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

DECLARATION

I, Amy Geard, 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 by candidate

Date: 27/09/2016

TABLE OF CONTENTS

LIST OF PUBLICATIONS FROM THIS WORK

1- Geard A, Pule G, Chelo D, Ngo Bitoungui V, and Wonkam A. Genetics of Sickle Cell Associated Cardiovascular Disease: An Expert Review with Lessons Learned in Africa. OMICS: A Journal of Integrative Biology 2016; 20(10) (in press). DOI: 10.1089/omi.2016.0125.
2- Geard A, Ngo Bitoungui V, Tchecha B, Pule G, Chimusa E, and Wonkam A. Association Of Variants In *APOL1*, *HMOX1* and Co-inheritance of 3.7 alpha-globin gene deletion With Micro-Albuminuria Among Sickle Cell Disease Patients From Cameroon, *in preparation*.

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 rss16996672), *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

LIST OF TABLES

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
Mg^{2+} - Magnesium ions
$MgCl_2$ – 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
T_a – 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).

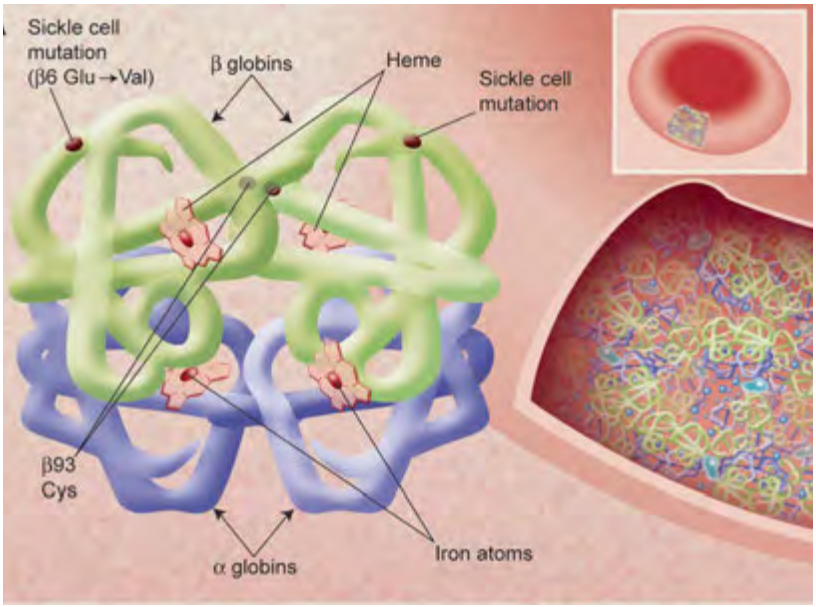


Figure 1: Structure of the adult hemoglobin molecule. The adult hemoglobin molecule consists of two α-globin (purple) and two β-globin (green) chains. The location of the sickle hemoglobin (HbS) mutation is depicted on β-globin chain, with the inset image displaying the high concentration of hemoglobin 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).

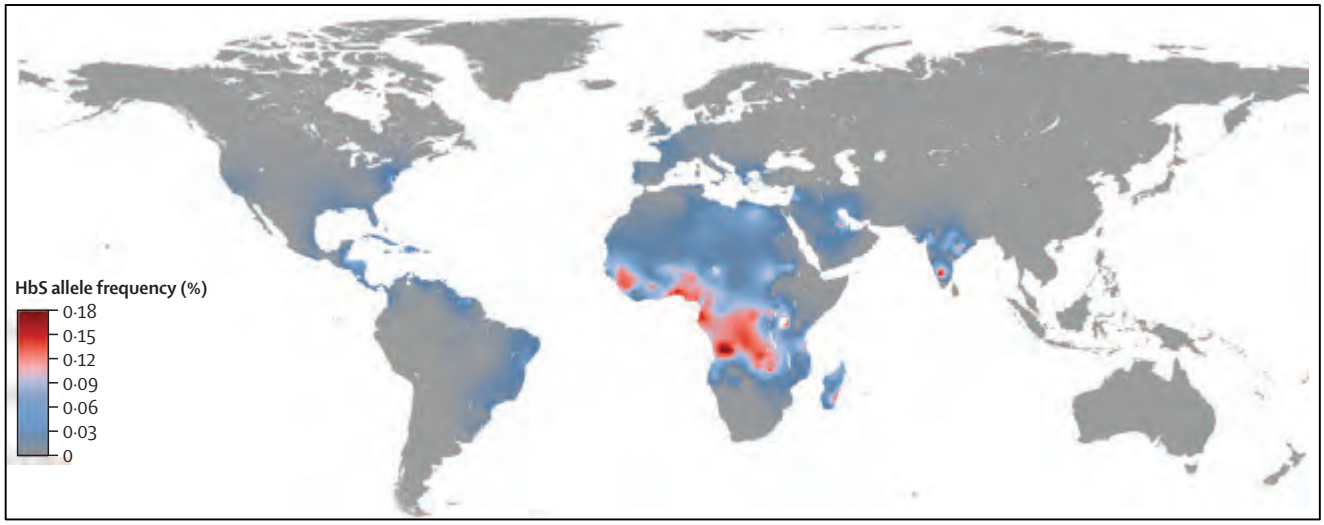


Figure 2: Global mean predicted HbS allele frequencies. The highest allele frequency corresponds to the malaria-endemic belt in sub-Saharan Africa due to the partial resistance conferred to heterozygotes for the sickle allele (HbS). Adapted from (Piel et al., 2013).

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.

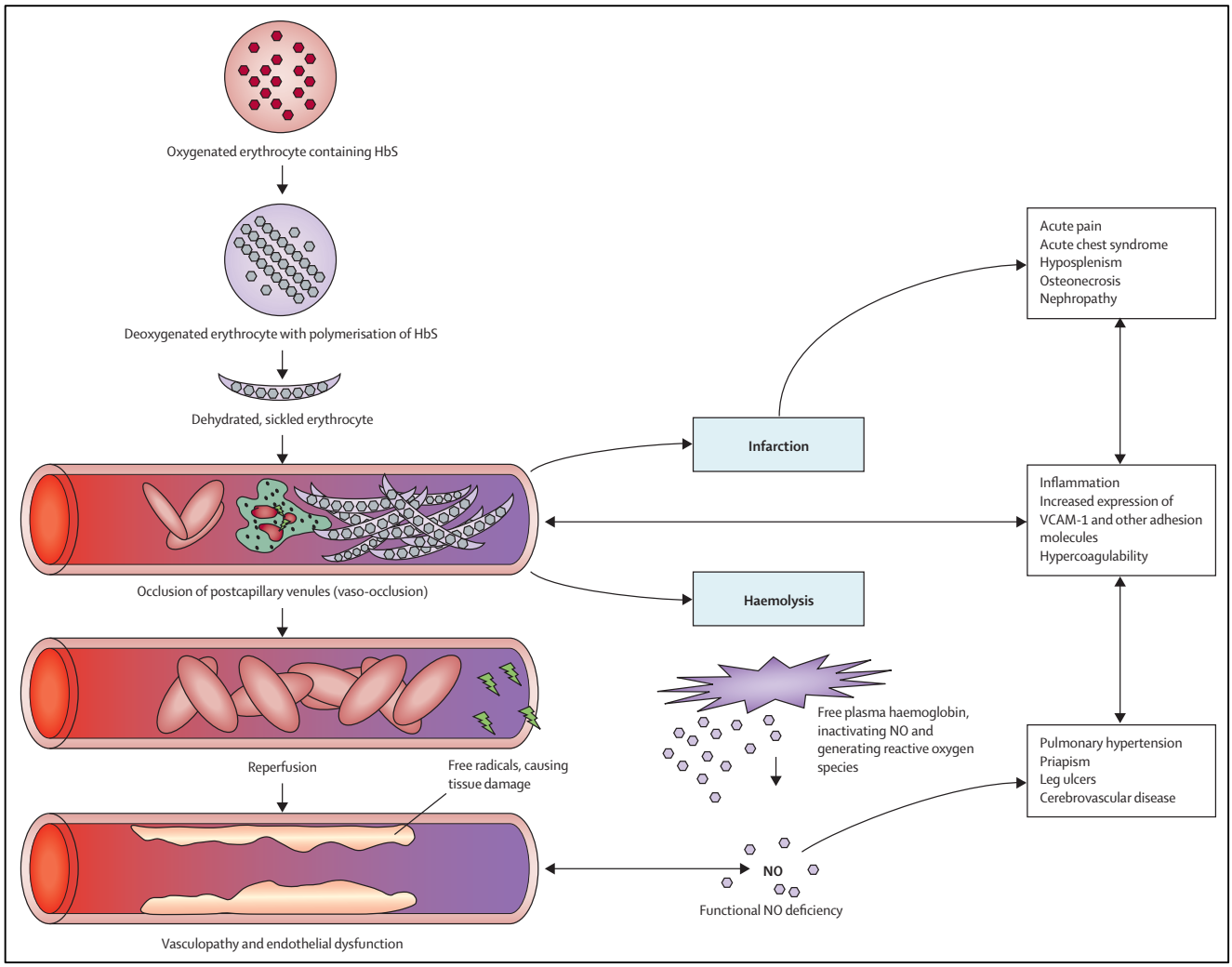


Figure 3: The pathophysiology of Sickle Cell Disease. The polymerization of hemoglobin during deoxygenation facilitates sickling of the erythrocytes, subsequently causing vaso-occlusion. This results in infarction and hemolysis, with reperfusion releasing free radicals. The repeated cycles of ischemia 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 micro-albuminuria 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

MYH9 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 macro-albuminuria. 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

Gene	SNPs	HGVS Nomenclature	Effect	References
MYH9	rs5750248	NC_000022.11:g.36306846T>C	Risk	(Ashley-Koch et al., 2011)
	rs8141189	NC_000022.11:g.36318665T>A	Risk	(Ashley-Koch et al., 2011)
	rs11912763	NC_000022.11:g.36288676G>A	Risk	(Ashley-Koch et al., 2011)
	rs16996648	NC_000022.11:g.36296706T>C	Risk	(Ashley-Koch et al., 2011)
	rs5756152	NC_000022.11:g.36316427A>G	Risk	(Ashley-Koch et al., 2011)
	rs16996672	NC_000022.11:g.36329925C>T	Risk	(Ashley-Koch et al., 2011)
	rs1557529	NC_000022.11:g.36309484A>G	Risk	(Ashley-Koch et al., 2011)
	rs1005570	NC_000022.11:g.36319229A>G	Risk	(Ashley-Koch et al., 2011)

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 2: Genetic variants in *APOL1* associated with the development of kidney disease in SCD

Gene	SNPs	HGVS Nomenclature	Effect	References
APOL1	rs73885319	NP_663318.1:p.Ser358Gly	Risk	(Ashley-Koch et al., 2011; Saraf et al., 2015)
	rs60910145	NP_663318.1:p.Ile400Met	Risk	(Saraf et al., 2015)
	rs71785313	NP_663318.1:p.Asn404_Tyr405del	Risk	(Saraf et al., 2015)

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).

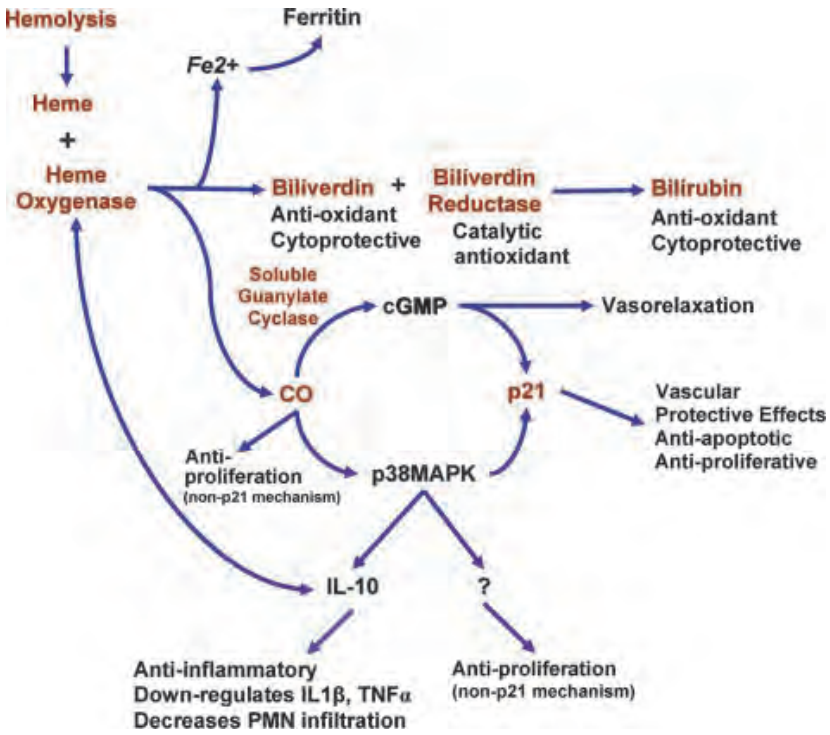


Figure 4: Role of heme oxygenase-1 and the actions of its products in response to repeated cycles of ischemia-reperfusion injury and hemolytic stress in SCD. Heme Oxygenase catalyzes the reduction of free heme to biliverdin, iron and Carbon Monoxide (CO). These molecules then proceed to induce effector 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 *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 *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). *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).

Table 3: Genetic variants in *HMOX1* associated with the development of kidney disease in SCD

Gene	SNPs	HGVS Nomenclature	Effect	References
HMOX1	rs3074372	NC_000022.11:g.35380894_35380895insGT	Protective/ Risk	(Saraf et al., 2015)
	rs743811	NC_000022.11:g.35396981T>C	Risk	(Saraf et al., 2015)

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 α-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.,

2009). 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) = $(175\times(Creatinine/10)^{-1.154})\times(Age^{-0.203})\times(0.742^{sex1})$

(2) Cockcroft-Gault (CG) = $\left(\left((140-Age)\times Weight\times\frac{10}{(Creatinine)\times 72}\right)\times sex2+\left((140-Age)\times Weight\times 10\times\frac{0.85}{(Creatinine)\times 72}\right)\times sex1\right)$

BSA = $((Weight^{0.425})\times((Height_{cm}/100)^{0.725})\times 0.20247)$

Calculate normalized Cockcroft-Gault:

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

(3) Compute the subcomponents for control purpose:

$$scr_k=\left(\left(\left(\frac{Creatinine}{10\times 0.7}\right)^{sex1}\right)\times\left(\left(\frac{Creatinine}{10\times 0.9}\right)^{sex2}\right)\right)$$

$$alph=(-\left((0.329^{sex1})\times(0.411^{sex2})\right))$$

$$\max\_sc=scr_k if\ scr_k\geq 1, otherwise=1$$

$$\min\_sc=scr_k\ if\ scr_k\leq 1, otherwise=1$$

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

The proportion of patients suffering from glomerular hyperfiltration (i.e. >130 ml/min/1.73m^2 for women and >140 ml/min/1.73m^2 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 m^2, 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.

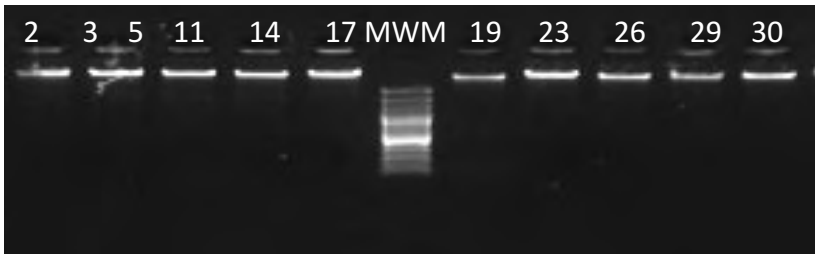


Figure 5: DNA Integrity gel of HbSS Cameroonian patients.
An electrophoretic gel for the resolution of HbSS Cameroonian patient DNA samples on a 1% (w/v) agarose gel. The lane numbers correspond to the respective patients from whom DNA was isolated. SYBR® Safe DNA Gel Stain was used to visualize the PCR products under UV light on the UVIGold Transilluminator. Lane MWM displays the GeneRuler^{TM}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 ($A_{260nm}/_{A280nm}$ 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

(http://www.ensembl.org/index.html) 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 (https://eu.idtdna.com/calc/analyzer), 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 H_2O 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.), rs16996672 (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	**Primers (rs5756129) (*MYH9*)**	
	Forward	**Reverse**
Sequence	5'-GGCTGCTTCTTGGAACC-3'	5'-GCCAACTACATATGAAGGTGCC-3'
Length (bases)	17	22
GC content (%)	58.8	50
Melting temperature (°C)	53.8	56
Molecular Weight (g/mole)	5177.4	6728.4

Expected PCR fragment size: 324bp

Table 4.2.

Properties	**Primers (rs11912763) (*MYH9*)**	
	Forward	**Reverse**
Sequence	5'-CCAGTTCTGCAGGATCCATG-3'	5'-CAGAGCGAGGAGAAGAAGAAG-3'
Length (bases)	20	21
GC content (%)	55	52.4
Melting temperature (°C)	55.7	54.7
Molecular Weight (g/mole)	6093	6611.4

Expected PCR fragment size: 268bp

Table 4.3.

Properties	**Primers (rs16996648) (*MYH9*)**	
	Forward	**Reverse**
Sequence	5'-AACGTGCCTGGTCAAGTC-3'	5'-CAGACCTCAGCGACCAGAT-3'
Length (bases)	18	19
GC content (%)	55.6	57.9
Melting temperature (°C)	555.4	56.6
Molecular Weight (g/mole)	5499.6	5766.8

Expected PCR fragment size: 420bp

Table 4.4.

Properties	**Primers (rs5750248) (*MYH9*)**	
	Forward	**Reverse**
Sequence	5’-CGTCGGACAGGAAAAGAGG-3’	5’-GAAAGACCCGTGCACCAG-3’
Length (bases)	19	18
GC content (%)	57.9	61.1
Melting temperature (°C)	55.4	56.3
Molecular Weight (g/mole)	5935.9	5502.6

Expected PCR fragment size: 579bp

Table 4.5.

Properties	**Primers (rs1557529) (*MYH9*)**	
	Forward	**Reverse**
Sequence	5’-CGTCAGCTTTGTAATCCACCTG-3’	5’-GCACGGTACTGTGTGGC-3’
Length (bases)	22	17
GC content (%)	50	64.7
Melting temperature (°C)	56	56.7
Molecular Weight (g/mole)	6661.4	5242.4

Expected PCR fragment size: 302bp

Table 4.6.

Properties	**Primers (rs8141189) (*MYH9*)**	
	Forward	**Reverse**
Sequence	5’-CCTTCTTTAGCAGGGCCAC-3’	5’-GCCAACATTGTGAAACCCC-3’
Length (bases)	19	19
GC content (%)	57.9	52.6
Melting temperature (°C)	55.9	54.8
Molecular Weight (g/mole)	5739.8	5741.8

Expected PCR fragment size: 478bp

Table 4.7.

Properties	**Primers (rs1005570) (*MYH9*)**	
	Forward	**Reverse**
Sequence	5’-CCTCAGAGTGTAAAAGCAGGTC-3’	5’-CTTGTGATCATGTGGACCAG-3’
Length (bases)	22	20
GC content (%)	50	50
Melting temperature (°C)	55.3	53.3
Molecular Weight (g/mole)	6768.5	6148

Expected PCR fragment size: 671bp

Table 4.8.

Properties	Primers (rs16996672) (*MYH9*)	
	Forward	**Reverse**
Sequence	5'-CAAACACACTCGAGGGCATG-3'	5'-GGTAAATTGGCTTCAGGAATTGG-3'
Length (bases)	20	23
GC content (%)	55	43.5
Melting temperature (°C)	56.5	54.5
Molecular Weight (g/mole)	6120	7158.7

Expected PCR fragment size: 447bp

Table 4.9.

Properties	Primers (rs73885319) (*APOL1*)	
	Forward	**Reverse**
Sequence	5'-CCAATCTTCAGTCAGTACCGC-3'	5'-GGTCCGCCTGCAGAATC -3'
Length (bases)	21	17
GC content (%)	52.4	64.7
Melting temperature (°C)	55.5	56.3
Molecular Weight (g/mole)	6326.2	5171.4

Expected PCR fragment size: 315bp

Table 4.10.

Properties	Primers (rs60910145) (*APOL1*)	
	Forward	**Reverse**
Sequence	5'-GGAAATGAGCAGAGGAGTCAA-3'	5'-TGTGCTCAGCTATGGAAATGC-3'
Length (bases)	22	21
GC content (%)	50	47.6
Melting temperature (°C)	55.4	55.7
Molecular Weight (g/mole)	6906.5	6461.2

Expected PCR fragment size: 666 bp

Table 4.11.

Properties	Primers (rs71785313) (*APOL1*)	
	Forward	**Reverse**
Sequence	5'-HEX-CTTCAGTCAGTACCGCATGC-3'	5'-AGTTTGCATTTTGTCCTGGC-3'
Length (bases)	20	20
GC content (%)	55	45
Melting temperature (°C)	56	54.3
Molecular Weight (g/mole)	6053	6105

Expected PCR fragment size: 384bp

Table 4.12.

