGS data of sufficiently high coverage were repeated to achieve the required level of data quality.

b) Template preparation

To prepare monoclonal fragments attached to ISPs, the Ion One-Touch 400 Template Kit v2 was used to prepare the template preparation reactions and the Ion One-Touch System was used to perform emulsion PCR reactions. Monoclonal fragments are recommended per ISP as they allow for amplification of that fragment which will be sequenced, thus providing enough coverage for that region. Through emulsion PCR, clonal ISPs were generated in water-in-oil microreactors containing ISPs, template, PCR components and primers. After setting up the Ion OneTouch™ 2 Instrument, 900 µl of amplification solution was prepared by adding Ion PGM™ Template OT2 400 Reagent Mix (500 µl), Ion PGM™ Template OT2 400 PCR reagent B (285 µl), Ion PGM™ Template OT2 400 enzyme mix (50 µl), ISPs Template OT2 400 reagent X (40 µl) and library (25 µl). Ion PGM™ Template OT2 400 Ion Sphere™ were added to the amplification solution. The mixture was immediately added / "installed" in the Ion PGM™ OneTouch plus reaction filter assembly and run using the PGM: Ion PGM™

Template OT2 400 kit protocol. Template-positive Ion PGM™ Template OT2 400 ISPs were recovered by spinning samples for 10 minutes (automatic) and removing the supernatant and leaving 100 µl of the recovery solution. Then 500 µl of Ion OneTouch™ wash solution was added to both recovery tube and mixed by pipetting, each suspension was transferred to the new labelled 1.5 mL Eppendorf LoBind® tube. The tube was heated at 50°C for 2 minutes and centrifuged for 2.5 minutes at 15,500 × g. All but 100 µl of the wash solution was removed from the tube and transferred into well 1 of the 8-well strip from the Ion OneTouch™ ES Supplies Kit. A 2µl aliquot of the sample was kept as an unenriched control for the Ion Sphere™ Quality control assay analysis, which offers a pre- and post-enrichment quantification of ISPs.

c) Enrichment and quality control

After the emulsion PCR, template-positive Ion Sphere™ particles were enriched with the Dynabeads® MyOne™ Streptavidin C1 beads. These beads (13 µl) were washed with 130 µl of MyOne™ Beads wash solution and resuspended in 130 µl of the same solution. A melting solution prepared with 280 µl of Tween®20 detergent and 40 µl of 1M NaOH (provided the Ion PGM™ Template OT2 400 Kit), so that the final concentrations are 0.1% Tween® (v/v) and 125 mM NaOH. Enrichment was performed using the automated Ion OneTouch^{TM} ES system which used melting solution, Dynabeads® MyOne™ Streptavidin C1 beads and Ion OneTouch™ Wash Solution, resulting in > 200 µl of the enriched template in neutralisation solution. The automated process takes approximately 35 minutes.

An Ion Sphere™ quality control assay (Life Technologies, California, USA), was employed to assess the enrichment efficiency and clonality of the enriched sample before sequencing, thus ensuring that the ISPs had monoclonal fragments attached. The quality control assays involve labelling the ISPs with two different fluorophores: Alexa Fluor® 488 (AF488) and Alexa Fluor® 647 (AF647). The ratio of the Alexa Fluor® 647 fluorescence (templated ISPs) to the Alexa Fluor® 488 fluorescence (all ISPs present) yields the percentage of templated ISPs. A starting volume of 2 µl of unenriched sample and 20 µl of enriched samples (centrifuged at 15 500 × g for 1.5 minutes, and 10 µl of supernatant removed to concentrate it down to 10 µl) were used. Ion AF488 and AF647 probes were bound to the ISPs by adding 19 µl annealing Buffer and 1 µl Ion probes directly to the ISPs and mixed. The AF488 and AF647 fluorescent probes were bound by incubating at 95^0C and 37^0C for two minutes at

each temperature. The unbound probes were subsequently washed 3X with 200 µl of quality control wash buffer, mixed, centrifuged at 15 500 × g for 1.5 minutes and the supernatant was removed leaving 10 µl. After the final wash 190 µl of quality control wash buffer was added for a total volume of 200µl. The samples and the negative control (200 µl of quality control wash buffer) were analysed with the Qubit® 2.0 fluorometer.

The AF488 probe is specific for all ISPs and the AF647 is specific for the adapter sequence ligated on each library fragment. The fluorescence of each probe was detected using the Ion Sphere™ Quality control assay and Qubit® fluorometer described previously. A Qubit® 2.0 Easy Calculator Microsoft® Excel® spreadsheet file containing the instrument specific Calibration Factor (Life Technologies, California, USA) was used to calculate the % of template ISPs; a cut-off range of 10-30% was used as a guideline based on the manufacturer's recommendations (Life Technologies, California, USA) to prevent polyclonality (>30%) and having under-templated ISPs (<10%).

After the Ion Sphere™ Particle preparation, MPS was performed with an Ion Torrent Personal Genome Machine (PGM) system using approximately 200 µl of templated ISP, the Ion PGM HiQ Sequencing Kit and Ion 318™ Chip v2 (318™). The 318™ chip was selected for use instead of the 314™ or 316™ chip size because of its superior loading capacity (Figure 2.2) of 5.5 million reads, its throughput (a measure of how many units of information a system can process in a given amount of time), and its relatively easier loading procedure of the templated ISPs. Each sequencing run took approximately seven hours to complete.

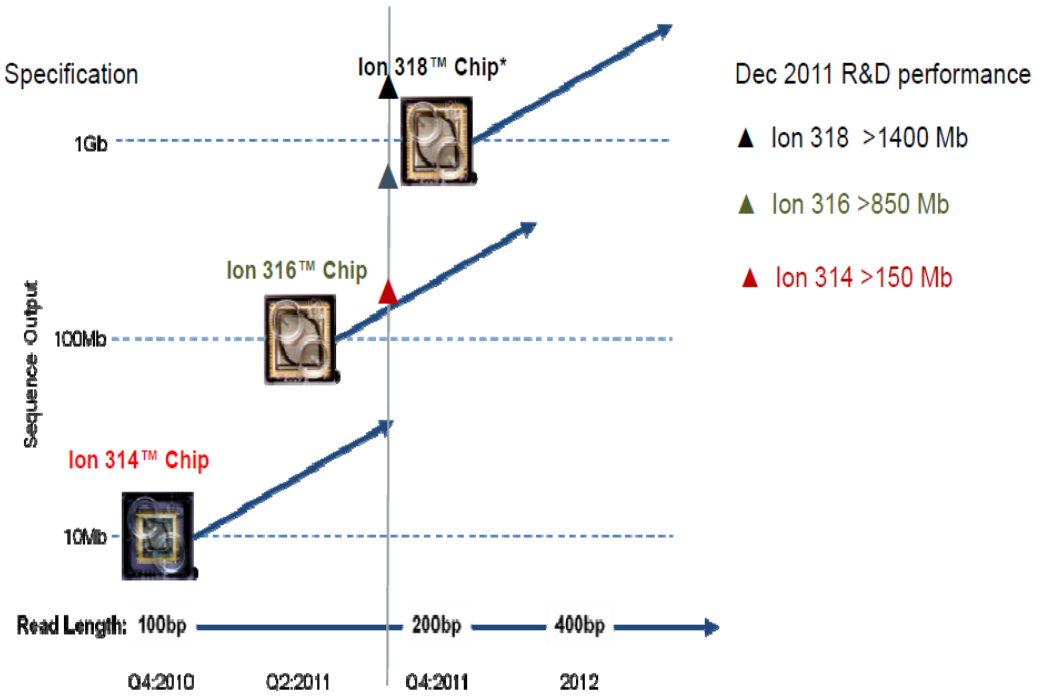


Figure 2.2. Ion PGM chip sequencing capacities and performance. (ThermoFisher, 2012)

In order to extract more data from the sequencing libraries that were created from the cohort, libraries were also sent to the National Genomics Infrastructure / SciLifeLab, Department of Immunology, Genetics and Pathology, Uppsala University, Sweden for targeted sequencing on an Ion Torrent S5 sequencer. The libraries were attached to the Ion Sphere™ particles via template preparation and emulsion PCR on the Ion One chef system as described previously. The Ion Sphere™ particles generated from the libraries sent to Uppsala University were sequenced on two 530 chips, which have a capacity of 3-8 Gb, producing up to 15-20 million reads each.

d) Data processing and filtering

Initial data processing was performed using a pre-defined automated workflow in Torrent Suite™ Software (version 4.4 and 4.4.2 for the last run) (Life Technologies, California, USA). This software was accessed using the Torrent Browser, which is a web-based programme to monitor, plan, and manage sequence data. The setup and configuration of this software were performed by Dr Natasha Beeton-Kempen, CSIR Biosciences Unit, CSIR. Following sequencing, the main data analysis steps in Torrent Suite™ involved aligning the sequences to the hg19 reference genome (Genome Reference Consortium, 2016) for base

calling and mapping of the sequence data to the reference genome. This was followed by variant calling for homozygous or heterozygous single nucleotide variants, and other changes such as insertions or deletions using an automated approach. Finally, variant data were exported from the software for downstream analysis using variant analysis software. Data were exported in the form of binary alignment map (BAM) files. The subsequent data processing steps that were performed using the BAM data files are summarised in Figure 2.3. BAM file analysis was initially achieved using the Ion Reporter™ software (version 4.4), which is an internet browser-based application that allows for automated analysis of samples, annotation of variants based on public and custom annotation databases, and the filtering of variants to identify those that are biologically relevant.

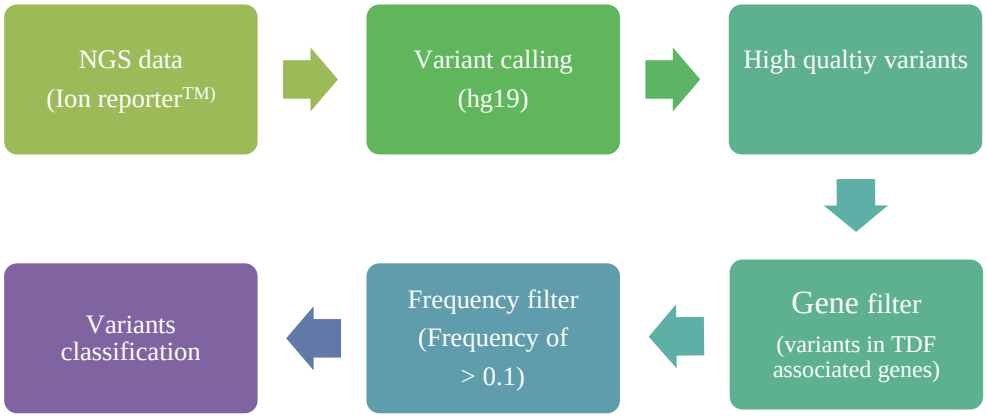


Figure 2.3. Summary of the NGS data filtering strategy that was employed for sequencing data analysis. Variants that lie in the eight selected genes of interest and that occur at a frequency of >0.1 in the cohort of the 40 sequenced individuals. Genes associated with TDF absorption (*ABCB1*) and elimination (*ABCC2*, *ABCC4*, *ABCC10*, *SLC22A6* and *SLC22A8*) were prioritized. Since TDF is elimination is less understood, *SLC22A1* and *SLC22A11* transporters were also included because they belong to the same class of transporters localized in the kidney.

For Ion Reporter™ analysis, BAM files were automatically transferred from Torrent Suite™ to a dedicated user account in Ion Reporter™. An automated workflow was then developed (by Dr Mancama) to analyse the sequencing data to identify germline variants in targeted regions for the CSIR PGx Panel 2.1b (IAD66518_197) (4.4) (AmpliSeq™ panel), this was done using the Hg19 reference genome, and applying variant filters and using different annotation databases available for version 4.4 (including for example SIFT, PolyPhen, ClinVar, cosmic, dbSNP, drugbank, GO and OMIM). The data were then downloaded in batches of 20 files and analysed using MS Excel. The software packages Golden Helix Genome Browser and SNP Variation Suite 8 (SVS) (Golden Helix™, Montana, USA) were

also employed to mine and visualise the sequencing data and to provide descriptive statistics of the data. For variant filtering prior to in-depth analysis, an NGS filtering strategy was employed (Figure 2.3). Although sequencing was done on 65 ADME genes, the current analysis for my dissertation was restricted to the eight genes coding for proteins that interact with TDF during absorption (*ABCB1)* and elimination (*ABCC2, ABCC4, ABCC10, SLC22A6 and SLC22A8*). Since TDF elimination remains poorly understood, *SLC22A1* and *SLC22A11* transporter genes were also included because they belong to the same class of transporters localised in the kidneys (Benjeddou 2010).

Following variant calling, a frequency filter was employed to select variants that occur at a frequency of ≥ 0.1 in the 40 individuals (i.e. occurred eight or more times among 80 alleles). This approach was used to analyse data using SVS (version 8) and Ion Reporter™. Subsequently, functional analysis through bioinformatic analysis was based on assessing amino acid changes and the subsequent effect of these changes on protein function. For variants where the consequence of a functional change has not been extensively demonstrated *in vitro* / *vivo*, additional information regarding the potential effect of such variants was gained using the algorithms PolyPhen-2 (Ion Reporter and http://genetics.bwh.harvard.edu/pph2/) and SIFT (Sorts Intolerant From Tolerant amino acid) (Ion reporter and http://sift.jcvi.org/) tools. SIFT and PolyPhen values represent the probability that an amino acid substitution is physiologically damaging in nature; the data is a score and a qualitative prediction, either 'possibly damaging' (0.8-0.9), 'probably damaging' (≥0.9), 'deleterious' (≤0.05) or 'tolerated' (www.ensembl.org). Further selection criteria included whether the variants have previously been reported to be associated with TDF treatment outcome.

2.1.4 NGS sequencing data validation

Sanger sequencing of patients with ambiguous or undetermined calls was employed for further investigation and confirmation. Two potentially novel variants in *ABCB1* and *ABCC2* were identified after the initial NGS data filtering and the most common variant in most populations, rs717620 (in *ABCC2*) (Izzedine *et al.* 2006, Rodríguez-Nóvoa *et al.* 2009, Nishijima *et al.* 2012), was shown to be rare in our population. The three SNPs were therefore assessed. The region flanking the SNP was amplified and purified and the PCR products were sent for sequencing to Inqaba Biotech.

a) Primer design

Polymerase chain reaction (PCR) reverse and forward primers were designed for the flanking sequence around the SNP, the sequence was retrieved from Ensembl (www. ensembl.org).

The region of interest (ROI) was selected and an 18-24 nucleotide forward sequence primer set in the 5' – 3' direction was selected in the beginning of the ROI. For the reverse primer, an 18-24 nucleotide reverse sequence primer in the 5' – 3' direction was selected at the end of the ROI. The reverse primer sequence was then input into a reverse complement tool to get the reverse sequence in the 5' – 3' direction. Both sequences were then put into a Tm calculator (New England Biosciences, NEB, Massachusetts, USA) (http://tmcalculator.neb.com/#!/) to calculate the melting temperature of each primer, GC content (%) and the Tm difference. A cut-off range for GC content was 40-60% and the Tm difference was set to below 5°C. The primer sequences were then sent off to Inqaba Biotech for synthesis. The primer sequences for NGS validation are shown in Table 2.4.

Table 2.4. Primers sequences and their expected amplicon size.

SNP	**Primer sequences (5'- 3')**	**Size (bps)**
rs717620	FWD: ATGGCCACTCCTACAGAG REV: GTTCTCCTTCCTCACACG	614
Novel *ABCC2*	FWD: CTTCCTTTTACCCCTCCC REV: CTCTTCAGGGCCTGTATG	551
Novel *ABCB1*	FWD: CTGTCAAGGCACTGGGAAC REV: CTGGGCACTGGTCCAAGTA	716

b) Polymerase chain reaction and amplicon purification

PCR was used to amplify DNA fragments that encompass the SNPs using a T100 thermocycler (Bio-Rad, California, USA). Taq polymerase (New England Biosciences, NEB, Massachusetts, USA) or iProof™ High-Fidelity DNA polymerase (Bio-Rad, California, USA) kits were tested initially to determine optimal annealing condition at annealing temperature (T_a) of melting temperature (T_m) minus 5°C. The reaction setup was based on the manufacture's recommendations (Figure 2.5 and Figure 2.6). To reduce primer dimers and non-specific amplicons DMSO (iProof™ kit, Bio-Rad, California, USA) was added at 4%. Gradient PCR was done to determine the optimal T_a for rs3742106; T_a was between 58°C and 64°C, with 60°C showing the best results.

c) Sequencing reaction set up

Amplified templates were purified using the QIAquick PCR Purification kits (Qiagen, Hilden, Germany). The amplicons were run on a 1% (w/v) agarose gel prior to sequencing. The sequencing reactions were set up using the BigDye® Terminator v3.1 Cycle Sequencing Kit (Life Technologies, California, USA) as illustrated in Figure 2.7 and Figure 2.8 and sent to Inqaba Biotech for sequencing. FinchTV 1.4 (Geospiza) or Bioedit (Ibis Biosciences, Carlsbad, USA) was employed to view the sequencing data and produce the chromatograms.

Table 2.5. The PCR reaction set up used to amplify sequences with the NEB kit

NEB Taq reagents	**Concentration**	**1X**
5X Buffer	1X Buffer	5 µl
DNTPs	0.2 µM	0.5 µl
Forward primer	0.5 µM	0.5 µl
Reverse primer	0.5 µM	0.5 µl
Template	44 ng	2 µl
Enzyme	0.025 units/µl	0.125 µl
ddH_2O	-	Up to 25 µl

Table 2.6. The PCR reaction set up used to amplify sequences using the iProof^{TM} kit

iProof™ reagents	**Concentration**	**1X**
10X Buffer	1X Buffer	2.5 µl
dNTPs	0.2 µM	0.5 µl
Forward primer	0.5 µM	1.25 µl
Reverse primer	0.5 µM	1.25 µl
Template	44 ng	2 µl
Enzyme	0.02 units/µl	0. 25 µl
ddH_2O	-	Up to 25 µl

Table 2.7. Sequencing reaction set up for the BigDye® terminator v3.1 cycle sequencing kit

Bigdye v3.1 reagents	**Concentration**	**1X**
Ready reaction mix	2.5 X	4
BigDye Sequencing buffer	5 X	2
Forward or reverse primer	3.2 pMol	6.4
Template	5-20 ng	1-2 µl
ddH_2O	-	Up to 20 µl

Table 2.8. Sequencing cycling conditions

Cycle sequencing conditions	
1	Place the tubes in a thermal cycler and set to the correct volume
2	Perform an initial denaturation
	a. Rapid thermal ramp to 96 °C
	b. 96° C for 1 min
3	Repeat the following for 25 cycles:
	• Rapid thermal ramp* to 96 °C
	• 96 °C for 10 sec
	• Rapid thermal ramp to 50 °C
	• 50 °C for 5 sec
4.	Rapid thermal ramp to 4 °C and hold
5.	Place the products in the fridge until ready to ship
***Rapid thermal ramp is 1 °C/second.**	

2.1.5 Characterization of identified variants

SVS (version 8.4.4) was employed to characterise and generate visual illustrations of the variants. The variants were classified in terms of gene, location relative to the gene (e.g. in the intron, upstream, downstream, coding, intergenic, untranslated regions or non-coding transcripts), and class (e.g. splicing, deletion, insertion, frame-shift insertion, frame-shift

deletion, initiator codon, substitution, synonymous or non-synonymous). Novel variants were identified using SNP Variation Suite 8.4.4 (SVS version 8). The 1000 Genomes Project Phase 3 (KGP), dbSNP common 149, NCBI (dbSNP 149), and Exome Aggregation Consortium (ExAC) databases were used to identify and filter out common variants. I also examined the data to identify variants in the published literature that had previously been associated with TDF related nephropathies (Table 1.2). These variants were further characterised by comparing the allele frequencies from the d4T cohort and commonly used databases. These included three African populations, a Caucasian population and a population of Chinese ancestry. Populations of African ancestry included a Bantu-speaking cohort from the African Genome Variation Project (Gurdasani *et al.* 2015), and two KGP's African populations, the Luhya (LWK) and Yoruba (YRI). The Bantu-speaking cohort comprised of 100 Zulu speakers from South Africa who were whole genome sequenced at low-coverage. The LWK is an East-African population from Webuye, Kenya and the YRI is a West-African population form Ibadan, Nigeria, (www.ensembl.org/)). Both the Caucasian and the Chinese populations are from the KGP study. The variants previously identified were associated with TDF intake, pharmacokinetics and toxicities and were also cross-checked against the pharmacogenetic mutation database (PGMD^{TM}) and pharmGKB database. These databases are resource tools used to identify all published genetic variants associated with drug response in patients. These databases also include pharmacogenetic studies that generate detailed phenotype to genotype relationship to facilitate the identification of relevant information from public sources.

2.2: RESULTS

2.2.1 Sequencing chip performance

Sample sequencing was initially performed on an Ion Torrent PGM^{TM} sequencer and in order to extract more data from the sequencing libraries, sequencing was repeated on an Ion Torrent S5^{TM} sequencer that was released by Thermo Fisher shortly after the initial sequencing had been performed.

Four chips (Ion 318™ Chip v2) were run on the Ion Torrent PGM^{TM} sequencer with libraries from 20, 20, 14 and 16 patient samples, respectively. The initial chip was run using 2 µl of a

combined patient library that had been created by combining 3.6 µl of each 200 pM library. The chip loading density for the initial chip was observed to be of low quality, with less than 15% of the chip wells utilised resulting in a low quality of final library with a read total count of 1 305 reads representing 1 305 representing 0.23 Mbp of data (Figure 2.4, Panel A). Data from this first chip was therefore excluded from the final dataset used for analysis.

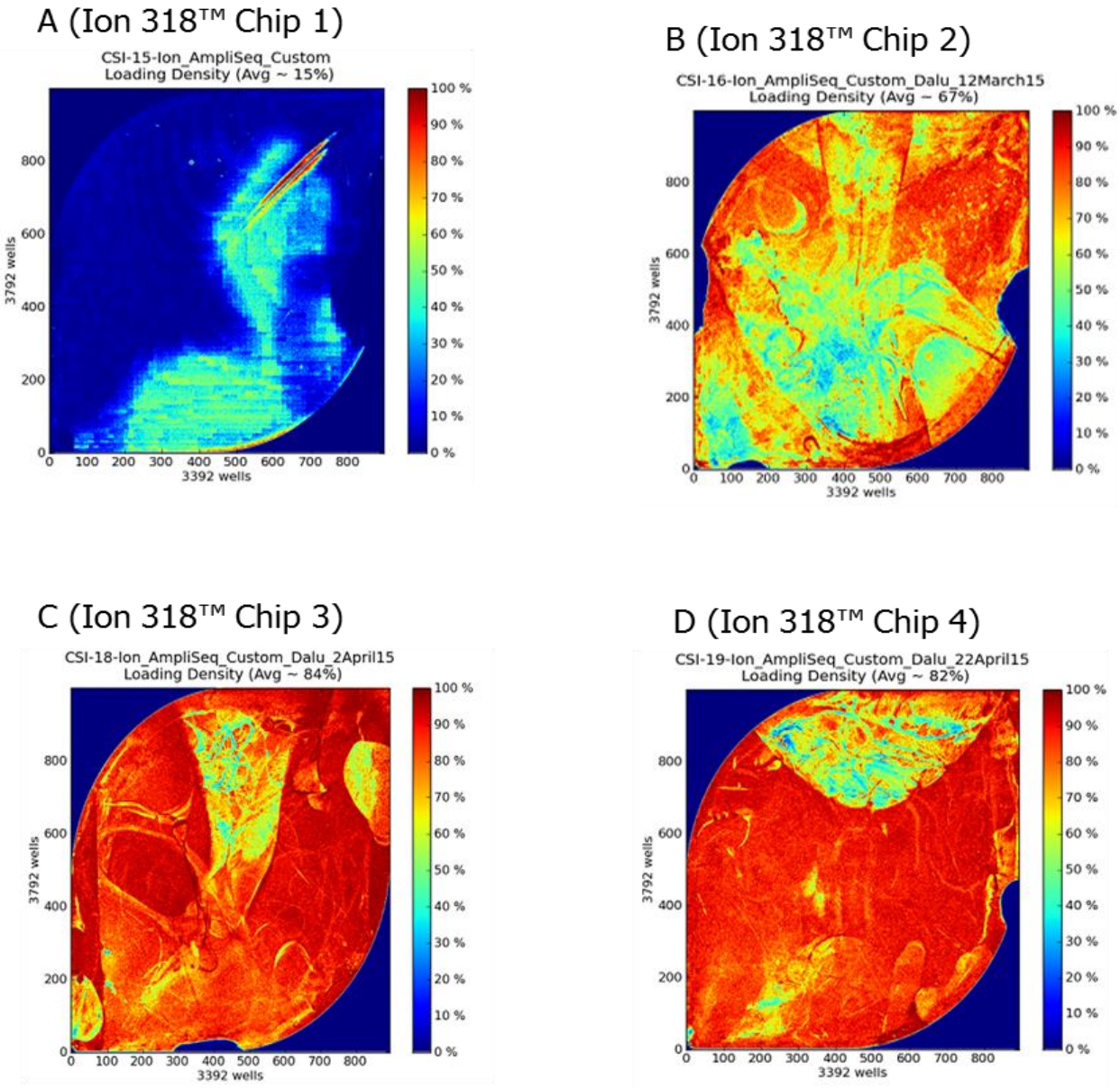


Figure 2.4. Loading density plots for the Ion 318™ chips used to produce the NGS data, in chronological order with (**A**) showing the first run followed by (**B**), (**C**) and (**D**). (**A**) showed poor loading density owing to insufficient starting library concentration. (**B**) was better but still not optimal, and (**C**) and (**D**) had excellent loading density reaching up to 100% in the majority of wells. Loading subsequently improved after the first attempt, with each chip generating an average of 3.9 million library reads.

The poor performance of the first chip was corrected by using 25 µl of the pooled 200 pM libraries for template preparation. This led to a significant increase in the loading density of the chip to 67% (Figure 2.4, Panel B), with a final library reads count of 2 643 682 reads representing 511.25 Mbp of data. The third chip (Figure2.4, Panel C) produced the highest loading density of 84% with a final library read count of 4 934 752 reads representing 1 021.27 Mbp. Similarly, the last chip (Figure2.4, Panel D) achieved a loading density of 82% and a final library read count of 4 119 973 reads representing 786.89 Mbp.

For each Ion 318™ run, sequencing quality settings were defined to ensure that only bases with a likelihood of a 99% chance of being correctly called (Q20) were included in the final data. In addition, only patient sample data that obtained sequencing nucleotide depth or coverage of greater than 50X were selected for further analysis to ensure the sequencing accuracy for each individual nucleotide. The coverage requirement for regional target resequencing typically ranges between 30-80X (https://genohub.com/next-generation-sequencing-guide/, 13/04/16). The average coverage for the final 40 patients included in the dataset described below was 92X. Individual 170253 (depicted as BAM file no. 8) was sequenced at an average read depth (coverage) of 320X and a coverage for target regions of 99%, and obtaining a read depth of 288 for SNP *ABCC4* 669T (rs899494), as depicted in Figure 2.5.

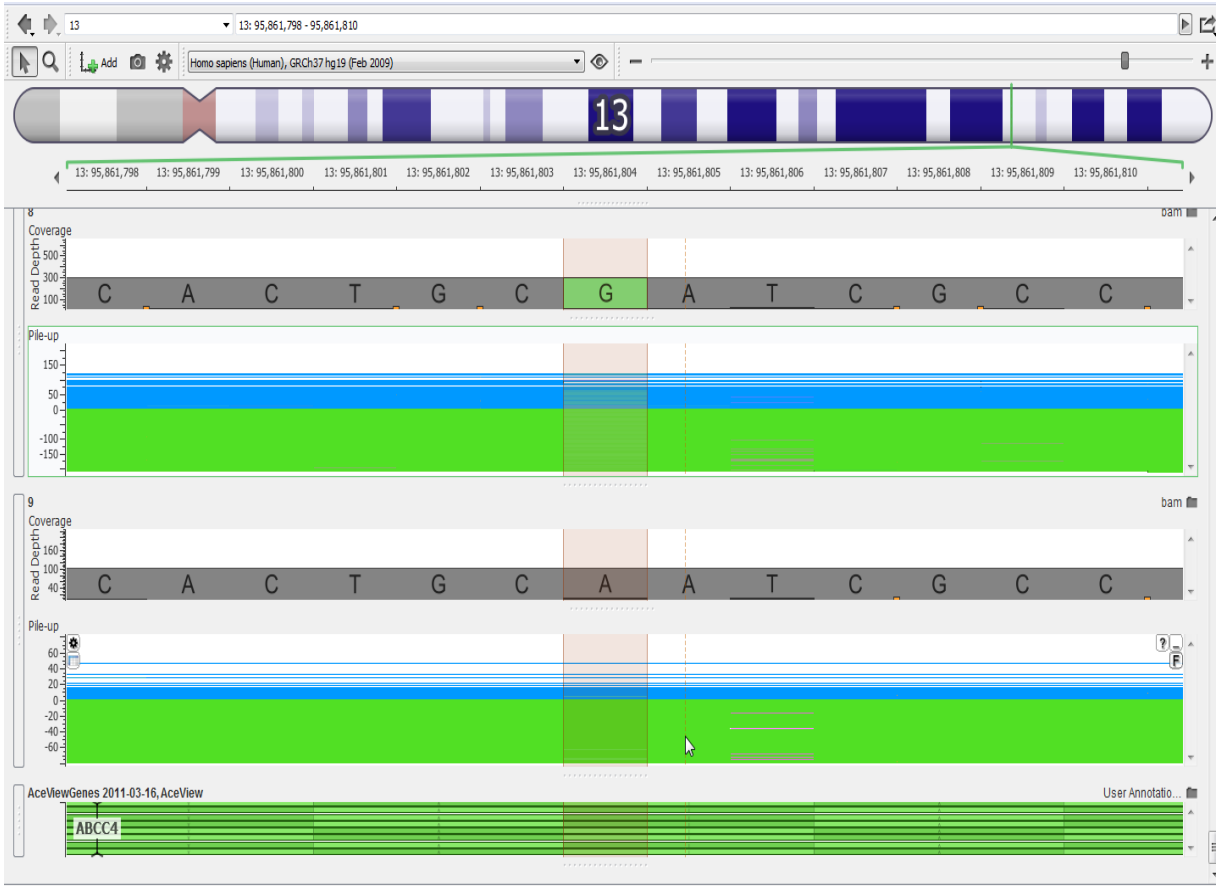


Figure 2.5. Next generation sequencing coverage for individuals 170253 (depicted as Bam file no. 8) and 170245 (Bam file no. 9) for SNP rs899494 located in gene *ABCC4*, position chr13:95861804. Individuals 170253 and 170245 had a read depth of 288 and 96, respectively. Individual 170253 carried the variant and could potentially be at risk of TDF-AKI (phosphorous wasting). The blue pile up represents the forward strand and the green pile up represents the reverse strand.

In order to extract more data from the sequencing libraries, libraries were also sent to Uppsala University, Sweden, for targeted sequencing on an Ion Torrent S5™ sequencer. Loading density for the two chips was observed to be 88% and 90%, respectively, with a final library read count of 16 144 778 wells and 17 991 586 reads respectively (Figure2.6). This respectively represented 3.76 Gbp and 4.09 Gbp of sequence data. Identical quality settings to those used on the Ion PGM™ system were employed. The average coverage for the same set of 40 individuals described above was 820X.

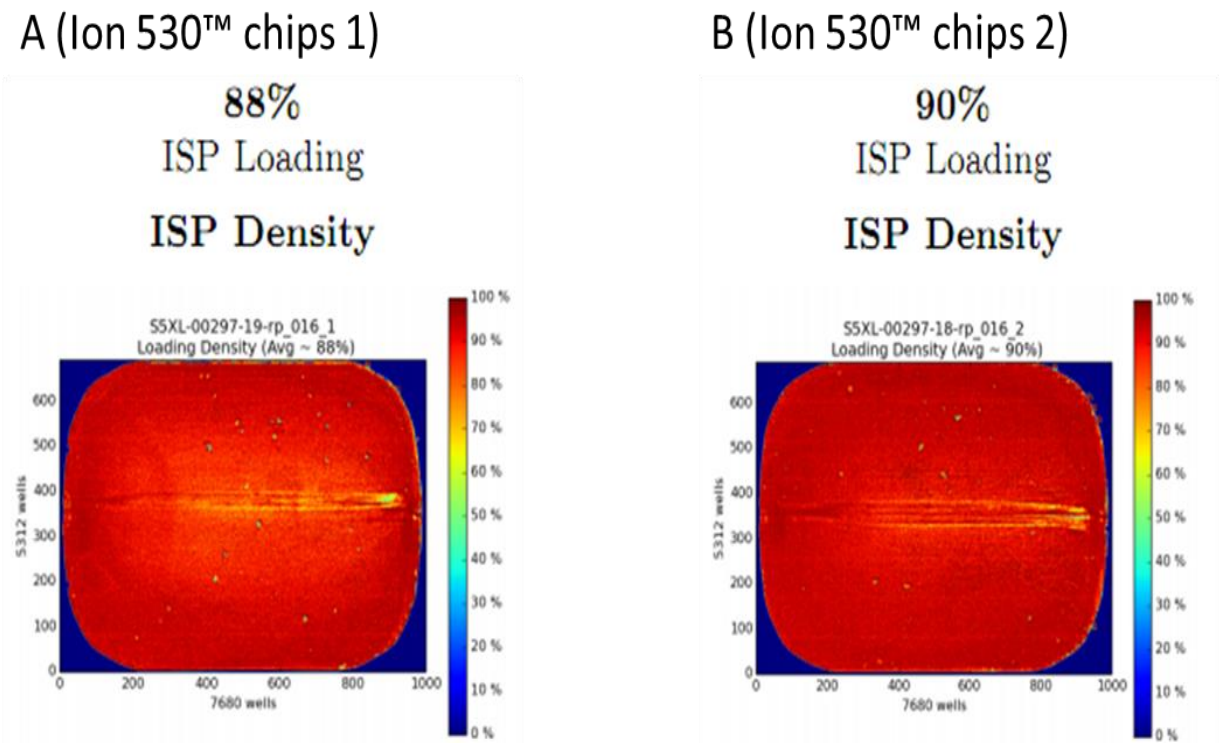


Figure 2.6. Loading density plots for the Ion 530™ chips used to produce the NGS data (**A-B**). The loading density for each chip was 88% and 90%, respectively. The chips produced an average of 17 million library reads due to their higher loading capacity compared to the Ion 318TM chip.

2.2.2 Sequencing data analysis

a) Ion PGM™ vs Ion S5™

The Ion reporter™ software was employed to detect, analyse and annotate approximately 2 600 variants across 65 genes that were present in the 40 sequenced patient samples obtained using Ion Torrent PGM™ sequencer. Golden helix™ (genome browser and SVS version 8.4.4 (SVS 8)) software was then employed to further visualise and mine the sequencing data. NGS filtering and characterisation strategies were applied to the high quality data (following QC) generated from both sequencers.

Ion PGM™ data were imported into the genome browser in BAM file format and into SVS 8 in VCF format. BAM is the binary format for storing sequence data and is therefore the binary version of a SAM file, a tab-delimited text file that contains sequence alignment data. VCF is a text file format that is stored in a compressed manner. The data obtained via the Ion Torrent PGM™ and Ion Torrent Ion S5™ sequencers were imported separately into SVS 8.

The files generated were analysed for the overlapping variants and 618 variants were identified by both sequencers (Figure 2.7). PGM™ sequencer generated 234 different variants and the S5™ generated 1 380 different variants, which were further analysed and filtered to identify possible variants to include in the panel for the case-control association study.

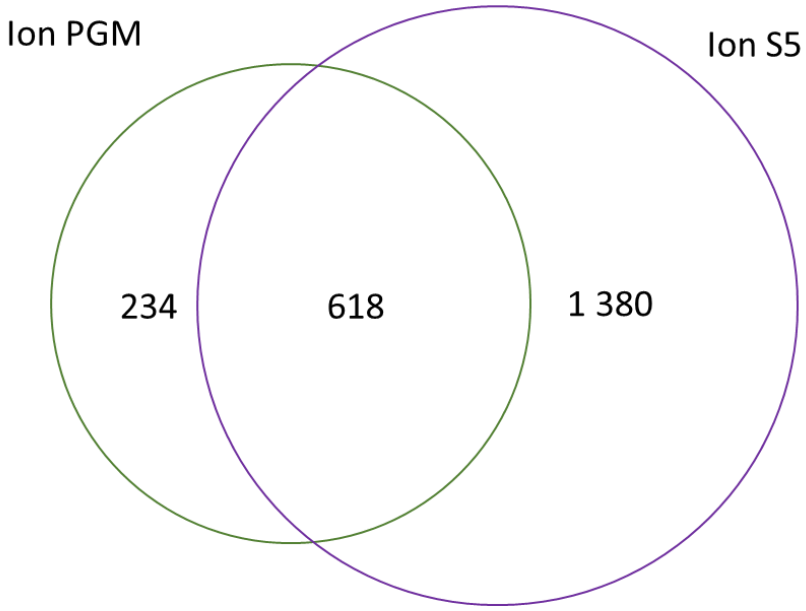


Figure 2.7. Venn diagram showing the total number of variants produced by Ion Torrent sequencers PGM™ and S5™. SVS 8 software was used to analyse both data sets and identified 618 overlapping variants.

Variants that were identified by PGM™ only (i.e. 234 variants) were filtered to eliminate variants that obtained a genotype quality score of ≤ 15 and a read depth of ≤10. This resulted in only 89 of the 234 variants having high quality reads generated via the PGM™ sequencer. Only the 89 high quality variants were analysed further. To gain an understanding of the nature of these variants, the variants were further characterised by assessing the nature of the underlying nucleic acid change (Figure 2.8).

To validate the presence of selected variants, these regions were PCR amplified and sequenced using Sanger sequencing. Two potentially novel *ABCC2* and *ABCB1* variants were identified after the initial NGS data filtering. In addition, an *ABCC2* variant (rs717620) (a commonly implicated TDF associated AKI (Izzedine *et al.* 2006, Rodríguez-Nóvoa *et al.*

2009, Nishijima *et al.* 2012)), was in fact rare in the d4T cohort. These three variants were assessed for robustness and the results are summarised in the table in Appendix E. Sanger sequencing indicated that *ABCC2* was rare and the two novel variants were not observed in our sample population. The was discordance between the two methods used to genotype the d4T patient samples indicated that the Ion PGM™ data was not accurate and unreliable. As a consequence, Ion PGM™ variants were excluded and only those generated by the Ion S5™ sequencer were examined further. Ion S5™ data includes the overlapping 618 variants identified in both sequencers.

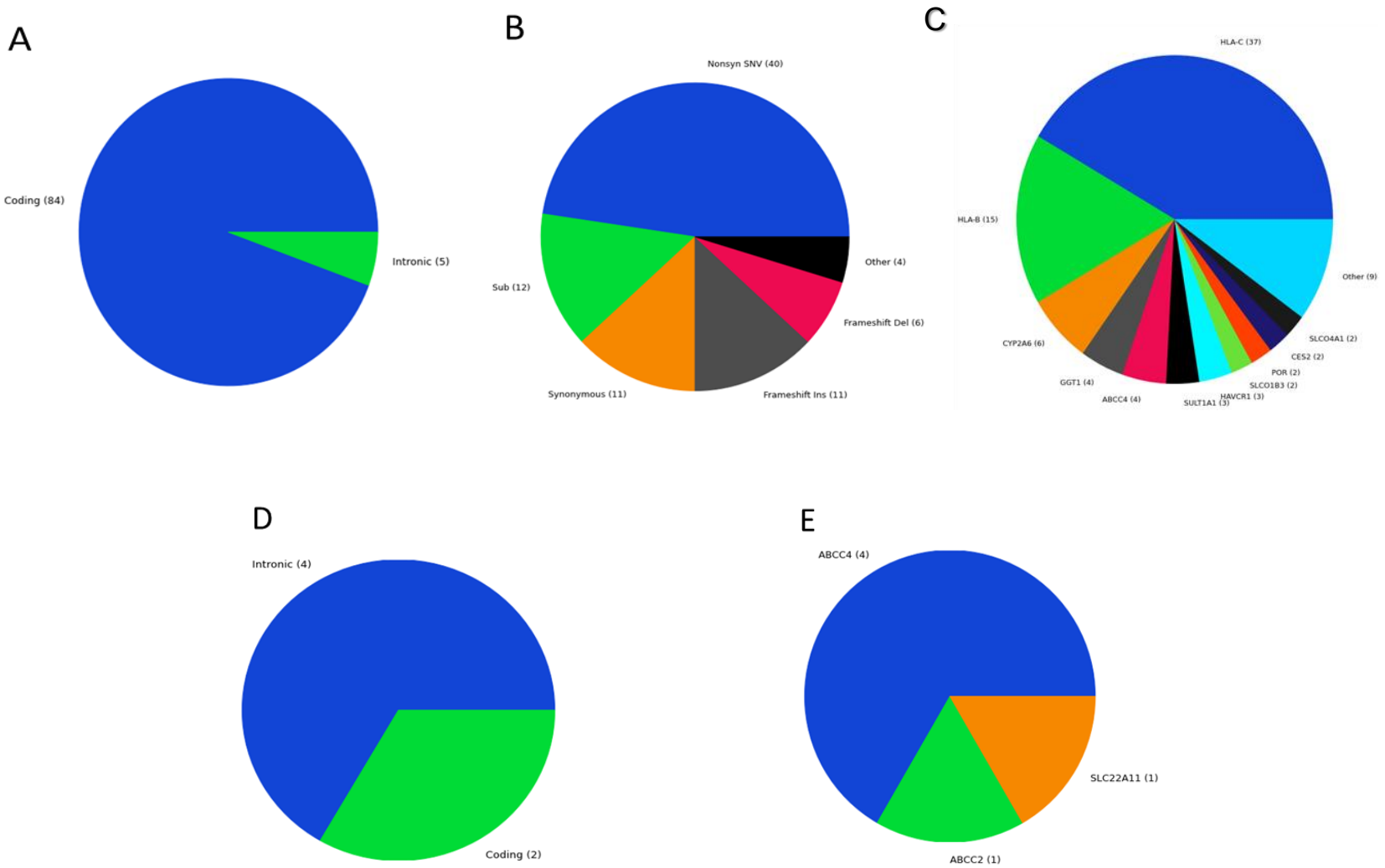


Figure 2.8. Variant classification for data generated using Ion Torrent PGM^{TM} sequencing. Pie chart (**A**) illustrates the classification based on the variant's position relative to the transcript of each variant (n=89). (**B**) illustrates the nature of the 84 coding variants, this includes variants that lie within a splicing region of the coding transcript, and those within 1 000 bp of the transcript start/stop site (upstream or downstream variants). 'Non syn' SNV refers to non-synonymous SNPs causing an amino acid change; 'Synonymous" refers to a coding SNP that does not cause an amino acid change; 'Sub' refers to other substitutions that do not cause a frameshift change. 'other' represents additional variant classes which include insertions, frameshift substitution, deletion-substitution and deletion. The 4 other variants were grouped together as their respective classification groups had less than 6 variants. (**C**) shows the corresponding number of variants that lie in each respective gene (n=64). (**D**) illustrates the classification class of 6 variants that lie in genes relevant to TDF nephropathy and (**E**) shows the number of variants that lie in genes relevant to TDF nephropathy. Please note that these data were not used for downstream analysis as several vatiants could not be validated with Sanger sequencing.

The data generated from the Ion Torrent Ion S5™ sequencer (the second sequencing experiment) was imported into SVS 8 and filtered identically to the Ion PGM™ data by eliminating variants that obtained a genotype quality score of ≤ 15 and a read depth ≤ 10. Ion S5 sequencing generated 1 687 high-quality variants from the 1 998 variants that were imported. The variants were further characterized (Figure 2.9). A total of 581 variants were identified to be coding, i.e. the DNA region that is comprised of protein coding exons, coding for proteins and bounded by a translation start and stop codon. These included 283 non-synonymous variants (which resulted in amino acid change), 260 synonymous variants (which results in no amino acid change), 14 other substitutions (where a nucleotide is substituted or replaced with another in a non-coding region) and 29 classified as other, which include frameshift mutations (where the variation resulted in a shift in the way the nucleotide sequence is coded and would change the amino acid sequences downstream of the mutation), deletions, splicing variants and insertions. The variants lie in 64 of the 65 genes selected during the primer design stage (Figure 2.9). Data generated from NGS sequencing was used to compile a manuscript that describes the deep sequencing of 40 HIV positive patients of Bantu ancestry (Appendix D). The manuscript aims to provide information that further broadens understanding of the inter-patient variability in drug response and metabolism.

