ASSOCIATION OF VARIANTS IN APOL1, MYH9 AND HMOX1 WITH MICRO-ALBUMINURIA AMONG SICKLE CELL DISEASE PATIENTS FROM CAMEROON

AMY GEARD
GRDAMY002

SUPERVISORS:
PROFESSOR AMBROISE WONKAM
DR EMILE CHIMUSA

DIVISION OF HUMAN GENETICS
PRESENTED FOR MSc (Med) IN HUMAN GENETICS
FACULTY OF HEALTH SCIENCES
UNIVERSITY OF CAPE TOWN
27/09/2016

Word Count: 40989
The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.
DECLARATION

I, ........................................, hereby declare that the work on which this dissertation/thesis is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university.

I empower the university to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever.

Signature: [Signed] ..................................

Date: ........................................
# TABLE OF CONTENTS

**LIST OF PUBLICATIONS FROM THIS WORK** ........................................................................................................... 6

**ABSTRACT** ............................................................................................................................................................... 7

**LIST OF FIGURES** ......................................................................................................................................................... 10

**LIST OF TABLES** .......................................................................................................................................................... 13

**LIST OF ABBREVIATIONS** ........................................................................................................................................... 14

**CHAPTER 1: INTRODUCTION AND LITERATURE REVIEW** .............................................................................................. 17

1.1. **HEMOGLOBINOPATHIES** ......................................................................................................................................... 17

1.2. **EPIDEMIOLOGY OF SICKLE CELL DISEASE (SCD)** ............................................................................................. 18

1.3. **PATHOPHYSIOLOGY AND CLINICAL MANIFESTATIONS OF SICKLE CELL DISEASE** ........................................... 18

1.4. **TREATMENT OF SCD** ............................................................................................................................................... 21

1.5. **GENETICS AND PHENOTYPIC VARIABILITY IN SICKLE CELL DISEASE** ............................................................... 22

1.6. **FOCUS ON GENETICS AND CARDIOVASCULAR PHENOTYPES OF SCD** ............................................................... 23

1.7. **KIDNEY DISEASE IN SCD** ........................................................................................................................................ 24

**Epidemiology and risk factors** ........................................................................................................................................ 24

**Genetics of kidney disease in SCD** ................................................................................................................................ 26

- **MYH9** ........................................................................................................................................................................ 26

- **APOL1** ........................................................................................................................................................................ 27

- **HMOX1** ....................................................................................................................................................................... 29

1.8. **SUMMARY OF LITERATURE REVIEW** .................................................................................................................... 31

1.9. **RATIONALE OF THE PRESENT STUDY** ................................................................................................................... 32

1.10. **AIM AND OBJECTIVES** ........................................................................................................................................... 32

**CHAPTER 2: METHODS AND MATERIALS** ................................................................................................................... 34

2.1. **PATIENT DATA** ........................................................................................................................................................ 34

2.1.1. **ETHICAL APPROVAL** .......................................................................................................................................... 34

2.1.2. **STUDY PARTICIPANTS** ........................................................................................................................................ 34

2.1.3. **RENAL FUNCTIONS DEFINITION AND PHENOTYPING** ..................................................................................... 35

2.1.3.1. **ALBUMINURIA AND THE ALBUMIN-TO-CREATININE RATIO (ACR) IN THE SCD COHORT** ........................................ 35

2.1.3.2. **THE ESTIMATED GLOMERULAR FILTRATION RATE** ............................................................................................ 35

2.1.3.3. **END STAGE KIDNEY DISEASE** .......................................................................................................................... 37

2.2. **MOLECULAR METHODS** ............................................................................................................................................ 37

2.2.1. **DNA QUANTIFICATION AND INTEGRITY ANALYSIS** .......................................................................................... 37

2.2.3. **POLYMERASE CHAIN REACTION (PCR)** ............................................................................................................. 43

2.2.4. **PCR OPTIMIZATION** ............................................................................................................................................. 44

2.2.4.1. **TEMPERATURE GRADIENT** ............................................................................................................................... 44

2.2.4.2. **MgCl₂ GRADIENT** ................................................................................................................................................ 52

2.2.4.3. **DIMETHYL SULFOXIDE PERCENTAGE GRADIENT** ............................................................................................. 53

2.2.4.4. **BETAINE CONCENTRATION GRADIENT** ............................................................................................................ 56

2.2.4.5. **SUMMARY OF THE OPTIMISATION PROCESS** .................................................................................................. 57
2.2.5. PCR PRODUCT DETECTION ................................................................. 59
2.2.6. PCR PRODUCT CLEAN-UP ......................................................... 59
2.2.7. SNaPSHOT® SEQUENCING ...................................................... 60
  2.2.7.1. SNaPSHOT PRODUCT CLEAN-UP .................................. 63
  2.2.7.2. CAPILLARY ELECTROPHORESIS ........................................ 64
2.2.8. FRAGMENT ANALYSIS ............................................................. 66
  2.2.8.1. PCR PRODUCT CLEAN UP .............................................. 66
  2.2.8.2. CAPILLARY ELECTROPHORESIS ........................................ 66
  2.2.9. TAQMAN SNP GENOTYPING .................................................. 68
  2.2.10. DIRECT CYCLE SEQUENCING VALIDATION .......................... 71
SUMMARY OF THE MOLECULAR METHODS .................................... 76

2.3. STATISTICAL ANALYSIS ............................................................... 77
  2.3.1. DESCRIPTIVE STATISTICAL PACKAGES ............................... 77
  2.3.2. DATA NORMALISATION ............................................................ 77
  2.3.3. ASSOCIATION STUDIES .......................................................... 78

CHAPTER 3: RESULTS ........................................................................... 80
SECTION 1: SCD PATIENTS AND CONTROLS ........................................... 80
  3.1. POPULATION DESCRIPTION AND CLINICAL PHENOTYPES .......... 80
    3.1.1. DESCRIPTION OF SOCIODEMOGRAPHIC (AGE AND SEX) VARIABLES OF SCD PATIENTS AND CONTROLS ................................................................. 80
    3.1.2. DESCRIPTION OF ANTHROPOMETRIC VARIABLES OF THE SCD COHORT .......... 80
    3.1.3. HEMATOLOGICAL VARIABLES OF THE SCD COHORT .................. 84
    3.1.4. DESCRIPTION OF KIDNEY FUNCTIONS OF THE SCD COHORT .......... 88
      3.1.4.1. ESTIMATED GLOMERULAR FILTRATION RATE (eGFR) ................. 88
      3.1.4.2. MICRO-ALBUMINURIA ...................................................... 91
      3.1.4.3. SUMMARY OF KIDNEY FUNCTIONS IN THE SCD COHORT ........... 96
    3.1.5. CLINICAL ASSOCIATIONS IN THE SCD COHORT ...................... 96
      Association between eGFR and blood pressure .................................. 96
      Association between eGFR and albuminuria .................................... 99
      Association between eGFR and sex, age and BMI .......................... 103
      Association between crude albuminuria, sex and age among SCD Patients ....... 107
    3.1.6. SUMMARY OF CLINICAL PHENOTYPE ASSOCIATIONS ............. 109
  3.2. DESCRIPTION OF GENETIC VARIANTS ...................................... 110
    3.2.1. COMPARATIVE MINOR ALLELE FREQUENCY (MAF) OF TARGETED SNPS WITH VARIOUS WORLD POPULATIONS USING DATA EXTRACTED FROM THE 1000 GENOME PROJECT .................................. 115
    3.2.2. GENOTYPE-TO-PHENOTYPE ASSOCIATION STUDIES ............. 120
      3.2.2.1. Association between HBB haplotype and crude albuminuria in the SCD patient cohort ................................................................. 120
      3.2.2.2. Association between 3.7 kb α-globin gene deletion and crude albuminuria in the SCD patient cohort ................................................................. 120
3.2.2.3. Association between four targeted SNPs in MYH9 and crude albuminuria, and ACR in the SCD patient cohort ................................................................. 121
3.2.2.4. Association between targeted polymorphisms in HMOX1 and APOL1 with select phenotypes of kidney disease; eGFR, crude albuminuria, macro-albuminuria, and ACR ................................................................. 122

SUMMARY OF THE RESULTS OF SECTION 1 ........................................................................ 129

SECTION 2. ESKD GROUP AND MATCHED CONTROLS ........................................................ 129
Description of the co-morbidities ...................................................................................... 129
Comparative sociodemographic cohort description ......................................................... 130
Molecular Association Studies ......................................................................................... 131

SUMMARY OF THE RESULT OF SECTION 2 .......................................................................... 133

CHAPTER 4: DISCUSSION. ......................................................................................................... 134
4.1. ORIGINALITY AND GENERAL COMMENTS ................................................................. 134
4.2. MICRO-ALBUMINURIA AND GLOMERULAR FILTRATION IN THE SCD POPULATION .. 135
4.3. CLINICAL FACTORS AFFECTING CRUDE ALBUMINURIA AND eGFR ................. 138
4.4. COMPARATIVE MINOR ALLELE FREQUENCIES IN MYH9, APOL1 AND HMOX1 IN VARIOUS WORLD POPULATIONS ................................................................. 140
4.5. VARIANTS IN MYH9, APOL1 AND HMOX1 AND SCD NEPHROPATHY ................. 141
  4.5.1. SNPs IN MYH9 AND APOL1 G1/G2 ......................................................................... 142
  4.5.2. VARIANTS IN HMOX1 .............................................................................................. 144
4.6. CO-INHERITANCE OF THE 3.7 kB α-GLOBIN GENE DELETION AND SCD NEPHROPATHY ......................................................................................................................... 145
4.7. HBB HAPLOTYPE AND SCD NEPHROPATHY ............................................................. 146
4.8. VARIANTS IN APOL1 AND HMOX1 IN THE NON-SCD ESKD POPULATION .......... 146
4.9. PRACTICAL IMPLICATIONS AND RESEARCH RECOMMENDATIONS .................. 148
4.10. LIMITATIONS OF THE STUDY ...................................................................................... 149

CHAPTER 5: CONCLUSION AND FUTURE PERSPECTIVES. .................................................. 150
REFERENCES .......................................................................................................................... 151

APPENDIX 1: REAGENTS AND BUFFERS.............................................................................. 166
APPENDIX 2: EQUIPMENT AND MACHINES ..................................................................... 172
APPENDIX 3: PCR CONSTITUENTS AND PROTOCOL FOR SNPS RELATED TO KIDNEY DISEASE IN SCD PATIENTS. ......................................................................................................................... 173
APPENDIX 4: SNAPSHOTS® REACTION CONSTITUENTS, PROTOCOL AND CAPILLARY ELECTROPHORESIS ......................................................................................................................... 183
APPENDIX 5: TAQMAN® CONSTITUENTS AND PROTOCOL ............................................ 184
APPENDIX 6: DIRECT CYCLE SEQUENCING ..................................................................... 185
LIST OF PUBLICATIONS FROM THIS WORK

ABSTRACT

Introduction
Sickle Cell Disease (SCD) is a monogenic, multi-organ hemoglobinopathy disorder that is highly prevalent in Africa, with nearly 300,000 newborn cases per year. The underlying pathophysiological mechanism of the disease involves alteration of the normal soft and biconcave disc shape of erythrocytes, to that of a rigid crescent. These abnormal red blood cells cause vaso-occlusion and intravascular hemolysis, resulting in a variety of clinical manifestations, including acute pain crises, anemia, and damage to various organs. Kidney disease is a clinical proxy of severity, developing only in a subset of patients, and is subject to modification by genetic variations. Indeed, reports have shown significant association between proteinuria and specific genetic variants in MYH9 and APOL1, and between estimated Glomerular Filtration Rate (eGFR) and End Stage Kidney Disease (ESKD) with HMOX1 variants among adult African Americans affected by SCD. However, the association between these variants and micro-albuminuria, a primary indicator of renal dysfunction, has not been investigated, nor has any study of these variants been performed among SCD patients in Africa.

Aim
The aim of this study was to investigate the association of targeted single nucleotide polymorphisms (SNPs) in APOL1, MYH9 and HMOX1, as well as a 5’ promoter dinucleotide repeat in HMOX1, with micro-albuminuria among SCD patients from Cameroon; and to compare the results to that from a cohort of non-SCD Cameroonian individuals affected by ESKD.

Methods
Patients and controls: Four groups of Cameroonian individuals were investigated; 1) a total of 413 SCD patients, phenotyped for crude albuminuria, the albumin-to-creatinine ratio (ACR) and eGFR, and characterized for HBB haplotypes; 2) 101 non-SCD controls, with both the SCD patients and this latter group of controls previously genotyped for the 3.7 kB alpha-globin gene deletion; 3) a total of 167 patients affected by ESKD of various etiologies (high blood
pressure, diabetes mellitus etc.) coupled to 4) an age-, sex- and comorbidity-matched control group of 162 individuals without ESKD.

**Molecular analysis:** PCR, SNaPshot sequencing, TaqMan SNP genotyping, capillary electrophoresis and Sanger sequencing were used to genotype 4, 3 and 2 target SNPs in *MYH9* (rs11912763, rs16996648, rs1005570 and rs16996672), *APOL1* (rs73885319, rs60910145 and rs71785313) and *HMOX1* (rs3074372 and rs743811), respectively.

**Statistical analysis:** Descriptive statistics were obtained for all quantitative data using SPSS (IBM, USA version 21.0). A Hardy-Weinberg Equilibrium (HWE) test was performed on the SNPs genotype results. Associations between targeted SNP genotypes in *MYH9*, *HMOX1*, and the combined *APOL1* G1/G2 variants, and kidney functions/diseases (crude albuminuria, ACR, eGFR, and ESKD outcome) were studied using a generalized linear or multinomial regression framework, adjusted for age and sex, using the R statistical package version 3.0.3 (The R Foundation for statistical computing, Vienna, Austria). The significance level was set at 0.05.

**Results**

All patients in the SCD cohort were homozygous for the HbS mutation, and consisted largely of children and adolescents with a median age of 15 years (range: 2-58). Among them, 49.2% (n=203) were female. After the analysis of 700 chromosomes, the Benin (72.9%; n = 510) and Cameroon haplotypes (20.7%; n = 145) were most prevalent. Among 338 SCD patients successfully genotyped for the 3.7 kb alpha-globin gene deletion, 41.4% (n = 140) had co-inherited alpha-thalassemia, a significantly higher proportion than among the control cohort without SCD (19%; n = 11) (p=0.003).

The SCD patients’ cohort had a prevalence of micro-albuminuria of 60.9% (n=248) and 73.4% (n=116) based on crude albuminuria and ACR data, respectively. Micro-albuminuria was significantly increased with age (p=0.004).

Among SCD patients, the development of macro-albuminuria was significantly associated with the co-inheritance of the 3.7 kb alpha-globin gene deletion (p=0.03), with alpha-thalassemia having a protective effect (3.6%, n=7 of SCD patients without the co-inherited 3.7 kb alpha-globin gene deletion developed macro-albuminuria, compared to 0.7%, n=1 of SCD patients heterozygous or homozygous for the deletion). *APOL1* G1/G2 was significantly associated with the ACR ratio in recessive or co-dominant models (p = 0.01), and tended to be associated with eGFR (CKD-EPI equation, p=0.07) in the SCD patient cohort, however not
with the development of ESKD in the non-sickle population (p = 0.98). As expected, the 5’ promoter dinucleotide repeat in HMOX1 displayed a trimodal distribution among both SCD patients and controls. Long (>25) 5’ promoter dinucleotide repeats for HMOX1 rs3074372 were significantly associated with the development of ESKD (p=0.002) in the general Cameroonian cohort, however not with micro-albuminuria (p=0.41) or ACR (p=0.81) in the SCD patients’ cohort. A near significant relationship was observed between HMOX1 rs743811 and eGFR (MDRD equation, p=0.09) in the SCD cohort.

Conclusions/perspectives
This study has revealed a high proportion of micro-albuminuria in this population of children and adolescents affected by SCD in Cameroon. The study confirmed micro-albuminuria’s association with increasing age, indicative of a directly proportional relationship between age and renal dysfunction, as well as the protective association of co-inheritance of alpha-thalassemia with proteinuria/macro-albuminuria in SCD. The significant association of APOL1 G1/G2 with ACR indicates its utility as a possible biomarker that could prove helpful in anticipatory guidance. Additional studies that include more adult SCD patients could provide further insight regarding the roles of both APOL1 G1/G2 and HMOX1 variants in kidney diseases in both SCD and general populations.
LIST OF FIGURES

Figure 1: Structure of the adult hemoglobin molecule .......................................................... 17
Figure 2: Global mean predicted HbS allele frequencies ....................................................... 18
Figure 3: The pathophysiology of Sickle Cell Disease .......................................................... 20
Figure 4: Role of heme oxygenase-1 and the actions of its products in response to repeated cycles of ischemia-reperfusion injury and hemolytic stress in SCD................................. 30
Figure 5: DNA Integrity gel of HbSS Cameroonian patients ................................................... 38
Figure 6: Initial temperature gradient for MYH9 SNP amplification ........................................ 45
Figure 7: Initial temperature gradient for APOL1 SNP amplification .................................... 46
Figure 8: Initial temperature gradient for HMOX1 SNP amplification ................................ 46
Figure 9: Secondary lower temperature gradient for rs5750248 .............................................. 49
Figure 10: Secondary higher temperature gradient for rs8141189 .......................................... 49
Figure 11: Secondary higher temperature gradient for rs3074372 (HMOX1) and rs71785313 (APOL1) ............................................................................................................................ 50
Figure 12: Reduced cycling conditions for rs73885319............................................................ 50
Figure 13: Temperature gradient for optimization of multiplexes ........................................... 52
Figure 14: Magnesium chloride gradient for rs5750248 optimization .................................... 53
Figure 15: DMSO gradient for rs5750248 optimization .......................................................... 55
Figure 16: Temperature gradient with 8% DMSO for rs5750248 optimization ......................... 55
Figure 17: Betaine concentration gradient for rs8141189 optimization ................................. 56
Figure 18: Temperature gradient for Multiplex C optimization .............................................. 57
Figure 19: Amplification of SCD patient samples for each PCR protocol for MYH9 SNPs .... 58
Figure 20: SNaPshot genotyping of SNPs rs11912763, rs16996648 and rs1005570 and rs16996672 for patient 472....................................................................................................................... 65
Figure 21: Fragment Analysis of rs3074372 and rs71785313 polymorphisms for SCD patient 65 ............................................................................................................................................. 67
Figure 22: TaqMan SNP Genotyping plot of SCD cohort plate 3 for rs73885319 ..................... 69
Figure 23: TaqMan SNP Genotyping plot of SCD cohort plate 3 for rs60910145 ..................... 70
Figure 24: Successful direct cycle sequencing validation of MYH9 SNPs ............................... 73
Figure 25: Successful direct cycle sequencing validation of HMOX1 SNP rs743811 ............. 74
Figure 26: Successful direct cycle sequencing validation of APOL1 G1/G2 allele ................. 75
Figure 27: Proximity of APOL1 G1/G2 alleles ....................................................................... 76
Figure 28: Histogram distribution of variables age and BMI in the SCD patient cohort .......... 82
Figure 29: Histogram distribution of variables SBP and DBP in the SCD patient cohort ....... 83
Figure 30: Histogram distribution of age in the non-SCD patient cohort............................... 84
Figure 31: Scatter plots illustrating the relationship between hematological indices and crude albuminurin in the SCD cohort ..................................................................................... 87
Figure 32: Histogram distribution of eGFR variables in the SCD cohort........................................ 89
Figure 33: eGFR estimates per age group for SCD patients .......................................................... 90
Figure 34: Histogram distribution of the prevalence of hyperfiltration in various age groups in the SCD cohort ................................................................................................................. 91
Figure 35: Prevalence of crude albuminuria and ACR in the SCD patient cohort .................... 92
Figure 36: Histogram distribution of micro-albuminuria .............................................................. 93
Figure 37: Histogram of the frequency of (A) ACR and (B) Log10ACR values ......................... 94
Figure 38: Albuminuria estimates per age group for SCD patients .......................................... 95
Figure 39: Histogram distribution of the prevalence of micro-albuminuria in various age groups in the SCD cohort ............................................................................................................. 95
Figure 40: Scatter plots illustrating the relationship between eGFR values and logSBP in the SCD patient cohort ................................................................................................................................... 97
Figure 41: Scatter plots illustrating the relationship between eGFR values and logDBP in the SCD patient cohort ................................................................................................................................... 98
Figure 42: Scatter plots illustrating the relationship between albuminuria values and eGFR in the SCD patient cohort ................................................................................................................. 100
Figure 43: Scatter plots illustrating the relationship between eGFR values and log(ACR) in the SCD patient cohort ................................................................................................................................... 102
Figure 44: Scatter plots illustrating the relationship between eGFR values and BMI in the SCD patient cohort ................................................................................................................................... 104
Figure 45: Scatter plots illustrating the relationship between eGFR values and age in the SCD patient cohort ................................................................................................................................... 105
Figure 46: Box and whisker plots showing the association of eGFR values with sex in the SCD cohort ................................................................................................................................... 106
Figure 47: Scatter plot illustrating the positive relationship between age and crude albuminuria in the SCD patient cohort ................................................................................................. 108
Figure 48: Scatter plot illustrating the negative relationship between albuminuria and hemoglobin in the SCD patient cohort .............................................................................................. 108
Figure 49: Fragment Analysis of rs3074372 and rs71785313 polymorphisms for SCD patient 102 ....................................................................................................................................... 110
Figure 50: TaqMan SNP Genotyping plot of SCD cohort plate 3 for rs743811 ......................... 111
Figure 51: Co-inheritance of alpha-thalassemia among SCD patients and non-SCD controls ............................................................................................................................................... 114
Figure 52: Prevalence of APOL1 G1/G2 variant among SCD patients and non-SCD controls ............................................................................................................................................... 114
Figure 53: Histogram of distribution of HMOX1 rs3074372 (GT)$_n$ promoter repeats .......... 115
Figure 54: Comparison of MAF of MYH9 SNPs between various populations ......................... 117
Figure 55: Comparison of MAF of HMOX1 and APOL1 SNPs between various populations .......................................................... 118
Figure 56: Comparison of genotype frequencies of HMOX1 rs3074372 between our Cameroonian cohorts ........................................................................................................ 119
Figure 57: Comparison of genotype frequencies of APOL1 G2 (rs71785313) between our Cameroonian cohorts .............................................................................................................................................. 120
Figure 58: Box and whisker plots showing the association of log$_{10}$ACR values with each of the four target SNPs in MYH9 in the general cohort ........................................................................................................ 122
Figure 59: Box and whisker plots showing the association of APOL1 G1/G2 with crude albuminuria, ACR, and eGFR values in the SCD patient cohort ........................................................................................................ 126
Figure 60: Box and whisker plots showing the association of HMOX1 S/L Genotypes with crude albuminuria and ACR values in the SCD patient cohort ........................................................................................................ 127
Figure 61: Box and whisker plots showing the association of HMOX1 rs743811 Genotypes with crude albuminuria and ACR values in the SCD patient cohort ........................................................................................................ 128
Figure 62: Histogram distribution of age in the ESKD cohort ..................................................... 131
LIST OF TABLES

Table 1: Genetic variants in MYH9 associated with the development of proteinuria in SCD patients .......................................................................................................................................................... 27
Table 2: Genetic variants in APOL1 associated with the development of kidney disease in SCD .......................................................................................................................................................... 29
Table 3: Genetic variants in HMOX1 associated with the development of kidney disease in SCD .......................................................................................................................................................... 31
Table 4: Nucleotide sequence and properties associated with PCR primers designed to amplify the following SNPs associated with kidney disease ................................................................................................................. 40
Table 5: Nucleotide sequence and properties associated with SNaPshot® primers designed to sequence the following SNPs associated with kidney disease ................................................................................................................. 61
Table 6: Socio-demographic variables of the SCD patient cohort and the non-SCD control cohort .......................................................................................................................................................... 80
Table 7: Anthropometric variables of the SCD patient cohort, including the BMI, systolic and diastolic blood pressure values of patients ................................................................................................................. 81
Table 8: Hematological indices of the SCD patient and non-SCD control cohorts ................................................................................................................. 86
Table 9: Estimated Glomerular Filtration Rates calculated using the MDRD, CG and CKD-API equations in the SCD patient cohort .......................................................................................................................................................... 88
Table 10: Summary of the kidney functioning measurements of the SCD patient cohort ................................................................................................................. 96
Table 11: Factors affecting crude albuminuria levels in Cameroonian SCD patients ................................................................................................................. 109
Table 12: Factors affecting eGFR levels in Cameroonian SCD patients for the MDRD, CKD-EPI and Cockcroft Gault equations. The relative equation is indicated in brackets ................................................................................................................. 109
Table 13: The frequency of genotypes of rs11912763, rs16996648, rs1005570, rs16996672 in MYH9 and HBB haplotypes within the SCD patient cohort .......................................................................................................................................................... 112
Table 14: The frequency of α-globin gene genotypes, genotypes of rs3074372, rs743811, rs73885319, rs60910145, rs71785313 and APOL1 G1/G2 alleles within the SCD patient and non-SCD control cohort .......................................................................................................................................................... 113
Table 15: The minor allele frequencies (MAF) of rs11912763, rs16996648, rs1005570, rs16996672 in the patient cohort (Cameroon) compared to various African populations, Africa, Europe, East Asia and South Asia .......................................................................................................................................................... 116
Table 16: P-values obtained from association analysis between the corresponding target SNPs and select kidney disease phenotypes for the SCD patient cohort ................................................................................................................. 124
Table 17: Socio-demographic variables of the ESKD case and control cohort ................................................................................................................. 130
Table 18: The frequency of the genotypes of rs3074372, rs743811, rs73885319, rs60910145, rs71785313 and APOL1 G1/G2 alleles within the ESKD case and control cohorts .......................................................................................................................................................... 132
LIST OF ABBREVIATIONS

°C – degree Celsius
ΔG – free energy change
µL – microliter
µM – micromolar
3’ – three prime
5’ – five prime
ACR – Albumin-to-creatinine ratio
ANOVA – analysis of variance
*APOL1* – Apolipoprotein 1 gene
*BCL11A* – B-cell lymphoma/leukemia 11A gene
BLAST – Basic Local Alignment Search Tool
BMI – body mass index
bp – base pairs
CAR – Central African
CG – Cockcroft-Gault
CKD – Chronic kidney disease
CKD-EPI – Chronic Kidney Disease Epidemiology Collaboration
CO – Carbon monoxide
DBP – diastolic blood pressure
Df – degrees of freedom
dL – deciliter
DMSO – Dimethyl sulfoxide
DNA – deoxyribonucleic acid
ddNTP – dideoxynucleotide triphosphates
dNTPs – deoxynucleotide triphosphates
eGFR – Estimated glomerular filtration rate
ESRD – end stage renal disease
*Exo*I – Exonuclease I
FastAP – Thermosensitive Alkaline Phosphatase
FSGS – focal segmental glomerulosclerosis
g – grams
GFR – glomerular filtration rate
Glu – Glutamic Acid
GWAS – Genome wide association study
Hb – Hemoglobin
HbA – adult hemoglobin
HbAS – heterozygous for the HbS allele.
*HBB* – hemoglobin beta
HbC – hemoglobin C
HbF – fetal hemoglobin
HbS – sickle hemoglobin
HBS1L-MYB – intergenic region between HBS1-Like Translational GTPase gene and V-Myb Avian Myeloblastosis gene
HbSS – homozygous for the sickle hemoglobin allele.
HMOX1 – Heme oxygenase 1 (gene)
HO-1 – Heme oxygenase 1 (protein)
HSCs – hematopoietic stem cells
HWE – Hardy Weinberg Equilibrium
IDT – Integrated DNA Technologies
Indel – insertion/deletion
IQR – interquartile range
kB – kilobases
Lys - Lysine
M – Molar
MAF – Minor Allele Frequency
MCHC - mean corpuscular hemoglobin concentration
MCV – mean corpuscular volume
MDRD – Modification of Diet in Renal Disease
Mg²⁺ - Magnesium ions
MgCl₂ – Magnesium Chloride
mL – milliliter
mV – millivolts
MYH9 – Myosin 9 gene
N – number
NCBI – National Centre for Biotechnology
NO – Nitric Oxide
p – probability value
PAH – pulmonary artery hypertension
PCR – Polymerase Chain Reaction
pmol – picoMolar
RHC – right heart catheterization
ROS – Reactive oxygen species
SBP – systolic blood pressure
SCA – Sickle Cell Anemia
SCD – Sickle Cell Disease
SCT – Sickle Cell Trait
SNP – Single Nucleotide Polymorphism
Tₘ – annealing temperature
TAMV – time-averaged mean velocities
TBE – Tris-borate EDTA
TCD – Transcranial Doppler Ultrasonography
\( T_m \) – melting temperature
TRV – tricuspid regurgitation jet velocity
TTE – transthoracic echocardiography
U – units
USA – United States of America
UV – Ultraviolet light
Val – Valine
VOC – vaso-occlusive crisis
w/v – weight volume ratio
WBC – white blood cells
WES – whole exome sequencing
WGS – whole genome sequencing
CHAPTER 1: INTRODUCTION AND LITERATURE REVIEW

1.1. HEMOGLOBINOPATHIES

The heme molecule is a protein moiety used ubiquitously for the transport and storage of oxygen (Tracz et al., 2007), making it indispensable for cell function and survival. Hemoglobinopathies are the most common monogenic disorders affecting humans, with an estimated 1/7 of the world population being unaffected carriers (Flint et al., 1998). These disorders can be due to quantitative or qualitative deficiencies in adult hemoglobin production, which normally comprises two α-globin and two β-globin chains (Figure 1). Quantitative deficiencies include the α- and β-thalassemia’s, where there is deficient synthesis of the respective globin protein, while qualitative deficiencies in hemoglobin production are generally due to single substitution structural mutations in the β-globin chain of adult hemoglobin (HbA) (Flint et al., 1998). These include sickle hemoglobin (HbS) (β6 Glu → Val), hemoglobin C (HbC) (β6 Glu → Lys), and other rarer forms of hemoglobin variants (Flint et al., 1998). These disorders display high prevalence in some human populations, particularly those exposed to the Plasmodium falciparum malaria parasite (Taylor et al., 2012). This is particularly relevant in an African setting, as this positive selection results in the mutant form reaching high allele frequencies in this region (Flint et al., 1998).

**Figure 1: Structure of the adult hemoglobin molecule.** The adult hemoglobin molecule consists of two α-globin (purple) and two β-globin (green) chains. The location of the sickle hemoglobin (HbS) mutation is depicted on β-globin chain, with the inset image displaying the high concentration of hemoglobin molecules packaged within erythrocytes. Adapted from (Schechter, 2008).
1.2. **EPIDEMIOLOGY OF SICKLE CELL DISEASE (SCD).**

Sickle Cell Disease (SCD) is a hematological, monogenic, multi-organ disorder of public health significance, with a global disease burden comparable to that of other non-communicable diseases, such as diabetes (Weatherall and Clegg, 2001). SCD is the most common hemoglobinopathy worldwide, with in excess of 300 000 SCA-affected births occurring annually (Piel et al., 2013). Roughly three quarters of this incidence occurs predominantly in sub-Saharan Africa (Aliyu et al., 2008; Piel et al., 2013) (Figure 2), due to the partial resistance conferred to heterozygotes for the allele against *Plasmodium Falciparum* (Malaria) infection (Flint et al., 1998). This is emphasized by evidence that the sickle hemoglobin (HbS) mutation originated independently in four different regions in Africa, defined by four beta-globin cluster haplotypes; Senegal, Benin, Cameroon and Bantu or Central African (CAR) (Stuart et al., 2004), as identified by specific RFLP patterns across this region, even though this multiple origin of the Beta-S mutation has recently been questioned (Bitoungui et al., 2015). Indian-Arab represents the fifth defined haplotype (Gabriel and Przybylski, 2010).  

**Figure 2:** Global mean predicted HbS allele frequencies. The highest allele frequency corresponds to the malaria-endemic belt in sub-Saharan Africa due to the partial resistance conferred to heterozygotes for the sickle allele (HbS). Adapted from (Piel et al., 2013).

1.3. **PATHOPHYSIOLOGY AND CLINICAL MANIFESTATIONS OF SICKLE CELL DISEASE**

Sickle Cell Anemia (SCA), the most prevalent and severe form of SCD (Piel et al., 2013), is caused by homozygosity for a substitution mutation (Glu→Val) in the 6th codon of the β-globin subunit of adult hemoglobin (HbA), facilitating the polymerization of hemoglobin during
deoxygenation. This polymerization subsequently disrupts the structure and flexibility of the erythrocytes, converting the normal biconcave disc shape to that of a sickle shape, which forms the primary basis of the molecular pathogenesis of SCD (Rees et al., 2010; Stuart et al., 2004). Patients are therefore susceptible to vaso-occlusion and subsequent organ damage (Bartolucci and Galacteros, 2012). Heterozygotes (HbAS) for the disease are generally asymptomatic, as the functional copy of the gene produces a sufficient amount of HbA (Bartolucci and Galacteros, 2012; Rees et al., 2010). In addition to episodic cycles of ischemia-reperfusion injury and subsequent organ damage seen in homozygotes for the HbS allele (i.e. HbSS), abnormal adhesion of leukocytes and platelets, coagulation, inflammation and hypoxia are also considered clinical manifestations of SCD. These manifestations of SCA are brought about through two major pathways: vaso-occlusion with ischemia-reperfusion injury and hemolytic anemia (Rees et al., 2010).