Properties	**Primers (rs3074372) (*HMOX1*)**	
	Forward	**Reverse**
Sequence	5'-FAM-AGAGCCTGCAGCTTCTCAGA-3'	5'-ACAAAGTCTGGCCATAGGAC-3'
Length (bases)	20	20
GC content (%)	55	50
Melting temperature (°C)	58.4	54.5
Molecular Weight (g/mole)	6102	6135

Expected PCR fragment size: 67bp +/- repeats

Table 4.13.

Properties	**Primers (rs743811) (*HMOX1*)**	
	Forward	**Reverse**
Sequence	5'-TAGCAGCTGGGACTCTCCTG-3'	5'-ACTGAACTCTCTTCCCTGTTGC-3'
Length (bases)	20	22
GC content (%)	60	50
Melting temperature (°C)	58.4	56.9
Molecular Weight (g/mole)	6109	6612.3

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 (T_a) 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 (T_a) 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).

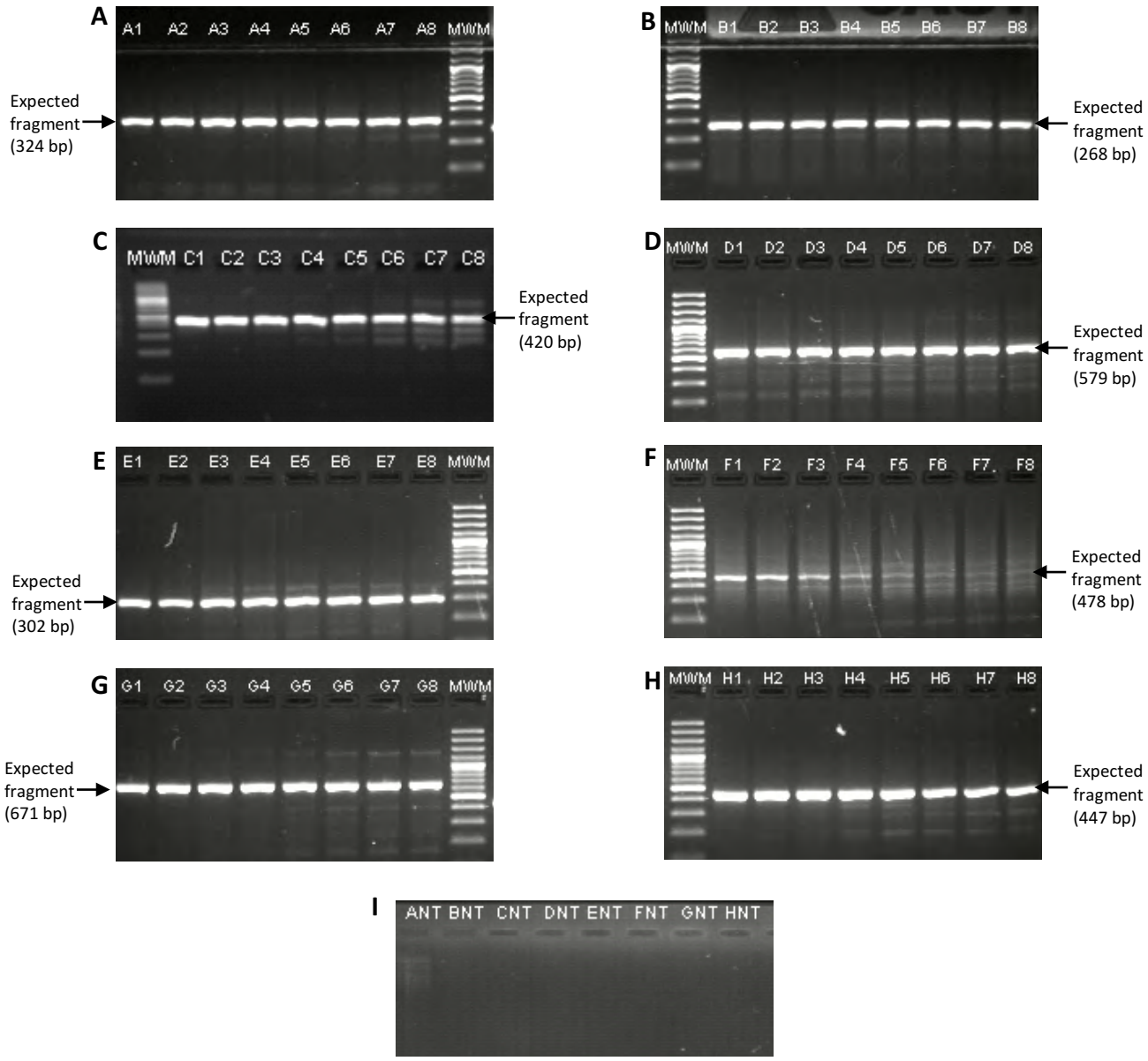


Figure 6: Initial temperature gradient for *MYH9* SNP amplification.

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^{TM}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.

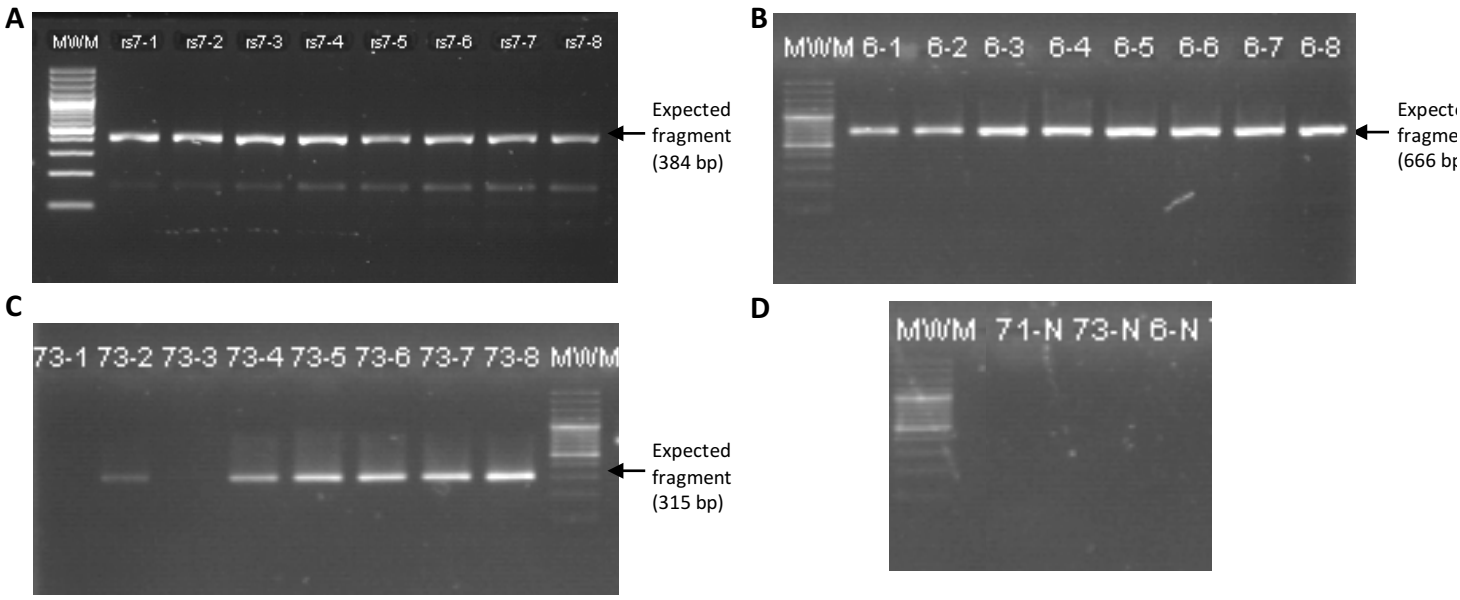


Figure 7: Initial temperature gradient for *APOL1* SNP amplification.

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^{TM}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. Lane 71-N, 73-N and 6-N represent the NTC's for rs71785313, rs73885319 and rs60910145, respectively.

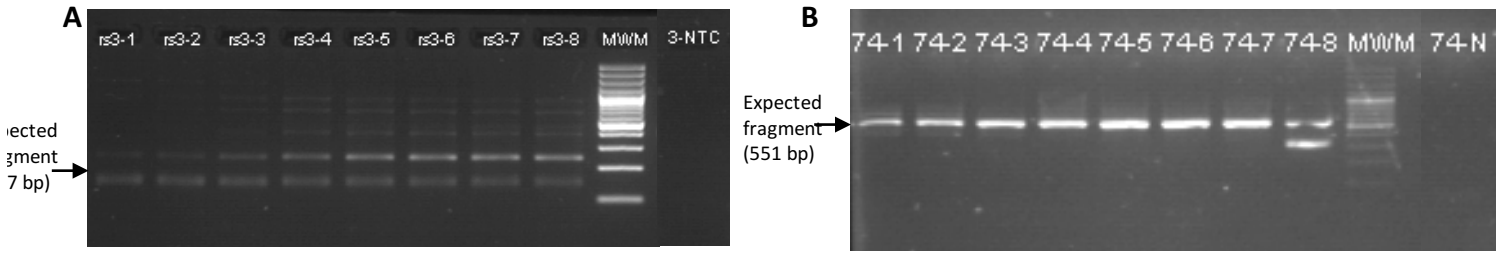


Figure 8: Initial temperature gradient for *HMOX1* SNP amplification.

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^{TM}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 $MgCl_2$ 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, $MgCl_2$ 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 T100™ (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.

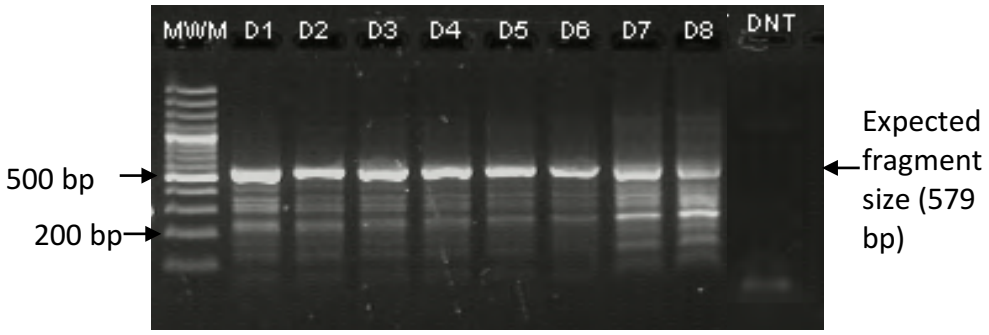


Figure 9: Secondary lower temperature gradient for rs5750248.

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^{TM} 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.

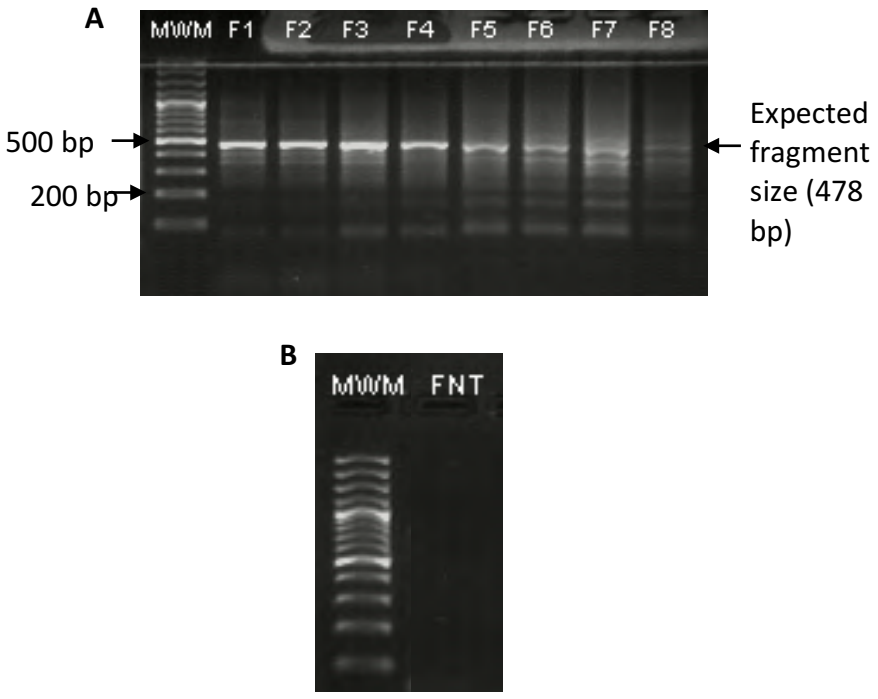


Figure 10: Secondary higher temperature gradient for rs8141189.

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^{TM} 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.

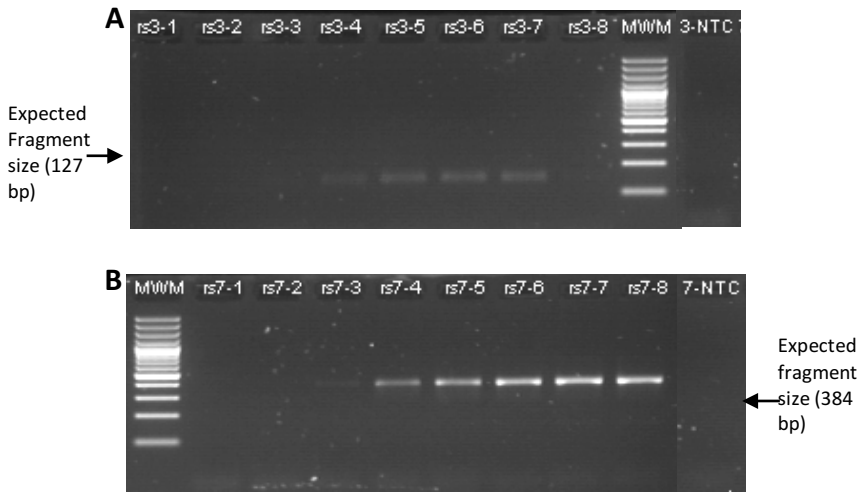


Figure 11: Secondary higher temperature gradient for rs3074372 (*HMOX1*) and rs71785313 (*APOL1*).

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^{TM} 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.

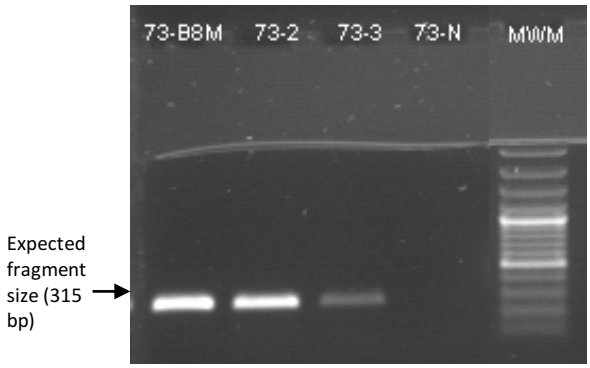


Figure 12: Reduced cycling conditions for rs73885319.

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^{TM} 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 ($MgCl_2$) 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.

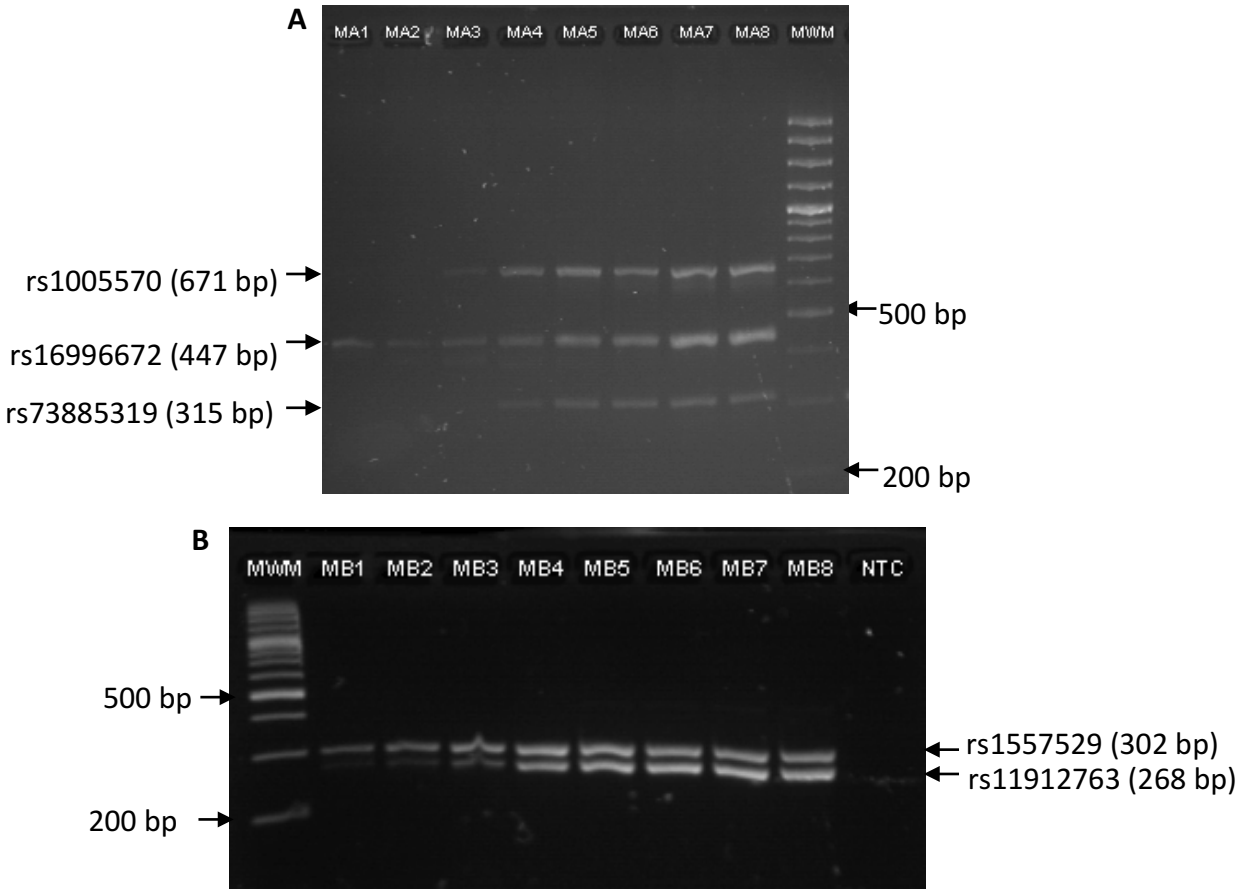


Figure 13: Temperature gradient for optimization of multiplexes.
An electrophoretic gel for the resolution of (A) Multiplex A, containing rs1005570 (671 bp), rs16996672 (447 bp) and rs73885319 (315 bp), and (B) Multiplex B, containing rs1557529 (302 bp) and rs11912763 (268 bp) on a (A) 2% (w/v) and (B) 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^{TM} 100bp Plus DNA molecular weight marker. Lanes MA1 to MA8 represent the annealing temperature gradient of 65°C-58°C (1 - 65°C, 2 - 64.5°C, 3 - 63.6°C, 4 - 62.3°C, 5 - 60.7°C, 6 - 59.4°C, 7 - 58.5°C, 8 – 58°C), and lanes MB1 to MB8 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 - 58.8°C, 6 - 56.9°C, 6 - 55.7°C, 8 - 55°C) using South African HbAA control DNA B3A. The NTC represents the non-template control, which was subjected to the lowest annealing temperature for each reaction.

2.2.4.2. $MgCl_2$ GRADIENT

The concentration of Magnesium Chloride ($MgCl_2$) in a polymerase chain reaction greatly influences the specificity of the enzymatic activity of the DNA *Taq* Polymerase. The Mg^{2+} 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.

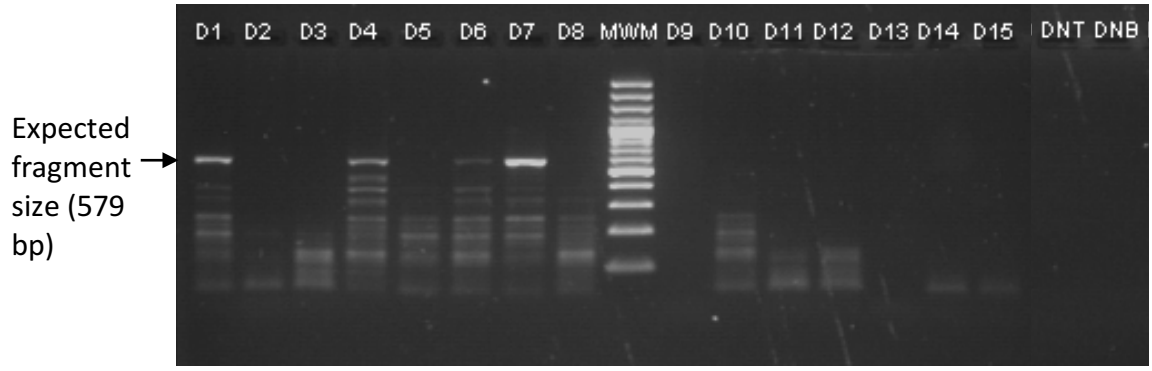


Figure 14: Magnesium chloride gradient for rs5750248 optimization.
An electrophoretic gel for the resolution of rs5750248 on a 2% (w/v) agarose gel. SYBR® Safe DNA Gel Stain was used to visualize the PCR products under UV light on the UVIGold Transilluminator. The expected size of the PCR product was 579 bp. Lane MWM displays the GeneRuler™ 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

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.