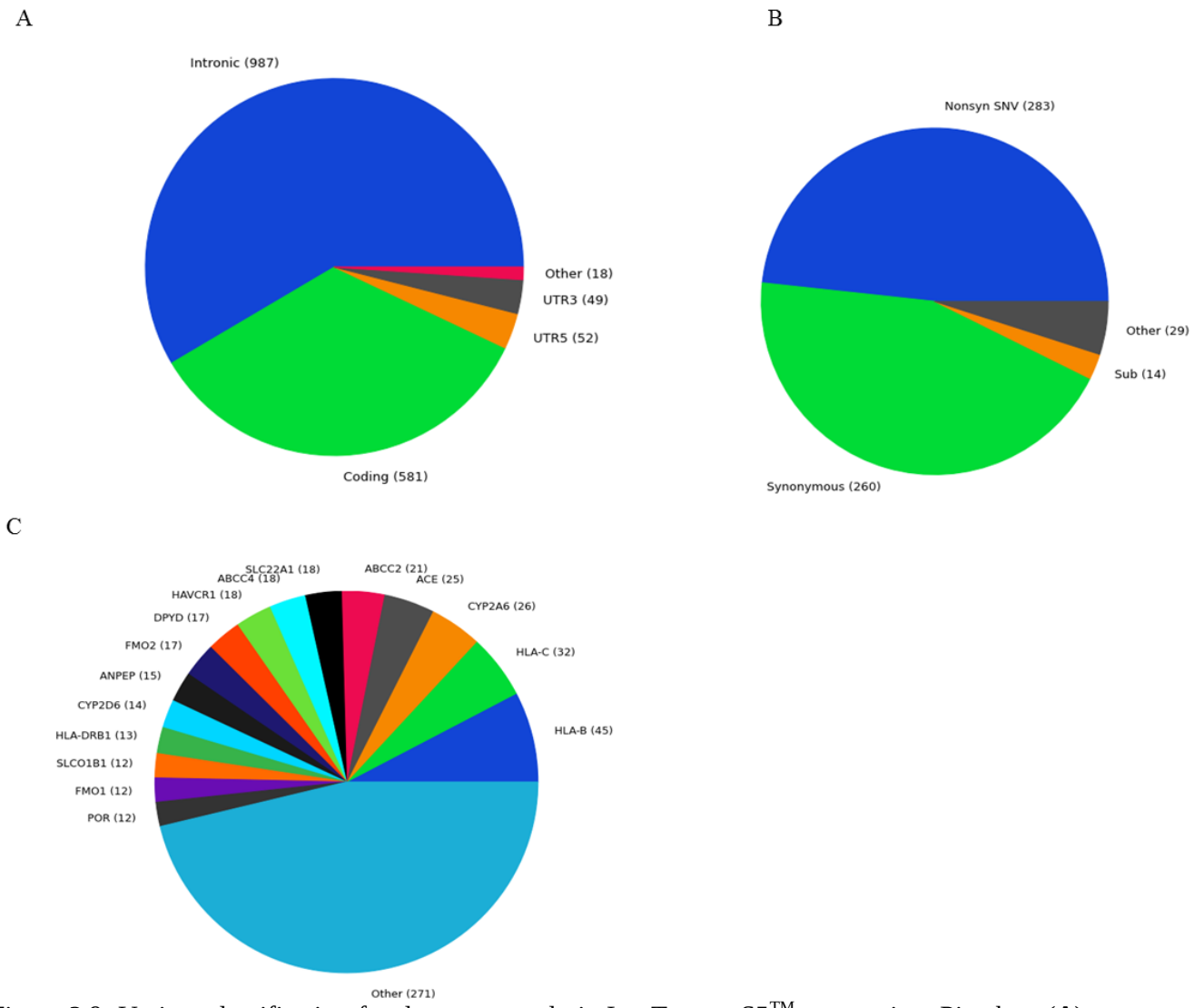


Figure 2.9. Variant classification for data generated via Ion Torrent S5™ sequencing. Pie chart (**A**) illustrates the classification based on the variant's position relative to the transcript of each variant (n= 1 687). (**B**) illustrates the nature of the 581 coding variants, this includes variants that lie within a splicing region of the coding transcript, and those within 1 000 bp of the transcript start/stop site (upstream or downstream variants). 'Non syn' SNV refers to non-synonymous SNPs causing an amino acid change; 'Synonymous" refers to a coding SNP that does not cause an amino acid change; 'Sub' refers to other substitutions that do not cause a frameshift change; additional variant classes are represented by 'other', which includes insertions, frameshift insertion, frame shift deletion, stop-loss, stop-gain, initiator codon and deletion. The 29 other variants were grouped together as their respective classification groups had less than 14 variants each. (**C**) shows the corresponding number of variants that lie in each respective gene.

2.2.3 Variants filtering and characterization

a) Gene panel filtering

For the specific purpose and focus of this study, the filtering strategy in Figure 2.10 was implemented. Gene panel filtering was employed to look for known and potentially novel

variants in genes coding for proteins that interact with TDF during absorption (*ABCB1)* and elimination (*ABCC2, ABCC4, ABCC10, SLC22A6, and SLC22A8*). Since TDF elimination is less well understood, *SLC22A1* and *SLC22A11* transporters were also included because they belong to the same class of transporters localised in the kidney and were included in other studies (Rodríguez-Nóvoa *et al.* 2009). Gene panel filtering of the S5^{TM} data resulted in 275 variants remaining and of these 86 variants lie in coding and splicing regions (Figure 2.11).

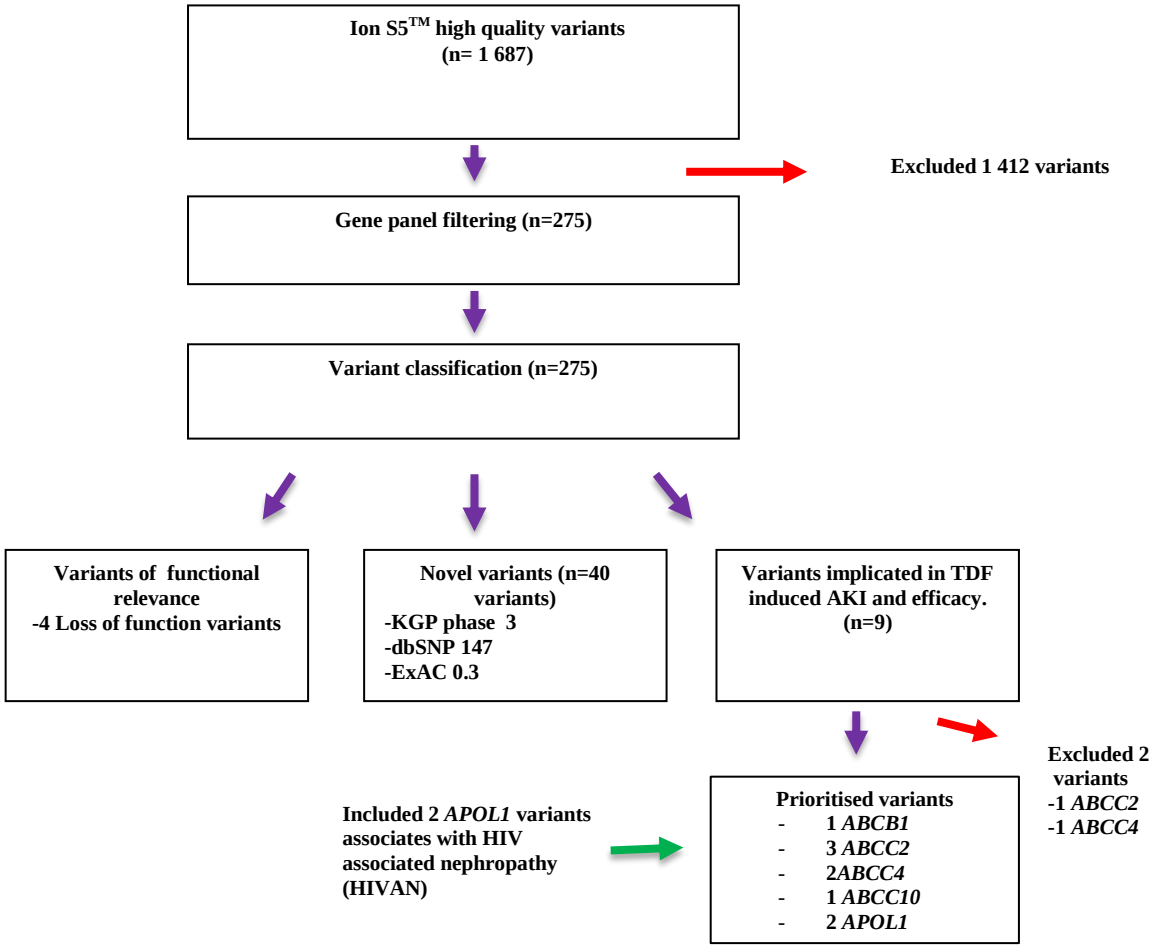


Figure 2.10: Summary of the filtering strategy that was employed for prioritising variants identified in the IonS5^{TM} data. The criteria for prioritizing the variants included selecting variants that lie in eight genes implicated in TDF induced AKI. These genes are associated with TDF absorption (*ABCB1*) and elimination (*ABCC2*, *ABCC4*, *ABCC10*, *SLC22A6* and *SLC22A8*). Since TDF is elimination is less understood, *SLC22A1* and *SLC22A11* transporters were also included because they belong to the same class of transporters localized in the kidney. The variants were futher prioritized based on functional impact, novelty and association with TDF induced AKI, efficacy and nephropathy.

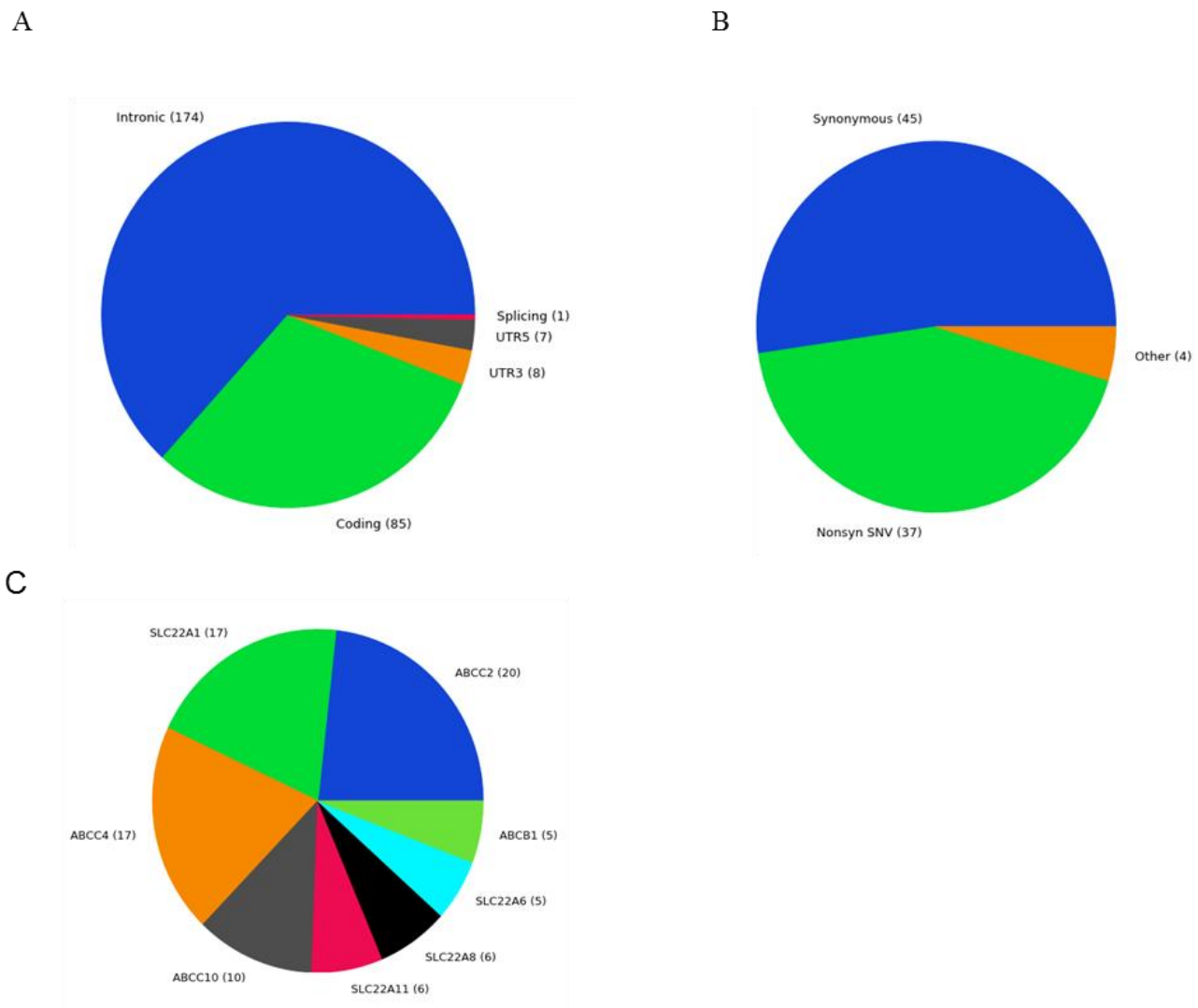


Figure 2.11. Gene panel and filtering of high quality variants identified by the IonS5^{TM}. Gene panel filtering was employed to identify variants localised in eight genes implicated in TDF induced AKI and efficacy, where 275 variants were identified and characterised. Pie chart (**A**) shows a variant classification according to gene position. Pie chart (**B**) illustrates the classification for each coding variant. 'Non syn' SNV refers to non-synonymous SNPs causing an amino acid change. Additional variant classes are represented by 'other', which includes frameshift insertion, frame shift deletion, deletion and splicing variants. The four other variants were grouped together as their respective classification groups had few variants each. (**C**) illustrates the number of variants in each gene.

Frequency filtering was employed where only variants that occurred at a frequency of ≤ 0.1 were filtered out. This resulted in the identification of 265 variants that occur at an allele frequency of greater than 0.1. The coding variants displayed a minor allele frequency range of 0.11-0.5 and a call rate range of 0.025-0.90 which is the total proportion of patient samples that were sequenced or called at those loci.

b) Bioinformatic functional analysis

Functional analysis included looking at variants from the Ion S5™ data that were indicated to be potentially functional, i.e. non-synonymous, where 41 (48%) of 86 coding variants were identified in genes implicated in TDF induced AKI and efficacy (Figure 2.11). Loss of function (LoF) variants are defined as variants that are predicted to disrupt protein-coding genes and include stop-gain, stop-loss, splicing, frame-shift insertion and frameshift deletion variants. From the variants identified three (7.3%) were potentially loss of function variants (Table 2.9). The three LoF variants included a frameshift deletion in *ABCC2*, a splicing variant and a frameshift insertion in *SLC22A1*. All three variants were novel and potentially unique in our population.

The potential functional impact of the all 275 variants identified in the eight genes implicated in TDF induced AKI was determined using SIFT (damaging threshold of ≤ 0.5), PolyPhen algorithms and ExAC VEP annotation 0.3 database (via SVS 8). SIFT indicated that 11 variants were potentially deleterious and PolyPhen indicated that three variants were possibly damaging and that five variants were probably damaging (Figure 2.12). Table 2.10 highlights the non-synonymous or missense variants with potential functional impact, detected by the ExAC VEP 0.3 database. A total of 26 variants were identified as deleterious, possibly or probably damaging. Of these, six variants classified as deleterious (based on the SIFT score), one as possibly damaging and four as probably damaging (based on the PolyPhen algorithm). From the 26, three variants were identified as deleterious (or probably damaging) by both tools (Table 2.10).

Table 2.9. Classification of Loss of function variants

Gene	**Location (hg19)**	**Classification**	**HGVS Coding**	**HGVS Protein**	**rs ID**
SLC22A1	6:160543329	Frameshift Insertion	c.361_362insG	p. Cys121fs	-
SLC22A1	6:160560898	Splicing	c.1275_1276+6delTGGTAAGT	-	-
ABCC2	10:101577129	Frameshift Deletion	c.2159_2160delTT	p. Phe721fs	-

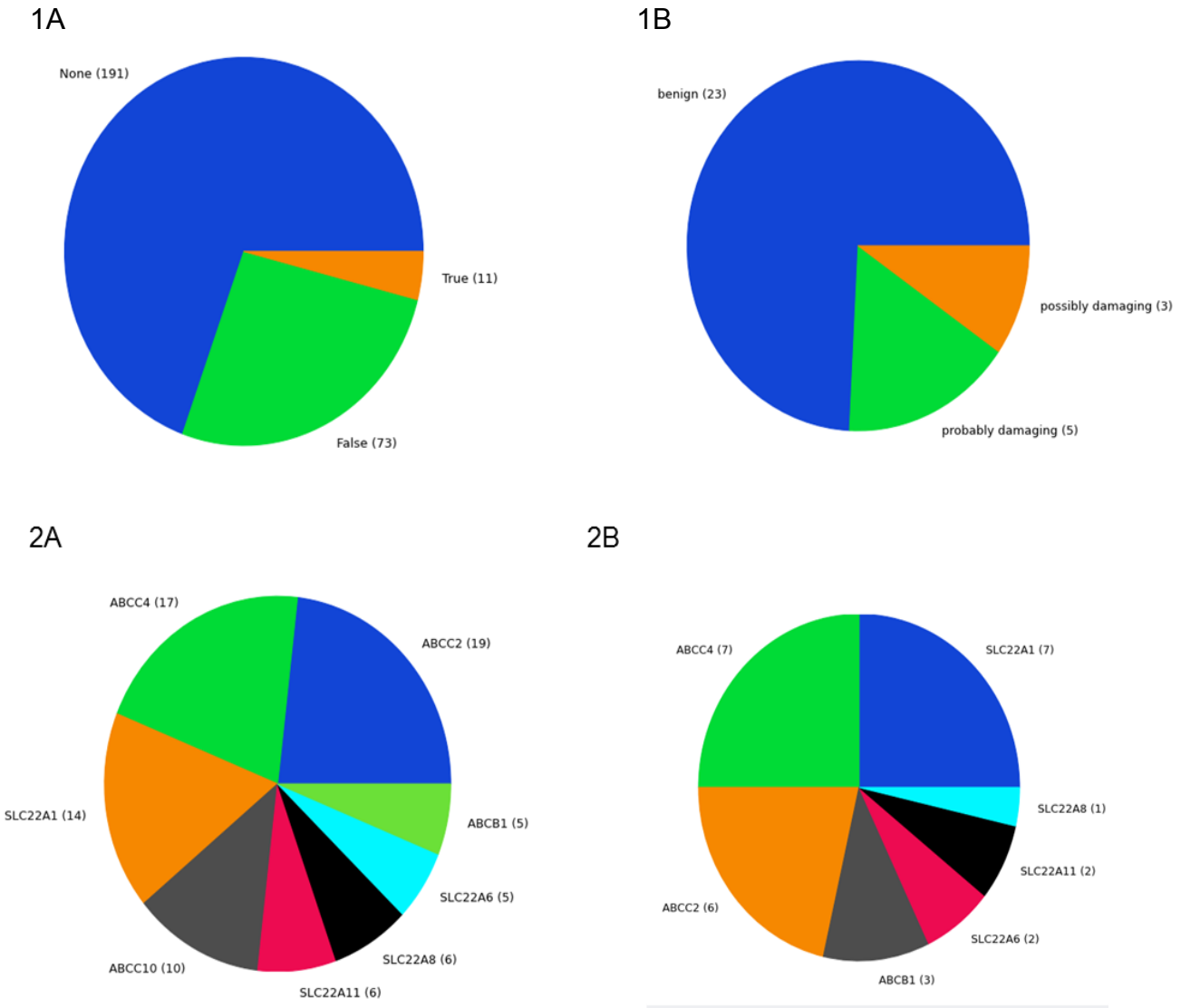


Figure 2.12. SVS version 8.4.4 functional effect analysis of 275 variants identified after gene panel filtering. SIFT (**A**) and PolyPhen (**B**) analysis of the total number of variants that were identified in genes relevant to TDF nephropathy. Pie charts (**1A** and **1B**) illustrates the number of variants classified as true indicates that the variants have a damaging threshold of ≤ 0.5, making them probably damaging and may exert a significant functional change. The corresponding genes in which the variants lie are indicated in the pie charts (**2A** and **2B**).

Table 2.10. Functional analysis of variants located in genes potentially associated with AKI

Gene Symbol	Marker	Ref/Alt	Consequence	Position in cDNA	Position in CDS	Position in Protein	Amino Acid Change	Co-located Variation	SIFT	PolyPhen
ABCC10	6:43400691	C/T	Missense variant	1203	844	282	Leu/Phe	rs73424611	Tolerated (0.1)	Benign (0.132)
ABCC10	6:43401027	A/G	Missense variant	1539	1180	394	Asn/Asp	rs61740187	Deleterious (0.01)	Probably damaging (0.999)
ABCC10	6:43411739	T/C	Missense variant	2760	2401	801	Ser/Pro	rs369336854	Tolerated (0.31)	Benign (0.003)
ABCC10	6:43412865	T/C	Missense variant, splice region variant	3118	2759	920	Ile/Thr	rs2125739	Tolerated (0.59)	Benign (0)
SLC22A1	6:160543008	C/T	Missense variant, NMD transcript variant	41	41	14	Ser/Phe	rs34447885, CM072043	Tolerated (1)	Benign (0)
SLC22A1	6:160557643	C/T	Missense variant, NMD transcript variant	1022	1022	341	Pro/Arg	rs2282143, CM041849	Deleterious (0)	Possibly damaging (0.72)
SLC22A1	6:160557643	C/T	Missense variant, NMD transcript variant	1022	1022	341	Pro/Leu	rs2282143, CM041849	Deleterious (0)	Probably damaging (0.96)
ABCB1	7:87138659	A/T	Missense variant	3229	3229	1077	Ser/Thr	rs2229107	Tolerated (0.43)	Benign (0.01)
ABCB1	7:87190625	T/C	Missense variant	589	589	197	Ile/Val	rs36008564	Tolerated (0.08)	Benign (0.116)
ABCC2	10:101557080	C/T	Missense variant	941	859	287	Leu/Phe	rs113972365	Deleterious (0)	Probably damaging (0.999)
ABCC2	10:101560169	G/A	Missense variant	1171	1058	353	Arg/His	rs7080681	Tolerated (0.16)	Benign (0.009)
ABCC2	10:101563815	G/A	Missense variant	1362	1249	417	Val/Ile	rs2273697, CM067967, COSM1179584	Tolerated (0.18)	Benign (0.044)
ABCC2	10:101590186	C/T	Missense variant	2856	2743	915	Arg/Cys	rs112758556	Tolerated (0.05)	Benign (0.009)
ABCC2	10:101590538	G/A	Missense variant	2926	2813	938	Ser/Asn	rs112339372	Tolerated (0.39)	Benign (0.002)
ABCC2	10:101591818	A/G	Missense variant	3301	3188	1063	Asn/Ser	rs17222540	Tolerated (0.1)	Benign (0.439)
ABCC2	10:101611294	G/A	Missense variant	4657	4544	1515	Cys/Tyr	rs8187710	Tolerated (0.31)	Benign (0)
SLC22A6	11:62749344	G/A	Missense variant	1060	767	256	Ala/Gly	rs11568624	Deleterious (0)	Benign (0.24)
SLC22A6	11:62749344	G/A	Missense variant	1060	767	256	Ala/Val	rs11568624	Tolerated (1)	Benign (0.001)
SLC22A6	11:62752014	C/T	Missense variant	442	149	50	Arg/Pro	rs11568626	Tolerated (0.07)	Benign (0.046)
SLC22A6	11:62752014	C/T	Missense variant	442	149	50	Arg/His	rs11568626	Tolerated (0.05)	Probably damaging (0.91)
SLC22A11	11:64338485	G/A	Missense variant	433	433	145	Ala/Ser	rs145855812	Tolerated (0.33)	Benign (0.085)
SLC22A11	11:64338485	G/A	Missense variant	433	433	145	Ala/Thr	rs145855812	Tolerated (0.52)	Benign (0.085)
ABCC4	13:95815454	T/C	Missense variant	2124	2005	669	Met/Val	rs9282570	Tolerated (0.32)	Benign (0.007)
ABCC4	13:95839008	T/C	Missense variant	1386	1267	423	Lys/Glu	rs11568669	Tolerated (0.81)	Benign (0.19)
ABCC4	13:95847145	G/A	Missense variant	1102	983	328	Pro/Leu	rs11568705, COSM1165207, COSM1165206	Tolerated (0.09)	Benign (0.009)
ABCC4	13:95859035	C/A	Missense variant, splice region variant	806	687	229	Lys/Asn	rs2274407	Deleterious (0.01)	Benign (0.3)

c) Novel variant identification in four genes and their classification

In order to determine the proportion of variants that are novel, known variants present in the KGP phase 3; dbSNP common 149, NCBI (dbSNP 149) and Exome Aggregation Consortium (ExAC) VEP annotation 0.3 databases were filtered out. Initial filtering to identify novel variants included filtering variants in the KGP database where 51 novel variants were identified from 275 high-quality variants identified in eight genes implicated TDF induced AKI. The next filtering step was filtering out variants in the dbSNP 149 database and 51 novel variants were identified. The last step involved filtering variants using the ExAC VEP annotation 0.3 database databases. A total of 40 novel variants were identified and classified in terms of the gene and location relative to the gene and class (Figure 2.13). Therefore 235 know variants were identified in the three commonly used databases selected and were filetered out to potentially identify novel variants associated with TDF induced AKI and efficacy.

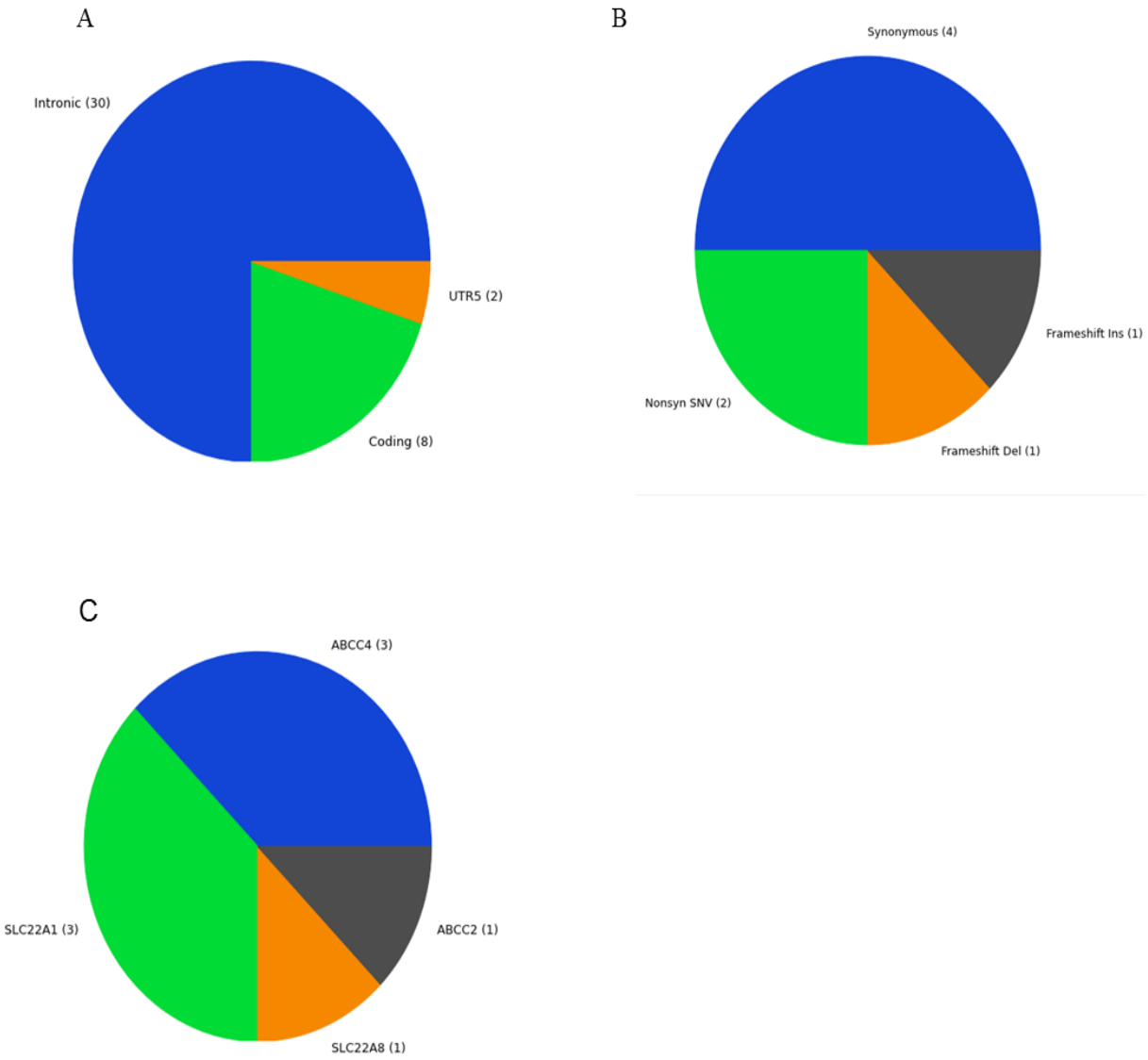


Figure 2.13. Novel variants classification. Pie chart (**A)** illustrates the classification based on the novel variant's position relative to the transcript of each variant (n= 40). (**B**) Illustrates the nature of the eight novel coding variants and (**C**) shows genes and the corresponding number of variants that lie in each respective gene.

2.2.4 Panel of polymorphisms selected for genotyping the TDF cohort

In order to select and prioritise a panel of polymorphisms that would be genotyped in the TDF cohort, further selection criteria, in addition to variant prevalence and novelty were applied. These included whether or not the variants result in functional changes and whether or not the variants have previously been associated with TDF treatment outcome. Based on this filtering step a total of seven prioritised variants were selected, as described in Table 3.3. These variants have all previously been shown in other populations to be associated with different aspects of TDF-induced nephrotoxicity (Table 3.3). The variants were identified in

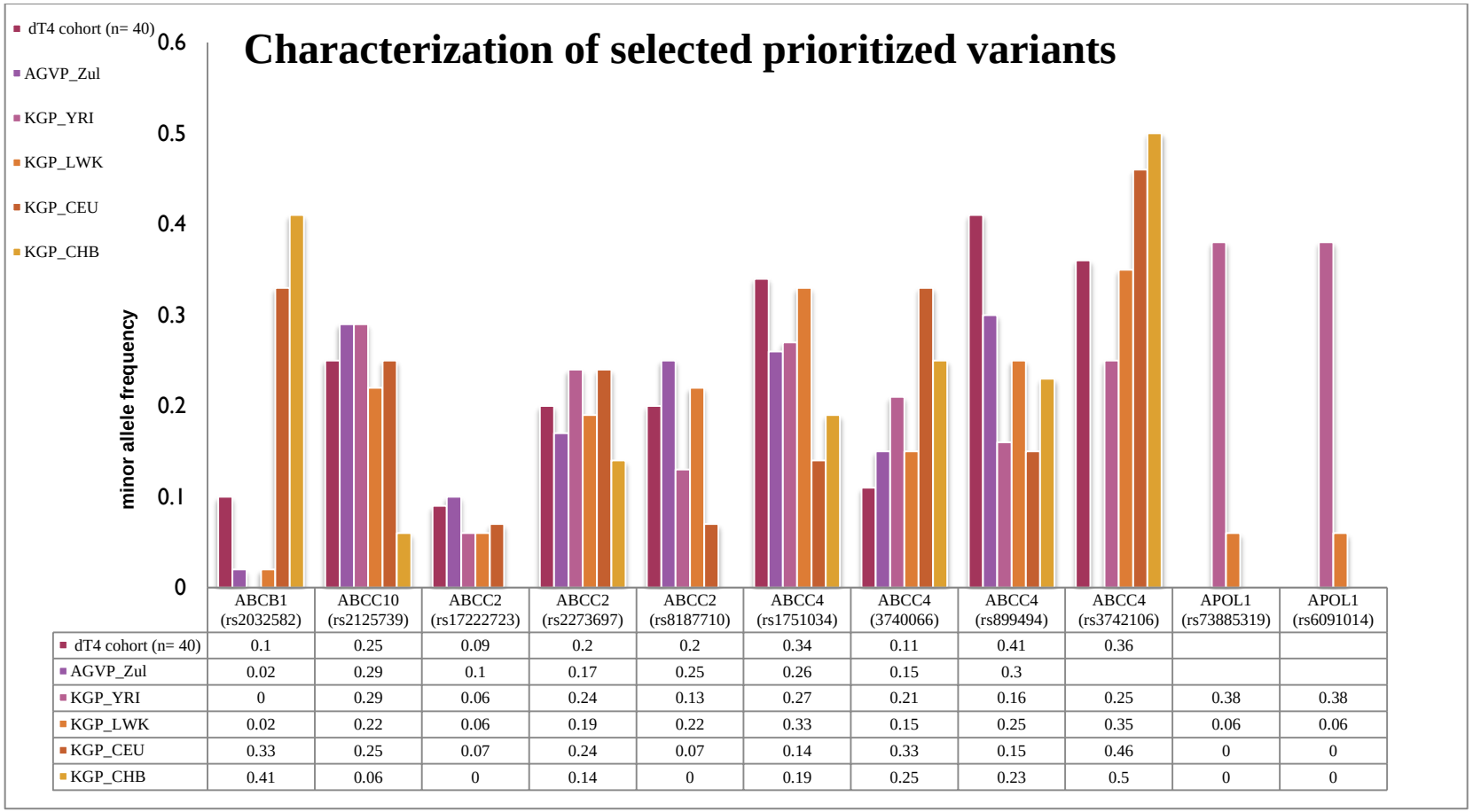
three different ethnic groups, namely Caucasian, American and Japanese. Two protective *ABCC2* alleles (3563A - rs17222723 and 4544A - rs8187710) (Izzedine *et al.* 2006) were observed to occur at a lower frequency in the cohort we sequenced (compared to the European population), warranting further investigation in the TDF cohort. The two protective variants were in linkage disequilibrium (LD); however this may not be the case as they occurred at different allele frequencies in our cohort. With the aim of selecting a more protective variant, ABCC2 3563A allele (rs17222723) was preferred as it occurs at a lower frequency in our cohort.

A commonly associated variant (*ABCC2* -24 C/T), along with two potentially novel variants, were excluded from the genotyping panel as Sanger sequencing indicated that they were not prevalent or were not observed in our sample population (Appendix E). Another commonly associated variant, *ABCC2* 1249 G/A (rs2273697), was identified in our population and occurs at a similar frequency compared to other populations (Table 2.11 and Figure 2.14). The African populations (our sequenced d4T cohort, Bantu-speaking (AGVP) and KGP's African populations) showed higher frequencies for most of the prioritised variants than the global population (KGP) (Figure 2.14). The variant allele frequencies were compared among different populations and the most frequent variant in our sequenced cohort was *ABCC4* 669 C/T (rs899494). This variant has been shown to be associated with increased risk of TDF-AKI (phosphorous wasting). Unfortunately, it was excluded from the final genotyping panel because it failed the QC steps during manufacturing.

Table 2.11. Nine variants previously reported to be associated with TDF associated nephropathy

	Polymorphism ID	Coding	Genotype associated w/ outcome	Protein	Clinical outcomes	Population	MAF in the AKI cohort	MAF in our cohort (n=40)	References
MPR2 (*ABCC2*)	rs2273697	c.1249G>A	GA/AA	p.Val417Ile	Associated with TDF-AKI (Aminoaciduria)	European	0.18	0.20	**Rodriguez et al 2009 & Pushpamkom et al 2011**
	rs17222723	c.3563T>A	AA	p.Val1188Glu	More frequent in the control group (protective)	29 European &1 African	0 (0.235 in the control group)	0.09	**Izzedine et al 2006**
	rs8187710	c.4544G>A	AA	p.Cys1515Tyr	More frequent in the control group (protective)	29 European &1 African	0 (0.235 in the control group)	0.2	**Izzedine et al 2006**
	rs3740066	c.3972C>T	CC (WT)	p.	Associated with TDF-KTD (β2- microglobulinuria)	European	0.68	0.11*	**Rodriguez et al 2009and Izzedine et al, 2006**
MPR4 (*ABCC4*)	rs899494	c. 669C>T	CC (WT)	p.	Increased risk of TDF-AKI (phosphorous wasting)	European	0.92	0.41	**Rodriguez et al, 2009**
	rs3742106	c.4131T>G	TG/GG	p.	Associated with TDF-KTD (altered uric acid excretion)	European	0.45	0.36*	**Rodriguez et al 2009**
	rs1751034	c.3463A>G	GG	p.	Higher intracellular TFV-DP (35%)	American (w/ 6 African)	0.201	0.34	**Kiser et al 2008**
MPR7 (*ABCC10*)	rs2125739	c.2759T>C	CC	p.Ile948Thr	Marginally associated with TDF-AKI	European (limited sample size)	0.388	0.25	**Pushpamkom et al 2011**
PgP-1 (*ABCB1*)	rs2032582	c.2677T>A /G	AA	p.Ser893Thr (A), p.Ser893Ala (G)	Significantly associated with TDF-AKI	Japanese	0.26	0.10	**Nishijima et al 2012**

MAF= Minor Allele Frequency
*Allele frequency determined using Golden Helix Genome Browser (GHGB)
Variants highlighted in purple are the variants that were prioritized for genotyping the tenofovir cohort..



	ABCB1 (rs2032582)	ABCC10 (rs2125739)	ABCC2 (rs17222723)	ABCC2 (rs2273697)	ABCC2 (rs8187710)	ABCC4 (rs1751034)	ABCC4 (3740066)	ABCC4 (rs899494)	ABCC4 (rs3742106)	APOL1 (rs73885319)	APOL1 (rs6091014)
dT4 cohort (n= 40)	0.1	0.25	0.09	0.2	0.2	0.34	0.11	0.41	0.36		
AGVP_Zul	0.02	0.29	0.1	0.17	0.25	0.26	0.15	0.3			
KGP_YRI	0	0.29	0.06	0.24	0.13	0.27	0.21	0.16	0.25	0.38	0.38
KGP_LWK	0.02	0.22	0.06	0.19	0.22	0.33	0.15	0.25	0.35	0.06	0.06
KGP_CEU	0.33	0.25	0.07	0.24	0.07	0.14	0.33	0.15	0.46	0	0
KGP_CHB	0.41	0.06	0	0.14	0	0.19	0.25	0.23	0.5	0	0

Figure 2.14. Characterisation of 11 prioritised variants. Nine variants have been associated with TDF associated acute kidney injury (TDF-AKI) (Table 2.11), and two additional variants in *APOL1* strongly associated with HIV-associated nephrotoxicity (HIVAN) are included. The risk alleles of TDF-AKI occur more frequently in the African populations and the *APOL1* variants are rare in the Caucasian (CEU) and Chinese (CHB) populations.

The variants previously associated with TDF induced AKI were cross-checked against the Pharmacogenetics mutation database (PGMD^{TM}), which is a large manually curated database of pharmacogenetics traits based on public databases and literature. The PGMD^{TM} database contained 10 genetic markers that are reported to be associated with TDF associated nephrotoxicity (TDF-AKI), eight variants and two haplotypes. This included ABCB1 variant 2759T/C (rs2125739), *ABCC2* variants 1249 G/A (rs2273697), 3972 C/T (rs3740066), 3563 T/A (rs17222723) and *ABCC4* variants 3463 A/G (rs1751034) and 669 C/T (rs899494). The haplotypes identified were *ABCC10* haplotype GGC [positions intron 4 or c.1791+526 (rs9349256), 2759 (rs2125739) and c.2044-182 (rs2487663)] and *ABCC2* haplotype CATC [positions -24 (rs717620), 1249 (rs2273697), 3972 (rs3740066) and 3563 (rs17222723)] (Figure 2.14)



Gene	Variation	dbSNP ID	Phenotype	Study Disease	HGMD	PMID
ABCC10	Ile948Thr	rs2125739	• Renal insufficiency	HIV Infections		57
ABCC10	c.1875+526G>A, c.1791+526G>A	rs9349256	• Renal insufficiency	HIV Infecti		57
ABCC10	View haplotype report c.1875+526G>A, c.1791+526G>A Ile948Thr c.2128-182A>G, c.2044-182A>G	rs9349256 rs2125739 rs2487663	• Renal insufficiency	HIV Infections		57
ABCC2	Val417Ile	rs2273697	• Proximal tubulopathy	HIV Infections	CM067967	96
ABCC2	Cys1515Tyr	rs8187710	• Proximal tubulopathy	HIV Infections	CM119536	96
ABCC2	c.-24C>T	rs717620	• Nephrotoxicity	HIV Infections	CR076688	80
ABCC2	Val1188Glu	rs17222723	• Proximal tubulopathy	HIV Infections	CM065960	96
ABCC2	View haplotype report c.-24C>T Val417Ile Ile1324Ile,Ile1324Met Val1188Glu	rs717620 rs2273697 rs3740066 rs17222723	• Proximal tubulopathy	HIV Infections		96
ABCC4	Lys1116Lys	rs1751034	• Accumulation of drug or metabolite	HIV Infections		87
ABCC4	Ile223Ile	rs899494	• Proximal tubulopathy	HIV Infections		96

Showing 1 to 10 of 10 entries First Previous 1 Next Last

Figure 2.15. An illustration of a PGMD^{TM} report. PGMD^{TM} identified ten variants reported to be associated with TDF associated nephrotoxicity, six of which were selected for genotyping in this study (outlined).

These were also cross-checked against the pharmGKB, pharmacogenetics knowledgebase, another manually curated database of pharmacogenetics traits based on public databases and literature. The database also provides clinical annotations, dosing guidelines and pharmacokinetics or pathways the drugs take thus providing pharmacogenetic knowledge regarding the impact of pharmacogenetics traits on drug response for clinicians and researchers. The database identified 10 variants and two haplotypes that were statistically associated with TDF-AKI. Thus, the database identified one additional variant associated with TDF-AKI. This variant was recently published [*ABCC4* 4976 T/C (rs1059751)] (Likanonsakul *et al.* 2016) but was not identified in our study population. Additional variants contained within these databases that were not selected for genotyping in this study were *ABCC10* variant c.1791+526 G/A (rs9349256), *ABCC2* variants -24 C/T (rs717620), c.2044-182 A/G (rs2487663) and 4544 G/A (rs8187710) and *ABCC4* 4976 T/C (rs1059751). These variants were excluded from the genotyping panel because of low allele frequency (-24 C/T and c.1791+526 G/A) and some were excluded because they were not observed in our cohort (ABCC4 4976 T/C and c.2044-182 A/G). Variant *ABCC2* 4544A was excluded because ABCC4 3563A was preferred, as described above. *ABCC4* 669 C/T (rs899494) was excluded from the final panel because a TaqMan assay probe to genotype this variant could not be manufactured.

The variants used for genotyping the TDF cohort in the case-control association study (highlighted in Table 2.11) included three *ABCC2* variants (rs2273697, rs3740066 and rs17222723), two *ABCC4* variants (rs1751034 and rs3742106), an *ABCB10* variant (rs2125739) and an *ABCB1* variant (rs2032582). In addition to the seven variants, two *APOL1* variants (1072 A/G and 1098 T/G) were included because they had previously been reported to be significantly associated with HIVAN in the African population (Behar *et al.* 2011, Tzur *et al.* 2012, Ko *et al.* 2013, Ulasi *et al.* 2013). The effect size for association with kidney disease for these variants was found to be higher in the black South African population compared to the African American population (Kasembeli *et al.* 2015a). Thus, the effect of these variants may also be relevant to ARV-induced nephrotoxicity. The genotyping panel therefore included nine SNPs in five genes, show in Figure 2.14.

2.3: DISCUSSION

HIV ARVs are highly effective and prevent patient progression to AIDS. Trea[tm]ent is based on highly active antiretroviral therapy (HAART) combination therapy and has improved patient prognosis. Some of the drugs used in HAART are, however, not available in developing countries. South Africa has the largest ARV program in the world with 3.5 million people receiving treatment (Reid 2016); due to the widespread use of HAART drugs hospital admissions due to adverse drug reactions (ADRs) are notably increased (Mouton *et al.* 2015, Peter *et al.* 2015, Mouton *et al.* 2016, Njuguna *et al.* 2016). Specifically, the use of TDF is commonly implicated in these adverse outcomes. It is well established that human genetic variation can influence drug outcome, but pharmacogenetic data is limited for African populations. Some research is now being undertaken in African populations and there is increasing evidence that points to notable differences compared to other populations in the metabolism and efficacy of drugs due to ethnic-specific genetic differences. In particular, higher allele frequencies of variants of clinical significance in ART, including TDF-induced nephrotoxicity (which presents clinically as AKI), have been reported, warranting their further investigation in local populations. African populations in general show higher genetic diversity offering the opportunity to uncover novel variants and to perform research to establish associations between such variants and outcome of treatment.

Given the significant differences observed across populations in variants associated with drug therapeutic outcome, the aim of the study was to comprehensively identify variants in eight TDF associated ADME genes to enable the selection of a subset of population-relevant SNPs for an association study of TDF-associated nephrotoxicity.

2.3.1 Next generation sequencing of a Bantu speaking cohort

Given the potential likely variant and variant allele frequency differences within the population used in this study compared to publicly available data, targeted sequencing of the genes in an existing South African cohort of Bantu ancestry was first undertaken. NGS technology was employed to comprehensively map (at high coverage) and characterise common and novel variants in eight genes implicated in TDF efficacy and nephrotoxicity in this population. In the experiment, data was generated on a total of 65 genes, but for the purpose of this dissertation the data on only eight of the genes are described. The 65 genes, including those implicated in TDF efficacy and nephrotoxicity, were investigated. The genes

included drug transporters, drug metabolising enzymes, genes implicated in hypersensitivity, kidney function, kidney development and kidney failure biomarkers. An amplicon library of the targeted regions within the 65 genes was prepared for each individual patient. The sequencing libraries were prepared for sequencing by attaching and creating monoclonal copies on Ion sphere particles (ISPs). The libraries prepared were sequenced using two Ion Torrent platforms (i.e. Ton PGM™ and Ion S5™).

a) Ion PGM™ vs Ion S5™

Two sequencing platforms, each using its own chip, were used. The templated ISPs were initially loaded onto a 318™ chip and massively parallel sequencing was performed on the Ion PGM™ sequencer. The initial chip setup using the manufacturer's recommendations (2 µl of 200 pM pooled libraries was added during the template preparation step) performed poorly on the Ion PGM™ sequencer and was subsequently omitted from data analysis. Per manufacture's recommendations, the library input was increased to target 25% templated ISPs, this was done by adding 25 µl of 200 pM pooled libraries and the chips performances improved. Three 318™ sequencing chips was used to sequence the 40 patient samples and they generated an average of 773.14 Mbp of data. To extract more data from the sequencing libraries, libraries were sequenced on an Ion Torrent S5™ sequencer, using Ion 530™ chips. As expected, the Ion 530™ chips performed much better than the Ion 318™ chips. With the 530™ chips generating 3.94 Gbp of data and obtaining an average sequencing depth of 820X.

NGS data generated from the Ion Torrent S5 and Ion PGM™ (Ion Reporter) was compared using SVS 8 (and Golden Helix Genome Browser) and both platforms generated similar results, i.e. call rates and allele frequencies. However, not surprisingly the Ion S5™ generated more variants (1 998 compared to the 852 variants generated by PGM™) and higher sequencing depth. This indicates that the regions targeted were covered much more deeply with the Ion S5, giving greater confidence in these results. The 1 146 variants not detected by Ion PGM™ sequencing was possibly due to poor quality data and the stringent QC process of variant calling and these variants failed to meet the cut-off or threshold. Another reason for the discrepancies is the increased chip capacity that allowed for more of the templated ISPs to be captured on the 530™ chips, resulting in increased coverage that meets the stringent QC process.