Studies in transgenic SCD murine models indicate that vascular inflammation, and subsequent leukocyte-endothelium adhesion stimulate vaso-occlusive episodes (Belcher, 2003; Kaul and Hebbel, 2000). This is further mediated by dynamic interactions between sickle erythrocytes and the vascular endothelium. Microvascular occlusion, ischemia and subsequent reperfusion promotes increased tissue injury (Rees et al., 2010). More severe phenotypes associated with the disease include stroke, cardiovascular disease, hemolytic anemia, pain, vascular-occlusive crises, acute chest syndrome, hypertension, pulmonary artery hypertension, and nephropathy (Flanagan et al., 2013; Piel et al., 2013; Rees et al., 2010; Stuart et al., 2004). These phenotypes are the result of various consequences of vascular occlusion, including infarction, hemolysis and the presence of free radicals (Figure 3). Furthermore, chronic hemolytic anemia or inflammation are major additional causes of chronic complications of the disease.
Figure 3: The pathophysiology of Sickle Cell Disease. The polymerization of hemoglobin during deoxygenation facilitates sickling of the erythrocytes, subsequently causing vaso-occlusion. This results in infarction and hemolysis, with reperfusion releasing free radicals. The repeated cycles of ischemia and reperfusion ultimately reduce medullary flow, subsequently causing kidney disease. These represent the major pathways through which complications of SCD arise. Adapted from Rees 2010 (Rees et al., 2010).

Sickle erythrocytes undergo extensive hemolysis in vivo due to the unstable nature of the sickle hemoglobin molecule, as demonstrated by increased bilirubin, plasma hemoglobin and reticulocyte counts, and reduced plasma haptoglobin in murine models (Belcher et al., 2006; Nath and Katusic, 2012). Plasma hemoglobin-mediated nitric oxide scavenging induces endothelial dysfunction (Jison et al., 2004), which acts with the production of reactive oxygen species (ROS) and cytokine activation to stimulate the inflammatory state characteristic of SCD. Furthermore, increased pulse pressure, recognized as a cardiovascular risk factor, may
be due to nitric oxide depletion as a result of increased hemolysis commonly seen in SCD patients, resulting in increased arterial stiffness (Novelli et al., 2014). Episodic cycles of ischemia-reperfusion injury, in combination with hemolytic anemia and the inflammatory state characteristic of SCD cumulatively contribute to progressive end organ damage.

1.4. TREATMENT OF SCD

Treatment approaches for SCD are generally adapted to the clinical phenotype of the patient, and include five broad approaches; preventative, supportive, symptomatic, curative and abortive (Ballas et al., 2012). The preventative approach aims to preclude the development of certain disease complications, such as vaccines to prevent pneumonia, blood transfusions to avert the development of primary and secondary stroke episodes, as well as hydroxyurea for the induction of fetal hemoglobin (HbF) (Platt et al., 1984; Ware et al., 2011). The supportive approach is the most common, as it promotes a healthy diet, folic acid supplementation and hydration for optimal management of the patient. The symptomatic approach, including blood transfusions, analgesia, and antibiotics (Alhashimi et al., 2010), aims to alleviate certain symptoms subsequent to their development. The only curative approach available for SCD is the transplantation of hematopoietic stem cells (HSCs) (Bernaudin et al., 2007). Nitric Oxide (NO) is the only current treatment accepted for the abortive approach, as it has been reported to permanently terminate the occurrence of chronic pain episodes in some SCD patients (Atz and Wessel, 1997). The treatment regime for SCD patients is therefore dependent on the clinical manifestation of the disease (Pule and Wonkam, 2014). Early identification of these clinical events, particularly cardiovascular phenotypes, may allow for the initiation of preventative treatment approaches to delay manifestations of more severe symptoms, such as stroke and kidney disease.

Recent improvement in the health-related quality of life of SCD patients in countries such as the US due to research and therapeutic advancements (Diallo and Tchernia, 2002) has decreased morbidity and mortality. Long term studies in Jamaica and the US have shown that early detection and treatment of acute events (Vichinsky et al., 1990), as well as hydroxyurea treatment (Charache et al., 1995) and blood-transfusion therapy (Adams et al., 1998) can reduce the severity of SCD and subsequently improve the quality of life of patients (Makani
et al., 2007). The early detection of possible complications that could involve genetic risk factors can inform anticipatory guidance for treatment of SCD, is a fundamental obstacle in low income and developing countries, particularly in Africa, as most patients have limited access to clinical care due a myriad of factors including distance and/or cost (Diallo and Tchernia, 2002). Low-income developing countries still require improvement in infrastructure to offer basic primary medical care and ensure disease awareness (Diallo and Tchernia, 2002; Makani et al., 2007).

1.5. GENETICS AND PHENOTYPIC VARIABILITY IN SICKLE CELL DISEASE

Despite the simple underlying causal factor, clinical manifestations of SCD vary greatly, and range from early childhood mortality to an almost undetectable condition where patients survive to late adulthood (Bartolucci and Galacteros, 2012; Sebastiani et al., 2005). The numerous possible phenotypes associated with SCD highlight the clinically heterogeneous nature of the disease, and indicates a possible role for genetic modifiers in development and predisposition towards specific phenotypes.

Genetic factors that alter the clinical course of the disease and underlie the varying clinical manifestations that occur in SCD represent a major focal area in current research, as the interaction between multiple genetic factors and the environment are considered the source of phenotypic variation in patients (Adams et al., 2003). The α-thalassemia polymorphism, a 3.7kb deletion (Flanagan et al., 2011), is proposed to ameliorate the clinical phenotype of SCD (Flanagan et al., 2013, 2011; Ohene-Frempong et al., 1998; Rees et al., 2010), with a particular protective association against the development of cardiovascular disease (Adams et al., 1994; Flanagan et al., 2013; Hsu et al., 2003; Lamarre et al., 2014; Nebor et al., 2010). The quantitative reduction of hemoglobin in each erythrocyte decreases the tendency and ability of HbS to polymerise, consequently diminishing the rate of hemolysis, however it may result in more frequent vaso-occlusive pain episodes (Steinberg, 2005).

Fetal hemoglobin (HbF) is another genetic factor that ameliorates the clinical phenotype of SCD, and has been shown to decrease stroke risk in both Nigerian and African-American cohorts (Enosolease et al., 2005; Fatunde and Scott-Emuakpor, 1992; Platt et al., 1991;
Yetunde and Anyaegbu, 2001). HbF levels vary between 1-30% in SCD patients, and three genetic loci have been associated with the variability and heritability of HbF expression; BCL11A, HBS1L-MYB and the HBB cluster (Lettre et al., 2008; Thein and Menzel, 2009; Wonkam et al., 2014a; Xu et al., 2010). Recent single nucleotide polymorphisms in a BCL11A erythroid-specific enhancer have been shown to explain significant variations in HbF levels in a Cameroonian SCD population (Pule et al., 2015). These variants may therefore serve as targets for therapeutic manipulation in an attempt to increase the HbF levels in adult SCD patients (Bauer et al., 2013; Canver et al., 2015; Xu et al., 2011). The Indian-Arab and Senegal β-globin gene haplotypes (Powars et al., 1990) are also associated with higher levels of HbF and milder disease severity in comparison to the Benin, Bantu/CAR and Cameroon haplotypes which are associated with lower HbF and more severe clinical manifestations (Diagne et al., 2000; Diop et al., 1999).

1.6. FOCUS ON GENETICS AND CARDIOVASCULAR PHENOTYPES OF SCD.

Cardiovascular phenotypes in SCD include complications involving the heart (e.g. heart failure), brain (e.g. stroke), lungs (e.g. pulmonary hypertension) and kidneys (e.g. proteinuria). Cardiovascular disease represents possibly the most catastrophic complication for children with SCD, and may cause overt stroke, transient ischemic attacks, silent infarcts, and neurocognitive dysfunctions, with approximately 11% under the age of 20 experiencing overt stroke (Adams et al., 2001; Flanagan et al., 2011; Ohene-Frempong et al., 1998). Longitudinal studies from the US have demonstrated that the peak incidence of stroke occurs between the ages of 2 to 5 years old (Adams et al., 2001; Ohene-Frempong et al., 1998), with SCD patients being 250 times more likely to develop stroke than individuals from the general population (Earley et al., 1998). Due to a lack of research in developing regions, the prevalence of overt stroke may be higher in Africa than that reported in these high-income countries. However, the prevalence of overt stroke in SCD patients in Africa may be higher than that reported in these high-income countries. In the context of SCD, occlusive vasculopathy has been identified as the most common cause of cardiovascular disease (Stockman et al., 1972).

An increased tricuspid regurgitation jet velocity (TRV > 2.5 m/s), pulmonary hypertension defined by right heart catheterization, and renal failure independently confer increased
mortality in SCD (Gladwin et al., 2004; Platt et al., 1994). We have recently reviewed the epidemiology, risk factors and genetic polymorphisms associated with the development of cardiovascular phenotypes; stroke, kidney disease, pulmonary hypertension and blood pressure in the context of SCD (Geard et al. 2016). This demonstrated the role genetic variants play in the predisposition to development of cardiovascular phenotypes in SCD.

1.7. KIDNEY DISEASE IN SCD.

Epidemiology and risk factors

Kidney disease occurs in 5-18% of patients, conferring increased risk of morbidity and mortality (Ashley-Koch et al., 2011; Platt et al., 1994), and is considered a characteristic clinical manifestation of SCD. Renal damage is common in SCD as a result of long-standing disruption of blood flow through the renal medullary capillaries caused by sickled erythrocytes (Sasongko et al., 2013) due to the hypoxic, hyperosmotic and acidotic environment (Pham et al., 2000). SCD nephropathy includes proteinuria, ischemia, impaired renal acidification, papillary necrosis, hyposthenuria and supranormal proximal tubular function (Pham et al., 2000; Stuart et al., 2004). Glomerular enlargement is possibly the earliest renal abnormality in SCD (Ware et al., 2010), with pediatric patients displaying increased glomerular filtration (Ware et al., 2010; Wigfall et al., 2000) and proteinuria that develops early in life (Wigfall et al., 2000). However, the earliest clinically detectable indication of renal dysfunction is an increase in albumin excreted in the urine (Guasch et al., 1997), with micro-albuminuria representing a more sensitive marker of glomerular injury than serum creatinine clearance (Guasch et al., 1996). Micro-albuminuria is therefore considered a primary indicator of glomerulopathy, preceding development of progressive renal insufficiency and ultimate end stage renal disease in children and adult patients (Becton et al., 2010; Guasch et al., 1999, 2006). Combined micro- and macro-albuminuria display an overall prevalence of 68% in adult SCD patients (Guasch et al., 2006). Proteinuria is likely mediated by chronic hemolytic anemia, and is associated with stroke (Ohene-Frempong et al., 1998), acute chest syndrome (Castro et al., 1994), hospitalization (Wigfall et al., 2000) and increased mortality (Platt et al., 1994) in pediatric patients with SCD. The presence of
Albuminuria is also associated with decreased glomerular filtration rates in SCD patients, indicating glomerular injury (Ashley-Koch et al., 2011; Guasch et al., 1996).

Transgenic murine models provide evidence for renal functional disturbances through increased sickling and subsequent congestion of the medullary vessels in the kidney (Nath et al., 2001). In addition, SCD results in anatomical alterations of the kidney, with renal tubular defects gradually developing from 7 years old (Pham et al., 2000). This process can be partially treated through multiple blood transfusions, however repeated thrombosis due to ‘sludging’ and necrosis of the inner medulla and papillae progressively reduces the ability to improve renal function (Pham et al., 2000). Angiotensin-converting enzyme (ACE) is used for treatment of renal disorders in the general population, however its utility and safety in the context of SCD is not known (Sasongko et al., 2013). Results of studies in 8 patients with sickle cell anemia indicated that enalapril, an ACE inhibitor, may be useful in reducing microalbuminuria in patients (Aoki and Saad, 1995). However, this does not apply to levels following discontinuation of the treatment, and further studies are required to determine the effectiveness of this treatment procedure in a larger cohort of patients. To this point, no concrete support for the reduction of micro-albuminuria and proteinuria through ACE treatment in people with SCD has been produced (Sasongko et al., 2013). Treatment with hydroxyurea in children with SCD resulted in significantly reduced glomerular filtration rate, albuminuria and lactate dehydrogenase levels, while significantly increasing fetal hemoglobin (Aygun et al., 2011; Bartolucci et al., 2015; Mckie et al., 2007). Examination of a single patient has indicated that treatment may reduce renal iron overload, common in SCD patients (Stehlé et al., 2016). Hydroxyurea treatment may also serve to prevent progression of sickle cell nephropathy to ESKD (Laurin et al., 2014). Hemoglobin-mediated oxidative toxicity is also a major contributor to renal damage in the context of SCD, however treatment with haptoglobin has been shown to reduce toxic effects following hemolysis (Chintagari et al., 2015).

In general, individuals of African ancestry suffer from the highest burden of ESRD, even in the absence of co-inherited SCD (Kopp et al., 2008), however the risk factors that predispose African populations to ESRD are largely unexplained and unknown (Kopp et al., 2008). Recent studies of African SCD populations, particularly in Nigeria (Adedoyin et al., 2012; Imuetinyan...
et al., 2011; Iwalokun et al., 2012), Ghana (Osei-Yeboah and Rodrigues, 2011) and Congo (Pakasa and Sumaili, 2012) provide amassing evidence of a high prevalence of renal disease in these populations.

Genetics of kidney disease in SCD.

Genetic factors involved in the development of kidney disease in SCA patients are largely unexplored, however α-thalassemia and HbF play protective roles, as is also seen with stroke (Steinberg et al., 2003). The co-inheritance of α-thalassemia deletions is specifically associated with a lower prevalence of macro-albuminuria (Guasch et al., 1999; Lamarre et al., 2014; Nebor et al., 2010) and is thought to confer protection through reduced erythrocyte hemoglobin concentration, lower mean corpuscular volume, reduced anemia, an increase in RBC deformability and reduction of RBC aggregate strength (Guasch et al., 1999; Lamarre et al., 2014). The effect of β-globin gene haplotypes on the severity of the kidney disease remains controversial, as some studies have identified no significant relationship between albuminuria and the different haplotypes (Guasch et al., 1999), whereas longitudinal studies of US populations state that the Bantu haplotype is associated with the highest incidence of renal failure in SCA patients (Powars et al., 1991).

Genetic variation in MYH9, APOL1 and HMOX1 has been associated with SCD nephropathy through various pathological pathways and mechanisms.

MYH9

*M*Y*H9* encodes the non-muscle myosin heavy chain IIA, a vital cytoskeletal nanomotor protein (Arrondel et al., 2002; Tzur et al., 2010). The protein is expressed widely in the kidney, particularly in podocyte cells found in the renal glomerulus (Arrondel et al., 2002; Tzur et al., 2010). Here, it likely plays a role in maintaining structures that are vital for glomerular filtration (Arrondel et al., 2002; Kopp et al., 2008; Tzur et al., 2010). Genetic variation at *MYH9* has been associated with increased susceptibility to develop hypertension-attributed end-stage kidney disease (H-ESKD), focal segmental glomerulosclerosis (FSGS) and HIV-associated nephropathy (HIVAN) in the general and SCD populations (Ashley-Koch et al., 2011; Kao et al.,
2008; Kopp et al., 2008). These genetic variations have also been previously associated with earlier kidney disease-related phenotypes in populations with varying levels of African ancestry inheritance (Behar et al., 2010; Barry I Freedman et al., 2009; Barry I. Freedman et al., 2009; Nelson et al., 2010) supporting the basis for a causative variant associated with African ancestry within MYH9.

In the SCD population, a single report put forward that there was stronger evidence for the role of MYH9 in the development of proteinuria and renal dysfunction than APOL1 (Ashley-Koch et al., 2011), identifying 8 SNPs (Table 1) associated with the development of macroalbuminuria. Among these variations, rs5750248 and rs11912763 had previously been strongly associated with kidney disease among African American and Hispanic populations (Behar et al., 2010; Nelson et al., 2010). Sequencing and fine mapping of the MYH9 region has not revealed any causal variants, however, thus the precise modus operandi of MYH9 has not yet been elucidated. Further genetic studies are required to determine the mechanism and identity of other environmental and genetic factors that could possibly interact to result in podocyte injury preceding development of renal failure (Kopp et al., 2008).

### Table 1: Genetic variants in MYH9 associated with the development of proteinuria in SCD patients

<table>
<thead>
<tr>
<th>Gene</th>
<th>SNPs</th>
<th>HGVS Nomenclature</th>
<th>Effect</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>MYH9</td>
<td>rs5750248</td>
<td>NC_000022.11:g.36306846T&gt;C</td>
<td>Risk</td>
<td>(Ashley-Koch et al., 2011)</td>
</tr>
<tr>
<td></td>
<td>rs8141189</td>
<td>NC_000022.11:g.36318665T&gt;A</td>
<td>Risk</td>
<td>(Ashley-Koch et al., 2011)</td>
</tr>
<tr>
<td></td>
<td>rs11912763</td>
<td>NC_000022.11:g.36288676G&gt;A</td>
<td>Risk</td>
<td>(Ashley-Koch et al., 2011)</td>
</tr>
<tr>
<td></td>
<td>rs16996648</td>
<td>NC_000022.11:g.36296706T&gt;C</td>
<td>Risk</td>
<td>(Ashley-Koch et al., 2011)</td>
</tr>
<tr>
<td></td>
<td>rs5756152</td>
<td>NC_000022.11:g.36316427A&gt;G</td>
<td>Risk</td>
<td>(Ashley-Koch et al., 2011)</td>
</tr>
<tr>
<td></td>
<td>rs16996672</td>
<td>NC_000022.11:g.36329925C&gt;T</td>
<td>Risk</td>
<td>(Ashley-Koch et al., 2011)</td>
</tr>
<tr>
<td></td>
<td>rs1557529</td>
<td>NC_000022.11:g.36309484A&gt;G</td>
<td>Risk</td>
<td>(Ashley-Koch et al., 2011)</td>
</tr>
<tr>
<td></td>
<td>rs1005570</td>
<td>NC_000022.11:g.36319229A&gt;G</td>
<td>Risk</td>
<td>(Ashley-Koch et al., 2011)</td>
</tr>
</tbody>
</table>

APOL1

In combination with MYH9, APOL1 variants have been associated with increased risk for FSGS and chronic kidney disease in African Americans (Ashley-Koch et al., 2011; Genovese et al.,
2010a; Kopp et al., 2008, 2011; Tzur et al., 2010). APOL1 encodes the trypanolytic factor found in human serum, which functions as an innate protection mechanism responsible for lysis of *Trypanosoma brucei rhodesiense* and *T. brucei gambiense*, the pathogens responsible for human African trypanosomiasis (i.e. African sleeping sickness) that is endemic to Africa (Kasembeli et al., 2015b; Limou et al., 2014; Vanhamme et al., 2003), with *T. brucei gambiense* being the prevalent strain occurring in Cameroon. There are two coding variants in the APOL1 gene, referred to as G1 and G2, which are common in sub-Saharan Africa (Limou et al., 2014). G1 consists of two non-synonymous variants rs73885319 (p. Ser358Gly) and rs60910145 (p.Ile400Met), while G2 is a 6bp deletion, rs71785313 (p.Asn404_Tyr405del) (Table 2). Due to the close proximity of the G1 and G2 variants, they are likely mutually exclusive, with recombination between them being very unlikely (Genovese et al., 2010a). The parasite is able to infect humans through a serum resistance-associated (SRA) protein that interacts with the C-terminal helix of APOL1, an interaction that is inhibited by the G1 and G2 variants, which has positively selected for the allele (Genovese et al., 2010a).

The G1 and G2 variants contribute to increased risk for the development of kidney disease in African Americans (FSGS and H-ESKD) in a recessive model (Genovese et al., 2010a), as well as with the development of HIV-associated nephropathy in Black South Africans (Kasembeli et al., 2015a, 2015b). Recent studies have found that APOL1 missense mutations were more strongly associated with risk of ESKD development than the leading MYH9 risk candidates (Nelson et al., 2010; Tzur et al., 2010). This possibly suggests that different disease and global contexts may be associated with different alleles. Social demographic status, the presence of other communicable diseases and lifestyle factors cumulatively act with APOL1 to determine the susceptibility of patients to the development of CKD (Kasembeli et al., 2015b).

APOL1 is a component of the plasma cell-free hemoglobin scavenging pathway, complexing with haptoglobin-related proteins to form the trypanosome lytic factor (Nielsen et al., 2006; Saraf et al., 2015). The G1 and G2 variants have reduced ability to bind and clear plasma-free hemoglobin (Saraf et al., 2015). Association studies have found a relationship between these variants and kidney disease in SCD patients, possibly through enhanced risk of hemoglobinuria (Saraf et al., 2014). Furthermore, these APOL1 variants may be important mediators of podocyte and kidney microvascular injury (Limou et al., 2014), with APOL1-
mediated risk of kidney disease following an autosomal-recessive pattern of inheritance (Tzur et al., 2010).

<table>
<thead>
<tr>
<th>Gene</th>
<th>SNPs</th>
<th>HGVS Nomenclature</th>
<th>Effect</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>APOL1</td>
<td>rs73885319</td>
<td>NP_663318.1:p.Ser358Gly</td>
<td>Risk</td>
<td>(Ashley-Koch et al., 2011; Saraf et al., 2015)</td>
</tr>
<tr>
<td></td>
<td>rs60910145</td>
<td>NP_663318.1:p.Ile400Met</td>
<td>Risk</td>
<td>(Saraf et al., 2015)</td>
</tr>
<tr>
<td></td>
<td>rs71785313</td>
<td>NP_663318.1:p.Asn404_Tyr405del</td>
<td>Risk</td>
<td>(Saraf et al., 2015)</td>
</tr>
</tbody>
</table>

**HMOX1**

Heme-oxygenase 1 is an inducible component with a major role in the catabolism of free heme (Tenhunen et al., 1968). Chronic hemolysis is a common clinical event in SCD due to the unstable nature of sickle hemoglobin (Nath and Katusic, 2012), with subsequent increased plasma levels of heme, a known pro-inflammatory molecule, postulated to preclude vascular injury and kidney damage in the context of SCD (Kanakiriya et al., 2003; Nath and Katusic, 2012; Tracz et al., 2007). Heme metabolizing pathways are differentially expressed between SCD patients and controls, with HMOX1 significantly induced in blood mononuclear cells of SCD patients (Jison et al., 2004). This correlated with increased levels of lactate and biliverdin reductases, as both are markers of increased hemolysis typical of SCD patients (Jison et al., 2004). HO-1, the protein produced by the HMOX1 gene, is the rate limiting step in catalyzing the oxidation of free heme to biliverdin, carbon monoxide and iron (Maines, 1997; Tenhunen et al., 1969, 1968; Yoshida and Kikuchi, 1979). Biliverdin reductase then reduces biliverdin to bilirubin, a potent antioxidant (Tenhunen et al., 1968). The HO-1 pathway exerts its anti-inflammatory and vascular protective effects through the actions of its products, carbon monoxide and bilirubin (Figure 4), in the context of repetitive ischemia-reperfusion injuries that occur with SCD (Jison et al., 2004). Carbon monoxide, a known cytoprotectant and vasodilator, has the ability to suppress proinflammatory cytokines and mitigate vascular-cellular interactions during hypoxia (Morita and Kourembanas, 1995; Otterbein et al., 1999). The induction of HO-1 ensures significantly reduced sickle erythrocyte-vascular adhesion (Bains et al., 2010).
**Figure 4: Role of heme oxygenase-1 and the actions of its products in response to repeated cycles of ischemia-reperfusion injury and hemolytic stress in SCD.** Heme Oxygenase catalyzes the reduction of free heme to biliverdin, iron and Carbon Monoxide (CO). These molecules then proceed to induce effector molecules cGMP (cyclic guanosine monophosphate), p38MAPK (p38 mitogen-activated protein kinase), and p21 (cyclin-dependent kinase inhibitor). These molecules stimulate the various vaso-protective effects listed. Red text represents upregulated molecules or genes identified in SCD patients (Jison et al., 2004). Adapted from (Jison et al., 2004).

*HMOX1* is induced in various tissues such as the spleen, liver, bone marrow, brain, lung, and the kidney (Tenhunen et al., 1968). Animal model studies in transgenic SCD mice have identified *HMOX1* as an ameliorator of both inflammation and vaso-occlusion, with induction occurring in the proximal tubules of the kidney in response to increased cell free hemoglobin (Belcher et al., 2006; Chintagari et al., 2015; Ghosh et al., 2011). Moreover, the overexpression of HO-1 reduced stasis induced by repeated cycles of hypoxia and re-oxygenation in transgenic SCD mice (Belcher et al., 2006) In the human sickle kidney, increased expression of HO-1 in circulating endothelial cells is observed, with induction occurring in the renal tubules, vasculature and interstitial cells (Nath et al., 2001).
Genetic variants in \textit{HMOX1} have been associated with increased risk of acute kidney injury (Table 3), and progression to chronic kidney disease (CKD) in the general population (Nath, 2014). Two promoter polymorphisms in \textit{HMOX1} (Table 3) are suggested to play functional roles in the regulation of HO-1 levels, and are possibly associated with the development of kidney disease through insufficient protection of the kidney from hemoglobin-mediated toxicity (Bean et al., 2012). \textit{HMOX1} rs743811 has been associated with CKD and ESRD in various cohorts (Saraf et al., 2015), with varying length repeats of rs3074372, a (GT)$_n$ dinucleotide repeat, having contrasting effects on the expression levels of heme oxygenase (Hirai et al., 2003; Saraf et al., 2015; Taha et al., 2010).

\begin{table}[h]
\centering
\caption{Genetic variants in \textit{HMOX1} associated with the development of kidney disease in SCD}
\begin{tabular}{|c|c|c|c|c|}
\hline
\textbf{Gene} & \textbf{SNPs} & \textbf{HGVS Nomenclature} & \textbf{Effect} & \textbf{References} \\
\hline
\textit{HMOX1} & rs3074372 & NC_000022.11:g.35380894_35380895insGT & Protective/Risk & (Saraf et al., 2015) \\
\hline
 & rs743811 & NC_000022.11:g.35396981T>C & Risk & (Saraf et al., 2015) \\
\hline
\end{tabular}
\end{table}

1.8. SUMMARY OF LITERATURE REVIEW

Global disparities in research stem not only from lack of resources but also from uneven distribution of research funding or poor targeting of scientific investments. In this context, sickle cell disease (SCD) vastly impacts the African continent and is associated with cardiovascular diseases. Stroke, kidney disease and pulmonary hypertension are considered as proxies of severity in SCD with several genomic loci implicated in their heritability. The aim of the present expert review was to examine data on epidemiology and genetic risk factors of stroke, pulmonary hypertension and kidney disease associated with SCD, and indexed in PubMed and Google Scholar. Multiple publications have identified genetic factors which contribute to the development risk of cardiovascular phenotypes, however little of this research has been conducted in an African setting. There is an urgent need for genomic (and multi-omics) research on cardiovascular diseases associated with SCD, particularly in Africa to allow for proportional, just, and ethical investment of global research funding on diseases that greatly impact the African continent and yet remain relatively underfunded. Ultimately, this will result in socially responsible research investments and identification of at-risk
individuals and improved preventive medicine which should be a cornerstone of global precision medicine.

1.9. RATIONALE OF THE PRESENT STUDY

The association between targeted single nucleotide polymorphisms (SNPs) in APOL1, MYH9 and HMOX1, as well as a 5′ promoter dinucleotide repeat in HMOX1 and micro-albuminuria, a sensitive primary indicator of glomerulopathy, has not yet been investigated in SCD, specifically in Africa. Moreover, kidney disease, and more specifically micro-albuminuria data, have never been reported among SCD patients in Cameroon. Furthermore, to the best of our knowledge there is no such study associating genetic polymorphisms with kidney disease in Africa, and more specifically association of these SNPS with an early indicator of glomerular dysfunction (micro-albuminuria). It is anticipated that this knowledge could be used to detect patients at risk of renal insufficiency, and to inform anticipatory guidance.

The high incidence of cardiovascular disease in SCD patients highlights the need for improved prognostic techniques for at-risk individuals, specifically in the case of renal dysfunction. The identification of genetic polymorphisms associated with these critical cardiovascular phenotypes of SCD could be used to create a panel of biomarkers to identify at-risk individuals and better inform surveillance and treatment approaches. Limited genetic studies associated with critical cardiovascular phenotypes in SCD have been reported in patients residing in Africa, indicating an urgent need to perform these studies that could inform, in a unique way, the global SCD communities about the values that genes and environment interactions play in the pathogenesis and hopefully the care of SCD. Understanding the genetic networks modulating the development of cardiovascular phenotypes may provide additional information regarding the underlying pathogenesis of SCD and subsequently lead to the identification of new therapeutic targets.

1.10. AIM AND OBJECTIVES

The aim of this study was to determine the association between selected SNPs in MYH9, APOL1 and HMOX1, as well as a 5′ promoter dinucleotide repeat in HMOX1 with micro-
albuminuria among a sample of Cameroonian SCD patients, and compare these results to that from a cohort of non-SCD Cameroonian individuals affected by End Stage Kidney Disease (ESKD) of various etiologies.

The objectives include studying:

1) The socio-demographic and anthropometric variables, micro-albuminuria, and estimated Glomerular filtration rate (eGFR) profile of SCD patients, as well as other clinical and hematological variables;

2) The frequency of specific variants in *MYH9, APOL1* and *HMOX1*, the 3.7 α-globin gene deletion, and the *HBB* haplotypes in a selected sample of SCD patients and non-SCD controls from Cameroon;

3) The association of these genetic variants in *MYH9, APOL1* and *HMOX1*, and the 3.7 α-globin gene deletion with micro-albuminuria and eGFR in SCD patients;

4) The comparison of the results to those of a cohort of non-SCD Cameroonian individuals affected by End Stage Kidney Disease (ESKD).
CHAPTER 2: METHODS AND MATERIALS

2.1. PATIENT DATA

2.1.1. ETHICAL APPROVAL

Ethical approval was granted by the University of Cape Town, Faculty of Health Sciences Human Research Ethics Committee (HREC REF: 661/2015) and the National Ethical Committee Ministry of Public Health, Republic of Cameroon (No 033/CNE/DNM/07). All participants older than 18 signed consent forms, while informed consent was given by the parents or guardians for participants younger than 18 years old.

2.1.2. STUDY PARTICIPANTS

Four groups of Cameroonian individuals were investigated in this study:

1) A total of 413 SCD patients were selected, based on available cardiovascular data, particularly micro-albuminuria and serum creatinine measurements, from a larger study conducted on Sickle Cell Disease patients recruited from hospitals in Yaoundé and Douala, the two largest cities in Cameroon.

2) A cohort of non-SCD patients (n=101) randomly selected from the same population of apparently healthy patients at Central Hospital, Yaoundé.

For the patients’ group, clinical and socio-demographic data were prospectively collected using a structured questionnaire administered to parents/guardians, as well as adult patients. SCD patient’s medical records were reviewed to define their clinical features over the past 3 years, specifically hospital outpatient visits, overt strokes, blood transfusions, vaso-occlusive crisis (VOC), and administration of hydroxyurea. Anthropometric variables such as height, weight, body mass index (BMI), and systolic and diastolic blood pressures (SBP and DBP) were also measured for patients.

The HbS carrier status of patients was determined using either alkaline hemoglobin electrophoresis or high-performance liquid chromatography (HPLC). The majority of both SCD
patients and controls were previously genotyped for \( \alpha \)-thalassemia genotypes, while only the SCD patients were previously genotyped for \( HBB \) haplotypes, and clinical data relating to renal functions were prospectively collected.

3) A total of 167 non SCD-patients affected by ESKD of various etiologies (diabetes mellitus, lupus, high blood pressure etc.) coupled to

4) an age-, sex- and comorbidity-matched control group of 162 non SCD-individuals without ESKD, from the general Cameroonian population.

2.1.3. RENAL FUNCTIONS DEFINITION AND PHENOTYPING

2.1.3.1. ALBUMINURIA AND THE ALBUMIN-TO-CREATININE RATIO (ACR) IN THE SCD COHORT

Laboratory tests were performed to measure serum creatinine concentration, urine creatinine concentration, and the level of albumin in the urine was measured with either a Siemens Clinitek Status (Erlangen®, Germany) test or the Hemocue Albumin 20 system (Angelholm®, Sweden). Urinary albumin quantitation was performed on first morning void samples at the day of a planned regular hospital visit when patients were not experiencing a SCD vaso-occlusive crisis. To determine the prevalence of albuminuria in the SCD cohort, the crude albuminuria and Albumin-to-creatinine ratio (ACR) values were divided into categories. The presence of albumin in the urine was defined as normal (<30 mg/mL), micro-albuminuria (30-300 mg/mL) and macro-albuminuria (>300 mg/mL). The ACR data, represented by milligrams albumin per millimole creatinine, was defined as normo-albuminuria (<3 mg/mmol), micro-albuminuria (3-30mg/mmol) or macro-albuminuria (>30 mg/mmol). This was defined previously in studies utilizing the same crude albuminuria detection systems (Ranque et al., 2014).