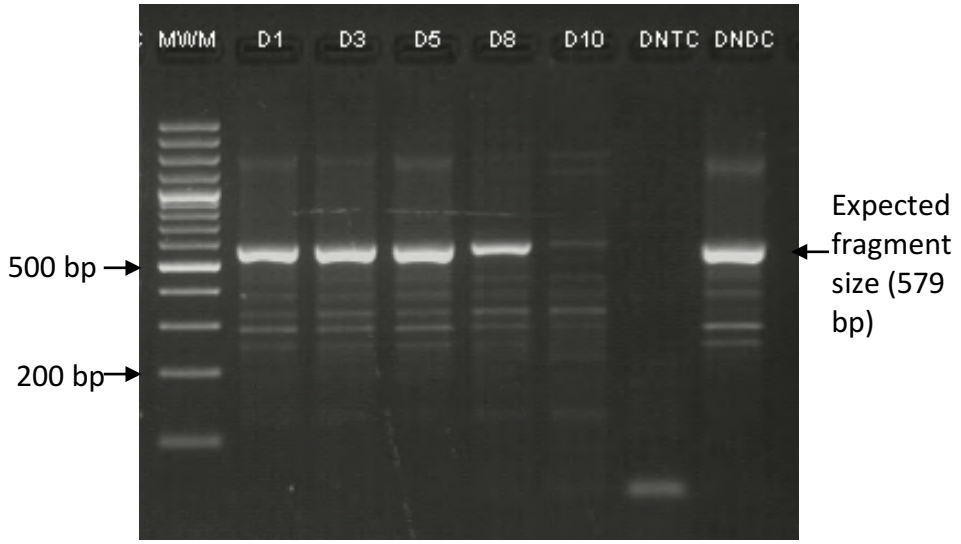


Figure 15: DMSO gradient for rs5750248 optimization.

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^{TM} 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.

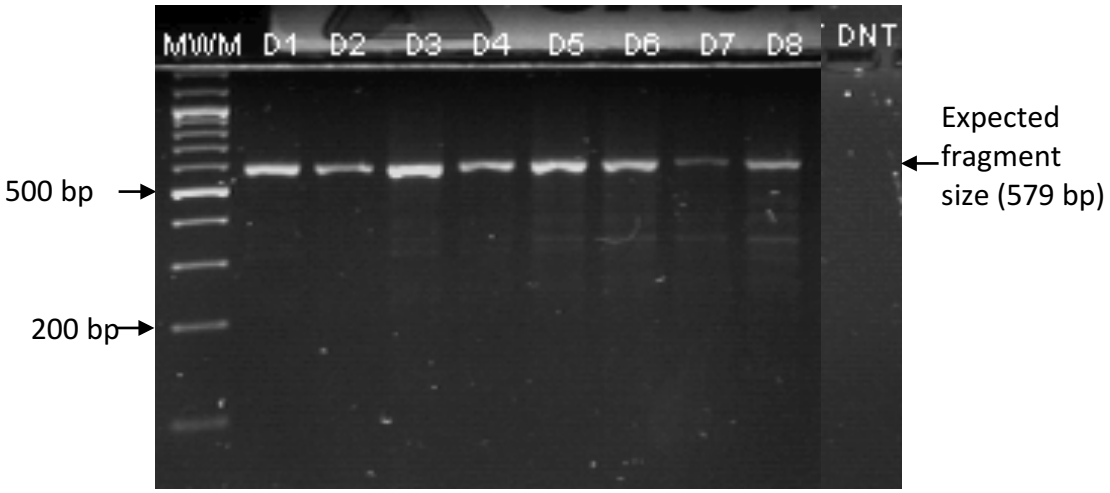


Figure 16: Temperature gradient with 8% DMSO for rs5750248 optimization.

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^{TM} 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*).

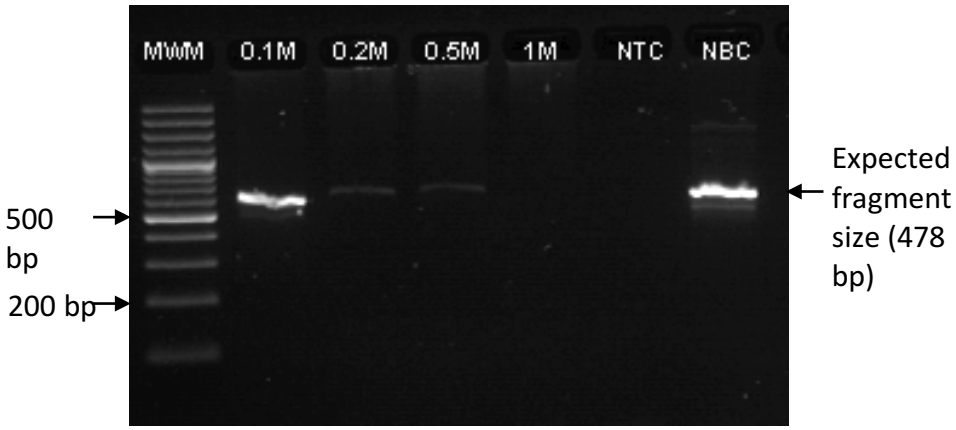


Figure 17: Betaine concentration gradient for rs8141189 optimization.
An electrophoretic gel for the resolution of rs8141189 on a 2% (w/v) agarose gel. SYBR® Safe DNA Gel Stain was used to visualize the PCR products under UV light on the UVIGold Transilluminator. The expected size of the PCR product was 478 bp. Lane MWM displays the GeneRuler^{TM} 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 T_a 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.

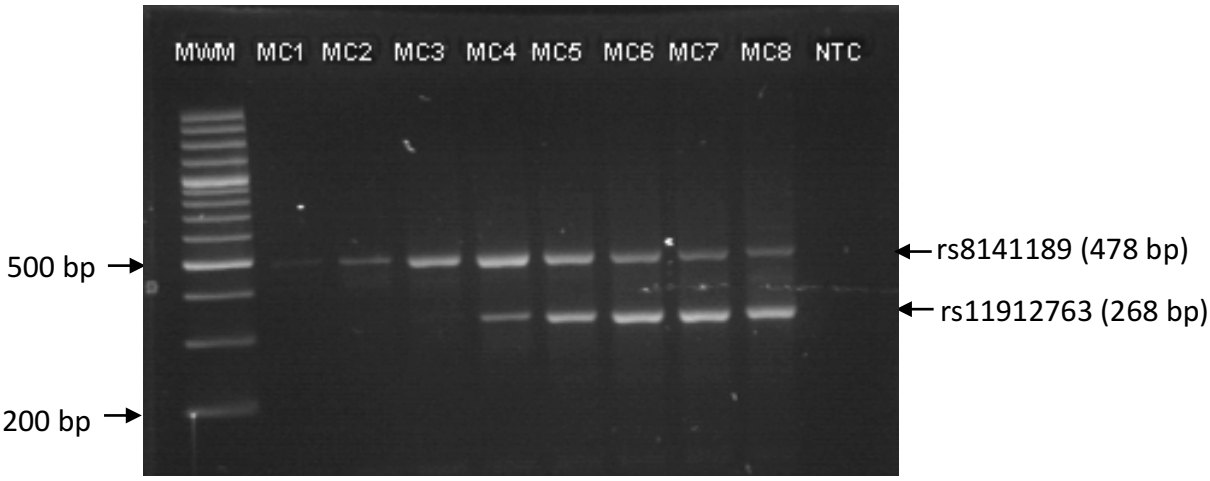


Figure 18: Temperature gradient for Multiplex C optimization.
An electrophoretic gel for the resolution of Multiplex C containing rs5756129 (324 bp) and rs8141189 (478 bp) on a 2% (w/v) agarose gel. SYBR® Safe DNA Gel Stain was used to visualize the PCR products under UV light on the UVIGold Transilluminator. Lane MWM displays the GeneRuler^{TM} 100bp Plus DNA molecular weight marker. Lanes MC1 to MC8 represent the 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) annealing temperature gradient with 0.1M Betaine using South African HbAA control DNA B3A. Lane NTC is the non-template control. The target fragment was efficiently and specifically amplified at 58.5°C (Temperature 5) and was subsequently used for the protocol for Multiplex C amplification.

2.2.4.5. SUMMARY OF THE OPTIMISATION PROCESS

The PCR optimization process produced optimal protocols for two single plexes, 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 T_a of 55°C for rs5750248, (ii) T_a of 60°C for rs16996646, (iii) T_a of 60.7°C for multiplex A (rs1005570, rs16996672, rs73885319) (iv) T_a of 58.8°C for Multiplex B (rs11912763, rs1557529) and (v) 0.1M Betaine and a T_a 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 T_a 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.

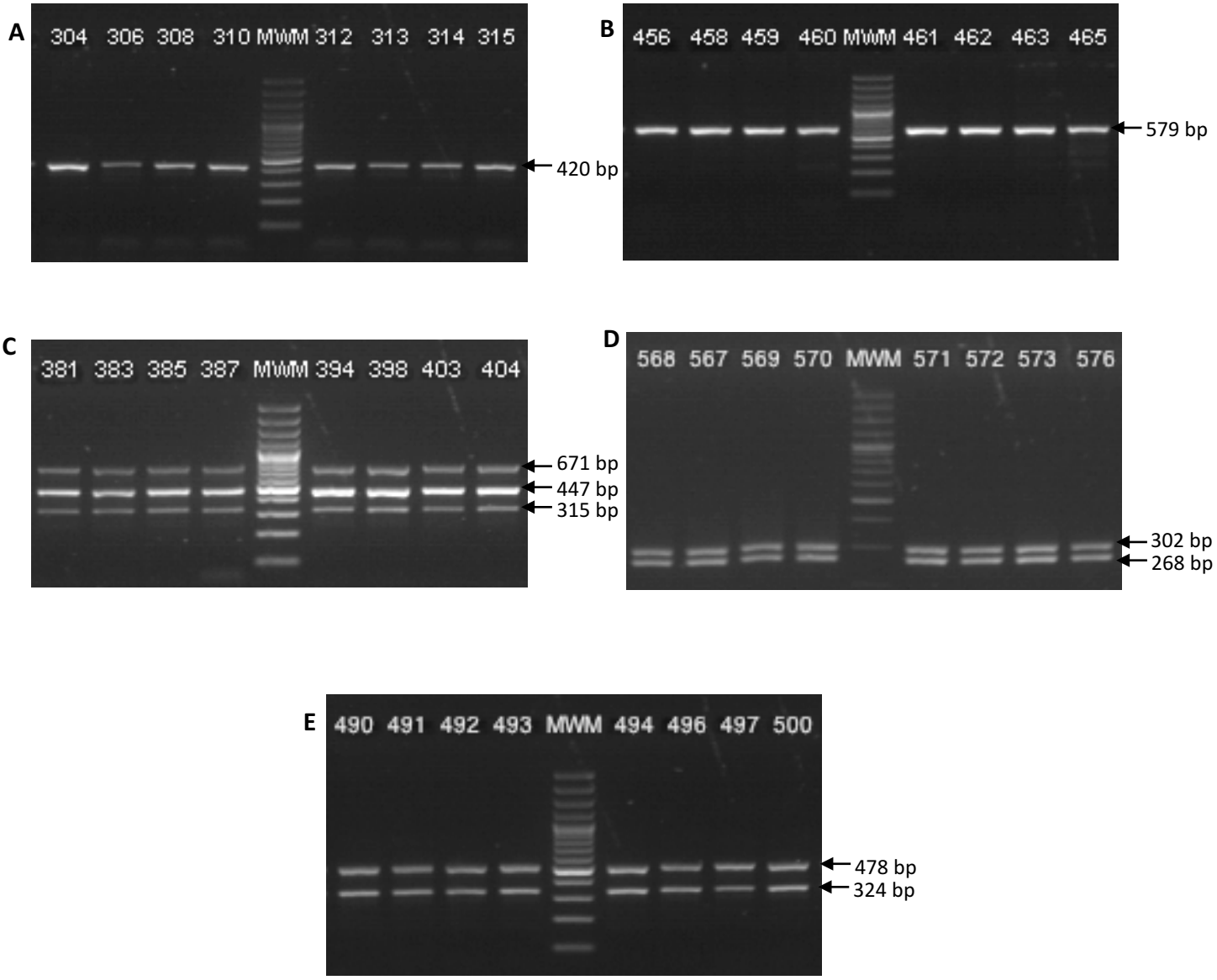


Figure 19: Amplification of SCD patient samples for each PCR protocol for *MYH9* SNPs.
Electrophoretic gels for the resolution of (A) rs16996648 (420 bp), (B) rs5750258 (579 bp), (C) Multiplex A, containing rs1005570 (671 bp), rs16996672 (447 bp) and rs73885319 (315 bp), (D) Multiplex B, containing rs1557529 (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^{TM} 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 (Invitrogen™, 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 (*Exo*I) 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 *Exo*I 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 *Exo*I, 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 T_m 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 3130*xl* 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 H_2O.

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.

Properties	**Primer (rs5756129) (*MYH9*)**
Direction	Forward
Sequence	5’-TTCTGTGTTTTATGTACCCAGTCATACACCGTTT-3’
Complement	5'-AAACGGTGTATGACTGGGTACATAAAACACAGAA-3'
Length (bases)	34
GC content (%)	38.2
Melting temperature (°C)	60.8
Molecular Weight (g/mole)	10339.7

Expected SNaPshot product size: 35 bp

Table 5.2.

Properties	**Primer (rs11912763) (*MYH9*)**
Direction	Forward
Sequence	5’-AACAGAAGCCTGCGT-3’
Complement	5'-ACGCAGGCTTCTGTT-3'
Length (bases)	15
GC content (%)	53.3
Melting temperature (°C)	51.1
Molecular Weight (g/mole)	4586

Expected SNaPshot product size: 16 bp

Table 5.3.

Properties	Primer (rs16996648) (*MYH9*)
Direction	Reverse
Sequence	5’-TCCACTGTTGTCTTAAAAGAGCCATTTATTTCAT-3’
Complement	5'-ATGAAATAAATGGCTCTTTTAAGACAACAGTGGA-3'
Length (bases)	34
GC content (%)	32.4
Melting temperature (°C)	58.7
Molecular Weight (g/mole)	10356.8

Expected SNaPshot product size: 35 bp

Table 5.4.

Properties	Primer (rs5750248) (*MYH9*)
Direction	Forward
Sequence	5’-GCCATGCATCTGCCA-3’
Complement	5'-TGGCAGATGCATGGC-3'
Length (bases)	15
GC content (%)	60
Melting temperature (°C)	53.0
Molecular Weight (g/mole)	4513

Expected SNaPshot product size: 16 bp

Table 5.5.

Properties	Primer (rs1557529) (*MYH9*)
Direction	Reverse
Sequence	5’-TCTTCACGTTTTCTGACTTTTCCA-3’
Complement	5'-TGGAAAAGTCAGAAAACGTGAAGA-3'
Length (bases)	24
GC content (%)	37.5
Melting temperature (°C)	54.9
Molecular Weight (g/mole)	7210.7

Expected SNaPshot product size: 25 bp

Table 5.6.

Properties	Primer (rs8141189) (*MYH9*)
Direction	Reverse
Sequence	5’-GGGACCTGCTTTTACACTCTGAGG-3’
Complement	5'-CCTCAGAGTGTAAAAGCAGGTCCC-3'
Length (bases)	24
GC content (%)	54.2
Melting temperature (°C)	59.2
Molecular Weight (g/mole)	7359.8

Expected SNaPshot product size: 25 bp

Table 5.7.

Properties	Primer (rs1005570) (*MYH9*)
Direction	Forward
Sequence	5’-CAAGGACATAATTTCAAGATCTCGACATAGTCAC-3’
Complement	5'-GTGACTATGTCGAGATCTTGAAATTATGTCCTTG-3'
Length (bases)	34
GC content (%)	38.2
Melting temperature (°C)	58.4
Molecular Weight (g/mole)	10402.8

Expected SNaPshot product size: 35 bp

Table 5.8.

Properties	Primer (rs16996672) (*MYH9*)
Direction	Reverse
Sequence	5’-CGTGTGCAGCTGTGT-3’
Complement	5'-ACACAGCTGCACACG-3'
Length (bases)	15
GC content (%)	60
Melting temperature (°C)	52.9
Molecular Weight (g/mole)	4615

Expected SNaPshot product size: 16 bp

Table 5.9.

Properties	Primer (rs73885319) (*APOL1*)
Direction	Reverse
Sequence	5’-CTACATCCAGCACAAGAAAGAAGC-3’
Complement	5'-GCTTCTTTCTTGTGCTGGATGTAG-3'
Length (bases)	24
GC content (%)	45.8
Melting temperature (°C)	56.1
Molecular Weight (g/mole)	7332.8

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 3130*xl* 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.

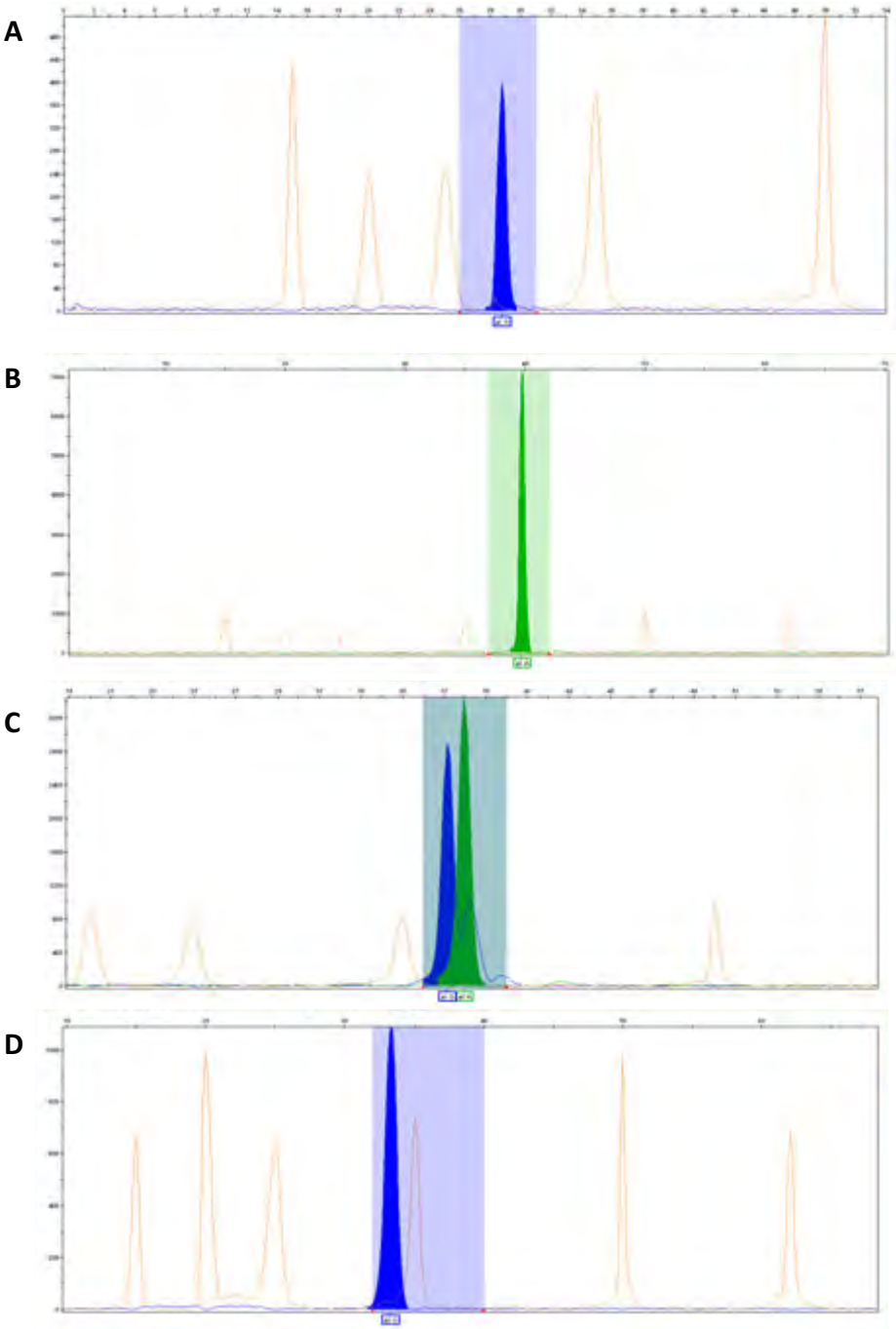


Figure 20: SNaPshot genotyping of SNPs rs11912763, rs16996648 and rs1005570 and rs16996672 for patient 472.

Electropherogram produced using the ABI Prism 3130*xl* 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.

2.2.8. FRAGMENT ANALYSIS

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 ABI3130*xl* (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 ExoI (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 3130*xl* , 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 3130*xl* Genetic Analyzer (Applied Biosystems, California, USA) for patient 65.