Ion PGM™ sequencer generated 234 variants that were not detected bu the Ion S5™ sequencer and this included 89 high quality variants, potential loss of functions (LoF), novel variants and six variants localised in TDF implicated genes. However, the data was not credible as a few of the variants identified by sequencing on the Ion PGM™ sequencer were evaluated by Sanger sequencing to validate and confirm their genotypic calls. Sanger sequencing was performed on the patients that possessed two novel variants in *ABCC2* and *ABCB1* genes and an *ABCC2* variant (rs717620) commonly implicated in TDF associated AKI (Appendix E). Sequencing analysis indicated that the calls were incorrect and the novel variants were not detected in our cohort. This cast doubt on the accuracy of some of the relatively low coverage data generated by the Ion PGM experiment. Sanger sequencing indicated that the *ABCC2* variant (rs717620) was present in our target population, however it was rare. The incorrect variant calls identified on the Ion PGM™ led to a decision to use only high coverage data generated from the Ion S5™ for downstream analysis and identifying variants to prioritise to screen the TDF cohort.

b) Variant discovery and characterization

Since the PGM™ generated data were unreliable, only Ion S5™ generated data were used to select variants for further analysis. The 1 687 high quality variants identified by the latter have been comprehensively analysed in a manuscript that is under revision (Appendix D). This manuscript discusses the implications of findings for future pharmacogenetic studies and strategies aimed at improving treatment outcome amongst patients of Bantu-ancestry.

My study was a subset of a bigger study and focused on genes implicated in TDF ADRs and efficacy. As a consequence, a subset of eight genes out of 65 were selected for downstream analysis. Gene panel filtering resulted in 275 variants that lie in genes associated with TDF absorption (*ABCB1*), elimination (*ABCC2*, *ABCC4*, *ABCC10*, *SLC22A6* and *SLC22A8)* and transport (*SLC22A1).* No variants were detected in the transporter gene *SLC22A11*. The 275 variants were further characterised and analysed for variants of functional significance and novelty. The functional impact of these variants was evaluated, via bioinformatic analysis, by first looking at coding variants of functional significance such as non-synonymous and loss of function (LoF) variants. This resulted in 86 coding variants (including a splicing variant) being identified and evaluated. From these 41 variants were predicted to potentially be

functional. This included three LoF variants, a frameshift insertion, a frameshift deletion and a splicing variant.

In order to identify potentially novel variants, KGP phase 3, dbSNP 149 and ExAC databases were employed to filter out variants identified in these three databases. This resulted in the identification of 40 potentially novel variants; three of the variants were LoF variants. The LoF and novel variants were not considered in the genotyping panel because of their low call rate, as the genotype calls would require validation using Sanger sequencing to ensure that they are not artefacts of the variant calling algorithm. This series of experiments was not performed due to time and cost constraints.

The NGS data provided more insight toward the genetic variation in ADME genes in a Bantu speaking cohort and concurs with recent studies which indicate that African populations reflect greater genetic diversity as compared to other populations (Aminkeng *et al.* 2013, Auton 2015). Genetic diversity provides an opportunity to discover entirely novel polymorphisms that affect therapeutic outcome and such variants were identified in patients of Bantu ancestry.

c) Prioritized panel of SNPs

For prioritisation of the panel of polymorphisms used to genotype the TDF cohort, further selection criteria, in addition to variant prevalence and novelty, included whether the variants result in functional changes and whether the variants have previously been associated with TDF treatment outcome. Based on the sequencing data from this study, nine variants that have previously been identified to be important in TDF induced nephrotoxicity in Caucasian, Japanese and American populations were confirmed to be present in the Bantu-speaking population. From the nine variants, two were excluded because one *ABCC2* (rs8187710) variant was in LD with another prioritised variant and was associated with the same outcome. The other variant (*ABCC4,* rs899494) failed to meet manufacturing QC requirements. Two additional variants were added to the remaining seven SNPs on the panel because of their strong associations with HIVAN, as confirmed in a South African cohort (Kasembeli *et al.* 2015a). This resulted in a final SNP genotyping panel of nine SNPs. The allele frequencies were compared with populations commonly used as references for populations of African, Caucasian and Chinese ancestry. This included Caucasian population and a Chinese population from the KGP, two African populations from the KGP (YRI and LWK) and a

Bantu speaking population (Zulu) from the African Genome Variation Project (AGVP) The African populations showed higher frequencies for most of the identified variants than the global population (Figure 2.11). Should the association with kidney disease be confirmed in Africans, it would suggest a potentially important genetic predisposition of African populations to nephropathy, HIVAN, and more significantly TDF associated nephropathy.

The most common variant implicated in TDF nephrotoxicity, *ABCC2* -24 C/A, was confirmed to be a rare variant in Bantu-speakers and as such was excluded from the final panel. The relatively low frequency suggests that in the broader population the effect of this variant may not be as significant (as other variants prioritised). In contrast, another commonly associated variant in this gene, 1249 G/A (rs2273697) was found to occur at similar prevalence (MAF=0.2) to that reported in other populations such as populations of Cape mixed ancestry (CMA) origin (MAF=0.18); this variant is prevalent at a relatively low frequency in the Xhosa (MAF=0.1) (Ikediobi *et al.* 2011) and CEU populations (MAF=0.14) while amongst the LWK population its frequency is relatively high (MAF=0.33). Thus, there would be merit in prioritising the analysis, in relation to AKI, of this variant in the broader population to establish its potential clinical relevance.

Likewise, a variant located in the gene *ABCC4*, 3463 T/C (rs1751034), which is associated with increased intracellular concentration of TDF (Kiser *et al.* 2008a) was also found, based on this study's sequencing data, to occur at a similar frequency (approximately 0.2) compared to the other populations, with the exception of the CEU population in which the MAF was 0.14. A pharmacogenetic study by Aminkeng and colleagues indicated that this *ABCC4* variant is more frequent in African populations as compared to Caucasian populations, however West Africa (Nigeria) has the highest frequency (MAF=0.44), Saharan Africa (Sudan) (MAF=0.24), with Southern Africa (SA) (MAF=0.29) exhibiting the lowest frequencies amongst the African populations (Aminkeng *et al.* 2013). This second variant thus represents another candidate for further studies to determine the potential effect of such variants on TDF outcome. The sequencing findings from this study to confirm the potential existence and establish the prevalence of important TDF induced AKI variants in the Bantu speaking population concur with findings from other pharmacogenetic studies conducted in African populations. There was therefore sufficient motivation to proceed to genotype a number of these in a larger cohort drawn from the same population.

The potential functional impact of the nine variants was evaluated using the algorithms SIFT and PolyPhen. Bioinformatic analysis indicated that *ABCC2* 3972 C/T (rs3740066) is deleterious and probably damaging while the remaining variants were classified as tolerated or benign. Of these, *ABCC10* 2759 T/C (rs2125739) is however a missense variant located in a splice site and may affect pre-mRNA splicing potentially leading to an altered protein. Thus, such variants are of potential relevance to patients receiving TDF.

d) Study limitations

Several limitations to this part of the study were identified. With regards to the variant discovery NGS, a major limitation was that only coding regions were prioritised during AmpliSeq^{TM} primer design for sequencing the 40 individuals, as such several potentially relevant (located in intronic, upstream, downstream, regulatory and untranslated regions) SNPs associated with TDF induced renal failure would have been missed. Such variants include *ABCC10* variant c.1791+526 G/A (rs9349256) (an intronic variant), associated with urine phosphorous wasting and β_2 microglobulin (β_2m) (Pushpakom *et al.* 2011), *ABCC2* variant -24 C/A (an intronic variant associated with phosphorous wasting and β_2m (Rodríguez-Nóvoa *et al.* 2009, Nishijima *et al.* 2012), and *SLC22A11* 64535478 T/A (rs11231809), a downstream variant associated with altered uric acid excretion (Rodríguez-Nóvoa *et al.* 2009). *ABCC4* (4976 T/C (rs1059751)) is a 3' UTR variant associated with proteinuria (β_2m) (Likanonsakul *et al.* 2016). These variants (including an intronic *ABCC10* variant, c.2044-182 A/G (rs2487663)) are commonly evaluated for association with TDF-AKI but are yet to be comprehensively tested in patients drawn from the Bantu-speaking population.

Novel and LoF variants were not considered because of their low call rate and frequency. The call rates could be increased by preparing new libraries for the same 40 patients sequenced in the d4T cohort as this could result in libraries which cover the targeted regions of interest for all the sequenced samples. There is a possibility that for some patient samples the amplicons that contain the novel and the LoF variants were not captured due to chip capacity. Sequencing the new libraries on a chip with more capcity or reducing the number of patient libraries per chip will subsequently increase the sequencing depth or coverage. Increasing the

chip capcity will not only futher increase the robustness of the genotypic calls, but could validate novel and LoF variants. Novel and LoF variants can also be validated by sequencing a bigger Bantu speaking cohort to not only identify and validate the novel and LoF variants, but could possibly result in the identification of more novel and LoF variants. These variants offer the opportunity to establish novel associations with TDF related nephropathies.

Despite these limitations, nine variants previously associated with TDF-AKI were identified in our Bantu speaking (d4T) cohort, thus warranting further investigation in a case-control association study to assess the potential association of these variants in a Bantu-speaking population.

CHAPTER 3: GENE VARIANT SCREENING FOR VARIANTS IMPLICATED IN TDF INDUCED AKI USING A CASE-CONTROL STUDY DESIGN

3.1 METHODOLOGY

3.1.1 Study participants

A representative cohort of patients who were undergoing TDF treatment was recruited to conduct a case-control association study to assess the potential association of genetic variants with TDF associated nephropathy.

The TDF cohort included patients from the inpatient and outpatient wards of the Tshepong (Klerksdorp) hospital and surrounding local clinics. The cases included black adult patients undergoing TDF combination therapy, who at the time of blood sampling presented with clinically-defined symptoms linked to drug-induced nephropathy, amongst other adverse effects. The controls included black adult HIV-positive patients undergoing TDF combination therapy, who at the time of blood sampling presented with no symptoms linked to drug-induced nephropathy and were considered otherwise clinically healthy. The age and sex distribution of cases and controls is shown in the results section and is roughly similar. Detailed clinical history and demographic information was collected for each patient by the residing clinician (see clinical response form in Appendix B), together with a quantitative and qualitative dataset of clinically pertinent variables. These variables included CD4 count, viral load counts, adverse effect assessment, concomitant medications, co-morbidity, and assessment of ARV metabolic function, effectiveness of treatment and adherence. Treatment-related nephropathy was defined by the collaborating clinicians according to the Kidney Disease Improving Outcomes (KDIGO) definition of acute kidney injury (AKI). This includes one or more of the following: an increase in serum creatinine by more than 1.5 times the baseline; creatinine increases >26.4umol/l above the upper limit of normal, urine output <0.5ml/kg/hr for 6 hrs, and defined reductions in eGFR. eGFR for each patient was calculated using the Modification of Diet in Renal Disease (MDRD) formula and is based on serum creatinine level (Creatinine), age, sex and race (eGFR= 186 x (Creatinine/88.4)-1.154 x (Age)-0.203 x (0.742 if female) x (1.210 if black)) (http://egfrcalc.renal.org/ accessed 28/02/17). Clinical variable cut-off and end-point values were defined according to existing rating scales in practice at the participating sites, for example AKI severity was calculated based on KDOQI (Kidney Disease Outcomes Quality Initiative) stages of kidney disease (i.e.

eGFR) (http://www.renal.org/information-resources/the-uk-eckd-guide/ckd-stages#sthash.KDtVYV37.dpbs accessed 28/02/17).

3.1.2 gDNA isolation and genotyping

SNP genotyping was performed to enable a case control association study of TDF-related nephrotoxicity. Nine SNPs were prioritised and genotyped in each individual. Genotype and allele frequencies were determined for the cases (n=53) and controls (n=84). TaqMan® SNP genotyping assays (Life Technologies, California, USA) were used to genotype the 137 patient samples, using Real time PCR. Each assay consisted of a forward primer, reverse primer and two TaqMan probes (specific to each allele) labelled with a fluorescent reporter dye and a quencher molecule (based on the principle of fluorescence resonance energy transfer (FRET), a low energy quencher that suppresses the signal of the high-energy reporter). The technology is based on the 5' nuclease activity of the Taq polymerase enzyme during DNA amplification (PCR). As the forward primer extends the template, Taq polymerase degrades the hybridized probe. This results in the release of the florescent dye and the signal is detected in real time. With successive amplification cycles, the signal intensifies overtime and a specific amplification signal is recorded. Sanger sequencing of patients with ambiguous or undetermined calls was employed for further investigation and confirmation of data.

a) Sample collection and genomic DNA isolation

Blood samples were obtained with consent from each participating patient, using conventional venepuncture. To allow optimal isolation of the gDNA to be used downstream, the salting out method of extracting nucleic acid from nucleated cells (buffy coat) was used (Miller et al., 1996). This was the preferred method as it produces good quality, pure DNA of high yield (refer to Appendix F and G).

Buffy coat samples were separated from whole blood by a member of the clinical team at the sites where the blood was collected. The red blood cell (RBC) membrane was lysed with the sucrose-triton-X lysis buffer. The supernatant (SNF) containing lysed RBCs was discarded and the reddish white pellet containing the nucleated white blood cells was kept and washed 2-3 times with the sucrose-triton X lysis buffer, incubating at -20°C for 5 minutes each time.

To digest the white blood cell membranes and to degrade the protein, 3 ml of TE solution ($T_{20}E_5$), 0.2 ml of 10% SDS (w/v) (Sigma-Aldrich, Missouri, United States) and 0.5 ml of proteinase K solution (2 mg/ml proteinase K (Merck, Darmstadt, Germany), 1% SDS (w/v) and 2 mM EDTA (Sigma-Aldrich, Missouri, USA) freshly prepared) were added and incubated overnight at 42 °C. To salt-out or precipitate the proteins and debris, 1 ml of saturated NaCl (40%, w/v) (Sigma-Aldrich, Missouri, USA) was used. After spinning for 10 mins @ 2 400 rpm and 4°C, the resultant SNF containing the gDNA was transferred to a new 50 ml Falcon tube and the precipitated protein pellet was left behind. To precipitate the DNA, two volumes of absolute ethanol were added and samples were agitated gently until the DNA precipitated out. The visible gDNA was then spooled out with a pipette and transferred to a 1.5 ml microcentrifuge tube containing 70% ice-cold ethanol (v/v) to wash out excess salts that could inhibit subsequent activities. The DNA pellet was air dried for approximately five mins, to avoid over drying, before being resuspended in 50-500 µl of TE buffer, depending on the size of the pallet, and stored at –20°C until required for genotyping.

The quality and quantity of each sample was assessed using the NanoDrop™ and the Qubit® fluorometer. The NanoDrop™ was used to assess the quality and quantity of the gDNA extracted and the Qubit® fluorometer was used to more accurately quantify the double stranded DNA. The samples were diluted to 2 ng/µl, and stored for sequencing and genotyping.

b) Genotyping TDF cohort samples

- *Real Time PCR and TaqMan® genotyping*

Genotyping of the final prioritised panel of nine SNPs was performed using TaqMan® probes (Table 3.1) and primers on a TaqMan® 7500 FAST system (Life Technologies, California, USA). A TaqMan® genotyping assay mix (which contains probes for each allele) for each SNP probe set and TaqMan® genotyping assay master mix (which contains the forward and reverse primers, dNTPs and the Taq polymerase) was used to make the master mix. Spare DNA obtained from the d4T cohort samples was initially used to assess and optimise each assay and known genotypes were used as positive controls to validate each assay (Appendix H and I). A master mix containing the genotyping assay master mix, genotyping assay mix and 10 ng of gDNA was prepared and aliquoted into each well with

volumes shown in Table 3.2. After optimization for four assays, 10 ng of gDNA was deemed sufficient to produce genotyping results (Appendix H).

Table 3.1. List of the nine TaqMan® SNP genotyping assays used in this study

Gene name	TaqMan® SNP genotyping ID	Alleles	dbSNP ID	SNP locus (Hg19)
ABCB1	C__11711720C_30	C/A	rs2032582	chr7:87160618
ABCC10	C__16173668_10	T/C	rs2125739	chr6:43412865
ABCC2	C__22272980_20	G/A	rs2273697	chr10:101563815
ABCC2	C__25291743_30	T/A	rs17222723	chr10:101595996
ABCC2	C__11214910_20	C/T	rs3740066	chr10:101603957
ABCC4	C__8059522_1_	A/C	rs3742106	chr13:95673541
ABCC4	C__1901918_30	T/C	rs1751034	chr13:95714976
APOL1	C__89555688_10	T/C	rs60910145	chr.22:36662034
APOL1	C__98253221_10	A/G	rs73885319	chr.22:36661906

Table 3.2. TaqMan® SNP genotyping assay reaction setup

Reagents	Reaction concentration	1X reaction (µl)
2x TaqMan® SNP Genotyping assay mix	2X	6
20x TaqMan® SNP Genotyping assay master mix	20X	0.6
ddH_2O	-	0.4
gDNA	10 ng	5
Final volume		12

- Genotyping validation: Sanger sequencing

Sanger sequencing was performed for patients with ambiguous or undetermined calls. The region flanking the SNP was amplified using the primers described below, purified and sent for sequencing to Inqaba Biotech, Pretoria.

PCR reverse and forward primers (Table 3.3) were designed for the flanking sequence around the SNP. The sequences flanking the SNPs were retrieved from Ensembl (www. ensembl.org) and were used for primer design. PCR was used for DNA amplification and the amplicons were purified using a QIAquick PCR Purification kit (Qiagen, Hilden, Germany). PCR conditions were similar to the conditions used for validation of the NGS calls as described earlier (Table 2.5 and Table 2.6).

Table 3.3. Primers sequences and their expected amplicon size

SNP	Primer sequences (5'- 3')	Size (bps)
rs2032582	FWD: GGAAGGAGTGGAGAGAGAGTGAAAGTG REV: TGGCTTGAGGGCACAGAATATTTCCAGG	1295
rs2125739	FWD: GGCCTATGAGAGAAGCCCACACTAGGG REV: CTCCTTCTTCTGTTGCTTTCCTCCTTCCC	1476
rs2273697	FWD: GAGGAAGGTGGGAATAAGAGGAAAGAGTGG REV: GACTTCTGGCAGGACTGAGGACACTGTCC	922
rs17222723	FWD: CTCATTGGTCTCCTCCTCGG REV: CCAACTCTCCTCTCTCCAACC	857
rs3740066	FWD: GCCATTAGGTAGCTGGGACACTGC REV: GAAGTGAAGAAGGGAAAGATTAGGTCTGCC	792
rs1751034	FWD: GCCCTTCATGGAGGTAATCACCTTTGAG REV: CCCAGAGGACGTTTGTCAGTGACA	880
rs3742106	FWD GCATTTAACTGGTGGCCTGCACC REV: GTTTGGCATTCAGCCCAGCC	557
rs60910145 & rs73885319	FWD: CAGCAGTACCATGGACTACGG REV: GCCATCTGCATTAACCCC	1001

3.1.3 Statistical analysis

Statistical analysis was performed using XLSTAT (Microsoft Corporation, Washington, USA) for descriptive statistical analysis. SNP variation suite software (SVS 8.4.4) (Golden Helix^{TM}, Montana, USA) was used for association analysis. Hardy-Weinberg equilibrium (HWE) tests were used to assess the distribution of the genotypes and $p \leq 0.05$ was considered a significant deviation from HWE. To assess the association of genotypes, allele frequencies and haplotypes with TDF outcome, the differences between the cases and controls were evaluated using a Chi-squared ($\chi 2$) test at 5% level of significance ($p \leq 0.05$). The $\chi 2$ tests were corrected for multiple testing (Bonferroni's correction), where applicable.

3.2 RESULTS

3.2.1 Characteristics of the study population

The complete study cohort consisted of HIV-positive adult patients who presented at Tshepong hospital. The patients were grouped into those with renal failure (designated as the case group) and a control group of patients who had not developed renal failure while on the same treatment regimen. An attempt was made to match by age (within +/- 3 years) and sex. However, since patient recruitment was still ongoing at the time of writing up this dissertation, the matching was not identical and it is not possible to refer to this cohort as a matched case-control group in the strictest sense. The cohort that was collected up to the point of writing up this dissertation consisted of 314 HIV-positive patients. However, a proportion of patients were excluded from genotyping (56%, n=177) and case-control association analysis because they did not fulfil the inclusion criteria for this study, more specifically they were not receiving TDF but were on other ARVs (Figure 3.1).

In order to assess the potential influence of patient genotype on clinical outcome specifically during TDF treatment, the inclusion criteria for the study were as follows: Black, Bantu-speaking, adults (older than 18 years), HIV-positive patients and patients who were undergoing TDF combination therapy at the time of sample collection. The patients were on a fixed dose combination (FDC) therapy of TDF/3TC/EFV. In total, 137 patients met the criteria for inclusion in the study and DNA samples from these individuals were genotyped (Figure 3.1).

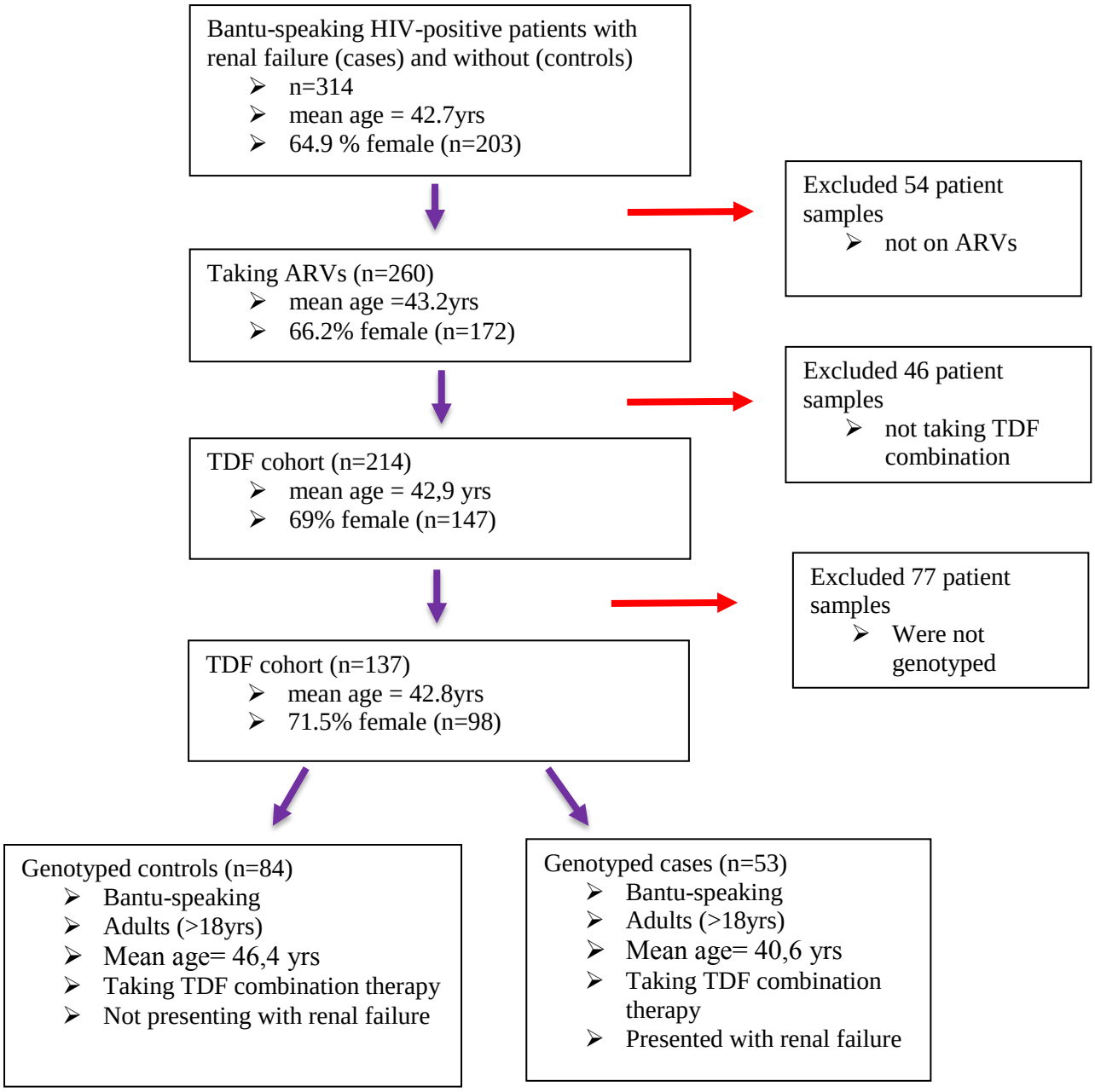


Figure 3.1. Summary of patients enrolled in the study.

Prior to statistical analysis, potential differences in the distribution frequencies between the cases and controls in sex and age were evaluated. To assess potential differences in sex, a $\chi 2$ test was employed and the outcome of this analysis indicated that the differences between the cases and controls were not statistically significant ($\chi 2=3.48$, $p=0.06$) (Table 3.4). Potential differences in age were assessed using the Mann-Whitney test (two-tailed test) and this analysis indicated that statistically significant differences exist between the mean ages of the cases and the control ($p=0.006$).

Table 3.4. Descriptive statistic of the cases vs the controls

Characteristics		Total (n=137)	Case (n=53)	Control (n=84)	Test significance (p value)
Sex*	**Female**	97 (71%)	33 (62%)	64 (77%)	0.060
	Male	39 (29%)	20 (38%)	19 (23%)	
Age		42.8 yrs (±10.8)	40.6 yrs (±12.3)	46.4 yrs (±9.2)	0.006

The differences between the sex and mean age of the cases and controls were evaluated using a $\chi 2$ test and a Mann-Whitney test (two-tailed) at 5% level of significance ($p \geq 0.05$). Standard deviation is given in the parenthesis.
*One genotyped control sample was not assigned a sex.

3.2.2 Genotyping results and association testing

A total of 137 patient samples were genotyped for nine variants using a TaqMan® 7500 FAST system and TaqMan SNP genotyping assays. Genotype results were not obtained for all 137 samples for all nine assays, since gDNA was depleted for some samples prior to the completion of the study. The numbers of samples genotyped for each SNP are shown in Table 3.6. The robustness of each assay was assessed and validated by comparing the genotypic calls obtained from NGS (Ion S5 sequencing), TaqMan® genotyping and Sanger sequencing (Appendix E) as well as comparing the genotypic calls from TaqMan® genotyping and Sanger sequencing (Appendix I). These findings are summarised in Table 3.5. There was excellent concordance between the genotypes when done using the three different methods for seven of the nine SNPs (Table 3.5 and Appendix H). As a result, not all samples needed to be validated for the assays. The genotyping results are shown in Appendix I and J. Tests for association between genotypes, alleles and haplotypes compared to categorical clinical data (case-control) were performed in SVS using $\chi 2$ tests.

Prior to association analysis, genotypes were tested for deviation from Hardy-Weinberg equilibrium (HWE). The distributions of the genotypes were evaluated for the cases and controls separately. For the cases, three SNPs, *ABCB1* 2677 C/A (rs2032582), *ABCC2* 1249 G/A (rs2273697) and *APOL1*1072 A/G (rs73885319), deviated from HWE ($p < 0.05$) (Table 3.6) and for the controls three SNPs, *ABCB1* 2677 C/A (rs2032582), APOL1 1098 T/G (rs60910145) and *APOL1* 1072 A/G (rs73885319), deviated from HWE (Table 3.6). For the full cohort, the genotypes for four SNPs (rs2032582, rs2273697, rs60910145 and

rs73885319) deviated from HWE. Possible reasons for deviation from HWE include small sample size with non-random sampling, errors in genotyping, or extreme genetic drift. Caution was therefore applied to interpreting potentially significant associations observed for these SNPs.

Table 3.5. TaqMan® SNP genotyping assay robustness and validation

Gene	SNP	Ion Torrent sequencing	TaqMan® assay	Sanger sequencing	Outcome
ABCB1	rs2032582	Concordance	Separation of the genotypes	Validated	Successful
ABCC2	rs2273697	Complete concordance	Good separation of 3 genotypes	Validated	Successful
	rs17222723	Complete concordance	Good separation of 3 genotypes	Validated	Successful
	rs3740066	Discordance	Good separation of 3 genotypes	Validated	Successful
ABCC4	rs3742106	90% concordance	Good separation of 3 genotypes	Validated	Successful
	rs1751034	Complete concordance	Good separation of 3 genotypes	N/A	Successful
ABCC10	rs2125739	Complete concordance	Good separation of 3 genotypes	Validated	Successful
APOL1	rs73885319	N/A	Good separation of 3 genotypes	Validated	Successful
	rs60910145	N/A	Separation of 3 genotypes	Validated	Successful

a) Chi-square tests for genotype association

Genotype (DD vs Dd vs dd, where D is the minor allele (variant) and d is the major allele) association testing was employed to assess the variants and their potential association with AKI. The outcome of the association studies is shown in Table 3.6. Genotype association testing indicated that *ABCC2*, where the 1249 AA genotype (rs2273697) was marginally associated with the presence of AKI (p=0.08) and was non-significant after Bonferroni's correction for multiple testing. Other genotype association tests performed included investigating whether additive (dd ≥ Dd ≥ DD), dominant (DD, Dd vs dd) or recessive (DD vs Dd, dd) models influenced association with TDF-induced AKI. Based on the additive and

recessive models, *ABCC2* 1249 AA was associated with this outcome (p=0.04 and p=0.03, respectively) but was non-significant after Bonferroni correction for multiple testing (p=0.22). On the basis of the above findings, this variant could be of interest in finding predictive markers for TDF-induced AKI amongst populations of Bantu-speaking origin. Independent *a priori* validation of this association in an adequately powered samples would however be required.

Table 3.6. Genotype distribution and statistics for association analysis of the variants assessed in this study in relation to acute kidney inury in the Bantu-speaking cohort

Gene	SNP	Genotype	Controls		Cases		Case: Control association ($\chi 2$ test, p-value)	Case: Control association ($\chi 2$ test with Bonferroni correction, p-value)
			no.	HWE (p value)	no.	HWE (p value)		
ABCB1	rs2032582	CC	71	<0.01	48	0.01	0.66	1.00
		CA	2		3			
		AA	1		1			
		Total	74		52			
ABCC10	rs2125739	TT	43	0.65	29	0.14	0.21	1.00
		TC	33		23			
		CC	8		1			
		Total	84		53			
ABCC2	rs2273697	GG	62	0.44	34	0.01	0.08	0.73
		GA	16		12			
		AA	2		6			
		Total	80		52			
	rs17222723	TT	51	0.60	27	0.36	0.23	1.00
		TA	24		19			
		AA	4		6			
		Total	79		52			
	rs3740066	CC	62	0.62	45	0.18	0.30	1.00
		CT	18		6			
		TT	2		1			
		Total	82		52			
ABCC4	rs1751034	TT	36	0.91	24	0.78	0.96	1.00
		TC	35		22			
		CC	8		6			
		Total	79		52			
	rs3742106	AA	34	0.11	17	0.86	0.42	1.00
		AC	31		26			
		CC	15		9			
		Total	80		52			
APOL1	rs73885319	AA	70	<0.01	38	0.10	0.28	1.00
		AG	10		11			
		GG	3		3			
		Total	83		52			
	rs60910145	TT	71	0.05	43	0.04	0.28	1.00
		TG	10		7			
		GG	2		2			
		Total	83		52			

The control group comprised HIV infected patients receiving tenofovir and presenting with normal tubular function); and the case group are HIV infected patients with AKI. Genotype frequency differences between the cases and controls were evaluated using a Chi-squared ($\chi 2$) testing at 5% level of significance ($p \leq 0.05$). p values for the corrected $\chi 2$ (Bonferroni correction for multiple testing) are also included.

b) Chi-square tests for allele association

Allele (D vs d) association testing was employed to assess the variants and their potential association with AKI. Table 3.7 summarises the allele frequencies (minor allele frequencies (MAF)) and p values of associations for each variant screened. Allele association testing indicated that of the SNPs tested, only one had a suggestive association (*ABCC2* (rs2273697) (p=0.024)) with AKI following TDF treatment, where the 1249A allele was associated with the increased risk of TDF associated AKI (MAF= 0.23 vs 0.13). The association was not significant after Bonferroni correction (p=0.23). The differences between the allele frequencies of *ABCC2* 3563A allele (rs17222723) was marginally significant (p=0.08) but, was non-significant after Bonferroni's correction (p=0.69). These findings suggest the importance of *ABCC2* and could be of interest in finding predictive markers for TDF-induced AKI amongst populations of Bantu-speaking origin.

Table 3.7. Allele distribution and statistics for association analysis of the variants assessed in this study in relation to acute kidney injury in the Bantu-speaking cohort

Gene	SNP	Minor allele	Minor allele frequency		Case-control association ($\chi 2$ test, p-value)	Case-control association ($\chi 2$ with Bonferroni correction, p-value)
			Controls	Cases		
ABCB1	**rs2032582**	A	0.03	0.05	0.38	1
ABCC10	**rs2125739**	C	0.29	0.24	0.31	1
ABCC2	**rs2273697**	A	0.13	0.23	0.02	0.22
	rs17222723	A	0.20	0.30	0.08	0.69
	rs3740066	T	0.13	0.08	0.15	1
ABCC4	**rs1751034**	C	0.32	0.33	0.94	1
	rs3742106	C	0.38	0.42	0.50	1
APOL1	**rs73885319**	G	0.08	0.11	0.55	1
	rs60910145	G	0.10	0.16	0.10	0.91

The control group comprised HIV infected patients receiving tenofovir and presenting with normal tubular function); and the case group are HIV infected patients with AKI. Allele frequency differences between the cases and controls were evaluated using a Chi-squared ($\chi 2$) testing at 5% level of significance ($p \leq 0.05$). p values for the corrected $\chi 2$ (Bonferroni correction for multiple testing) are also included.

c) Chi-square tests for haplotype association

In addition to determining whether individual SNPs were associated with TDF-induced AKI, haplotype combinations were investigated using haplotype association tests. Haplotypes were only constructed for SNPs that occur in the same gene and are therefore closely linked. Association tests were performed in SVS using individual haplotypes inferred by the haplotype blocks defined by markers that lie in the same gene. The haplotype frequency estimates of each haplotype were calculated using the Expectation-Maximization algorithm (Gabriel *et al.* 2002).

Three *ABCC2* haplotypes (constructed based on positions 1249, 3563 and 3972) had suggestive associations (Table 3.8). The frequency of the *ABCC2* GTT haplotype was higher in the controls as it was found in 13% of the controls and 5% of the cases (p=0.02). However, the association was non-significant after correcting for multiple testing (p=0.15). The frequencies of two additional *ABCC2* haplotypes were significantly different between the cases and the controls. *ABCC2* haplotype AAC was found in 2% of the controls and in 8% of the cases (p=0.02, p-value with Bonferroni correction =0.13). These results propose that this haplotype could potentially be associated with susceptibility towards the development of renal failure post TDF administration. *ABCC2* haplotype AAT was found in 3% of the cases and not in any of the controls (p=0.03, p-value with Bonferroni correction =0.19). This analysis suggests that this haplotype could be a potential genetic biomarker as none of the cases had the genotype. These results are only suggestive and depict the possible significance of *ABCC2* variants. *ABCC2* variant rs2273697 deviated from HWE and therefore analysis must be replicated in a larger cohort. No significant associations were detected with *ABCB1*, *ABCC4* and *APOL1*. The haplotypes and their frequencies are summarised in Table 3.8 .

Table 3.8. Haplotype frequencies and associations with TDF induced acute kidney injury in a Bantu-speaking cohort

Markers Used	Haplotype	EM Freq	EM Controls Freq	EM Cases Freq	Case: control association ($\chi 2$ test, p-value)	Case: control association ($\chi 2$ with Bonferroni correction, p-value)
***ABCC2* haplotypes**						
rs2273697, rs17222723, rs3740066	GTC	0.55	0.56	0.53	0.66	1.00
	GAC	0.19	0.18	0.19	0.87	1.00
	ATC	0.11	0.11	0.12	0.67	1.00
	GTT	0.10	0.13	0.05	0.02	0.15
	AAC	0.04	0.02	0.08	0.02	0.13
	AAT	0.01	0.00	0.03	0.03	0.19
***ABCC4* haplotypes**						
rs3742106, rs1751034	AT	0.43	0.45	0.40	0.45	1.00
	CT	0.25	0.23	0.27	0.43	1.00
	AC	0.17	0.17	0.18	0.84	1.00
	CC	0.15	0.16	0.15	0.9	1.00
***APOL1* haplotypes**						
rs73885319, rs60910145	AT	0.87	0.90	0.84	0.11	0.34
	GG	0.09	0.07	0.11	0.36	1.00
	GT	0.04	0.02	0.06	0.16	0.49

The control group comprised HIV infected patients receiving tenofovir and presenting with normal tubular function; and the case group are HIV infected patients receiving tenofovir and with AKI. Expectation-Maximization (EM) algorithm was used to compute each haplotype frequency via SNP Variation Suite (SVS 8). Haplotypes frequency differences between the cases and controls were evaluated using a Chi-squared ($\chi 2$) testing at 5% level of significance ($p \leq 0.05$). p values for the corrected $\chi 2$ (Bonferroni correction for multiple testing) are also included.

3.3: DISCUSSION

3.3.1 Robustness of genotyping assays

The TaqMan assays for the nine prioritised variants were initially evaluated and compared to the genotypic calls from Ion Torrent sequencing and then to Sanger sequencing. The analysis showed complete concordance between the three methods for eight of the nine SNPs illustrating the robustness of the assaays. Several samples from the TDF cohort were used to validate the genotypes using Sanger sequencing (Appendix K-L). Allelic discrimination plots indicated that the genotypic calls for the assays were well distinguished (Appendix I). A few of the 'undetermined' genotypic calls from TDF cohort screening were confirmed with Sanger sequencing (Appendix M). Four SNPs deviated from HWE possibly due small sample size with non-random sampling, errors in genotyping, or extreme genetic drift. Non-random

sampling results in a biased sample population (or cohort) and selection criteria for a specific group of people (i.e HIV positive, adults of Bantu-ancestry) could introduce bias. It is possible that bias was introduced when an attempt was made to match the cohort for sex and age. It is unlikely that there was an error in genotyping as there was confirmed validation. Genetic drift is the random fluctuation in allele frequency due to chance events and could lead to short term disequilibrium. Genetic drift can make a big difference in small populations as alleles can become more frequent by chance and in extreme cases, fixed in a population. Genetic drift which reduces the genetic diversity of a population, can have important consequences when the minor allele frequency is relatively low. *ABCB1* 2677 C/A (rs2032582) was relatively rare (MAF=0.04) in our study cohort and a larger population is required to detect the effect of this variant. As such, any associations detected with those variants that deviated from HWE will need to be evaluated in a larger cohort.

3.3.2 TDF cohort descriptive statistics

Based on the variants that were prioritised following sequencing analysis, case-control association analysis was performed to evaluate their potential influence on TDF treatment outcome in relation to AKI. The TDF cohort that was genotyped for this purpose consisted of 137 patients of Bantu-ancestry who were HIV positive and were undergoing TDF combination therapy at the time of sample collection. The patients were on a fixed-dose combination regimen of TDF, emtricitabine (FTC) and EFV. EFV and FTC induced adverse drug reactions include rashes and hepatotoxicity. Other clinical indicators associated with adverse drug reactions to other ARVs were recorded and evaluated, but for the purpose of this study, the focus was on patients presenting with renal failure. Initially patients on TDF who presented with renal failure (cases) at the participating hospital were recruited onto the study. Simultaneously, otherwise healthy patients undergoing TDF combination therapy were recruited (from the hospital and surrounding clinics) and were allocated to the control group. The cases and controls were not matched and therefore it was necessary to assess the potential influence of confounders. Prior to performing case-control association analysis, descriptive statistics of the study cohort were generated. Sex was not statistically different between cases and controls ($p=0.060$); however, the mean ages of the cases and controls was statistically different ($p=0.006$). This analysis indicated that age is a potential confounder and was corrected for in the association tests.

3.3.3 Chi-squared association testing

a) Genotype and allele association testing

With to the pharmacogenetic of TDF induced AKI, *ABCC2* variants are the most studied variants; they have been studied in Caucasian and Asian populations (Izzedine *et al.* 2006, Rodríguez-Nóvoa *et al.* 2009, Nishijima *et al.* 2012, Giacomet *et al.* 2013, Manosuthi *et al.* 2014, Rungtivasuwan *et al.* 2015, Likanonsakul *et al.* 2016). In this current study *ABCC2* SNPs were evaluated at a genotype and allele frequency level using chi-squared tests for potential association with TDF-induced AKI. At a cut-off of $p \leq 5$ %, the *ABCC2* 1249 A allele (rs2273697) was observed to be associated with TDF induced AKI as it was the only variant that was statistically significantly higher (allele frequency) in the cases (p=0.02) compared to the controls. The association was not significant after multiple testing corrections. This variant has been previously shown to be statistically significantly higher amongst cases (presented with kidney tubular dysfunction) (p<0.02, OR= 6.11[95% CI= 1.19-31.15]) (Izzedine *et al.* 2006) and is associated with aminoaciduria (Rodríguez-Nóvoa *et al.* 2009) in patients of Caucasian ancestry and in patients of Japanese ancestry (p=0.12, OR=16.1 [95% CI= 1.63-161.1) (Nishijima *et al.* 2012). Since the genotype distribution for this SNP deviated from HWE in our population, a larger cohort should be tested to determine whether the data does support the previous findings or not. Similar results were obtained when evaluating the SNP association using additive ($dd \geq Dd \geq DD$) and recessive (DD vs Dd, dd) models.

ABCC4 variants are significantly studied in pharmacogenomics investigations (Izzedine *et al.* 2006, Imaoka *et al.* 2007, Kiser *et al.* 2008a, Rodríguez-Nóvoa *et al.* 2009, Pushpakom *et al.* 2011, Nishijima *et al.* 2012, Giacomet *et al.* 2013, Rungtivasuwan *et al.* 2015, Likanonsakul *et al.* 2016) and in the current study two variants were evaluated in relation to TDF induced AKI. *ABCC4* 3463 C/C (rs1751034) was previously shown to be associated with increased intracellular TDF concentration (35%) (p=0.04) and lower clearance of TDF as compared to TT carriers, who had lower TDF concentration and higher TDF clearance (Kiser *et al.* 2008a). However, in our study cohort the differences in genotype and allele frequencies of the cases and controls were non-significant. Previous studies indicated that *ABCC4* 4131 A/C and C/C were associated with altered uric acid excretion (Rodríguez-Nóvoa *et al.* 2009) and

carriers of the C/C genotype had increased mean plasma concentration of TDF (Rungtivasuwan *et al.* 2015). However, no associations between this variant and TDF induced AKI were observed in our study cohort.

ABCC10 is another gene that was evaluated in the current study. Pushpamkom and colleagues (Pushpakom *et al.* 2011) have previously reported that the allele frequency of *ABCC10* 2759 C/C genotype (rs2125739) was marginally associated with patients with kidney tubular dysfunction, KTD (p=0.05, OR=2.0). Their haplotype analysis also indicated that the *ABCC10* haplotype, GGC (rs9349256G, rs2487663G and rs2125739C) was marginally significantly higher in patients with TDF induced kidney tubular dysfunction (KTD) (p=0.05, OR=2.1). The potential influence of rs2125739 was evaluated in our Bantu-speaking cohort. There was no significant difference between the 2759 C allele and CC genotype frequencies of the cases and the controls. *ABCC10* genotype 2759 C/C is also associated with nevirapine (NVP) plasma concentrations associated with NVP induced hepatotoxicity (Liptrott *et al.* 2012); however none of the patients in this study were receiving NVP at the time of sample collection. *ABCC10* 1791+526 A/A (rs9349256) was previously shown to be associated with KTD, phosphorous wasting and β_2 microglobulin (p=0.02 [OR=2.3], 0.02 and 0.04) (Pushpakom *et al.* 2011); however this variant was not prioritised for investigation as it was not detected during sequencing of our representative population due to its intronic location.

ABCB1 variants have been previously evaluated in relation to TDF induced AKI and Nishijima and colleagues initially identified the *ABCB1* T2677A polymorphism in a Japanese population (Nishijima *et al.* 2012); however the same ethnic group did not display any association between variant genotype status at *ABCB1* 2677 C/A and TDF induced AKI (Nishijima et al 2015). *ABCB1* 2677 C/A was also evaluated in our study cohort; however no significant association was observed. Interestingly, *ABCB1* T2677A has been reported to be associated with increased survival and response in patients with multiple myeloma, and has been proposed for use as a prognostic biomarker (Maggini *et al.* 2008, Buda *et al.* 2009, Jamroziak *et al.* 2009, Penna 2011). *ABCB1* 2677 C/A is also in strong LD with *ABCB1* 3435 C/T (rs1045642), associated with DILI (rifampicin-EFV) in TB-HIV patients of African ancestry (Haas *et al.* 2005, Yimer *et al.* 2011).

APOL1 variants were evaluated in the TDF cohort because of their association with structural kidney dysfunction (Genovese *et al.* 2010, Tzur *et al.* 2010) and strong association (p=< 0.001 and OR=89) with HIV associated nephropathy in South African patients (Kasembeli *et al.* 2015a). Kasembeli and collegues evaluated the influence of G1 allele (*APOL1* 1072 G and 1098 G) and G2 allele (c.1212_1217delTTATAA (rs71785313)) and carriers of two risk alleles were estimated to exhibit 89-fold higher odds of developing HIVAN. G1 allele variants were evaluated in our study cohort and, to my knowledge, this is the first study to assess the potential influence of theses variants on TDF induced renal failure. No significant associations were detected with *APOL1* alleles, genotypes or its haplotypes.