2.1.3.2. THE ESTIMATED GLOMERULAR FILTRATION RATE

The glomerular filtration rate (GFR) was estimated using three equations; (1) the Modification of Diet in Renal Disease (MDRD) (Levey et al., 1999), (2) Cockcroft-Gault (Cockcroft and Gault, 1976) and the (3) Chronic Kidney Disease Epidemiology Collaboration (CKD-EPI) (Levey et al.,
These equations estimate the GFR of patients based on variables such as plasma creatinine levels, age, weight, and ethnicity (Cockcroft and Gault, 1976; Levey et al., 2009, 1999). The formulae for the respective equations are displayed below, where sex1 takes the value 1 for women and 0 for men, while sex2 takes the value 1 for men and 0 for women.

1. eGFR (ml/min) = \((175 \times (\text{Creatinine}/10)^{-1.154}) \times (\text{Age}^{-0.203}) \times (0.742^{\text{sex1}})\)

2. Cockcroft-Gault (CG) = \[
\left(\left((140 - \text{Age}) \times \text{Weight} \times \frac{10}{(\text{Creatinine} \times \text{BSA})} \right) \times \text{sex2} + \left((140 - \text{Age}) \times \text{Weight} \times 10 \times \frac{0.85}{(\text{Creatinine} \times \text{BSA})} \right) \times \text{sex1}\right)
\]

BSA = \((\text{Weight}^{0.425}) \times ((\text{Height}_{cm}/100)^{0.725}) \times 0.20247\)

Calculate normalized Cockcroft-Gault:

eGFR (ml/min) = \((CG \times \frac{1.73}{\text{BSA}})\)

3. Compute the subcomponents for control purpose:

\[scr_k = \left(\left((\text{Creatinine}^{\text{sex1}}/10 \times 0.7)^{\text{sex1}}\right) \times \left((\text{Creatinine}^{\text{sex2}}/10 \times 0.9)^{\text{sex2}}\right)\right)\]

\[\text{alph} = -((0.329^{\text{sex1}} \times (0.411^{\text{sex2}})))\]

\[\text{max}_sc = \text{scr}_k \text{ if } \text{scr}_k \geq 1, \text{otherwise} = 1\]

\[\text{min}_sc = \text{scr}_k \text{ if } \text{scr}_k \leq 1, \text{otherwise} = 1\]

\[\text{eGFR} = (141 \times (\text{min}_sc^{\text{alph}}) \times (\text{max}_sc^{-1.209}) \times (0.993^{\text{Age}}) \times (1.018^{\text{sex1}}))\]

The proportion of patients suffering from glomerular hyperfiltration (i.e. >130 ml/min/1.73m² for women and >140 ml/min/1.73m² for men), as defined previously (Haymann et al., 2010; Vazquez et al., 2014) was also investigated.
2.1.3.3. **END STAGE KIDNEY DISEASE**

End Stage Kidney Disease, also known as renal failure, frequently leads to death in patients unless replacement by dialysis and transplant procedures are successfully implemented (Levey et al., 2005). Classification of ESKD includes a GFR < 15 ml/min/1.73 m², with increased albuminuria values corresponding to increased risk of mortality (Levey et al., 2005).

2.2. **MOLECULAR METHODS**

The SCD cohort was genotyped for the MYH9 SNPs, while all four cohorts were genotyped for variants in APOL1 and HMOX1 following the methods described below.

2.2.1. **DNA QUANTIFICATION AND INTEGRITY ANALYSIS**

Blood samples that were previously taken from the patients in Cameroon had DNA samples were prospectively isolated following the manufacturer’s instructions (Puregene Blood Kit, Qiagen, Alameda, CA, USA). The process of DNA isolation can subject the DNA to possible mechanical damage, resulting in compromised DNA integrity and subsequent fragmentation. This may hinder downstream applications such as the Polymerase Chain reaction (PCR) and sequencing, therefore it is necessary to verify the integrity of the DNA. This was achieved using a 1% (w/v) agarose gel (Appendix 1). Two microliters (2 µL) of DNA and 5 µL of Fermentas 1X loading dye was electrophoresed at 80 mV for 45 minutes.
The short and bright DNA smear seen under ultraviolet (UV) light for samples of the SCD cohort (Figure 5) is indicative of good quality DNA that was not fragmented by the isolation procedure.

![Figure 5: DNA Integrity gel of HbSS Cameroonian patients.](image)

An electrophoretic gel for the resolution of HbSS Cameroonian patient DNA samples on a 1% (w/v) agarose gel. The lane numbers correspond to the respective patients from whom DNA was isolated. SYBR® Safe DNA Gel Stain was used to visualize the PCR products under UV light on the UVIGold Transilluminator. Lane MWM displays the GeneRuler™ 100bp Plus DNA molecular weight marker. The single bright band of large molecular weight is indicative of a high yield of good quality DNA.

The concentration and purity ($A_{260nm}/A_{280nm}$ ratio) of the DNA stock solutions was determined on the NanoDrop® ND-1000 Spectrophotometer using the NanoDrop® 1000 computer software (Appendix 2), and subsequent dilutions were made to produce working samples of 100 ng/µl in a final reaction volume of 100 µL for SNPs amplified for MYH9 SNAPSHOT sequencing, and 10 ng/µl in a final reaction volume of 50 µl for fragment analysis and TaqMan SNP genotyping of polymorphisms in APOL1 and HMOX1.

2.2.2. PRIMER DESIGN

To ensure that specific regions containing the target SNPs are amplified during PCR, oligonucleotide primer pairs must bind to particular flanking sequences of DNA with high fidelity. The 500bp reference sequence flanking each of the target SNPs in MYH9 (rs5756129, rs11912763, rs16996648, rs5750248, rs1557529, rs8141189, rs1005570, rs16996672), APOL1 (rs73885319, rs60910145 and rs71785313) and HMOX1 (rs3074372 and rs743811) was obtained and extracted in FastA format from the Ensembl Genome Browser website.
using their respective rs numbers. The primer pairs were designed manually using this sequence.

The properties of each primer pair were analysed using the OligoAnalyzer tool v3.1, specifically the GC content (%), melting temperature (°C), and the free energy change (ΔG, kcal.mole⁻¹) and melting temperature (°C) of hairpin and self-dimer intramolecular structures, as well as for hetero-dimer structures formed between primer pairs. The PCR product size was limited to 200-700bp to ensure sufficient amplification during PCR, and allow the combination of specific MYH9 primer sets in multiplex. The GC content preferably ranged from 40-60%, and the melting temperature was set to between 54-58°C, with a maximum 2°C difference between primer pairs. The free energy change (ΔG) in the formation of all inter- and intramolecular structures was confirmed to be greater than -6 kcal/mole, with a melting temperature below 20°C to ensure that these structures would not exist at the reaction temperature of the PCR and hinder the efficiency and specificity of amplification.

Primer pairs that met the aforementioned criteria where then subjected to an NCBI (National Centre for Biotechnology) Primer-BLAST search (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) to ensure that the selected oligonucleotide primers did not hybridize to nonspecific regions in the human genome that may be preferentially amplified during PCR. Oligonucleotide primer pairs that met all the criteria for the target SNPs for SNaPshot sequencing and validation of TaqMan® SNP genotyping were ordered from Whitehead Scientific®SA (IDT). The oligonucleotide primer stocks were reconstituted in TE Buffer (7M Tris and 0.3M EDTA, pH 7.3) to a concentration of 100 μM, and then diluted to 20 μM working concentrations using SABAX H₂O and stored at 4°C.

Fragment analysis was employed to determine the length of the polymorphic (GT)ₙ tandem repeat (rs3074372) in the promoter region of HMOX1, and to genotype patients for a 6bp deletion in APOL1 (rs71785313, referred to as G2). The sequence of the primers used for rs3074372 have been described previously (Kimpara et al., 1997), while the G2 primers were manually designed from the Ensembl sequence. 6-FAM™ (Applied Biosystems, California, USA) was used as the fluorescent tag for rs3074372, and HEX™ (Applied Biosystems, California, USA) as the fluorescent tag for rs71785313. These primer sets were ordered from
DNA Synthesis Service at UCT (synthesized using the Applied Biosystems 394 DNA synthesizer). A non-fluorescent forward primer for rs71785313 of the same sequence was ordered for validation of the fragment analysis results. Table 4 shows the nucleotide sequence and properties of each primer for the target SNPs.

**Table 4:** Nucleotide sequence and properties associated with PCR primers designed to amplify the following SNPs associated with kidney disease; rs5756129 (Table 4.1.), rs11912763 (Table 4.2.), rs16996648 (Table 4.3.), rs5750248 (Table 4.4.), rs1557529 (Table 4.5.), rs8141189 (Table 4.6.), rs1005570 (Table 4.7.), rs1699672 (Table 4.8.), rs73885319 (Table 4.9.), rs60910145 (Table 4.10), rs71785313 (Table 4.11), rs3074372 (table 4.12) and rs743811 (table 4.13).

**Table 4.1. Properties of Primers (rs5756129) (MYH9)**

<table>
<thead>
<tr>
<th>Properties</th>
<th>Forward</th>
<th>Reverse</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequence</td>
<td>5'-GGCTGCTTCTGTGAAACC-3’</td>
<td>5'-GCCAAGCTCATATGAAAGGTGCC-3’</td>
</tr>
<tr>
<td>Length (bases)</td>
<td>17</td>
<td>22</td>
</tr>
<tr>
<td>GC content (%)</td>
<td>58.8</td>
<td>50</td>
</tr>
<tr>
<td>Melting temperature (°C)</td>
<td>53.8</td>
<td>56</td>
</tr>
<tr>
<td>Molecular Weight (g/mole)</td>
<td>5177.4</td>
<td>6728.4</td>
</tr>
</tbody>
</table>

Expected PCR fragment size: 324bp

**Table 4.2. Properties of Primers (rs11912763) (MYH9)**

<table>
<thead>
<tr>
<th>Properties</th>
<th>Forward</th>
<th>Reverse</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequence</td>
<td>5'-CCAGTCTGCTGGCATCATG-3’</td>
<td>5'-CAGAGCGAGGAGAAGAAGAAG-3’</td>
</tr>
<tr>
<td>Length (bases)</td>
<td>20</td>
<td>21</td>
</tr>
<tr>
<td>GC content (%)</td>
<td>55</td>
<td>52.4</td>
</tr>
<tr>
<td>Melting temperature (°C)</td>
<td>55.7</td>
<td>54.7</td>
</tr>
<tr>
<td>Molecular Weight (g/mole)</td>
<td>6093</td>
<td>6611.4</td>
</tr>
</tbody>
</table>

Expected PCR fragment size: 268bp

**Table 4.3. Properties of Primers (rs16996648) (MYH9)**

<table>
<thead>
<tr>
<th>Properties</th>
<th>Forward</th>
<th>Reverse</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequence</td>
<td>5'-AACGTTCTGGCTCAAGTC-3’</td>
<td>5'-CAGACCCCATACGGACCAGAT-3’</td>
</tr>
<tr>
<td>Length (bases)</td>
<td>18</td>
<td>19</td>
</tr>
<tr>
<td>GC content (%)</td>
<td>55.6</td>
<td>57.9</td>
</tr>
<tr>
<td>Melting temperature (°C)</td>
<td>555.4</td>
<td>56.6</td>
</tr>
<tr>
<td>Molecular Weight (g/mole)</td>
<td>5499.6</td>
<td>5766.8</td>
</tr>
</tbody>
</table>

Expected PCR fragment size: 420bp
<table>
<thead>
<tr>
<th>Table 4.4.</th>
<th>Properties</th>
<th>Primers (rs5750248) (MYH9)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sequence</strong></td>
<td>Forward</td>
<td>Reverse</td>
</tr>
<tr>
<td><strong>Length (bases)</strong></td>
<td>19</td>
<td>18</td>
</tr>
<tr>
<td><strong>GC content (%)</strong></td>
<td>57.9</td>
<td>61.1</td>
</tr>
<tr>
<td><strong>Melting temperature (°C)</strong></td>
<td>55.4</td>
<td>56.3</td>
</tr>
<tr>
<td><strong>Molecular Weight (g/mole)</strong></td>
<td>5935.9</td>
<td>5502.6</td>
</tr>
</tbody>
</table>

Expected PCR fragment size: 579bp

<table>
<thead>
<tr>
<th>Table 4.5.</th>
<th>Properties</th>
<th>Primers (rs1557529) (MYH9)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sequence</strong></td>
<td>Forward</td>
<td>Reverse</td>
</tr>
<tr>
<td><strong>Length (bases)</strong></td>
<td>22</td>
<td>17</td>
</tr>
<tr>
<td><strong>GC content (%)</strong></td>
<td>50</td>
<td>64.7</td>
</tr>
<tr>
<td><strong>Melting temperature (°C)</strong></td>
<td>56</td>
<td>56.7</td>
</tr>
<tr>
<td><strong>Molecular Weight (g/mole)</strong></td>
<td>6661.4</td>
<td>5242.4</td>
</tr>
</tbody>
</table>

Expected PCR fragment size: 302bp

<table>
<thead>
<tr>
<th>Table 4.6.</th>
<th>Properties</th>
<th>Primers (rs8141189) (MYH9)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sequence</strong></td>
<td>Forward</td>
<td>Reverse</td>
</tr>
<tr>
<td><strong>Length (bases)</strong></td>
<td>19</td>
<td>19</td>
</tr>
<tr>
<td><strong>GC content (%)</strong></td>
<td>57.9</td>
<td>52.6</td>
</tr>
<tr>
<td><strong>Melting temperature (°C)</strong></td>
<td>55.9</td>
<td>54.8</td>
</tr>
<tr>
<td><strong>Molecular Weight (g/mole)</strong></td>
<td>5739.8</td>
<td>5741.8</td>
</tr>
</tbody>
</table>

Expected PCR fragment size: 478bp

<table>
<thead>
<tr>
<th>Table 4.7.</th>
<th>Properties</th>
<th>Primers (rs1005570) (MYH9)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Sequence</strong></td>
<td>Forward</td>
<td>Reverse</td>
</tr>
<tr>
<td><strong>Length (bases)</strong></td>
<td>22</td>
<td>20</td>
</tr>
<tr>
<td><strong>GC content (%)</strong></td>
<td>50</td>
<td>50</td>
</tr>
<tr>
<td><strong>Melting temperature (°C)</strong></td>
<td>55.3</td>
<td>53.3</td>
</tr>
<tr>
<td><strong>Molecular Weight (g/mole)</strong></td>
<td>6768.5</td>
<td>6148</td>
</tr>
</tbody>
</table>

Expected PCR fragment size: 671bp
<table>
<thead>
<tr>
<th>Properties</th>
<th>Primers (rs16996672) (MYH9)</th>
<th>Primers (rs73885319) (APOL1)</th>
<th>Primers (rs60910145) (APOL1)</th>
<th>Primers (rs71785313) (APOL1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Sequence</strong></td>
<td>5’-CAACACACTCGAGGGCATG-3’</td>
<td>5’-GGTAAATTGGCTTCAGGAATTGG-3’</td>
<td>5’-GGAAATGAGCAGAGGAGTCAA 5’-TGTGCTCAGCTATGGAAATGC-3’</td>
<td>5’-HEX-CTTCAGTCAGTACCAGCATGC-3’ 5’-AGTTTGCATTITGTCTGCGC-3’</td>
</tr>
<tr>
<td><strong>Length (bases)</strong></td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>20</td>
</tr>
<tr>
<td><strong>GC content (%)</strong></td>
<td>55</td>
<td>52.4</td>
<td>50</td>
<td>55</td>
</tr>
<tr>
<td><strong>Melting temperature (°C)</strong></td>
<td>56.5</td>
<td>55.5</td>
<td>55.4</td>
<td>56</td>
</tr>
<tr>
<td><strong>Molecular Weight (g/mole)</strong></td>
<td>6120</td>
<td>6326.2</td>
<td>6906.5</td>
<td>6053</td>
</tr>
<tr>
<td></td>
<td>Expected PCR fragment size: 447bp</td>
<td>Expected PCR fragment size: 315bp</td>
<td>Expected PCR fragment size: 666 bp</td>
<td>Expected PCR fragment size: 384bp</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 4.12.

<table>
<thead>
<tr>
<th>Properties</th>
<th>Primers (rs3074372) (HMOX1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Forward</td>
</tr>
<tr>
<td>Sequence</td>
<td>5’-FAM-AGAGCCTGCAGCTTCTCAGA-3’</td>
</tr>
<tr>
<td>Length (bases)</td>
<td>20</td>
</tr>
<tr>
<td>GC content (%)</td>
<td>55</td>
</tr>
<tr>
<td>Melting temperature (°C)</td>
<td>58.4</td>
</tr>
<tr>
<td>Molecular Weight (g/mole)</td>
<td>6102</td>
</tr>
</tbody>
</table>

Expected PCR fragment size: 67bp +/- repeats

Table 4.13.

<table>
<thead>
<tr>
<th>Properties</th>
<th>Primers (rs743811) (HMOX1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Forward</td>
</tr>
<tr>
<td>Sequence</td>
<td>5’-TAGCAGCTGGACTCTCCTG-3’</td>
</tr>
<tr>
<td>Length (bases)</td>
<td>20</td>
</tr>
<tr>
<td>GC content (%)</td>
<td>60</td>
</tr>
<tr>
<td>Melting temperature (°C)</td>
<td>58.4</td>
</tr>
<tr>
<td>Molecular Weight (g/mole)</td>
<td>6109</td>
</tr>
</tbody>
</table>

Expected PCR fragment size: 551bp

2.2.3. POLYMERASE CHAIN REACTION (PCR)

Successful amplification of the target sequence was achieved through the polymerase chain reaction, consisting of repeated thermal cycling intervals of DNA denaturation, primer annealing to its complementary DNA sequence, and subsequent Taq polymerase-mediated extension in the 5’ to 3’ direction from both the forward and reverse primers. The total reaction volume was 25 µL, with the final concentrations of each oligonucleotide primer, each dNTP, Taq polymerase, and colourless GoTaq® buffer being 10 pM, 200 µM, 0.02 U/µL and 1X, respectively. 100 ng of DNA template was then added to each reaction, with Sabax sterile water being used to make up the remaining reaction volume. A non-template control (NTC) was included in each reaction to ensure no contamination by secondary DNA sources had occurred.

The necessary conditions to facilitate amplification were achieved using the Bio-Rad® thermal cycler T100™ (Bio-Rad laboratories, Hercules, CA, USA). The thermal cycling
conditions included an initial three minute denaturation step, typically at 95°C. This was succeeded by 35 cycles of a second denaturation step, primer annealing, and extension at 95°C for 30 sec, primer-specific annealing temperature (Tₐ) for 30 sec, and 72°C for one minute, respectively. A final elongation step of 72°C for 5 minutes completed the reaction.

2.2.4. PCR OPTIMIZATION

PCR optimization is the process whereby the optimal reaction conditions for each primer pair in the PCR reaction are established. These included the cycling temperatures, concentration of specific additives, and establishment of primer multiplexes to increase the efficiency and decrease the cost of the amplification procedure. The various optimization procedures were performed on the Bio-Rad® thermal cycler T100™ using the several procedures described below, which consisted of slight adjustments of the standard PCR protocol (section 2.2.3.) and/or the cycling conditions.

2.2.4.1. TEMPERATURE GRADIENT

The annealing temperature (Tₐ) of a PCR is the temperature at which the oligonucleotide primers hybridize to the target template DNA. It must be high enough to inhibit imperfect and nonspecific complementarity, yet low enough to allow sufficient primer-template DNA hybrids and subsequent ample amplification of the target region. An initial eight-point temperature gradient from 50°C-60°C was used (50.0°C, 50.7°C, 51.9°C, 53.8°C, 56.1°C, 58°C, 59.2°C and 60°C) to determine the optimum annealing temperature for the PCR protocol of each SNP in MYH9 (Figure 6), APOL1 (Figure 7) and HMOX1 (Figure 6).
Figure 6: Initial temperature gradient for MYH9 SNP amplification.
Electrophoretic gels for the resolution of SNPs (A) rs5756129 (324 bp), (B) rs11912763 (268 bp), (C) rs16996648 (420 bp), (D) rs5750248 (579 bp), (E) rs1557529 (302 bp), (F) rs8141189 (478 bp), (G) rs1005570 (671 bp), and (H) rs16996672 (447 bp) on a 2% (w/v) agarose gel. SYBR® Safe DNA Gel Stain was used to visualize the PCR products under UV light on the UVIGold Transilluminator. Lane MWM displays the GeneRuler™100bp Plus DNA molecular weight marker. Lanes 1 to 8 represent the annealing temperature gradient of 60°C-50°C (1 - 60°C, 2 - 59.2°C, 3 - 58°C, 4 - 56.1°C, 5 - 53.8°C, 6 - 51.9°C, 7 - 50.7°C and 8 - 50°C) using South African HbAA control DNA B3A. Figure 4.I. contains the NTC’s for the respective SNPs.
Figure 7: Initial temperature gradient for APOL1 SNP amplification.

Electrophoretic gels for the resolution of SNPs (A) rs71785313 (384 bp), (B) rs60910145 (666 bp), (C) rs73885319 (315 bp), and (D) water blanks for each of the respective SNPs on a 2% (w/v) agarose gel. SYBR® Safe DNA Gel Stain was used to visualize the PCR products under UV light on the UVIGold Transilluminator. Lane MWM displays the GeneRuler™ 100bp Plus DNA molecular weight marker. Lanes 1 to 8 represent the annealing temperature gradient of 60°C - 50°C (1 - 60°C, 2 - 59.2°C, 3 - 58°C, 4 - 56.1°C, 5 - 53.8°C, 6 - 51.9°C, 7 - 50.7°C and 8 - 50°C) using South African HbAA control DNA B8M. Lanes 71-N, 73-N and 6-N represent the NTC’s for rs71785313, rs73885319 and rs60910145, respectively.

Figure 8: Initial temperature gradient for HMOX1 SNP amplification.

Electrophoretic gels for the resolution of SNPs (A) rs3074372 (±127 bp depending on number of repeats), and (B) rs743811 (551 bp) on a 2% (w/v) agarose gel. SYBR® Safe DNA Gel Stain was used to visualize the PCR products under UV light on the UVIGold Transilluminator. Lane MWM displays the GeneRuler™ 100bp Plus DNA molecular weight marker. Lanes 1 to 8 represent the annealing temperature gradient of 60°C - 50°C (1 - 60°C, 2 - 59.2°C, 3 - 58°C, 4 - 56.1°C, 5 - 53.8°C, 6 - 51.9°C, 7 - 50.7°C and 8 - 50°C) using South African HbAA control DNA B8M. Lanes (A) 3-NTC and (B) 74-N represent the water blanks for rs3074372 and rs71785313, respectively.
The MYH9 SNPs rs5756129, rs11912763, rs16996646, rs1557529, rs1005570 and rs16996672 displayed satisfactory specific amplification of the target DNA region at 60°C (Temperature 1, Figure 6). SNP rs5750248 displayed amplification of nonspecific regions at temperatures 1-8, with increasing specificity at lower temperatures, therefore a second eight-point temperature gradient was performed from 45°C-55°C (45.0°C, 45.7°C, 46.9°C, 48.9°C, 51.2°C, 52.9°C, 54.2°C, 55°C) (Figure 9). Rs8141189 displayed preferential amplification of the target fragment and decreased non-specificity at higher temperatures of the initial temperature gradient, therefore a second gradient was performed from 55°C-65°C (55.0°C, 55.7°C, 56.9°C, 58.5°C, 61.1°C, 63.0°C, 64.3°C, 65.0°C) (Figure 10).

The initial temperature gradient for APOL1 rs71785313 and HMOX1 rs3074372 yielded little to no amplification for either primer set, despite repetition of the reaction to ensure experimental error was not the underlying cause. A lower temperature gradient of 48°C-55°C (48°C, 48.5°C, 49.3°C, 50.7°C, 52.3°C, 53.6°C, 54.5°C, 55°C), and doubling of the primer concentrations (final concentration of 20pm per reaction) yielded no positive results. Next, a MgCl₂ gradient (section 2.2.4.2) was performed in an attempt to catalyse the incorporation of dNTPs to initiate amplification, however this was also unsuccessful in producing any amplicons for either polymorphism. KAPA2G Fast HotStart (KAPA Biosystems) (Cape Town, South Africa OR Wilmington, Massachusetts) enzyme offers increased processivity and speed of reaction in comparison to wild-type Taq DNA polymerase. This is due to a process of directed evolution used for manufacturing of the enzyme. The enzyme is supplied with a specially formulated buffer (ReadyMix) allowing improved specificity, yield and overall reaction efficiency. Due to its improved robustness in comparison to normal Taq, the KAPA2G Fast HotStart ReadyMix was used in the optimization procedures in an attempt to encourage amplification of the target fragment. The ReadyMix is supplied at a 2X concentration containing KAPA2G Fast HotStart DNA Polymerase, dNTPs, MgCl₂ and stabilizers at a concentration of 0.5 U per 25ul reaction, 0.2 mM at 1X and 1.5 mM at 1X, respectively. The reaction was optimized using 50ng of control DNA template (Sample B8M at 10ng/µl). Cycling conditions were optimized on the Bio-Rad® thermal cycler T100TM (Bio-Rad laboratories, Hercules, CA, USA). The thermal cycling conditions, as specified by the protocol, included an initial three minute denaturation step at 95°C, succeeded by 40 cycles of denaturation, annealing and extension at 95°C for 15 seconds, annealing temperature for 15 seconds, and
72°C for 1 second, respectively. The reaction was concluded by a final extension cycle at 72°C for 1 minute. Rs71785313 (APOL1) and rs3074372 (HMOX1) displayed large amounts of nonspecific amplification (Figure 7A and 8A, respectively), particularly at lower temperatures, therefore a secondary higher temperature gradient was performed from 60°C-68°C in an attempt to reduce non-specificity (Figure 11). The secondary temperature gradient identified 60.6°C (Temperature 7, Figure 11) as the optimum annealing temperature for both rs3074372 and rs71785313, and was subsequently utilized in downstream amplification of samples for fragment analysis and direct cycle sequencing validation. Rs60910145 and rs743811 displayed satisfactory specific amplification of the target region at 60°C (Temperature 1, Figure 7) and 58°C (Temperature 3, Figure 8), respectively, which were then selected for use in further amplification procedures. Electrophoresis of rs73885319 displayed ‘smearing’ surrounding the target region. Reduction of the cycle number to 30 cycles in comparison to 35 in the standard PCR procedure (Appendix 3) was performed in attempt to improve specificity at 56.1°C (Figure 12). This was seen to reduce the nonspecific amplification and 30 cycles were used in downstream sample amplification for rs73885319.
Figure 9: Secondary lower temperature gradient for rs5750248.
An electrophoretic gel for the resolution of rs5750248 on a 2% (w/v) agarose gel. SYBR® Safe DNA Gel Stain was used to visualize the PCR products under UV light on the UVIGold Transilluminator. The expected size of the amplicon was 579 bp. Lane MWM displays the GeneRuler™ 100bp Plus DNA molecular weight marker. Lanes D1 to D8 represent the annealing temperature gradient of 55°C-45°C (1 - 55°C, 2 - 54.2°C, 3 - 52.9°C, 4 - 51.2°C, 5 - 48.9°C, 6 - 46.9°C, 7 - 45.7°C, 8 - 45°C) using South African HbAA control DNA B3A. Lane DNT represents the non-template control.

Figure 10: Secondary higher temperature gradient for rs8141189.
An electrophoretic gel for the resolution of (A) rs8141189 on a 2% (w/v) agarose gel. SYBR® Safe DNA Gel Stain was used to visualize the PCR products under UV light on the UVIGold Transilluminator. The expected size of the rs8141189 amplicon was 478 bp. Lane MWM displays the GeneRuler™ 100bp Plus DNA molecular weight marker. Lanes 1 to 8 represent the annealing temperature gradient of 65°C-55°C (1 - 65°C, 2 - 64.3°C, 3 - 63°C, 4 - 61.1°C, 5 - 58.5°C, 6 - 56.9°C, 7 - 55.7°C, 8 - 55°C) using South African HbAA control DNA B3A. (B) The non-template control rs8141189 (FNT), shows that no contamination occurred.
Figure 11: Secondary higher temperature gradient for rs3074372 (HMOX1) and rs71785313 (APOL1).

An electrophoretic gel for the resolution of (A) rs3074372 and (B) rs71785313 on a 2% (w/v) agarose gel. SYBR® Safe DNA Gel Stain was used to visualize the PCR products under UV light on the UVIGold Transilluminator. Lane MWM displays the GeneRuler™ 100bp Plus DNA molecular weight marker. Lanes 1 to 8 represent the annealing temperature gradient of 68°C-60°C (1 – 68°C, 2 – 67.4°C, 3 – 66.4°C, 4 – 64.9°C, 5 – 63.1°C, 6 – 61.6°C, 7 – 60.6°C, 8 - 60°C) using South African HbAA control DNA B8M. The lanes 3-NTC (A) and 7-NTC (B) contain the no-template controls for rs3074372 and rs71785313, respectively, indicating that no contamination has occurred.

Figure 12: Reduced cycling conditions for rs73885319.

An electrophoretic gel for the resolution of (A) rs73885319 on a 2% (w/v) agarose gel. SYBR® Safe DNA Gel Stain was used to visualize the PCR products under UV light on the UVIGold Transilluminator. Lane MWM displays the GeneRuler™ 100bp Plus DNA molecular weight marker. Lanes 73-B8M, 73-2 and 73-3 represent amplification using DNA samples South African HbAA Control B8M, SCD 2 and SCD 3, respectively. Lane 73-N contains the no-template control for rs73885319 indicating that no contamination has occurred. The cycling conditions were reduced from 35 cycles to 30.
SNP rs5750248 continued to display amplification of nonspecific DNA regions (Figure 9), therefore a Magnesium Chloride (MgCl₂) gradient was performed. SNP rs8141189 at 65°C (Temperature 1, Figure 10) showed specific, yet low amplification of the target DNA region and was then subjected to a Betaine concentration gradient in an attempt to increase the yield of the target amplicon.

Primers rs1005570 (MYH9), rs16996672 (MYH9) and rs73885319 (APOL1) were tested in a multiplex in an attempt to reduce the cost and time of PCR amplification of all the patient samples. This was due to a satisfactory difference in the size of their expected PCR fragments, necessary for resolution of the independent amplicons, and their similar annealing temperatures. Rs73885319 (APOL1) was originally intended for SNaPshot sequencing, however this was not successful and the SNP was subsequently genotyped using TaqMan® SNP genotyping methods. A temperature gradient of 58°C-65°C (58°C, 58.5°C, 59.4°C, 60.7°C, 62.3°C, 63.6°C, 64.5°C, 65°C) was performed. All three fragments amplified successfully at 60.7°C (Temperature 5, figure 13A). This multiplex was referred to as multiplex A.

Primers rs1557529 (MYH9) and rs11912763 (MYH9) were tested in a second multiplex, and a temperature gradient of 55°C-65°C was performed. The two target amplicons amplified successfully at 58.8°C (Temperature 5, Figure 13B). This was referred to as multiplex B.
The concentration of Magnesium Chloride (MgCl₂) in a polymerase chain reaction greatly influences the specificity of the enzymatic activity of the DNA Taq Polymerase. The Mg²⁺ ions
are enzyme cofactors that enhance the incorporation of the freely available dNTPs during extension of the oligonucleotide primers. High concentrations of Mg$^{2+}$ can therefore result in imperfect hybridization of the primer pairs and nonspecific PCR amplification, while low concentrations of the cofactor result in decreased yield of the target amplicon, possibly limiting downstream applications of the PCR product. The MgCl$_2$ gradient incorporated various Mg$^{2+}$ concentrations, as well as various pH values for the buffer. The gradient ranged from 1.0mM to 2.5mM MgCl$_2$ in increments of 0.5mM, with each concentration value being replicated at pH 8.3, pH 8.6, pH 8.9 and pH 9.2 at a $T_a$ of 46.9°C (figure 14) as this showed the least nonspecific amplification in the previous temperature gradient. 1.0mM at pH 8.3, 2.5mM at pH 8.3 and 2.0mM at pH 8.6 were the only MgCl$_2$ buffer parameters that amplified the fragment of interest, however non-specific amplification was still abundant and therefore the gradient was unsuccessful. GoTaq® buffer (7.5mM MgCl$_2$) was used in subsequent optimization procedures, and a Dimethyl sulfoxide concentration gradient was performed to further attempt to enhance the specificity of the reaction.

![Expected fragment size (579 bp)](image)

**Figure 14: Magnesium chloride gradient for rs5750248 optimization.**
An electrophoretic gel for the resolution of rs5750248 on a 2% (w/v) agarose gel. SYBR® Safe DNA Gel Stain was used to visualize the PCR products under UV light on the UVIGold Transilluminator. The expected size of the PCR product was 579 bp. Lane MWM displays the GeneRule™ 100bp Plus DNA molecular weight marker. Lanes D1 to D15 represent the gradient of Mg$^{2+}$ concentrations ranging from 1.0mM to 2.5mM in 0.5mM increments at pH 8.3 (Lanes D1-D4), pH 8.6 (Lanes D5-D8), pH 8.9 (Lanes D9-D12) and from 1.0mM to 2.0mM in 0.5mM increments at pH 9.2 (Lanes D13-D15) at an annealing temperature of 46.9°C using South African HbAA control DNA B3A. Lane DNT represents the non-template control for the reaction, and DNB is the no buffer control for the reaction.