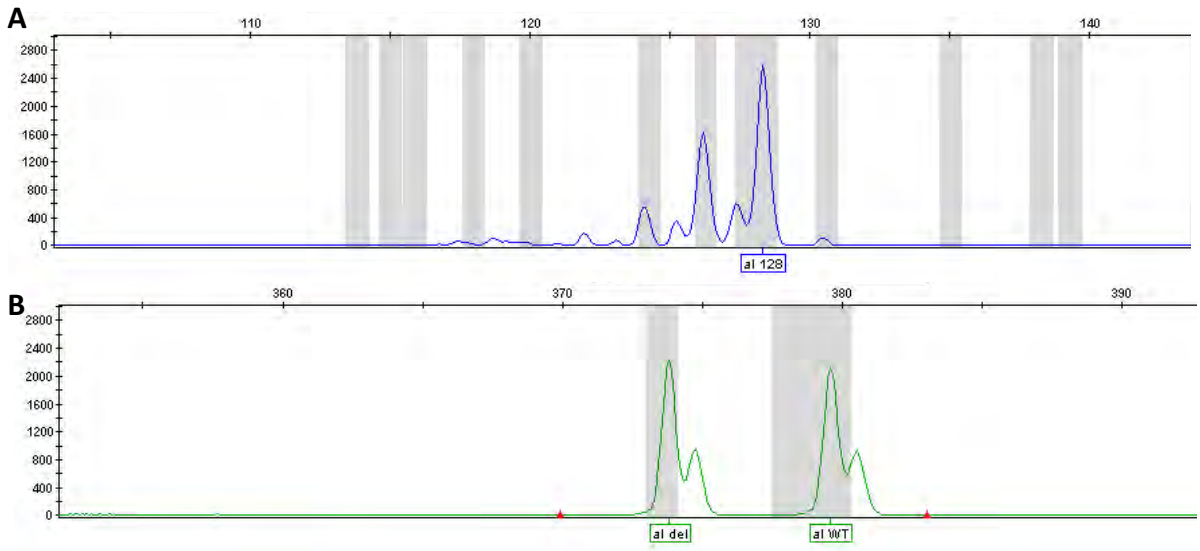


Figure 21: Fragment Analysis of rs3074372 and rs71785313 polymorphisms for SCD patient 65.

Electropherogram produced using the ABI Prism 3130*xl* 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 (*APOL1*), rs73885319 (*APOL1*) and rs743811 (*MYH9*). Rs73885319 was initially sequenced using SNaPshot®, 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).

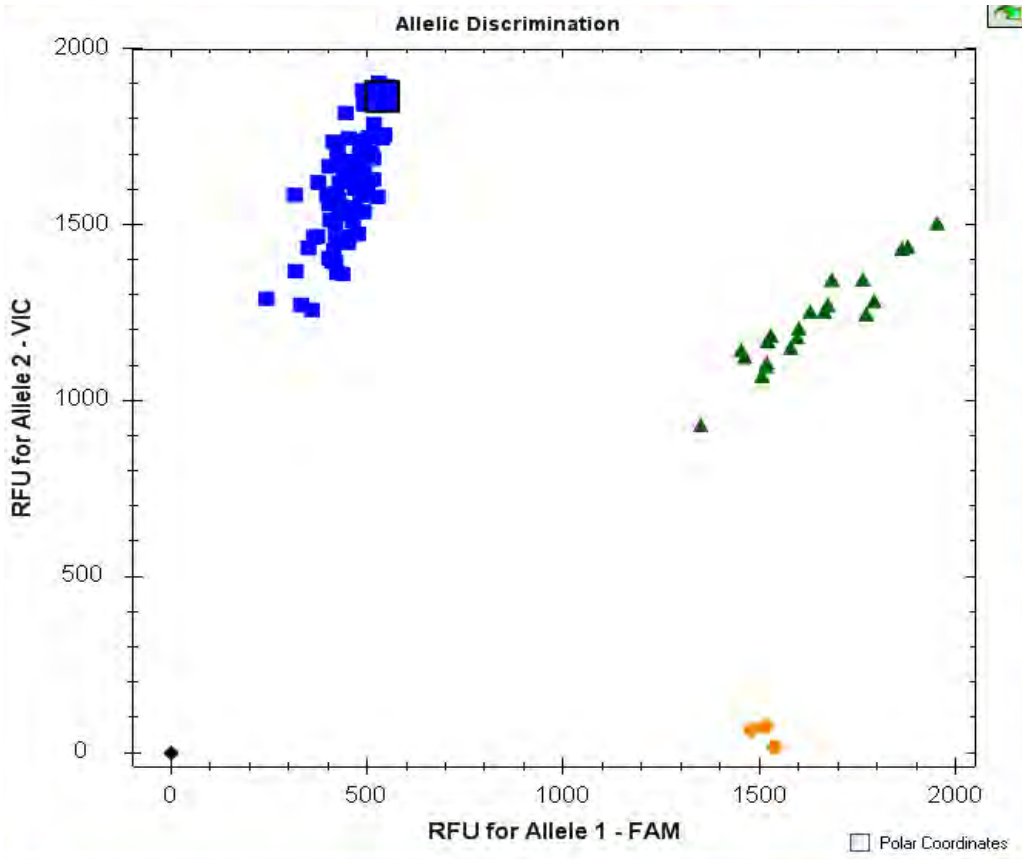


Figure 22: TaqMan SNP Genotyping plot of SCD cohort plate 3 for rs73885319.

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.

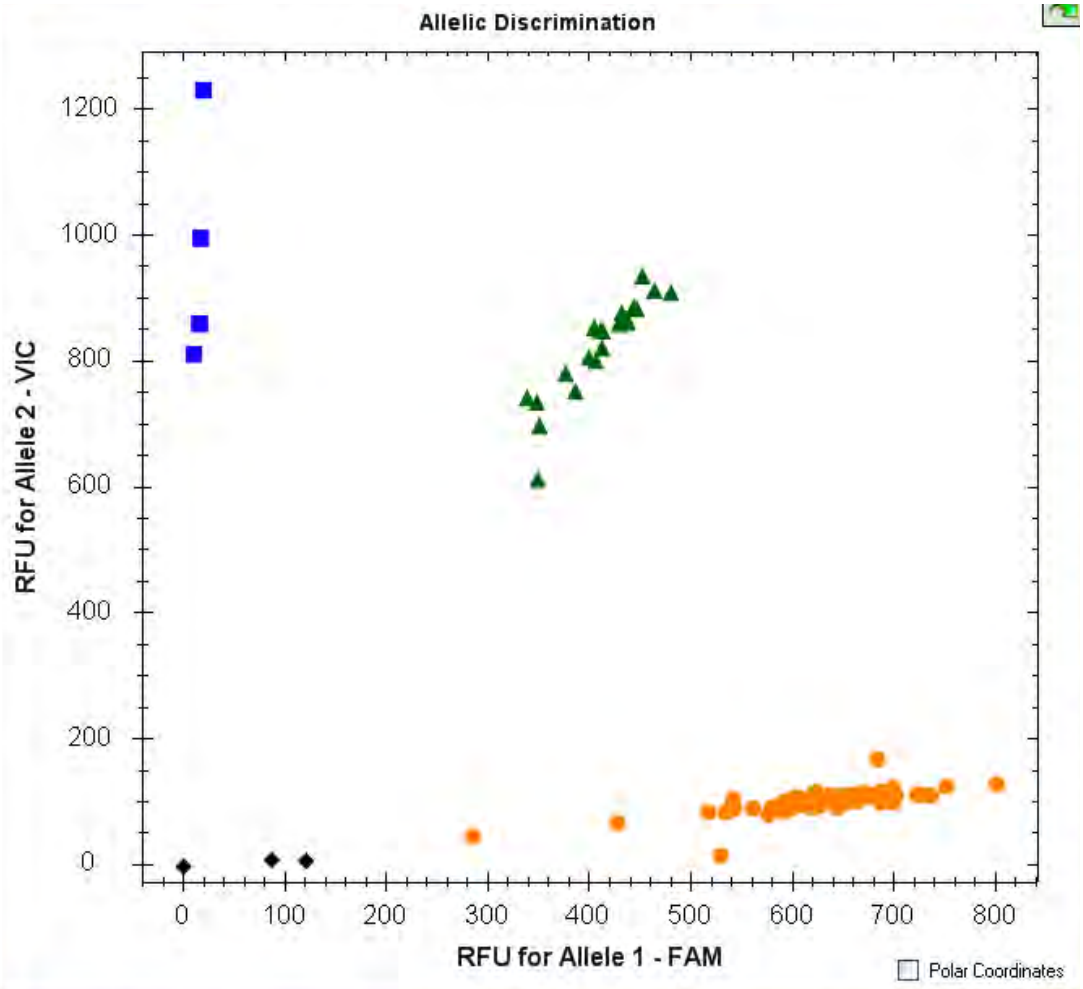


Figure 23: TaqMan SNP Genotyping plot of SCD cohort plate 3 for rs60910145.

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 3130*xl* 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 3130*xl* 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.

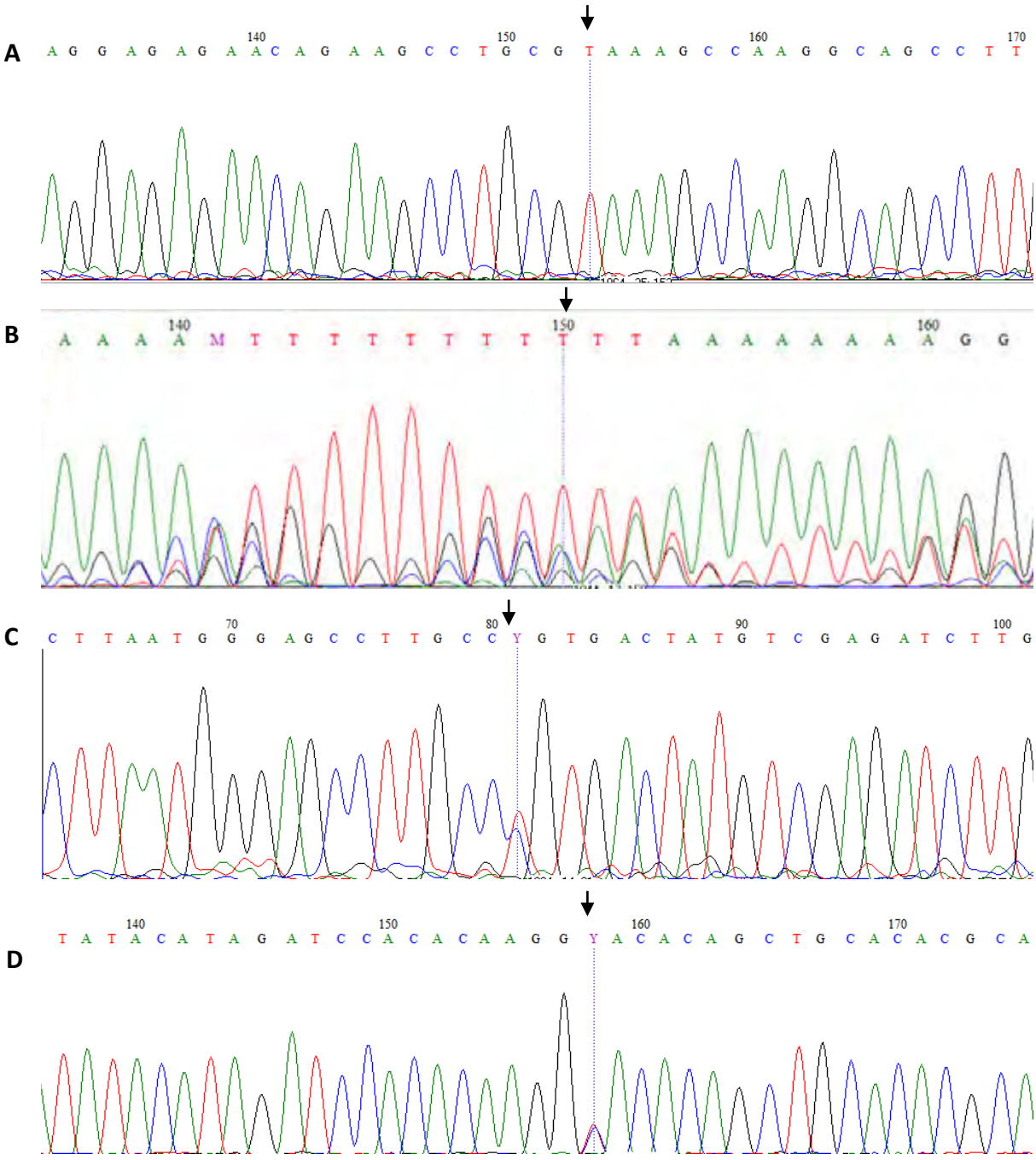


Figure 24: Successful direct cycle sequencing validation of *MYH9* SNPs.
Electropherogram produced using the ABI Prism 3130*xl* Genetic Analyzer (Applied Biosystems) analyzed using the BioEdit® software program (version 7.0.5.2; 6/5/05) for the base calling of the genotype of (A) patient 606 for rs11912763, (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.

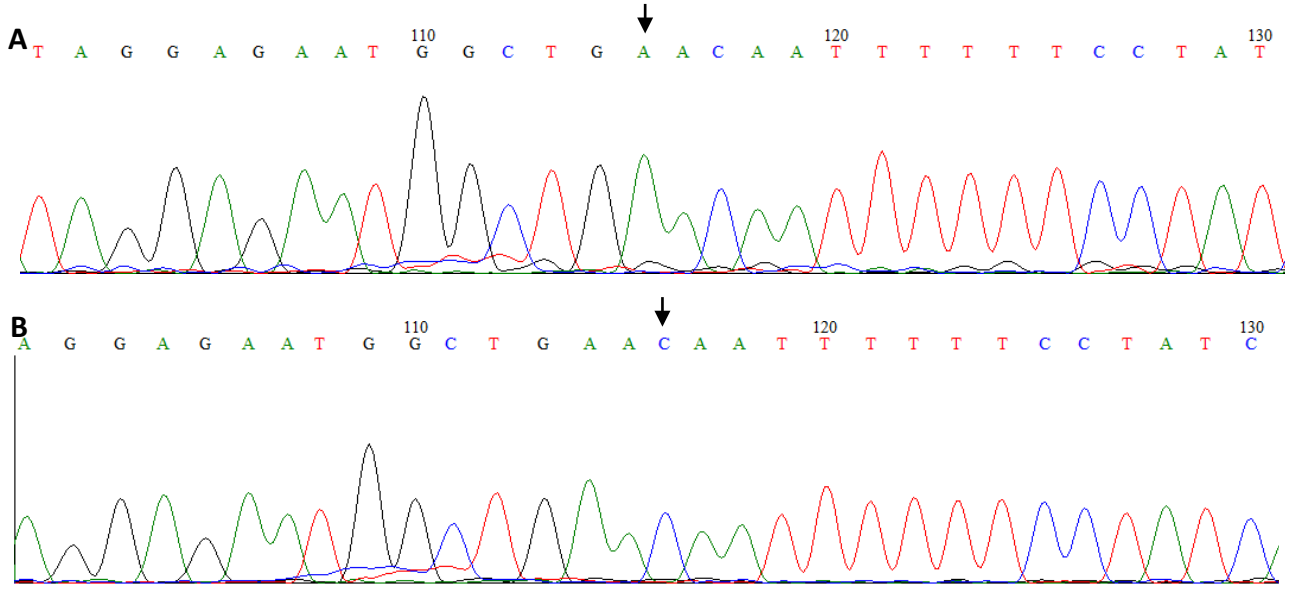


Figure 25: Successful direct cycle sequencing validation of *HMOX1* SNP rs743811.

Electropherogram produced using the ABI Prism 3130*xl* 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.

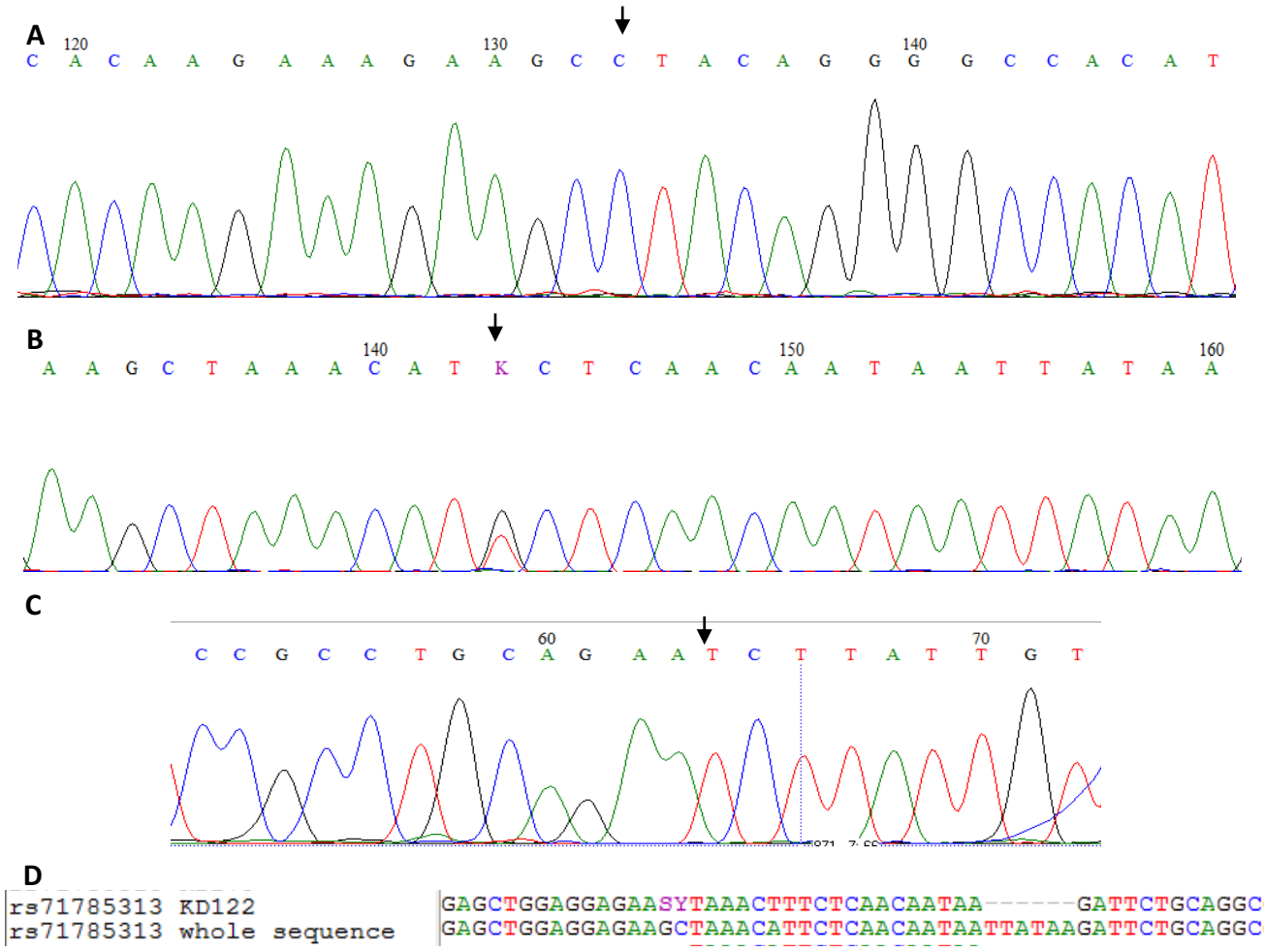


Figure 26: Successful direct cycle sequencing validation of *APOL1* G1/G2 allele.
Electropherogram produced using the ABI Prism 3130*xl* Genetic Analyzer (Applied Biosystems) analyzed using the BioEdit® software program (version 7.0.5.2; 6/5/05) for the base calling of the genotype of (A) SCD patient 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.

```
>22 dna:chromosome chromosome:GRCh38:22:36265696:36266168:1
GACGAGCCAGAGCCAATCTTCAGTCAGTACCGCATGCCTCAGCCTCACGCCCCCGGGTCACTGAGCCA
ATCTCAGCTGAAAGCGGTGAACAGGTGGAGAGGGTTAATGAACCCAGCATCCTGGAAATGAGCAGAGG
                                *
AGTCAAGCTCACGGATGTGGCCCCTGTA[A/G]GCTTCTTTCTTGTGCTGGATGTAGTCTACCTCGTG
TACGAATCAAAGCACTTACATGAGGGGGCAAAGTCAGAGACAGCTGAGGAGCTGAAGAAGGTGGCTCA
                  **                        ***
GGAGCTGGAGGAGAAGCTAAACAT[T/G]CTCAACAATAA[TTATAA/-]GATTCTGCAGGCGGACCA
AGAACTGTGACCACAGGGCAGGGCAGCCACCAGGAGAGATATGCCTGGCAGGGGCCAGGACAAAATGC
AAACTTTTTTTTTTTTCTGAGACAGAGTCTTGCTCTGTCGCCAAGTTGGAGTGCAATGGTGCGATCTC
```

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).

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 albuminuria, 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, α-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 *APOL1* G1/G2 variants, *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 *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. DESCRIPION 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.