The frequencies of the allele and genotypes for the other four genes (*ABCB1*, *ABCC4*, *ABCC10* and *APOL1*) between the cases and controls were statistically similar. The *ABCC2* 1249A allele was the only allele that was found to be potentially important due to the detection of a suggestive association. Further work may therefore be warranted to investigate this variant in larger cohorts for its potential contribution to AKI during TDF use, and furthermore its ability to contribute to predicting those most likely to have this outcome in this population.

b) Haplotype association testing

At some loci, haplotypes have better discriminatory power to identify associations with a gene locus. This could be because the causal variant has not been identified and a specific haplotype, rather than a single SNP, is co-inherited. In order to analyse multiple variants in a single gene for association with TDF induced AKI, haplotype analysis was performed with a cut-off of ≤ 5%. The Expectation-Maximization (EM) algorithm (implemented in SVS) was used to calculate the haplotype frequency estimations. Suggestive associations were detected between TDF induced AKI and *ABCC2* haplotypes. The haplotypes were constructed on the basis of three SNPs at positions at 1249, 3463 and 3972. The AAC and AAT haplotypes were associated with TDF induced AKI (p=0.02 and 0.03) whilst haplotype GTT was associated with controls status, i.e. protective (p=0.02). These associations were however non-significant after correcting for multiple testing. AAT haplotype may result in susceptibility towards the development of TDF induced AKI as it was only detected in the cases, however due to the size of the cohort that was tested a likely possibility is that the cohort was insufficiently powered to detect the presence of this haplotype amongst the control group. A

previous study detected significant associations with *ABCC2* haplotypes (at positions -24, 1249, 3563, and 3972), whereby CATC increased risk of TDF induced AKI ($p<0.01$) whereas CGAC has protective effects against AKI ($p<0.01$) (Izzedine *et al.* 2006). The same study also reported that *ABCC2* variants 3563 A (rs17222723) and 4544 A alleles have protective effects (decreased susceptibility to TDF-KTD) as the alleles were only detected in the control groups (Izzedine *et al.* 2006). In contrast, a previous study reported an association between variants at *ABCC2* 3563 A/A and cardiotoxicity in non-Hodgkin lymphoma patients treated with doxorubicin (Wojnowski *et al.* 2005). An *ABCC2* 3972 T/C (rs3740066) variant has previously been shown to be statistically associated with β2- microglobulinuria (Rodríguez-Nóvoa *et al.* 2009) however; the C/C genotype (wild type) was associated with AKI indicators compared to the influence of the T/T genotype that was evaluated in our cohort. No significant association was detected with this variant in our cohort ($p= 0.3$, corrected $p=1$). Studies, including ours, have suggested and demonstrated that variants in this gene are associated with increased risk of TDF induced AKI and some haplotypes have protective effects. These studies demonstrate the necessity of evaluating these variants in bigger and suitably powered cohorts to elucidate the exact effect on TDF outcomes and TDF induced nephropathy.

c) *ABCC2* and TDF associated nephrotoxicity

The mechanism through which TDF causes nephrotoxicity is not well understood. The uptake of TDF into the renal tubular cells is well characterised (Cihlar *et al.* 2001, Motohashi *et al.* 2002, Ray *et al.* 2006) (Appendix N) however how TDF is eliminated during active tubular secretion is less known. P glycoprotein 1 (P-gp 1), multidrug resistance protein (MPR) type 2 (MRP2) and type 4 (MPR4) proteins are possible transporters of TDF as they are expressed in the apical side of the tubular cells (Miller 2001). MRP2 (*ABCC2)* protein is implicated in the secretion of TDF (Izzedine *et al.* 2005b, Zimmermann *et al.* 2006). This is supported by the fact that MRP2 is located in the apical side of proximal-tubule cells and an *in-vitro* study demonstrated that MRP2 eliminates cidofovir and adefovir (acyclic analogues which are structurally related to TDF) (Miller 2001). TDF and other ayclic nucleotide phosphonates (such as adefovir) are eliminated by MRP4 (Ray *et al.* 2006, Imaoka *et al.* 2007). The elimination of TDF by MRP2 has not been experimentally established and Ray and colleagues demonstrated that TDF is not a substrate of P-gp1 or MPR2 (Ray *et al.* 2006).

Even though TDF is not a substrate for human MRP2 *in vitro* however MPR2 appears to affect hepatobiliary elimination in rats (Mallants *et al.* 2005). TDF usage is associated with tubular dysfunction and tubular damage is multifactorial. Age and body weight have been shown to be confounders for TDF induced AKI (Jullien *et al.* 2005, Rodríguez-Nóvoa *et al.* 2009, Calcagno *et al.* 2013). Studies have suggested a genetic predisposition to TDF induced AKI. Their findings, and ours, indicate the importance of the transporters localised in kidneys that transport TDF (Appendix N), especially *ABCC2*. This gene contains two of the most commonly implicated variants (-24C and 1249A) (Izzedine *et al.* 2006, Rodríguez-Nóvoa *et al.* 2009, Nishijima *et al.* 2012).

The mechanism through which MPR2 mediates TDF induced AKI is not established and several hypotheses have been suggested. Active tubular secretion could be less efficient in tubular cells of carriers of genotypes associated with the adverse outcome. This leads to increased intracellular concentration of TDF and subsequent nephrotoxicity due damaged tubular cells. This mechanism was reported to be associated with the MPR4 transporter where individuals with the *ABCC4* 3463 G/G genotype had 35% higher concentration of TDF than carriers of the wild type genotype (Kiser *et al.* 2008a). Another mechanism is that variants associated with TDF induced AKI are in LD with variants that influence TDF induced AKI. It is also plausible is that MPR2 may transport other xenobiotics or chemicals that influence TDF induced nephrotoxicity. Rodríguez-Nóvoa and collgues motivated for this hypothesis as TDF is not a MPR2 substrate but appears to affect hepatobilay elimination (Rodríguez-Nóvoa *et al.* 2009). TDF associated tubular dysfunction is multifactorial as TDF is an ATP analogue that inhibits transcription and results in mitochondrial toxicity. Since TDF requires active transport concentration dependent toxicity is postulated to be the main cause of renal dysfunction (Hall. *et al.* 2011). Therefore, the build-up of TDF in the renal cells is postulated to be due to structural changes in the proteins that mediate the influx and efflux of TDF. TDF results in the reduced activity of energy dependent transporters. The extent to which variants investigated in this study may influence the efficacy of Tenofovir alafenamide (TAF), a prodrug of tenofovir (Gibson *et al.* 2016), is unclear. This drug was approved by the FDA in 2015 and clinical studies have shown that the drug is safer and has higher rates of virological suppression (Sax *et al.* 2015, Gibson *et al.* 2016, Greig and Deeks 2016, Mills *et al.* 2016, Pozniak *et al.* 2016, Ray *et al.* 2016). This drug (TAF) could potentially replace TDF in the future.

d) Study limitations

Overall, no significant associations between genetic variant status and with TDF induced AKI were detected and this may be a reflection of the limitations of this sub study. A possible reason for the lack of association could be due the study's relatively low sample power. This study only found suggestive associations between *ABCC2* variants and TDF-induced AKI; at the same time four of the nine SNPs deviated from HWE. The sample size has a significant impact on the outcome of the results and increasing the sample size can improve the study's power to detect important associations. Another limitation could potentially be study design, only patients that presented with acute renal failure were recruited while for several studies the cases were exposed to TDF for a relatively longer period (Zimmermann *et al.* 2006, Rodríguez-Nóvoa *et al.* 2009, Rungtivasuwan *et al.* 2015, Likanonsakul *et al.* 2016). Several of the studies were matched for age, sex and race (based on the descriptive statistics of some) and the differences between the cases and controls were not significant (Kiser *et al.* 2008a, Rodríguez-Nóvoa *et al.* 2009, Pushpakom *et al.* 2011, Nishijima *et al.* 2012, Likanonsakul *et al.* 2016), except for one (Nishijima *et al.* 2012). These studies were adjusted for age and sex. Descriptive statistical analysis indicating that age was a potential confounder in the current study ($p=0.006$) and may have influenced outcome. Despite these limitations, evidence was observed that suggests a potential influence of *ABCC2* variants on TDF induced AKI in HIV-positive patients of Bantu ancestry. Therefore, findings from this work will need to be replicated in more suitably powered cohorts.

CHAPTER 4: FUTURE PERSPECTIVES

This study can be used as the basis for potential future research. Such research includes repeating the study in a bigger cohort to potentially obtain a more normal distribution of genotypes of all variants evaluated. This bigger study could also potentially result in a suitably powered cohort, through which statistically significant association can be detected and validated. At the same time meta-analysis would enable data from these multiple studies to be analysed simultaneously, in order to confidently identify important associations. Data generated from such studies would provide more knowledge towards the development of pharmacogenetic interventions designed to guide optimal, more precise, ARV therapy.

This study focused on coding regions of genes implicated in TDF induced AKI, primarily due to the nature of the sequencing technology that was employed. Variants in intronic, upstream, downstream, untranslated and regulatory regions could therefore also be characterised and evaluated in a suitably powered Bantu-speaking cohort. Novel and LoF variants detected using NGS could be validated using an alternative method and their influence on TDF induced AKI can be evaluated. The influence of variants associated with TDF induced AKI could be evaluated where a genetic model can be developed to determine the influence of these variants and to determine how well they can predict outcome. The accuracy of the model can be evaluated and validated in a Bantu-speaking cohort. Once the variants are validated they can be used to develop a genetic panel that can potentially be used in the future to predict AKI prior to TDF initiation. Initially such a test would serve predominantly patients of Bantu-ancestry who represent a large proportion of those receiving such medication, given the initial development work in this population. However similar studies on other population groups within the country could lead to the development of similar tests.

CHAPTER 5: CONCLUSION

Adverse drug events are a concern in South Africa. South African reports indicated that 8.33% of admissions and 2.9% of hospital deaths are attributed to adverse events associated with TDF, EFV, rifampicin and co-trimoxazole (an antibiotic). Approximately 3.5 million people are on ARVs and with more HIV positive people expected to be put on the TDF-based regimen, the number of patients affected by TDF induced AKI is expected to increase. Prior pharmacogenetic studies have noted significant differences across populations in variants associated with drug therapeutic outcome, including TDF induced AKI. In this study I aimed to comprehensively identify variants present in eight TDF implicated ADME genes and to select a subset of SNPs for a genetic association study for TDF-associated nephrotoxicity.

Targeted ultra-deep sequencing of 40 black South Africans confirmed the presence in this population of 275 variants in eight ADME genes implicated in TDF efficacy. These included three loss of function variants and 40 novel variants. The study also detected the presence of at least nine variants that have previously been shown in other populations to be associated with TDF-induced nephrotoxicity.

A panel of nine carefully chosen SNPs were evaluated in a case-control genetic association study to determine their association with TDF associated nephrotoxicity. Suggestive allelic associations with *ABCC2 (*rs2272369*)* were detected, such that the A allele segregated more often amongst cases (p=0.02). The *ABCC2* haplotypes (GTT and AAT) were associated with the cases (p=0.02 and 0.03, respectively), where ATT was over-represented. Suggestive associations were also detected between the *ABCC2* AAC haplotype and reduced susceptibility to TDF induced AKI in the controls (p=0.02).

A limitation of the study was its relatively low sample size and reduced power to detect small to moderate effects. Despite this, we observed suggestive evidence for a potential influence of a number of genetic variants on TDF induced AKI, as has been reported in current literature (Izzedine *et al.* 2006, Rodríguez-Nóvoa *et al.* 2009, Nishijima *et al.* 2012, Manosuthi *et al.* 2014). The findings from this study will need to be replicated in more suitably powered cohorts. In summary, this study provided more insight into the genetic variance of eight ADME genes in a Bantu-speaking cohort, and explored clinically

significant variants in two cohorts of Bantu ancestry. It is hoped that the knowledge generated from this study will add value towards the development of Bantu-specific pharmacogenetic-based interventions.

CHAPTER 6: REFERENCES

Abecasis, G. A., A . Brooks, LD . DePristo, MA . Durbin, RM. Handsaker, RE. Kang, HM. Marth, GT. McVean, GA. (1000 Genome Project Consortium) (2012) 'An integrated map of genetic variation from 1,092 human genomes.', *Nature*, 491(7422), 56-65.

Abecasis, G. A., D . Auton, A . Brooks, LD .Durbin, RM . Gibbs, RA . Hurles, ME . McVean, GA. (1000 Genome Project Consortium) (2010) 'A map of human genome variation from population-scale sequencing.', *Nature*, 467(7319), 1061-1073.

Abel, S., Russell, D., Taylor-Worth, R. J., Ridgway, C. E. and Muirhead, G. J. (2008) 'Effects of CYP3A4 inhibitors on the pharmacokinetics of maraviroc in healthy volunteers', *Br J Clin Pharmacol*, 65 Suppl 1, 27-37.

AIDSinfo (2016) *HIV and Lipodystrophy Understanding HIV/AIDS* [online], available: https://www.ncbi.nlm.nih.gov/pubmed/ [accessed 26.09.17].

Altshuler, D. M., Gibbs, R. A., Peltonen, L., Dermitzakis, E., Schaffner, S. F., Yu, F., Bonnen, P. E., de Bakker, P. I., Deloukas, P., Gabriel, S. B., Gwilliam, R., Hunt, S., Inouye, M., Jia, X., Palotie, A., Parkin, M., Whittaker, P., Chang, K., Hawes, A., Lewis, L. R., Ren, Y., Wheeler, D., Muzny, D. M., Barnes, C., Darvishi, K., Hurles, M., Korn, J. M., Kristiansson, K., Lee, C., McCarrol, S. A., Nemesh, J., Keinan, A., Montgomery, S. B., Pollack, S., Price, A. L., Soranzo, N., Gonzaga-Jauregui, C., Anttila, V., Brodeur, W., Daly, M. J., Leslie, S., McVean, G., Moutsianas, L., Nguyen, H., Zhang, Q., Ghori, M. J., McGinnis, R., McLaren, W., Takeuchi, F., Grossman, S. R., Shlyakhter, I., Hostetter, E. B., Sabeti, P. C., Adebamowo, C. A., Foster, M. W., Gordon, D. R., Licinio, J., Manca, M. C., Marshall, P. A., Matsuda, I., Ngare, D., Wang, V. O., Reddy, D., Rotimi, C. N., Royal, C. D., Sharp, R. R., Zeng, C., Brooks, L. D. and McEwen, J. E. (2010) 'Integrating common and rare genetic variation in diverse human populations', *Nature*, 467(7311), 52-8.

Aminkeng, F., Ross, C. J. D., Rassekh, S. R., Brunham, L. R., Sistonen, J., Dube, M.-P., Ibrahim, M., Nyambo, T. B., Omar, S. A., Froment, A., Bodo, J.-M., Tishkoff, S., Carleton, B. C. and Hayden, M. R. (2013) 'Higher frequency of genetic variants conferring increased risk for ADRs for commonly used drugs treating cancer, AIDS and tuberculosis in persons of African descent', *The Pharmacogenomics Journal*, 14(2), 160-170.

An, P., Bleiber, G., Duggal, P., Nelson, G., May, M., Mangeat, B., Alobwede, I., Trono, D., Vlahov, D., Donfield, S., Goedert, J. J., Phair, J., Buchbinder, S., O'Brien, S. J., Telenti, A. and Winkler, C. A. (2004) 'APOBEC3G Genetic Variants and Their Influence on the Progression to AIDS', *virology*, 78(20), 11070-11076.

Andiman, G., Andiman, W. A., Chernoff, M. C., Mitchell, C., Purswani, M., Oleske, J., Williams, P. L., Spiegel, H., Gona, P. and Seage, G. R. (2009) 'Incidence of persistent renal dysfunction in human immunodeficiency virus-infected children: associations with the use of antiretrovirals, and other nephrotoxic medications and risk factors.', *The Pediatric infectious disease journal*, 28(7), 619-625.

Andrew, C. (2003) 'HIV lipodystrophy: risk factors, pathogenesis, diagnosis and management', *AIDS*, 17.

Andrew, C., Katherine, S., Samantha, B., Matthew, L., Judith, F., J., C. D. and David, C. (2016) 'A syndrome of peripheral lipodystrophy, hyperlipidaemia and... : AIDS'.

Arribas, J. R., Pozniak, A. L., Gallant, J. E., Dejesus, E., Gazzard, B., Campo, R. E., Chen, S. S., McColl, D., Holmes, C. B., Enejosa, J., Toole, J. J. and Cheng, A. K. (2008) 'Tenofovir disoproxil fumarate, emtricitabine, and efavirenz compared with zidovudine/lamivudine and efavirenz in treatment-naive patients: 144-week analysis', *J Acquir Immune Defic Syndr*, 47(1), 74-8.

Aurpibul, L., Chotirosniramit, N., Sugandhavesa, P., Kosashunhanan, N., Thetket, S., Supindham, T., Piyamongkol, W. and Supparatpinyo, K. (2012) 'Correlation of CYP2B6-516G > T Polymorphism with Plasma Efavirenz Concentration and Depression in HIV-Infected Adults in Northern Thailand', *Curr HIV Res*, 10(8), 653-60.

Auton, A. B., LD . Durbin, RM . Garrison, EP. Kang, HM . Korbel, JO. Marchini, JL . McCarthy, S. McVean, GA. Abecasis, GR. (1000 Genomes Project Consortium) (2015) 'A global reference for human genetic variation.', *Nature*, 526(7571), 68-74.

Azuma, J., Ohno, M., Kubota, R., Yokota, S., Nagai, T., Tsuyuguchi, K., Okuda, Y., Takashima, T., Kamimura, S., Fujio, Y. and Kawase, I. (2013) 'NAT2 genotype guided regimen reduces isoniazid-induced liver injury and early treatment failure in the 6-month four-drug standard treatment of tuberculosis: a randomized controlled trial for pharmacogenetics-based therapy', *Eur J Clin Pharmacol*, 69(5), 1091-101.

Baeten, J. M., Donnell, D., Ndase, P., Mugo, N. R., Campbell, J. D., Wangisi, J., Tappero, J. W., Bukusi, E. A., Cohen, C. R., Katabira, E., Ronald, A., Tumwesigye, E., Were, E., Fife, K. H., Kiarie, J., Farquhar, C., John-Stewart, G., Kakia, A., Odoyo, J., Mucunguzi, A., Nakku-Joloba, E., Twesigye, R., Ngure, K., Apaka, C., Tamooh, H., Gabona, F., Mujugira, A., Panteleeff, D., Thomas, K. K., Kidoguchi, L., Krows, M., Revall, J., Morrison, S., Haugen, H., Emmanuel-Ogier, M., Ondrejcek, L., Coombs, R. W., Frenkel, L., Hendrix, C., Bumpus, N. N., Bangsberg, D., Haberer, J. E., Stevens, W. S., Lingappa, J. R. and Celum, C. (2012) 'Antiretroviral Prophylaxis for HIV Prevention in Heterosexual Men and Women', *The New England Journal od medicine*, 367(5), 399-410.

Bains, R. K. (2013) 'African variation at Cytochrome P450 genes'.

Barco, E. Á., and Nóvoa, S. R. (2013) 'The Pharmacogenetics of HIV Treatment: A Practical Clinical Approach.', *Journal of Pharmacogenomics & Pharmacoproteomics.*

Barco, E. Á. and Nóvoa, S. R. (2013) 'The Pharmacogenetics of HIV Treatment: A Practical Clinical Approach.The Pharmacogenetics of HIV Treatment: A Practical Clinical Approach.', *Journal of Pharmacogenomics & Pharmacoproteomics.Journal of Pharmacogenomics & Pharmacoproteomics.*

Baroncelli, S., Pirillo, M. F., Tamburrini, E., Guaraldi, G., Pinnetti, C., Degli Antoni, A., Galluzzo, C. M., Stentarelli, C., Amici, R. and Floridia, M. (2015) 'Full Viral Suppression, Low-Level Viremia, and Quantifiable Plasma HIV-RNA at the End of Pregnancy in HIV-Infected Women on Antiretroviral Treatment', *AIDS Res Hum Retroviruses*, 31(7), 673-8.

Barré-Sinoussi, F., Chermann, J. C., Rey, F., Nugeyre, M. T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Vézinet-Brun, F., Rouzioux, C., Rozenbaum, W. and Montagnier, L. (1983) 'Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS)', *Science*, 220(4599), 868-71.

Beach, J. W. (1998) 'Chemotherapeutic agents for human immunodeficiency virus infection: mechanism of action, pharmacokinetics, metabolism, and adverse reactions', *Clin Ther*, 20(1), 2-25; discussion l.

Behar, D. M., Kedem, E., Rosset, S., Haileselassie, Y., Tzur, S., Kra-Oz, Z., Wasser, W. G., Shenhar, Y., Shahar, E., Hassoun, G., Maor, C., Wolday, D., Pollack, S. and Skorecki, K. (2011) 'Absence of APOL1 risk variants protects against HIV-associated nephropathy in the Ethiopian population', *Am J Nephrol*, 34(5), 452-9.

Benjeddou, M. (2010) 'Solute carrier transporters: Pharmacogenomics research opportunities in Africa', *African Journal of Biotechnology*, 9(2).

Berniell-Lee, G., Calafell, F., Bosch, E., Heyer, E., Sica, L., Mouguiama-Daouda, P., Veen, L. v. d., Hombert, J.-M., Quintana-Murci, L. and Comas, D. (2009) 'Genetic and Demographic Implications of the Bantu Expansion: Insights from Human Paternal Lineages', *Molecular Biology and Evolution*.

Best, B. M., Burchett, S., Li, H., Stek, A., Hu, C., Wang, J., Hawkins, E., Byroads, M., Watts, D. H., Smith, E., Fletcher, C. V., Capparelli, E. V. and Mirochnick, M. (2015) 'Pharmacokinetics of tenofovir during pregnancy and postpartum', *HIV Med*, 16(8), 502-11.

Bickel, M., Khaykin, P., Stephan, C., Schmidt, K., Buettner, M., Amann, K., Lutz, T., Gute, P., Haberl, A., Geiger, H., Brodt, H. R. and Jung, O. (2013) 'Acute kidney injury caused by tenofovir disoproxil fumarate and diclofenac co-administration', *HIV Med*, 14(10), 633-8.

Bihl, G. (2003) 'HIV-related renal disease - A clinical and practical approach in the South African context', *South African Journal of Medicine*, 11–14.

Bolhaar, M. G. and Karstaedt, A. S. (2007) 'A high incidence of lactic acidosis and symptomatic hyperlactatemia in women receiving highly active antiretroviral therapy in Soweto, South Africa', *Clin Infect Dis*, 45(2), 254-60.

Brennan, A., Evans, D., Maskew, M., Naicker, S., Ive, P., Sanne, I., Maotoe, T. and Fox, M. (2011) 'Relationship between renal dysfunction, nephrotoxicity and death among HIV adults on tenofovir', *AIDS*, 25(13), 1603-9.

Buda, G., Martino, A., Maggini, V., Orciuolo, E., Galimberti, S., Gentile, M., Morabito, F., Antognoni, F., Brunori, M., Barale, R., Rossi, A. M. and Petrini, M. (2009) 'MDR1 C3435T Polymorphism Indicates a Different Outcome in Advanced Multiple Myeloma', *Acta Haematologica*, 122(1), 42-45.

Busby, G. B., Band, G., Si Le, Q., Jallow, M., Bougama, E., Mangano, V. D., Amenga-Etego, L. N., Enimil, A., Apinjoh, T., Ndila, C. M., Manjurano, A., Nyirongo, V., Doumba, O., Rockett, K. A., Kwiatkowski, D. P. and Spencer, C. C. (2016) 'Admixture into and within sub-Saharan Africa', *Elife*, 5.

Calcagno, A., Gonzalez de Requena, D., Simiele, M., D'Avolio, A., Tettoni, M. C., Salassa, B., Orofino, G., Bramato, C., Libanore, V., Motta, I., Bigliano, P., Orsucci, E., Di Perri, G. and Bonora, S. (2013) 'Tenofovir plasma concentrations according to companion drugs: a cross-sectional study of HIV-positive patients with normal renal function', *Antimicrob Agents Chemother*, 57(4), 1840-3.

Caldwell, J., Gardner, I. and Swales, N. (1995) 'An introduction to drug disposition: the basic principles of absorption, distribution, metabolism, and excretion', *Toxicol Pathol*, 23(2), 102-14.

Carr, A. (2003) 'HIV lipodystrophy: risk factors, pathogenesis, diagnosis and management', *AIDS*, 17.

Carr, A., Samaras, K., Burton, S., Law, M., Freund, J., Chisholm, D. J. and Cooper, D. A. (1998) 'A syndrome of peripheral lipodystrophy, hyperlipidaemia and insulin resistance in patients receiving HIV protease inhibitors', *AIDS*, 12(7).

Cassetti, I., Madruga, J. V. R., Suleiman, J. M. A. H., Etzel, A., Zhong, L., Cheng, A. K., Enejosa, J. and Team, f. t. S. E. (2007) 'The Safety and Efficacy of Tenofovir DF in Combination with Lamivudine and Efavirenz Through 6 Years in Antiretroviral-Naïve HIV-1—Infected Patients', *http://dx.doi.org/10.1310/hct0803-164*.

Chigutsa, E., Visser, M. E., Swart, E. C., Denti, P., Pushpakom, S., Egan, D., Holford, N. H. G., Smith, P. J., Maartens, G., Owen, A. and McIlleron, H. (2011) 'The SLCO1B1 rs4149032 Polymorphism Is Highly Prevalent in South Africans and Is Associated with Reduced Rifampin Concentrations: Dosing Implications▽' in *Antimicrob Agents Chemother*4122-7.

Chirac, P. T., Els. (2006) 'Global framework on essential health R&D', *The Lancet*, 367(9522), 1560-1.

Christine Njuguna, A. S., Johannes P. Mouton, Marc Blockman, Gary Maartens, Annoesjka Swart, Briony Chisholm, Jackie Jones, Mukesh Dheda, Ehimario U. Igumbor, Karen Cohen (2016) 'Adverse Drug Reactions Reported to a National HIV & Tuberculosis Healt', *Drug Safety*, 39(2), 159–169.

Cihlar, T., Ho, E. S., Lin, D. C. and Mulato, A. S. (2001) 'Human renal organic anion transporter 1 (hOAT1) and its role in the nephrotoxicity of antiviral nucleotide analogs', *Nucleosides Nucleotides Nucleic Acids*, 20(4-7), 641-8.

Cihlar, T. and Ray, A. S. (2010) 'Nucleoside and nucleotide HIV reverse transcriptase inhibitors: 25 years after zidovudine', *Antiviral Res*, 85(1), 39-58.

Coghlan, M. E., Sommadossi, J. P., Jhala, N. C., Many, W. J., Saag, M. S. and Johnson, V. A. (2001) 'Symptomatic lactic acidosis in hospitalized antiretroviral-treated patients with human immunodeficiency virus infection: a report of 12 cases', *Clin Infect Dis*, 33(11), 1914-21.

Conradie, F., D Wilson, and On behalf of the Southern African HIV Clinicians Society (2012) 'The 2012 southern African ARV drug resistance testing guidelines', *Southern African Journal of HIV Medicine[Online]Southern African Journal of HIV Medicine[Online]*, 13.4 (2012): 162. Web. 18 Feb. 2014(4), 162-164.

Créput, C., Gonzalez-Canali, G., Hill, G., Piketty, C., Kazatchkine, M. and Nochy, D. (2003) 'Renal lesions in HIV-1-positive patient treated with tenofovir', *AIDS*, 17(6), 935-937.

Currier, J. S. (2007) 'Sex differences in antiretroviral therapy toxicity: lactic acidosis, stavudine, and women', *Clin Infect Dis*, 45(2), 261-2.

De Clercq, E. (1998) 'The role of non-nucleoside reverse transcriptase inhibitors (NNRTIs) in the therapy of HIV-1 infection', *Antiviral Res*, 38(3), 153-79.

Dickinson, L., Khoo, S. and Back, D. (2010) 'Pharmacokinetics and drug-drug interactions of antiretrovirals: an update', *Antiviral Res*, 85(1), 176-89.

DoHSA, T. d. o. h. S. A. (2013) *The South African antiretroviral treatment guidelines:2013*http://www.kznhealth.gov.za/medicine/2013_art_guidelines.pdf [accessed 19/12/16].

Eilers, M., Roy, U. and Mondal, D. (2008) 'MRP (ABCC) transporters-mediated efflux of anti-HIV drugs, saquinavir and zidovudine, from human endothelial cells', *Exp Biol Med (Maywood)*, 233(9), 1149-60.

Erickson, D. A., Mather, G., Trager, W. F., Levy, R. H. and Keirns, J. J. (1999) 'Characterization of the in vitro biotransformation of the HIV-1 reverse transcriptase inhibitor nevirapine by human hepatic cytochromes P-450', *Drug Metab Dispos*, 27(12), 1488-95.

Eugen-Olsen, J., Iversen, A. K., Garred, P., Koppelhus, U., Pedersen, C., Benfield, T. L., Sorensen, A. M., Katzenstein, T., Dickmeiss, E., Gerstoft, J., Skinhoj, P., Svejgaard, A., Nielsen, J. O. and Hofmann, B. (1997) 'Heterozygosity for a deletion in the CKR-5 gene leads to prolonged AIDS-free survival and slower CD4 T-cell decline in a cohort of HIV-seropositive individuals', *Aids*, 11(3), 305-10.

Fabian, J. and Naicker, S. (2009) 'HIV and kidney disease in sub-Saharan Africa', *Nat Rev Nephrol*, 5(10), 591-8.

Fabian, J., Naicker, S., Goetsch, S. and Venter, W. D. (2013) 'The clinical and histological response of HIV-associated kidney disease to antiretroviral therapy in South Africans', *Nephrol Dial Transplant*, 28(6), 1543-54.

Falcó, V., Crespo, M. and Ribera, E. (2003) 'Lactic acidosis related to nucleoside therapy in HIV-infected patients', *Expert Opin Pharmacother*, 4(8), 1321-9.

Falcó, V., Rodríguez, D., Ribera, E., Martínez, E., Miró, J. M., Domingo, P., Diazaraque, R., Arribas, J. R., González-García, J. J., Montero, F., Sánchez, L. and Pahissa, A. (2002) 'Severe nucleoside-associated lactic acidosis in human immunodeficiency virus-infected patients: report of 12 cases and review of the literature', *Clin Infect Dis*, 34(6), 838-46.

Fatti, G., Bock, P., Grimwood, A. and Eley, B. (2010) 'Increased vulnerability of rural children on antiretroviral therapy attending public health facilities in South Africa: a retrospective cohort study', *J Int AIDS Soc*, 13, 46.

Fellay, J., Marzolini, C., Meaden, E. R., Back, D. J., Buclin, T., Chave, J. P., Decosterd, L. A., Furrer, H., Opravil, M., Pantaleo, G., Retelska, D., Ruiz, L., Schinkel, A. H., Vernazza, P., Eap, C. B., Telenti, A. and Study, S. H. C. (2002) 'Response to antiretroviral treatment in HIV-1-infected individuals with allelic variants of the multidrug resistance transporter 1: a pharmacogenetics study', *Lancet*, 359(9300), 30-6.

Franceschini, N., Napravnik, S., Eron, J. J. J., Szczech, L. A. and Finn, W. F. (2005) 'Incidence and etiology of acute renal failure among ambulatory HIV-infected patients', *Kidney International*, 67, 1526–1531.

Frasco, M. A., Mack, W. J., Van Den Berg, D., Aouizerat, B. E., Anastos, K., Cohen, M., De Hovitz, J., Golub, E. T., Greenblatt, R. M., Liu, C., Conti, D. V. and Pearce, C. L. (2012) 'Underlying genetic structure impacts the association between CYP2B6 polymorphisms and response to efavirenz and nevirapine', *AIDS*, 26(16), 2097-106.

Freedman, B. I., Kopp, J. B., Langefeld, C. D., Genovese, G., Friedman, D. J., Nelson, G. W., Winkler, C. A., Bowden, D. W. and Pollak, M. R. (2010) 'The apolipoprotein L1 (APOL1) gene and nondiabetic nephropathy in African Americans', *J Am Soc Nephrol*, 21(9), 1422-6.

Fux, C. A., Rauch, A., Simcock, M., Bucher, H. C., Hirschel, B., Opravil, M., Vernazza, P., Cavassini, M., Bernasconi, E., Elzi, L., Furrer, H. and Study, S. H. C. (2008) 'Tenofovir use is

associated with an increase in serum alkaline phosphatase in the Swiss HIV Cohort Study', *Antivir Ther*, 13(8), 1077-82.

Gabriel, S. B., Schaffner, S. F., Nguyen, H., Moore, J. M., Roy, J., Blumenstiel, B., Higgins, J., DeFelice, M., Lochner, A., Faggart, M., Liu-Cordero, S. N., Rotimi, C., Adeyemo, A., Cooper, R., Ward, R., Lander, E. S., Daly, M. J. and Altshuler, D. (2002) 'The structure of haplotype blocks in the human genome', *Science*, 296(5576), 2225-9.

Gafni, R. I., Hazra, R., Reynolds, J. C., Maldarelli, F., Tullio, A. N., DeCarlo, E., Worrell, C. J., Flaherty, J. F., Yale, K., Kearney, B. P. and Zeichner, S. L. (2006) 'Tenofovir disoproxil fumarate and an optimized background regimen of antiretroviral agents as salvage therapy: impact on bone mineral density in HIV-infected children', *Pediatrics*, 118(3), e711-8.

Gallant, J. E., DeJesus, E., Arribas, J. R., Pozniak, A. L., Gazzard, B., Campo, R. E., Lu, B., McColl, D., Chuck, S., Enejosa, J., Toole, J. J., Cheng, A. K. and Group, S. (2006) 'Tenofovir DF, emtricitabine, and efavirenz vs. zidovudine, lamivudine, and efavirenz for HIV', *N Engl J Med*, 354(3), 251-60.

Gallant, J. E., Staszewski, S., Pozniak, A. L., DeJesus, E., Suleiman, J. M., Miller, M. D., Coakley, D. F., Lu, B., Toole, J. J. and Cheng, A. K. (2004) 'Efficacy and safety of tenofovir DF vs stavudine in combination therapy in antiretroviral-naive patients: a 3-year randomized trial', *Jama*, 292(2), 191-201.

Gallo, R. C. (1988) 'HIV--the cause of AIDS: an overview on its biology, mechanisms of disease induction, and our attempts to control it', *J Acquir Immune Defic Syndr*, 1(6), 521-35.

Geddes, R., Knight, S., Moosa, M. Y., Reddi, A., Uebel, K. and Sunpath, H. (2006) 'A high incidence of nucleoside reverse transcriptase inhibitor (NRTI)-induced lactic acidosis in HIV-infected patients in a South African context', *S Afr Med J*, 96(8), 722-4.

Genovese, G., Friedman, D. J., Ross, M. D., Lecordier, L., Uzureau, P., Freedman, B. I., Bowden, D. W., Langefeld, C. D., Oleksyk, T. K., Uscinski Knob, A. L., Bernhardy, A. J., Hicks, P. J., Nelson, G. W., Vanhollebeke, B., Winkler, C. A., Kopp, J. B., Pays, E. and Pollak, M. R. (2010) 'Association of trypanolytic ApoL1 variants with kidney disease in African Americans', *Science*, 329(5993), 841-5.

Giacomet, V., Cattaneo, D., Viganò, A., Nannini, P., Manfredini, V., Ramponi, G., Clementi, E. and Zuccotti, G. V. (2013) 'Tenofovir-induced renal tubular dysfunction in vertically HIV-infected patients associated with polymorphisms in ABCC2, ABCC4 and ABCC10 genes', *Pediatr Infect Dis J*, 32(10), e403-5.

Giacomet, V., Mora, S., Martelli, L., Merlo, M., Sciannamblo, M. and Viganò, A. (2005) 'A 12-month treatment with tenofovir does not impair bone mineral accrual in HIV-infected children', *J Acquir Immune Defic Syndr*, 40(4), 448-50.

Gibson, A. K., Shah, B. M., Nambiar, P. H. and Schafer, J. J. (2016) 'Tenofovir Alafenamide: A Review of Its Use in the Treatment of HIV-1 Infection', *Ann Pharmacother*.

Goicoechea, M., Liu, S., Best, B., Sun, S., Jain, S., Kemper, C., Witt, M., Diamond, C., Haubrich, R. and Louie, S. (2008) 'Greater tenofovir-associated renal function decline with protease inhibitor-based versus nonnucleoside reverse-transcriptase inhibitor-based therapy', *J Infect Dis*, 197(1), 102-8.

Gorny, M., Rohm, S., Laer, S., Morali, N. and Niehues, T. (2010) 'Pharmacogenomic adaptation of antiretroviral therapy: overcoming the failure of lopinavir in an African infant with CYP2D6 ultrarapid metabolism', *Eur J Clin Pharmacol*, 66(1), 107-8.

Gounden, V., van Niekerk, C., Snyman, T. and George, J. A. (2010) 'Presence of the CYP2B6 516G> T polymorphism, increased plasma Efavirenz concentrations and early neuropsychiatric side effects in South African HIV-infected patients', *AIDS Res Ther*, 7, 32.

Greig, S. L. and Deeks, E. D. (2016) 'Elvitegravir/Cobicistat/Emtricitabine/Tenofovir Alafenamide: A Review in HIV-1 Infection', *Drugs*, 76(9), 957-68.

Gurdasani, D., Carstensen, T., Tekola-Ayele, F., Pagani, L., Tachmazidou, I., Hatzikotoulas, K., Karthikeyan, S., Iles, L., Pollard, M. O., Choudhury, A., Ritchie, G. R. S., Xue, Y., Asimit, J., Nsubuga, R. N., Young, E. H., Pomilla, C., Kivinen, K., Rockett, K., Kamali, A., Doumatey, A. P., Asiki, G., Seeley, J., Sisay-Joof, F., Jallow, M., Tollman, S., Mekonnen, E., Ekong, R., Oljira, T., Bradman, N., Bojang, K., Ramsay, M., Adeyemo, A., Bekele, E., Motala, A., Norris, S. A., Pirie, F., Kaleebu, P., Kwiatkowski, D., Tyler-Smith, C., Rotimi, C., Zeggini, E. and Sandhu, M. S. (2015) 'The African Genome Variation Project shapes medical genetics in Africa', *Nature*, 517, 327-332.

Haas, D. W., Bartlett, J. A., Andersen, J. W., Sanne, I., Wilkinson, G. R., Hinkle, J., Rousseau, F., Ingram, C. D., Shaw, A., Lederman, M. M., Kim, R. B. and Group, A. A. C. T. (2006) 'Pharmacogenetics of nevirapine-associated hepatotoxicity: an Adult AIDS Clinical Trials Group collaboration', *Clin Infect Dis*, 43(6), 783-6.

Haas, D. W., Gebretsadik, T., Mayo, G., Menon, U. N., Acosta, E. P., Shintani, A., Floyd, M., Stein, C. M. and Wilkinson, G. R. (2009) 'Associations between CYP2B6 polymorphisms and pharmacokinetics after a single dose of nevirapine or efavirenz in African americans', *J Infect Dis*, 199(6), 872-80.

Haas, D. W., Ribaudo, H. J., Kim, R. B., Tierney, C., Wilkinson, G. R., Gulick, R. M., Clifford, D. B., Hulgan, T., Marzolini, C. and Acosta, E. P. (2004) 'Pharmacogenetics of efavirenz and central nervous system side effects: an Adult AIDS Clinical Trials Group study', *AIDS*, 18(18), 2391-400.

Haas, D. W., Smeaton, L. M., Shafer, R. W., Robbins, G. K., Morse, G. D., Labbe, L., Wilkinson, G. R., Clifford, D. B., D'Aquila, R. T., De Gruttola, V., Pollard, R. B., Merigan, T. C., Hirsch, M. S., George, A. L., Donahue, J. P. and Kim, R. B. (2005) 'Pharmacogenetics of long-term responses to antiretroviral regimens containing Efavirenz and/or Nelfinavir: an Adult Aids Clinical Trials Group Study', *J Infect Dis*, 192(11), 1931-42.

Hall., A. M., Hendry., B. M., Nitsch., D. and Connolly., J. O. (2011) 'Tenofovir-Associated Kidney Toxicity in HIV-Infected Patients: A Review of the EvidenceTenofovir-Associated Kidney Toxicity in HIV-Infected Patients: A Review of the Evidence', *American Journal of Kidney Diseases*, 57(5), 773-780.

Han, T. M., Naicker, S., Ramdial, P. K. and Assounga, A. G. (2006) 'A cross-sectional study of HIV-seropositive patients with varying degrees of proteinuria in South Africa', *Kidney Int*, 69(12), 2243-50.

Hawkins, S. and Ball, C. (2007) 'Adverse events experienced by three children taking tenofovir and didanosine in combination', *HIV Med*, 8(6), 411.

Hazra, R., Gafni, R. I., Maldarelli, F., Balis, F. M., Tullio, A. N., DeCarlo, E., Worrell, C. J., Steinberg, S. M., Flaherty, J., Yale, K., Kearney, B. P. and Zeichner, S. L. (2005) 'Tenofovir

disoproxil fumarate and an optimized background regimen of antiretroviral agents as salvage therapy for pediatric HIV infection', *Pediatrics*, 116(6), e846-54.

Hirani, V. N., Raucy, J. L. and Lasker, J. M. (2004) 'Conversion of the HIV protease inhibitor nelfinavir to a bioactive metabolite by human liver CYP2C19', *Drug Metab Dispos*, 32(12), 1462-7.

Hoehe, M. R., Timmermann, B. and Lehrach, H. (2003) 'Human inter-individual DNA sequence variation in candidate genes, drug targets, the importance of haplotypes and pharmacogenomics', *Curr Pharm Biotechnol*, 4(6), 351-78.

Hussain, S., Khayat, A., Tolaymat, A. and Rathore, M. H. (2006) 'Nephrotoxicity in a child with perinatal HIV on tenofovir, didanosine and lopinavir/ritonavir', *Pediatr Nephrol*, 21(7), 1034-6.

Ikediobi, O., Aouizerat, B., Xiao, Y., Gandhi, M., Gebhardt, S. and Warnich, L. (2011) 'Analysis of pharmacogenetic traits in two distinct South African populations', *Hum Genomics*, 5(4), 265-82.

Imaoka, T., Kusuhara, H., Adachi, M., Schuetz, J. D., Takeuchi, K. and Sugiyama, Y. (2007) 'Functional involvement of multidrug resistance-associated protein 4 (MRP4/ABCC4) in the renal elimination of the antiviral drugs adefovir and tenofovir', *Mol Pharmacol*, 71(2), 619-27.

Izzedine, H., Hulot, J. S., Villard, E., Goyenvalle, C., Dominguez, S., Ghosn, J., Valantin, M. A., Lechat, P. and Deray, A. G. (2006) 'Association between ABCC2 gene haplotypes and tenofovir-induced proximal tubulopathy', *J Infect Dis*, 194(11), 1481-91.

Izzedine, H., Hulot, J. S., Vittecoq, D., Gallant, J. E., Staszewski, S., Launay-Vacher, V., Cheng, A., Deray, G. and Team, S. (2005a) 'Long-term renal safety of tenofovir disoproxil fumarate in antiretroviral-naive HIV-1-infected patients. Data from a double-blind randomized active-controlled multicentre study', *Nephrol Dial Transplant*, 20(4), 743-6.

Izzedine, H., Launay-Vacher, V. and Deray, G. (2005b) 'Antiviral drug-induced nephrotoxicity', *Am J Kidney Dis*, 45(5), 804-17.

Izzedine, H., Launay-Vacher, V. and Deray, G. (2005c) 'Fanconi syndrome associated with didanosine therapy', *AIDS*, 19(8), 844-5.

Jallow, M., Teo, Y. Y., Small, K. S., Rockett, K. A., Deloukas, P., Clark, T. G., Kivinen, K., Bojang, K. A., Conway, D. J., Pinder, M., Sirugo, G., Sisay-Joof, F., Usen, S., Auburn, S., Bumpstead, S. J., Campino, S., Coffey, A., Dunham, A., Fry, A. E., Green, A., Gwilliam, R., Hunt, S. E., Inouye, M., Jeffreys, A. E., Mendy, A., Palotie, A., Potter, S., Ragoussis, J., Rogers, J., Rowlands, K., Somaskantharajah, E., Whittaker, P., Widden, C., Donnelly, P., Howie, B., Marchini, J., Morris, A., SanJoaquin, M., Achidi, E. A., Agbenyega, T., Allen, A., Amodu, O., Corran, P., Djimde, A., Dolo, A., Doumbo, O. K., Drakeley, C., Dunstan, S., Evans, J., Farrar, J., Fernando, D., Hien, T. T., Horstmann, R. D., Ibrahim, M., Karunaweera, N., Kokwaro, G., Koram, K. A., Lemnge, M., Makani, J., Marsh, K., Michon, P., Modiano, D., Molyneux, M. E., Mueller, I., Parker, M., Peshu, N., Plowe, C. V., Puijalon, O., Reeder, J., Reyburn, H., Riley, E. M., Sakuntabhai, A., Singhasivanon, P., Sirima, S., Tall, A., Taylor, T. E., Thera, M., Troye-Blomberg, M., Williams, T. N., Wilson, M. and Kwiatkowski, D. P. (2009) 'Genome-wide and fine-resolution association analysis of malaria in West Africa', *Nature Genetics*, 41(6), 657-665.

Jamroziak, K., Balcerczak, E., Calka, K., Piaskowski, S., Urbanska-Rys, H., Salagacka, A., Mirowski, M. and Robak, T. (2009) 'Polymorphisms and haplotypes in the multidrug resistance 1 gene (MDR1/ABCB1) and risk of multiple myeloma', *Leuk Res*, 33(2), 332-5.

Jullien, V., Treluyer, J. M., Rey, E., Jaffray, P., Krivine, A., Moachon, L., Lillo-Le Louet, A., Lescoat, A., Dupin, N., Salmon, D., Pons, G. and Urien, S. (2005) 'Population pharmacokinetics of tenofovir in human immunodeficiency virus-infected patients taking highly active antiretroviral therapy', *Antimicrob Agents Chemother*, 49(8), 3361-6.

Kalra, K., Jarmal, G. and Mishra, N. (2009) 'Drugs impact on CYP-450 enzyme family: A pharmacogenetical study of response variation' in *Indian J Hum Genet*78-83.

Karim, A. Q., Karim, A. S. S., Frohlich, J. A., Grobler, A. C., Baxter, C., Mansoor, L. E., Kharsany, A. B., Sibeko, S., Mlisana, K. P., Omar, Z., Gengiah, T. N., Maarschalk, S., Arulappan, N., Mlotshwa, M., Morris, L. and Taylor, D. (2010) 'Effectiveness and Safety of Tenofovir Gel, an Antiretroviral Microbicide, for the Prevention of HIV Infection in Women', *Science*, 329(5996), 1168-74.

Karim, Q. A., Kharsany, A. B., Frohlich, J. A., Baxter, C., Yende, N., Mansoor, L. E., Mlisana, K. P., Maarschalk, S., Arulappan, N., Grobler, A., Sibeko, S., Omar, Z., Gengiah, T. N., Mlotshwa, M., Samsunder, N. and Karim, S. S. (2011) 'Recruitment of high risk women for HIV prevention trials: baseline HIV prevalence and sexual behavior in the CAPRISA 004 tenofovir gel trial', *Trials*, 12, 67.

Karras, A., Lafaurie, M., Furco, A., Bourgarit, A., Droz, D., Sereni, D., Legendre, C., Martinez, F. and Molina, J. M. (2003) 'Tenofovir-related nephrotoxicity in human immunodeficiency virus-infected patients: three cases of renal failure, Fanconi syndrome, and nephrogenic diabetes insipidus', *Clin Infect Dis*, 36(8), 1070-3.