### 2.2.4.3. DIMETHYL SULFOXIDE PERCENTAGE GRADIENT

53
Dimethyl sulfoxide (DMSO) is a common PCR enhancing agent that disrupts secondary DNA structures and enhances the specificity of amplification (Frackman et al., 1998). A DMSO percentage gradient of 1%, 3%, 5%, 8% and 10% of the final reaction volume was performed for SNP rs5750248 (figure 15) to reduce nonspecific amplification, with Sabax distilled water added to make up a final reaction volume of 25 µL. Lane D8 (8% DMSO) displays the least nonspecificity with sufficient amplification of the target region (579 bp). A temperature gradient of 45°C-55°C with 8% DMSO was performed to re-optimize the annealing temperature in light of the PCR additive (Figure 16), with 55°C (Temperature 1) displaying specific amplification of the target region.
Figure 15: DMSO gradient for rs5750248 optimization.
An electrophoretic gel for the resolution of rs5750248 on a 2% (w/v) agarose gel. SYBR® Safe DNA Gel Stain was used to visualize the PCR products under UV light on the UVIGold Transilluminator. The expected size of the PCR product was 579 bp. Lane MWM displays the GeneRuler™ 100bp Plus DNA molecular weight marker. Lanes D1, D3, D5, D8 and D10 represent the 1%, 3%, 5%, 8% and 10% DMSO concentration percentages, respectively, at an annealing temperature of 46.9°C using South African HbAA control DNA B3A. Lane DNTC shows the non-template control which indicates no contamination occurred, as the low molecular weight band is primer dimers. Lane DNDC shows the sample with no DMSO added. 8% DMSO produced the least nonspecific amplification and highest yield of PCR product.

Figure 16: Temperature gradient with 8% DMSO for rs5750248 optimization.
An electrophoretic gel for the resolution of rs5750248 on a 2% (w/v) agarose gel. SYBR® Safe DNA Gel Stain was used to visualize the PCR products under UV light on the UVIGold Transilluminator. The expected size of the PCR product was 579 bp. Lane MWM displays the GeneRuler™ 100bp Plus DNA molecular weight marker. Lanes D1 to D8 represent the annealing temperature gradient of 55°C-45°C (1 - 55°C, 2 - 54.2°C, 3 - 52.9°C, 4 - 51.2°C, 5 - 48.9°C, 6 - 46.9°C, 7 - 45.7°C, 8 - 45°C) with 8% DMSO using South African HbAA control DNA B3A. Lane DNT represents the non-template control. The target fragment was efficiently and specifically amplified at 55°C and 8% DMSO concentration.
2.2.4.4. BETAINE CONCENTRATION GRADIENT

Betaine acts as an isostabilizing agent during PCR (Rees et al., 1993), and it functions to equalize the effect of AT- and GC-base pairing to the stability of the DNA duplex. This results in an increase in the amplification of the target amplicon. The concentration gradient used was 0.1M, 0.2M, 0.5M and 1M in a final reaction volume of 25 µL (Figure 17). A Betaine concentration of 0.1M resulted in the largest increase in amplification of the target amplicon for SNP rs8141189 (MYH9).

Figure 17: Betaine concentration gradient for rs8141189 optimization. An electrophoretic gel for the resolution of rs8141189 on a 2% (w/v) agarose gel. SYBR® Safe DNA Gel Stain was used to visualize the PCR products under UV light on the UVIGold Transilluminator. The expected size of the PCR product was 478 bp. Lane MWM displays the GeneRuler™ 100bp Plus DNA molecular weight marker. Lanes 0.1M, 0.2M, 0.5M and 1M represent the concentrations of Betaine at 65°C using South African HbAA control DNA B3A. Lane NTC is the non-template control, while lane NBC is the no Betaine control. The target fragment was efficiently and specifically amplified at 0.1M Betaine concentration and was chosen as the optimum concentration for further optimization procedures.

SNP rs5756129 (MYH9) was tested with SNP rs8141189 in a multiplex due to the similar Ta of both primer pairs. A temperature gradient of 55°C-65°C at 0.1M Betaine concentration for Multiplex C (Figure 18) revealed that both SNPs were sufficiently and specifically amplified at 58.5°C. This multiplex was referred to as multiplex C. SNP rs16996646 (MYH9) did not amplify efficiently within this multiplex and was therefore optimized as a single plex.
2.2.4.5. SUMMARY OF THE OPTIMISATION PROCESS

The PCR optimization process produced optimal protocols for two singleplexes, two multiplexes containing two primer pairs, and a single multiplex containing three primer pairs for the MYH9 SNPs. The protocols included; (i) 8% DMSO at a Ta of 55°C for rs5750248, (ii) Ta of 60°C for rs16996646, (iii) Ta of 60.7°C for multiplex A (rs1005570, rs16996672, rs73885319) (iv) Ta of 58.8°C for Multiplex B (rs11912763, rs1557529) and (v) 0.1M Betaine and a Ta of 58.4°C for multiplex C (rs5756129, rs8141189) (Appendix 3). The PCR protocols for amplification of rs3074372 and rs71785313 for fragment analysis were optimized using KAPA2G Fast HotStart ReadyMix at a Ta of 60.1°C (Appendix 5). The protocols for direct cycle sequencing validation of rs60910145 (APOL1) and rs743811 (HMOX1) were optimized at 60°C and 58°C, respectively, while rs73885319 (APOL1) was optimized at 56.1°C with the number
of cycles reduced from 35 to 30. Figure 19 displays successful PCR amplification of a subset of patient samples for the respective MYH9 protocols.

Figure 19: Amplification of SCD patient samples for each PCR protocol for MYH9 SNPs. Electrophoretic gels for the resolution of (A) rs16996648 (420 bp), (B) rs5750258 (579 bp), (C) Multiplex A, containing rs1005570 (671 bp), rs16996672 (447 bp) and rs73885319 (315 bp), (D) Multiplex B, containing rs1557529 (302 bp) and rs11912763 (268 bp) and (E) Multiplex C containing rs5756129 (324 bp) and rs8141189 (478 bp) on a (A, B, C, E) 2% and (D) 2.5% (w/v) agarose gel. SYBR® Safe DNA Gel Stain was used to visualize the PCR products under UV light on the UVIGold Transilluminator. Lane MWM displays the GeneRuler™ 100bp Plus DNA molecular weight marker. Each lane contains the indicated number of the HbSS Cameroonian patient’s DNA sample amplified for the respective PCR protocols.
2.2.5. PCR PRODUCT DETECTION

The PCR products were all resolved through agarose gel electrophoresis (Appendix 1) using the electrophoretic power pack (Appendix 2). The agarose gels used to resolve single plex C, single plex D, multiplex A and multiplex C for MYH9 SNPs, as well as the products of amplification for fragment analysis (rs3074372 and rs71785313), and direct cycle sequencing validation (all SNPs excluding rs3074372 (HMOX1)) were prepared to 2% (w/v) using Seakem® LE agarose dissolved in 1X TBE buffer (Appendix 1). Multiplex B required a 2.5% (w/v) gel in order to resolve the more similarly sized fragments. SYBR® Safe DNA Gel Stain (InvitrogenTM, Thermo Fisher Scientific) was utilized to visualize DNA fragments. Five microliters (5 µL) of PCR product and loading dye, respectively (Appendix 1) was loaded into the gel. The DNA loading dye contained glycerol to increase the density of the sample, ensuring it remained in its allocated well, as well as bromophenol blue which assisted in loading and tracking the migration of the sample. A molecular weight marker, GeneRuler™ 100bp DNA Ladder Plus (Appendix 1), was electrophoresed alongside the samples to size the PCR product. Gel electrophoresis was performed for between 45 and 90 minutes, depending on product size, at 120mV voltage. The DNA fragments were visualized using the UVIPro UVIGold TransIlluminator (UVItec Limited), in conjunction with the UVIPro software (version 12.3 for Windows 1995-2005 ©) for editing and labelling of the gel images.

2.2.6. PCR PRODUCT CLEAN-UP

To facilitate the use of the PCR products in downstream sequencing reactions, Thermo Scientific FastAP Thermosensitive Alkaline Phosphatase (FastAP) and Exonuclease I (ExoI) clean-up was performed on all PCR products (Appendix 1). FastAP removes unincorporated dNTPs through dephosphorylation, thereby preventing any interference in downstream applications such as sequencing, while ExoI degrades the remaining single-stranded DNA molecules and oligonucleotide primers, preventing the need for further purification methods. The final reaction volume contained 1.5 units and 2 units of FastAP and ExoI, respectively, with Sabax distilled water being used to make up the final reaction volume to 25 µL. The
cycling conditions included a 37°C incubation for 60 minutes, with a subsequent 75°C incubation for 15 minutes.

2.2.7. SNaPSHOT® SEQUENCING

SNaPshot® sequencing is a single nucleotide primer extension based SNP genotyping technique that utilizes the incorporation of a fluorescently labelled dideoxyribonucleotide triphosphate (ddNTP) directly downstream of the SNaPshot® primer. These primers for MYH9 SNPs and rs73885319 (APOL1) were designed from the same sequence extracted from the Ensembl browser their properties analyzed using the OligoAnalyzer tool used previously (Section 2.2.2.). The primer parameters included a Tm of approximately 55°C, and an absence of primer-dimer or hairpin structures. Table 5 shows the nucleotide sequence and properties of each primer. Despite careful design of the primers to facilitate multiplex sequencing reactions, single plex sequencing was performed due to the inability of the ABI Prism 3130xl Genetic Analyzer to resolve the various DNA fragments. The oligonucleotide primer stocks were reconstituted in TE Buffer (Appendix 1) to a concentration of 100 µM, and then diluted to 20 µM working concentrations using Sabax H2O.

The SNaPshot® reaction comprised of the SNaPshot® primer (forward or reverse), cleaned-up PCR product (section 2.2.6.) and the SNaPshot® ready reaction mix containing the fluorescently labelled ddNTPs required for SNP genotyping, AmpliTaq® DNA Polymerase for single nucleotide extension of the DNA-primer template in the 5’ to 3’ direction, and the reaction buffer. Each reaction contained 20pmol of the respective oligonucleotide primer, 250 ng of the corresponding PCR product, and 1 µL of the SNaPshot® ready reaction mix. Sabax sterile water was used to make up the total reaction volume of 10 µL (Appendix 4).

The thermal cycling conditions for the SNaPshot® reaction included 25 cycles of fragment denaturation, primer annealing and product extension at 96°C for 10 seconds, 50°C for 5 seconds and 60°C for 30 seconds, respectively. SNaPshot sequencing was not performed for
the non-SCD controls and non-SCD ESKD cohort due to the inaccuracy of this genotyping method, as only four SNPs out of the nine (rs11912763, rs16996648, rs1005570 and rs16996672) were successfully genotyped and validated, and subsequently considered pertinent for immediate analysis. Furthermore, preliminary analysis indicated no significant association with crude albuminuria, thus MYH9 genotyping was not performed for the remaining SNPs.

Table 5: Nucleotide sequence and properties associated with SNaPshot® primers designed to sequence the following SNPs associated with kidney disease; rs5756129 (Table 5.1.), rs11912763 (Table 5.2.), rs16996646 (Table 5.3.), rs5750248 (Table 5.4.), rs1557529 (Table 5.5.), rs8141189 (Table 5.6.), rs1005570 (Table 5.7.), rs16996672 (Table 5.8.), and rs73885319 (Table 5.9.)

Table 5.1.

<table>
<thead>
<tr>
<th>Properties</th>
<th>Primer (rs5756129) (MYH9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direction</td>
<td>Forward</td>
</tr>
<tr>
<td>Sequence</td>
<td>5’-TTCTGTGTATATGACCGACCCTACACCGCTTT-3’</td>
</tr>
<tr>
<td>Complement</td>
<td>5’-AAACGGGTATGACTGGGTACATAAAAACACAGAA-3’</td>
</tr>
<tr>
<td>Length (bases)</td>
<td>34</td>
</tr>
<tr>
<td>GC content (%)</td>
<td>38.2</td>
</tr>
<tr>
<td>Melting temperature (°C)</td>
<td>60.8</td>
</tr>
<tr>
<td>Molecular Weight (g/mole)</td>
<td>10339.7</td>
</tr>
</tbody>
</table>

Expected SNaPshot product size: 35 bp

Table 5.2.

<table>
<thead>
<tr>
<th>Properties</th>
<th>Primer (rs11912763) (MYH9)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direction</td>
<td>Forward</td>
</tr>
<tr>
<td>Sequence</td>
<td>5’-AACAGAAGCCCTGGGT-3’</td>
</tr>
<tr>
<td>Complement</td>
<td>5’-ACGCAGGCTTCTGTT-3’</td>
</tr>
<tr>
<td>Length (bases)</td>
<td>15</td>
</tr>
<tr>
<td>GC content (%)</td>
<td>53.3</td>
</tr>
<tr>
<td>Melting temperature (°C)</td>
<td>51.1</td>
</tr>
<tr>
<td>Molecular Weight (g/mole)</td>
<td>4586</td>
</tr>
</tbody>
</table>

Expected SNaPshot product size: 16 bp
Table 5.3. Primer (rs16996648) (MYH9)

<table>
<thead>
<tr>
<th>Properties</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direction</td>
<td>Reverse</td>
</tr>
<tr>
<td>Sequence</td>
<td>5’-TCCACTGTTGTCTTTAAAAGAGCCATTTATTCAT-3’</td>
</tr>
<tr>
<td>Complement</td>
<td>5’-ATGAAATAAAATGGCTTTTAAGACAAACAGTGGA-3’</td>
</tr>
<tr>
<td>Length (bases)</td>
<td>34</td>
</tr>
<tr>
<td>GC content (%)</td>
<td>32.4</td>
</tr>
<tr>
<td>Melting temperature (°C)</td>
<td>58.7</td>
</tr>
<tr>
<td>Molecular Weight (g/mole)</td>
<td>10356.8</td>
</tr>
</tbody>
</table>

Expected SNaPshot product size: 35 bp

Table 5.4. Primer (rs5750248) (MYH9)

<table>
<thead>
<tr>
<th>Properties</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direction</td>
<td>Forward</td>
</tr>
<tr>
<td>Sequence</td>
<td>5’-GCCATGCAATGCTGCA-3’</td>
</tr>
<tr>
<td>Complement</td>
<td>5’-TGCCAGATGCAATGCC-3’</td>
</tr>
<tr>
<td>Length (bases)</td>
<td>15</td>
</tr>
<tr>
<td>GC content (%)</td>
<td>60</td>
</tr>
<tr>
<td>Melting temperature (°C)</td>
<td>53.0</td>
</tr>
<tr>
<td>Molecular Weight (g/mole)</td>
<td>4513</td>
</tr>
</tbody>
</table>

Expected SNaPshot product size: 16 bp

Table 5.5. Primer (rs1557529) (MYH9)

<table>
<thead>
<tr>
<th>Properties</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direction</td>
<td>Reverse</td>
</tr>
<tr>
<td>Sequence</td>
<td>5’-TCCTCAGCTTTCTGACTTTTCCA-3’</td>
</tr>
<tr>
<td>Complement</td>
<td>5’-TGGAATGCAAAACATGGGA-3’</td>
</tr>
<tr>
<td>Length (bases)</td>
<td>24</td>
</tr>
<tr>
<td>GC content (%)</td>
<td>37.5</td>
</tr>
<tr>
<td>Melting temperature (°C)</td>
<td>54.9</td>
</tr>
<tr>
<td>Molecular Weight (g/mole)</td>
<td>7210.7</td>
</tr>
</tbody>
</table>

Expected SNaPshot product size: 25 bp

Table 5.6. Primer (rs8141189) (MYH9)

<table>
<thead>
<tr>
<th>Properties</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Direction</td>
<td>Reverse</td>
</tr>
<tr>
<td>Sequence</td>
<td>5’-GGGACCTGCTTTTACACTCTGAGG-3’</td>
</tr>
<tr>
<td>Complement</td>
<td>5’-CCTCAGAGTGTAAGAGCAGGCTCC-3’</td>
</tr>
<tr>
<td>Length (bases)</td>
<td>24</td>
</tr>
<tr>
<td>GC content (%)</td>
<td>54.2</td>
</tr>
<tr>
<td>Melting temperature (°C)</td>
<td>59.2</td>
</tr>
<tr>
<td>Molecular Weight (g/mole)</td>
<td>7359.8</td>
</tr>
</tbody>
</table>

Expected SNaPshot product size: 25 bp
Table 5.7.  
<table>
<thead>
<tr>
<th>Properties</th>
<th>Primer (rs1005570) (MYH9)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Direction</strong></td>
<td>Forward</td>
</tr>
<tr>
<td><strong>Sequence</strong></td>
<td>5’-CAAGGACATAATTTCAGATCTCGACATAGTCAC-3’</td>
</tr>
<tr>
<td><strong>Complement</strong></td>
<td>5’-GTGACTATGTGGAGATCTTTGAAATTATGTCCTTG-3’</td>
</tr>
<tr>
<td><strong>Length (bases)</strong></td>
<td>34</td>
</tr>
<tr>
<td><strong>GC content (%)</strong></td>
<td>38.2</td>
</tr>
<tr>
<td><strong>Melting temperature (°C)</strong></td>
<td>58.4</td>
</tr>
<tr>
<td><strong>Molecular Weight (g/mole)</strong></td>
<td>10402.8</td>
</tr>
</tbody>
</table>

Expected SNaPshot product size: 35 bp

Table 5.8.  
<table>
<thead>
<tr>
<th>Properties</th>
<th>Primer (rs16996672) (MYH9)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Direction</strong></td>
<td>Reverse</td>
</tr>
<tr>
<td><strong>Sequence</strong></td>
<td>5’-CGTGTGCAGCTGTGTG-3’</td>
</tr>
<tr>
<td><strong>Complement</strong></td>
<td>5’-ACACAGCTGCACACG-3’</td>
</tr>
<tr>
<td><strong>Length (bases)</strong></td>
<td>15</td>
</tr>
<tr>
<td><strong>GC content (%)</strong></td>
<td>60</td>
</tr>
<tr>
<td><strong>Melting temperature (°C)</strong></td>
<td>52.9</td>
</tr>
<tr>
<td><strong>Molecular Weight (g/mole)</strong></td>
<td>4615</td>
</tr>
</tbody>
</table>

Expected SNaPshot product size: 16 bp

Table 5.9.  
<table>
<thead>
<tr>
<th>Properties</th>
<th>Primer (rs73885319) (APOL1)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Direction</strong></td>
<td>Reverse</td>
</tr>
<tr>
<td><strong>Sequence</strong></td>
<td>5’-CTACATCCAGCACAAGAAAAGAC-3’</td>
</tr>
<tr>
<td><strong>Complement</strong></td>
<td>5’-GCTTTTCTTGTGGATGATG-3’</td>
</tr>
<tr>
<td><strong>Length (bases)</strong></td>
<td>24</td>
</tr>
<tr>
<td><strong>GC content (%)</strong></td>
<td>45.8</td>
</tr>
<tr>
<td><strong>Melting temperature (°C)</strong></td>
<td>56.1</td>
</tr>
<tr>
<td><strong>Molecular Weight (g/mole)</strong></td>
<td>7332.8</td>
</tr>
</tbody>
</table>

Expected SNaPshot product size: 25 bp

**2.2.7.1. SNaPSHOT PRODUCT CLEAN-UP**

A single unit of FastAP was added to each reaction to remove any unincorporated ddNTP’s that may compromise the use of the SNaPshot® product in downstream application, such as sequencing, as excess fluorescence may hinder the correct base calling of the SNP genotype. The thermal cycling conditions included a 37°C incubation for 60 minutes, and a subsequent 75°C incubation for 15 minutes.
2.2.7.2. **CAPILLARY ELECTROPHORESIS**

The ABI Prism 3130xl Genetic Analyzer (Applied Biosystems) (Appendix 2) was used for base calling of the SNP genotype for the MYH9 SNPs. The fluorescently-labelled fragments migrate through the capillary gel polymer matrix and are detected by laser-activated fluorescence. Five microliters (5 µL) of single plex cleaned-up SNaPshot® product was added to 4.7 µL of Hi-Di™ Formamide and 0.3 µL GeneScan™ Liz120 dye-labelled size standard (Appendix 1) into each well of a 96-well microtiter plate. The Hi-Di acts to stabilize the single-stranded DNA molecules, while the size standard functions analogously to the GeneRuler™ 100bp DNA ladder (Section 2.2.5.), helping to approximate the length of the SNaPshot fragment. The final solutions were denatured at 95°C for two minutes on the Hybaid Touchdown Thermocycler, followed by a snap cooling on ice to preserve the DNA molecules in their single-stranded form. The analysis and identification of the SNP allele was performed using GeneMapper® Software version 4.0 (Appendix 2).

SNPs rs16996648 and rs16996672 were SNaPshot sequenced in the reverse direction, therefore their genotype results represent the complementary bases to those found on the forward strand. Figure 20 shows the SNaPshot genotype results for patient 472 for each SNP. The patient was homozygous for the major allele of rs11912763 and rs16996648, heterozygous for rs1005570, and homozygous for the minor allele of rs16996672.
Figure 20: SNaPshot genotyping of SNPs rs11912763, rs16996648 and rs1005570 and rs16996672 for patient 472.

Electropherogram produced using the ABI Prism 3130xl Genetic Analyzer (Applied Biosystems) for the base calling of the genotype of patient 472 for SNPs (A) rs11912763, (B) rs16996648, (C) rs1005570 and (D) rs16996672. Dye-labelled GeneScan™ Liz120 was used as the size standard for the reaction. The analysis of each genotype was performed using GeneMapper® Software version 4.0. The genotype results showed (A) homozygosity for the major allele of rs11912763, (B) homozygosity for the major allele of rs16996648, (C) heterozygosity for rs1005570 and (D) homozygosity for the minor allele of rs16996672. SNPs rs16996648 and rs16996672 were sequenced in the reverse direction, therefore the genotype of each SNP on the forward strand is the complement of that shown above.
Fragment analysis was employed to determine the length of the polymorphic (GT)$_n$ tandem repeat (rs3074372) in the promoter region of HMOX1, and to genotype patients for a 6bp deletion in APOL1 (rs71785313, referred to as G2). The 5’ flanking region of HMOX1 containing the polymorphism of interest, and the target region in APOL1 was amplified by polymerase chain reaction (PCR) using a fluorescently labelled forward primer and a standard reverse primer. 6-FAM (Applied Biosystems, California, USA) was used as the fluorescent tag for rs3074372, and HEX (Applied Biosystems, California, USA) as the fluorescent tag for rs71785313 as these dye labels are successfully detected by the Standard Dye Set D installed on the ABI3130xl (Applied Biosystems). This technique determines the length of the fragments, of which the expected length is known, allowing determination of the number of repeats, and detection of the 6bp deletion, through the size difference observed in the electropherogram images.

2.2.8.1. PCR PRODUCT CLEAN UP

PCR product clean up using FastAp and Exol (section 2.2.6.) was performed on 16 of the 32 samples used in the initial optimization step of capillary electrophoresis. The clean-up did not improve the resulting electropherograms and was deemed unnecessary for inclusion in the standard operating procedure for fragment analysis.

2.2.8.2. CAPILLARY ELECTROPHORESIS

Equal volumes (2 µl) of rs3074372 and rs71785313 PCR products for the corresponding sample were pooled and added to each well of a 96-well microtiter plate, along with 2 µl of GeneScan™ 500 ROX™ dye-labelled size standard (Applied Biosystems, Warrington, UK). Hi-Di™ Formamide was added to ensure a final reaction volume of 10 µl. Hi-Di™ Formamide acts to stabilize the single stranded DNA molecules, while GeneScan™ ROX500 functions similarly to the GeneRuler™ 100bp DNA ladder (Section 2.2.5) and the GeneScan™ Liz120
size standard used for SNaPshot sequencing (section 2.2.7.2.). The final solutions were
denatured at 95°C for five minutes using the Bio-Rad® thermal cycler T100™ (Bio-Rad
laboratories, Hercules, CA, USA), followed by snap cooling on ice to preserve the DNA
molecules in their single-stranded form. Capillary electrophoresis was performed on the ABI
3130x/!, with analysis and identification of the fragment length and corresponding genotype
for each polymorphism detected using GeneMapper® Software version 4.0 (Appendix __).
Figure 21 displays the electropherograms produced by the ABI Prism 3130x/ Genetic Analyzer
(Applied Biosystems, California, USA) for patient 65.

![Electropherogram](image)

**Figure 21: Fragment Analysis of rs3074372 and rs71785313 polymorphisms for SCD patient 65.**

Electropherogram produced using the ABI Prism 3130x/ Genetic Analyzer (Applied Biosystems) for the
genotyping of patient 65 for polymorphisms (A) rs3074372, and (B) rs71785313. Dye-labelled GeneScan™
500 ROX was used as the size standard for the reaction. The analysis of each genotype was performed
using GeneMapper® Software version 4.0. The genotype results showed (A) homozygosity for a fragment
size of 128 bp and (B) heterozygosity for the 6bp APOL1 G2 deletion. The fluorescent dyes (A) 6-FAM™
and (B) HEX™ were used for labelling of the forward primers of rs3074372 and rs71785313, respectively.


2.2.9. TAQMAN SNP GENOTYPING

TaqMan® is a SNP genotyping method utilizing differentially fluorescently labelled, allele-specific minor groove binding (MGB) probes to allow for allelic discrimination. These validated, pre-designed assays (Applied Biosystems, ThermoFisher Scientific, California, USA) were used for genotyping of rs60910145 (APO1), rs73885319 (APO1) and rs743811 (MYH9). Rs73885319 was initially sequenced using SNaphshot®, however this variant was not validated using direct cycle sequencing (section 2.2.10) and was subsequently genotyping using the TaqMan® method. The assays are supplied at 40X working stock concentration containing region-specific PCR primers, as well as two MGB probes with an allele-specific 5′ fluorescent label and 3′ nonfluorescent quencher (NFQ) to eliminate background signal detection. Exonuclease cleavage of the 5′ dye label during amplification allows for allelic discrimination through detection of the relative fluorescent units using real-time PCR technology. VIC and FAM 5′ dyes were utilized for discriminatory allelic labelling.

The working stock assays were diluted to a 20X concentration in a final reaction volume of 120 µl using TE Buffer (7M Tris and 0.3M EDTA, pH 7.3). 2X TaqMan® Universal PCR MasterMix (Applied Biosystems, ThermoFisher Scientific, Warrington, UK) was utilized as the buffer to facilitate optimal reaction conditions. The final reaction volume was optimized to 10 µl, containing a 1X final reaction concentration of both the working stock assay and TaqMan® Universal PCR MasterMix. 20 ng of template DNA was added to each reaction, with Sabax sterile water added to make up the reaction volume. The reactions were performed in 96-well white shell/clear well PCR plates (Bio-Rad laboratories, Hercules, CA, USA) and sealed with optically clear adhesive covers (ThermoFisher Scientific, Warrington, UK). Thermal cycling conditions were performed on the Bio-Rad CFX96 (Bio-Rad laboratories, Hercules, CA, USA) and included enzyme activation at 95 °C for 10 minutes, followed by 40 cycles of denaturation and annealing/extension at 95°C for 15 seconds and 60°C for 1 minute, respectively. Analysis of the amplification charts and allelic discrimination plots was performed on the Bio-Rad CFX96 (Bio-Rad laboratories, Hercules, CA, USA) Manager Software (version 3.1.).
Figures 22 and 23 display the allelic discrimination plots produced and analyzed using Bio-Rad CFX96 Manager Software (version 3.1.) (Bio-Rad laboratories, Hercules, CA, USA) for the third plate of samples for the SCD cohort for rs73885319 and rs60910145, respectively. VIC and FAM 5’-dyes were utilized for allelic discrimination. For rs73885319, VIC and FAM 5’-dyes correspond to the major (A) and minor alleles (G), respectively, while for rs60910145, VIC is associated with the minor allele (G), and FAM the major allele (T).

Figure 22: TaqMan SNP Genotyping plot of SCD cohort plate 3 for rs73885319.

TaqMan allelic discrimination plot produced using the Bio-Rad CFX96 (Bio-Rad laboratories, Hercules, CA, USA) for the genotyping of plate 1 SCD cohort patients for rs73885319. The x-axis corresponds to the relative fluorescent units (RFU) for the FAM 5’-dye label, and the y-axis corresponds to the relative fluorescent units (RFU) for the VIC 5’-due label. The analysis of each genotype was performed using Bio-Rad CFX96 (Bio-Rad laboratories, Hercules, CA, USA) Manager Software version 3.1. The FAM and VIC 5’-dyes correspond to the minor (G) and major (A) alleles, respectively. The blue squares therefore represent patients homozygous for the major allele (i.e. AA), the green triangles represent individuals heterozygous for rs73885319 (i.e. AG), and the orange circles represent patients homozygous for the minor allele (i.e. GG) The black diamond represents the NTC, validating that no contamination occurred.
Figure 23: TaqMan SNP Genotyping plot of SCD cohort plate 3 for rs60910145.

TaqMan allelic discrimination plot produced using the Bio-Rad CFX96 (Bio-Rad laboratories, Hercules, CA, USA) for the genotyping of plate 3 SCD cohort patients for rs60910145. The x-axis corresponds to the relative fluorescent units (RFU) for the FAM 5’-dye label, and the y-axis corresponds to the relative fluorescent units (RFU) for the VIC 5’-due label. The analysis of each genotype was performed using Bio-Rad CFX96 (Bio-Rad laboratories, Hercules, CA, USA) Manager Software version 3.1. The FAM and VIC 5’-dyes correspond to the major (T) and minor (G) alleles, respectively. The blue squares therefore represent patients homozygous for the minor allele (i.e. GG), the green triangles represent individuals heterozygous for rs60910145 (i.e. TG), and the orange circles represent patients homozygous for the major allele (i.e. TT) The black diamond represents the NTC, and unsuccessfully amplified samples, validating that no contamination occurred.
2.2.10. DIRECT CYCLE SEQUENCING VALIDATION

Direct cycle sequencing was performed on a subset of the SCD patient, non-SCD controls and the non-SCD ESKD cohorts in order to validate the SNaPshot sequencing, fragment analysis and TaqMan® SNP genotyping results. An example of each genotype for all of the respective SNPs, excluding rs3074372, were validated in each of the studied cohorts. This is considered the gold standard of genotyping, and a subset of the direct cycle sequencing for MYH9 SNPs was performed by an independent laboratory (Central Analytical Facilities, University of Stellenbosch), that confirmed the SNP genotyping results of four SNPs (rs11912763, rs16996648, rs11005570, and rs16996672). Each cycle sequencing reaction comprised 5 µL of the cleaned-up PCR product (Section 2.3.4.), 1X BigDye® dilution buffer, 10 pmol of the forward/reverse (unlabeled) oligonucleotide primer, and 2 µL of the BigDye® Terminator mix. Sabax sterile water was used to make the final reaction volume of 20 µL.

The cycling parameters included an initial denaturation step of 96°C for 5 minutes, which was succeeded by 30 cycles, each containing sub-steps of fragment denaturation, primer annealing, and primer extension at 96°C for 30 seconds, 50°C for 15 seconds, and 60°C for 4 minutes, respectively. This was performed on the GeneAmp® PCR System 9700 (Applied Biosystems). The DNA samples were then precipitated and prepared (Appendix 5) for capillary electrophoresis on the ABI (Appendix 2) to determine the nucleotide sequence of the PCR product through detection of the differentially fluorescently labelled ddNTP’s. This produced an electropherogram of the nucleotide sequence. Subsequent to detection and data collection, the sequencing results were analyzed using the BioEdit® software program (version 7.0.5.2; 6/5/05).

Validation confirmed the SNP genotyping results of four SNPs in MYH9 (rs11912763, rs16996648, rs11005570, and rs16996672) from SNaPshot sequencing that were considered for analysis. SNaPshot sequencing of rs73885319, one of two alleles of APOL1 G1, was not validated by direct cycle sequencing and subsequently sequenced using TaqMan® SNP genotyping. Figure 24 displays the electropherograms for the MYH9 SNPs produced by the
ABI Prism 3130xl Genetic Analyzer (Applied Biosystems, California, USA) and analyzed using the BioEdit® software program (version 7.0.5.2; 6/5/05), for patients 456 (SNP rs16996672), 345 (rs16996648), 604 (rs11912763) and 620 (rs1005570). SNP rs1005570 was sequenced using the reverse primer, while SNPs rs11912763, rs16996648 and rs16996672 were sequenced using the forward primer. Figure 25 displays the electropherograms for validation of the HMOX1 SNP rs743811. Figure 26 displays the electropherograms for validation of APOL1 SNPs rs73885319, rs60910145, and rs71785313. The nucleotide sequences, produced by the ABI Prism 3130xl Genetic Analyzer (Applied Biosystems, California, USA) and analyzed using the BioEdit® software program, for each sample amplified for rs71785313 are shown in alignment for confirmation of homozygosity or heterozygosity for the 6bp rs71785313 (APOL1) deletion. SNPs rs73885319 (APOL1) and rs71785313 (APOL1) were sequenced using the reverse primer, while rs743811 (MYH9) and rs60910145 (APOL1) were sequenced using the forward primer.
Figure 24: Successful direct cycle sequencing validation of *MYH9* SNPs.