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

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.

Variables	Median (IQR)	Value range	Number of observations
BMI (kg/m^2)	17.361 (14.863-20.505)	10.61-31.029	413
Systolic blood pressure	107 (99-114)	77-156	413
Diastolic blood pressure	57 (53-63)	37-93	413

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).

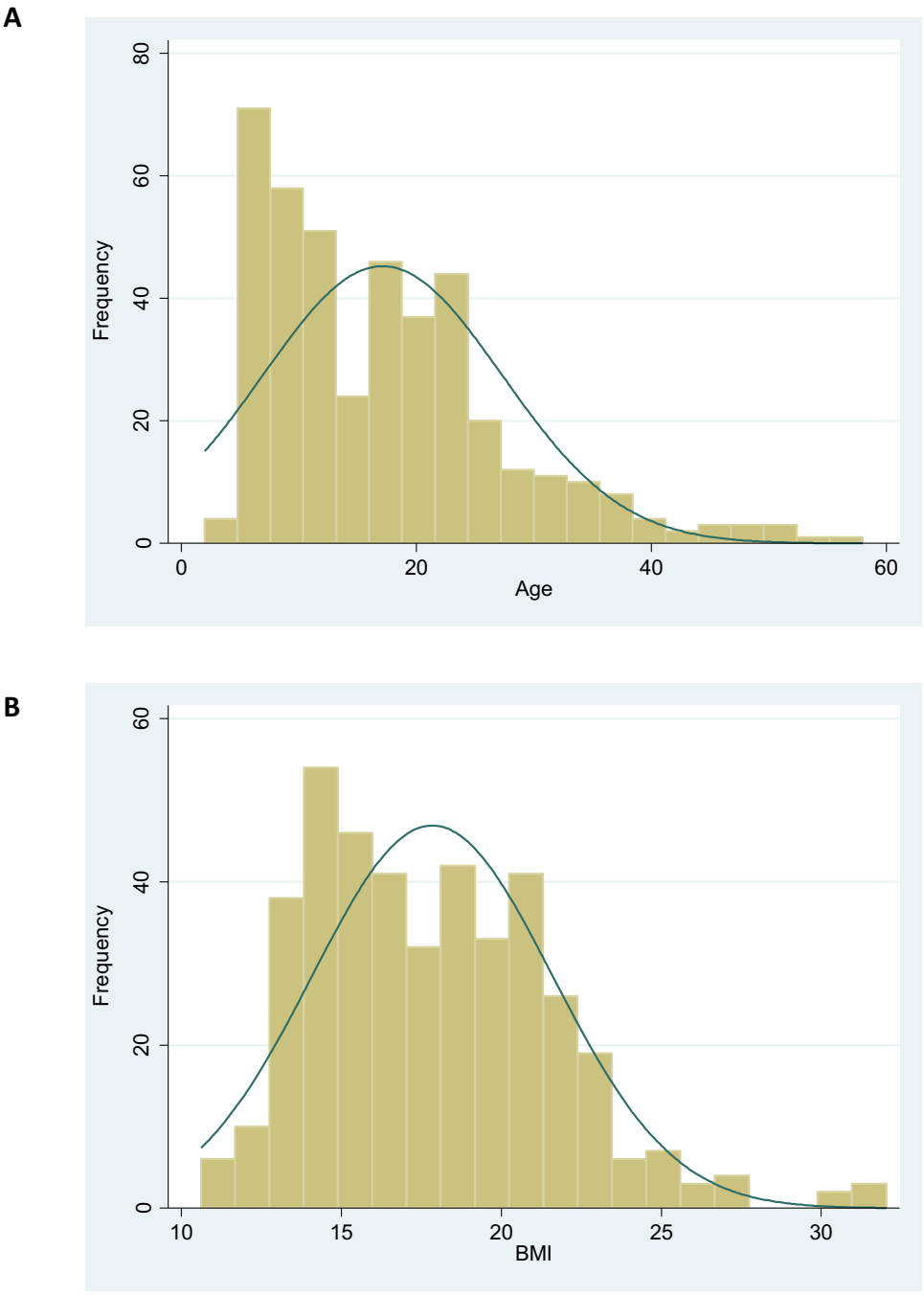
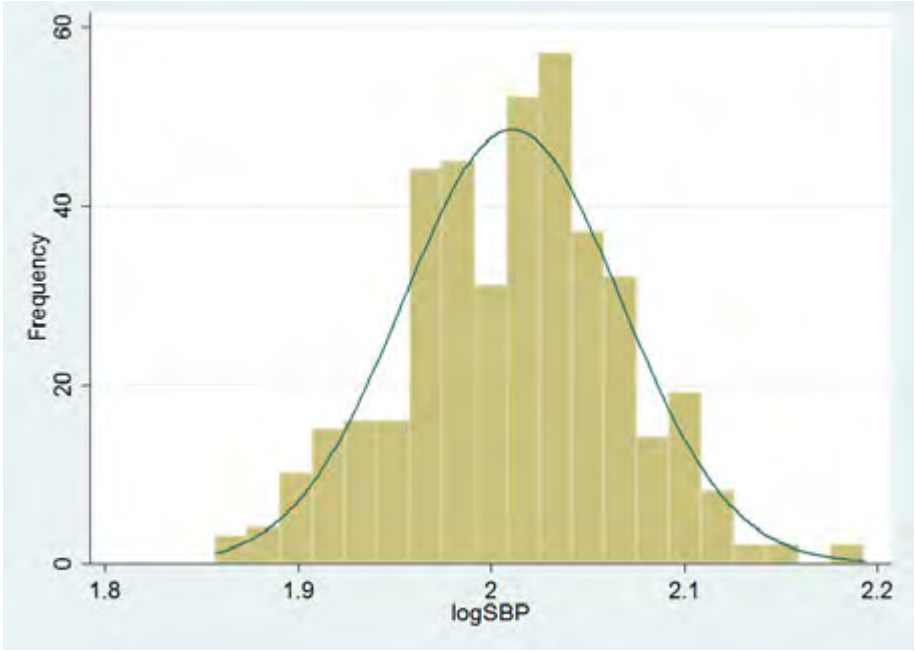


Figure 28: Histogram distribution of variables age and BMI in the SCD patient cohort.

Histogram overlaid with the expected skewed distribution curve of variables (A) age and (B) BMI for the 413 patients included in the cohort.

A



B

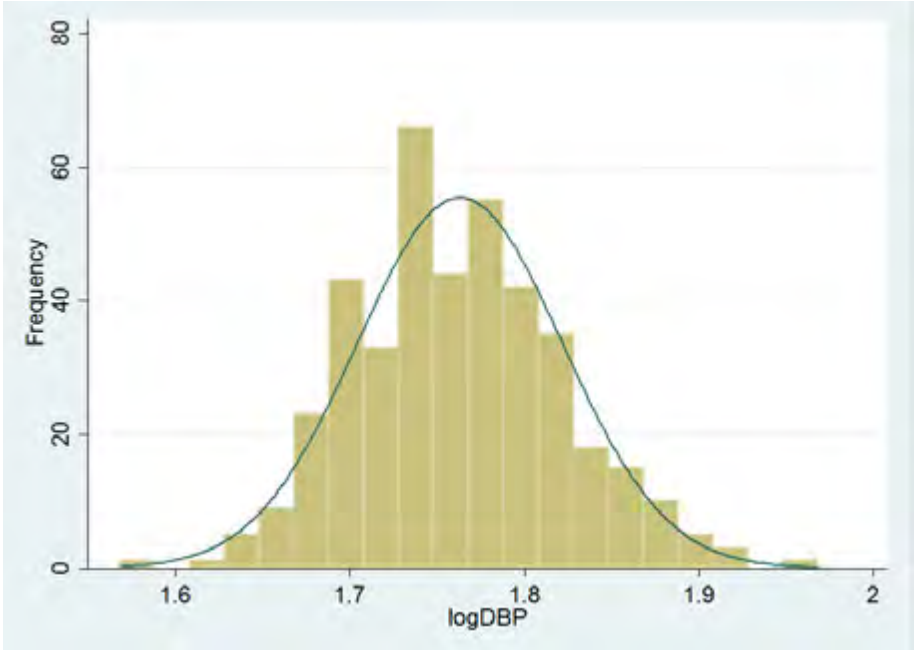


Figure 29: Histogram distribution of variables SBP and DBP in the SCD patient cohort.

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).

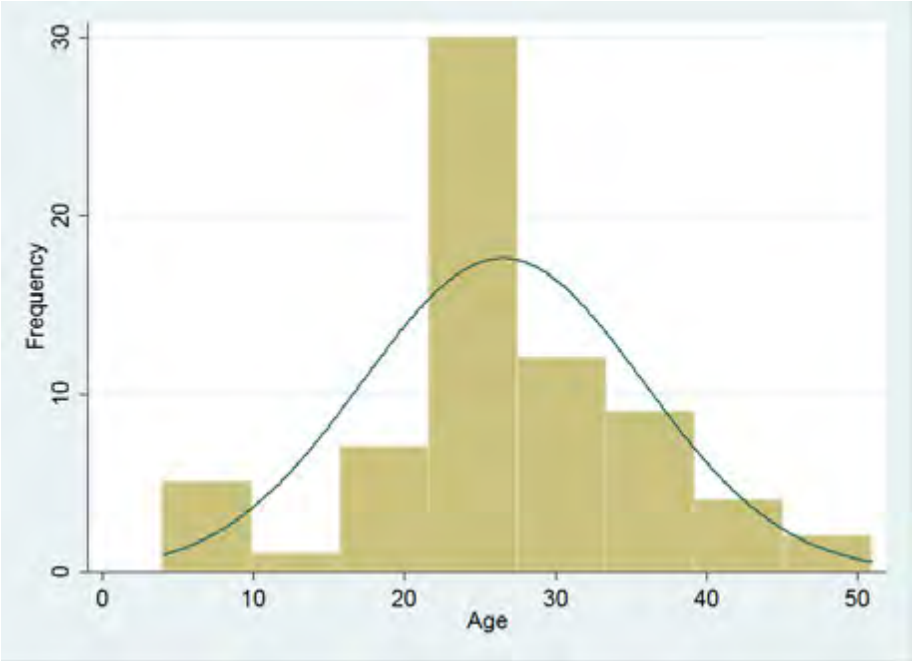


Figure 30: Histogram distribution of age in the non-SCD patient cohort

Histogram overlaid with the expected skewed distribution curve of age and for the 101 patients included in the non-SCD control cohort.

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

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).

Variables	Cohort or *HBB Haplotype**	Median (IQR) or Frequency (%)**	Value range	Number of observations
Hb (g/dL)	SCD patients	7.5 (6.7-8.4)	3.5 - 3090.5	407
	Non-SCD controls	13 (11.6-14.6)	8.1-18.7	71
MCV (fl)	SCD patients	85 (78-92)	60-117	413
	Non-SCD controls	81 (76.5-84)	56-102	71
MCHC (g/dL)	SCD patients	33.5 (31.6-35.5)	21.5-54.3	413
	Non-SCD controls	36 (32.3-37.7)	29.8-45.2	70
WBC (10^9/l)	SCD patients	12.7 (10.2-16.2)	4-49.8	413
	Non-SCD controls	5.03 (4.5-5.9)	2-10	71
Lymphocytes (10^9/l)	SCD patients	5 (3.8 - 6.7)	1.4 - 22.1	406
	Non-SCD controls	2.2 (1.8-2.6)	0.4-6.1	71
Monocytes (10^9/l)	SCD patients	1.4 (0.9-2.1)	0.1-8.2	404
	Non-SCD controls	0.6 (0.4-0.8)	0.2-94	71
Platelets (10^9/l)	SCD patients	370 (284-466)	29 - 1078	402
	Non-SCD controls	242 (181.5-290.5)	73-651	71
HbA2 (%)	SCD patients	3.6 (2.85 - 4.1)	0-18.2	408
	Non-SCD controls	3.2 (2-3.9)	0.1-5	66
HbF (%)	SCD patients	8.05 (2.15-13.35)	0-37.4	408
HBB Haplotype	Atypical	38 (5.43)		700^{+}
	Bantu/CAR	3 (0.43)		
	Cameroon	145 (20.71)		
	Indian-Arab	3 (0.43)		
	Senegal	1 (0.14)		
	Benin	510 (72.86)		

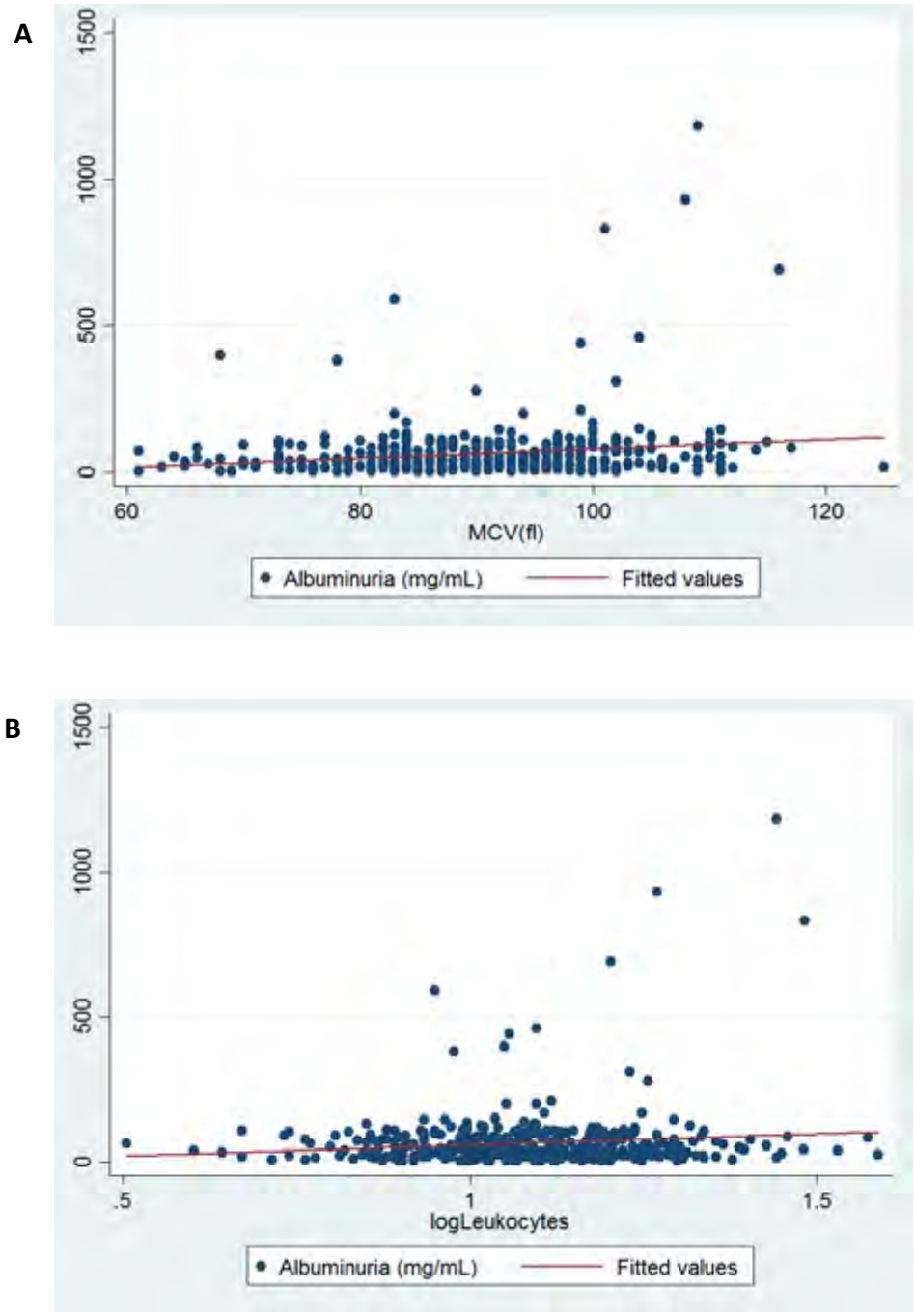


Figure 31: Scatter plots illustrating the relationship between hematological indices and crude albuminuria in the SCD cohort, with a significant positive relationship observed between (A) MCV (r=0.1667, p=0.008) and (B) log(Leukocytes) (r=0.1160, p=0.0202) with crude albuminuria. The (A) MCV and (B) 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/1.73m^2 and >140 ml/min/1.73m^2 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/1.73m^2 in women and 140 ml/min/1.73m^2 in men.

Variables	Equation	Median (IQR)	Value range	Glomerular Hyperfiltration (%)	Number of observations
eGFR (ml/min)	MDRD	140.5 (101.4-205.5)	44.6-779.2	214 (52.97)	404
	Cockcroft-Gault	111.8 (90.5-162)	55.1-368.7	172 (42.57)	404
	CKD-EPI	135.1 (112-154.4)	50.8-250.8	200 (49.5)	404

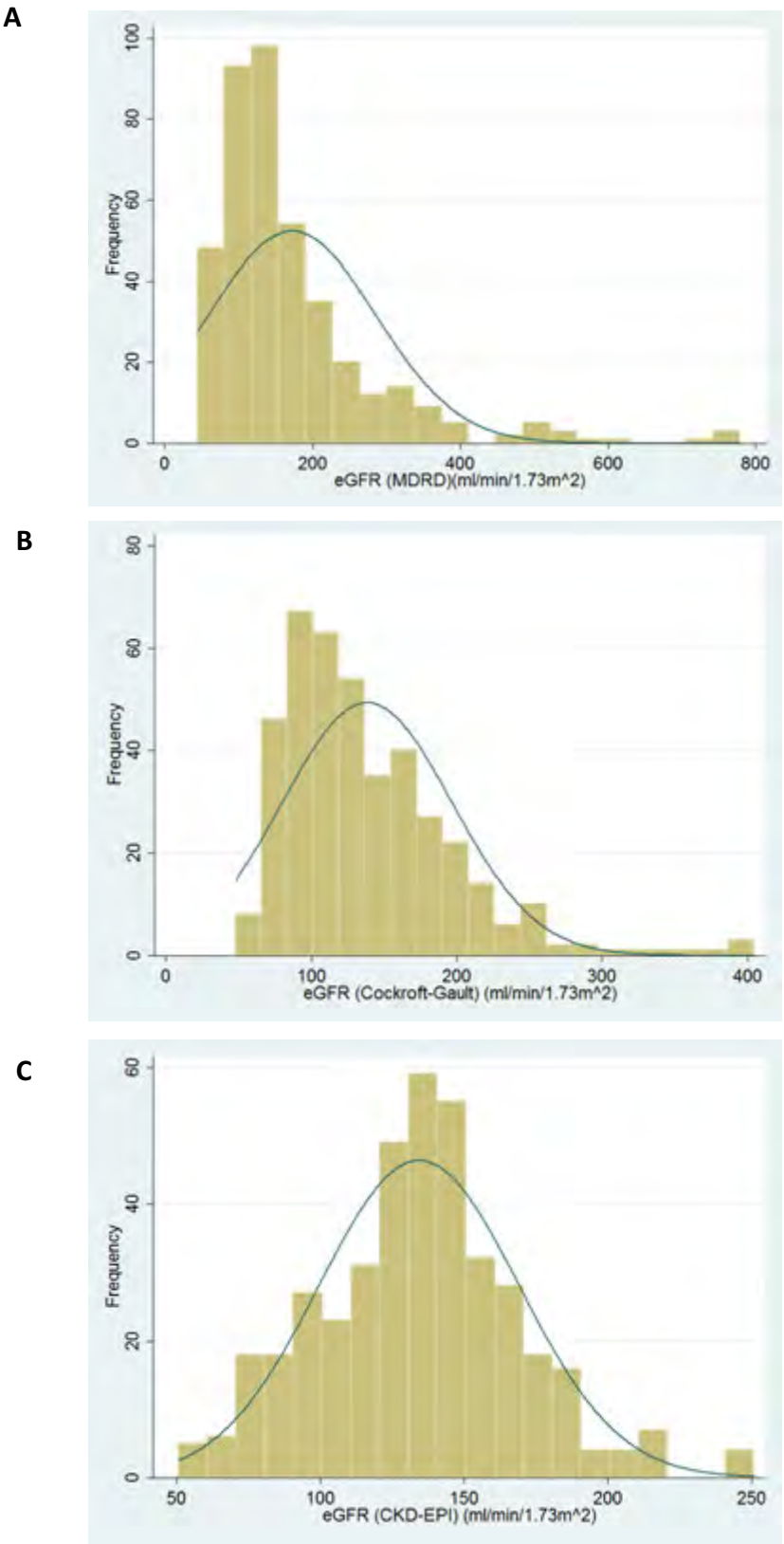


Figure 32: Histogram distribution of eGFR variables in the SCD cohort.

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.