Kasembeli, A. N., Duarte, R., Ramsay, M., Mosiane, P., Dickens, C., Dix-Peek, T., Limou, S., Sezgin, E., Nelson, G. W., Fogo, A. B., Goetsch, S., Kopp, J. B., Winkler, C. A. and Naicker, S. (2015a) 'APOL1 Risk Variants Are Strongly Associated with HIV-Associated Nephropathy in Black South Africans', *J Am Soc Nephrol*, 26(11), 2882-90.

Kasembeli, A. N., Duarte, R., Ramsay, M. and Naicker, S. (2015b) 'African origins and chronic kidney disease susceptibility in the human immunodeficiency virus era', *World J Nephrol*, 4(2), 295-306.

Kassahun, K., McIntosh, I., Cui, D., Hreniuk, D., Merschman, S., Lasseter, K., Azrolan, N., Iwamoto, M., Wagner, J. A. and Wenning, L. A. (2007) 'Metabolism and disposition in humans of raltegravir (MK-0518), an anti-AIDS drug targeting the human immunodeficiency virus 1 integrase enzyme', *Drug Metab Dispos*, 35(9), 1657-63.

Kauf, T. L., Farkouh, R. A., Earnshaw, S. R., Watson, M. E., Maroudas, P. and Chambers, M. G. (2010) 'Economic efficiency of genetic screening to inform the use of abacavir sulfate in the treatment of HIV', *Pharmacoeconomics*, 28(11), 1025-39.

Kerb, R., Fux, R., Mörike, K., Kremsner, P. G., Gil, J. P., Gleiter, C. H. and Schwab, M. (2009) 'Pharmacogenetics of antimalarial drugs: effect on metabolism and transport', *Lancet Infectious Diseases*, 9(12), 760-774,.

Kimmel, P. L. (2003) 'HIV-associated nephropathy: virologic issues related to renal sclerosis', *Nephrol Dial Transplant*, 18 Suppl 6, vi59-63.

Kimmel, P. L., Barisoni, L. and Kopp, J. B. (2003) 'Pathogenesis and treatment of HIV-associated renal diseases: lessons from clinical and animal studies, molecular pathologic correlations, and genetic investigations', *Ann Intern Med*, 139(3), 214-26.

Kimmel, P. L., Bosch, J. P. and Vassalotti, J. A. (1998) 'Treatment of human immunodeficiency virus (HIV)-associated nephropathy', *Semin Nephrol*, 18(4), 446-58.

Kinai, E. and Hanabusa, H. (2005) 'Renal tubular toxicity associated with tenofovir assessed using urine-beta 2 microglobulin, percentage of tubular reabsorption of phosphate and alkaline phosphatase levels', *AIDS*, 19(17), 2031-3.

Kinai, E. and Hanabusa, H. (2009) 'Progressive renal tubular dysfunction associated with long-term use of tenofovir DF', *AIDS Res Hum Retroviruses*, 25(4), 387-94.

Kiser, J. J., Aquilante, C. L., Anderson, P. L., King, T. M., Carten, M. L. and Fletcher, C. V. (2008a) 'Clinical and genetic determinants of intracellular tenofovir diphosphate concentrations in HIV-infected patients', *J Acquir Immune Defic Syndr*, 47(3), 298-303.

Kiser, J. J., Fletcher, C. V., Flynn, P. M., Cunningham, C. K., Wilson, C. M., Kapogiannis, B. G., Major-Wilson, H., Viani, R. M., Liu, N. X., Muenz, L. R., Harris, D. R., Havens, P. L. and Interventions, A. T. N. f. H. A. (2008b) 'Pharmacokinetics of antiretroviral regimens containing tenofovir disoproxil fumarate and atazanavir-ritonavir in adolescents and young adults with human immunodeficiency virus infection', *Antimicrob Agents Chemother*, 52(2), 631-7.

Klotman, M. E. and Klotman, P. E. (1998) 'AIDS and the kidney. Introduction', *Semin Nephrol*, 18(4), 371-2.

Knobler, S. L., Lemon, S. M., Najafi, M., Burroughs, T. and Infections), I. o. M. U. F. o. E. (2003) 'Factors Contributing to the Emergence of Resistance' in National Academies Press (US).

Ko, W. Y., Rajan, P., Gomez, F., Scheinfeldt, L., An, P., Winkler, C. A., Froment, A., Nyambo, T. B., Omar, S. A., Wambebe, C., Ranciaro, A., Hirbo, J. B. and Tishkoff, S. A. (2013) 'Identifying Darwinian selection acting on different human APOL1 variants among diverse African populations', *Am J Hum Genet*, 93(1), 54-66.

Kopp, J. B., Nelson, G. W., Sampath, K., Johnson, R. C., Genovese, G., An, P., Friedman, D., Briggs, W., Dart, R., Korbet, S., Mokrzycki, M. H., Kimmel, P. L., Limou, S., Ahuja, T. S., Berns, J. S., Fryc, J., Simon, E. E., Smith, M. C., Trachtman, H., Michel, D. M., Schelling, J. R., Vlahov, D., Pollak, M. and Winkler, C. A. (2011) 'APOL1 Genetic Variants in Focal Segmental Glomerulosclerosis and HIV-Associated Nephropathy'.

Li, J., Menard, V., Benish, R. L., Jurevic, R. J., Guillemette, C., Stoneking, M., Zimmerman, P. A. and Mehlotra, R. K. (2012) 'Worldwide variation in human drug-metabolism enzyme genes CYP2B6 and UGT2B7: implications for HIV/AIDS treatment', *Pharmacogenomics*, 13(5), 555-70.

Li, S., Jiao, H., Yu, X., Strong, A. J., Shao, Y., Sun, Y., Altfeld, M. and Lu, Y. (2007) 'Human leukocyte antigen class I and class II allele frequencies and HIV-1 infection associations in a Chinese cohort', *J Acquir Immune Defic Syndr*, 44(2), 121-31.

Likanonsakul, S., Suntisuklappon, B., Nitiyanontakij, R., Prasithsirikul, W., Nakayama, E. E., Shioda, T. and Sangsajja, C. (2016) 'A Single-Nucleotide Polymorphism in ABCC4 Is Associated with Tenofovir-Related Beta2-Microglobulinuria in Thai Patients with HIV-1 Infection', *PLoS One*, 11(1), e0147724.

Lim, S. E., Ponamarev, M. V., Longley, M. J. and Copeland, W. C. (2003) 'Structural determinants in human DNA polymerase gamma account for mitochondrial toxicity from nucleoside analogs', *J Mol Biol*, 329(1), 45-57.

Lim, Y., Lyall, H. and Foster, C. (2015) 'Tenofovir-Associated Nephrotoxicity in Children with Perinatally-Acqui'.

Limou, S., Le Clerc, S., Coulonges, C., Carpentier, W., Dina, C., Delaneau, O., Labib, T., Taing, L., Sladek, R., Deveau, C., Ratsimandresey, R., Montes, M., Spadoni, J. L., Lelievre, J. D., Levy, Y., Therwath, A., Schachter, F., Matsuda, F., Gut, I., Froguel, P., Delfraissy, J. F., Hercberg, S. and Zagury, J. F. (2009) 'Genomewide association study of an AIDS-nonprogression cohort emphasizes the role played by HLA genes (ANRS Genomewide Association Study 02)', *J Infect Dis*, 199(3), 419-26.

Lipkowitz, M. S., Freedman, B. I., Langefeld, C. D., Comeau, M. E., Bowden, D. W., Kao, W. L., Astor, B. C., Bottinger, E. P., Iyengar, S. K., Klotman, P. E., Freedman, R. G., Zhang, W., Parekh, R. S., Choi, M. J., Nelson, G. W., Winkler, C. A. and Kopp, J. B. (2013) 'Apolipoprotein L1 gene variants associate with hypertension-attributed nephropathy and the rate of kidney function decline in African Americans', *Kidney Int*, 83(1), 114-20.

Liptrott, N. J., Pushpakom, S., Wyen, C., Fatkenheuer, G., Hoffmann, C., Mauss, S., Knechten, H., Brockmeyer, N. H., Hopper-Borge, E., Siccardi, M., Back, D. J., Khoo, S. H., Pirmohamed, M. and Owen, A. (2012) 'Association of ABCC10 polymorphisms with nevirapine plasma concentrations in the German Competence Network for HIV/AIDS', *Pharmacogenet Genomics*, 22(1), 10-9.

Liu, R., Paxton, W. A., Choe, S., Ceradini, D., Martin, S. R., Horuk, R., MacDonald, M. E., Stuhlmann, H., Koup, R. A. and Landau, N. R. (1996) 'Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection', *Cell*, 86(3), 367-77.

Lucas, G. M., Lau, B., Atta, M. G., Fine, D. M., Keruly, J. and Moore, R. D. (2008) 'Chronic Kidney Disease Incidence, and Progression to End-Stage Renal Disease, in HIV-Infected Individuals: A Tale of Two Races', *The Journal of Infectious Diseases*, 197(11), 1548-1557.

Lucey, J. M., Hsu, P. and Ziegler, J. B. (2013) 'Tenofovir-related Fanconi's syndrome and osteomalacia in a teenager with HIV', *BMJ Case Rep*, 2013.

Ma, Q. and Lu, A. Y. H. (2011) 'Pharmacogenetics, Pharmacogenomics, and Individualized Medicine', *Pharmacological Reviews*, 63(2), 437-59.

Maarseveen, E. M. v., Rogers, C. C., Trofe-Clark, J., Zuilen, A. D. v. and Mudrikova, T. (2012) 'Drug–Drug Interactions Between Antiretroviral and Immunosuppressive Agents in HIV-Infected Patients After Solid Organ Transplantation: A Review', *http://www.liebertpub.com/apc*.

Maga, G., Radi, M., Gerard, M. A., Botta, M. and Ennifar, E. (2010) 'HIV-1 RT Inhibitors with a Novel Mechanism of Action: NNRTIs that Compete with the Nucleotide Substrate', *Viruses*, 2(4), 880-99.

Maggini, V., Buda, G., Martino, A., Presciuttini, S., Galimberti, S., Orciuolo, E., Barale, R., Petrini, M. and Rossi, A. M. (2008) 'MDR1 diplotypes as prognostic markers in multiple myeloma', *Pharmacogenet Genomics*, 18(5), 383-9.

Maiers, M., Gragert, L. and Klitz, W. (2007) 'High-resolution HLA alleles and haplotypes in the United States population', *Hum Immunol*, 68(9), 779-88.

Mallal, S., Nolan, D., Witt, C., Masel, G., Martin, A. M., Moore, C., Sayer, D., Castley, A., Mamotte, C., Maxwell, D., James, I. and Christiansen, F. T. (2002) 'Association between presence of HLA-B*5701, HLA-DR7, and HLA-DQ3 and hypersensitivity to HIV-1 reverse-transcriptase inhibitor abacavir', *Lancet*, 359(9308), 727-32.

Mallal, S., Phillips, E., Carosi, G., Molina, J. M., Workman, C., Tomazic, J., Jägel-Guedes, E., Rugina, S., Kozyrev, O., Cid, J. F., Hay, P., Nolan, D., Hughes, S., Hughes, A., Ryan, S., Fitch, N., Thorborn, D., Benbow, A. and Team, P.-S. (2008) 'HLA-B*5701 screening for hypersensitivity to abacavir', *N Engl J Med*, 358(6), 568-79.

Mallants, R., Van Oosterwyck, K., Van Vaeck, L., Mols, R., De Clercq, E. and Augustijns, P. (2005) 'Multidrug resistance-associated protein 2 (MRP2) affects hepatobiliary elimination but not the intestinal disposition of tenofovir disoproxil fumarate and its metabolites', *Xenobiotica*, 35(10-11), 1055-66.

Manosuthi, W., Sukasem, C., Thongyen, S., Nilkamhang, S. and Sungkanuparph, S. (2014) 'ABCC2*1C and plasma tenofovir concentration are correlated to decreased glomerular filtration rate in patients receiving a tenofovir-containing antiretroviral regimen', *J Antimicrob Chemother*, 69(8), 2195-201.

Margolis, A. M., Heverling, H., Pham, P. A. and Stolbach, A. (2013) 'A Review of the Toxicity of HIV Medications', *J Med Toxicol*.

Marrazzo, J. M., Ramjee, G., Richardson, B. A., Gomez, K., Mgodi, N., Nair, G., Palanee, T., Nakabiito, C., van der Straten, A., Noguchi, L., Hendrix, C. W., Dai, J. Y., Ganesh, S., Mkhize, B., Taljaard, M., Parikh, U. M., Piper, J., Mâsse, B., Grossman, C., Rooney, J., Schwartz, J. L., Watts, H., Marzinke, M. A., Hillier, S. L., McGowan, I. M. and Chirenje, Z. M. (2015) 'Tenofovir-Based Preexposure Prophylaxis for HIV Infection among African Women', *The New England Journal of medicine*.

Martin, A. M., Nolan, D., Gaudieri, S., Almeida, C. A., Nolan, R., James, I., Carvalho, F., Phillips, E., Christiansen, F. T., Purcell, A. W., McCluskey, J. and Mallal, S. (2004) 'Predisposition to abacavir hypersensitivity conferred by HLA-B*5701 and a haplotypic Hsp70-Hom variant', *Proc Natl Acad Sci U S A*, 101(12), 4180-5.

Martin, A. M., Nolan, D., James, I., Cameron, P., Keller, J., Moore, C., Phillips, E., Christiansen, F. T. and Mallal, S. (2005) 'Predisposition to nevirapine hypersensitivity associated with HLA-DRB1*0101 and abrogated by low CD4 T-cell counts', *AIDS*, 19(1).

Martin, J. L., Brown, C. E., Matthews-Davis, N. and Reardon, J. E. (1994) 'Effects of antiviral nucleoside analogs on human DNA polymerases and mitochondrial DNA synthesis', *Antimicrob Agents Chemother*, 38(12), 2743-9.

Marzolini, C., Paus, E., Buclin, T. and Kim, R. B. (2004a) 'Polymorphisms in human MDR1 (P-glycoprotein): recent advances and clinical relevance', *Clin Pharmacol Ther*, 75(1), 13-33.

Marzolini, C., Tirona, R. G. and Kim, R. B. (2004b) 'Pharmacogenomics of the OATP and OAT families', *Pharmacogenomics*, 5(3), 273-82.

Masebe, T. M., Bessong, P. O., Nwobegahay, J., Ndip, R. N. and Meyer, D. (2012) 'Prevalence of MDR1 C3435T and CYP2B6 G516T polymorphisms among HIV-1 infected South African patients', *Dis Markers*, 32(1), 43-50.

Masimirembwa, C. and Hasler, J. A. (2013) 'Pharmacogenetics in Africa, an Opportunity for Appropriate Drug Dosage Regimens: on the Road to Personalized Healthcare', *CPT: Pharmacometrics & Systems Pharmacology*, 2(5), e45.

May, A., Hazelhurst, S., Li, Y., Norris, S. A., Govind, N., Tikly, M., Hon, C., Johnson, K. J., Hartmann, N., Staedtler, F. and Ramsay, M. (2013) 'Genetic diversity in black South Africans from Soweto', *BMC Genomics*, 14(1), 1.

McDermott, D. H., Peter A. Zimmerman, Florence Guignard, Cynthia A. Kleeberger, Susan F. Leitman, and Philip M. Murphy. (1998) ' "CCR5 promoter polymorphism and HIV-1 disease progression." ', *The Lancet*, 352(9131), 866-870.

McIlleron, H., Meintjes, G., Burman, W. J. and Maartens, G. (2007) 'Complications of antiretroviral therapy in patients with tuberculosis: drug interactions, toxicity, and immune reconstitution inflammatory syndrome', *J Infect Dis*, 196 Suppl 1, S63-75.

Mehta, U., Durrheim, D. N., Blockman, M., Kredo, T., Gounden, R. and Barnes, K. I. (2008) 'Adverse drug reactions in adult medical inpatients in a South African hospital serving a community with a high HIV/AIDS prevalence: prospective observational study', *Br J Clin Pharmacol*, 65(3), 396-406.

Michaud, V., Bar-Magen, T., Turgeon, J., Flockhart, D., Desta, Z. and Wainberg, M. A. (2012) 'The dual role of pharmacogenetics in HIV treatment: mutations and polymorphisms regulating antiretroviral drug resistance and disposition', *Pharmacol Rev*, 64(3), 803-33.

Miller, D. S. (2001) 'Nucleoside phosphonate interactions with multiple organic anion transporters in renal proximal tubule', *J Pharmacol Exp Ther*, 299(2), 567-74.

Mills, A., Arribas, J. R., Andrade-Villanueva, J., DiPerri, G., Van Lunzen, J., Koenig, E., Elion, R., Cavassini, M., Madruga, J. V., Brunetta, J., Shamblaw, D., DeJesus, E., Orkin, C., Wohl, D. A., Brar, I., Stephens, J. L., Girard, P. M., Huhn, G., Plummer, A., Liu, Y. P., Cheng, A. K. and McCallister, S. (2016) 'Switching from tenofovir disoproxil fumarate to tenofovir alafenamide in antiretroviral regimens for virologically suppressed adults with HIV-1 infection: a randomised, active-controlled, multicentre, open-label, phase 3, non-inferiority study', *Lancet Infect Dis*, 16(1), 43-52.

Mini, E. and Nobili, S. (2009) 'Pharmacogenetics: implementing personalized medicine', *Clin Cases Miner Bone Metab*, 6(1), 17-24.

Mofenson, L. M. (2010) 'Protecting the next generation--eliminating perinatal HIV-1 infection' in *N Engl J Med*, United States: 2316-8.

Mok, S. and Minson, Q. (2008) 'Drug-related problems in hospitalized patients with HIV infection', *Am J Health Syst Pharm*, 65(1), 55-9.

Monforte, A. d. A., Testa, L., Adorni, F., Chiesa, E., Bini, T., Moscatelli, G. C., Abeli, C., Rusconi, S., Sollima, S., Balotta, C., Musicco, M., Galli, M. and Moroni, M. (1998) 'Clinical outcome and predictive factors of failure of highly... : AIDS', 12(13), 1631–1637.

Montagnier, L., Chermann, J. C., Barré-Sinoussi, F., Klatzmann, D., Wain-Hobson, S., Alizon, M., Clavel, F., Brun-Vezinet, F., Vilmer, E. and Rouzioux, C. (1984) 'Lymphadenopathy associated virus and its etiological role in AIDS', *Princess Takamatsu Symp*, 15, 319-31.

Motohashi, H., Sakurai, Y., Saito, H., Masuda, S., Urakami, Y., Goto, M., Fukatsu, A., Ogawa, O. and Inui, K. (2002) 'Gene expression levels and immunolocalization of organic ion transporters in the human kidney', *J Am Soc Nephrol*, 13(4), 866-74.

Mouton, J. P., Mehta, U., Parrish, A. G., Wilson, D. P. K., Stewart, A., Njuguna, C. W., Kramer, N., Maartens, G., Blockman, M. and Cohen, K. (2015) 'Mortality from adverse drug reactions in adult medical inpatients at four hospitals in South Africa: a cross-sectional survey', *British Journal of Clinical Pharmacology*, 80(4), 818-826.

Mouton, J. P., Njuguna, C., Kramer, N., Stewart, A., Mehta, U., Blockman, M., Fortuin-De Smidt, M., MSc, De Waal, R., Parrish, A. G., Wilson, D. P. K., Igumbor, E. U., Aynalem, G., Dheda, M., Maartens, G. and Cohen, K. (2016) 'Adverse Drug Reactions Causing Admission to Medical Wards: A Cross-Sectional Survey at 4 Hospitals in South Africa', *medicine*, 95(19), e3437.

Mukonzo, J. K., Owen, J. S., Ogwal-Okeng, J., Kuteesa, R. B., Nanzigu, S., Sewankambo, N., Thabane, L., Gustafsson, L. L., Ross, C. and Aklillu, E. (2014) 'Pharmacogenetic-based efavirenz dose modification: suggestions for an African population and the different CYP2B6 genotypes', *PLoS One*, 9(1), e86919.

Myer, L., Kamkuemah, M., Kaplan, R. and Bekker, L. G. (2013) 'Low prevalence of renal dysfunction in HIV-infected pregnant women: implications for guidelines for the prevention of mother-to-child transmission of HIV', *Trop Med Int Health*.

Naicker, S. (2013) 'Integrated management: chronic kidney disease, diabetes mellitus, hypertension', *African Journal of Nephrology*, 16(1).

Naicker, S. and Fabian, J. (2010) 'Risk factors for the development of chronic kidney disease with HIV/AIDS', *Clin Nephrol*, 74 Suppl 1, S51-6.

Nebert, D. W. and Russell, D. W. (2002) 'Clinical importance of the cytochromes P450', *The Lancet*, 360(9340), 1155-1162.

Nelson, M. R., Katlama, C., Montaner, J. S., Cooper, D. A., Gazzard, B., Clotet, B., Lazzarin, A., Schewe, K., Lange, J., Wyatt, C., Curtis, S., Chen, S. S., Smith, S., Bischofberger, N. and Rooney, J. F. (2007) 'The safety of tenofovir disoproxil fumarate for the treatment of HIV infection in adults: the first 4 years', *AIDS*, 21(10), 1273-81.

Ngaimisi, E., Habtewold, A., Minzi, O., Makonnen, E., Mugusi, S., Amogne, W., Yimer, G., Riedel, K. D., Janabi, M., Aderaye, G., Mugusi, F., Bertilsson, L., Aklillu, E. and Burhenne, J. (2013) 'Importance of ethnicity, CYP2B6 and ABCB1 genotype for efavirenz pharmacokinetics and treatment outcomes: a parallel-group prospective cohort study in two sub-Saharan Africa populations', *PLoS One*, 8(7), e67946.

Nishijima, T., Komatsu, H., Higasa, K., Takano, M., Tsuchiya, K., Hayashida, T., Oka, S. and Gatanaga, H. (2012) 'Single nucleotide polymorphisms in ABCC2 associate with tenofovir-induced kidney tubular dysfunction in Japanese patients with HIV-1 infection: a pharmacogenetic study', *Clin Infect Dis*, 55(11), 1558-67.

Njuguna, C., Stewart, A., Mouton, J. P., Blockman, M., Maartens, G., Swart, A., Chisholm, B., Jones, J., Dheda, M., Igumbor, E. U. and Karen., C. (2016) 'Adverse Drug Reactions Reported to a National HIV & Tuberculosis Healt', *Drug Safety*, 39(2), 159–169.

Nolan, D., Phillips, E. and Mallal, S. (2006) 'Efavirenz and CYP2B6 polymorphism: implications for drug toxicity and resistance', *Clin Infect Dis*, 42(3), 408-10.

Novembre, J., Galvani, A. P. and Slatkin, M. (2005) 'The geographic spread of the CCR5 Delta32 HIV-resistance allele', *PLoS Biol*, 3(11), e339.

Nyakutira, C., Roshammar, D., Chigutsa, E., Chonzi, P., Ashton, M., Nhachi, C. and Masimirembwa, C. (2008) 'High prevalence of the CYP2B6 516G-->T(*6) variant and effect on the population pharmacokinetics of efavirenz in HIV/AIDS outpatients in Zimbabwe', *Eur J Clin Pharmacol*, 64(4), 357-65.

Núñez, M. J., Martín-Carbonero, L., Moreno, V., Valencia, E., Garcia-Samaniego, J., Castillo, J. G. D., Barreiro, P., González-Lahoz, J. and Soriano, V. (2006) 'Impact of Antiretroviral Treatment-Related Toxicities on Hospital Admissions in HIV-Infected Patients', *http://www.liebertpub.com/aid*.

Okwundu, C. I., Uthman, O. A. and Okoromah, C. A. (2012) 'Antiretroviral pre-exposure prophylaxis (PrEP) for preventing HIV in high-risk individuals', *Cochrane Database Syst Rev*, (7), Cd007189.

Parathyras, J., Gebhardt, S., Hillermann-Rebello, R., Grobbelaar, N., Venter, M. and Warnich, L. (2009) 'A pharmacogenetic study of CD4 recovery in response to HIV antiretroviral therapy in two South African population groups', *J Hum Genet*, 54(5), 261-5.

Patel, D., Desai, M., Shah, A. N. and Dikshit, R. K. (2013) 'Early outcome of second line antiretroviral therapy in treatment-experienced human immunodeficiency virus positive patients', *Perspect Clin Res*, 4(4), 215-20.

Penna, G. A., A. Alonci, A. Aguennouz, M. Garufi, A. Cannavò, A. Gerace, D. Alibrandi, A. Musolino, C. (2011) 'MDR-1 polymorphisms (G2677T and C3435T) in B-chronic lymphocytic leukemia: an impact on susceptibility and prognosis | SpringerLink', *medical oncology*, 28(4), 1549–1554.

Penzak, S. R., Kabuye, G., Mugyenyi, P., Mbamanya, F., Natarajan, V., Alfaro, R. M., Kityo, C., Formentini, E. and Masur, H. (2007) 'Cytochrome P450 2B6 (CYP2B6) G516T influences nevirapine plasma concentrations in HIV-infected patients in Uganda', *HIV Med*, 8(2), 86-91.

Peter, J., Mouton, J. and Lehloenya, R. (2015) 'Immune-mediated adverse drug reactions in South Africa : review article', *Current Allergy & Clinical Immunology*, 28(3), 186-190.

Phillips, E., Bartlett, J. A., Sanne, I., Lederman, M. M., Hinkle, J., Rousseau, F., Dunn, D., Pavlos, R., James, I., Mallal, S. A. and Haas, D. W. (2013) 'Associations between HLA-DRB1*0102, HLA-B*5801, and hepatotoxicity during initiation of nevirapine-containing regimens in South Africa', *J Acquir Immune Defic Syndr*, 62(2), e55-7.

Piketty, C., Gerard, L., Chazallon, C., Marcelin, A. G., Clavel, F., Taburet, A. M., Calvez, V., Madelaine-Chambrin, I., Molina, J. M., Aboulker, J. P. and Girard, P. M. (2006) 'Salvage therapy with atazanavir/ritonavir combined to tenofovir in HIV-infected patients with multiple treatment failures: randomized ANRS 107 trial', *Antivir Ther*, 11(2), 213-21.

Pillai, G., Davies, G., Denti, P., Steimer, J. L., McIlleron, H., Zvada, S., Chigutsa, E., Ngaimisi, E., Mirza, F., Tadmor, B. and Holford, N. H. (2013) 'Pharmacometrics: opportunity for reducing disease burden in the developing world: the case of Africa', *CPT Pharmacometrics Syst Pharmacol*, 2, e69.

Pontrelli, G., Cotugno, N., Amodio, D., Zangari, P., Tchidjou, H. K., Baldassari, S., Palma, P. and Bernardi, S. (2012) 'Renal function in HIV-infected children and adolescents treated with tenofovir disoproxil fumarate and protease inhibitors', *BMC Infect Dis*, 12, 18.

Pozniak, A., Arribas, J. R., Gathe, J., Gupta, S. K., Post, F. A., Bloch, M., Avihingsanon, A., Crofoot, G., Benson, P., Lichtenstein, K., Ramgopal, M., Chetchotisakd, P., Custodio, J. M., Abram, M. E., Wei, X., Cheng, A., McCallister, S., SenGupta, D. and Fordyce, M. W. (2016) 'Switching to Tenofovir Alafenamide, Coformulated With Elvitegravir, Cobicistat, and Emtricitabine, in HIV-Infected Patients With Renal Impairment: 48-Week Results From a Single-Arm, Multicenter, Open-Label Phase 3 Study', *J Acquir Immune Defic Syndr*, 71(5), 530-7.

Purswani, M. U., Patel, K., Winkler, C. A., Spector, S. A., Hazra, R., Seage, G. R., Mofenson, L., Karalius, B., Scott, G. B., Van Dyke, R. B., Kopp, J. B. and Study, P. H. C. (2016) 'APOL1 Renal Risk Variants are Associated with Chronic Kidney Disease in Children and Youth with Perinatal HIV infection', *J Acquir Immune Defic Syndr*.

Pushpakom, S. P., Liptrott, N. J., Rodríguez-Nóvoa, S., Labarga, P., Soriano, V., Albalater, M., Hopper-Borge, E., Bonora, S., Di Perri, G., Back, D. J., Khoo, S., Pirmohamed, M. and Owen, A. (2011) 'Genetic variants of ABCC10, a novel tenofovir transporter, are associated with kidney tubular dysfunction', *J Infect Dis*, 204(1), 145-53.

Ramsay, M., Tiemessen, C. T., Choudhury, A. and Soodyall, H. (2011) 'Africa: the next frontier for human disease gene discovery?', *Hum Mol Genet*, 20(R2), R214-20.

Rathbun, R. C. and Liedtke, M. D. (2011) 'Antiretroviral drug interactions: overview of interactions involving new and investigational agents and the role of therapeutic drug monitoring for management', *Pharmaceutics*, 3(4), 745-81.

Ray, A. S., Cihlar, T., Robinson, K. L., Tong, L., Vela, J. E., Fuller, M. D., Wieman, L. M., Eisenberg, E. J. and Rhodes, G. R. (2006) 'Mechanism of Active Renal Tubular Efflux of Tenofovir' in *Antimicrob Agents Chemother*3297-304.

Ray, A. S., Fordyce, M. W. and Hitchcock, M. J. (2016) 'Tenofovir alafenamide: A novel prodrug of tenofovir for the treatment of Human Immunodeficiency Virus', *Antiviral Res*, 125, 63-70.

Ray, P. E., Liu, X. H., Henry, D., Dye, L., Xu, L., Orenstein, J. M. and Schuztbank, T. E. (1998) 'Infection of human primary renal epithelial cells with HIV-1 from children with HIV-associated nephropathy', *Kidney Int*, 53(5), 1217-29.

Reeves-Daniel, A. M., DePalma, J. A., Bleyer, A. J., Rocco, M. V., Murea, M., Adams, P. L., Langefeld, C. D., Bowden, D. W., Hicks, P. J., Stratta, R. J., Lin, J. J., Kiger, D. F., Gautreaux, M. D., Divers, J. and Freedman, B. I. (2011) 'The APOL1 gene and allograft survival after kidney transplantation', *Am J Transplant*, 11(5), 1025-30.

Reid, A. (2016) *South Africa takes bold step to provide HIV treatment for all | UNAIDS* [press release], available: http://www.unaids.org/en/resources/presscentre/pressreleaseandstatementarchive/2016/may/20160513_UTT [accessed 27 Sep 2016].

Reid, A., Stöhr, W., Walker, A. S., Williams, I. G., Kityo, C., Hughes, P., Kambugu, A., Gilks, C. F., Mugyenyi, P., Munderi, P., Hakim, J., Gibb, D. M. and Trial, D. o. A. T. (2008) 'Severe Renal Dysfunction and Risk Factors Associated with Renal Impairment in HIV-Infected Adults in Africa Initiating Antiretroviral Therapy', *Clinical Infectious Diseases*, 46(8), 1271-1281.

Roden, D. M., Jr, A. L. G., Roden, D. M. and Jr, A. L. G. (2002) 'The genetic basis of variability in drug responses', *Nature Reviews Drug Discovery*, 11(11), 37-44.

Rodriguez-Novoa, S., Barreiro, P., Rendón, A., Jiménez-Nacher, I., González-Lahoz, J. and Soriano, V. (2005) 'Influence of 516G>T polymorphisms at the gene encoding the CYP450-2B6 isoenzyme on efavirenz plasma concentrations in HIV-infected subjects', *Clin Infect Dis*, 40(9), 1358-61.

Rodríguez-Nóvoa, S., Barreiro, P., Jiménez-Nacher, I., Rendón, A. and Soriano, V. (2005) 'Pharmacogenetics in HIV therapy', *AIDS Rev*, 7(2), 103-12.

Rodríguez-Nóvoa, S., Barreiro, P., Jiménez-Nácher, I. and Soriano, V. (2006) 'Overview of the pharmacogenetics of HIV therapy', *Pharmacogenomics J*, 6(4), 234-45.

Rodríguez-Nóvoa, S., Cuenca, L., Morello, J., Córdoba, M., Blanco, F., Jiménez-Nácher, I. and Soriano, V. (2010) 'Use of the HCP5 single nucleotide polymorphism to predict hypersensitivity reactions to abacavir: correlation with HLA-B*5701', *J Antimicrob Chemother*, 65(8), 1567-9.

Rodríguez-Nóvoa, S., García-Gascó, P., Blanco, F., González-Pardo, G., Castellares, C., Moreno, V., Jiménez-Nácher, I., González-Lahoz, J. and Soriano, V. (2007) 'Value of the HLA-B*5701 allele to predict abacavir hypersensitivity in Spaniards', *AIDS Res Hum Retroviruses*, 23(11), 1374-6.

Rodríguez-Nóvoa, S., Labarga, P., Soriano, V., Egan, D., Albalater, M., Morello, J., Cuenca, L., González-Pardo, G., Khoo, S., Back, D. and Owen, A. (2009) 'Predictors of kidney tubular dysfunction in HIV-infected patients treated with tenofovir: a pharmacogenetic study', *Clin Infect Dis*, 48(11), e108-16.

Roe, J., Campbell, L. J., Ibrahim, F., Hendry, B. M. and Post, F. A. (2008) 'HIV Care and the Incidence of Acute Renal Failure', *Clinical Infectious Diseases*, 47(2), 242-249.

Rollot, F., Nazal, E.-M., Chauvelot-Moachon, L., Kélaïdi, C., Daniel, N., Saba, M., Abad, S. and Blanche, P. (2003) 'Tenofovir-Related Fanconi Syndrome with Nephrogenic Diabetes Insipidus in a Patient with Acquired Immunodeficiency Syndrome: The Role of Lopinavir-Ritonavir-Didanosine', *Oxford JournalsMedicine & Health Clinical Infectious Diseases*, 37(12), e174-e176.

Rosen, S., Long, L., Fox, M. and Sanne, I. (2008) 'Cost and cost-effectiveness of switching from stavudine to tenofovir in first-line antiretroviral regimens in South Africa', *J Acquir Immune Defic Syndr*, 48(3), 334-44.

Rotger, M., Colombo, S., Furrer, H., Bleiber, G., Buclin, T., Lee, B. L., Keiser, O., Biollaz, J., Décosterd, L., Telenti, A. and Study, S. H. C. (2005) 'Influence of CYP2B6 polymorphism on plasma and intracellular concentrations and toxicity of efavirenz and nevirapine in HIV-infected patients', *Pharmacogenet Genomics*, 15(1), 1-5.

Rotger, M., Tegude, H., Colombo, S., Cavassini, M., Furrer, H., Décosterd, L., Blievernicht, J., Saussele, T., Günthard, H. F., Schwab, M., Eichelbaum, M., Telenti, A. and Zanger, U. M. (2007) 'Predictive value of known and novel alleles of CYP2B6 for efavirenz plasma concentrations in HIV-infected individuals', *Clin Pharmacol Ther*, 81(4), 557-66.

Rungtivasuwan, K., Avihingsanon, A., Thammajaruk, N., Mitruk, S., Burger, D. M., Ruxrungtham, K., Punyawudho, B. and Pengsuparp, T. (2015) 'Influence of ABCC2 and ABCC4 polymorphisms on tenofovir plasma concentrations in Thai HIV-infected patients', *Antimicrob Agents Chemother*, 59(6), 3240-5.

Saag, M., Balu, R., Phillips, E., Brachman, P., Martorell, C., Burman, W., Stancil, B., Mosteller, M., Brothers, C., Wannamaker, P., Hughes, A., Sutherland-Phillips, D., Mallal, S. and Shaefer, M. (2008) 'High sensitivity of human leukocyte antigen-b*5701 as a marker for immunologically confirmed abacavir hypersensitivity in white and black patients', *Clin Infect Dis*, 46(7), 1111-8.

Saitoh, A., Fletcher, C. V., Brundage, R., Alvero, C., Fenton, T., Hsia, K. and Spector, S. A. (2007a) 'Efavirenz pharmacokinetics in HIV-1-infected children are associated with CYP2B6-G516T polymorphism', *J Acquir Immune Defic Syndr*, 45(3), 280-5.

Saitoh, A., Sarles, E., Capparelli, E., Aweeka, F., Kovacs, A., Burchett, S. K., Wiznia, A., Nachman, S., Fenton, T. and Spector, S. A. (2007b) 'CYP2B6 genetic variants are associated with nevirapine pharmacokinetics and clinical response in HIV-1-infected children', *AIDS*, 21(16), 2191-9.

Sax, P. E., Wohl, D., Yin, M. T., Post, F., DeJesus, E., Saag, M., Pozniak, A., Thompson, M., Podzamczer, D., Molina, J. M., Oka, S., Koenig, E., Trottier, B., Andrade-Villanueva, J., Crofoot, G., Custodio, J. M., Plummer, A., Zhong, L., Cao, H., Martin, H., Callebaut, C., Cheng, A. K., Fordyce, M. W. and McCallister, S. (2015) 'Tenofovir alafenamide versus tenofovir disoproxil fumarate, coformulated with elvitegravir, cobicistat, and emtricitabine, for initial treatment of HIV-1 infection: two randomised, double-blind, phase 3, non-inferiority trials', *Lancet*, 385(9987), 2606-15.

Schaaf, B., Aries, S. P., Kramme, E., Steinhoff, J. and Dalhoff, K. (2003) 'Acute Renal Failure Associated with Tenofovir Treatment in a Patient with Acquired Immunodeficiency Syndrome', *Clinical Infectious Diseases*, 37(3), e41-e43.

Scherzer, R., Estrella, M., Li, Y., Choi, A. I., Deeks, S. G., Grunfeld, C. and Shlipak, M. G. (2012) 'Association of tenofovir exposure with kidney disease risk in HIV infection', *AIDS*, 26(7), 867-75.

Schuster, S. C., Miller, W., Ratan, A., Tomsho, L. P., Giardine, B., Kasson, L. R., Harris, R. S., Petersen, D. C., Zhao, F., Qi, J., Alkan, C., Kidd, J. M., Sun, Y., Drautz, D. I., Bouffard, P., Muzny, D. M., Reid, J. G., Nazareth, L. V., Wang, Q., Burhans, R., Riemer, C., Wittekindt, N. E., Moorjani, P., Tindall, E. A., Danko, C. G., Teo, W. S., Buboltz, A. M., Zhang, Z., Ma, Q., Oosthuysen, A., Steenkamp, A. W., Oostuisen, H., Venter, P., Gajewski, J., Zhang, Y., Pugh, B. F., Makova, K. D., Nekrutenko, A., Mardis, E. R., Patterson, N., Pringle, T. H., Chiaromonte, F., Mullikin, J. C., Eichler, E. E., Hardison, R. C., Gibbs, R. A., Harkins, T. T. and Hayes, V. M. (2010) 'Complete Khoisan and Bantu genomes from southern Africa', *Nature*, 463(7283), 943-7.

Schwartz, E. J. and Klotman, P. E. (1998) 'Pathogenesis of human immunodeficiency virus (HIV)-associated nephropathy', *Semin Nephrol*, 18(4), 436-45.

Seedat, F., Martinson, N. A., Motlhaoleng, K., Abraham, P., Mancama, D., Naicker, S. and Variava, E. (2016) 'Acute kidney injury, risk factors and prognosis in hospitalized HIV-infected adults in South Africa, compared by Tenofovir exposure', *AIDS Res Hum Retroviruses*.

Smith, K. Y., Patel, P., Fine, D., Bellos, N., Sloan, L., Lackey, P., Kumar, P. N., Sutherland-Phillips, D. H., Vavro, C., Yau, L., Wannamaker, P., Shaefer, M. S. and team, f. t. H. s. (2009) 'Randomized, double-blind, placebo-matched, multicenter trial... : AIDS'.

Spence, R. A., Kati, W. M., Anderson, K. S. and Johnson, K. A. (1995) 'Mechanism of inhibition of HIV-1 reverse transcriptase by nonnucleoside inhibitors', *Science*, 267(5200), 988-93.

Stöhr, W., Reid, A., Walker, A. S., Ssali, F., Munderi, P., Mambule, I., Kityo, C., Grosskurth, H., Gilks, C. F., Gibb, D. M., Hakim, J. and Team, D. T. (2011) 'Glomerular dysfunction and associated risk factors over 4-5 years following antiretroviral therapy initiation in Africa', *Antivir Ther*, 16(7), 1011-20.

Subbaraman, R., Chaguturu, S. K., Mayer, K. H., Flanigan, T. P. and Kumarasamy, N. (2007) 'Adverse effects of highly active antiretroviral therapy in developing countries', *Clin Infect Dis*, 45(8), 1093-101.

Susantitaphong, P., Cruz, D. N., Cerda, J., Abulfaraj, M., Alqahtani, F., Koulouridis, I. and Jaber, B. L. (2013) 'World incidence of AKI: a meta-analysis', *Clin J Am Soc Nephrol*, 8(9), 1482-93.

Swart, M., Ren, Y., Smith, P. and Dandara, C. (2012) 'ABCB1 4036A>G and 1236C>T Polymorphisms Affect Plasma Efavirenz Levels in South African HIV/AIDS Patients', *Front Genet*, 3, 236.

Swart, M., Skelton, M., Ren, Y., Smith, P., Takuva, S. and Dandara, C. (2013) 'High predictive value of CYP2B6 SNPs for steady-state plasma efavirenz levels in South African HIV/AIDS patients', *Pharmacogenet Genomics*, 23(8), 415-27.

Tishkoff, S. A., Reed, F. A., Friedlaender, F. R., Ehret, C., Ranciaro, A., Froment, A., Hirbo, J. B., Awomoyi, A. A., Bodo, J. M., Doumbo, O., Ibrahim, M., Juma, A. T., Kotze, M. J., Lema, G., Moore, J. H., Mortensen, H., Nyambo, T. B., Omar, S. A., Powell, K., Pretorius, G. S., Smith, M. W., Thera, M. A., Wambebe, C., Weber, J. L. and Williams, S. M. (2009) 'The Genetic Structure and History of Africans and African Americans', *Science*, 324(5930), 1035-44.

Townsend, C. L., Cortina-Borja, M., Peckham, C. S., de Ruiter, A., Lyall, H. and Tookey, P. A. (2008) 'Low rates of mother-to-child transmission of HIV following effective pregnancy interventions in the United Kingdom and Ireland, 2000-2006', *Aids*, 22(8), 973-81.

Tozzi, V. (2010) 'Pharmacogenetics of antiretrovirals', *Antiviral Res*, 85(1), 190-200.

Tozzi, V., Libertone, R. and Liuzzi, G. (2008) 'HIV pharmacogenetics in clinical practice: recent achievements and future challenges', *Curr HIV Res*, 6(6), 544-54.

Tsiodras, S., Mantzoros, C., Hammer, S. and Samore, M. (2000) 'Effects of protease inhibitors on hyperglycemia, hyperlipidemia, and lipodystrophy: A 5-year cohort study', *Archives of Internal Medicine*, 160(13), 2050-2056.

Tzur, S., Rosset, S., Shemer, R., Yudkovsky, G., Selig, S., Tarekegn, A., Bekele, E., Bradman, N., Wasser, W. G., Behar, D. M. and Skorecki, K. (2010) 'Missense mutations in the APOL1 gene are highly associated with end stage kidney disease risk previously attributed to the MYH9 gene', *Hum Genet*, 128(3), 345-50.

Tzur, S., Rosset, S., Skorecki, K. and Wasser, W. G. (2012) 'APOL1 allelic variants are associated with lower age of dialysis initiation and thereby increased dialysis vintage in African and Hispanic Americans with non-diabetic end-stage kidney disease', *Nephrol Dial Transplant*, 27(4), 1498-505.

Ulasi, I. I., Tzur, S., Wasser, W. G., Shemer, R., Kruzel, E., Feigin, E., Ijoma, C. K., Onodugo, O. D., Okoye, J. U., Arodiwe, E. B., Ifebunandu, N. A., Chukwuka, C. J., Onyedum, C. C., Ijoma, U. N., Nna, E., Onuigbo, M., Rosset, S. and Skorecki, K. (2013) 'High population frequencies of APOL1 risk variants are associated with increased prevalence of non-diabetic chronic

kidney disease in the Igbo people from south-eastern Nigeria', *Nephron Clin Pract*, 123(1-2), 123-8.

UNAIDS (2016a) *AIDS by the numbers 2015 | UNAIDS* [online], available: http://www.unaids.org/en/resources/documents/2015/AIDS_by_the_numbers_2015 [accessed 08 Jun 2016].

UNAIDS (2016b) *HIV and AIDS estimates (2015)* [online], available: http://www.unaids.org/en/regionscountries/countries/southafrica [accessed 08/08/17].

Uwai, Y., Ida, H., Tsuji, Y., Katsura, T. and Inui, K. (2007) 'Renal transport of adefovir, cidofovir, and tenofovir by SLC22A family members (hOAT1, hOAT3, and hOCT2)', *Pharm Res*, 24(4), 811-5.

Van Damme, L., Corneli, A., Ahmed, K., Agot, K., Lombaard, J., Kapiga, S., Malahleha, M., Owino, F., Manongi, R., Onyango, J., Temu, L., Monedi, M. C., Mak'Oketch, P., Makanda, M., Reblin, I., Makatu, S. E., Saylor, L., Kiernan, H., Kirkendale, S., Wong, C., Grant, R., Kashuba, A., Nanda, K., Mandala, J., Fransen, K., Deese, J., Crucitti, T., Mastro, T. D. and Taylor, D. (2012) 'Preexposure prophylaxis for HIV infection among African women', *N Engl J Med*, 367(5), 411-22.

Verhelst, D., Monge, M., Meynard, J. L., Fouqueray, B., Mougenot, B., Girard, P. M., Ronco, P. and Rossert, J. (2002) 'Fanconi syndrome and renal failure induced by tenofovir: a first case report', *Am J Kidney Dis*, 40(6), 1331-3.

Viganò, A., Bedogni, G., Manfredini, V., Giacomet, V., Cerini, C., di Nello, F., Penagini, F., Caprio, C. and Zuccotti, G. V. (2011) 'Long-term renal safety of tenofovir disoproxil fumarate in vertically HIV-infected children, adolescents and young adults: a 60-month follow-up study', *Clin Drug Investig*, 31(6), 407-15.

Wang, J., Sönnerborg, A., Rane, A., Josephson, F., Lundgren, S., Ståhle, L. and Ingelman-Sundberg, M. (2006) 'Identification of a novel specific CYP2B6 allele in Africans causing impaired metabolism of the HIV drug efavirenz', *Pharmacogenet Genomics*, 16(3), 191-8.