Electropherogram produced using the ABI Prism 3130x/ Genetic Analyzer (Applied Biosystems) analyzed using the BioEdit® software program (version 7.0.5.2; 6/5/05) for the base calling of the genotype of (A) patient 606 for rs11912763, (B) patient 345 for SNP 16996648, (C) patient 620 for rs1005570 and (D) patient 456 for SNP rs16996672. SNP rs1005570 was sequenced using the reverse primer, while the other SNPs were sequenced using the forward primers. The genotype results showed (A) homozygosity for the minor allele of rs11912763, (B) homozygosity for the major allele of rs16996648, (C) heterozygosity for rs1005570. And (D) heterozygosity for rs16996672. The black arrow indicates the nucleotide of interest.
Figure 25: Successful direct cycle sequencing validation of HMOX1 SNP rs743811.

Electropherogram produced using the ABI Prism 3130xl Genetic Analyzer (Applied Biosystems) analyzed using the BioEdit® software program (version 7.0.5.2; 6/5/05) for the base calling of the genotype for rs743811 of (A) SCD patient 602, and (B) SCD patient 605. Rs743811 was sequenced using the reverse primer. The genotype results showed (A) homozygosity for the major allele (TT) and (B) homozygosity for the minor allele (GG) of rs743811. The black arrow indicates the nucleotide of interest.
Electropherogram produced using the ABI Prism 3130xl Genetic Analyzer (Applied Biosystems) analyzed using the BioEdit® software program (version 7.0.5.2; 6/5/05) for the base calling of the genotype of (A) SCD patient 461 for rs73885319, (B) non-SCD ESKD patient 143 for rs60910145, and (C) non-SCD ESKD patient 122 for rs71785313. (D) displays the alignment of non-SCD ESKD patient 143’s sequencing result against the whole reference sequence, extracted from Ensembl Genome Browser website (http://www.ensembl.org/index.html). Rs73885319 and rs71785313 were sequenced using the reverse primer, while rs60910145 was sequenced using the forward primer. The genotype results showed (A) homozygosity for the minor allele of rs73885319 (i.e. GG), (B) heterozygosity for rs60910145 (i.e. TG) and (C) homozygosity for the 6bp deletion (TTATAA), as validated by (D) alignment against the reference sequence. The black arrow indicates the nucleotide of interest.
Direct Cycle sequencing confirmed the results acquired previously for majority of the validated samples. Homozygotes for the minor allele of rs60910145, as identified by TaqMan SNP genotyping, were not validated using direct cycle sequencing. This may be due to the close proximity of rs71785313, a 6bp deletion which may affect binding of the reverse TaqMan® probe, as seen in figure 27. All samples identified as homozygotes for the rs60919145 minor allele were subsequently correctly genotyped using Direct Cycle Sequencing.

```
GACGAGCCAGAGCCAATCTTCAGTCAGTACCGCATGCTCAGTTCTCACGCCTCACGCCC CGGTTCACTGAGCCA
ATTCAGCTGAAAGGAGGTTGAAGGTTGAACATGCTCAGCTGGAAATGAGGAGCTGAGGAGCTGAGAGGTTAATGAACCC
AGTCAAGTCAAGTGAGGCTGCCCTGTA [A/G] GTCTCTCTTTCTGCTGGGATCTAGCTACCTCGTG
TACGAATCAAAGCACTTACATGAGGGGCAGGATGCAGAGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTG
GGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTGAGGAGCTG
AGAATCTGAGGAGAAGCTAAACAT [T/G] CTCAACAATAA [TTATAA/-] GATTCTGCAGCGCGACCA
AGAATCTGAGGAGAAGCTAAACAT [T/G] CTCAACAATAA [TTATAA/-] GATTCTGCAGCGCGACCA
AAACTTTTTTTTTTCTGAGACAGAGCTCTCTGCTGCTGCCAAGTTGGAGTGCAATGGTGCGATCTC
```

**Figure 27: Proximity of APOL1 G1/G2 alleles.**

Nucleotide sequence in FastA format displaying the proximity of the APOL1 G1/G2 SNPs to each other. (*) represents rs73885319, (**) represents rs60910145, and (***) represents rs71785313, a 6 bp deletion. The combination of rs73885319 and rs60910145 constitute the G1 allele, while rs71785313 is referred to as the G2 allele. The sequences were extracted from Ensembl Genome Browser website [http://www.ensembl.org/index.html](http://www.ensembl.org/index.html).

**SUMMARY OF THE MOLECULAR METHODS**

Analysis of the integrity and purity of the DNA indicated that all the samples were uncontaminated and suitable for downstream PCR and sequencing applications. These samples were amplified in multiplex and single plex reactions to increase the efficiency and decrease the cost of the assay. SNaPshot® sequencing was used for genotyping of the target SNPs in MYH9, fragment analysis for two indel’s (insertion/deletion) in APOL1 (rs71785313) and HMOX1 (rs3074372), and TaqMan® sequencing of the remaining SNPs in APOL1 (rs73885319 and rs60910145) and HMOX1 (rs743811). Direct cycle sequencing of a subset of the cohort for all polymorphisms excluding rs3074372 was performed, and confirmed the
genotypes of only four SNPs in MYH9 (rs1191376, rs16996648, rs1005570 and rs16996672). Moreover, preliminary analysis identified no significant association between these SNPs in our SCD cohort, and accumulating evidence have mechanistically shown that APOL1 is more strongly linked to the development of kidney disease, with MYH9 previously identified as a well of it being in linkage disequilibrium with APOL1. For this reason, MYH9 genotyping was not performed for the remainder of the cohorts. On the contrary, all genotypes for the APOL1 and HMOX1 variations were validated. Due to the polymorphic and repetitive nature of the rs3074372 variant, validation by direct cycle sequencing could not be performed, however this was not deemed necessary as indicated by previous studies (Bean et al., 2012; Kimpara et al., 1997; Ono et al., 2004; Saraf et al., 2015). The minor allele for rs60910145, sequenced using a predesigned validated TaqMan® assay, was not validated using direct cycle sequencing. This may be due to the close proximity of rs71785313, a 6bp deletion, possibly affecting the binding of TaqMan® probes or primers. This is currently being addressed with the TaqMan® assay international design team.

2.3. STATISTICAL ANALYSIS

2.3.1. DESCRIPTIVE STATISTICAL PACKAGES

Descriptive statistical analysis was performed using STATA® SE-64 software program (version 14.0.370 for Windows) using the prospectively collected patient data (Section 2.1.2.) Descriptive statistics were obtained for all quantitative data, particularly sociodemographic variables and hematological indices, haplotypes, as well as SNP and α-Thalassemia genotypes. PLINK (Appendix 2) (Purcell et al., 2007) was employed to examine the MAF and HWE for each of the target SNPs. The MAF for SNP results in our Cameroonian cohort was compared to those reported in other populations in the 1000 Genome project. To examine the Hardy-Weinberg equilibrium (HWE), a chi-squared test (df = 1) was performed on the respective SNP and α-Thalassemia genotypes using PLINK (Purcell et al., 2007).

2.3.2. DATA NORMALISATION.

Normality was determined by the Shapiro-Wilk Test, followed by the use of parametric (Chi-squared test and t-test) or non-parametric tests (Mann-Whitney U test for 2 samples or the
Kruskal-Wallis ONE way ANOVA for more than 2 samples). To correct for the skewness of the albuminurias, eGFR, age, BMI, systolic blood pressure, diastolic blood pressure, serum creatinine, and the hematological indices distributions, we log10-transformed the data in an attempt to normalize the variables in order to obtain the quantitative trait used in the association analysis. The ACR, leukocyte, systolic and diastolic variables were normalized using log transformation (p<0.05) and could be analyzed using parametric methods.

2.3.3. ASSOCIATION STUDIES

Spearman’s test of correlation was performed to determine the relationship between albuminuria and the numerical variables; age, hemoglobin levels, eGFR, hematological indices, and serum creatinine. Pearson’s correlation coefficient, the parametric equivalent of Spearman’s test of correlation, was performed to test the association between log(ACR), log(SBP), log(DBP) and log(leukocytes) with other non-numerical variables. The Kruskal-Wallis test was performed to compare the difference in medians between albuminuria and categorical variables; sex, \( \alpha \)-Thalassemia, HBB haplotype cluster and MYH9 SNP genotypes using STATA® SE-64 software program (version 14.0.370 for Windows). The parametric ANOVA (analysis of variance) test was performed to compare the mean values between the log(ACR) and MYH9 SNP genotypes.

General linear and multinomial regression frameworks, adjusted for age, and sex, using the statistical package R (version 3.0.3, The R foundation for statistical computing, Vienna, Austria) were employed for the analysis of combined \( \text{APOL1} \) G1/G2 variants, \( \text{HMOX1} \) polymorphisms, and the 3.7 kB alpha-globin gene deletion with the development of ESKD in the general Cameroonian cohort, and crude albuminuria, ACR and eGFR data in the SCD cohort. A genetic association case-control test was also performed to test the frequency of the various genotypes between the SCD patient cohort and the non-SCD controls, as well as between the ESKD cases and relative controls. For the ESKD cohort, the analysis was also split up for individuals with and without diabetic CKD, as the \( \text{APOL1} \) G1/G2 variants are associated with non-diabetic CKD (Limou et al., 2014), and association analysis was re-performed for all SNPs. To test for association between non-diabetic CKD status and each SNP, we fitted logistic regression models in which each SNP was presented as a predictor variable whose values
were equal to the number of copies of the minor allele (0, 1, 2) (i.e. additive model), the presence of at least one copy of the minor allele (0, 1) (i.e. dominant model) or the presence of two copies of the minor allele (0,1) (i.e. recessive model).

The significance level was set at 0.05.
CHAPTER 3: RESULTS

SECTION 1: SCD PATIENTS AND CONTROLS
3.1. POPULATION DESCRIPTION AND CLINICAL PHENOTYPES

3.1.1. DESCRIPTION OF SOCIODEMOGRAPHIC (AGE AND SEX) VARIABLES OF SCD PATIENTS AND CONTROLS.

Table 6 describes the socio-demographic variables of the SCD patients that were majority children and adolescents with a median age of 15 years (range: 2-58), with equal proportions of males and females. The non-SCD control group, including those with ESKD, were mainly adult patients with significantly increased median age values in comparison to the SCD patient cohort (p<0.0001). There was no significant difference in the age between males and female patients (p=0.5216) in the SCD cohort.

Table 6: Socio-demographic variables of the SCD patient cohort and the non-SCD control cohort.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Cohort</th>
<th>Group</th>
<th>Number (%)</th>
<th>Median (IQR)</th>
<th>Value range</th>
<th>Number of observations</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>SCD patients</td>
<td>Female</td>
<td>203 (49.15)</td>
<td>15 (9-23)</td>
<td>2-58</td>
<td>413</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td></td>
<td>Non-SCD controls</td>
<td>Female</td>
<td>26 (23-30)</td>
<td>26 (23-30)</td>
<td>4-51</td>
<td>70</td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td>SCD patients</td>
<td>Female</td>
<td>51 (71.83)</td>
<td>210 (50.85)</td>
<td>413</td>
<td>71</td>
<td>0.0005</td>
</tr>
<tr>
<td></td>
<td>Non-SCD controls</td>
<td>Male</td>
<td>20 (28.17)</td>
<td>20 (28.17)</td>
<td>70</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3.1.2. DESCRIPTION OF ANTHROPOMETRIC VARIABLES OF THE SCD COHORT
SCD patients presented with relatively low BMI, diastolic and systolic blood pressure values (Table 7).

Table 7: Anthropometric variables of the SCD patient cohort, including the BMI, systolic and diastolic blood pressure values of patients.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Median (IQR)</th>
<th>Value range</th>
<th>Number of observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>BMI (kg/m²)</td>
<td>17.361 (14.863-20.505)</td>
<td>10.61-31.029</td>
<td>413</td>
</tr>
<tr>
<td>Systolic blood pressure</td>
<td>107 (99-114)</td>
<td>77-156</td>
<td>413</td>
</tr>
<tr>
<td>Diastolic blood pressure</td>
<td>57 (53-63)</td>
<td>37-93</td>
<td>413</td>
</tr>
</tbody>
</table>

Figure 28 displays the distributions of age and BMI in the SCD patients’ cohort, while figure 29 displays the distributions of systolic and diastolic blood pressures, normalized using log transformation. Figure 30 displays the age distribution of the non-SCD control cohort. The age and BMI distributions were positively skewed in the SCD cohort, and age was negatively skewed in the non-SCD controls cohort and could therefore not be analyzed under the assumption of a normal underlying distribution. The skewed distributions informed the use of non-parametric methods to describe and analyze the data, therefore the median and interquartile range (IQR), were used to describe the socio-demographic variables of the SCD patient and non-SCD control cohort (Table 6).
Figure 28: Histogram distribution of variables age and BMI in the SCD patient cohort.

Histogram overlaid with the expected skewed distribution curve of variables (A) age and (B) BMI for the 413 patients included in the cohort.
Figure 29: Histogram distribution of variables SBP and DBP in the SCD patient cohort.

Histogram overlaid with the expected normal distribution curve of variables (A) log(SBP) and (B) log(DBP) for the 413 patients included in the cohort. These variables were normalized using log transformation, as confirmed by the Shapiro-Wilk test (p=0.6849 and p=0.6151, respectively).
3.1.3. HEMATOLOGICAL VARIABLES OF THE SCD COHORT

SCD patients presented with normocytic anemia with higher white blood cell count in comparison to the non-SCD controls (Table 8).

The Benin (73%; n=510) and Cameroon haplotypes (21%, n=144) were the most prevalent per number of chromosomes in the SCD patient cohort. In combination, the Benin/Cameroon haplotype represented 26.3% (n=92) and the Benin/Benin haplotype 56% (n=196) of the patient cohort.

Table 8 displays descriptive statistics for the hematological indices of the SCD patient and non-SCD control cohort for whom data was available. The non-parametric spearman’s correlation coefficient test was performed to determine the relation between the mean corpuscular volume (MCV) of red blood cells with albuminuria, while a Pearson’s correlation test was performed between leukocyte levels and albuminuria. Both MCV and leukocyte

Figure 30: Histogram distribution of age in the non-SCD patient cohort

Histogram overlaid with the expected skewed distribution curve of age and for the 101 patients included in the non-SCD control cohort.
count displayed a significant positive relationship with crude albuminuria ($r=0.1667$, $p=0.008$ and $r=0.1160$, $p=0.0202$, respectively) (figure 31), however no significant association with ACR was observed ($p=0.2511$ and $p=0.5025$, respectively).
Table 8: Hematological indices of the SCD patient and non-SCD control cohorts. The HBB haplotype (*) and frequency (%) (**) statistic relates to the HBB haplotype for the SCD cohort. (+) indicates the number of chromosomes analysed.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Cohort or HBB Haplotype*</th>
<th>Median (IQR) or Frequency (%)**</th>
<th>Value range</th>
<th>Number of observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hb (g/dL)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCD patients</td>
<td>7.5 (6.75-8.4)</td>
<td>3.5 - 3090.5</td>
<td>407</td>
<td></td>
</tr>
<tr>
<td>Non-SCD controls</td>
<td>13 (11.6-14.6)</td>
<td>8.1-18.7</td>
<td>71</td>
<td></td>
</tr>
<tr>
<td>MCV (fl)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCD patients</td>
<td>85 (78-92)</td>
<td>60-117</td>
<td>413</td>
<td></td>
</tr>
<tr>
<td>Non-SCD controls</td>
<td>81 (76.5-84)</td>
<td>56-102</td>
<td>71</td>
<td></td>
</tr>
<tr>
<td>MCHC (g/dL)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCD patients</td>
<td>33.5 (31.6-35.5)</td>
<td>21.5-54.3</td>
<td>413</td>
<td></td>
</tr>
<tr>
<td>Non-SCD controls</td>
<td>36 (32.3-37.7)</td>
<td>29.8-45.2</td>
<td>70</td>
<td></td>
</tr>
<tr>
<td>WBC (10^9/l)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCD patients</td>
<td>12.7 (10.2-16.2)</td>
<td>4-49.8</td>
<td>413</td>
<td></td>
</tr>
<tr>
<td>Non-SCD controls</td>
<td>5.03 (4.5-5.9)</td>
<td>2-10</td>
<td>71</td>
<td></td>
</tr>
<tr>
<td>Lymphocytes (10^9/l)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCD patients</td>
<td>5 (3.8 - 6.7)</td>
<td>1.4 - 22.1</td>
<td>406</td>
<td></td>
</tr>
<tr>
<td>Non-SCD controls</td>
<td>2.2 (1.8-2.6)</td>
<td>0.4-6.1</td>
<td>71</td>
<td></td>
</tr>
<tr>
<td>Monocytes (10^9/l)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCD patients</td>
<td>1.4 (0.9-2.1)</td>
<td>0.1-8.2</td>
<td>404</td>
<td></td>
</tr>
<tr>
<td>Non-SCD controls</td>
<td>0.6 (0.4-0.8)</td>
<td>0.2-94</td>
<td>71</td>
<td></td>
</tr>
<tr>
<td>Platelets (10^9/l)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCD patients</td>
<td>370 (284-466)</td>
<td>29 - 1078</td>
<td>402</td>
<td></td>
</tr>
<tr>
<td>Non-SCD controls</td>
<td>242 (181.5-290.5)</td>
<td>73-651</td>
<td>71</td>
<td></td>
</tr>
<tr>
<td>HbA2 (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCD patients</td>
<td>3.6 (2.85 - 4.1)</td>
<td>0-18.2</td>
<td>408</td>
<td></td>
</tr>
<tr>
<td>Non-SCD controls</td>
<td>3.2 (2-3.9)</td>
<td>0.1-5</td>
<td>66</td>
<td></td>
</tr>
<tr>
<td>HbF (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SCD patients</td>
<td>8.05 (2.15-13.35)</td>
<td>0-37.4</td>
<td>408</td>
<td></td>
</tr>
<tr>
<td>HBB Haplotype</td>
<td>Atypical</td>
<td>38 (5.43)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bantu/CAR</td>
<td>3 (0.43)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cameroon</td>
<td>145 (20.71)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Indian-Arab</td>
<td>3 (0.43)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Senegal</td>
<td>1 (0.14)</td>
<td>700+</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Benin</td>
<td>510 (72.86)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 31: Scatter plots illustrating the relationship between hematological indices and crude albuminuria in the SCD cohort, with a significant positive relationship observed between (A) MCV ($r=0.1667$, $p=0.008$) and (B) log(Leukocytes) ($r=0.1160$, $p=0.0202$) with crude albuminuria. The (A) MCV and (B) log(Leukocyte) variables are displayed on the x-axis for the respective graphs, with crude albuminuria values on the y-axis. The red line indicates a line of best fit, fitted to the data.
3.1.4. DESCRIPTION OF KIDNEY FUNCTIONS OF THE SCD COHORT

3.1.4.1. ESTIMATED GLOMERULAR FILTRATION RATE (eGFR)

Table 9 describes the eGFR values calculated for the cohort using the MDRD, CG, and CKD-EPI equations, with these equations indicating that 52.9% (n=214), 42.57% (n=172) and 49.5% (n=200) of patients suffered from glomerular hyperfiltration, defined as a filtration rate >130 ml/min/1.73m² and >140 ml/min/1.73m² in women and men, respectively. Figure 32 displays the distribution of eGFR values calculated using the various equations. All three variables were not normally distributed despite log transformation (Shapiro-Wilkes test p-value<0.05). This informed the use of non-parametric statistical analysis, with figure 33 displaying the median eGFR values, by equation, for various age groups in the SCD cohort. Albuminuria appears to decrease, while glomerular filtration rate appears to increase in the oldest two age groups for all equations, however this is likely due to the small size of these cohorts (n=6 and n=4, respectively). The prevalence of hyperfiltration, as determined using the CKD-EPI equation for its increased accuracy in SCD patients (Arlet et al., 2012), per age group of SCD patients is displayed in figure 34. The high prevalence in the age group 44-49 years old was due to one individual in a small cohort (n=6) displaying hyperfiltration.

**Table 9**: Estimated Glomerular Filtration Rates calculated using the MDRD, CG and CKD-API equations in the SCD patient cohort. Glomerular hyperfiltration is defined as a filtration rate greater than 130 ml/min/1.73m² in women and 140 ml/min/1.73m² in men.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Equation</th>
<th>Median (IQR)</th>
<th>Value range</th>
<th>Glomerular Hyperfiltration (%)</th>
<th>Number of observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>eGFR (ml/min)</td>
<td>MDRD</td>
<td>140.5 (101.4-205.5)</td>
<td>44.6-779.2</td>
<td>214 (52.97)</td>
<td>404</td>
</tr>
<tr>
<td></td>
<td>Cockcroft-Gault</td>
<td>111.8 (90.5-162)</td>
<td>55.1-368.7</td>
<td>172 (42.57)</td>
<td>404</td>
</tr>
<tr>
<td></td>
<td>CKD-EPI</td>
<td>135.1 (112-154.4)</td>
<td>50.8-250.8</td>
<td>200 (49.5)</td>
<td>404</td>
</tr>
</tbody>
</table>
Figure 32: Histogram distribution of eGFR variables in the SCD cohort. Histogram overlaid with the expected skewed distribution curve of eGFR values calculated using the (A) MDRD, (B) Cockcroft-Gault and (C) CKD-EPI equations the 413 patients included in the SCD patient cohort.
Figure 33: eGFR estimates per age group for SCD patients.

eGFR values calculated using (A) MDRD, (B) Cockcroft-Gault and (C) CKD-EPI equations for various age groups of patients with SCD. The eGFR values are increased in children and adolescents in comparison to adults. The values represent the median GFR per age group, while the error bars represent 1 average deviation from the median.
The prevalence of micro-albuminuria in the SCD patient population was 60.9% (n=248), with macro-albuminuria seen in 2.5% (n=10) of the cohort (Figure 35). The prevalence of micro- and macro-albuminuria, as defined by the ACR value, was 73.4% (n=116) and 5.1% (n=8), respectively. The youngest patient to display micro-albuminuria as defined using crude albuminuria was four years old. Based on the CKD-EPI equation, 34.2% (n=67) of patients developed hyperfiltration alone, while 61.7% (n=121) and 4.1% (n=8) displayed associated micro- and macro-albuminuria, respectively.

**Figure 34:** Histogram distribution of the prevalence of hyperfiltration in various age groups in the SCD cohort. The prevalence of hyperfiltration, defined as filtration values >130 ml/min/1.73m² and >140 ml/min/1.73m² in women and men, respectively, calculated using the CKD-EPI equation in various age groups of the SCD cohort.

### 3.1.4.2. MICRO-ALBUMINURIA
The crude albuminuria distribution for the SCD cohort was positively skewed (Figure 36) and not normally distributed, despite log transformation (Shapiro-Wilkes test for normality; p<0.00001). The non-parametric Kruskal Wallis and spearman’s correlation test were therefore employed to determine the relationship between albuminuria and categorical and numerical variables, respectively.

The distribution of the ACR data was positively skewed (Figure 37A), and subsequently normalized using log transformation (Shapiro-Wilkes test for normality; p=0.152) (Figure 37B). Parametric methods of analysis were therefore employed for statistical analysis.

Figure 35: Prevalence of crude albuminuria and ACR in the SCD patient cohort.
Pie chart displaying the prevalence of normal albuminuria, micro-albuminuria and macro-albuminuria as determined using (A) crude albuminuria values (n=407) and (B) the Albumin-to-Creatinine ratio (ACR) (n=158) in the SCD patient cohort.
Figure 36: Histogram distribution of micro-albuminuria
Histogram of the distribution of micro-albuminuria in the general patient cohort, overlaid with the skewed distribution of the variable. The distribution is positively skewed and subsequently not normally distributed.
Figure 38 displays a graphical representation of the median albuminuria values per age group in the SCD cohort, with figure 39 displaying the prevalence of micro- and macro-albuminuria per age group. Albuminuria was reduced in older patients due to the low number of individuals in the three oldest age group (n=7, n=6 and n=5, respectively). This is also likely due to the fact that the patients have survived to this age. All cohorts displayed a high prevalence of micro-albuminuria (figure 39)
Figure 38: Albuminuria estimates per age group for SCD patients.
The median albuminuria values per age group for SCD patients. There is a general positive association between crude albuminuria and increasing age until the fourth decade. The values represent the median albuminuria values per age group, while the error bars represent 1 average deviation from the median.

Figure 39: Histogram distribution of the prevalence of micro-albuminuria in various age groups in the SCD cohort. The prevalence of micro-albuminuria, defined as albuminuria values of 30-300 mg/ml in various age groups of the SCD cohort.
3.1.4.3. SUMMARY OF KIDNEY FUNCTIONS IN THE SCD COHORT

Table 10 describes the summary of the various kidney functions measured in the SCD patients’ cohort, namely the ACR, crude albuminuria, serum creatinine, and eGFR values calculated using the three equations. The median and IQR statistical measures are used to describe the various functional renal measurements in the SCD cohort as all variables are not normally distributed.

**Table 10:** Summary of the kidney functioning measurements of the SCD patient cohort.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Median (IQR)</th>
<th>Value range</th>
<th>Number of observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Albumin-to-creatinine ratio (ACR) (mg/mmol)</td>
<td>7.16 (3.4-12.5)</td>
<td>0.4-236.5</td>
<td>152</td>
</tr>
<tr>
<td>Crude albuminuria (mg/ml)</td>
<td>41 (23-83)</td>
<td>3-1180</td>
<td>407</td>
</tr>
<tr>
<td>eGFR (MDRD) (ml/min/1.73m²)</td>
<td>140.5 (101.4-205.5)</td>
<td>44.6-779.3</td>
<td>404</td>
</tr>
<tr>
<td>eGFR (CKD-EPI) (ml/min/1.73m²)</td>
<td>135.1 (112-154.4)</td>
<td>50.8-250.8</td>
<td>404</td>
</tr>
<tr>
<td>eGFR (Cockcroft-Gault) (ml/min/1.73m²)</td>
<td>123.5 (97.4-167.2)</td>
<td>47.9-404.6</td>
<td>404</td>
</tr>
<tr>
<td>Serum creatinine (µmol/L)</td>
<td>7 (5-8.5)</td>
<td>2-13.8</td>
<td>404</td>
</tr>
</tbody>
</table>

The SCD patient cohort displayed median micro-albuminuric values based on crude albuminuria (30-300 mg/ml) and ACR data (3-30 mg/mmol), respectively. The median eGFR values were close to the hyperfiltration definition (>130 ml/min/1.73m² in women and >140 ml/min/1.73m² in men) for all three equations.

3.1.5. CLINICAL ASSOCIATIONS IN THE SCD COHORT

Association between eGFR and blood pressure

The log transformed variables of systolic blood pressure and diastolic blood pressure followed a normal distribution and could therefore be analysed using Pearson’s correlation coefficient. A significantly negative association was observed between eGFR and log(SBP) and log(DBP) values, respectively, for the MDRD (p<0.0001 and p<0.0001), CKD-EPI (p<0.0001 and p<0.0001) and Cockcroft-Gault equations (p=0.0001 and p=0.001) (Figures 40 and 41)
Figure 40: Scatter plots illustrating the relationship between eGFR values and logSBP in the SCD patient cohort, calculated using the (A) MDRD (r=-0.4071, p<0.0001) (B) CKD-EPI (r=-0.3674, p<0.0001) and (C) Cockcroft-Gault (r=p=0.0182) equations. The log(SBP) variable is displayed on the x-axis for all three graphs, with the eGFR values on the y-axis. The red line indicates a line of best fit, fitted to the data.
Figure 41: Scatter plots illustrating the relationship between eGFR values and logDBP in the SCD patient cohort, calculated using the (A) MDRD ($r=-0.3004, p<0.0001$) (B) CKD-EPI ($r=-0.2956, p<0.0001$) and (C) Cockcroft-Gault ($r=-0.1644, p=0.001$) equations. The log(DBP)variable is displayed on the x-axis for all three graphs, with the eGFR values on the y-axis. The red line indicates a line of best fit, fitted to the data.
Association between eGFR and albuminuria

Neither the albuminuria nor eGFR variables were normally distributed, despite log transformation. The non-parametric spearman’s coefficient was therefore calculated to determine the association between these variables. There was no significant association between albuminuria and eGFR values, as calculated using the MDRD (p=0.9159), CKD-EPI (p=0.8706) and Cockcroft-Gault equations (p=0.1935) (Figure 42).
Figure 42: Scatter plots illustrating the relationship between albuminuria values and eGFR in the SCD patient cohort, calculated using the (A) MDRD (p=0.9159), (B) CKD-EPI (p=0.8706) and (C) Cockcroft-Gault (p=0.1935) equations. The eGFR values are displayed on the x-axis for all three graphs, with the albuminuria values on the y-axis. No significant association was observed between albuminuria and eGFR for all three equations. The red line indicates a line of best fit, fitted to the data.
A significant association was observed between the log(ACR) and eGFR values using the Pearson’s correlation coefficient, as log transformation of the ACR normalized the distribution of the variable. The eGFR values were significantly positively associated with log(ACR) in the SCD cohort (p=0.0059 for MDRD equation, p=0.0057 for CKD-EPI equation and p=0.0042 for the Cockcroft-Gault method) (figure 43).
Figure 43: Scatter plots illustrating the relationship between eGFR values and log(ACR) in the SCD patient cohort, calculated using the (A) MDRD (r=0.2232, p=0.0059), (B) CKD-EPI (r=0.02238, p=0.0057) and (C) Cockcroft-Gault (r=0.2317, p=0.0042) equations. The log(ACR) variable is displayed on the x-axis for all three graphs, with the eGFR values on the y-axis. The red line indicates a line of best fit, fitted to the data.
Association between eGFR and sex, age and BMI

The eGFR, age and BMI variables were not normally distributed, therefore Spearman's rank correlation coefficient was performed to determine the significance of the association between these variables. The non-parametric Wilcoxon Rank sum test was performed to test for the equality of eGFR values for all three equations between males and females.

BMI was significantly negatively associated with eGFR values for the MDRD (p<0.0001), CKD-EPI (p<0.0001) and Cockcroft-Gault equations (p=0.0182) (Figure 44). This relationship was also observed between age and eGFR values (p<0.0001, p<0.0001 and p=0.0009, respectively) (figure 45). Results of the Wilcoxon Rank sum test indicated that eGFR values were significantly increased in male patients compared to females (p<0.0001 for all three equations) (figure 46).
Figure 44: Scatter plots illustrating the relationship between eGFR values and BMI in the SCD patient cohort, calculated using the (A) MDRD ($r=-0.52$, $p<0.0001$) (B) CKD-EPI ($r=0.529$, $p<0.0001$) and (C) Cockcroft-Gault ($r=-0.164$, $p=0.0182$) equations. The BMI (kg/m$^2$) variable is displayed on the x-axis for all three graphs, with the eGFR values on the y-axis. The red line indicates a line of best fit, fitted to the data.
Figure 45: Scatter plots illustrating the relationship between eGFR values and age in the SCD patient cohort, calculated using the (A) MDRD ($r=0.5233$, $p<0.0001$) (B) CKD-EPI ($r=0.5501$, $p<0.0001$) and (C) Cockcroft-Gault ($r=0.1684$, $p<0.0001$) equations. Age is displayed on the x-axis for all three graphs, with the eGFR values on the y-axis. The red line indicates a line of best fit, fitted to the data.
Figure 46: Box and whisker plots showing the association of eGFR values with sex in the SCD cohort.

Box and whisker plots illustrating the distribution of eGFR values conditioned on sex for the entire patient cohort. The eGFR values were calculated using (A) MDRD equations, (B) CKD-EPI equation and (C) Cockcroft-Gault equations. The sex group is indicated on the x-axis of each graph, including females (F) and males (M). The horizontal lines that constitute the ‘box’ correspond to the lower quartile, median, and upper quartile parameters. The length of the ‘whiskers’ that extend from the box in the upwards and downwards direction represent a distance 1.5 times the interquartile range. Values that lie outside this distance are considered outliers, or extreme values in the data, and are indicated by diamonds. Significant results are indicated using * (p<0.05).
Association between crude albuminuria, sex and age among SCD Patients

Analytical statistics were performed for all quantitative data using STATA® SE-64 software program (version 14.0.370 for Windows). Spearman’s correlation test was performed to investigate the significance of the relationships between crude albuminuria, and age and hemoglobin levels in the SCD patient cohort. A Kruskal-Wallis test was employed to determine the significance of the relationship between crude albuminuria and sex. Age was significantly associated with increased crude albuminuria in SCD patient cohort \( (r=0.1421, \ p = 0.0041) \) (Figure 47). The relationship of sex with crude albuminuria was not significant \( (p=0.184) \). Conversely, the relationship between crude albuminuria and lower hemoglobin tended to a significant in the SCD patient population \( (p=0.0691) \) (Figure 48).
Figure 47: Scatter plot illustrating the positive relationship between age and crude albuminuria in the SCD patient cohort, with age displayed on the x-axis and crude albuminuria data on the y-axis. The red line indicates a line of best fit, fitted to the data, indicating the slightly positive nature of the relationship \(r=0.1421, \ p=0.004\).