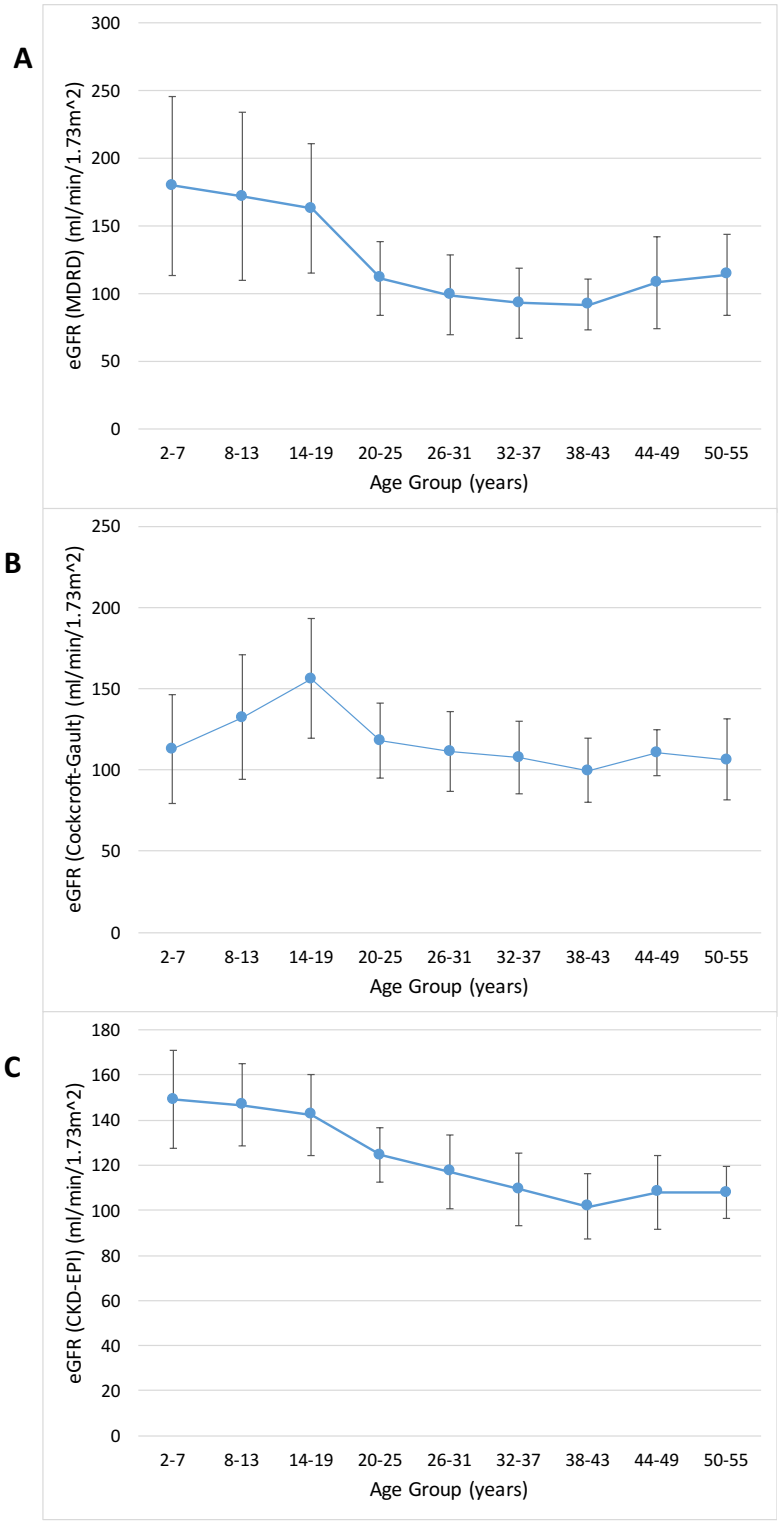


Figure 33: eGFR estimates per age group for SCD patients.

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.

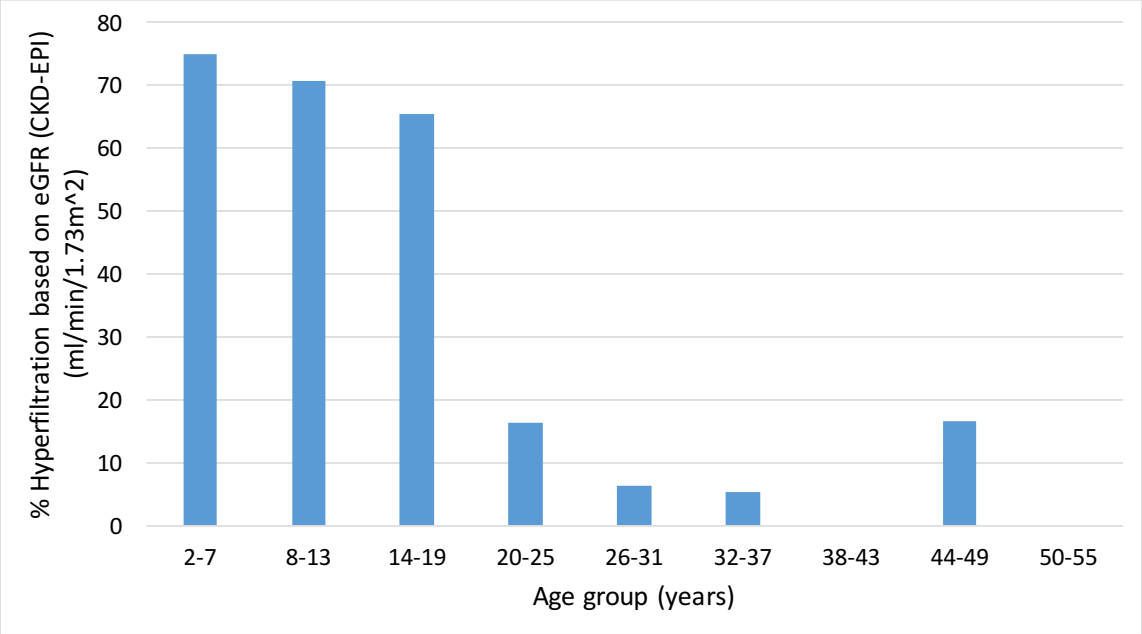


Figure 34: Histogram distribution of the prevalence of hyperfiltration in various age groups in the SCD cohort. The prevalence of hyperfiltration, defined as filtration values >130 ml/min/1.73m^2 and >140 ml/min/1.73m^2 in women and men, respectively, calculated using the CKD-EPI equation in various age groups of the SCD cohort.

3.1.4.2. MICRO-ALBUMINURIA

The 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.

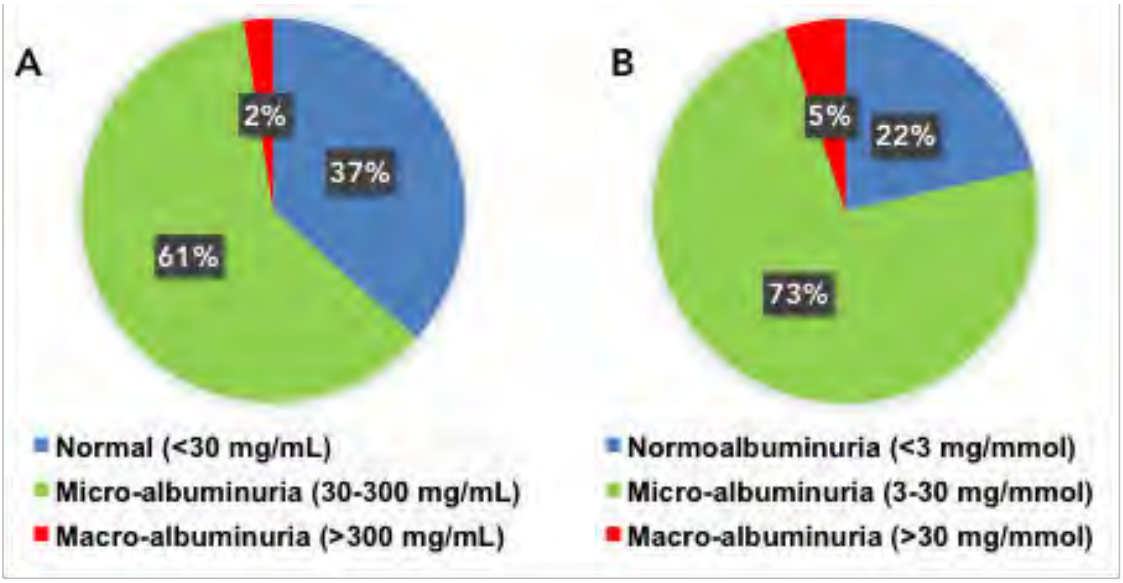


Figure 35: Prevalence of crude albuminuria and ACR in the SCD patient cohort.

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.

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.

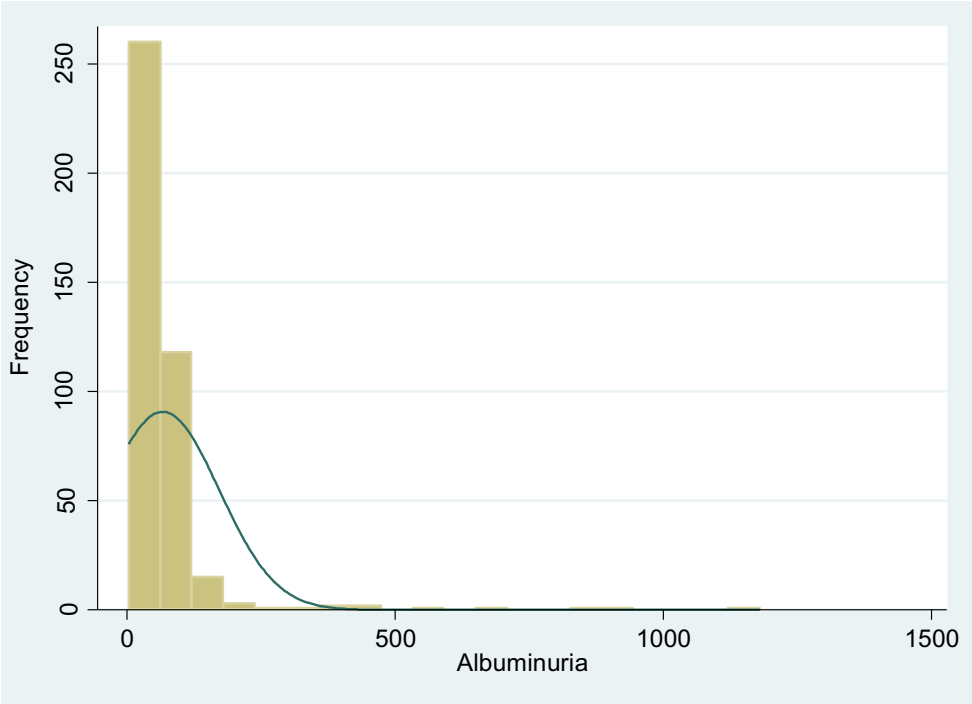


Figure 36: Histogram distribution of micro-albuminuria

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.

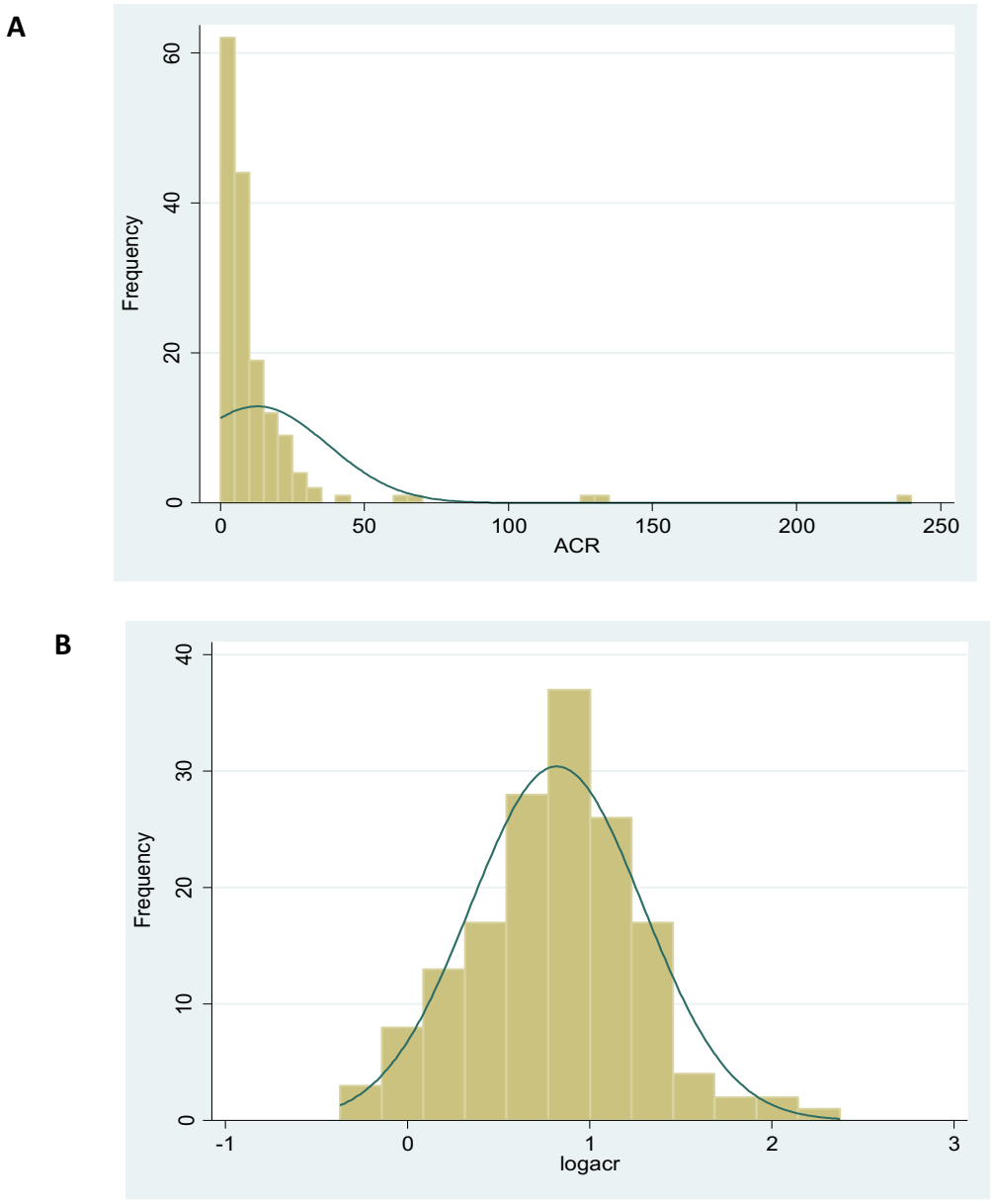


Figure 37: Histogram of the frequency of (A) ACR and (B) Log_{10}ACR values, overlaid with the expected normal distribution for the SCD patient cohort

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)

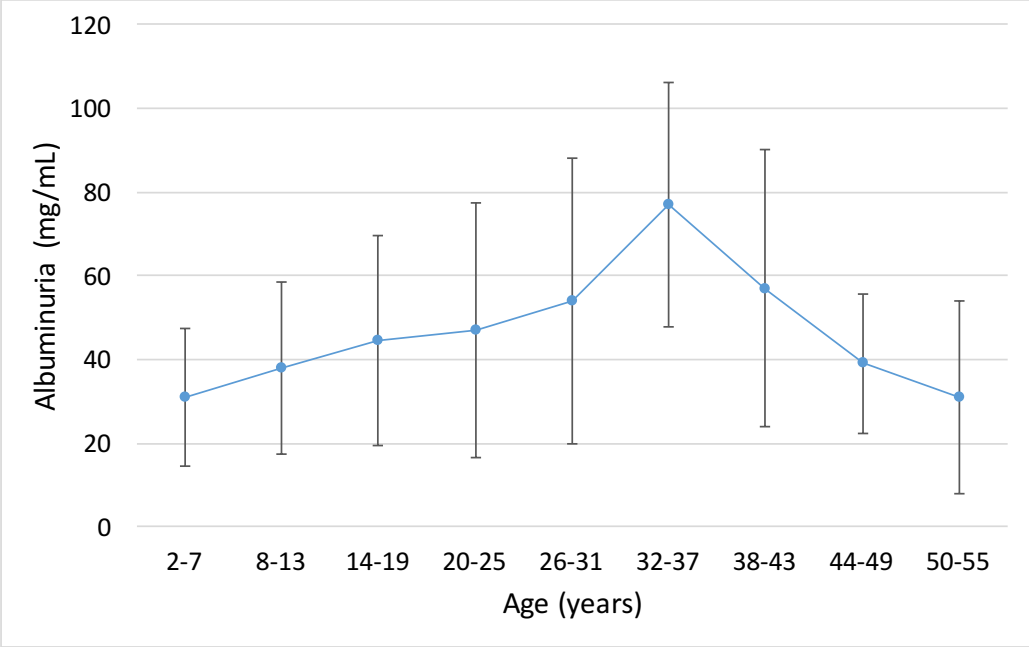


Figure 38: Albuminuria estimates per age group for SCD patients.
The median albuminuria values per age group for SCD patients. There is a general positive association between crude albuminuria and increasing age until the fourth decade. The values represent the median albuminuria values per age group, while the error bars represent 1 average deviation from the median.

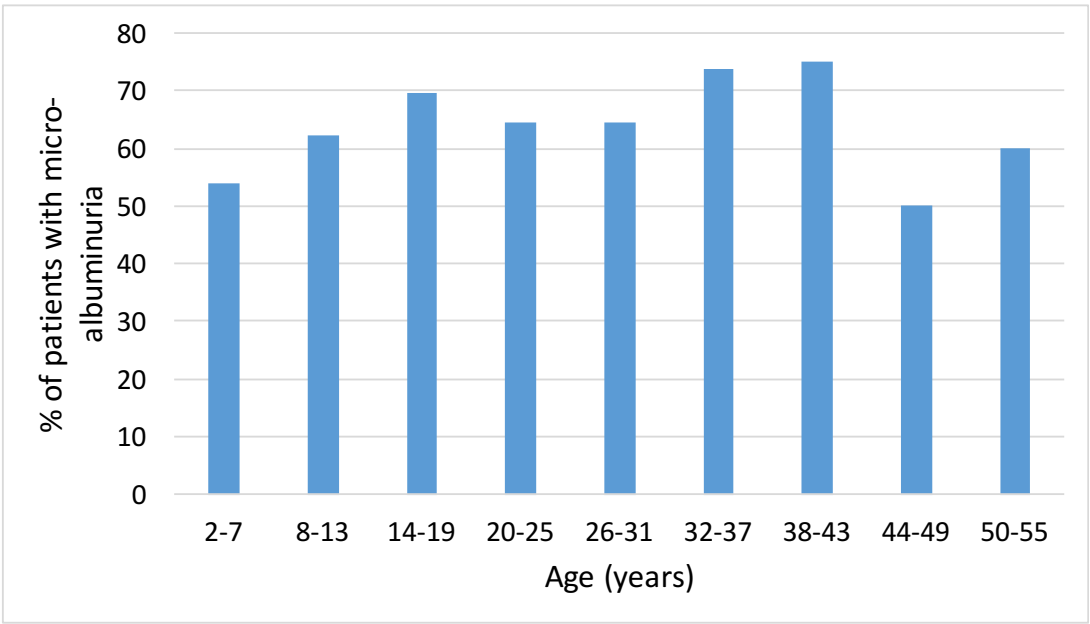


Figure 39: Histogram distribution of the prevalence of micro-albuminuria in various age groups in the SCD cohort. The prevalence of micro-albuminuria, defined as albuminuria values of 30-300 mg/ml in various age groups of the SCD cohort.

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.