Whirl-Carrillo, M., USA, S. U. D. o. G. P. A. C., McDonagh, E. M., USA, S. U. D. o. G. P. A. C., Hebert, J. M., USA, S. U. D. o. G. P. A. C., Gong, L., USA, S. U. D. o. G. P. A. C., Sangkuhl, K., USA, S. U. D. o. G. P. A. C., Thorn, C. F., USA, S. U. D. o. G. P. A. C., Altman, R. B., USA, S. U. D. o. G. P. A. C., USA, S. U. D. o. B. a. M. P. A. C., Klein, T. E. and USA, S. U. D. o. G. P. A. C. (2017) 'Pharmacogenomics Knowledge for Personalized Medicine', *Clinical Pharmacology & Therapeutics*, 92(4), 414-417.

WHO, UNICEF and UNAID (2013) 'Global update on HIV treatment 2013: Results, impact and opportunitiesSharePrint '.

Wojnowski, L., Kulle, B., Schirmer, M., Schluter, G., Schmidt, A., Rosenberger, A., Vonhof, S., Bickeboller, H., Toliat, M. R., Suk, E. K., Tzvetkov, M., Kruger, A., Seifert, S., Kloess, M., Hahn, H., Loeffler, M., Nurnberg, P., Pfreundschuh, M., Trumper, L., Brockmoller, J. and Hasenfuss, G. (2005) 'NAD(P)H oxidase and multidrug resistance protein genetic polymorphisms are associated with doxorubicin-induced cardiotoxicity', *Circulation*, 112(24), 3754-62.

Woodward, C. L., Hall, A. M., Williams, I. G., Madge, S., Copas, A., Nair, D., Edwards, S. G., Johnson, M. A. and Connolly, J. O. (2009) 'Tenofovir-associated renal and bone toxicity', *HIV Med*, 10(8), 482-7.

Yimer, G., Ueda, N., Habtewold, A., Amogne, W., Suda, A., Riedel, K. D., Burhenne, J., Aderaye, G., Lindquist, L., Makonnen, E. and Aklillu, E. (2011) 'Pharmacogenetic & pharmacokinetic biomarker for efavirenz based ARV and rifampicin based anti-TB drug induced liver injury in TB-HIV infected patients', *PLoS One*, 6(12), e27810.

Yombi, J. C., Pozniak, A., Boffito, M., Jones, R., Khoo, S., Levy, J. and Post, F. A. (2014) 'Antiretrovirals and the kidney in current clinical practice: renal pharmacokinetics, alterations of renal function and renal toxicity', *Official journal of the international AID society*, 28(5), 621-632.

Zaidan, M., Lescure, F. X., Brochériou, I., Dettwiler, S., Guiard-Schmid, J. B., Pacanowski, J., Rondeau, E., Pialoux, G., Girard, P. M., Ronco, P. and Plaisier, E. (2013) 'Tubulointerstitial nephropathies in HIV-infected patients over the past 15 years: a clinico-pathological study', *Clin J Am Soc Nephrol*, 8(6), 930-8.

Zimmermann, A. E., Pizzoferrato, T., Bedford, J., Morris, A., Hoffman, R. and Braden, G. (2006) 'Tenofovir-Associated Acute and Chronic Kidney Disease: A Case of Multiple Drug Interactions', *Clinical Infectious Diseases*, 42(2), 283-290.

APPENDIX A: ETHICS CLEARANCE

CSIR Research Ethics Committee
PO Box 395 Pretoria 0001 South Africa
Tel: +27 12 841 4060
Fax: +27 12 841 2478
Email:R&DEthics@csir.co.za

29 November 2016

Dear: Dr Dalu Mancama

Extension of the validity of the previously approved Protocol: **A feasibility study to investigate the potential influence of known and novel genetic and protein biomarkers on the clinical outcome of first-line Tenofovir-based Antiretroviral (ARV) drug treatment.**

This is to confirm that the ethical approval validity for this Protocol that has since expired has been extended for another year. The reference number of this research project is REF: 58/2013.

This approval of the extension is granted under the condition that:

1. The researcher remains within the procedures and protocols indicated in the proposal, as well as the additions made to the procedures and protocols as indicated in the responses submitted to the questions of the REC, particularly in terms of any undertakings made and guarantees given.
2. The researcher notes that **any deviations to the approved project/protocol must be submitted to the REC for approval before implementation**.
3. The researcher remains within the parameters of any applicable national legislation, institutional guidelines and scientific standards relevant to the specific field of research.
4. This **approval is valid for one calendar year from the date of this letter.**
5. The researcher submit bi-annual progress reports to the REC
6. The researcher immediately alert the REC of any adverse events that have occurred during the course of the study, as well as the actions that were taken to immediately respond to these events.
7. The researcher alert the REC of any new or unexpected ethical issues that emerged during the course of the study, and how these ethical issues were addressed. If unsure how to respond to these unexpected or new ethical issues as they emerge, the researcher should immediately consult with the REC for advice.
8. The researcher submit a short report to the REC on completion of the research in which it is indicated (i) that the research has been completed; (ii) if any new or unexpected ethical issues emerged during the course of the study; and if so, (iii) how these ethical issues were addressed.

We wish you all of the best with your research project.

Kind regards

Dr Shenuka Singh — (CSIR REC Chair)

Dr Sandile Ncanana — (CSIR REC Secretariat)

APPENDIX B: LIST OF 65 GENES IN NEXT GENERATION SEQUENCING (NGS) PANEL USED FOR THE AmpliSeq™ PRIMER PANEL DESIGN

Table A: AmpliSeq™ gene list

Category or function	Genes
Transporters	*ABCC10, ABCB1, ABCC2, ABCC4, ABCG2, SLC15A2, SLC22A1, SLC22A11, SLC22A6, SLC22A8, SLC35D1, SLCO1A2, SLCO1B1, SLCO1B3, SLCO2A1, SLCO2B1, SLCO3A1, SLCO4A1, SLCO4C1*
Metabolizers or enzyme	*CYP1A1, CYP1A2, CYP2A6, CYP2B6, CYP2C19, CYP2C8, CYP2C9, CYP2D6, CYP2E1, CYP3A4, CYP3A5,CYP3A7, CYP8B1, DPYD, ANPEP, GGT1, GSTM1, UGT1A1, CES1, CES2, FMO1,FMO2,FMO3, POR, NAGA, NAT1, NAT2, SULT1A1*
Kidney function, development and kidney failure biomarker	*B2M, CLU, LCN2, TIMP1, HAVCR1 (or KIM)*
Immune function	*HLAB, HLAC, HLADRB1,IL18*
Transcription factor	*AHR, NR1I2, NR1I3*
other	*AGTR2, AGTRAP, ACE, FGA, FGG, TFF3*

APPENDIX C: CLINICAL RESPONSE FORM (CRF)

Part of this data will be used for a separate project to investigate the proteome of patients undergoing TDF ART

INFORMED CONSENT / DEMOGRAPHICS

Date Informed Consent Signed: ☐☐ ☐☐☐ ☐☐ (DAY / MONTH / YEAR)

Study duration: 1st August 2012 – 31st July 2015

Patient demographics:

Date of Birth: ☐☐ ☐☐☐ ☐☐ (DAY / MONTH / YEAR)

Race: *(check one)* **Ethnicity (if known and consented to by patient)**

- ☐ Black _____
- ☐ White _____
- ☐ Coloured _____
- ☐ Indian _____
- ☐ Asian _____
- ☐ Other (state) _____

Gender

- ☐ Male
- ☐ Female

1

INCLUSION CRITERIA: ***Please check Yes or No.***
If any of the responses below are checked "No ", the patient is NOT eligible for the study.

	Yes	No
1. Of adult age (>18 years)	❑	❑
2. Positive diagnosis for HIV	❑	❑
3. Meets eligibility criteria for ART / already on ART	❑	❑
4. Willing to have blood taken for research	❑	❑
5. Of otherwise sound health	❑	❑

EXCLUSION CRITERIA: ***Please check Yes or No.***
If any of the responses below are checked "Yes ", the patient is NOT eligible for the study.

1. Medication for chronic condition/s; suspected concomitant illness e.g. TB	❑	❑
2. Unwilling to have blood taken for research	❑	❑
3. Poor venous access	❑	❑
4. Unwilling to comply with the study protocols	❑	❑
5. Poor treatment compliance	❑	❑

2

PHYSICAL ATTRIBUTES / BIOLOGICAL TISSUE SAMPLE OBTAINED

Physical attributes

Please enter all information using leading zeros as applicable.

Weight (Kgs)/ BMI ☐☐☐ . ☐

Height (metric) ☐☐☐ . ☐

CD4 cell count (viral load)

Original count & date (if available)……

Current count……

BIOLOGICAL TISSUE SAMPLE

Venous blood sample for gene / protein analysis ☐

Volume (MLs)

Other (describe) _____

_____

_____

Visit Date: ☐☐ DAY ☐☐☐ MONTH ☐☐ YEAR

BIOLOGICAL TISSUE SAMPLE - Renal Function

Venous blood or urine sample for kidney function analysis ☐

Creatinine level…… Glycosuria level……
Proteinuria level…… Bicarbonate level……..
Uric acid level……..... Phosphorus level……..
Potassium level……. Glomerular filtration rate ……

Visit Date: ☐☐ DAY ☐☐☐ MONTH ☐☐ YEAR

Consenting patients who present with ART-induced adverse events during study, or are a control (good responders) volunteer

Venous blood sample taken ☐

Volume (MLs)

State adverse event(s) and intervention taken _____

_____

Visit Date: ☐☐ DAY ☐☐☐ MONTH ☐☐ YEAR

3

MEDICAL HISTORY / MEDICATION

Existing Disease Condition or Allergy *(p/ease enter only ONE abnormal/per line)*	**Year of Diagnosis**
LIST DISEASE UNDER STUDY, IF APPLICABLE	
_____	☐☐☐☐
_____	☐☐☐☐
_____	☐☐☐☐
_____	☐☐☐☐

CONCOMITANT MEDICATIONS

Please list all medications taken during the course of the study. ARVs taken prior to study should be listed as well. Generic and trade names are both acceptable. If medication is ongoing, _X' the Ongoing box. Complete date fields as completely as possible.

Medication **Indication**

Start Date ☐☐ DAY ☐☐☐ MONTH ☐☐ YEAR **Stop Date** ☐☐ DAY ☐☐☐ MONTH ☐☐ YEAR **Ongoing** ☐

Medication **Indication**

Start Date ☐☐ DAY ☐☐☐ MONTH ☐☐ YEAR **Stop Date** ☐☐ DAY ☐☐☐ MONTH ☐☐ YEAR **Ongoing** ☐

Medication **Indication**

Start Date ☐☐ DAY ☐☐☐ MONTH ☐☐ YEAR **Stop Date** ☐☐ DAY ☐☐☐ MONTH ☐☐ YEAR **Ongoing** ☐

Medication **Indication**

Start Date ☐☐ DAY ☐☐☐ MONTH ☐☐ YEAR **Stop Date** ☐☐ DAY ☐☐☐ MONTH ☐☐ YEAR **Ongoing** ☐

4

ADVERSE EVENT REPORT

Onset Date ☐☐ ☐☐☐ ☐☐ **End Date** ☐☐ ☐☐☐ ☐☐ **Ongoing** ☐
DAY MONTH YEAR DAY MONTH YEAR

Severity and event type _____

Action taken _____

Onset Date ☐☐ ☐☐☐ ☐☐ **End Date** ☐☐ ☐☐☐ ☐☐ **Ongoing** ☐
DAY MONTH YEAR DAY MONTH YEAR

Severity and event type _____

Action taken _____

COMPLIANCE TO TREATMENT

Compliancecounselling undertakenprior totherapeutic intervention

☐ Yes ☐ No

On a scale of 0-100 (0 being non-compliant, 100 being fully compliant) at what frequency do you think the patient is taking their medication as prescribed during the preceding months

Months 0- 3	**Months 3-6**	**Months 6-9**	**Months 9-12**
_____	_____	_____	_____

If non-compliant state reason(s) given, duration, and intervention taken

5

END OF STUDY / EARLY TERMINATION

Date of Last study treatment intake: ☐☐ ☐☐☐ ☐☐
DAY MONTH YEAR

Did the patient complete the study completely as described in the protocol? ☐ Yes ☐ No

Date of Withdrawal: ☐☐ ☐☐☐ ☐☐
DAY MONTH YEAR

If the answer to the above question was _No', indicate primary reason for early withdrawal from study:
(Check one box only if the patient did NOT fully complete the study)

Protocol Violation: *(spec)*

Adverse Event/s
Poor Compliance
Lack of Efficacy
Concomitant illness
Other e.g. withdrew consent

6

APPENDIX D: ARTICLE

Targeted ultra-deep sequencing of forty black South Africans to comprehensively map and characterise common and novel variants in the ADME genes

Sibongile Tshabalala BHSc. (Hons)[1,2,3], Ananyo Choudhury Ph.D.[2] Natasha Beeton-Kempen Ph.D.[3], Neil Martinson MBBCh[4,5], Michèle Ramsay Ph.D.[1,2] and Dalu Mancama Ph.D.[3]

1) Department of Human Genetics, School of Pathology, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa.
2) Sydney Brenner Institute for Molecular Bioscience (SBIMB), Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa.
3) CSIR: Biosciences Unit, Pretoria, South Africa.
4) Perinatal HIV Research Unit, Baragwanath Hospital and Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa.
5) Department of Internal Medicine, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa.
Correspondence: Dalu Mancama, CSIR: Biosciences, PO Box 395, Pretoria, South Africa, 0001, +27 12 841 4533, dmancama@csir.co.za

ABSTRACT:

African populations are characterised by high genetic diversity, which provides opportunities for discovering and elucidating novel variants and polymorphisms of clinical importance, especially those affecting therapeutic outcome. Significantly more detailed knowledge is however needed before such populations can begin to take full advantage of increasingly mainstream advances in precision and personalised medicine. Coupled with the need to concisely map and better understand the pharmacological implications of genetic changes in populations of sub-Sharan African ancestry, the aim of this study was to identify and characterize with high confidence the polymorphisms present within 65 of the most important ADME (absorption, disposition, metabolism and excretion) genes. This included elucidating numerous polymorphisms that have not previously been reported. Using ultra-deep next-generation sequencing to screen 40 South African individuals of Bantu ancestry, we identified a total of 1 687 variants. Functional analysis revealed 152 that are novel, of which 20 represent potential loss-of-function mutations across 15 ADME genes. A high level of allele frequency differentiation for these variants was observed when compared with other populations. Notably, many of these variants appear to be pharmacologically important in the pharmacokinetics of a broad range of drugs, including antiretrovirals, chemotherapeutic drugs, antiepileptics, antidepressants and anticoagulants. In-depth analysis of our high-resolution data enabled us to extensively interrogate the pharmacogenetic implications of this genetic diversity. Furthermore, our study illustrated that a more comprehensive understanding of these population-specific differences is needed to facilitate the development of pharmacogenetic-based interventions for optimal drug therapy in patients of African ancestry.

APPENDIX E: NGS SEQUENCING VALIDATION OF d4T COHORT

Sanger sequencing of patients with ambiguous or undetermined calls was employed for further investigation and confirmation. Two potentially novel variants and one known variant were identified after the initial NGS data filtering. The most common variant of rs717620 in most population, (Izzedine *et al.* 2006, Rodríguez-Nóvoa *et al.* 2009, Nishijima *et al.* 2012), was the rare variant in our population. The region flanking the SNP was amplified, purified and sent off for sequencing to Inqaba Biotech. The results are summarised in the table below where there is discordance between the two methods used to genotype the d4T patient samples. The Sanger sequencing electropherograms shown below were of high quality and therefore we conclude that the genotype calling of the Ion Torrent data was incorrect. Unfortunately the two anticipated novel variants appear to be due to an artefact.

A comparison of next generation sequencing (NGS) and Sanger sequencing genotypic calls.

Gene/SNP	Patient sample ID	Genotype		
		NGS		Sanger sequencing
		Ion reporter	Golden Helix genome browser	
ABCC2 c.-24C/T (rs717620)	170 244	C/C	T/T*	C/T
ABCB1 c.360T/C (novel)	170 274	C/T	C/T	T/T
	170 192	T/T	C/T	T/T
	170 357	T/T	T/T	T/T
	170 564	T/T	T/T	T/T
	170 287	T/T	Deletion	T/T
ABCC2 c.2437A/C (novel)	170 077	A/C	A/C	A/A
	170 186	A/C	A/C	A/A
	170 300	A/A	A/A	A/A
	170237	A/A	/	A/A
	170 287	A/A	Deletion	A/A

ABCC2 (rs717620)

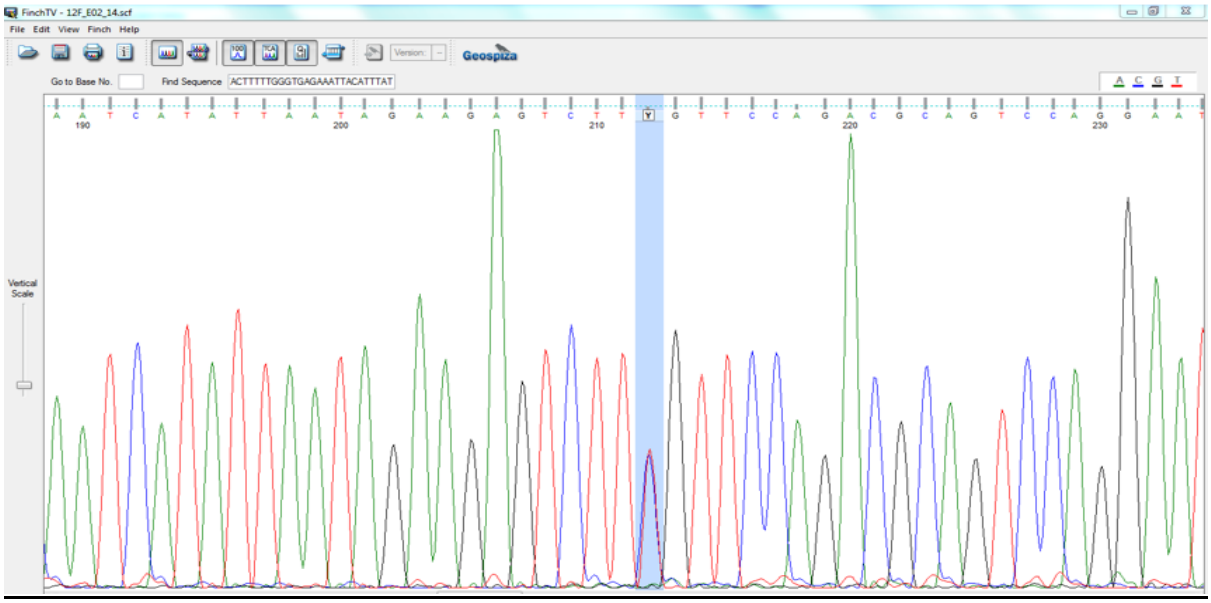
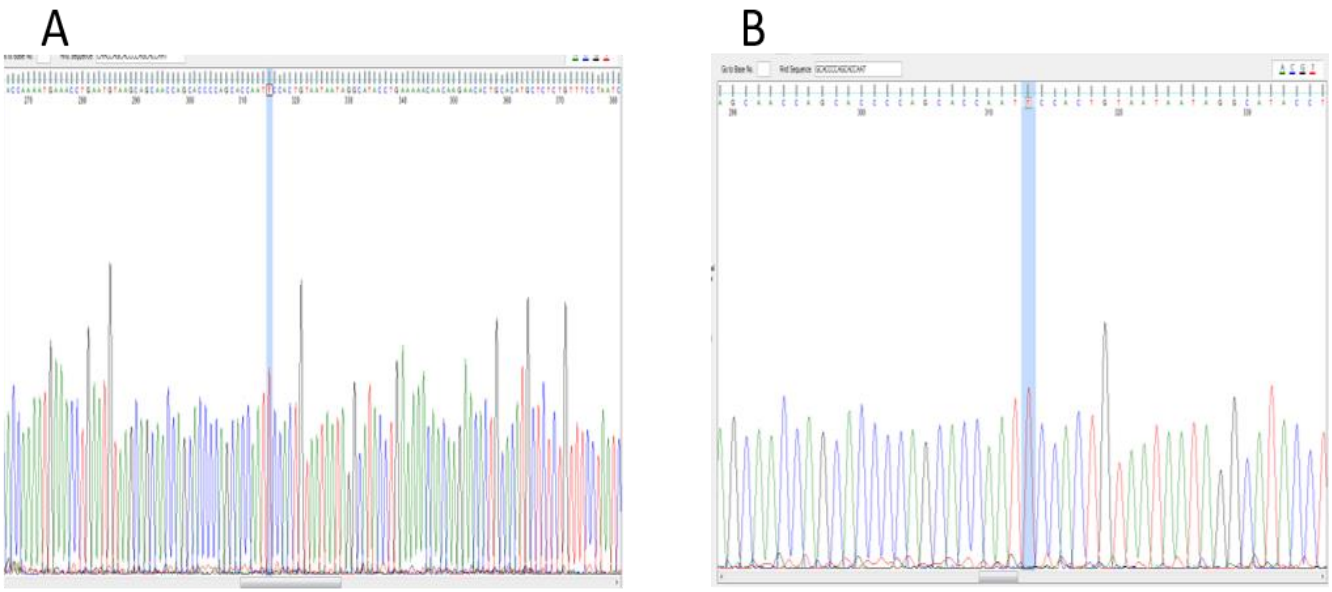


Figure 1.3. Sanger sequencing results (A).
Chromatogram images from the Sanger sequencing analysis of patient sample (170 244); highlighted nucleotide shows the position of variant identified by Ion PGM™ sequencing (Ion reporter and golden helix), (rs717620, ABCC2,c.-24C>T) identified by Ion PGM™ sequencing (Ion reporter and golden helix). Sequencing results indicates that the patient sample has a heterozygous genotype.

ABCB1 novel variant



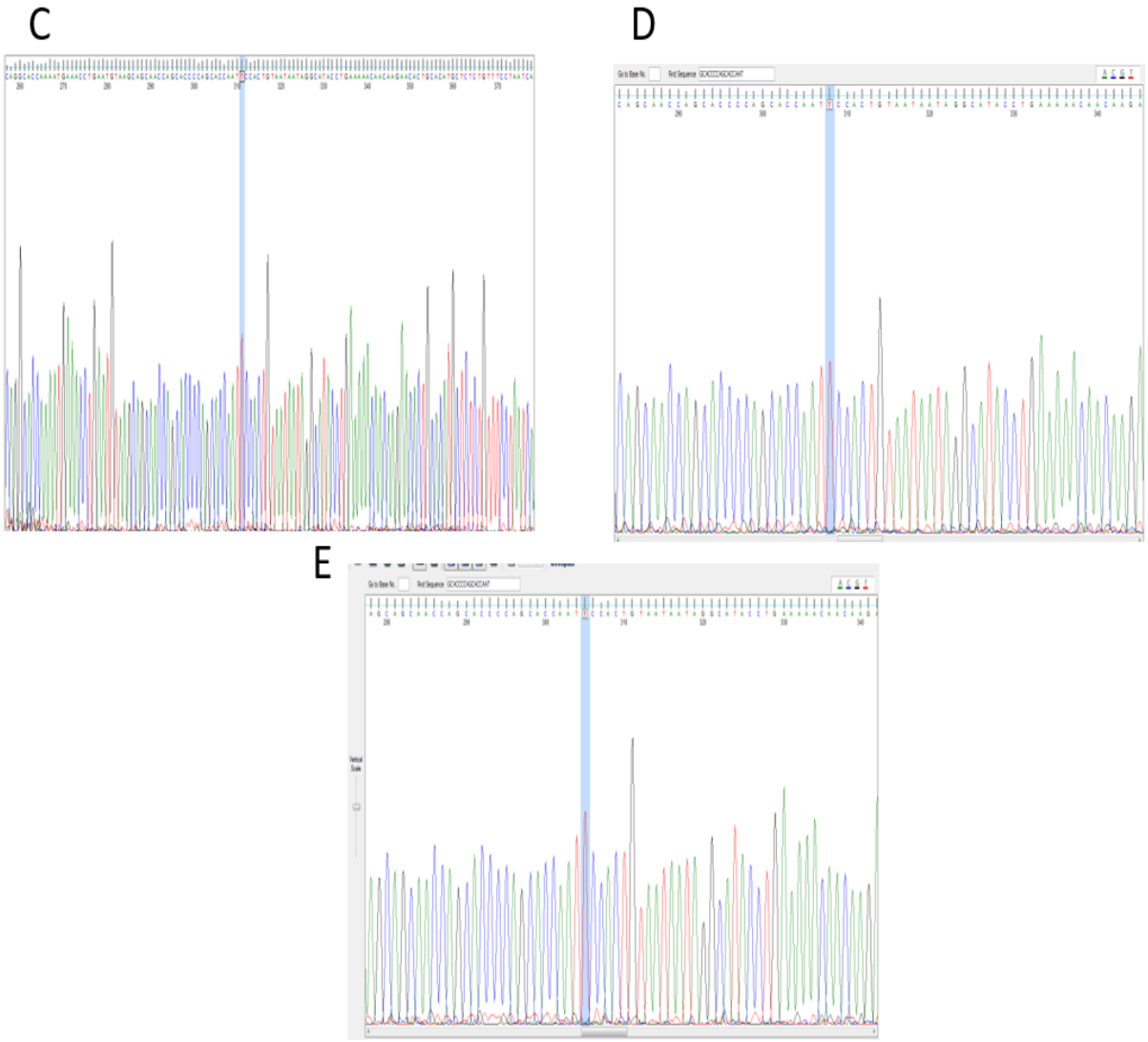


Figure 1.1. Sanger sequencing results (A-E).
Chromatogram images from the sanger sequencing analysis of five patient samples A-E (170 274, 170 192, 170 357, 170 564 and 170 287); the highlighted nucleotide shows the position where a potentially novel variant was identified by Ion PGM^{TM} sequencing (Ion reporter and golden helix), located in ABCB1, c.360A>G. Sequencing results indicated that all selected patient samples don't possess the variant, i.e. were homozygous for the T allele.

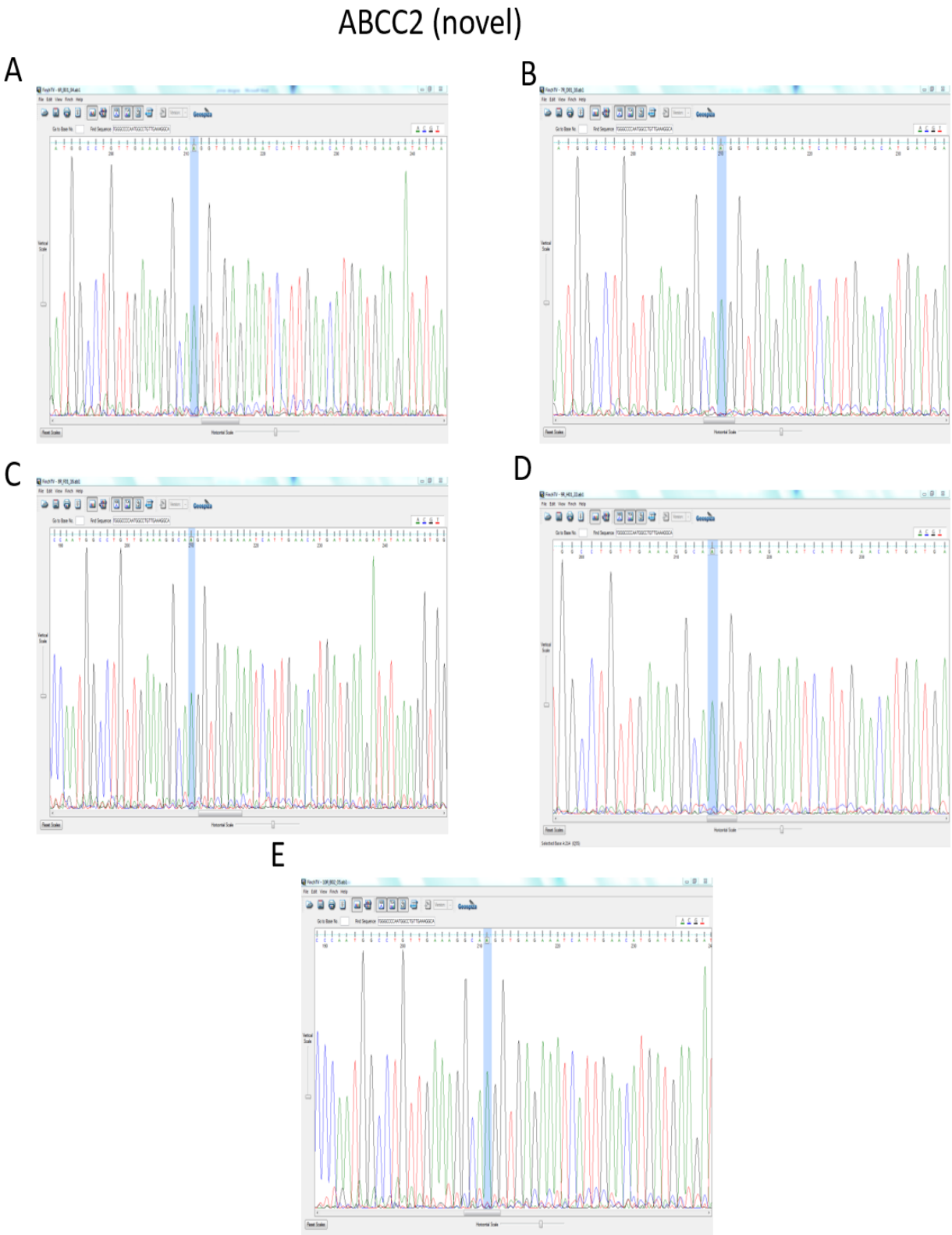
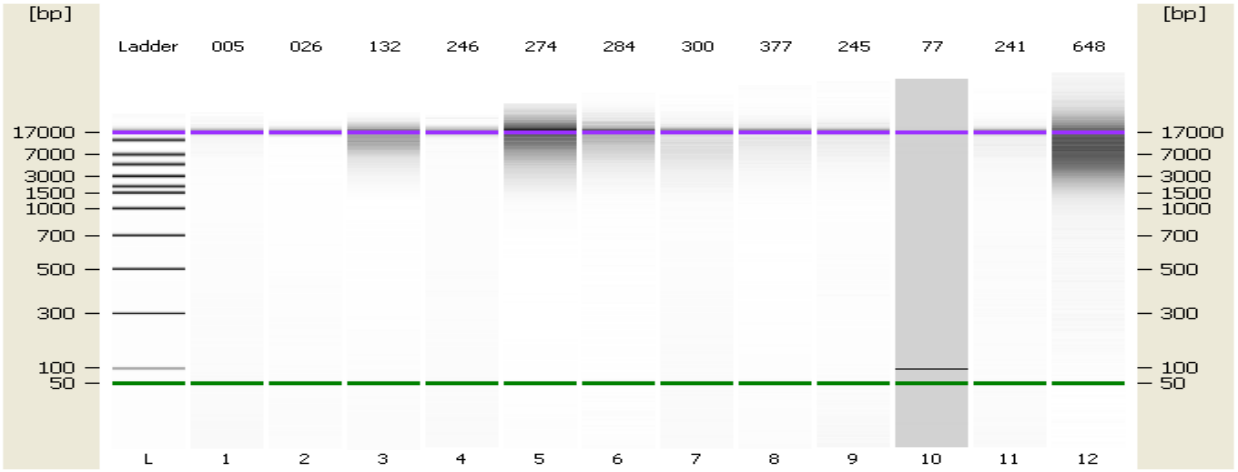


Figure 1.2. Sanger sequencing results (A-E).
Chromatogram images from the Sanger sequencing analysis of five patient sample A-E (170 077, 170 186, 170 300, 170237, 170 287); highlighted nucleotide shows the position where the potentially novel variant was identified by Ion PGM™ sequencing (Ion reporter and golden helix), (located in ABCC2, c.2437A>G). Sequencing results indicates that all patient samples don't possess the variant, i.e. were homozygous for the A allele.

APPENDIX F: gDNA EXTRACTION & QUALITY CONTROL OF d4T COHORT SAMPLES:

gDNA was extracted from buffy coat samples from a cohort of 40 South African HIV positive patients who were undergoing d4T or EFV treatment regimes at the time of blood collection. gDNA was extracted using the QIAgen kit (Qiagen, Hilden, Germany). The quality and quantity of extracted gDNA was analysed using the Nanodrop spectrophotometer (ThermoFisher Scientific, Massachusetts, USA). To improve purity; gDNA was purified with using the ethanol (EtOH) purification method, where 0.1X volume of 3 M sodium acetate and 2X volume of cold 100% EtOH (v/v) was added, mixed well and incubated in the freezer (-20°C) for 20 mins. The samples were then spun at max speed ($14X10^3$ for 15 mins), washed with 70% EtOH (v/v) and air dried for 5-20 mins. The pellets were resuspened in AE elution buffer (QIAgen kit) and kept at 4°C. DNA 12 000 DNA kit and the bioanalyzer (Agilent Technologies, California, USA) were also employed to assess purity and concentration. Since the Nanodrop detects all nucleic acid and Qubit detects only double stranded DNA, the qualities do not correlate exactly (with coefficient of 0.17). Since double stranded DNA is relevant in this case, the Qubit reading for DNA concentration was used. For final assessment of the quantity of each sample the Qubit® dsDNA HS Assay Kit and a Qubit® 2.0 Fluorometer were used.

i.



ii.

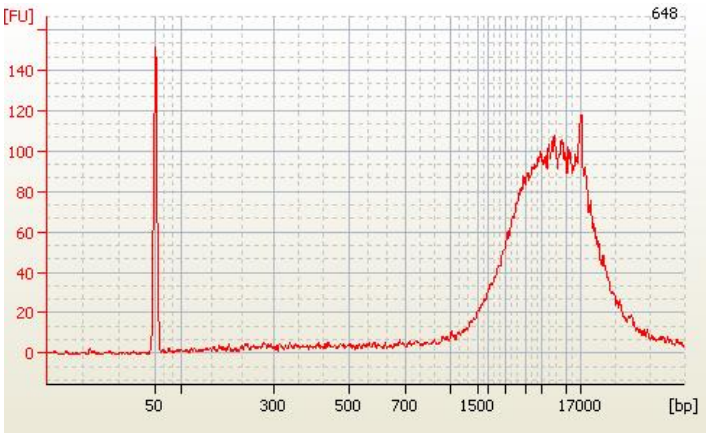


Figure 1: Bioanalyzer analysis of EtOH purified gDNA samples. Gel image (i) generated by the analysis of 12 samples and the electropherogram of sample (ii) sample 648; illustrating the presence of high molecular weight gDNA and the lack of high molecular weight band illustrates that the samples gDNA was intact and not fragmented.

APPENDIX G: EVALUATING THE SALTING OUT METHOD FOR gDNA EXTRACTION

Since the dT4 cohort gDNA samples were extracted using a QIAgen kit (Qiagen, Hilden, Germany) and the yield was low and the samples needed EtOH purification to improve purity; an alternative method of extraction was required. The salting out method (Miller et al., 1996) was selected and evaluated. The salting out method produced high quality DNA and a better yield (Table 1 and Figure 1.1). Restriction enzyme digestion with the *EcoRl* (NEB) restriction enzyme was employed to validate the integrity of the gDNA structure of gDNA extracted using the salting out method and DNA 1000 and the bioanalyzer (Agilent Technologies, California, USA) were used for analysis (Figure 1.2)

Table 1: Salting out purification of gDNA from nucleated cells

Sample ID	Nanodrop™			Qubit (ng/ µl)
	ng/µl	260/280	260/230	
RC49	938.42	1.91	2.14	1140
RC57	369.06	1.57	0.62	1040
RC65	222.50	1.9	2.15	1180
R16	403.58	1.92	2.24	3084
RC41	346.21	1.9	2.23	158.8
R60	75.26	1.9	2.54	356
RC111	96.56	1.35	0.36	116.4

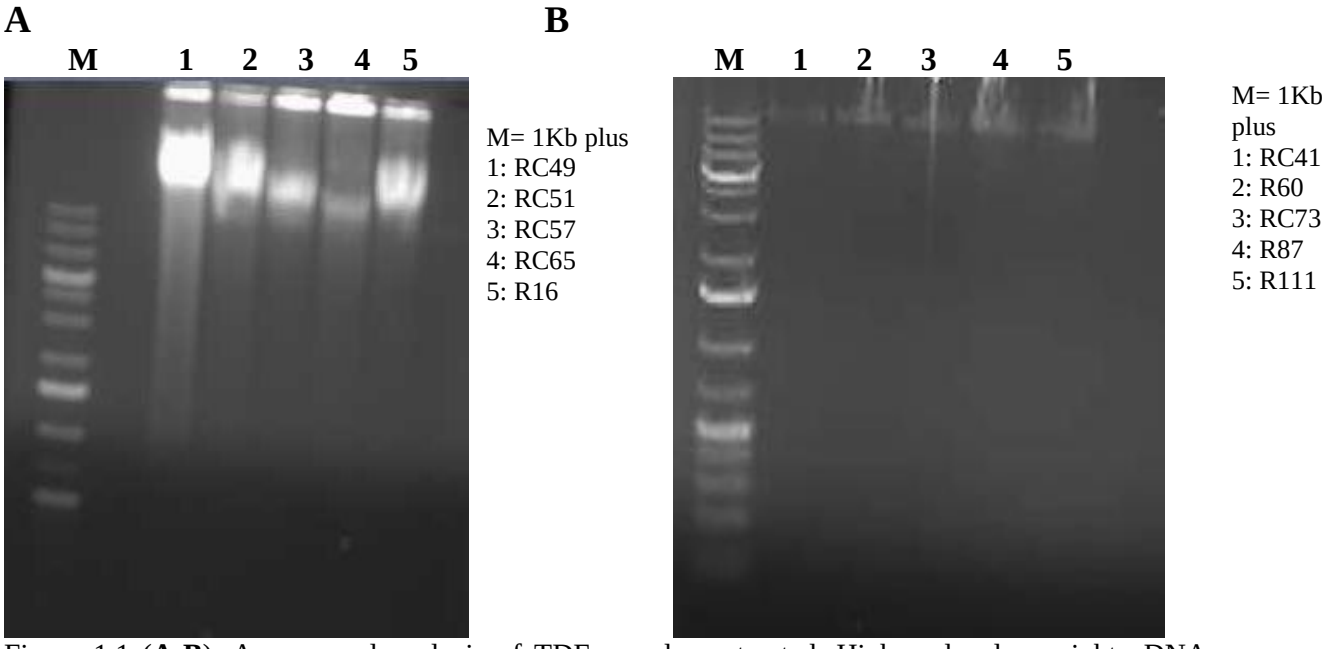
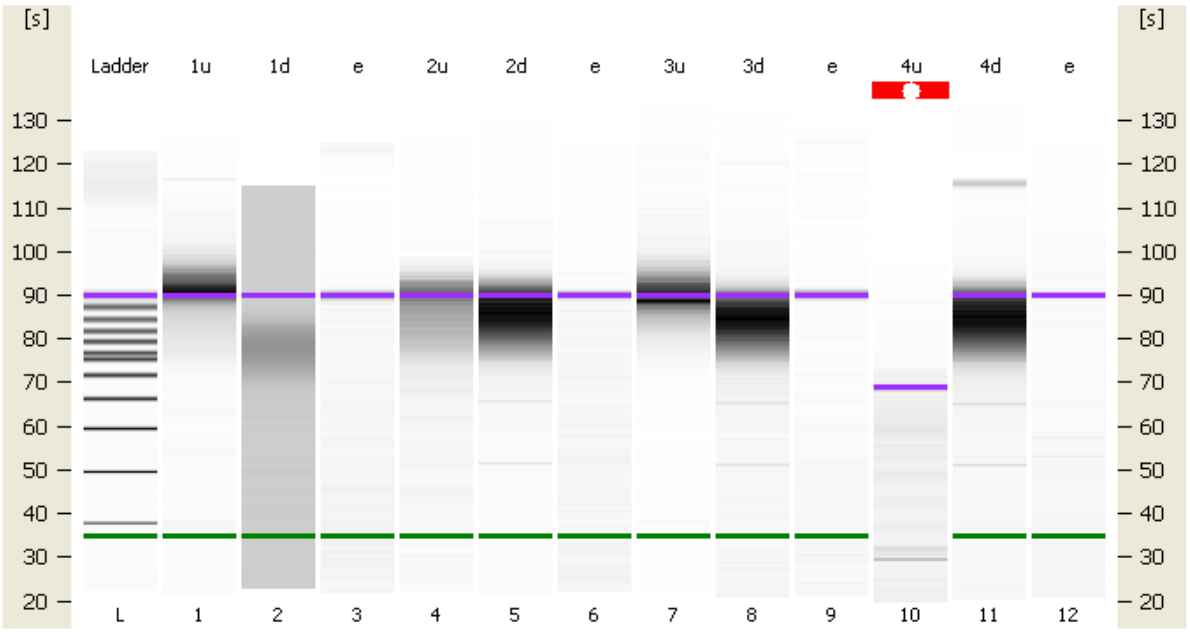
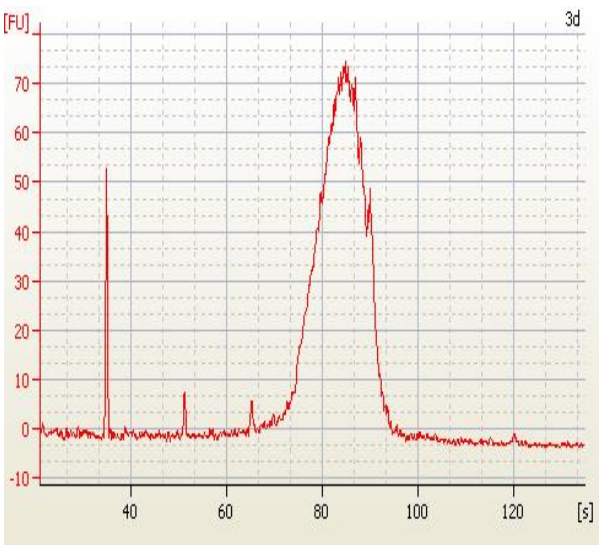


Figure 1.1 (**A-B**). Agarose gel analysis of TDF samples extracted. High molecular weight gDNA produced by salting out extraction from buffy coat samples from the TDF cohort. The high molecular weight single band indicates the absence of degraded fragments, showing the structural integrity of the gDNA extracted. 10µL of extracted gDNA was loading in gel A and 1 µL of gDNA was loaded in gel B.

i.



ii.



iii.

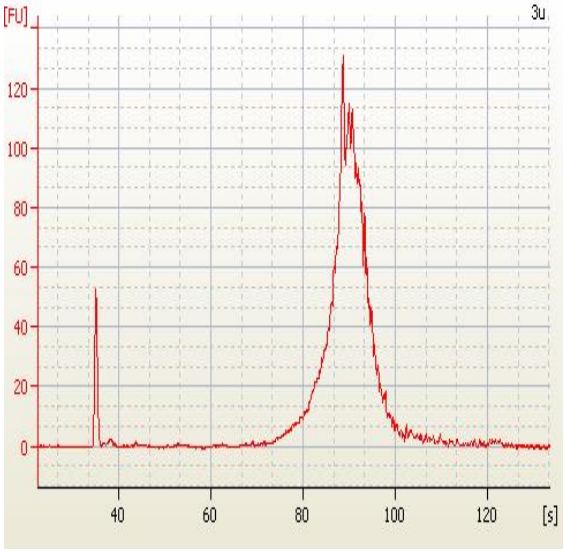


Figure1.2. gDNA structural analysis: Gel image (**i**) generated by the analysis of samples that were digested with EcoRl to assess the structural integrity of the genomic DNA. Four different samples (1-4) were evaluated at different phased (d=digested and u=undigested). Electropherograms of digested (**ii**) and undigested (**iii**) gDNA sample3 are shown.

APPENDIX H: ASSAY OPTIMIZATION

Inorder to improve allelelic discrimination for each assay and produce more defined clusters for each of the genotypes in the assay, increase amplification and fluorescent signals is required. This can be achieved by increasing the starting amount of gDNA. The starting gDNA was increased from 10ng to 20 ng however this did not significantlly lower the Ct (cycle no. at which amplification begins), and therefore 10ng was deamed sufficient as the starting amount gDNA for each assay.