Figure 48: Scatter plot illustrating the negative relationship between albuminuria and hemoglobin in the SCD patient cohort, with hemoglobin displayed on the x-axis and crude albuminuria data on the y-axis. The red line indicates a line of best fit, fitted to the data, indicating the near significant negative nature of the relationship \(r=-0.0909, \ p=0.0691\).
**3.1.6. SUMMARY OF CLINICAL PHENOTYPE ASSOCIATIONS**

Tables 11 and 12 summarise the factors associated with crude albuminuria and eGFR development, respectively, in Cameroonian HbSS patients.

**Table 11:** Factors affecting crude albuminuria levels in Cameroonian SCD patients.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Co-efficient</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.1241</td>
<td>0.0041</td>
</tr>
<tr>
<td>Hemoglobin (g/dl)</td>
<td>-0.091</td>
<td>0.0289</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>0.0977</td>
<td>0.0489</td>
</tr>
<tr>
<td>Leukocytes (10⁹/l)</td>
<td>0.1160</td>
<td>0.0202</td>
</tr>
<tr>
<td>Mean arterial pressure (mm Hg)</td>
<td>0.1181</td>
<td>0.0175</td>
</tr>
<tr>
<td>MCV (fl)</td>
<td>0.1667</td>
<td>0.008</td>
</tr>
</tbody>
</table>

**Table 12:** Factors affecting eGFR levels in Cameroonian SCD patients for the MDRD, CKD-EPI and Cockcroft Gault equations. The relative equation is indicated in brackets.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Co-efficient</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>-0.523 (MDRD)</td>
<td>&lt;0.0001 (MDRD)</td>
</tr>
<tr>
<td></td>
<td>-0.550 (CKD-EPI)</td>
<td>&lt;0.0001 (CKD-EPI)</td>
</tr>
<tr>
<td></td>
<td>-0.165 (Cockcroft-Gault)</td>
<td>0.0009 (Cockcroft-Gault)</td>
</tr>
<tr>
<td>ACR (mg/mmol)</td>
<td>0.2232 (MDRD)</td>
<td>0.0059 (MDRD)</td>
</tr>
<tr>
<td></td>
<td>0.2238 (CKD-EPI)</td>
<td>0.0057 (CKD-EPI)</td>
</tr>
<tr>
<td></td>
<td>0.2317 (Cockcroft-Gault)</td>
<td>0.0042 (Cockcroft-Gault)</td>
</tr>
<tr>
<td>Male sex</td>
<td>risk</td>
<td>&lt;0.0001 (all equations)</td>
</tr>
<tr>
<td>SBP (mm Hg)</td>
<td>-0.407 (MDRD)</td>
<td>&lt;0.0001 (MDRD)</td>
</tr>
<tr>
<td></td>
<td>-0.367 (CKD-EPI)</td>
<td>&lt;0.0001 (CKD-EPI)</td>
</tr>
<tr>
<td></td>
<td>-0.186 (Cockcroft-Gault)</td>
<td>0.0001 (Cockcroft-Gault)</td>
</tr>
<tr>
<td>DBP (mm Hg)</td>
<td>-0.3 (MDRD)</td>
<td>&lt;0.0001 (MDRD)</td>
</tr>
<tr>
<td></td>
<td>-0.296 (CKD-EPI)</td>
<td>&lt;0.0001 (CKD-EPI)</td>
</tr>
<tr>
<td></td>
<td>-0.164 (Cockcroft-Gault)</td>
<td>0.0001 (Cockcroft-Gault)</td>
</tr>
<tr>
<td>Hemoglobin (g/dl)</td>
<td>-0.128 (MDRD)</td>
<td>0.0106 (MDRD)</td>
</tr>
<tr>
<td></td>
<td>-0.142 (CKD-EPI)</td>
<td>0.0046 (CKD-EPI)</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>-0.52 (MDRD)</td>
<td>&lt;0.0001 (MDRD)</td>
</tr>
<tr>
<td></td>
<td>-0.529 (CKD-EPI)</td>
<td>&lt;0.0001 (CKD-EPI)</td>
</tr>
<tr>
<td></td>
<td>-0.118 (Cockcroft-Gault)</td>
<td>0.0182 (Cockcroft-Gault)</td>
</tr>
<tr>
<td>Serum Creatinine (mg/dl)</td>
<td>-0.915 (MDRD)</td>
<td>&lt;0.0001 (all equations)</td>
</tr>
<tr>
<td></td>
<td>-0.901 (CKD-EPI)</td>
<td>&lt;0.0001 (MDRD)</td>
</tr>
<tr>
<td></td>
<td>-0.931 (Cockcroft-Gault)</td>
<td>&lt;0.0001 (CKD-EPI)</td>
</tr>
<tr>
<td>Mean arterial pressure (mm Hg)</td>
<td>-0.321 (MDRD)</td>
<td>&lt;0.0001 (all equations)</td>
</tr>
<tr>
<td></td>
<td>-0.337 (CKD-EPI)</td>
<td>&lt;0.0001 (MDRD)</td>
</tr>
</tbody>
</table>
3.2. DESCRIPTION OF GENETIC VARIANTS

The SCD patients’ and non-SCD controls’ cohorts were previously genotyped for the 3.7 kB α-globin gene deletion. In addition, the SCD cohort was genotyped for HBB haplotypes. In the present study, the SCD patient cohort was genotyped for four targeted SNPs in MYH9 (table 13) and along with the non-SCD control cohort, genotyped for rs3074372 (HMOX1) and rs71785313 (APOL1) (figure 49) using fragment analysis, and for rs733885319 (APOL1), rs60910145 (APOL1) and rs743811 (HMOX1) (figure 50) using TaqMan® SNP genotyping (Table 14). The results of genotyping of targeted variants in MYH9 for SCD patients are displayed in table 13, while overall frequency of the 3.7 kB α-globin gene deletion and relevant genotypes for each of the chosen SNPs in APOL1 and HMOX1 of the SCD patient and non-SCD control cohorts are displayed in Table 14.

Figure 49: Fragment Analysis of rs3074372 (HMOX1) and rs71785313 (APOL1) polymorphisms for SCD patient 102.

Electropherogram produced using the ABI Prism 3130xl Genetic Analyzer (Applied Biosystems) for the genotyping of patient 102 for polymorphisms (A) rs3074372 (HMOX1), and (B) rs71785313 (APOL1). Dye-labelled GeneScan™ 500 ROX was used as the size standard for the reaction. The analysis of each genotype was performed using GeneMapper® Software version 4.0. The genotype results showed (A) heterozygosity for fragment sizes of 127 bp and 151 bp, and (B) homozygosity for the wild type APOL1 G2 allele. The fluorescent dyes (A) 6-FAM™ and (B) HEX™ were used for labelling of the forward primers of rs3074372 (HMOX1) and rs71785313 (APOL1), respectively.
Figure 50: TaqMan SNP Genotyping plot of SCD cohort plate 3 for rs743811 (HMOX1).

TaqMan allelic discrimination plot produced using the Bio-Rad CFX96 (Bio-Rad laboratories, Hercules, CA, USA) for the genotyping of plate 1 SCD cohort patients for rs743811 (HMOX1). The x-axis corresponds to the relative fluorescent units (RFU) for the FAM 5’-dye label, and the y-axis corresponds to the relative fluorescent units (RFU) for the VIC 5’-dye label. The analysis of each genotype was performed using Bio-Rad CFX96 (Bio-Rad laboratories, Hercules, CA, USA) Manager Software version 3.1. The FAM and VIC 5’-dyes correspond to the major (T) and minor (C) alleles, respectively. The blue squares therefore represent patients homozygous for the minor allele (i.e. CC), the green triangles represent individuals heterozygous for rs43811 (i.e. TC), and the orange circles represent patients homozygous for the major allele (i.e. TT). The black diamond represents the NTC, validating that no contamination occurred.
The results of the chi-squared test (df = 1) performed to detect deviation from the HWE is displayed in both tables, with table 10 also containing results of the chi-squared test (df = 1) of significant deviations in SNP genotype frequencies between the SCD patient and non-SCD control cohorts. No p-value for rs73885319 (APOL1), rs60910145 (APOL1) and rs71785313 (APOL1) was calculated as these were analysed as the combined APOL1 G1/G2 variant. The results indicate that a much higher proportion of SCD patients had co-inherited alpha-thalassemia’s than controls (41.4 % (n=141) vs. 19 % (n=11); p=0.0028) (Figure 51). There was also a significant difference in variant frequency in APOL1 G1/G2 between SCD patients and controls (p=0.0097) (figure 52).

Table 13: The frequency of genotypes of rs11912763, rs16996648, rs1005570, rs16996672 in MYH9 and HBB haplotypes within the SCD patient cohort. The p-value obtained from the chi-squared test for deviation from Hardy-Weinberg equilibrium is displayed for each SNP.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Frequency (%)</th>
<th>HWE (p-value)</th>
<th>Number of observations</th>
</tr>
</thead>
<tbody>
<tr>
<td>MYH9 - rs11912763 (G/A)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>305 (78.01)</td>
<td>0.924</td>
<td>391</td>
</tr>
<tr>
<td>AG</td>
<td>80 (20.46)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>6 (1.53)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MYH9 - rs16996648 (T/C)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TT</td>
<td>207 (56.25)</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>CT</td>
<td>138 (37.5)</td>
<td></td>
<td>368</td>
</tr>
<tr>
<td>CC</td>
<td>23 (6.25)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MYH9 - rs1005570 (G/A)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>147 (38.79)</td>
<td>0.992</td>
<td>379</td>
</tr>
<tr>
<td>AG</td>
<td>178 (46.97)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>54 (14.25)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MYH9 - rs16996672 (C/T)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC</td>
<td>211 (54.38)</td>
<td>0.122</td>
<td>388</td>
</tr>
<tr>
<td>CT</td>
<td>158 (40.72)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TT</td>
<td>19 (4.9)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 14: The frequency of α-globin gene genotypes, genotypes of rs3074372, rs743811, rs73885319, rs60910145, rs71785313 and APOL1 G1/G2 alleles within the SCD patient and non-SCD control cohort. The p-value obtained from the chi-squared test for deviation from Hardy-Weinberg equilibrium is displayed for each SNP, as well as the p-value obtained from the chi-squared test for significance of genotype frequency difference between cases and controls. Significant results * comparing allele frequency between SCD patients and controls; (p<0.05) and those approaching significance (p<0.05) are highlighted in bold.

<table>
<thead>
<tr>
<th>Variables</th>
<th>SCD Patient cohort</th>
<th></th>
<th>Non-SCD control cohort</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Freq (%)</td>
<td>HWE (p-value)</td>
<td>Number of observations</td>
<td>Freq (%)</td>
</tr>
<tr>
<td>3.7del α-globin gene genotypes</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>αα/αα</td>
<td>198 (58.58)</td>
<td></td>
<td></td>
<td>47 (81.03)</td>
</tr>
<tr>
<td>αα/α3.7</td>
<td>105 (31.07)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>α3.7/α3.7</td>
<td>35 (10.36)</td>
<td></td>
<td></td>
<td>3 (5.17)</td>
</tr>
<tr>
<td>HMOX1 – rs3074372 (S/L)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S/S</td>
<td>2 (0.54)</td>
<td>0.67</td>
<td>368</td>
<td>3 (3.49)</td>
</tr>
<tr>
<td>S/L</td>
<td>45 (12.23)</td>
<td></td>
<td></td>
<td>11 (12.79)</td>
</tr>
<tr>
<td>L/L</td>
<td>321 (87.23)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HMOX1 – rs743811 (T/C)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TT</td>
<td>290 (78.59)</td>
<td>0.598</td>
<td>369</td>
<td>77 (79.38)</td>
</tr>
<tr>
<td>TC</td>
<td>76 (20.60)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC</td>
<td>3 (0.81)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>APOL1 – rs73885319 (A/G)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>285 (75.6)</td>
<td>1</td>
<td>377</td>
<td>82 (81.19)</td>
</tr>
<tr>
<td>AG</td>
<td>86 (22.81)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>6 (1.59)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>APOL1 – rs60910145 (T/G)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TT</td>
<td>278 (76.8)</td>
<td>0.81</td>
<td>362</td>
<td>82 (82)</td>
</tr>
<tr>
<td>TG</td>
<td>78 (21.55)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>6 (1.66)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>APOL1 – rs71785313 (WT/del)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WT/WT</td>
<td>318 (84.35)</td>
<td>0.73</td>
<td>377</td>
<td>72 (78.26)</td>
</tr>
<tr>
<td>Del/WT</td>
<td>56 (12.85)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Del/del</td>
<td>3 (0.8)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>APOL1 – G1/G2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G/G</td>
<td>333 (88.1)</td>
<td>p=0.005</td>
<td>378</td>
<td>72 (78.26)</td>
</tr>
<tr>
<td>G/A</td>
<td>40 (10.58)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A/A</td>
<td>5 (1.32)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

There was no deviation from HWE (p<0.05) for each of the MYH9 SNPs, however the APOL1 G1/G2 allele displayed significant deviation from HWE (p=0.005) in the SCD patient cohort, as did HMOX1 rs3074372 in the non-SCD control cohort (p=0.03). The 3.7del α-globin gene genotypes showed significant deviation from the equilibrium in both the SCD (p=0.0005) and non-SCD control (p=0.007) cohorts.
Figure 51: Co-inheritance of alpha-thalassemia among SCD patients and non-SCD controls. The SCD patients display a much higher prevalence of co-inherited 3.7 kb α-globin gene deletion than the non-SCD controls (p=0.0028).

Figure 52: Prevalence of APOL1 G1/G2 variant among SCD patients and non-SCD controls. The SCD patients display a prevalence of 2 minor alleles for the APOL1 G1/G2 variant in comparison to the controls (p=0.0097).
The distribution of the *HMOX1* dinucleotide GT promoter repeats for the SCD patient, non-SCD controls and non-SCD ESKD cohort are displayed in figure 53. The distribution of repeats followed a trimodal pattern.

**Figure 53: Histogram of distribution of *HMOX1* rs3074372 (GT)_n promoter repeats** in the (A) SCD patient and (B) non-SCD controls, with short repeats (<= 25) shown in red, and the long repeats (>25) shown in blue.

### 3.2.1. COMPARATIVE MINOR ALLELE FREQUENCY (MAF) OF TARGETED SNPS WITH VARIOUS WORLD POPULATIONS USING DATA EXTRACTED FROM THE 1000 GENOME PROJECT

The calculated minor allele frequency (MAF) for each of the SNPs was compared to frequencies in other African populations, as well as populations from Europe and East Asia.
(Table 15). These MAFs were obtained from the 1000 Genome Project Data published in the Ensembl Genome browser website (http://www.ensembl.org/index.html).

**Table 15**: The minor allele frequencies (MAF) of rs11912763, rs16996648, rs1005570, rs16996672 in the patient cohort (Cameroon) compared to various African populations, Africa, Europe, East Asia and South Asia.

<table>
<thead>
<tr>
<th>Region</th>
<th>rs11912763</th>
<th>rs16996648</th>
<th>rs1005570</th>
<th>rs16996672</th>
<th>rs743811</th>
<th>rs73885319</th>
<th>rs60910145</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCD patient cohort (Cameroon)</td>
<td>0.119</td>
<td>0.249</td>
<td>0.373</td>
<td>0.255</td>
<td>0.111</td>
<td>0.130</td>
<td>0.124</td>
</tr>
<tr>
<td>Non-SCD control cohort (Cameroon)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.113</td>
<td>0.010</td>
<td>0.090</td>
</tr>
<tr>
<td>ESKD case cohort (Cameroon)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.146</td>
<td>0.146</td>
<td>0.116</td>
</tr>
<tr>
<td>ESKD control cohort (Cameroon)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.122</td>
<td>0.106</td>
<td>0.084</td>
</tr>
<tr>
<td>African American (Southwest US)</td>
<td>0.156</td>
<td>0.23</td>
<td>0.385</td>
<td>0.262</td>
<td>0.156</td>
<td>0.221</td>
<td>0.221</td>
</tr>
<tr>
<td>Esan (Nigeria)</td>
<td>0.399</td>
<td>0.49</td>
<td>0.591</td>
<td>0.419</td>
<td>0.162</td>
<td>0.495</td>
<td>0.495</td>
</tr>
<tr>
<td>Yoruba (Kenya)</td>
<td>0.324</td>
<td>0.407</td>
<td>0.5</td>
<td>0.37</td>
<td>0.120</td>
<td>0.375</td>
<td>0.375</td>
</tr>
<tr>
<td>Europe</td>
<td>0.001</td>
<td>0.021</td>
<td>0.101</td>
<td>0.002</td>
<td>0.253</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>East Asia</td>
<td>0</td>
<td>0</td>
<td>0.067</td>
<td>0</td>
<td>0.664</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
The MAF of SNPs in MYH9 (Figure 54) rs11912763 (0.118), rs16996648 (0.249), rs1005570 (0.373) and rs16996672 (0.254) were similar to that of the African American population. The MAF for rs11912763 in the Yoruba (0.324) and Esan (0.399) populations were approximately three times greater than that of the Cameroonian cohort, and there was no similarity between the MAF of each of the SNPs in the Cameroonian cohort compared to the European and Asian populations. This highlights the great variation seen between individual African regions.

Figure 54: Comparison of MAF of MYH9 SNPs between various populations.
Bar graph comparing the minor allele frequencies (MAF) of MYH9 SNPs rs11912763, rs16996648, rs1005570, rs16996672 in the SCD patient cohort (Cameroon) compared to various African populations, Europe and East Asia. These MAFs were obtained from the 1000 Genome Project Data published in the Ensembl Genome browser website (http://www.ensembl.org/index.html).
The MAF of SNPs rs743811 (*HMOX1*), and rs60910145 (*APOL1*) were similar between all of our Cameroonian cohorts. The MAF of rs73885319 (*APOL1*) was similar in the SCD patient cohort (0.13), non-SCD ESKD case cohort (0.146) and the non-SCD ESKD control cohort (0.106), however differed slightly to the non-SCD control cohort (0.01). For rs743811 (*HMOX1*), the populations of African descent displayed a MAF approximately six times less than that seen in East Asian populations.

The MAF for rs73885319 (*APOL1*) and rs60910145 (*APOL1*) was approximately three times greater in the Yoruba (0.375 and 0.375, respectively) and Esan populations (0.495 and 0.495, respectively), in comparison to our Cameroonian cohorts. While there was no prevalence of the SNPs reported among the European (0) and Asian (0) populations.
No MAF data is available for the *HMOX1* rs3074372 and *APOL1* G2 (rs71785313) variants in various world populations. Figures 54 and 55 therefore display the frequency of the genotypes for the respective variants in our Cameroonian SCD patients, non-SCD controls, non-SCD ESKD cases and non-SCD ESKD controls. The S/S genotype for *HMOX1* rs3074372 was more common in non-SCD and non-SCD ESKD controls, while homozygotes for the 6 bp deletion in *APOL1* (rs71785313) were most common in the non-SCD ESKD case cohort.

![Figure 56: Comparison of genotype frequencies of *HMOX1* rs3074372 between our Cameroonian cohorts.](image)

Bar graph comparing the genotype frequencies for the dinucleotide repeat in the promoter of *HMOX1* (rs3074372) on the basis of short ( </=25 repeats) and long (>25 repeats) classification of the number of repeats between our Cameroonian cohorts; SCD patients, non-SCD controls, non-SCD ESKD cases and non-SCD ESKD controls.
3.2.2. GENOTYPE-TO-PHENOTYPE ASSOCIATION STUDIES

3.2.2.1. Association between HBB haplotype and crude albuminuria in the SCD patient cohort.

The Kruskal Wallis was employed to determine the relationship between the HBB haplotype cluster and crude albuminuria. No significant relationship was found in the SCD patient cohort (p=0.7534).

3.2.2.2. Association between 3.7 kb α-globin gene deletion and crude albuminuria in the SCD patient cohort.

The Kruskal-Wallis test was also performed to determine the nature of the relationship between the 3.7kB α-globin gene deletion genotype and crude albuminuria, also displaying a non-significant relationship in the SCD patient cohort (p=0.541).
3.2.2.3. Association between four targeted SNPs in MYH9 and crude albuminuria, and ACR in the SCD patient cohort.

A Kruskal-Wallis equality-of-populations rank test was performed for each of the four MYH9 SNPs with crude albuminuria in the SCD patient cohort. There was no significant difference (p<0.05) in the medians of each of the SNP genotype’s albuminuria values, therefore none of the MYH9 SNPs were found to be significantly associated with crude albuminuria in our SCD patient cohort.

A one-way ANOVA was performed to establish the significance of each of the SNPs genotypes with their corresponding ACR values the SCD patient cohort. No significant results were obtained (p>0.05). Figure 58 illustrates the distribution of ACR values conditioned on the SNP genotypes in the SCD patient cohort. ANOVA tests were also performed to determine the relationship between the ACR values and the 3.7kB α-globin gene deletion genotypes. The results were not significant for the SDC patient cohort (p=0.2243).
3.2.2.4. Association between targeted polymorphisms in *HMOX1* and *APOL1* with select phenotypes of kidney disease; eGFR, crude albuminuria, macro-albuminuria, and ACR.

Statistical analysis for the combined *APOL1* G1/G2 alleles, rs3074372 and rs743811, as well as the 3.7 kB alpha-thalassemia deletion was performed using the statistical package, R.
Association tests, adjusted for sex and age, were performed to determine the significance of the association between the target SNPs and our chosen phenotypes for kidney disease in the SCD patient cohort; crude albuminuria, macro-albuminuria, ACR, and the estimated Glomerular Filtration Rate (eGFR), calculated using three equations; CKD-EPI, MDRD and CGN. Further associations were performed to test the association of these SNPs with the development of End Stage Kidney Disease in the non-SCD ESKD case and control cohort. The resulting p-values obtained from association analysis between target SNPs and the phenotype of choice in the SCD cohort are displayed in table 16.
Table 16: P-values obtained from association analysis between the corresponding target SNPs and select kidney disease phenotypes for the SCD patient cohort, using the statistical package ‘R’, with the relevant genetic model indicated in brackets. Statistically significant results (p<0.05), or results approaching significance (p≥0.05) are highlighted in bold.

<table>
<thead>
<tr>
<th>Polymorphism</th>
<th>Crude Albuminuria</th>
<th>Macroalbuminuria</th>
<th>Log(ACR)</th>
<th>eGFR (CKD-EPI)</th>
<th>eGFR (MDRD)</th>
<th>eGFR (CG)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>APOL1 G1/G2</strong></td>
<td>0.9239 (recessive)</td>
<td><strong>0.08537</strong></td>
<td>0.01829 (recessive)</td>
<td>0.07283 (recessive)</td>
<td>0.2434 (recessive)</td>
<td>0.1069 (recessive)</td>
</tr>
<tr>
<td><strong>HMOX1 – rs3074372</strong></td>
<td>0.4095 (codominant)</td>
<td>0.9218 (codominant)</td>
<td>0.8142 (codominant)</td>
<td>0.8521 (codominant)</td>
<td>0.6920 (codominant)</td>
<td>0.6946 (codominant)</td>
</tr>
<tr>
<td><strong>HMOX1 – rs743811</strong></td>
<td>0.07372 (codominant)</td>
<td>0.1549 (codominant)</td>
<td>0.8455 (codominant)</td>
<td>0.2081 (codominant)</td>
<td><strong>0.09865</strong> (codominant)</td>
<td>0.1836 (codominant)</td>
</tr>
<tr>
<td><strong>3.7del α-globin gene genotypes</strong></td>
<td>0.1371 (recessive)</td>
<td><strong>0.03441</strong> (codominant)</td>
<td>0.8098 (recessive)</td>
<td>0.4516 (codominant)</td>
<td>0.7236 (codominant)</td>
<td>0.6123 (codominant)</td>
</tr>
</tbody>
</table>

*APOL1 G1/G2* had no statistically significant association with crude albuminuria in the SCD cohort, however it was statistically significantly associated with log(ACR) values in a recessive model (p=0.01829). *APOL1* therefore holds potential to anticipate early development of kidney disease in the context of SCD, which likely has a selective association with these SNPs in comparison to the general population. The association of *APOL1* G1/G2 with the development of macro-albuminuria approaches significance in a recessive model (p=0.08537). The association of *APOL1* G1/G2 tends to significance with eGFR calculated using...
the CKD-EPI and CG equations in a recessive genetic model (p=0.07283 and p=0.1069, respectively). Figure 59 displays the distribution of crude albuminuria, ACR and eGFR values (calculated using the CKD-EPI equation) conditioned on the APOL1 G1/G2 genotypes in the SCD cohort.

The HMOX1 dinucleotide promoter polymorphism (rs3074372) showed no significant association with crude albuminuria, macro-albuminuria, or increased ACR levels in the SCD cohort (Figure 60). Furthermore, no significant association with eGFR was observed. The HMOX1 SNP (rs743811) tends to a significant association with crude albuminuria in a codominant (p=0.07372) and recessive (p=0.08013) model (Figure 61), as well as with macro-albuminuria in a recessive model (p=0.0629), but shows no significant association with ACR (p=0.6159), despite log transformation (p=0.8455). A near significant association was observed with eGFR calculated using the MDRD equation in a codominant (p=0.09865) model.

The 3.7 kB alpha-thalassemia deletion displayed no significant association with crude albuminuria or ACR data, however it was significantly associated with the development of macro-albuminuria in a co-dominant (p=0.034410), dominant (p=0.051535) and over-dominant model (p=0.009707). No significant association was observed with eGFR values for all three equations utilized.
Figure 59: Box and whisker plots showing the association of APOL1 G1/G2 with crude albuminuria, ACR, and eGFR values in the SCD patient cohort.

Box and whisker plots illustrating the distribution of (A) crude albuminuria, (B) ACR and (C) eGFR (calculated using the CKD-EPI equation) values conditioned on the APOL1 G1/G2 polymorphism, based on the presence of zero, one or two minor alleles. The horizontal lines that constitute the ‘box’ correspond to the lower quartile, median, and upper quartile parameters. The length of the ‘whiskers’ that extend from the box in the upwards and downwards direction represent a distance to the maximum and minimum values, respectively. Significant results are indicated using * (p<0.05).
Figure 60: Box and whisker plots showing the association of HMOX1 S/L Genotypes with crude albuminuria and ACR values in the SCD patient cohort.

Box and whisker plots illustrating the distribution of (A) crude albuminuria and (B) ACR values conditioned on the HMOX1 rs3074372 genotype based on the presence of short (S, ≤25) or long (L, >25) repeats. The horizontal lines that constitute the ‘box’ correspond to the lower quartile, median, and upper quartile parameters. The length of the ‘whiskers’ that extend from the box in the upwards and downwards direction represent a distance to the maximum and minimum values, respectively. Significant results are indicated using * (p<0.05).
Figure 61: Box and whisker plots showing the association of \textit{HMOX1} rs743811 Genotypes with crude albuminuria and ACR values in the SCD patient cohort.

Box and whisker plots illustrating the distribution of (A) crude albuminuria and (B) ACR values conditioned on the \textit{HMOX1} rs743811 genotype (T/C). The horizontal lines that constitute the ‘box’ correspond to the lower quartile, median, and upper quartile parameters. The length of the ‘whiskers’ that extend from the box in the upwards and downwards direction represent a distance to the maximum and minimum values, respectively. Significant (p<0.05) or near significant results (p\geq 0.05) are indicated using *.
SUMMARY OF THE RESULTS OF SECTION 1

The SCD patient cohort consisted largely of children and adolescent individuals (15 [9-23]), while the non-SCD patient cohort majorly consisted of adults (26 [23-30]). Micro-albuminuria and glomerular hyperfiltration were common in patients with SCD with prevalence’s of 61% (n=248) and 49.5% (n=200) (based on the CKD-EPI equation), respectively. There was no significant association between target SNPs in MYH9 and any early indicator of glomerular dysfunction, while the APOL1 G1/G2 were significantly associated with the ACR (p=0.01), and tended to a significant association with eGFR data as determined using the CKD-EPI equation (p=0.07). The dinucleotide promoter polymorphism of HMOX1 (rs3074372) displayed no association with any early indicator of renal dysfunction in SCD patients, however rs743811 (HMOX1) tended to a significant association with eGFR values calculated using the MDRD equation (p=0.09) in the SCD cohort. The co-inheritance of alpha-thalassemia was significantly increased in the SCD cohort in comparison to the controls (p=0.003), and displayed a protective association with the development of macro-albuminuria in SCD patients (p=0.03).

SECTION 2. ESKD GROUP AND MATCHED CONTROLS

Description of the co-morbidities

The studied cohort included:

1. 95 individuals with hypertensive nephropathy and a sub-population of 95 non-CKD hypertensives;
2. 36 patients with diabetic nephropathy and 35 non-CKD diabetics;
3. 16 HIV-associated nephropathy (HIVAN) patients and 15 non-CKD HIV controls;
4. 7 individuals affected by lupus nephritis and 11 non-CKD SLE controls;
5. 7 patients affected by Loa loa associated nephropathy and 4 non-CKD individuals with Loa loa filarial infection;
6. 8 individuals affected by focal segmental glomerulosclerosis (FSGS) and 2 controls.
Comparative sociodemographic cohort description

Table 17 describes the sociodemographic variables of the ESKD case cohort and matched controls. Both cohorts consisted of adult patients in comparison to the SCD cohort (table 6), and validates the age and sex-matched control cohort. The age distribution of the case and control cohorts are displayed in figure 62, with both distributions being negatively skewed.

Table 17: Socio-demographic and clinical characteristics of the ESKD case and control cohort (n=329)

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Cases (n=167)</th>
<th>Controls (n=162)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean age* (SD)</td>
<td>51.65 (12.60)</td>
<td>53.01 (11.67)</td>
<td>0.31</td>
</tr>
<tr>
<td><strong>Age groups</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤40 years, n (%)</td>
<td>29 (17.4)</td>
<td>26 (16.0)</td>
<td>0.749</td>
</tr>
<tr>
<td>&gt;40 years, n (%)</td>
<td>138 (82.6)</td>
<td>136 (84.0)</td>
<td></td>
</tr>
<tr>
<td><strong>Sex (F), n (%)</strong></td>
<td>49 (29.3)</td>
<td>54 (33.3)</td>
<td>0.435</td>
</tr>
<tr>
<td><strong>Ethnic Group</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bantu, n (%)</td>
<td>77 (46.1)</td>
<td>92 (56.8)</td>
<td>0.149</td>
</tr>
<tr>
<td>Semi-Bantu, n (%)</td>
<td>72 (43.1)</td>
<td>57 (35.2)</td>
<td></td>
</tr>
<tr>
<td>Sudanese, n (%)</td>
<td>18 (10.8)</td>
<td>13 (8.0)</td>
<td></td>
</tr>
<tr>
<td><strong>Blood pressure</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean SBP (SD)</td>
<td>147.36 (22.35)</td>
<td>140.92 (28.43)</td>
<td>0.023</td>
</tr>
<tr>
<td>Mean DBP (SD)</td>
<td>85.46 (12.53)</td>
<td>82.84 (13.86)</td>
<td>0.073</td>
</tr>
<tr>
<td><strong>SES</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low, n (%)</td>
<td>5 (3.0)</td>
<td>1 (0.6)</td>
<td></td>
</tr>
<tr>
<td>Middle, n (%)</td>
<td>48 (28.7)</td>
<td>65 (40.1)</td>
<td></td>
</tr>
<tr>
<td>High, n (%)</td>
<td>114 (68.3)</td>
<td>96 (59.3)</td>
<td></td>
</tr>
<tr>
<td>Family history of CKD, yes n (%)</td>
<td>19 (11.4)</td>
<td>12 (7.4)</td>
<td>0.218</td>
</tr>
</tbody>
</table>

CKD = chronic kidney disease; SBP = systolic blood pressure; DBP = diastolic blood pressure; SES = socio-economic status. * = years.
**Molecular Association Studies**

The genotypic frequencies of the *APOL1* and *HMOX1* SNPs, and the p-value resulting from chi-squared tests to investigate the differences in these frequencies between the ESKD cases and controls are shown in table 18.
Table 18: The frequency of the genotypes of rs3074372, rs743811, rs73885319, rs60910145, rs71785313 and APOL1 G1/G2 alleles within the ESKD case and control cohorts. The p-value obtained from the chi-squared test for deviation from Hardy-Weinberg equilibrium is displayed for each SNP, as well as the p-value obtained from the test of association between genotype frequency and ESKD outcome (adjusted for age and sex). Statistically significant results (p<0.05), or results approaching significance (p≤0.05) are highlighted in bold.