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

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/1.73m^2 in women and >140 ml/min/1.73m^2 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)

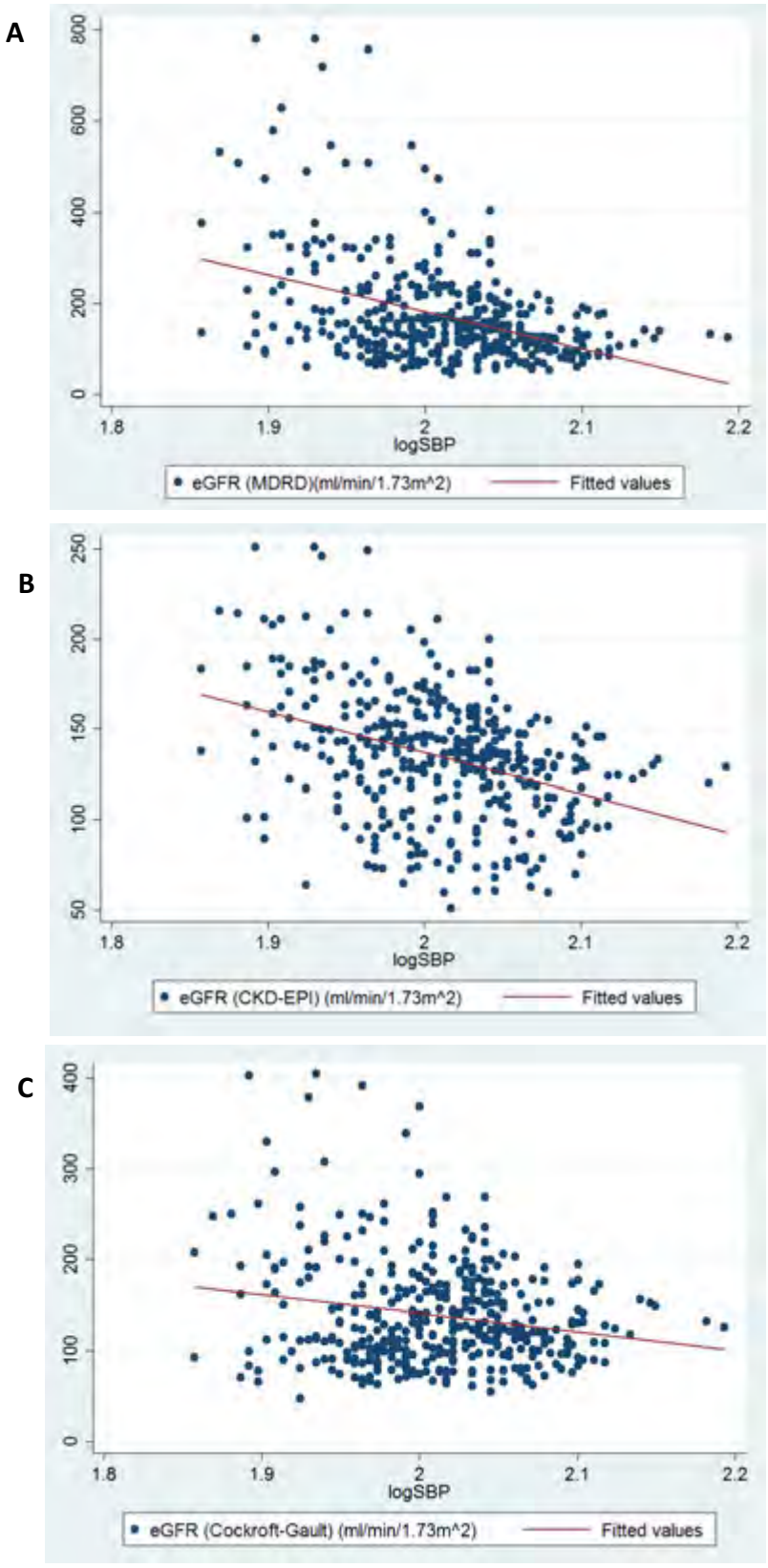


Figure 40: Scatter plots illustrating the relationship between eGFR values and logSBP in the SCD patient cohort, calculated using the (A) MDRD (r=-0.4071, p<0.0001) (B) CKD-EPI (r=-0.3674, p<0.0001) and (C)Cockcroft-Gault (r=p=0.0182) equations. The log(SBP) variable is displayed on the x-axis for all three graphs, with the eGFR values on the y-axis. The red line indicates a line of best fit, fitted to the data.

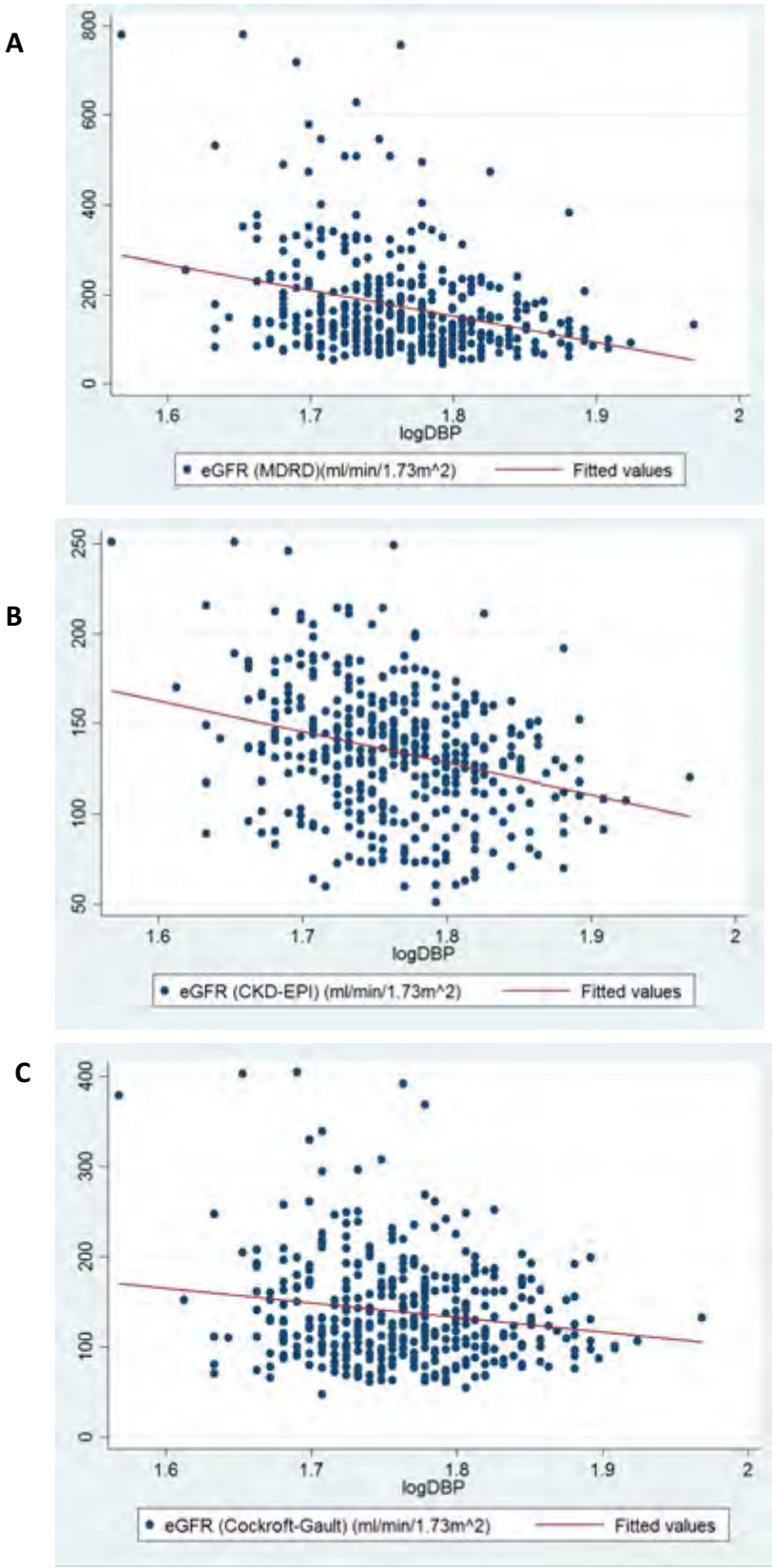
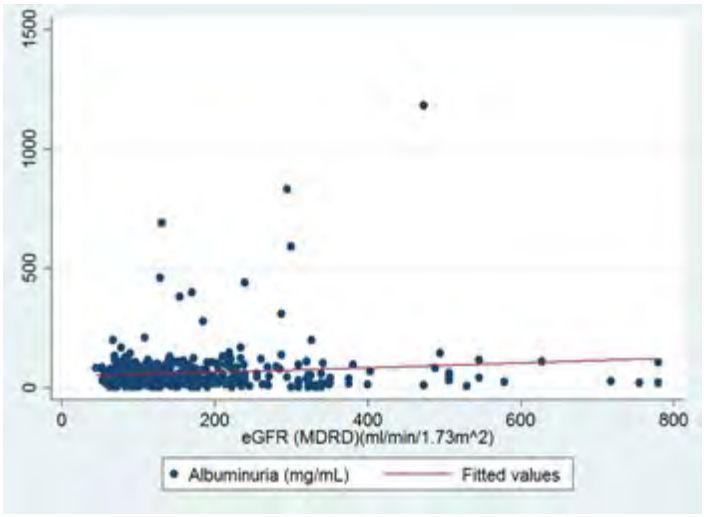


Figure 41: Scatter plots illustrating the relationship between eGFR values and logDBP in the SCD patient cohort, calculated using the (A) MDRD (r=-0.3004, p<0.0001) (B) CKD-EPI (r=-0.2956, p<0.0001) and (C) Cockcroft-Gault (r=-0.1644, p=0.001) equations. The log(DBP)variable is displayed on the x-axis 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.

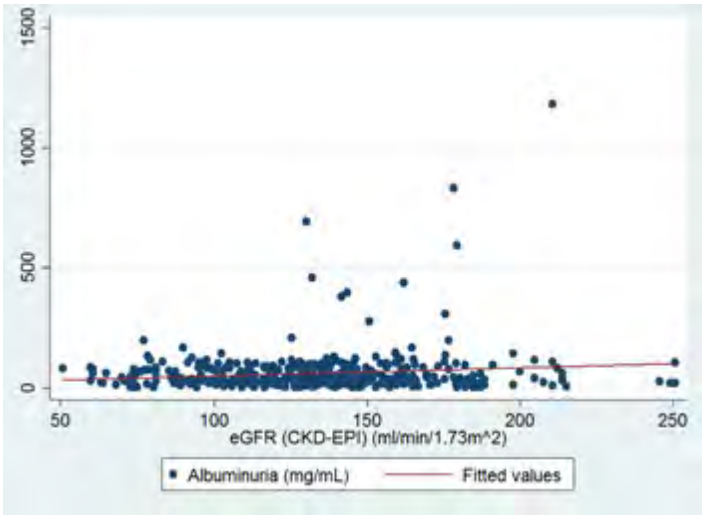
Association between eGFR and albuminuria

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).

A



B



C

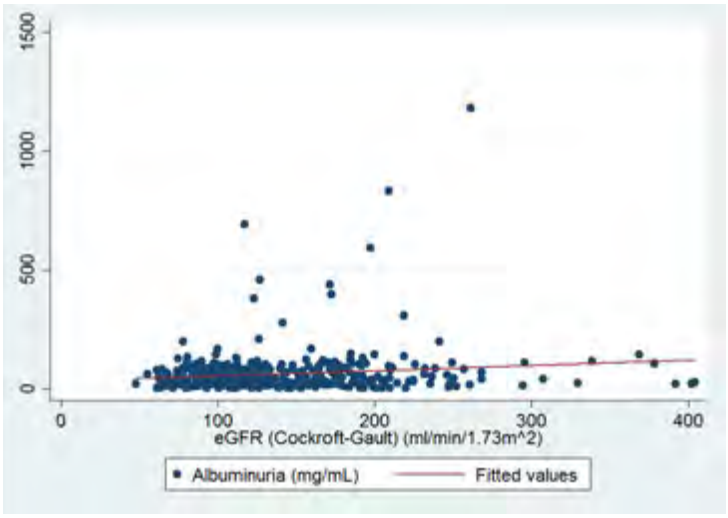
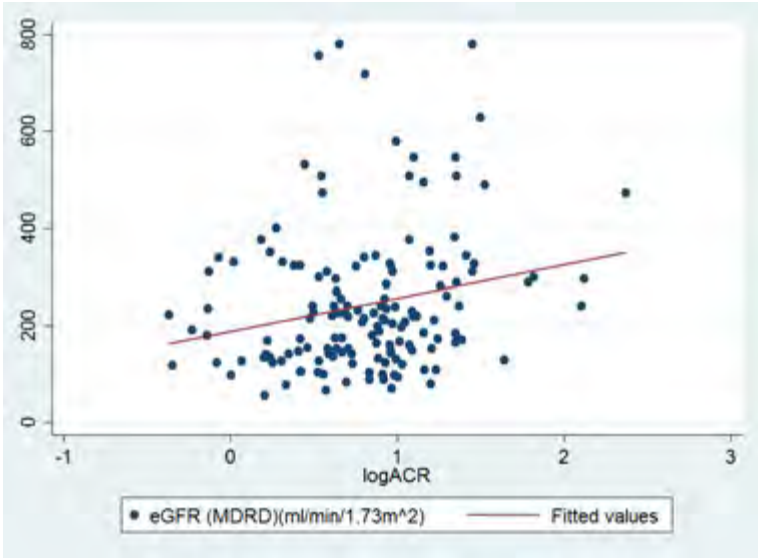


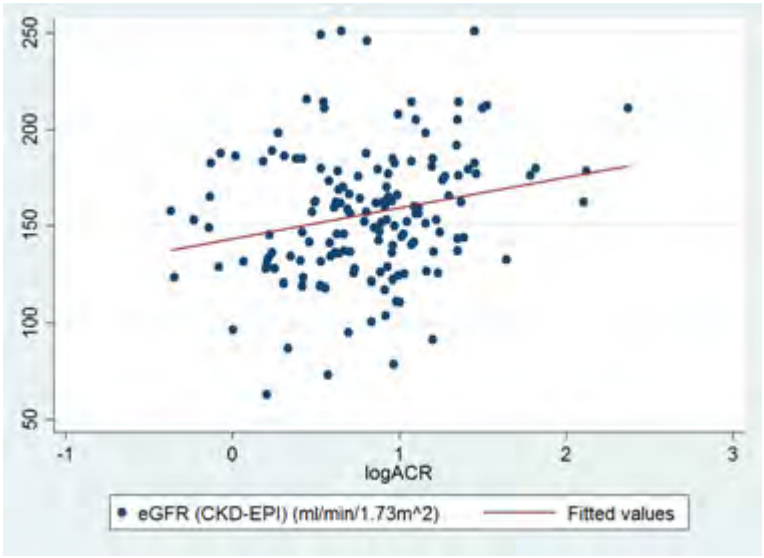
Figure 42: Scatter plots illustrating the relationship between albuminuria values and eGFR in the SCD patient cohort, calculated using the (A) MDRD (p=0.9159), (B) CKD-EPI (p=0.8706) and (C) Cockcroft-Gault (p=0.1935) equations. The eGFR values are displayed on the x-axis for all three graphs, with the 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).

A



B



C

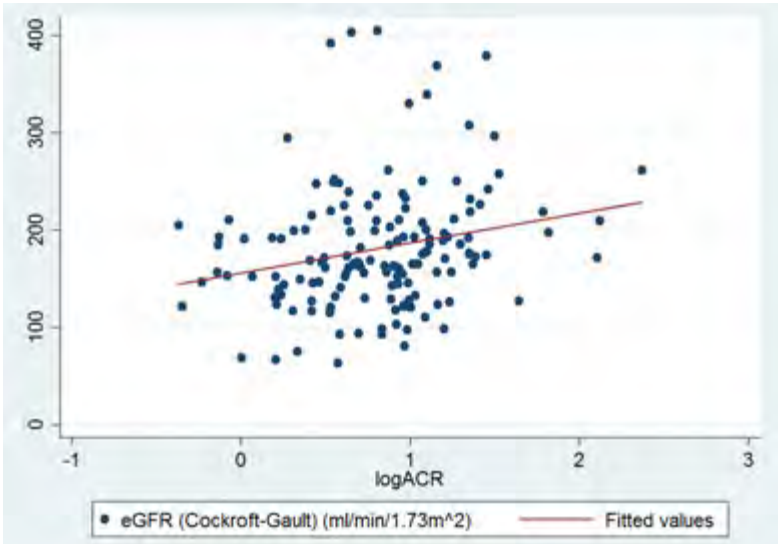


Figure 43: Scatter plots illustrating the relationship between eGFR values and log(ACR)in the SCD patient cohort, calculated using the (A) MDRD (r=0.2232, p=0.0059), (B) CKD-EPI (r=0.02238, p=0.0057) and (C) Cockcroft-Gault (r=0.2317, p=0.0042) equations. The log(ACR) variable is displayed on the x-axis 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).

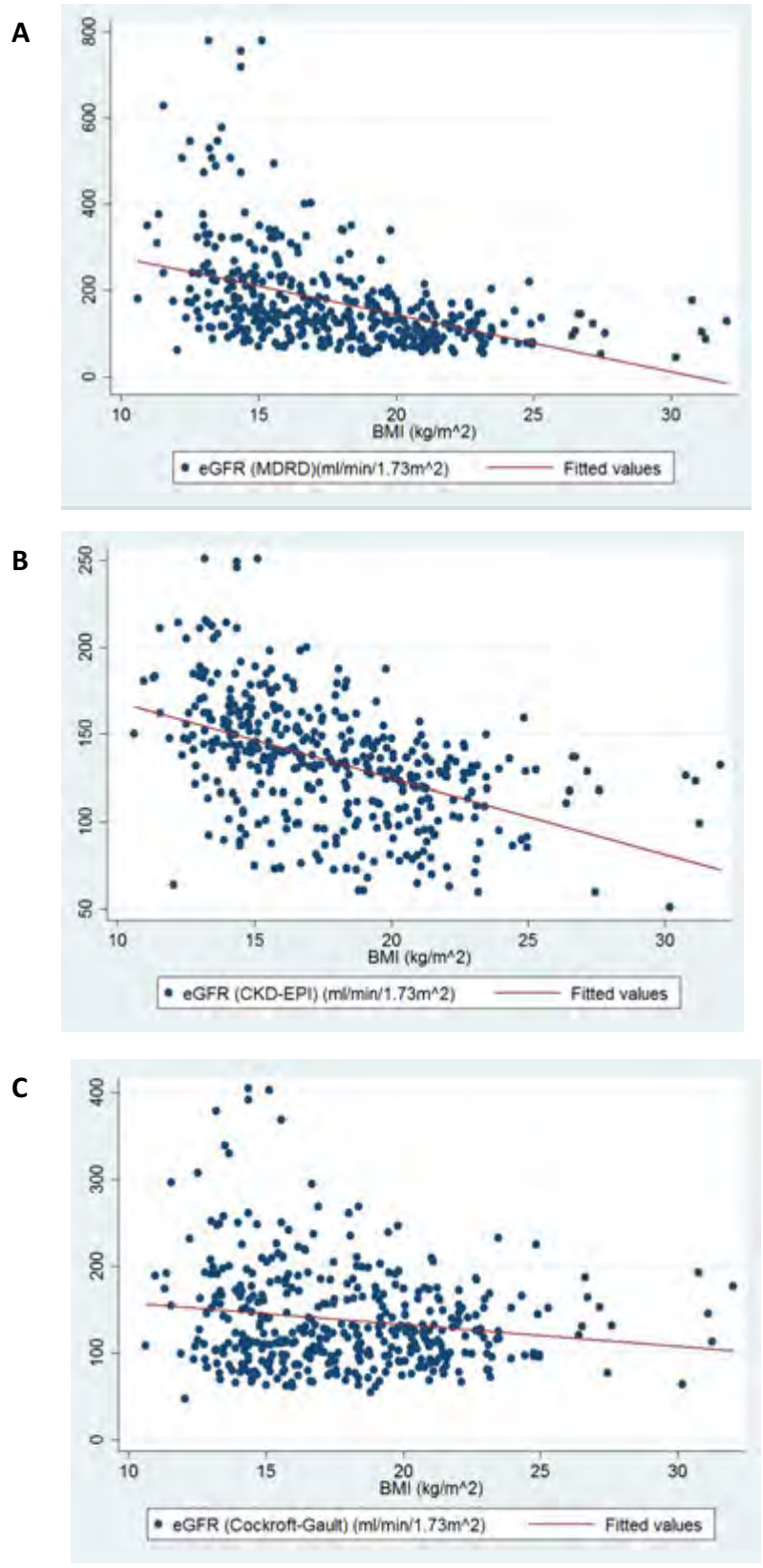


Figure 44: Scatter plots illustrating the relationship between eGFR values and BMI in the SCD patient cohort, calculated using the (A) MDRD (r=-0.52, p<0.0001) (B) CKD-EPI (r=0.529, p<0.0001) and (C) Cockcroft-Gault (r=-0.164, p=0.0182) equations. The BMI (kg/m^2) variable is displayed on the x-axis 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.