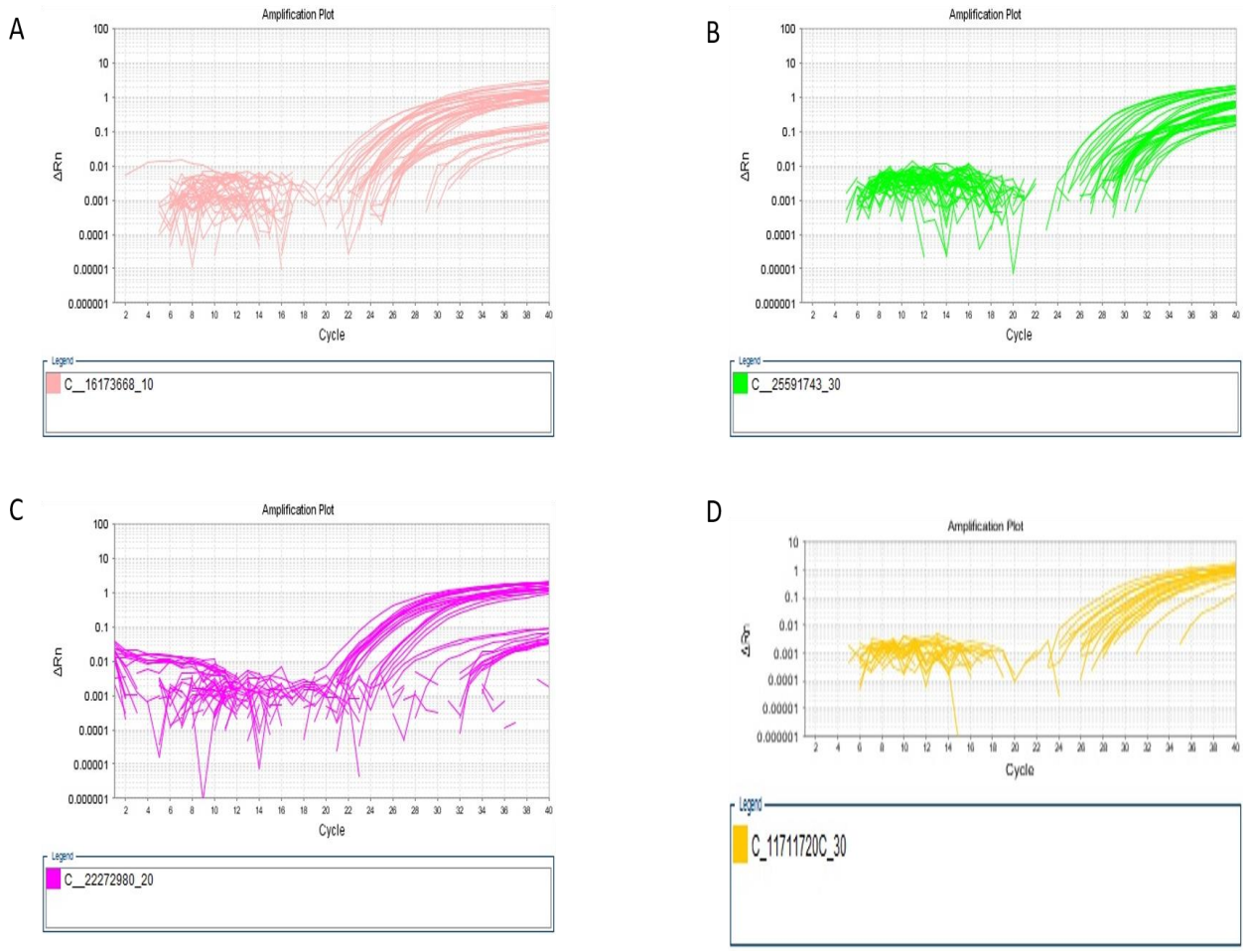


Figure A-D. Amplification plots for four assays tested. **A** was obtained with 10 ng of gDNA; and **B-D** was obtained with 20ng of gDNA. Increasing the strarting gDNA did not significantlly lower the Ct (cycle no. at which amplification begins) therefore increasing the starting gDNA did not increase amplification and fluorescent signals that results in more defined clusters of the genotypes

APPENDIX I: TAQMAN® SNP GENOTYPING ASSAY ROBUSTNESS AND VALIDATION

Stavudine (dT4) cohort samples were used to validate each TaqMan® assay. Genotypic calls from the ion reporter software and golden helix were used as a reference. This comparison for each assay is displayed in the Tables 1-7. These results show concordance for 6 of the 7 assays. The TaqMan® assays were further evaluated by assessing the genotypic. Calls obtained from screening the TDF cohort where both the TaqMan® and Sanger sequencing calls were compared and the outcomes are displayed in the figures below. The results indicate concordance between the TaqMan® and Sanger sequencing genotypic calls.

1. *ABCB1* rs2032582

Assessment and validation of TaqMan SNP genotyping assay C_11711720C_30: ABCB1 rs2032582

SAMPLES ID	GENOTYPE		
	Ion reporter	Golden helix	TaqMan
456	C/C	-	C/C
4	C/C	C/C	C/C
314	C/C	C/C	C/C
351	C/C	C/C	C/C
245	-	C/C	C/C
263	-	-	C/C
246	-	C/C	C/C
349	-	-	C/C
213	C/C	C/C	C/C
631	C/C	C/C	C/C

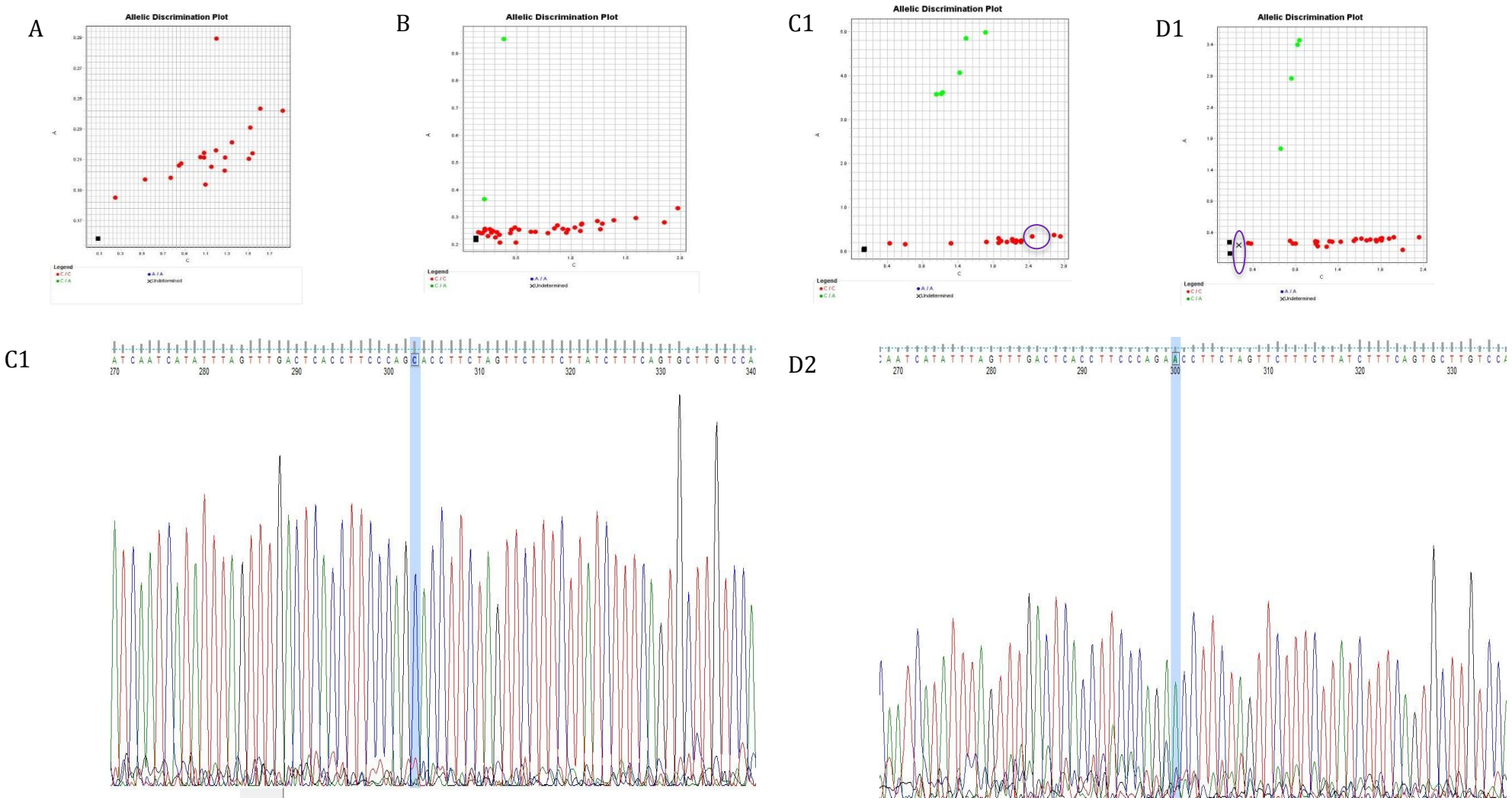


Figure 1. rs2032582 assay validation. Allelic discrimination plot for the d4T cohort (**A**), No clusters were formed because this variant is rare and a larger sample size was required to identify the variant. The calls were in complete concordance with the genotypic calls for Ion importer and Golden Helix. Allelic discrimination plot for the TDF cohort (**B**), homozygous individuals with the wild-type (WT) genotype (CC) are clustered in red. Those possessing heterozygous form of the genotype (AC) are cluster in green. The squares represent experimental control (Negative control). Patients with the 22 forms of the genotype were separated but due to the lack of the variant genotype (AA), clustering was not good. Allele discrimination plots (**C1** & **D1**) and chromatogram (**C2** & **D2**) images for samples RC46 (1) and R73 (2) from th TDF cohort. Tthe highlighted nucleotide shows the position where the variant rs2032582 lied and sequencing. Sanger sequencing confirmed the TaqMan genotypic calls. For patient RC46. The TaqMan assay signals (FAM/VIC) for R73 were low but the assays were sensitive enough to cluster accurately and the genotypic call was confirmed by Sanger seqiencing. rs2032582 islocated in ABCB1 position c.2677.

2. *ABCC10* rs2125739

Assessment and validation of TaqMan® SNP genotyping assay C__16173668_10 : ABCC10 rs2125739

SAMPLES ID	GENOTYPE		
	Ion reporter	Golden helix	TaqMan
192	T/C	T/C	T/C
286	T/C	T/C	T/C
259	T/C	T/C	T/C
9	T/C	T/C	T/C
249	C/C	C/C	C/C
274	C/C	C/C	C/C
476	C/C	C/C	C/C
648	C/C	C/C	C/C
12	T/T	T/T	T/T
564	T/T	T/T	T/T

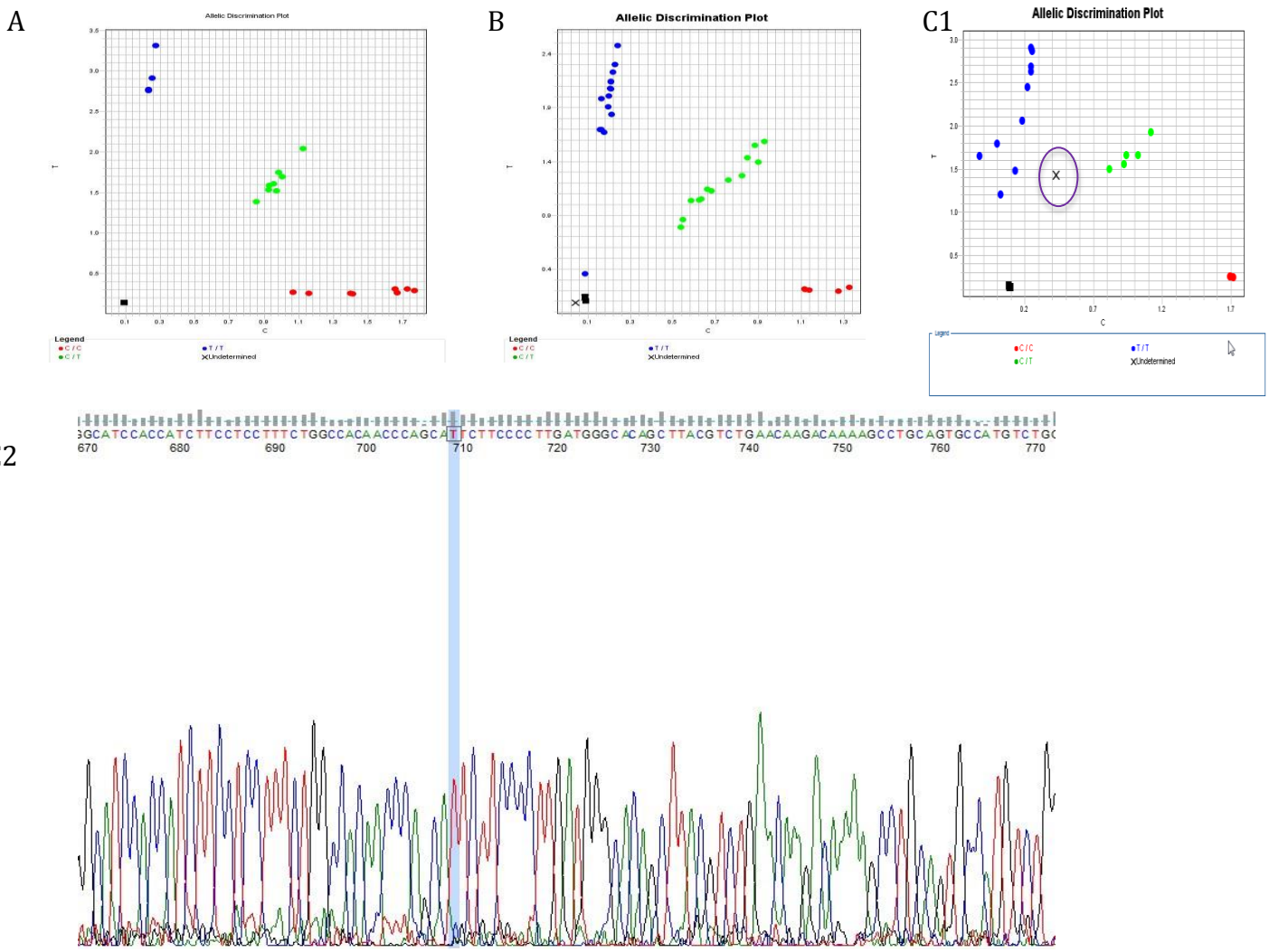


Figure 2. rs212739 assay validation. Allelic discrimination plot for the d4T cohort (**A**) and TDF cohort (**B**). Homozygous individuals with the wild-type allele (TT) are clustered in blue. Those possessing heterozygous form of the allele (CT) are cluster in green. The red cluster indicates the patients that are homozygous for the variant allele (CC). The squares represent experimental control (Negative control) and the individual marked X could not be analysed due to low or no amplification and were therefore repeated. The genotypic calls from the d4T cohort were in accordance with the genotypic calls for Ion reporter and Golden Helix. Allele discrimination plots (**C1**) and chromatogram (**C2**) images for RC77; the highlighted nucleotide shows the position where the variant rs212739 lied. The TaqMan assay signals (FAM/VIC) for R77 were undeterminded (X) low but the assays were sensitive

3. *ABCC2* rs2273697

Assessment and validation of TaqMan SNP genotyping assay C__22272980_20 : *ABCC2* rs2273697

SAMPLES ID	GENOTYPE		
	Ion reporter	Golden helix	TaqMan
294	A/A	A/A	A/A
12	G/A	G/A	G/A
249	G/A	G/A	G/A
383	G/A	G/A	G/A
213	G/A	G/A	G/A
12	G/A	G/A	G/A
648	G/G	G/G	G/G
192	G/G	G/G	G/G
9	G/G	G/G	G/G
241	G/G	G/G	G/G

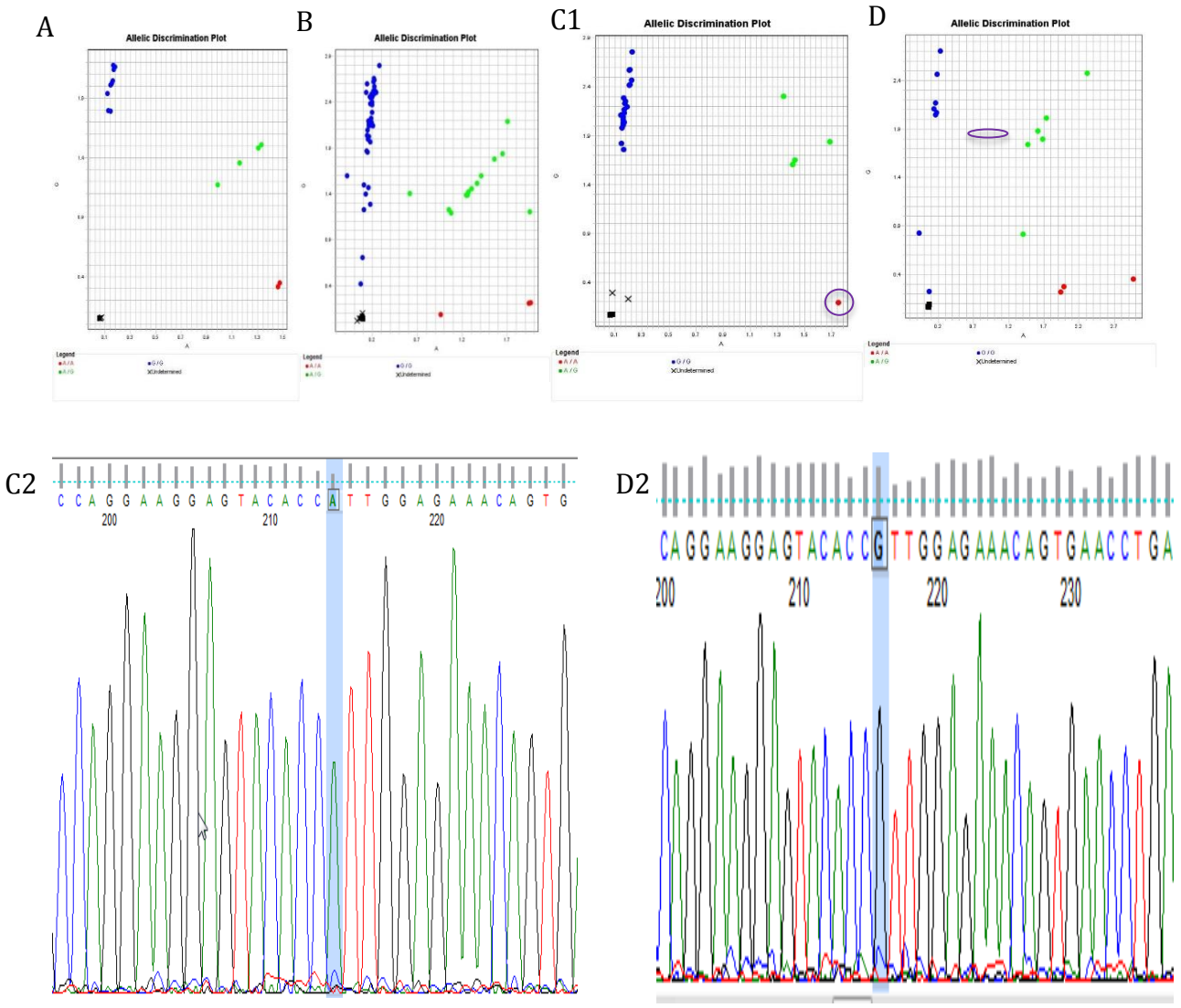


Figure 3. rs2273697 assay validation. Allelic discrimination plot for the d4T cohort (**A**) and TDF cohort (**B**). Individuals homozygous for the wild-type allele (GG) are clustered in blue. Those possessing heterozygous form of the allele (GA) cluster in green. Individuals homozygous for the mutated allele (AA) cluster in red. The squares represent experimental control (Negative control). X represents empty wells, as the genotype cannot be clearly distinguished. Genotyping calls from the d4T cohort where the calls were in accordance with the genotypic calls for Ion importer and Golden Helix. Allele discrimination plots (**C1** & **D1**) and chromatogram (**C2** & **D2**) images for samples RC66 (1) and R26 (2) from th TDF cohort. The highlighted nucleotide shows the position of the variant rs2273697. Sanger sequencing confirmed the TaqMan genotypic calls.

4. *ABCC2* rs17222723

Assessment and validation of TaqMan SNP genotyping assay C__25591743_30 : *ABCC2* rs17222723

SAMPLES ID	GENOTYPE		
	Ion reporter	Golden helix	TaqMan
284	T/A	T/A	T/A
631	T/A	T/A	T/A
241	T/A	T/A	T/A
245	T/A	T/A	T/A
192	T/T	T/T	T/T
249	T/T	T/T	T/T
648	T/T	T/T	T/T
213	T/T	T/T	T/T
351	T/T	T/T	T/T
9	T/T	T/T	T/T

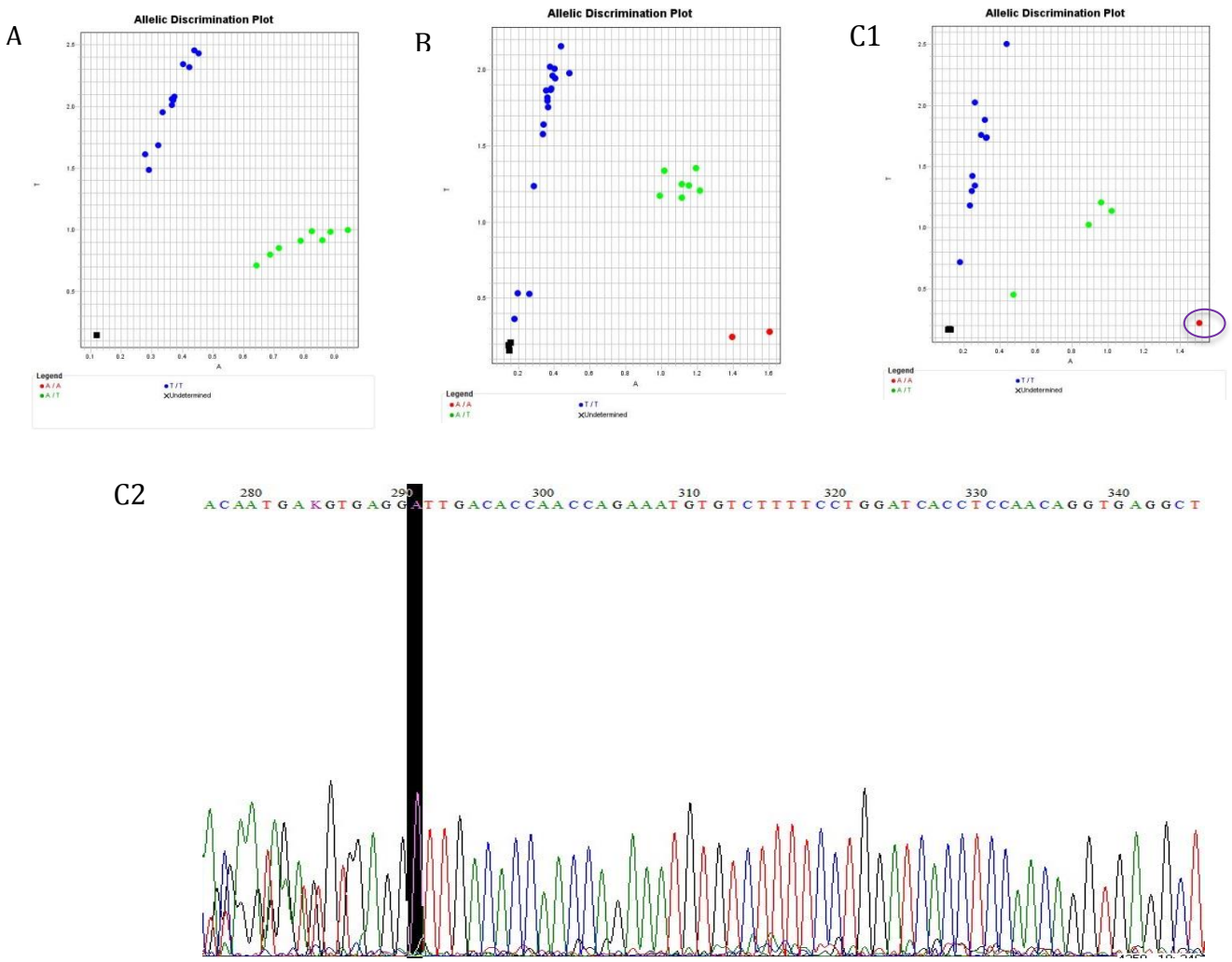


Figure 4. rs17222723 assay validation. Allelic discrimination plot for the d4T cohort (**A**). Genotyping calls from the d4T cohort where the calls were in accordance with the genotypic calls for Ion importer and Golden Helix. Individuals homozygous for the wild-type allele (TT) clustered in blue. Those possessing heterozygous forms of the allele (AT) cluster in green. The squares represent experimental control (Negative control). Allelic discrimination plot for the TDF cohort (**B**), where an increase in sample size resulted in the variants genotype being identified and the plot displayed good clustering of the patients based on their genotype. Allele discrimination plots (C1) and chromatogram (C2) images for samples R26 (1) from th TDF cohort. Tthe highlighted nucleotide shows the position of the variant rs17222723 and Sanger sequencing confirmed the TaqMan genotypic calls.

5. *ABCC2* rs3740066

Assessment and validation of TaqMan SNP genotyping assay C__11214910_20 : *ABCC2* rs3740066

SAMPLES ID	GENOTYPE		
	Ion reporter	Golden helix	TaqMan
259	-	C/T (2X)	C/T
192	-	T/T (1X)	C/C
274	-	T/T (3X)	C/C
319	-	T/T (1C/3T)	C/C
77	-	T/T (1C/5T)	C/C
314	-	C/C (5X)	C/C
241*	-	C/C (1X)	T/T
631	-	C/C (10X)	C/C
249	-	C/C (10X)	C/C
12	-	C/C (1X)	C/C

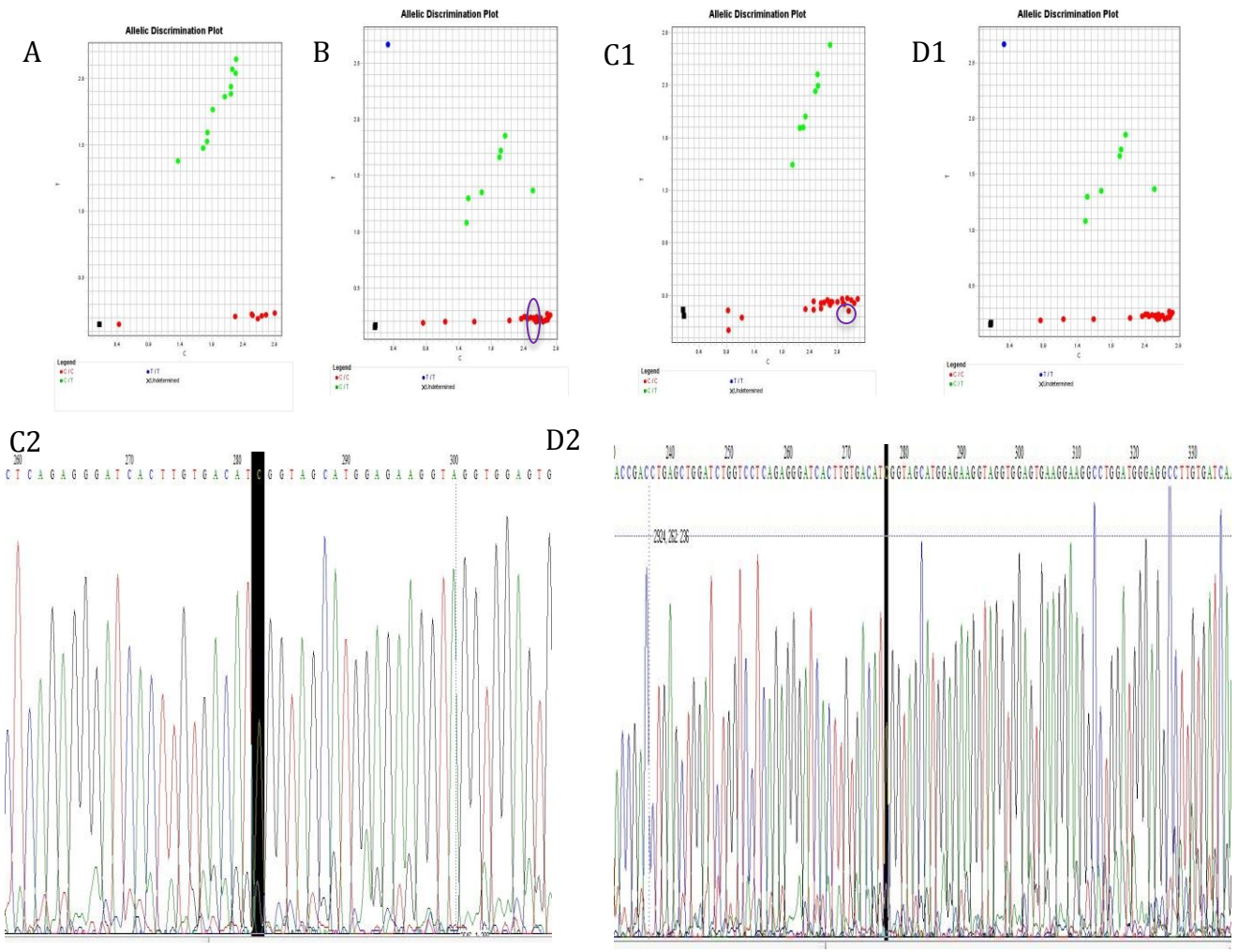


Figure 5. rs3740066 assay validation. Allelic discrimination plot for the d4T cohort (A). Individuals homozygous for the wild-type allele (CC) clustered in red. Those possessing heterozygous form of the allele (CT) cluster in green. The squares represent experimental control (Negative control). The genotypic calls were in accordance with 50 % of the Ion reporter and Golden Helix calls and this is due very low sequencing depth or coverage. The coverage is listed in the parenthesis and the samples with low coverage were incorrectly called. Allelic discrimination plot for the TDF cohort (B), where the plot displayed good clustering of the patients based on their genotype. Allele discrimination plots (C1) and chromatogram (C2) images for samples RC91 (1) and RC66 (2) from the TDF cohort. The highlighted nucleotide shows the position of the variant rs3740066 and Sanger sequencing confirmed the TaqMan genotypic calls.

6. *ABCC4* rs3742106

Assessment and validation of TaqMan SNP genotyping assay C___8059522_1_:
ABCC4 rs3742106

SAMPLES ID	GENOTYPE		
	Ion reporter	Golden helix	TaqMan
456	A/A	-	A/A
300*	A/A	A/A	-
319	A/C	A/C	A/C
631	A/A	A/A	A/A
648	A/C	A/C	A/C
274	A/C	A/C	A/C
357	C/C	C/C	A/A
12*	C/C	C/C	-
333*	C/C	C/C	-
314	C/C	C/C	C/C

*omitted replicates with no amplification to produce good clustering.

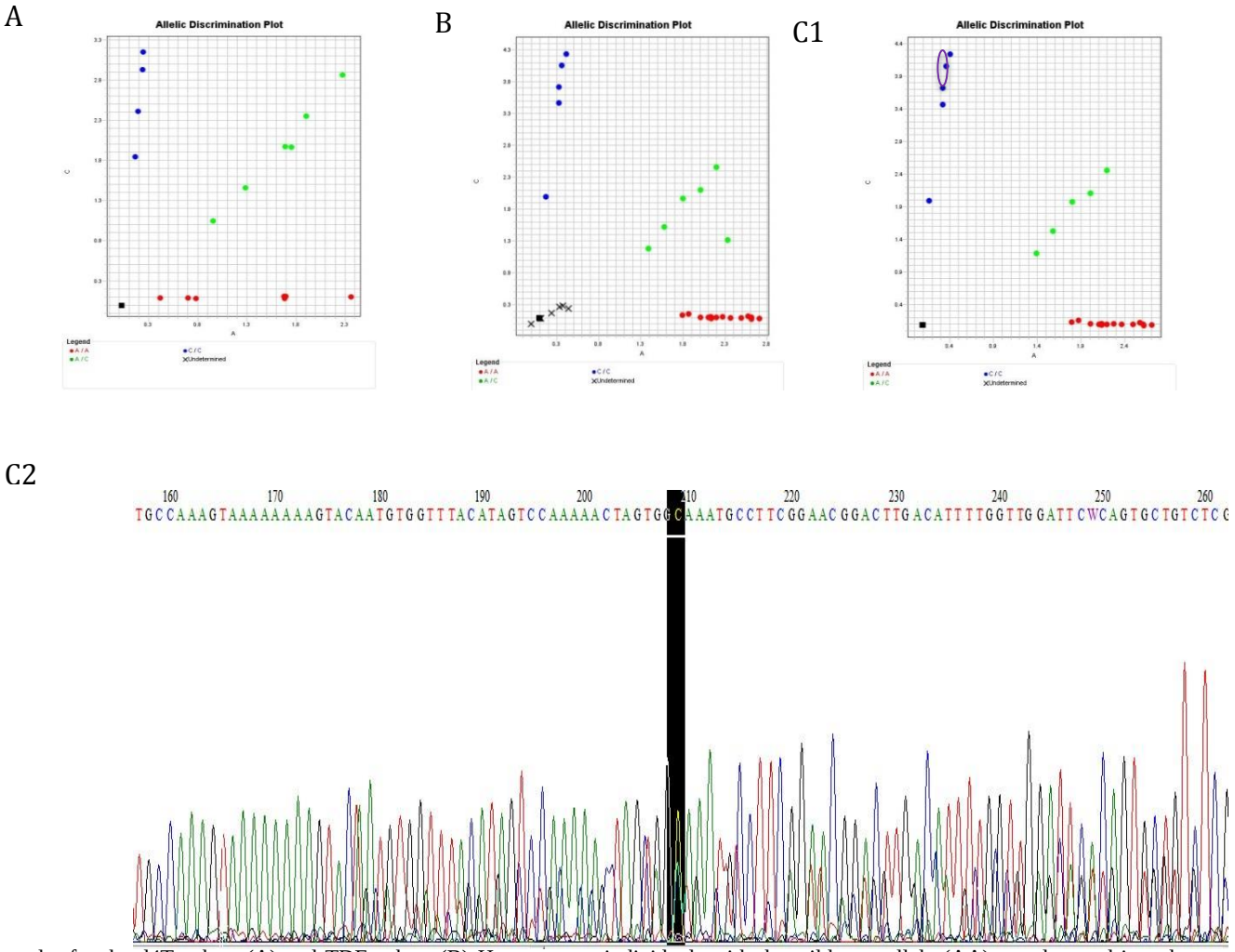


Figure 6. rs3742106 assay validation. Allelic discrimination plot for the d4T cohort (A) and TDF cohort (B) Homozygous individuals with the wild-type allele (AA) are clustered in red. Those possessing heterozygous form of the allele (AC) cluster in green. The blue cluster indicates the patients that are homozygous for the variant allele (CC). The squares represent experimental control (Negative control) and the individuals marked X could not be analysed due to low or no amplification and were therefore repeated. The calls were in accordance with 90% the genotypic calls for Ion importer and Golden Helix. Sample 357 was incorrectly called AA and 3 samples were omitted replicates with no amplification to produce good clustering. Allele discrimination plots (C1) and chromatogram (C2) images for patient samples RC46; the highlighted nucleotide shows the position where the variant rs3742106 was identified by Sanger sequencing, which confirmed the TaqMan genotypic calls for RC46.

7. *ABCC4* rs1751034

Assessment and validation of TaqMan SNP genotyping assay C___1901918_30: *ABCC4* rs1751034

SAMPLES ID	GENOTYPE		
	Ion reporter	**golden helix**	**TaqMan**
383*	CC	CC	-
564*	CC	CC	CC
9	TC	TC	TC
314*	TC	TC	TC
249	TC	TC	TC
186	TC	TC	-
631	TT	TT	TT
274	TT	TT	TT
648	TT	TT	-
12	TT	TT	-

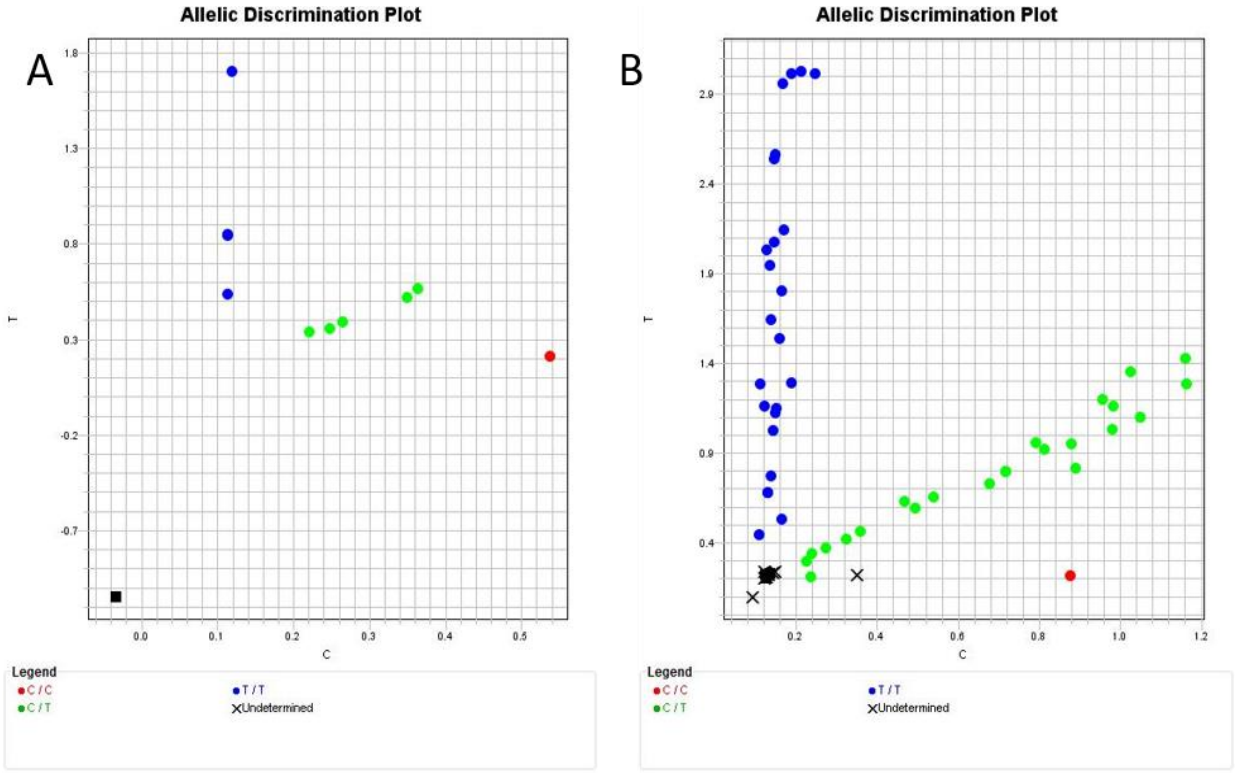


Figure 7. rs1751034 assay validation. Allelic discrimination plot for the d4T cohort (A) and TDF cohort (B) Homozygous individuals with the wild-type allele (TT) are clustered in blue. Those possessing heterozygous form of the allele (AC) cluster in green. The red cluster indicates the patient that are homozygous for the variant allele (CC). The squares represent experimental control (Negative control) and the individuals marked X could not be analysed due to low or no amplification and were therefore repeated. The calls were in accordance with 40% of the genotypic calls for Ion importer and Golden Helix.

8. *APOL1* rs73885319

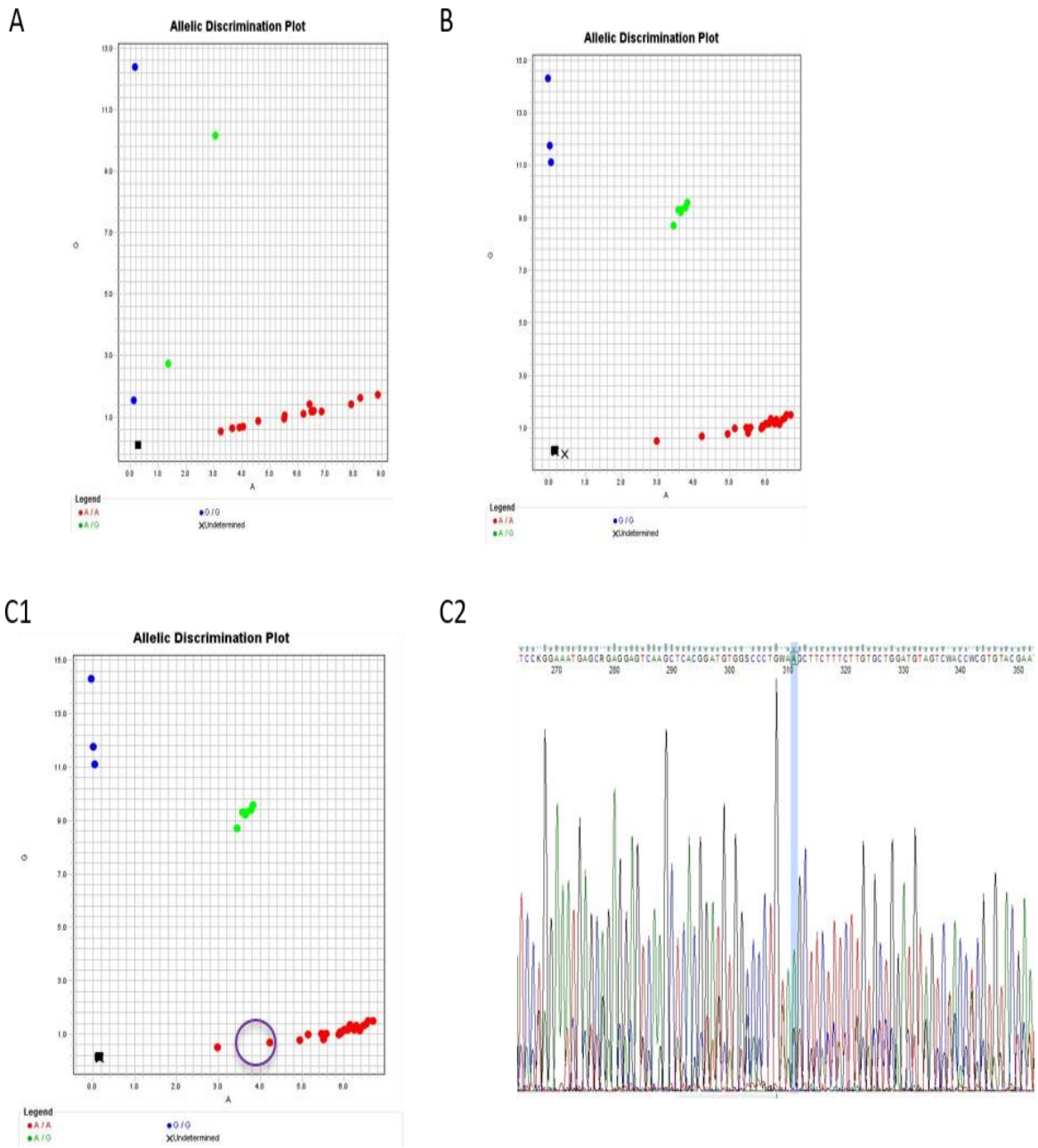


Figure8. rs73885319assay validation. Allelic discrimination plot for the d4T cohort (A) and TDF cohort (B) Individuals homozygous for the wild-type allele (AA) are clustered in red. Those possessing heterozygous form of the allele (AG) are clustered in green. Individuals with both forms of the mutated allele (GG) are clustered in blue. The squares represent experimental control (Negative control). Allele discrimination plots (C1) and chromatogram (C2) image for patient samples RC46; the highlighted nucleotide shows the position where the variant rs3742106 was identified by Sanger sequencing, confirming the TaqMan genotypic calls for RC46.

9. *APOL1* rs60910145

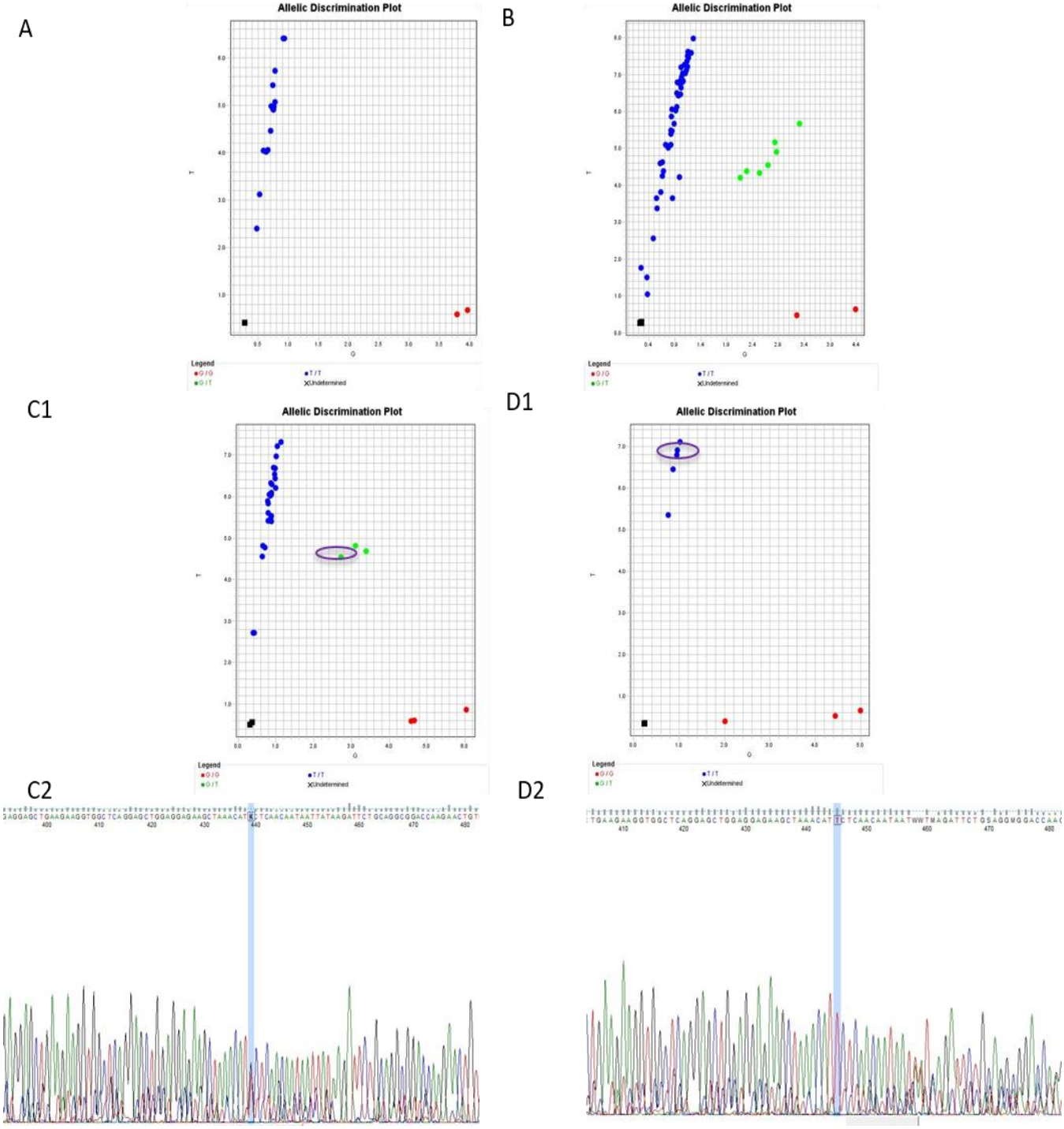


Figure 9. rs60910145 assay validation. Allelic discrimination plot for the d4T cohort (A) and TDF cohort (B) Genotyping calls from the d4T cohort. Individuals both forms of the wild-type allele (AA) are clustered in red. Those possessing homozygous variant allele (GG) are cluster in green. The squares represent experimental control (Negative control) Allele discrimination plots (C1 & D1) and chromatogram (C2 & D2) images for patient samples RC46 (1) and R26 (2); the highlighted nucleotide shows the position where the variant rs60910145 was identified by Sanger sequencing confirmed the TaqMan genotypic calls for RC46 and R26.