<table>
<thead>
<tr>
<th>Variables</th>
<th>ESKD Cases</th>
<th>ESKD Controls</th>
<th>Association test p-value (model)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Freq (%)</td>
<td>HWE (p-value) Number of observations</td>
<td>Freq (%)</td>
</tr>
<tr>
<td>HMOX1 – rs3074372 (S/L)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S/S</td>
<td>2 (1.45)</td>
<td>0.19</td>
<td>138</td>
</tr>
<tr>
<td>S/L</td>
<td>18 (13.04)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>L/L</td>
<td>118 (85.51)</td>
<td>0.19</td>
<td>144</td>
</tr>
<tr>
<td>HMOX1 – rs743811 (T/C)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TT</td>
<td>107 (74.31)</td>
<td>0.19</td>
<td>144</td>
</tr>
<tr>
<td>TC</td>
<td>32 (22.22)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC</td>
<td>5 (3.47)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>APOL1 – rs73885319 (A/G)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AA</td>
<td>115 (76.16)</td>
<td>0.0017</td>
<td>151</td>
</tr>
<tr>
<td>AG</td>
<td>28 (18.54)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>8 (5.30)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>APOL1 – rs60910145 (T/G)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TT</td>
<td>119 (80.95)</td>
<td>0.0011</td>
<td>147</td>
</tr>
<tr>
<td>TG</td>
<td>22 (14.97)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td>6 (4.08)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>APOL1 – rs71785313 (WT/del)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WT/WT</td>
<td>105 (76.09)</td>
<td>0.27</td>
<td>138</td>
</tr>
<tr>
<td>Del/WT</td>
<td>29 (21.01)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Del/del</td>
<td>4 (2.90)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>APOL1 – G1/G2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G/G</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>G/A</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A/A</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The association of target SNPs with development of ESKD displayed no significant association between homozygotes and compound heterozygotes for APOL1 G1/G2 allele and cases and controls, with and without end stage renal disease.
Individuals with non-diabetic ESKD as the comorbidity (with hypertensive nephropathy in the majority (n=95)) were analyzed separately. There was no significant association between the *APOL1* G1/G2 variant and the development of non-diabetic ESKD (p=0.857). However, the high risk G1 variant alone was significantly associated with the development of non-diabetic ESKD in a recessive genetic model (p=0.014), while G2 was not (p=0.2847). The *HMOX1* dinucleotide promoter polymorphism was significantly associated with the development of ESKD in a codominant (p=0.0020401), dominant (p=0.0005690), and overdominant (p=0.0049646) model, with a recessive model (p=0.0634278) approaching significance. Further analysis identified the repeat polymorphism as being significantly associated with non-diabetic ESKD (p<0.0001), however not with diabetes-associated ESKD (p=0.770). There was no significant association between the *HMOX1* SNP rs743811 and the development of ESKD, despite separation into non-diabetic and diabetic ESKD cases and controls (p=0.412 and p=0.873, respectively).

**SUMMARY OF THE RESULT OF SECTION 2**

The cohort of patients with ESKD of various etiologies, and the age-, sex- and comorbidity-matched control group consisted majorly of adult patients (52 [46-59] and 54 [47-61], respectively). There was no significant association between the *APOL1* G1/G2 variants or rs743811 (*HMOX1*) and the development of ESKD in our Cameroonian cohort, however analysis of *APOL1* G1 alone produced identified a significant association with non-diabetic CKD in this group of Cameroonian patients. On the contrary, the dinucleotide promoter repeats in *HMOX1* (rs3074372) was associated with the development of ESKD, more specifically in patients with non-diabetic associated nephropathy.
CHAPTER 4: DISCUSSION.

4.1. ORIGINALITY AND GENERAL COMMENTS

This study is, to our knowledge, the first of its kind to investigate homozygosity and compound heterozygosity for APOL1 G1/G2 and variants in HMOX1 in African SCD cohorts from Cameroon and to compare them to that of various World populations. The study further provides some unique associations between these genomic variants with indicators of renal dysfunction in this Cameroonian group of mainly children and adolescents affected by SCD. The result indicates a high proportion of Cameroonian SCD infant and adolescent patients with micro-albuminuria that increased with age, as well as a high proportion of patients presenting with hyperfiltration as evaluated by eGFR values, that decreases with age and was higher among males. The study also provides and confirms important clinical predictors of renal function; hemoglobin level was a significant positive predictor, whereas serum creatinine, systolic blood pressure and urinary ACR were negative predictors of eGFR.

The significant association of variants at APOL1 and near-significant association of rs743811 (HMOX1) with crude albuminuria or glomerular hyperfiltration, both early indicators of renal dysfunction, are novel findings. This acts as an indication of the possibility of using these markers, and those that still need to be discovered, to screen for individuals at the risk of developing renal disease, for anticipatory guidance, e.g. after newborn screening for SCD. The variable MAF of these variants among African populations, specifically for APOL1, with the data extracted from the 1000G project, supporting the known huge genomic variations that exists among African populations (Gurdasani et al., 2015), that in this specific case is confounded and greatly influenced by various trypanosome endemicity. Moreover, these results indicate the need to extend the study with further genotyping e.g. using whole exome and genome sequencing to refine the genomic architectures in target regions such as the MYH9, APOL1 and HMOX1 genes, in many SCD groups from various regions of Africa. This will help to refine and define appropriate tagging variants for micro-albuminuria in these populations. The comparison with a non-SCD group of older patients with ESKD indicates another level of complexity with a surprising insignificant association between APOL1 G1/G2 variants with the development of ESKD, in this small group of patients with various co-
morbidities. In anticipation of increased sample size, it is possible that other variants may be identified in APOL1, that could better serve as tagging variants for ESKD in the Cameroonian population. Alternatively, these preliminary results could indicate that variants in HMOX1 gene, rather than APOL1, are more prominently associated with in the development of ESKD in this non-SCD population from sub-Saharan Africa. Future studies require larger sample sizes of older SCD patient populations, as well as complete genotyping within the non-sickle population of study, such as alpha-thalassemia and HbS carrier status, HMOX1, APOL1, and the occurrence of kidney disease. Sickle Cell trait (SCT) is associated with reduced eGFR, increased risk of CKD and albuminuria, in comparison to homozygotes for wild type adult hemoglobin (i.e. HbAA), and may account for increased risk of kidney disease observed in African Americans (Naik et al., 2014), therefore it is important to consider.

4.2. MICRO-ALBUMINURIA AND GLOMERULAR FILTRATION IN THE SCD POPULATION

In our cohort, 60.9% (n=248) and 2.5% (n=10) of patients (n=404) were found to have albuminuria values of between 3 and 30 mg/mL or above 30 mg/mL, respectively (Table 8, section 3.2.2.). Micro-albuminuria is considered the primary measure for the early detection of renal dysfunction (Ataga et al., 2014; Gosmanova et al., 2014). Chronic kidney disease (CKD) and end stage renal disease (ESRD) are a leading cause of morbidity and mortality in older individuals with SCD, and hence routine monitoring for early detection of renal problems is recommended during maintenance health visits (M. R. Asnani and Reid, 2015; M. Asnani and Reid, 2015). Micro-albuminuria is preferred to proteinuria measurement due to difficult standardization resulting from the variable composition of protein in the urine (Levey et al., 2013). The urine albumin-to-creatinine ratio (ACR) resulting from urine albumin and creatinine measurement in an untimed urine collection is the preferred approach for detecting the presence of crude albuminuria, and is comparable to results obtained from albumin excretion rate calculated from timed urine collections (Eknoyan and Levin, 2002; Levey et al., 2013). Micro- and macro-albuminuria, as defined by the ACR (Table 8, Section 3.2.2.) were present in 73.4% (n=121) and 5.1% (n=8) of the patient cohort (n=158), respectively. This differs to the 40% micro- and 19% macro-albuminuria seen in American adult SCD patients (Haymann et al., 2010), 25.9% and 16.5% prevalence seen in a Jamaican
SCD population (Asnani et al., 2011), 28.2% among children with SCD in Uganda (Mawanda et al., 2011), 44% in Barbados (Quimby et al., 2014) and the 50% combined micro- and macro-albuminuria prevalence reported in Nigerian patients (Bolarinwa et al., 2012). This may be due primarily to the age structures of the 3 cohorts, with the American, Jamaican and Nigerian cohorts comprising a majority of adult patients, compared to high prevalence of children and adolescents in the present study. However, our result is still substantially higher than the 18.5% and 6.2% prevalence of micro-albuminuria and proteinuria among Nigerian and Congolese children with SCD, respectively (Aloni et al., 2014; Eke et al., 2012). The prevalence of micro-albuminuria in our cohort is also higher compared to previous results of 19% (McBurney et al., 2002) and 15.8% (Alvarez et al., 2006), which may be partly due to discrepancies in hemoglobin values between the cohorts. The children in these studies had mean hemoglobin values of 8.1 and 9.6 g/dl, respectively, while our cohort had a median hemoglobin value of 7.4 g/dl. Decreased hemoglobin is a significant predictor of increased albuminuria in SCD patients (Alvarez et al., 2006; Mawanda et al., 2011; McBurney et al., 2002), thus may act to increase the prevalence of micro-albuminuria in our cohort. In African-American patients (Dharnidharka et al., 1998), micro-albuminuria was not seen in any children younger than 7, contradictory to our cohort where 54% of children aged 2-7 years old (n=40) had developed micro-albuminuria. The relatively high proportion of patients with micro-albuminuria in the present study could be attributed to the hospital-based recruitment employed in the design of the study that could have favored the recruitment of sicker individuals. A high proportion (63.6%) of micro-albuminuria was equally observed in a population of hospital-based SCD patients in Saudi Arabia (Abo-Zenah et al., 2009). The distribution of median albuminuria values per age group in our SCD cohort (figure 38, section 3.1.4.2.) appears to indicate that crude albuminuria increases with age in children as previously reported in various African populations from Cameroon, Cote d’Ivoire, Mali and Senegal, with SCD SS, SC and S-beta(0)-thalassemia types (Ranque et al., 2014). However, in the present cohort, we have observed that crude albuminuria seems to reduce with age after the fourth decade of life in this group of patients. This figure, however, may be misleading as the age groups of 38-43, 44-49 and 50-55 have a minority of participants and contained only eight, six and five individuals, respectively. Alternatively, individuals that survive to this age in developing countries are possibly protected against risk of kidney disease on the basis of their genetic constitution, and it could therefore be unsurprising that the older patients
display reduced median albuminuria values, as this is likely a contributing factor to their survival rate.

The glomerular filtration rate (GFR) is thought to be the most ideal marker of renal function in humans, both in health and disease, however its measurements tend to be cumbersome and expensive (Xie et al., 2008). Therefore, estimated GFR values are being utilized in clinical practice as per Kidney Disease Outcomes Quality Initiative (KDOQI) guidelines (Levey et al., 2003). Glomerular hyperfiltration, defined as an estimated filtration rate >130 ml/min/1.73m² and >140 ml/min/1.73m² in women and men, respectively, in accordance with previous studies of SCD cohorts (Haymann et al., 2010; Hirschberg, 2010; Vazquez et al., 2014), was prevalent in 52.97% (n=214), 42.57% (n=172) and 49.5% (n=200) of the SCD cohort based on calculations using the MDRD, CG and CKD-EPI calculations, respectively. This is comparable with previous prevalence rates of 51%, 53.1% and 41.5% in adolescent French, adult sub-Saharan African and French West Indian natives, respectively (Arlet et al., 2012; Haymann et al., 2010), however much higher than what has been previously reported (30.6%) among adult Nigerians with SCD (Bolarinwa et al., 2012). This study further validates previous findings that glomerular damage is a common phenotype in SCD, and is frequently associated with the development of micro- or macro-albuminuria.

The MDRD equation for estimation of the glomerular filtration rate was postulated to be a more robust predictor of hyperfiltration in adult SCD patients, in comparison to the Cockcroft-Gault equation, despite its systemic overestimation of GFR values (Haymann et al., 2010). Recent studies contradict these findings through identification of the CKD-EPI equation, without correction for ethnicity, as being more precise and displaying the least bias in estimation, particularly for increased estimation values (Arlet et al., 2012).

Owing to the young age of the cohort, a possible limitation of the eGFR calculations is the use of the CKD-EPI, Cockcroft-Gault and MDRD equations, where the Schwartz formula could have applied. This formula is postulated to be most accurate in patients younger than 19 (Schwartz et al., 2009), however new reports indicate that the original Schwartz formula overestimates eGFR values, with the recommendation that a new adjusted Schwartz equation offers the best approximate calculation (Selistre et al., 2016). Nevertheless, the adjusted Schwartz equation includes the cystatin C and blood urea nitrogen variables (Selistre et al., 2016) that
were not available for the SCD cohort. Moreover, the CKD-EPI equation has been reported to be the most appropriate for SCD populations (Arlet et al., 2012), and we trust that the outcomes of the present study are appropriate.

4.3. CLINICAL FACTORS AFFECTING CRUDE ALBUMINURIA AND eGFR

The study also validated the significant positive association of crude albuminuria with increasing age, as previously reported in American (Guasch et al., 2006; Saraf et al., 2014) and Nigerian patients (Eke et al., 2012). Furthermore, this study indicates a hyperfiltration prevalence based on eGFR values among SCD patients in Cameroon comparable to those of French (51%) adult Nigerian (30.6%), adult sub-Saharan African and French West Indian natives (53.1%), respectively (Arlet et al., 2012; Bolarinwa et al., 2012; Haymann et al., 2010). Crude albuminuria displayed a near significant association with reduced hemoglobin (p=0.0691), while increased BMI (p=0.0489), leukocyte (p=0.0202), mean arterial pressure (p=0.0175) and MCV (p=0.008) values identified as positive predictors of increasing albuminuria in our SCD cohort. This validates the significant negative association between hemoglobin and albuminuria reported in previous studies (Alvarez et al., 2008; Becton et al., 2010; McKie et al., 2007; Thompson et al., 2007). Alpha-thalassemia is known to reduce the MCV of erythrocytes, and exerts a protective effect against the development of macro-albuminuria, as discussed earlier (Guasch et al., 1999; Lamarre et al., 2014; Nebor et al., 2010). This is further validated by our findings that elevated MCV values, that are associated with the non-coinheritance of alpha-thalassemia, are associated with increased-risk for the development of micro-albuminuria. Our findings agree with previous association studies of adult Jamaican SCD patients and African American children with SCD whereby lower Hb levels, increased white blood cell counts and higher MCV’s were significant clinical indicators of impending kidney disease (Asnani et al., 2011; Wigfall et al., 2000). We did not observe the significant association between rising blood pressure (p=0.06659 for SBP, p=0.6679 for DBP) and increased albuminuria seen previously (Asnani et al., 2016; Aygun et al., 2011), even though a significant association was observed between ACR values and increased SBP (p=0.0217), however not with DBP (p=0.6902). No significant association with albuminuria and serum creatinine (p=0.2951) was observed, similar to previous reports in African American children with SCD (Wigfall et al., 2000). Our results showed no significant
association between albuminuria and hyperfiltration, contrary to previous reports in adult SCD patients (Arlet et al., 2012; Thompson et al., 2007), which may be due to the high prevalence of children and adolescents, with consequent low proportions of macroalbuminuria in our cohort. The relationship between albuminuria and reduced glomerular filtration is, however, controversial as previous studies of Ugandan children (Mawanda et al., 2011) and Brazilian adults (Barros et al., 2006) found no significant association between the clinical phenotypes. In the present study, significant positive associations between eGFR and ACR values were observed for all three equations, agreeing with previous studies of adult SCD patients (Asnani et al., 2016; McPherson Yee et al., 2011). Among Jamaicans living with SCD, up to half of the hyperfiltrating individuals might also have micro-albuminuria (M. R. Asnani and Reid, 2015; M. Asnani and Reid, 2015), further supporting the fact that the state of hyperfiltration is an early pathology in those likely to decline to CKD. Vazquez et al. (2014) have similarly described increasing rates of micro-albuminuria development with increasing levels of hyperfiltration. The surprising positive association of albuminuria with GFR was also recently shown in a prospective longitudinal study (Asnani et al., 2016). Nevertheless, the contradictory findings between micro-albuminuria and ACR with declining eGFR, as reported in the present study, indicate the need for a prospective cohort study to unravel the role of albuminuria as a predictor of GFR and the progression of sickle glomerulopathy and worsening CKD.

Glomerular filtration rate, another early indicator of renal dysfunction, was increased in younger patients, however reduced with increasing age (p<0.0001), similar to previous reports (Aygun et al., 2011; Wigfall et al., 2000). Other studies have also indicated increasing albumin excretion with increasing age (Asnani et al., 2016, 2011; Guasch et al., 2006; Thompson et al., 2007). Renal functional decline with age is anticipated in individuals with HbSS disease, and is one of the most common causes of morbidity and mortality in older adults with SCD, with the annual rate of decline in GFR estimated to be 3.2 ml/min/1.73 m² among Jamaican adults with SCD (Asnani et al., 2016), that is much higher the normal expected levels of 0.75 ml/min/1.73 m² (Lindeman et al., 1985). The positive association of albuminuria and GFR was also shown in Jamaican adults with SCD (Asnani et al., 2016). In the present study hyperfiltration is shown to be very prevalent in the Cameroonian population with SCD, indicating that it is possible the first step to ensuing renal damage and a precursor
to eventual focal segmental glomerulosclerosis, which is the most commonly found pathology in SCD (Guasch et al., 1997; Scheinman, 2009).

Elevated hemoglobin, systolic and diastolic blood pressure, BMI, serum creatinine and mean arterial pressure were all significant predictors for decreased eGFR, replicating some results of previous studies (Asnani et al., 2016; Becker et al., 2014). In the present study systolic and diastolic blood pressure and serum creatinine levels both showed significant negative associations with eGFR. Cross-sectional studies have reported similar associations (Asnani et al., 2011; Powars et al., 1991), but also a recent longitudinal study that points to a predictive role as well (Asnani et al., 2016), however this is in contradiction with previous associations of positive correlations between eGFR and blood pressure (Aygun et al., 2011). The negative association of hemoglobin with estimated GFR, reported here reflects, as previously speculated by previous studies, the relationship between hemoglobin and elevated eGFR values (Asnani et al., 2016; Haymann et al., 2010). Worsening anemia is known to be associated with renal insufficiency in sickle and non-sickle populations (Norris, 2006; Powars et al., 1991). Furthermore, in the present study male patients displayed a significantly increased risk of developing hyperfiltration, equal to findings recently shown among Jamaican adult SCD patients in a study that investigated longitudinal changes in GFR over time (Asnani et al., 2016).

4.4. COMPARATIVE MINOR ALLELE FREQUENCIES IN MYH9, APOL1 AND HMOX1 IN VARIOUS WORLD POPULATIONS

The MAF for the selected MYH9 SNPs show great variation among the African populations (Table 12, section 3.2.). This is evident in the MAF results for rs11912763 (Table 12, section 3.2.). The MAF in the Yoruban and Esan populations (0.324 and 0.399, respectively) was approximately three times greater than in the Cameroonian patient cohort (0.118). The results of MAF of the studied SNPS in the present cohort are most similar to those obtained from an African-American population in Southwest USA. This is likely a result of the flow of West African natives through the slave trade in the America’s, as substantiated by the high prevalence of Benin haplotypes in these regions (Bitoungui et al., 2015). The MAF of rs743811 in HMOX1 was similar among African countries, however was approximately two and four times greater in European (0.25) and Asian (0.66) populations, respectively. In addition, European and Asian populations from the 1000G displayed no presence of the APOL1
rs73885319 and rs60910145 minor allele, validating hypotheses that these variants originated and exceptionally increased their frequency in Africa, following a selective pressure by trypanosomes parasites (Genovese et al., 2010a; Kasembeli et al., 2015b; Limou et al., 2014). The Yoruban and Esan (0.375 and 0.495 for both SNPs, respectively) populations have an approximate four-fold increase in the MAF of these variants in comparison to our Cameroonian cohorts and other African populations. Similar to the genetic persistence of the HbS mutation, which protects against mortality from malaria, the APOL1 G1 and G2 variants are believed to have been selected by affording protection from Trypanosoma brucei rhodesiense infection (Genovese et al., 2010a; Limou et al., 2014; Van Xong et al., 1998). ApoL1 complexes with haptoglobin-related protein to form the trypanosome lytic factor, and this complex can scavenge cell-free hemoglobin (Nielsen et al., 2006), however there is still much of the evolutionary history that is still to be learned. The current hypothesis indicates that the APOL1 variants arose in sub-Saharan Africa, likely West Africa (Genovese et al., 2010a), with results from the Yoruban population in Nigeria from the 1000G project providing convincing evidence of a selective sweep in these regions. Indeed, the APOL1 G1/G2 alleles are widely found throughout Africa (Limou et al., 2014), although with widely differing frequencies, as supported by the results of our study.

The difference in MAF among various African populations, in relation to historical and present trypanosomiasis endemicity, informs us that the association of APOL1 variants with kidney diseases and its effect size are unlikely to be identical to the range of evidence seen so far in African Americans. This further highlights the wide genetic diversities and variations that exist within Africa (Gurdasani et al., 2015), substantiating the need for increased research in various populations on the continent to uncover the underlying regional specificities for specific polymorphisms (Pule et al., 2015), that will be useful in precision medicine.

4.5. VARIANTS IN MYH9, APOL1 AND HMOX1 AND SCD NEPHROPATHY

The SCD patient cohort contained roughly equivalent numbers of males and females (210 and 203, respectively), with a relatively young median age of 15 years and a low proportion of adult patients (40%, n=162) in comparison to the non-SCD control cohort with a median age of 26, and 89% (n=62) prevalence of adults. The young age of this selected group of SCD patients could be an indirect indicator of the poor survival rates of SCD patients in Cameroon.
Previous studies have shown an association of homozygous or compound heterozygous
*APOL1* G1/G2 variants with kidney disease in adult patients living with sickle cell disease, possibly through increased risk of hemoglobinuria, and associations of *HMOX1* variants with kidney disease, possibly through reduced protection of the kidney from hemoglobin-mediated toxicity (Saraf et al., 2015). The relatively young age of the cohort could account for a lower proportion of patients without macro-albuminuria that was previously associated with variants in *MYH9* in adult patients (Ashley-Koch et al., 2011; Kopp et al., 2008). This indicates the need to perform a separate study in the near future, including adult patients only, to have a better idea of the proportion of macro-albuminuria as well its association with *APOL1* and *HMOX1* variants.

Alternatively, the relatively high proportion of micro-albuminuria in the majority of patients could have contributed to the lack of significant associations for some *APOL1* and *HMOX1* variants with micro-albuminuria, as the present group of SCD patients is virtually a cohort of patients with micro-albuminuria, an indirect proxy of clinical severity of SCD in the sample studied.

4.5.1. SNPs IN *MYH9* AND *APOL1* G1/G2

Many genetic variants have been associated with CKD in African-Americans, including *MYH9*, which encodes the non-muscle myosin IIA heavy chain and is a component of the cytoskeleton of podocytes (Palmer et al., 2014). A single report found that individuals with the *APOL1* G1/G2 risk variants (defined as being homozygous or compound heterozygous for the G1 and/or G2 risk variants using a recessive model) were 3.4-times more likely to have dipstick-defined proteinuria and that variants in *MYH9* were independently associated with proteinuria after adjusting for *APOL1* variant status (Ashley-Koch et al., 2011). Furthermore, a significant interaction between the *APOL1* G1/G2 risk variants and a *MYH9* risk haplotype were observed in predicting eGFR (Ashley-Koch et al., 2011). In the present research, no significant association was found between the four SNPs in *MYH9* and the development of micro-albuminuria based on crude albuminuria or the ACR values. These findings cannot be compared to those reported by Ashley Koch et al (Ashley-Koch et al., 2011), which state that the minor allele predisposes SCD patients to increased risk of proteinuria (and not micro-albuminuria) development and subsequently higher risk of developing kidney disease among
adult patients. However, the findings in the present study could indicate that the SNPs that were explored in \textit{MYH9} in the present study are not appropriate markers to detect early stage renal dysfunctions among infant and adolescent patients living with SCD in Cameroon. The lack of significant results, as well as the growing body of work identifying \textit{APOL1}, and not \textit{MYH9}, as contributing to the development of kidney disease (Kopp et al., 2011; Nelson et al., 2010; Tzur et al., 2010).

The G1 and G2 alleles of \textit{APOL1} are mutually exclusive, never occurring on the same chromosome and have comparable effect sizes for kidney disease, therefore their independent frequencies can be combined to determine the burden of risk variants for a particular cohort (Genovese et al., 2010b; Kopp et al., 2011). A significant difference was observed with homozygosity or compound heterozygosity for \textit{APOL1} G1/G2 between SCD patients and the non-SCD control cohorts (p=0.0097), which could tend to indicate a selective force and enrichment of favorable and protective alleles among SCD patients. We have previously shown such a trend for the 3.7 kB alpha-thalassemia deletion which has a proportion that is double in SCD patients compared to non-sickle Cameroonian controls, likely a result of its association with improved hematological indices, protective effect against stroke or kidney disease, and possibly increased survival of SCD patients (Rumaney et al., 2014; Wonkam et al., 2014b). Interestingly, homozygosity or compound heterozygosity for \textit{APOL1} G1/G2 displayed a significant association with ACR values in a recessive model (p=0.018) in the present study, as observed previously in African American adult patients, where the variants displayed a prevalence of 14% (Saraf et al., 2015). In addition, \textit{APOL1} G1/G2 tended to a significant association with increased eGFR as calculated using the CKD-EPI equation (p=0.073) in a recessive genetic model. Moreover, the association with macroalbuminuria also approached significance in a recessive model (p=0.085). These results indicate that variants in \textit{APOL1} could be used in the context of SCD to anticipate early development of kidney disease and offer interesting perspectives for the future practice of precision medicine for SCD in Africa. It could, for example, be anticipated that typing variants will guide the indication of the use of Hydroxyurea that has been shown to be beneficial in reducing albuminuria (Alvarez et al., 2008; Bartolucci et al., 2015; Laurin et al., 2014; Mckie et al., 2007) or the glomerular filtration rate (Aygun et al., 2013) in SCD. The possible association of \textit{APOL1} G1/G2 with eGFR agrees with previous studies of a group of adult
African-American SCD patients (Saraf et al., 2015), and urges for further studies among a much larger group of Cameroonian adults living with SCD. Previous studies have shown that the CKD-EPI equation, in estimation of the glomerular filtration rate, was the most precise, particularly with regards to high GFR values, and demonstrated the least bias in adult SCD patients and has validated the CKD-EPI equation to be the most accurate of those that currently exist (Arlet et al., 2012; Asnani et al., 2013). This highlights its possible utility in clinical settings, and the importance of its possible association with genetic markers, as we have shown in the present study.

4.5.2. VARIANTS IN \textit{HMOX1}

When erythrocytes are lysed, extracellular hemoglobin is released and easily oxidized from ferrous (Fe\textsuperscript{2+}) to ferric (Fe\textsuperscript{3+}) hemoglobin (methemoglobin), which in turn readily releases heme into the vasculature, and subsequently circulating cell-free hemoglobin is increased more than 10-fold in SCD patients (Muller-Eberhard et al., 1968; Naumann et al., 1971). Therefore, intravascular hemolysis is a potential cause of oxidative injury and endothelial damage in SCD (Gladwin and Sachdev, 2012; Nath and Katusic, 2012; Tracz et al., 2007). Exposing human proximal tubular cells to increasing cell-free hemoglobin led to increasing concentrations of supernatant kidney injury molecule, reduced viability and induction of \textit{HMOX1} (Saraf et al., 2015). \textit{HMOX1} rs743811 was associated with chronic kidney disease stage, and longer \textit{HMOX1} GT-tandem repeats (>25) were associated with a lower estimated glomerular filtration rate in African Americans (Saraf et al., 2015). In the present study, the \textit{HMOX1} GT-tandem repeat polymorphism in the promoter (rs3074372) displayed a trimodal distribution of repeats in all cohorts, replicating findings of previous studies (Bean et al., 2012; Ono et al., 2004; Saraf et al., 2015). The repeats were classified as short (≤25 repeats) or long (>25 repeats) in accordance with previous investigations of SCD cohorts (Bean et al., 2012; Saraf et al., 2015), with evidence suggesting that lower numbers of repeats have higher inducible heme oxygenase expression levels and a subsequent protective effect in the face of increased hemolysis (Hirai et al., 2003; Kimpara et al., 1997; Taha et al., 2010). The dinucleotide promoter polymorphism showed no significant association with crude albuminuria, macro-albuminuria, increased ACR levels, or eGFR for all equations utilized for
estimation in the SCD patient cohort. This did not replicate previous findings of a significant association between long \textit{HMOX1} GT dinucleotide repeats (>25) and lower eGFR rates in African American SCD patients (Saraf et al., 2015), which may be due to a lack of statistical power, and the young age of our cohort, with a relatively low proportion of adults. In African-American children with SCD, shorter \textit{HMOX1} GT repeats were associated with lower rates of hospitalization for acute chest syndrome consistent with the shorter GT-tandem repeat length polymorphism having a protective association in SCD (Bean et al., 2012), and will deserve further investigation in Africa.

However, in the present report, the SNP in \textit{HMOX1} (rs743811) tended to a significant association with crude albuminuria in a codominant (\(p=0.07372\)) and recessive (\(p=0.08013\)) model in the SCD cohort, as well as with macro-albuminuria in a recessive model (\(p=0.0629\)). Moreover, the SNP in \textit{HMOX1} (rs743811) displays a near significant association with GFR estimated using the MDRD equation in a codominant (\(p=0.09865\)) model. Glomerular hyperfiltration is considered an early indicator of renal dysfunction in the context of SCD, thus the results indicated that this variant may be potentially also use for anticipatory guidance for kidney disease in children and adolescents with SCD, and encourage additional studies to both confirm the association and also explore other tag variants in multiple African settings. The nearly significant association of \textit{HMOX1} (rs743811) with albuminuria and eGFR is in line with previous findings of \textit{HMOX1} rs743811 significantly contributing to the development of chronic kidney disease and ESKD in an African American population (Saraf et al., 2015). However, it remains to be investigated if a different tagging variant would be suited to the Cameroonian population.

4.6. CO-INHERITANCE OF THE 3.7 kB \(\alpha\)-GLOBIN GENE DELETION AND SCD NEPHROPATHY

The 3.7kB \(\alpha\)-globin gene deletion, a known modifier of SCD, was found to be co-inherited in 41.4% (n=140) of patients, a significantly higher proportion than the 19% (n=11) of individuals among the non-SCD control cohort (\(p=0.003\)), in line with data that we previously reported (Rumaney et al., 2014; Wonkam et al., 2014b). Alpha-thalassemia, improved hematological indices and reduction of hemolysis due to a decrease in the quantity of hemoglobin in SCD, is
thought to be associated with reduced severity of clinical manifestation and improved survival among Cameroonian with SCD (Rumaney et al., 2014; Wonkam et al., 2014b). This high prevalence of the co-inherited 3.7 kB alpha-globin gene deletion and SCD as compared with non-SCD controls (p=0.00276), was also reported in Saudi Arabia (40%) (Guasch et al., 1999), Tanzania (58%) (Cox et al., 2013), and in the USA among African Americans (41%) (Guasch et al., 1999). The 3.7kB α-globin gene deletion was also significantly deviated from the Hardy Weinberg equilibrium in SCD patients (p=0.0005), and could therefore be attributed to this selective association, as no genotyping issues were noticed. As replicated in the present study, co-inheritance of SCD and alpha-thalassemia was also previously shown to have a protective effect against macro-albuminuria development (Day et al., 2012; Lamarre et al., 2014; Nebor et al., 2010).

4.7. HBB HAPLOTYPE AND SCD NEPHROPATHY

The Benin HBB haplotype was the most common among the SCD cohort, present in 73% of patient chromosomes (n=510), that is also associated with a milder and more moderate course of SCD (Steinberg et al., 1998), No significant differences between the HBB haplotypes and micro-albuminuria were observed in the present, validating results seen in previous studies (Guasch et al., 2006). This may also be likely due to the high proportion effect of the Benin haplotype in the present cohort, leaving little room for comparison and detection of significant associations, in comparison to the Cameroon haplotype.

4.8. VARIANTS IN APOL1 AND HMOX1 IN THE NON-SCD ESKD POPULATION

APOL1 and MYH9 genes were identified using mapping by admixture disequilibrium as a risk locus for non-diabetic forms of kidney disease including idiopathic and HIV-associated focal segmental glomerulosclerosis (FSGS) and kidney disease clinically attributed to hypertension (Genovese et al., 2010a; Kopp et al., 2008; Rosenberg et al., 2015; Tzur et al., 2010). APOL1 risk variants were also associated with non-diabetic forms of CKD among Nigerians of Yoruban ethnicity (Tayo et al., 2013; Ulasi et al., 2013) as well as Southern African patients with HIV-associated nephropathy (Kasembeli et al., 2015a), and the absence of risk variants protected against HIV-associated nephropathies among Ethiopians (Behar et al., 2011). Unexpectedly, there was no significant association (p=0.98) among homozygosity and compound
heterozygosity for APOL1 G1/G2 variants and the development of ESKD in the non-SCD non-diabetic cohort from Cameroon, in contrast to previously reported significant associations in adult African American patients living with (Saraf et al., 2015) and without SCD with non-diabetic ESKD (Genovese et al., 2010a; Kopp et al., 2011; Tzur et al., 2010). The APOL1 G1/G2 variants have been strongly associated with non-diabetic ESKD in non-sickle African American populations (Genovese et al., 2010a; Kopp et al., 2011; Tzur et al., 2010). Patients who had inherited two risk alleles displayed earlier onset of focal segmental glomerulosclerosis, and accelerated progression to ESKD, as well as increased risk for HIV-associated nephropathy and hypertension-attributed ESKD (Genovese et al., 2010a; Kopp et al., 2011). Studies in a Nigerian population of non-sickle patients with non-diabetic CKD indicated a significant association with APOL1 G1 variant (Tayo et al., 2013). Our lack of reproducible results may be due to insufficient statistical power as a result of small sample size used in Cameroon, as analysis of only the patients successfully genotyped for all three variants constituting APOL1 G1/G2 was performed (n=242). The lower MAF of APOL1 G1/G2 variants seen in our Cameroonian cohort in comparison to the Nigerian population indicates the need for increased sample sizes of Cameroonian patients. This result may also indicate that the APOL1 G1/G2 variants have a lesser association with ESKD in non-diabetic patients (compare to e.g. HMOX1 variants in the present study) and do not have the same effect in Africans as previously reported in African American patients, as this could be dependent on APOL1 G1/G2 MAF, as well as past and present trypanosomiasis endemicity, as discussed earlier (Limou et al., 2014). In the context of Africa, where Trypanosoma brucei gambiense is not equally common in all the regions as it is in Cameroon, the frequency of the protective allele may be too high or too low to detect a significant association with the development of ESKD, highlighting the need for further studies of these variants in African populations (Kasembeli et al., 2015b). However, the effect of HbS carrier status as well as the carrier status of the 3.7 kB alpha-globin gene deletion, in relation to APOL1 G1/G2 or HMOX1 variants, remains to be investigated, specifically in the context of malaria endemicities. Intravascular hemolysis that is associated with both malaria and SCD, is associated with chronic renal dysfunctions and its progression in SCD (Saraf et al., 2014), but have not been investigated in non-sickle patients, more specifically those with sickle cell trait (SCT). Even though SCT does not seem to increase the risk of microvascular complications with diabetes mellitus (Bleyer et al., 2010; Oli et al., 2004), more recent convincing data has shown that the presence of SCT is associated with an
increased risk of CKD, decline in eGFR, and albuminuria, compared with non-carriers among African Americans (Naik et al., 2014)

4.9. PRACTICAL IMPLICATIONS AND RESEARCH RECOMMENDATIONS

Micro-albuminuria’s association with increasing age substantiates the prevalence of renal dysfunction in older SCD patients. Furthermore, the selective and protective effect of the co-inheritance of alpha-thalassemia with macro-albuminuria in the context of SCD was replicated. The unique finding of the present research that have shown a significant association of APOL1 G1/G2 variants with ACR suggests the possible utility of these polymorphisms to be used as proxies for early detection of kidney dysfunction in SCD, requiring further investigation to validate its usefulness as a predictive biomarker of kidney disease. Our results further support the growing body of work supporting the role of APOL1 and not MYH9 as being significantly associated with risk of development and progression of kidney dysfunction of various etiologies in individuals of African descent (Genovese et al., 2010a; Kasembeli et al., 2015a; Kopp et al., 2011; Saraf et al., 2015). The near significant association of APOL1 G1/G2 and rs743811 with eGFR, another early indicator of glomerular dysfunction, requires further investigation into their utility as markers for screening individuals at risk for kidney damage. The lack of significant results associated with crude micro-albuminuria may be due to the high prevalence of micro-albuminuria in our cohort, or that the SNPs analysed may be an inappropriate choice for tagging our phenotype of choice in the population studied. These results also suggest that specific variants may contribute to the development of some forms of kidney dysfunction, but not others. This highlights the need for additional studies that include more adult SCD patients in order to provide further insight into the roles that each of these variants play in the development of renal dysfunction and kidney disease in both SCD and general populations.

In addition to clinical factors used to predict the development of sickle nephropathy, it is important to consider the role that biomarkers may play in early detection of renal dysfunction. Kidney injury molecule-1 (KIM-1) and N-acetyl-b-D-glucosaminidase (NAG) are biomarkers significantly associated with albuminuria in SCA patients (Sundaram et al., 2011),
which, in combination with genetic markers, may be used to identify patients at risk of developing sickle nephropathy.

4.10. LIMITATIONS OF THE STUDY

Limitations of this study include 1) the small number of adult patients and the correlated low proportion of SCD patients with macro-albuminuria, 2) the small sample of controls that were not properly age and sex matched with the SCD group; and 3) The hospital-based recruitment could have allowed disproportionate enrollment of sicker patients presenting micro-albuminuria, which may thwart the detection of significant associations. Moreover, while analysis of a single urine sample by first morning void is considered more reliable than spot urine samples in the diagnosis and monitoring of crude-albuminuria levels (Witte et al., 2009), timed urine analysis or measurement of urinary albumin excretion determine in a 24-hour collection period is considered the gold standard and should be employed in future studies. The cross-sectional nature of the present study is also considered a limitation, particularly in the analysis of the APOL1 variants and their association with the development of kidney disease. Recent reports indicate that APOL1 risk variants contributed to a more rapid progression of diabetic nephropathy (Parsa et al., 2013), indicating these variants may play a more important role in the progression, rather than initiation, of CKD. Future studies should therefore aim to recruit among the clinically stable SCD population to include individuals that display a wider range of clinical presentations of the disease and follow these individuals longitudinally to monitor the progression and development of CKD in African populations.
CHAPTER 5: CONCLUSION AND FUTURE PERSPECTIVES.

Our results show that Cameroonian children and adolescent SCD patients suffer from a high burden of micro-albuminuria, which anticipates that a high proportion of adult patients will be affected by renal disease in this setting, indicated by the significant association with increasing age that validates the progressive development of renal dysfunctions. Validation of the protective effect of alpha-thalassemia, particularly with regards to macro-albuminuria development, highlights the ability of genetic polymorphisms to modulate the development of kidney disease in SCD patients. These results indicate the potential of selected variants such as \textit{APOL1} G1/G2 and \textit{HMOX1} rs743811 to be used to identify patients at risk for development of early indicators of kidney disease, such as micro-albuminuria and glomerular hyperfiltration, while \textit{MYH9} SNPs and the \textit{HMOX1} dinucleotide promoter polymorphism (rs3074372) may be unsuitable to detect any early stage of renal impairment in a Cameroonian SCD population. \textit{HMOX1} rs3074372, however, is associated with the development of ESKD in the general Cameroonian population with various comorbidities, implying that genetic polymorphisms may contribute to the development of kidney dysfunction and failure through diverse functional pathways, and in different disease contexts, and it is anticipated that similar association would be found in an adult population of SCD patients with kidney disease. These results also indicate the need for further studies among adult patients to allow appropriate comparisons with African American SCD cohorts. Further studies are also required as there is limited evidence for the contribution of genetic polymorphisms influencing the vast clinical heterogeneity and disease severity of SCD, particularly in Africa.

More advanced techniques such as genome-wide association studies (GWAS), whole exome sequencing (WES), and whole genome sequencing (WGS) are also required to identify novel loci that may contribute to the predisposition to disease-related cardiovascular phenotypes.
ACKNOWLEDGEMENTS

The Wonkam Group; Gift Pule, Kamogelo Lebeko, Noluthando Manyisa, Tsephiso Masekoameng and Khuthala Mnika for their support and friendship.

Professor Collet Dandara for his assistance and advice.

Dr Emile Chimusa for his statistical support and recommendations.

The members of the division of Human Genetics for their words of insight and support.

Professor Ambroise Wonkam for his unwavering supervision, guidance and wisdom.

The University of Cape Town for funding.

Mom, Dad and my brother David for their undying love, encouragement and belief in my ability.
REFERENCES


Levey Andrew S, Becker Cassandra, Inker Lesley A, 2013. Glomerular Filtration Rate and


32, 811–816.


Osei-Yeboah CT, Rodrigues O, 2011. Renal Status of Children with Sickle Cell Disease in
Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J, Sklar P, de Bakker PIW, Daly MJ, Sham PC, 2007. PLINK: A tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81, 559–575.


163


Wonkam Ambrose, Rumaney Maryam Bibi, Ngo Bitoungui Valentina Josiane, Vorster Anna Alvera, Ramesar Raj, Ngogang Jeanne, 2014b. Coinheritance of sickle cell anemia and α-thalassemia delays disease onset and could improve survival in cameroonian’s patients (Sub-Saharan Africa). Am. J. Hematol. doi:10.1002/ajh.23711


APPENDIX 1: REAGENTS AND BUFFERS

TE Buffer
7 M Tris-Cl (Sigma Aldrich)
0.3 M EDTA (B&M Scientific)
pH 7.3

Sabax Water for injections
Adcock Ingram

Polymerase Chain Reaction

5M Betaine Solution
SIGMA-ALDRICH® (St. Louis MO, USA)
Lot no.: SLBH9055V

Dimethyl Sulfoxide (DMSO)
C₂H₆SO
MW: 78.13

dNTPs
Kapa Biosystems (Wilmington, Massachusetts, USA)

GoTaq® DNA Polymerase
500 units (5u/μl)
Promega (Madison WI, USA)
Lot no.: 0000160698

GoTaq® Buffer
5X colourless GoTaq®
Promega (Madison WI, USA)
Agarose Gel Electrophoresis

This was used for the resolution of DNA fragments as the negatively charged DNA phosphate (PO₄⁻) backbone causes DNA fragments to migrate to the positive electrode (anode). Agarose gel electrophoresis is a procedure used for the separation of biological molecules such as nucleic acids on the basis of charge, conformation and size. The pore size is inversely proportional to the percentage agarose (w/v) of the gel, therefore higher percentage gels are able to resolve smaller biological molecules.

SeaKem®LE Agarose LONZA (ME, USA)

Lot no.: 0000436471

1% (w/v) Agarose gel (100 ml)
1 g agarose
100 ml 1x TBE buffer
Dissolve agarose powder by microwave heating
Allow to cool, add 10 μl of SYBR® Safe DNA Gel Stain (Invitrogen™, Thermo Fisher Scientific) and pour into a gel electrophoresis casting tray with a well-comb.

2% (w/v) Agarose gel (150 ml)
3 g agarose
150 ml 1x TBE buffer
Dissolve agarose powder by microwave heating
Allow to cool, add 15 μl of SYBR® Safe DNA Gel Stain (Invitrogen™, Thermo Fisher Scientific) and pour into a gel electrophoresis casting tray with a well-comb.

2.5% (w/v) Agarose gel (150 ml)
3.75 g agarose
150 ml 1x TBE buffer
Dissolve agarose powder by microwave heating

Allow to cool, add 15 μl of SYBR® Safe DNA Gel Stain (Invitrogen™, Thermo Fisher Scientific) and pour into a gel electrophoresis casting tray with a well-comb.

6X DNA Loading Dye (Fermentas Life Sciences (USA))

0.03% Bromophenol blue 0.03% xylene cyanol FF
60% glycerol
60 mM EDTA
10 mM Tris-HCl (pH 7.6)

10X TBE Buffer (Tris-borate-EDTA [Ethylene-diamine tetra-acetic acid])

0.9 M Tris Base (216 g) (Melford Biolaboratories Ltd.)
0.9 M Boric Acid (110g) (Sigma Aldrich)
0.03 M EDTA (14.6g) (B&M Scientific)
Made up to 100 mL and distilled dH2O made the total volume up to 1 L.

1X TBE Buffer was used to dissolve the agarose, as well as the electrophoretic agent for agarose gel electrophoresis.

GeneRuler™ 100bp DNA Ladder Plus (Fermentas Life Sciences, USA)

#SM0321
Lot no: 00267549
SYBR® Safe DNA Gel Stain
Lot no.: 1567474
Invitrogen™, Thermo Fisher Scientific
1 µL used per 10 mL agarose gel

**PCR Product Clean Up**

FastAP Thermosensitive Alkaline Phosphatase (FastAP) (Thermo Scientific)
Lot no.: 00247968
1U/µL

Exonuclease I (ExoI)
Lot no.: 00254712
20U/µL

**Cycle sequencing**

70% Ethanol (EtOH)
350 mL absolute ethanol (100%) (Kimix) mixed with 150 mL of autoclaved H₂O.

Sodium Acetate (3 M)
246.1 g sodium acetate in 1 L autoclaved dH₂O

**BigDye® Terminator (v3.1)**

Austin TX, USA
Cycle Sequencing Kit
Lot no: 1106093
**SNaPSHOT® Sequencing**

**SNaPshot® Sequencing Mix**
Thermo Fisher Scientific Inc.

**GeneScan™ 120 Liz™ Size standard**
Applied Biosystems (California, USA)
Lot no.: 1506053

![GeneScan™ 120 Size Standard](image)

The size standard consists of nine single-stranded Liz dye-labelled DNA fragments between 15 and 120 base pairs.

**Hi-Di™ Formamide**
Lot no: 1506386
Applied Biosystems (Warrington, UK)

**Fragment Analysis**

**KAPA2G Fast HotStart ReadyMix**
KAPA Biosystems (Wilmington, USA)

**GeneScan™ 500 ROX™ Size standard**
The size standard consists of nine single-stranded ROX dye-labelled DNA fragments between 35 and 500 base pairs.

**TaqMan® SNP Genotyping**

**TaqMan® Universal PCR Master Mix**

Lot number: 1601107

Applied Biosystems by Life Technologies (Warrington, UK)
APPENDIX 2: EQUIPMENT AND MACHINES

DNA Quantification
NanoDrop® ND-1000 Spectrophotometer

ND-1000 Software v3.8.1.

PCR and agarose gel electrophoresis
Electrophoresis power pack (Hoefer Scientific Instruments, USA)
PS 500X DC Power Supply 500V, 400MA, 200W.

GeneAmp® (Applied Biosystems) PCR System 9700

Thermal Cycler T100TM (BIORAD®)

UVIPro UVIGold Transilluminator (UVItec Limited, UK)

UVIPro software (version 12.3 for Windows 1995-2005©)

Genotyping and Sequencing
ABI Prism 3130xl Genetic Analyzer (Applied Biosystems, UK)

GeneMapper® Software version 4.0.

Bio-Rad CFX96™ Real-Time system (Bio-Rad Laboratories, Hercules, CA, USA)

Bio-Rad CFX96™ Manager software (version 3.1.) (Bio-Rad Laboratories, Hercules, CA, USA)

BioEdit® Software Version 7.2.5.

TOUCHDOWN Heating Block (HYBAID® Limited, UK)
APPENDIX 3: PCR CONSTITUENTS AND PROTOCOL FOR SNPS RELATED TO KIDNEY DISEASE IN SCD PATIENTS.

Standard PCR reaction:

<table>
<thead>
<tr>
<th>REAGENTS</th>
<th>STOCK CONCENTRATION</th>
<th>REQUIRED FINAL CONCENTRATION</th>
<th>VOLUME (µL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sabax distilled H₂O</td>
<td>16.9</td>
<td></td>
<td>16.9</td>
</tr>
<tr>
<td>GoTaq® Buffer</td>
<td>5X</td>
<td>1X</td>
<td>5</td>
</tr>
<tr>
<td>dNTPs</td>
<td>5 mM</td>
<td>200 µM</td>
<td>1</td>
</tr>
<tr>
<td>Forward Primer</td>
<td>20 µM</td>
<td>10 pM</td>
<td>0.5</td>
</tr>
<tr>
<td>Reverse Primer</td>
<td>20 µM</td>
<td>10 pM</td>
<td>0.5</td>
</tr>
<tr>
<td>GoTaq® DNA Polymerase</td>
<td>5 U/µL</td>
<td>0.02 U/µL</td>
<td>0.1</td>
</tr>
<tr>
<td>Template DNA</td>
<td>100 ng/µL</td>
<td>100 ng</td>
<td>1</td>
</tr>
</tbody>
</table>

Total: 25
PCR CONSTITUENTS AND CYCLING PROCEDURES FOR THE AMPLIFICATION OF RS16996648

<table>
<thead>
<tr>
<th>REAGENTS</th>
<th>STOCK CONCENTRATION</th>
<th>REQUIRED FINAL CONCENTRATION</th>
<th>VOLUME (µL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sabax distilled H₂O</td>
<td></td>
<td></td>
<td>16.9</td>
</tr>
<tr>
<td>GoTaq® Buffer</td>
<td>5X</td>
<td>1X</td>
<td>5</td>
</tr>
<tr>
<td>dNTPs</td>
<td>5 mM</td>
<td>200 µM</td>
<td>1</td>
</tr>
<tr>
<td>Forward Primer</td>
<td>20 µM</td>
<td>10 pM</td>
<td>0.5</td>
</tr>
<tr>
<td>Reverse Primer</td>
<td>20 µM</td>
<td>10 pM</td>
<td>0.5</td>
</tr>
<tr>
<td>GoTaq® DNA Polymerase</td>
<td>5 U/µL</td>
<td>0.02 U/µL</td>
<td>0.1</td>
</tr>
<tr>
<td>Template DNA</td>
<td>100ng/µL</td>
<td>100 ng</td>
<td>1</td>
</tr>
<tr>
<td><strong>Total:</strong></td>
<td></td>
<td></td>
<td><strong>25</strong></td>
</tr>
</tbody>
</table>

CYCLING PARAMETERS:

![Cycling Parameters Diagram]
PCR CONSTITUENTS AND CYCLING PROCEDURES FOR THE AMPLIFICATION OF RS5750248

<table>
<thead>
<tr>
<th>REAGENTS</th>
<th>STOCK CONCENTRATION</th>
<th>REQUIRED FINAL CONCENTRATION</th>
<th>VOLUME (µL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sabax distilled H₂O</td>
<td></td>
<td></td>
<td>14.9</td>
</tr>
<tr>
<td>GoTaq® Buffer</td>
<td>5X</td>
<td>1X</td>
<td>5</td>
</tr>
<tr>
<td>dNTPs</td>
<td>5 mM</td>
<td>200 µM</td>
<td>1</td>
</tr>
<tr>
<td>DMSO</td>
<td>100%</td>
<td>8%</td>
<td>2</td>
</tr>
<tr>
<td>Forward Primer</td>
<td>20 µM</td>
<td>10 pM</td>
<td>0.5</td>
</tr>
<tr>
<td>Reverse Primer</td>
<td>20 µM</td>
<td>10 pM</td>
<td>0.5</td>
</tr>
<tr>
<td>GoTaq® DNA Polymerase</td>
<td>5 U/µL</td>
<td>0.02 U/µL</td>
<td>0.1</td>
</tr>
<tr>
<td>Template DNA</td>
<td>100ng/µL</td>
<td>100 ng</td>
<td>1</td>
</tr>
<tr>
<td><strong>Total:</strong></td>
<td></td>
<td></td>
<td><strong>25</strong></td>
</tr>
</tbody>
</table>

CYCLING PARAMETERS:

![Cycling Parameters Diagram]
PCR CONSTITUENTS AND CYCLING PROCEDURES FOR THE AMPLIFICATION OF MULTIPLEX A (RS1005570, RS16996672, RS73885319)

<table>
<thead>
<tr>
<th>REAGENTS</th>
<th>STOCK CONCENTRATION</th>
<th>REQUIRED FINAL CONCENTRATION</th>
<th>VOLUME (µL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sabax distilled H₂O</td>
<td></td>
<td></td>
<td>14.4</td>
</tr>
<tr>
<td>GoTaq® Buffer</td>
<td>5X</td>
<td>1X</td>
<td>5</td>
</tr>
<tr>
<td>dNTPs</td>
<td>5 mM</td>
<td>200 µM</td>
<td>1.5</td>
</tr>
<tr>
<td>Forward Primer (rs1005570)</td>
<td>20 µM</td>
<td>10 pM</td>
<td>0.5</td>
</tr>
<tr>
<td>Reverse Primer (rs1005570)</td>
<td>20 µM</td>
<td>10 pM</td>
<td>0.5</td>
</tr>
<tr>
<td>Forward Primer (rs16996672)</td>
<td>20 µM</td>
<td>10 pM</td>
<td>0.5</td>
</tr>
<tr>
<td>Reverse Primer (rs16996672)</td>
<td>20 µM</td>
<td>10 pM</td>
<td>0.5</td>
</tr>
<tr>
<td>Forward Primer (rs73885319)</td>
<td>20 µM</td>
<td>10 pM</td>
<td>0.5</td>
</tr>
<tr>
<td>Reverse Primer (rs73885319)</td>
<td>20 µM</td>
<td>10 pM</td>
<td>0.5</td>
</tr>
<tr>
<td>GoTaq® DNA Polymerase</td>
<td>5 U/µL</td>
<td>0.02 U/µL</td>
<td>0.1</td>
</tr>
<tr>
<td>Template DNA</td>
<td>100ng/µL</td>
<td>100 ng</td>
<td>1</td>
</tr>
</tbody>
</table>

Total: 25

CYCLING PARAMETERS:

- 95°C for 3:00 min
- 60.7°C for 0:30 min
- 72°C for 1:00 min
- 72°C for 5:00 min

GO TO 34X
## PCR Constituents and Cycling Procedures for the Amplification of Multiplex B (RS1557529, RS11912763)

<table>
<thead>
<tr>
<th>Reagents</th>
<th>Stock Concentration</th>
<th>Required Final Concentration</th>
<th>Volume (µL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sabax distilled H₂O</td>
<td></td>
<td></td>
<td>15.4</td>
</tr>
<tr>
<td>GoTaq® Buffer</td>
<td>5X</td>
<td>1X</td>
<td>5</td>
</tr>
<tr>
<td>dNTPs</td>
<td>5 mM</td>
<td>200 µM</td>
<td>1.5</td>
</tr>
<tr>
<td>Forward Primer (rs1557529)</td>
<td>20 µM</td>
<td>10 pM</td>
<td>0.5</td>
</tr>
<tr>
<td>Reverse Primer (rs1557529)</td>
<td>20 µM</td>
<td>10 pM</td>
<td>0.5</td>
</tr>
<tr>
<td>Forward Primer (rs11912763)</td>
<td>20 µM</td>
<td>10 pM</td>
<td>0.5</td>
</tr>
<tr>
<td>Reverse Primer (rs11912763)</td>
<td>20 µM</td>
<td>10 pM</td>
<td>0.5</td>
</tr>
<tr>
<td>GoTaq® DNA Polymerase</td>
<td>5 U/µL</td>
<td>0.02 U/µL</td>
<td>0.1</td>
</tr>
<tr>
<td>Template DNA</td>
<td>100ng/µL</td>
<td>100 ng</td>
<td>1</td>
</tr>
</tbody>
</table>

Total: 25

### Cycling Parameters:

- **Temperature:**
  - 95°C for 3.00 min
  - 95°C for 0.30 min
  - 58.5°C for 0.30 min
  - 72°C for 1.00 min
  - 72°C for 5.00 min
  - Go to 34X

- **GO TO 34X**
### PCR Constituents and Cycling Procedures for the Amplification of Multiplex C (RS5756129, RS8141189)

<table>
<thead>
<tr>
<th>REAGENTS</th>
<th>STOCK CONCENTRATION</th>
<th>REQUIRED FINAL CONCENTRATION</th>
<th>VOLUME (µL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sabax distilled H₂O</td>
<td></td>
<td></td>
<td>14.9</td>
</tr>
<tr>
<td>GoTaq® Buffer</td>
<td>5X</td>
<td>1X</td>
<td>5</td>
</tr>
<tr>
<td>dNTPs</td>
<td>5 Mm</td>
<td>200 µM</td>
<td>1.5</td>
</tr>
<tr>
<td>Betaine</td>
<td>5M</td>
<td>0.1M</td>
<td>0.5</td>
</tr>
<tr>
<td>Forward Primer (rs5756129)</td>
<td>20 µM</td>
<td>10 pM</td>
<td>0.5</td>
</tr>
<tr>
<td>Reverse Primer (rs5756129)</td>
<td>20 µM</td>
<td>10 pM</td>
<td>0.5</td>
</tr>
<tr>
<td>Forward Primer (rs8141189)</td>
<td>20 µM</td>
<td>10 pM</td>
<td>0.5</td>
</tr>
<tr>
<td>Reverse Primer (rs8141189)</td>
<td>20 µM</td>
<td>10 pM</td>
<td>0.5</td>
</tr>
<tr>
<td>GoTaq® DNA Polymerase</td>
<td>5 U/µL</td>
<td>0.02 U/µL</td>
<td>0.1</td>
</tr>
<tr>
<td>Template DNA</td>
<td>100 ng/µL</td>
<td>100 ng</td>
<td>1</td>
</tr>
<tr>
<td><strong>Total:</strong></td>
<td></td>
<td></td>
<td><strong>25</strong></td>
</tr>
</tbody>
</table>

### Cycling Parameters:

- 95°C for 3:00 minutes
- 95°C for 0:30 minutes
- 58.4°C for 0:30 minutes
- 72°C for 1:00 minutes
- 72°C for 5:00 minutes

Go to 34x
PCR CONSTITUENTS AND CYCLING PROCEDURES FOR THE AMPLIFICATION OF SNPs FOR FRAGMENT ANALYSIS (rs3074372 and rs71785313)

<table>
<thead>
<tr>
<th>REAGENTS</th>
<th>STOCK CONCENTRATION</th>
<th>REQUIRED FINAL CONCENTRATION</th>
<th>VOLUME (µL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sabax distilled H₂O</td>
<td></td>
<td></td>
<td>8.7</td>
</tr>
<tr>
<td>KAPA2G Fast HotStart ReadyMix</td>
<td></td>
<td>2X</td>
<td>1X</td>
</tr>
<tr>
<td>Forward primer</td>
<td>20 µM</td>
<td>10 pM</td>
<td>0.4</td>
</tr>
<tr>
<td>Reverse primer</td>
<td>20 µM</td>
<td>10 pM</td>
<td>2.4</td>
</tr>
<tr>
<td>Template DNA</td>
<td>10 ng/µL</td>
<td>5 ng</td>
<td>0.5</td>
</tr>
<tr>
<td><strong>Total:</strong></td>
<td></td>
<td></td>
<td><strong>20</strong></td>
</tr>
</tbody>
</table>

Cycling conditions:

[Diagram of cycling conditions with temperatures and times]
PCR CONSTITUENTS AND CYCLING PROCEDURES FOR THE AMPLIFICATION OF rs73885319 FOR DIRECT CYCLE SEQUENCING

<table>
<thead>
<tr>
<th>REAGENTS</th>
<th>STOCK CONCENTRATION</th>
<th>REQUIRED FINAL CONCENTRATION</th>
<th>VOLUME (µL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sabax distilled H₂O</td>
<td></td>
<td></td>
<td>11.9</td>
</tr>
<tr>
<td>GoTaq® Buffer</td>
<td>5X</td>
<td>1X</td>
<td>5</td>
</tr>
<tr>
<td>dNTPs</td>
<td>5 mM</td>
<td>200 µM</td>
<td>1</td>
</tr>
<tr>
<td>Forward Primer (rs73885319)</td>
<td>20 µM</td>
<td>10 pM</td>
<td>0.5</td>
</tr>
<tr>
<td>Reverse Primer (rs73885319)</td>
<td>20 µM</td>
<td>10 pM</td>
<td>0.5</td>
</tr>
<tr>
<td>GoTaq® DNA Polymerase</td>
<td>5 U/µL</td>
<td>0.02 U/µL</td>
<td>0.1</td>
</tr>
<tr>
<td>Template DNA</td>
<td>10 ng/µL</td>
<td>50 ng</td>
<td>5</td>
</tr>
<tr>
<td><strong>Total:</strong></td>
<td></td>
<td></td>
<td><strong>25</strong></td>
</tr>
</tbody>
</table>

Cycling procedures:

![Cycling procedure diagram](image-url)
PCR CONSTITUENTS AND CYCLING PROCEDURES FOR THE AMPLIFICATION OF rs60910145 FOR DIRECT CYCLE SEQUENCING

<table>
<thead>
<tr>
<th>REAGENTS</th>
<th>STOCK CONCENTRATION</th>
<th>REQUIRED FINAL CONCENTRATION</th>
<th>VOLUME (µL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sabax distilled H₂O</td>
<td></td>
<td></td>
<td>11.9</td>
</tr>
<tr>
<td>GoTaq® Buffer</td>
<td>5X</td>
<td>1X</td>
<td>5</td>
</tr>
<tr>
<td>dNTPs</td>
<td>5 mM</td>
<td>200 µM</td>
<td>1</td>
</tr>
<tr>
<td>Forward Primer (rs60910145)</td>
<td>20 µM</td>
<td>10 pM</td>
<td>0.5</td>
</tr>
<tr>
<td>Reverse Primer (rs60910145)</td>
<td>20 µM</td>
<td>10 pM</td>
<td>0.5</td>
</tr>
<tr>
<td>GoTaq® DNA Polymerase</td>
<td>5 U/µL</td>
<td>0.02 U/µL</td>
<td>0.1</td>
</tr>
<tr>
<td>Template DNA</td>
<td>10 ng/µL</td>
<td>50 ng</td>
<td>5</td>
</tr>
</tbody>
</table>

Total: 25

Cycling procedures:

\[\begin{array}{c}
\text{Temperature (°C)} \\
\hline
95°C & 95°C & 72°C & 72°C \\
95°C & 95°C & 72°C & 72°C \\
60°C & 72°C & 72°C & 72°C \\
34X & 34X & 34X & 34X \\
\end{array}\]
PCR CONSTITUENTS AND CYCLING PROCEDURES FOR THE AMPLIFICATION OF rs743811 FOR DIRECT CYCLE SEQUENCING

<table>
<thead>
<tr>
<th>REAGENTS</th>
<th>STOCK CONCENTRATION</th>
<th>REQUIRED FINAL CONCENTRATION</th>
<th>VOLUME (µL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sabax distilled H₂O</td>
<td></td>
<td></td>
<td>11.9</td>
</tr>
<tr>
<td>GoTaq® Buffer</td>
<td>5X</td>
<td>1X</td>
<td>5</td>
</tr>
<tr>
<td>dNTPs</td>
<td>5 mM</td>
<td>200 µM</td>
<td>1</td>
</tr>
<tr>
<td>Forward Primer (rs743811)</td>
<td>20 µM</td>
<td>10 pM</td>
<td>0.5</td>
</tr>
<tr>
<td>Reverse Primer (rs743811)</td>
<td>20 µM</td>
<td>10 pM</td>
<td>0.5</td>
</tr>
<tr>
<td>GoTaq® DNA Polymerase</td>
<td>5 U/µL</td>
<td>0.02 U/µL</td>
<td>0.1</td>
</tr>
<tr>
<td>Template DNA</td>
<td>10 ng/µL</td>
<td>50 ng</td>
<td>5</td>
</tr>
</tbody>
</table>

**Total:** 25

Cycling procedures:

- 95°C for 3:00
- 95°C for 0:30
- 58°C for 0:30
- 72°C for 1:00
- 72°C for 5:00
APPENDIX 4: SNAPSHOT® REACTION CONSTITUENTS, PROTOCOL AND CAPILLARY ELECTROPHORESIS

<table>
<thead>
<tr>
<th>REAGENTS</th>
<th>STOCK CONCENTRATION</th>
<th>REQUIRED FINAL CONCENTRATION</th>
<th>VOLUME (µL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sabax distilled H$_2$O</td>
<td></td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>SNaPshot® Primer</td>
<td>20 µM</td>
<td>2 µM</td>
<td>1</td>
</tr>
<tr>
<td>SNaPshot® mix</td>
<td></td>
<td></td>
<td>0.5</td>
</tr>
<tr>
<td>PCR Product</td>
<td>100ng</td>
<td>10ng</td>
<td>2.5</td>
</tr>
<tr>
<td><strong>Total:</strong></td>
<td></td>
<td></td>
<td><strong>10</strong></td>
</tr>
</tbody>
</table>

CYCLING PARAMETERS:
APPENDIX 5: TAQMAN® CONSTITUENTS AND PROTOCOL

TaqMan® SNP genotyping was performed for rs73885319 (APOL1), rs60910145 (APOL1) and rs743811 (HMOX1)

<table>
<thead>
<tr>
<th>REAGENTS</th>
<th>STOCK CONCENTRATION</th>
<th>REQUIRED FINAL CONCENTRATION</th>
<th>VOLUME (µL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TaqMan® Universal PCR MasterMix</td>
<td>2X</td>
<td>1X</td>
<td>5</td>
</tr>
<tr>
<td>TaqMan® Assay Working stock</td>
<td>20X</td>
<td>1X</td>
<td>0.5</td>
</tr>
<tr>
<td>Sabax distilled H2O</td>
<td></td>
<td></td>
<td>2.5</td>
</tr>
<tr>
<td>Template DNA</td>
<td>10 ng/µL</td>
<td>20 ng</td>
<td>2</td>
</tr>
<tr>
<td><strong>Total:</strong></td>
<td></td>
<td></td>
<td><strong>20</strong></td>
</tr>
</tbody>
</table>

Cycling conditions: Performed on the BioRad CFX96
APPENDIX 6: DIRECT CYCLE SEQUENCING

Steps performed to prepare the samples for capillary electrophoresis subsequent to the direct cycle sequencing reaction.

1. Added two microliters (2 µL) of sodium acetate (Appendix 1) and 2.5X the total sample reaction volume of absolute ethanol (100%) to each sample.
2. The samples were vortexed.
3. Samples were left at -20°C overnight.
4. The following day, the samples were centrifuged at 10 000g for ten minutes.
5. Discarded the supernatant.
6. A volume of 2.5X the reaction volume of 70% ethanol (100% ethanol diluted with distilled water) was added to each sample.
7. The samples were vortexed.
8. Centrifuged samples at 10 000g for ten minutes.
9. Discarded the supernatant.
10. Samples were left to air-dry for six hours.
11. Once dry, the samples were resuspended in 10 µL Hi-Di™ Formamide.
12. Five microliters (5 µL) of the sequencing product was added in combination with 8 µL Hi-Di™ Formamide to a 96-well microtiter plate (Appendix 1).
13. Loaded the samples into the ABI (Appendix 2).