APPENDIX J: TDF GENOTYPING DATABSE

Table1: Database containing information and genotypic calls for patients included in the TDF cohort

Sample ID	Age	Sex	Ethnicity	AKI Status	*ABCB1*	*ABCC10*	*ABCC2*			*ABCC4*		*APOL1*	
					rs2032582	rs2125739	rs2273697	rs17222723	rs3740066	rs1751034	rs3742106	rs73885319	rs60910145
R012	55	Male	Black	Case	C_C	T_C	G_G	T_T	C_T	T_T	C_C	A_A	T_T
R019	50	Female	Black	Case	C_C	T_T	A_A	T_T	C_C	T_T	A_C	A_G	T_G
R040	37	Female	Indian	Case	C_C	T_C	G_G	T_A	C_C	T_C	A_A	A_A	T_T
R052	46	Female	Black	Case	C_C	T_T	G_G	T_A	C_C	T_C	A_A	A_A	T_T
R068	32	Female	Black	Case	C_C	T_T	G_A	T_T	C_C	T_T	C_C	A_A	T_T
R086	36	Female	Black	Case	C_C	T_C	G_G	T_T	C_C	T_T	A_C	A_A	T_T
R089	47	Female	Black	Case	C_C	T_C	G_G	T_A	C_C	T_C	A_C	A_A	T_T
R090	37	Male	Black	Case	C_C	T_C	G_G	T_T	C_C	T_T	A_A	A_G	T_G
R094	62	Female	Black	Case	C_C	T_T	G_G	T_A	C_C	T_T	A_C	A_A	T_T
R096	18	Female	Black	Case	C_C	T_C	G_G	T_A	C_C	T_C	A_C	A_G	T_G
R098	66	Female	Black	Case	C_C	T_T	G_G	T_A	C_C	T_C	A_A	A_G	T_T
R100	52	Female	Black	Case	C_C	T_T	G_G	A_A	C_C	C_C	C_C	A_A	T_T
R105	63	Male	Black	Case	C_A	T_T	G_G	T_A	C_C	T_C	A_A	A_A	T_T
R106	38	Male	Black	Case	C_C	T_T	G_G	T_A	C_C	T_T	A_C	A_A	T_T
R107	42	Male	Black	Case		T_T							
R110	33	Female	Black	Case	C_C	T_T	G_A	T_T	C_C	C_C	A_A	A_A	T_T
R114	65	Female	Black	Case	C_C	T_T	A_A	T_A	C_C	T_C	A_C	G_G	G_G
R117	44	Female	Black	Case	C_C	T_C	A_A	T_A	C_C	C_C	C_C	A_A	T_T
R127	50	Female	Black	Case	C_C	T_T	G_G	T_A	C_C	T_T	A_C	A_A	T_T
R130	43	Female	Black	Case	C_C	T_T	G_A	T_T	C_C	T_C	A_A	A_A	T_T
R132	41	Male	Black	Case	C_C	T_C	G_G	A_A	C_C	T_T	A_C	A_A	T_T
R133	33	Female	Black	Case	C_C	T_C	G_G	T_T	C_C	T_T	C_C	A_A	T_T
R134	35	Female	Black	Case	C_C	T_T	G_A	T_T	C_C	T_T	A_C	A_A	T_T
R140	27	Female	Black	Case	C_C	T_C	G_G	T_T	C_T	T_T	A_C	A_A	T_T
R141	42	Male	Black	Case	C_C	T_T	G_G	T_T	C_C	T_T	A_A	A_A	T_T
R144	46	Female	Black	Case	C_C	C_C	G_A	T_A	C_C	T_T	C_C	A_A	T_T
R146	62	Male	Black	Case	C_C	T_C	G_A	T_T	C_C	T_T	A_A	A_A	T_T
R147	48	Female	Black	Case	A_A	T_C	G_G	T_A	C_C	T_T	A_C	A_A	T_T
R148	71	Male	Black	Case	C_C	T_T	G_G	T_T	C_C	C_C	C_C	A_A	T_T
R151	52	Male	Black	Case	C_C	T_C	G_G	T_T	C_C	T_C	A_C	A_A	T_T
R152	45	Male	Black	Case	C_C	T_T	G_G	T_A	C_C	T_C	A_A	A_G	T_T
R161	44	Female	Coloured	Case	C_C	T_C	A_A	A_A	T_T	T_C	A_C	A_A	T_T
R165	40	Male	Black	Case	C_C	T_C	G_A	T_A	C_C	T_C	C_C	G_G	G_G
R166	65	Male	Black	Case	C_C	T_T	G_G	T_A	C_C	T_T	A_C	A_G	T_G
R172	54	Male	Black	Case	C_A	T_T	G_A	T_A	C_C	T_T	A_A	A_G	T_T
R176	56	Female	Black	Case	C_C	T_T	G_G	T_T	C_C	T_T	A_C	A_A	T_T
R177	64	Female	Black	Case	C_C	T_T	G_G	T_T	C_C	C_C	A_A	G_G	T_T
R182	51	Male	Black	Case	C_C	T_C	A_A	A_A	C_T	T_C	C_C	A_A	T_T
R183	60	Female	Black	Case	C_C	T_T	G_G	T_T	C_C	T_C	A_C	A_G	T_G

Sample ID	Age	Sex	Ethnicity	AKI Status	*ABCB1*	*ABCC10*	*ABCC2*			*ABCC4*		*APOL1*	
					rs2032582	**rs2125739**	**rs2273697**	**rs17222723**	**rs3740066**	**rs1751034**	**rs3742106**	**rs73885319**	**rs60910145**
R184	49	Male	Black	Case	C_C	T_C	G_A	T_T	C_C	T_C	A_C	A_A	T_T
R185	23	Female	Black	Case	C_C	T_C	G_G	T_T	C_C	T_C	A_C	A_G	T_T
R187	28	Female	Black	Case	C_C	T_T	G_G	T_T	C_T	T_T	A_A	A_A	T_T
R188	35	Female	Black	Case	C_C	T_T	G_G	T_T	C_C	T_T	A_C	A_A	T_T
R189	38	Male	Black	Case	C_C	T_C	G_A	T_A	C_C	T_T	A_C	A_G	T_G
R190	39	Male	Black	Case	C_C	T_T	G_G	T_T	C_C	T_C	A_A	A_G	T_G
R192	33	Female	Black	Case	C_A	T_T	G_A	T_T	C_C	T_C	A_C	A_A	T_T
R193	37	Female	Black	Case	C_C	T_C	G_G	T_T	C_T	T_C	A_A	A_A	T_T
R194	54	Female	Black	Case	C_C	T_C	A_A	A_A	C_C	C_C	A_C	A_A	T_T
R195	41	Female	Black	Case	C_C	T_C	G_G	T_T	C_C	T_C	A_A	A_A	T_T
R091	69	Male	Black	Case	C_C	T_T	G_G	T_A	C_C	T_T	A_A	A_A	T_T
R116	54	Female	Black	Case	C_C	T_C	G_A	A_A	C_C	T_C	A_C	A_A	T_T
R142	53	Female	Black	Case	C_C	T_T	G_G	T_T	C_C	T_C	A_C	A_A	T_T
R169	55	Male	Black	Case	C_C	T_T	G_G	T_T	C_T	T_T	A_C	A_A	T_T
RC001	39	Male	Black	Control	C_C	T_T	G_G	T_A	C_C	C_C	A_C	G_G	T_G
RC002	53	Male	Coloured	Control		T_T	G_G	T_A	C_C	T_T	C_C	A_A	T_T
RC003	35	Female	Black	Control	C_C	T_T	G_G	T_T	T_T	C_C	A_C	A_A	T_T
RC004	37	Female	Black	Control	C_C	T_C	G_G	T_T	C_C	T_T	A_A	A_A	T_T
RC005	50	Male	Black	Control	C_C	T_T	G_G	T_A	C_C	T_C	A_C	A_A	T_T
RC006	40	Female	Black	Control	C_C	T_T	G_G	T_T	C_C	T_T	A_C	A_A	T_T
RC009	33	Female	Black	Control	C_C	T_C	G_G	T_A	C_T	T_C	A_C	A_A	T_T
RC010	42	Female	Black	Control	C_A	C_C	G_A	T_T	C_C	T_C	A_A	A_G	G_G
RC011	37	Male	Black	Control	C_C	T_T	G_A	T_T	C_C	T_C	A_A	A_A	T_T
RC013	46		Black	Control	C_C	C_C	A_A	A_A	C_C	T_C	A_C	A_A	T_T
RC041	47	Male	Black	Control	C_C	T_T	G_G	T_T	C_T	T_T	C_C	A_A	T_T
RC042	43	Female	Black	Control	C_C	T_T	G_G	T_A	C_C	T_T	A_C	A_A	T_T
RC043	28	Female	Black	Control	C_C	T_T	G_G	T_T	C_C	T_T	A_A	A_G	T_T
RC044	34	Male	Black	Control	C_C	T_T	G_G	T_T	C_T	T_T	A_C	A_A	T_T
RC045	40	Female	Black	Control	C_C	T_T	G_G	T_T	C_C	T_C	A_A	A_G	T_T
RC046	32	Female	Black	Control	C_C	T_T	A_A	T_T	C_C	T_C	C_C	A_G	T_G
RC047	30	Female	Black	Control	A_A	C_C	G_G	T_A	C_C	T_C	A_A	A_A	T_T
RC048	56	Male	Black	Control	C_C	T_T	G_G	T_A	C_C	T_C	A_C	A_A	T_T
RC049	38	Male	Black	Control	C_C	C_C	G_G	T_T	C_C	T_C	A_A	A_G	T_G
RC050	40	Female	Black	Control	C_C	T_T	G_G	T_T	C_C	T_T	A_A	A_A	T_T
RC051	50	Female	Black	Control	C_C	T_T	G_G	T_T	C_T	T_T	A_A	A_A	T_T
RC052	28	Female	Black	Control	C_C	T_T	G_G	T_A	C_C	T_T	A_C	A_A	T_T
RC053	34	Female	Black	Control	C_C	T_C	G_G	T_T	C_T	T_C	C_C	A_G	T_G
RC054	55	Male	Black	Control	C_C	T_T	G_A	T_T	C_C	C_C	C_C	A_A	T_T
RC055	59	Male	Black	Control	C_C	T_T	G_A	T_T	C_C	T_C	A_C	A_A	T_T
RC056	44	Female	Black	Control	C_C	T_C	G_G	T_A	C_C	T_T	A_A	A_A	T_T
RC057	43	Male	Black	Control	C_C	T_C	G_G	A_A	C_C	T_C	A_C	A_A	T_T
RC058	28	Female	Black	Control		T_C	G_A	T_T	C_C			A_A	T_T
RC059	36	Female	Black	Control	C_C	T_C	G_G	T_T	C_T	T_T	A_A	A_A	T_T

Sample ID	Age	Sex	Ethnicity	AKI Status	*ABCB1*	*ABCC10*	*ABCC2*			*ABCC4*		*APOL1*	
					rs2032582	rs2125739	rs2273697	rs17222723	rs3740066	rs1751034	rs3742106	rs73885319	rs60910145
RC060	35	Female	Black	Control	C_C	T_T	G_G	T_T	C_C	T_T	C_C	A_A	T_T
RC061	43	Female	Black	Control	C_C	T_T	G_G	T_T	C_C	T_C	A_A	A_A	T_T
RC062	35	Male	Black	Control	C_C	T_C	G_G	T_A	C_C	T_C	A_C	A_A	T_T
RC063	36	Female	Black	Control	C_C	T_T	G_G	T_T	C_C	T_T	A_A	A_A	T_T
RC064	34	Male	Black	Control	C_C	T_T	G_G	T_T	C_C	T_C	A_A	A_A	T_T
RC065	37	Female	Black	Control	C_C	T_T	G_G	T_A	C_C	T_T	A_A	A_A	T_T
RC066	39	Female	Black	Control	C_C	T_C	G_G	A_A	C_C	T_T	A_A	A_A	T_T
RC067	38	Female	Black	Control	C_C	T_C	G_A	T_T	C_T	T_C	A_A	A_A	T_T
RC068	42	Female	Black	Control	C_C	T_C	G_G	T_A	C_C	T_T	A_A	A_A	T_T
RC069	26	Female	Black	Control	C_C	T_T	G_G	T_T	T_T	T_T	C_C	A_A	T_T
RC070	30	Female	Black	Control	C_C	T_C	G_A	T_T	C_T	T_C	A_C	A_A	T_T
RC071	40	Male	Black	Control	C_A	T_C	G_A	T_T	C_C	T_C	A_C	A_A	T_T
RC072	49	Male		Control	C_C	T_T	G_G	T_T	C_C	T_C	A_C	A_G	T_G
RC073	42	Female	Black	Control	C_C	T_C	G_G	T_T	C_T	T_C	C_C	A_A	T_T
RC074	45	Female	Black	Control	C_C	C_C	G_A	T_T	C_C	T_C	C_C	A_A	T_T
RC075	19	Female	Black	Control	C_C	T_T			C_T				T_G
RC076	35	Female	Black	Control	C_C	T_C	G_G	T_T	C_C	T_C	A_C	A_A	T_T
RC077	31	Female	Black	Control	C_C	T_C	G_G	T_T	C_T	T_C	A_C	A_A	T_T
RC078	38	Male	Black	Control	C_C	C_C	G_G	T_A	C_C	C_C	A_C	A_A	T_T
RC079	39	Male	Black	Control		T_T	G_G	A_A	C_C	T_T	A_A	A_A	T_T
RC080	38	Female	Black	Control	C_C	T_T	G_G	T_T	C_T	C_C	A_A	A_A	T_T
RC081	27	Female	Black	Control		T_T	G_G	T_A	C_C	T_T	C_C	A_A	T_T
RC082	47	Male	Black	Control	C_C	T_C	G_G	T_T	C_C	T_T	A_C	A_A	T_T
RC083	54	Female	Black	Control	C_C	T_C	G_G	T_T	C_C	T_T	A_A	A_A	T_T
RC084	56	Female	Black	Control	C_C	T_C	G_G	T_A	C_C	C_C	A_C	G_G	G_G
RC085	30	Female	Black	Control	C_C	T_C	G_G	T_T	C_C	T_C	A_C	G_G	T_G
RC086	39	Female	Black	Control	C_C	T_C	G_G	T_T	C_C	T_C	C_C	A_G	T_G
RC088	39	Female	Black	Control	C_C	T_C	G_G	T_A	C_C	T_T	C_C	A_A	T_T
RC089	53	Female	Black	Control	C_C	T_T	G_G	T_A	C_C	T_C	A_C	A_A	T_T
RC090	58	Female	Black	Control	C_C	T_C	G_A	T_T	C_C	T_T	A_C	A_A	T_T
RC091	48	Male	Black	Control	C_C	T_C	G_G	T_T	C_C	T_C	A_A	A_A	T_T
RC092	57	Female	Black	Control		T_T						A_A	T_T
RC093	36	Female	Black	Control		T_T						A_G	T_G
RC094	41	Female	Black	Control		T_C	G_G	T_T	C_T	T_C	C_C	A_A	T_T
RC095	30	Female	Black	Control		T_T	G_G	T_T	C_C	T_C	A_A	A_G	T_G
RC096	41	Female	Black	Control	C_C	T_C	G_A	T_T	C_C	T_T	A_A	A_A	T_T
RC097	52	Female	Black	Control	C_C	T_T	G_G	T_T	C_T	T_T	C_C	A_A	T_T
RC098	49	Female	Black	Control	C_C	T_C	G_G	T_T	C_T	T_T	A_A	A_A	T_T
RC099	36	Female	Black	Control		T_C	G_A	T_T	C_C	T_T	A_C	A_A	T_T
RC100	51	Female	Black	Control	C_C	T_C	G_G	T_T	C_C	T_C	A_C	A_A	T_T
RC101	21	Female	Black	Control	C_C	C_C	G_G	T_T	C_C	T_T	A_C	A_A	T_T
RC102	57	Female	Black	Control	C_C	T_T	G_G	T_A	C_C	T_C	C_C	A_A	T_T
RC103	26	Female	Black	Control	C_C	T_T	G_A	T_T	C_C	T_T	A_A	A_A	T_T

Sample ID	Age	Sex	Ethnicity	AKI Status	*ABCB1*	*ABCC10*	*ABCC2*			*ABCC4*		*APOL1*	
					rs2032582	**rs2125739**	**rs2273697**	**rs17222723**	**rs3740066**	**rs1751034**	**rs3742106**	**rs73885319**	**rs60910145**
RC104	25	Female	Black	Control	C_C	C_C	G_G	T_T	C_C	T_C	A_C	A_A	T_T
RC105	49	Female	Black	Control	C_C	T_T	G_A	T_A	C_C	T_T	A_A	A_A	T_T
RC106	53	Female	Black	Control		T_T	G_G		C_C		A_A	A_A	
RC107	53	Female	Black	Control	C_C	T_T	G_A	T_T	C_C	C_C	A_C	A_A	T_T
RC108	45	Female	Black	Control	C_C	T_C			C_C	C_C	A_C	A_A	T_T
RC109	35	Female	Black	Control	C_C	T_T	G_G	T_T	C_T	T_T	A_A	A_A	T_T
RC110	50	Female	Black	Control	C_C	T_C	G_G	T_A	C_C	T_T	A_C	A_A	T_T
RC111	37	Female	Black	Control	C_C	T_C	G_G	T_A	C_T	T_T	A_A	A_A	T_T
RC112	46	Female	Black	Control	C_C	T_T	G_G	T_A	C_C	T_C	A_A	A_A	T_T
RC113	41	Female	Black	Control	C_C	T_T	G_G	T_T	C_T	T_T	A_A	A_A	T_T
RC114	37	Female	Black	Control	C_C	T_T	G_G	T_A	C_C	T_T	A_A	A_A	T_T
RC115	41	Female	Black	Control	C_C	T_C	G_A	T_A	C_C	T_C	A_A	A_A	T_T
					C_C	T_C	G_G	T_T	C_T	T_C	C_C	A_G	G_G

APPENDIX K: ASSESSMENT OF THE PERFORMANCE OF THE PRIMERS USED FOR PCR PRIOR TO SANGER SEQUENCING TO VALIDATE THE GENOTYPING CALLS

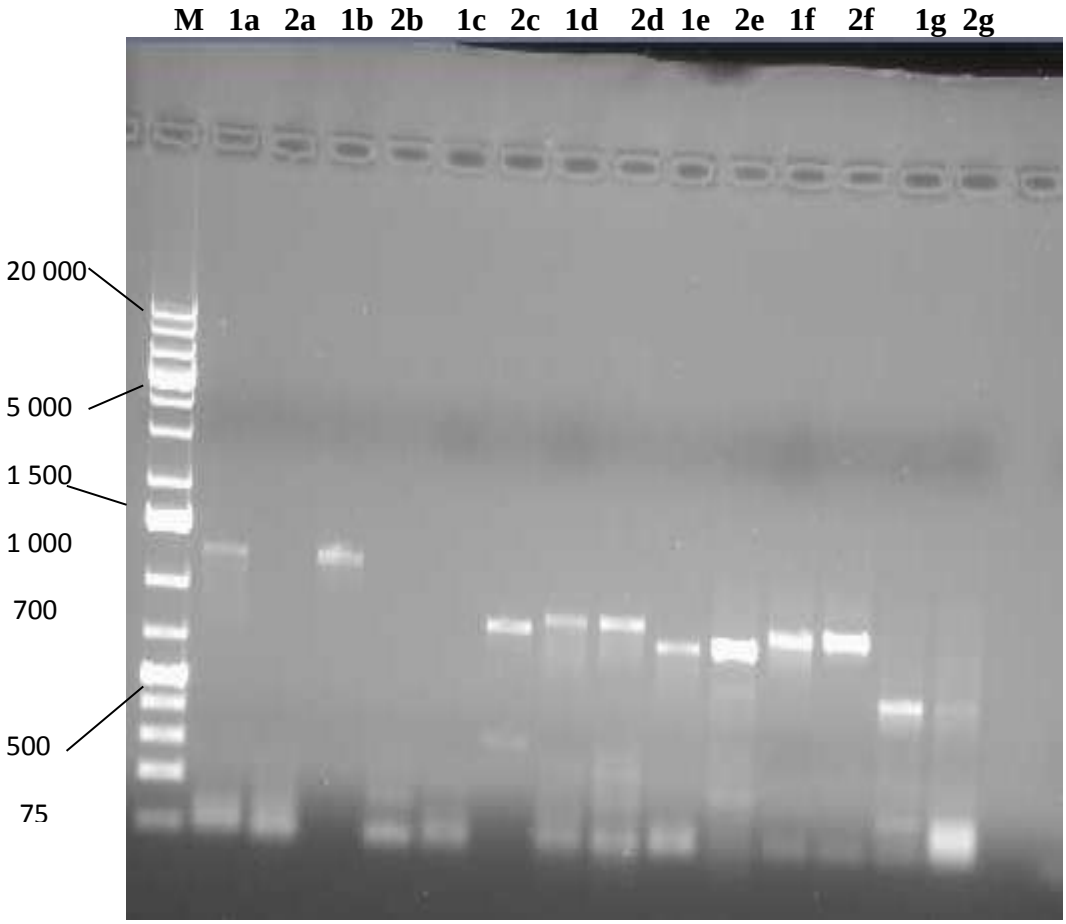


Figure 1. PCR amplicons of DNA fragments that encompass the SNPs of interest (a-g). Initially, PCR reactions were set up using Taq polymerase (**1**) and iProof™ polymerase (**2**). The corresponding fragment sizes are indicated in brackets with SNP b not being amplified. Therefore optimised using Iproof polymerase for **a, c, d, e, f** and Taq polymerase for **g.** About 10µL loaded on a 1% gel, with 10µL of 1Kb plus marker (ThermoFisher Scientific, Massachusetts, USA).

Key: a=rs2032582 (1295 bp) b= rs2125739 (1476 bp), c= rs17222273 (857 bp), d= rs227369 (992 bp), e= rs374066 (792bp), f= rs1751034 (880bp) and g= rs37402106 (557 bp).

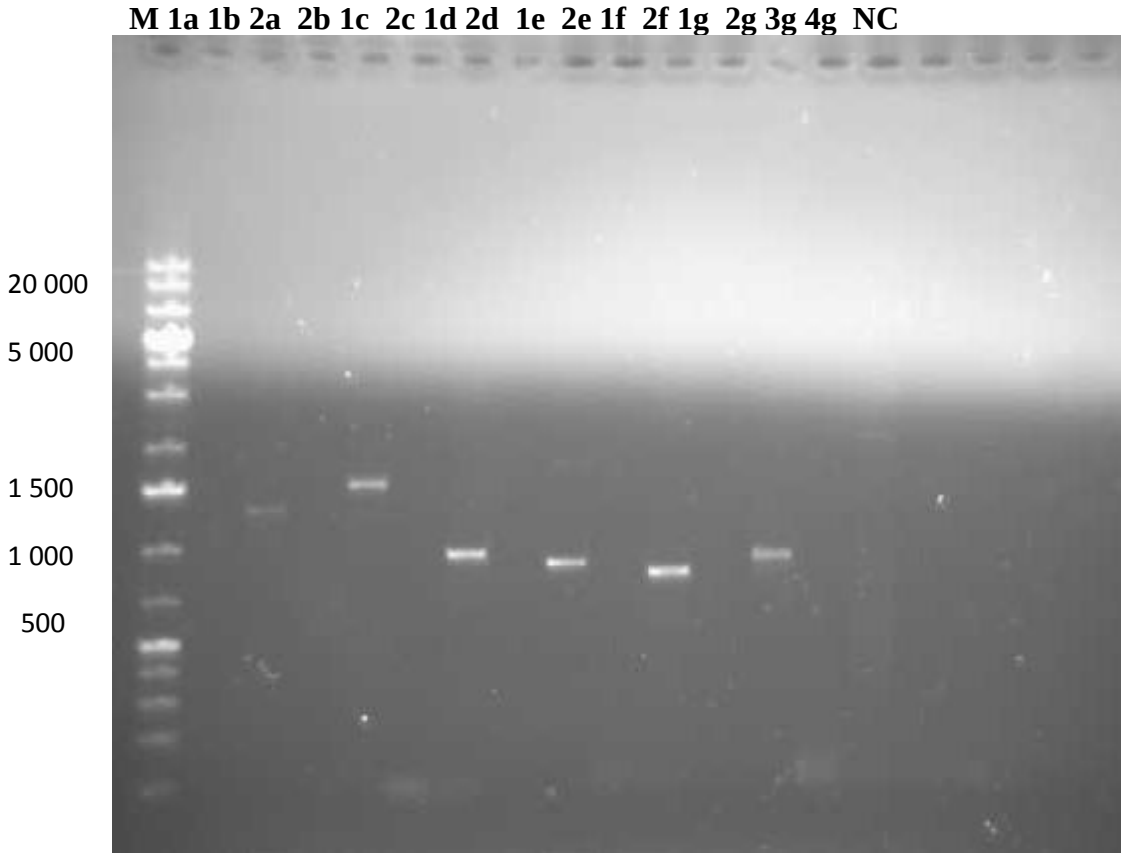


Figure 2: PCR amplicons of DNA fragments that encompass the SNPs of interest (a-g). Initially PCR reactions were set up using iProof^{TM} polymerase (**1**), iProof^{TM} polymerase with DMSO (**2**), Taq polymerase (**3**) and Taq polymerase with DMSO (**4**). The corresponding fragment sizes are indicated in brackets with the region for SNP g not amplified. The amount gDNA was doubled to 2ul (44ng). About 10ul loaded on a 1% gel, with 10ul of 1Kb plus marker (ThermoFisher Scientific, Massachusetts, USA).

Key: a=rs2032582 (1295 bp) b= rs2125739 (1476 bp), c= rs17222273 (857 bp), d= rs227369 (992 bp), e= rs374066 (792 bp), f= rs1751034 (880 bp) and g= rs37402106 (557 bp).

APPENDIX L: ASSESSMENT AND VALIDATION OF TAQMAN SNP GENOTYPING ASSAY

Table 1: Validation of TaqMan SNP genotyping assays.

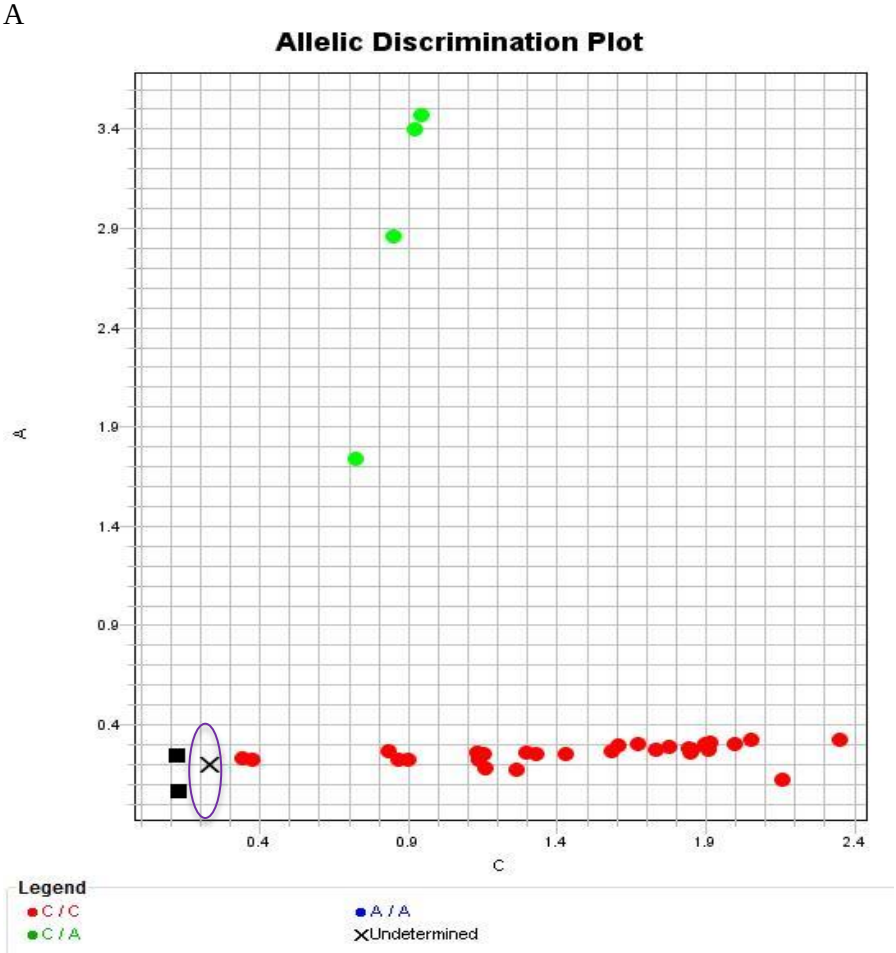
Gene	SNP ID	Assay	sample ID	TaqMan	Sanger sequencing
ABCB1	rs2032582	C_11711720_30	RC46	C/C	C/C
			R 73	A/A (X)	A/A
ABCC10	rs212739	C__16173668_10	RC77	X	T/T
ABCC2	rs17222723	C__25591743_30	R 26	T/T	T/T
	rs227369	C__22272980_20	RC46	A/A	A/A
			R 26	G/G	G/G
	rs3740066	C__11214910_20	RC91	C/C	C /C
			RC66	C/C	C /C
ABCC4	rs3742106	C___8059522_1_	RC46	C/C	C/C
APOL1	rs60910145	C__98253221_10	RC46	A/A	A/A
			R 26	A/A	A/A
	rs73885319	C__89555688_10	RC46	T/G	T/G (K)
			R 26	T/T	T/T

X= undetermined, RC= renal control sample, R= renal sample

APPENDIX M: CONFIRMATION OF AMBIGUOUS GENOTYPIC CALLS

Sanger sequencing was performed for patients with ambiguous or undetermined calls for further investigation and confirmation. The region flanking the SNP was amplified using the primers described chapter 2, the amplicons were purified and sent for sequencing to Inqaba Biotech, Pretoria. The results are summarised in the figures below where Sanger sequencing was used to determine the genotype of one of the TDF patient sample (Figure 1) and confirmed the ambiguous call of the other (Figure 2).

A



B

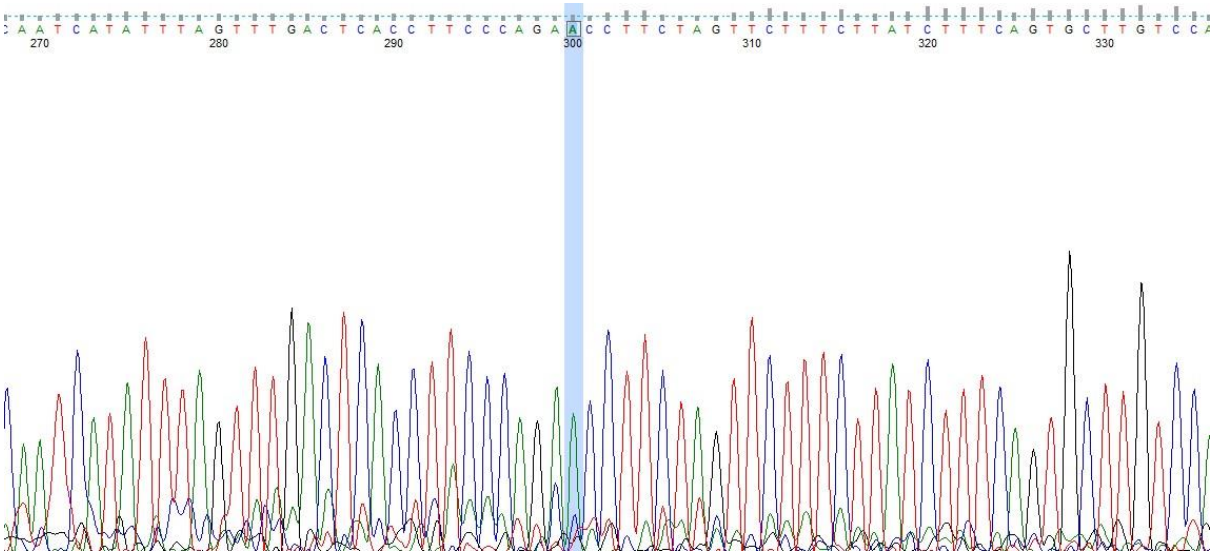
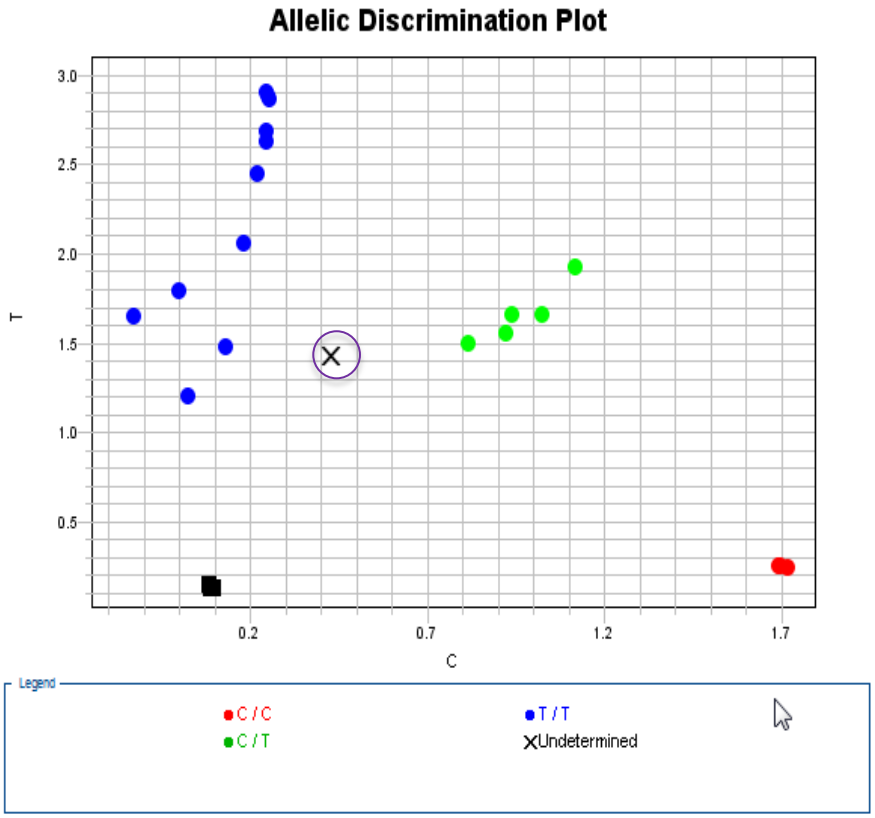


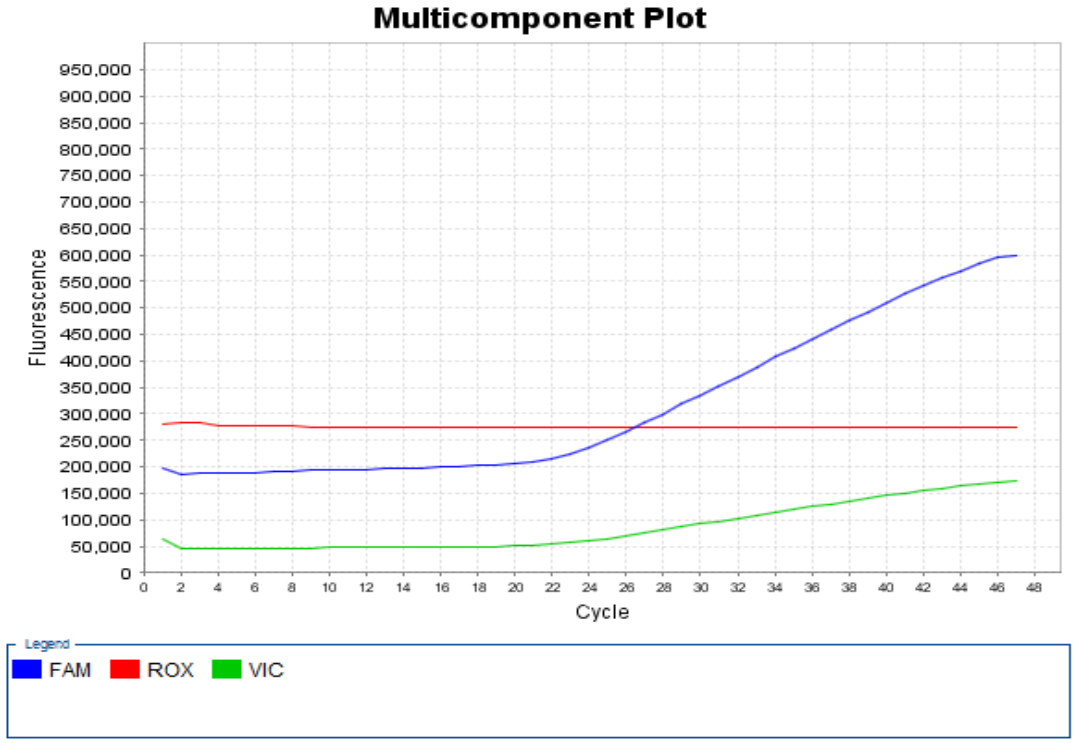
Figure 1. TaqMan SNP genotyping assay C_11711720_30.

Allele discrimination plots (A) and chromatogram (B) images for sample R73; the highlighted nucleotide shows the position where the variant rs2032582 was identified by Ion S5^{TM} sequencing, located in *ABCB1*, c.2677C>A. Sanger sequencing confirmed the TaqMan genotypic calls. The TaqMan assay signals (FAM/VIC) for R73 were low but the assays were sensitive enough to call accurately.

A



B



C

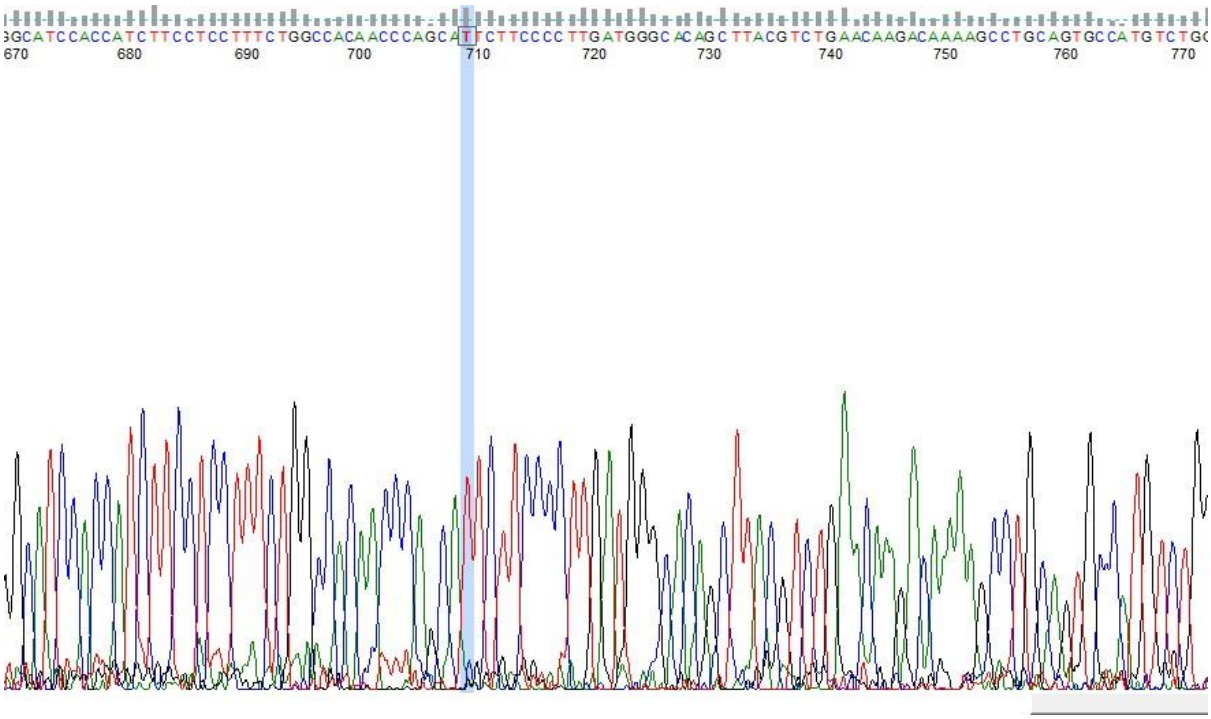
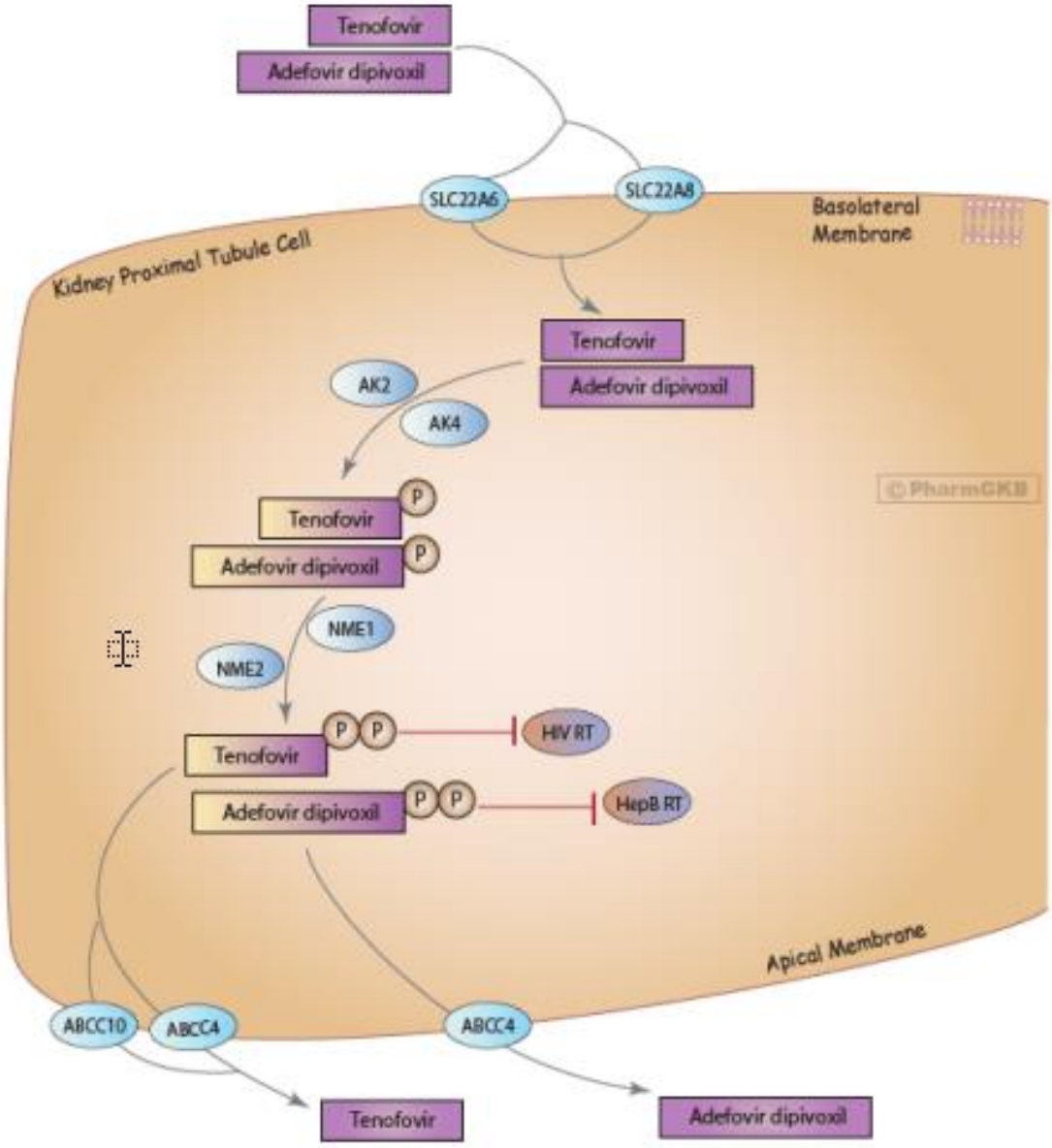


Figure 2. TaqMan SNP genotyping assay C__16173668_10.
Allele discrimination plots (A) and chromatogram (C) images for patient sample RC77 (1) ; the highlighted nucleotide shows the position where the variant rs212739 was identified by Ion S5^{TM} sequencing, located in *ABCC10*, c.2759 T>C. The allelic discrimination plots show representative genotypic calls and the multicomponent plot illustrates the fluorescent signalling for each probe in the assay. The TaqMan assay signals (FAM/VIC) for R77 were undeterminded (X) low but the assays were sensitive enough to call accurately, as indicated by the multicomponent plot (B) and confirmed by Sanger sequencing (C).

APPENDIX N: TENOFOVIR/ ADEFOVIR PATHWAY, PHARMACOKINETICS



Tenofovir pathway. Tenofovir is administered as a prodrug tenofovir _ adefovir) dipivoxil and is absorbed into the gastrointestinal tract via p-gp (encoded by *ABCB1*). P-gp transport the pro-drug Tenofovir/adefovir dipivoxil, which require initial diester hydrolysis for conversion to tenofovir/adefovir. Phosphorylation occurs twice, by adenylate kinases (AK2 and AK4) and (NME1 and NEM1) nucleotide diphosphate kinases, for the pro-drug to be active and have anti-viral properties. Tenofovir /adenovir are not metabolized by cytochrome P450 enzymes. They are hardly metabolised as they are eliminated in their active form. Elimination occurs through the kidneys via tubular secretion and glomerular filtration. Elimination occurs in the proximal tubule of the renal nephron, the functional units of the kidneys. Clearance of these drugs is regulated mostly by membrane transport proteins. Organic anion transporter 1 (OAT1) (SLC22A6) and organic anion transporter 3 (OAT3) (SLC22A8) are located on the basolateral membrane of the human proximal tubule (kidney) cell, and multidrug resistance protein 4 (MRP4) (*ABCC4*) and MRP 7 (*ABCC10*) on the apical side (Miller 2001). SLC22A6 and SLC22A8 act in the uptake of tenofovir/adefovir into the renal proximal tubule cell while *ABCC4* controls the active efflux of the drug across the apical membrane (Uwai *et al.* 2007). MRP2 protein (*ABCC2*) has also been implicated in the efflux process. Genetic variation in SLC22A6, SLC22A8 and *ABCC4* may alter regulation and exposure of the kidney to tenofovir/adefovir and thus, play a major role in tenofovir/adefovir associated renal toxicity. Other drugs that interact with these proteins may influence how tenofovir/adefovir is eliminated and may possibly play a role in tenofovir/adefovir associated toxicity (Whirl (Whirl-Carrillo *et al.* 2017)Carrillo et al, 2012).