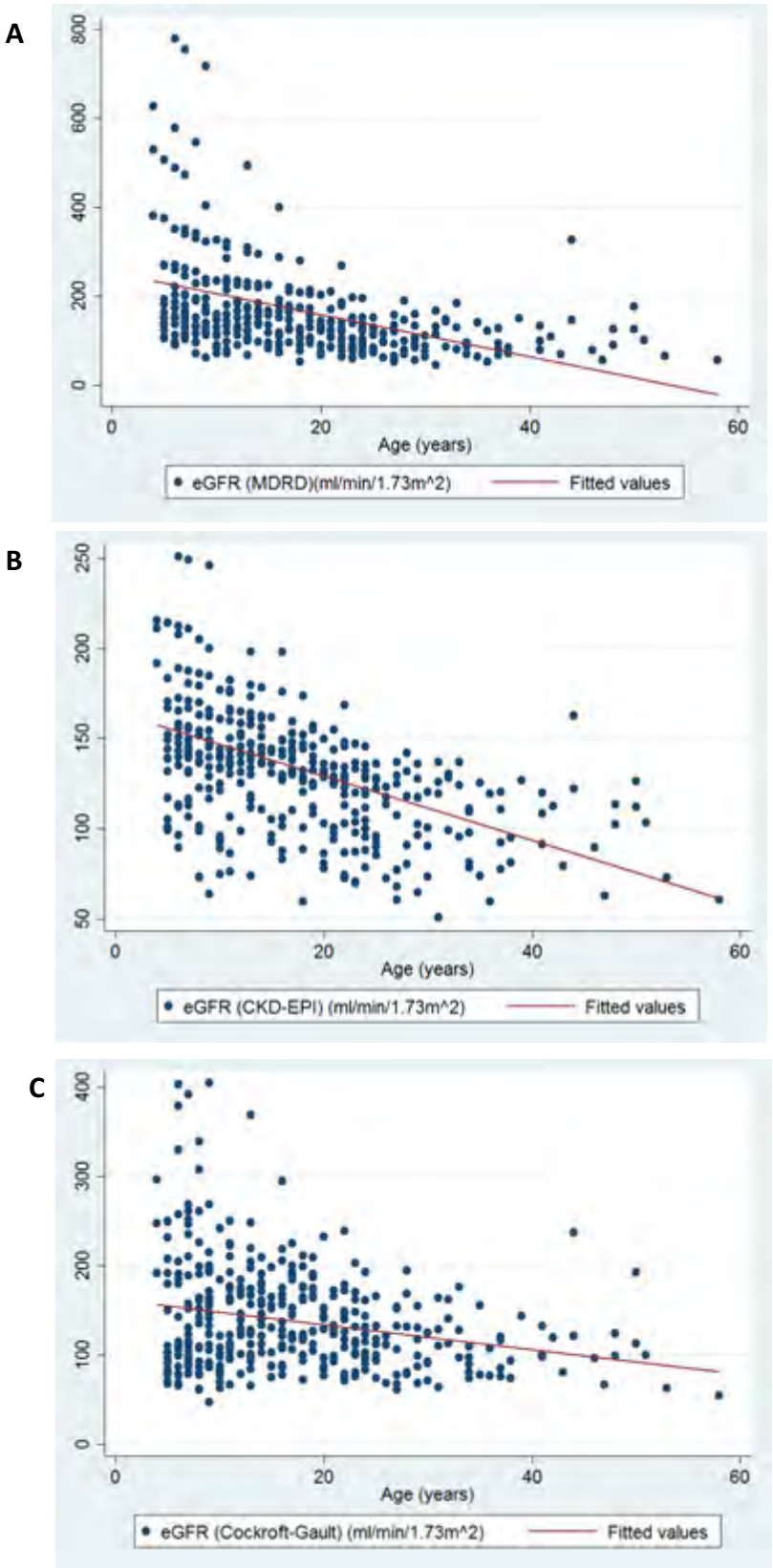


Figure 45: Scatter plots illustrating the relationship between eGFR values and age in the SCD patient cohort, calculated using the (A) MDRD (r=-0.5233, p<0.0001) (B) CKD-EPI (r=-0.5501, p<0.0001 and (C) Cockcroft-Gault (r=-0.1684, p<0.0001) equations. Age is displayed on the x-axis for all three graphs, with the eGFR values on the y-axis. The red line indicates a line of best fit, fitted to the data.

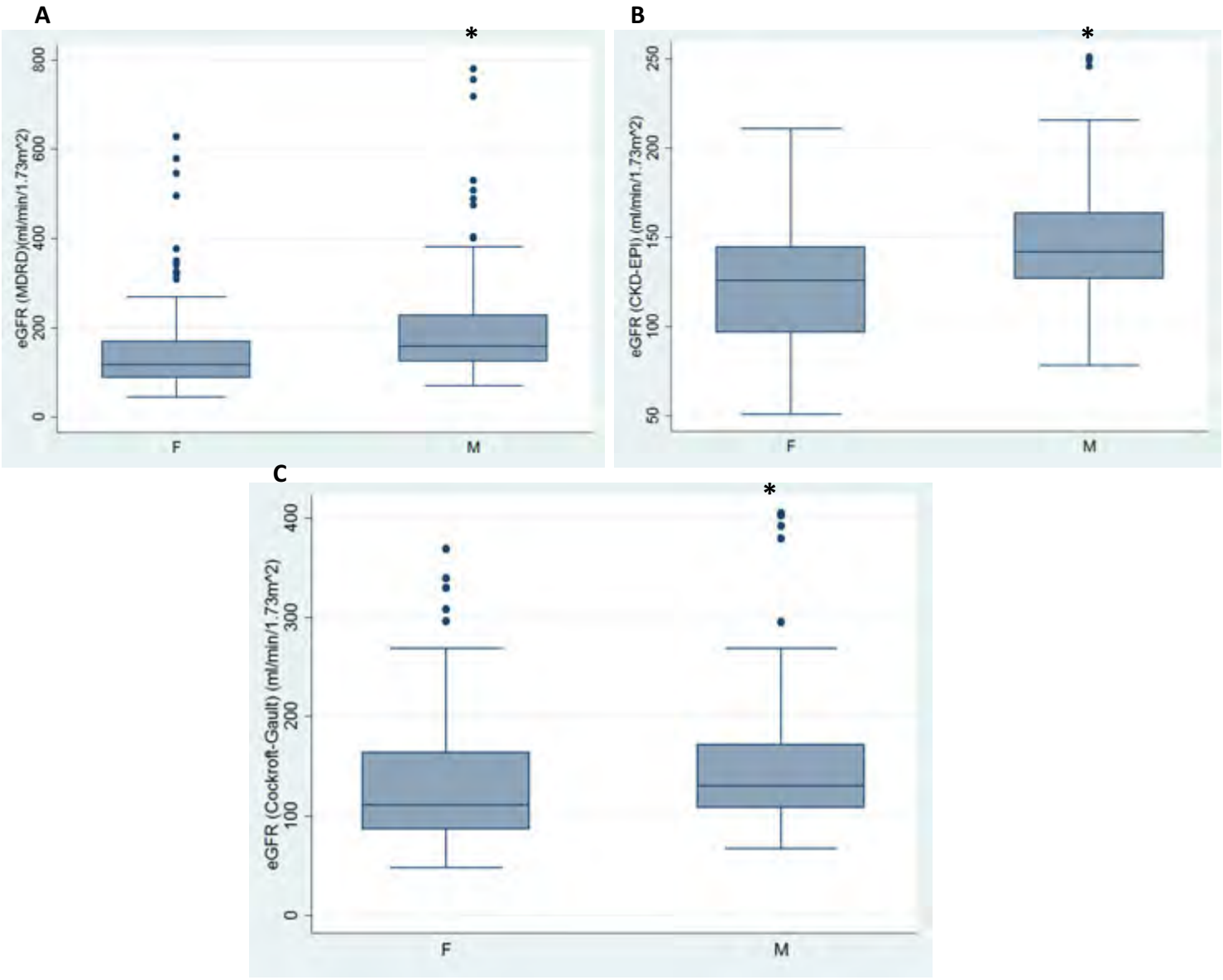


Figure 46: Box and whisker plots showing the association of eGFR values with sex in the SCD cohort.

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).

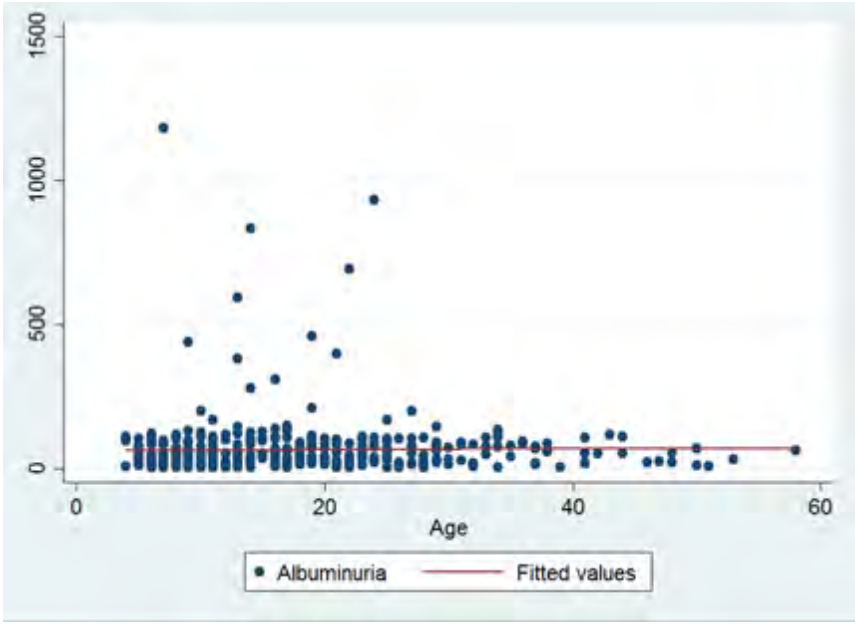


Figure 47: Scatter plot illustrating the positive relationship between age and crude albuminuria in the SCD patient cohort, with age displayed on the x-axis and crude albuminuria data on the y-axis. The red line indicates a line of best fit, fitted to the data, indicating the slightly positive nature of the relationship (r=0.1421, p=0.004)

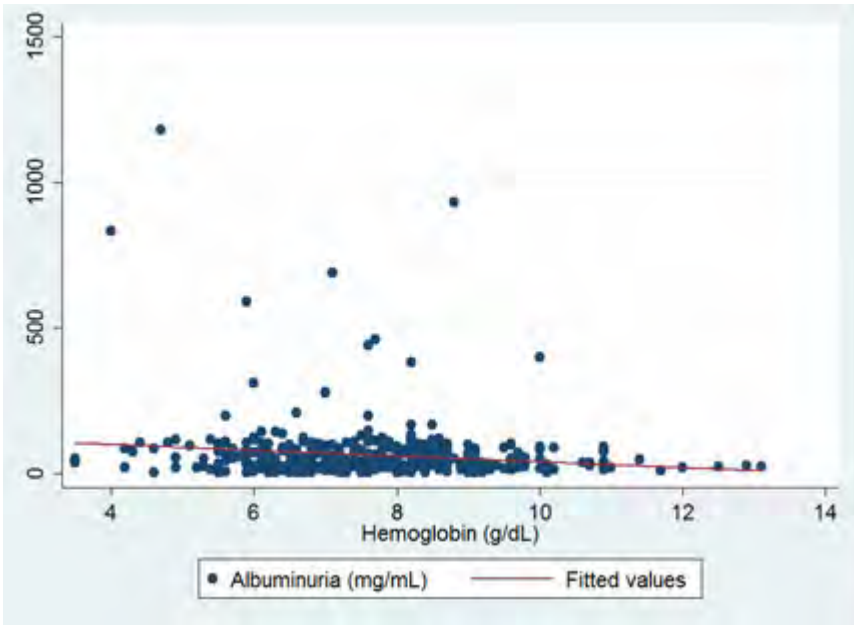


Figure 48: Scatter plot illustrating the negative relationship between albuminuria and hemoglobin in the SCD patient cohort, with hemoglobin displayed on the x-axis and crude albuminuria data on the y-axis. The red line indicates a line of best fit, fitted to the data, indicating the near significant 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.

Variable	Co-efficient	P-value
Age	0.1241	0.0041
Hemoglobin (g/dl)	-0.091	0.0289
BMI (kg/m^2)	0.0977	0.0489
Leukocytes ($10^9/l$)	0.1160	0.0202
Mean arterial pressure (mm Hg)	0.1181	0.0175
MCV (fl)	0.1667	0.008

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.

Variable	Co-efficient	P-value
Age	-0.523 (MDRD) -0.550 (CKD-EPI) -0.165 (Cockcroft-Gault)	<0.0001 (MDRD) <0.0001 (CKD-EPI) 0.0009 (Cockcroft-Gault)
ACR (mg/mmol)	0.2232 (MDRD) 0.2238 (CKD-EPI) 0.2317 (Cockcroft-Gault)	0.0059 (MDRD) 0.0057 (CKD-EPI) 0.0042 (Cockcroft-Gault)
Male sex	risk	<0.0001 (all equations)
SBP (mm Hg)	-0.407 (MDRD) -0.367 (CKD-EPI) -0.186 (Cockcroft-Gault)	<0.0001 (MDRD) <0.0001 (CKD-EPI) 0.0001 (Cockcroft-Gault)
DBP (mm Hg)	-0.3 (MDRD) -0.296 (CKD-EPI) -0.164 (Cockcroft-Gault)	<0.0001 (MDRD) <0.0001 (CKD-EPI) 0.0001 (Cockcroft-Gault)
Hemoglobin (g/dl)	-0.128 (MDRD) -0.142 (CKD-EPI)	0.0106 (MDRD) 0.0046 (CKD-EPI)
BMI (kg/m^2)	-0.52 (MDRD) -0.529 (CKD-EPI) -0.118 (Cockcroft-Gault)	<0.0001 (MDRD) <0.0001 (CKD-EPI) 0.0182 (Cockcroft-Gault)
Serum Creatinine (mg/dl)	-0.915 (MDRD) -0.901 (CKD-EPI) -0.931 (Cockcroft-Gault)	<0.0001 (all equations)
Mean arterial pressure (mm Hg)	-0.321 (MDRD) -0.337 (CKD-EPI)	<0.0001 (MDRD) <0.0001 (CKD-EPI)

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.

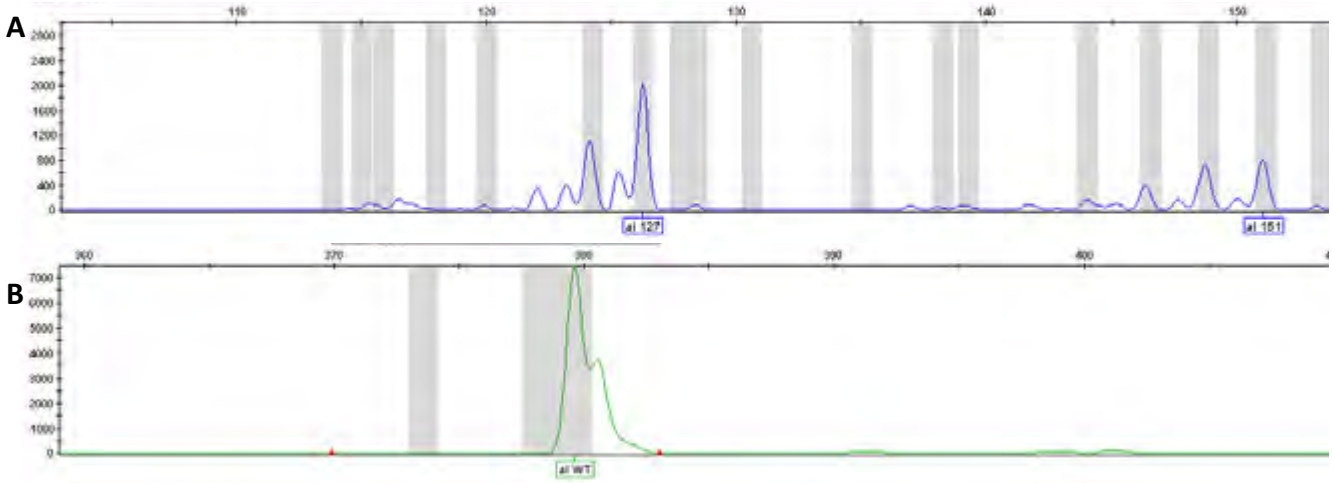


Figure 49: Fragment Analysis of rs3074372 (*HMOX1*) and rs71785313 (*APOL1*) polymorphisms for SCD patient 102.

Electropherogram produced using the ABI Prism 3130*xl* 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.

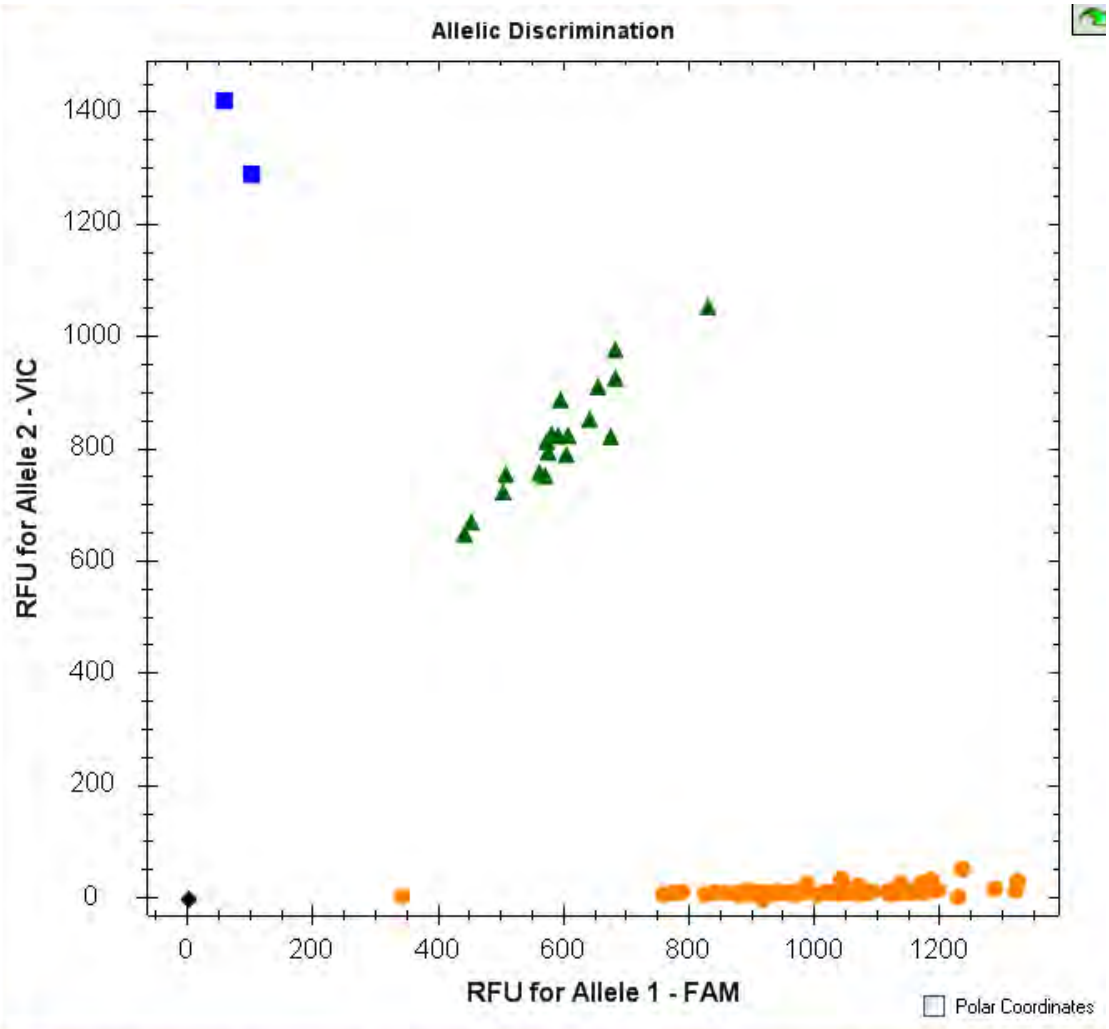


Figure 50: TaqMan SNP Genotyping plot of SCD cohort plate 3 for rs743811 (*HMOX1*).

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'-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 (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.

Variables		**Frequency (%)**	**HWE (p-value)**	**Number of observations**
MYH9 - rs11912763 (G/A)	GG	305 (78.01)	0.924	391
	AG	80 (20.46)		
	AA	6 (1.53)		
MYH9 - rs16996648 (T/C)	TT	207 (56.25)	1	368
	CT	138 (37.5)		
	CC	23 (6.25)		
MYH9 - rs1005570 (G/A)	GG	147 (38.79)	0.992	379
	AG	178 (46.97)		
	AA	54 (14.25)		
MYH9 - rs16996672 (C/T)	CC	211 (54.38)	0.122	388
	CT	158 (40.72)		
	TT	19 (4.9)		

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.

Variables		**SCD Patient cohort**			**Non-SCD control cohort**			
		Freq (%)	**HWE (p-value)**	**Number of observations**	**Freq (%)**	**HWE (p-value)**	**Number of observations**	**Chi-squared test (p-value)***
3.7del α-globin gene genotypes	αα/αα	198 (58.58)	**0.0005**	338	47 (81.03)	**0.007**	58	**0.0028**
	αα/α3.7	105 (31.07)			8 (13.79)			
	α3.7/α3.7	35 (10.36)			3 (5.17)			
HMOX1 – rs3074372 (S/L)	S/S	2 (0.54)	0.67	368	3 (3.49)	**0.03**	86	**0.06062**
	S/L	45 (12.23)			11 (12.79)			
	L/L	321 (87.23)			72 (83.72)			
HMOX1 – rs743811 (T/C)	TT	290 (78.59)	0.598	369	77 (79.38)	0.34	97	0.527
	TC	76 (20.60)			18 (18.56)			
	CC	3 (0.81)			2 (2.06)			
APOL1 – rs73885319 (A/G)	AA	285 (75.6)	1	377	82(81.19)	1	101	
	AG	86 (22.81)			18 (17.82)			
	GG	6 (1.59)			1 (0.99)			
APOL1 – rs60910145 (T/G)	TT	278 (76.8)	0.81	362	82 (82)	1	100	
	TG	78 (21.55)			18 (18)			
	GG	6 (1.66)			0 (0)			
APOL1 – rs71785313 (WT/del)	WT/WT	318 (84.35)	0.73	377	72 (78.26)	1	92	
	Del/WT	56 (12.85)			19 (20.65)			
	Del/del	3 (0.8)			1 (1.09)			
APOL1 – G1/G2	G/G	333 (88.1)	p=0.005	378	72 (78.26)	0.24	92	**0.0097**
	G/A	40 (10.58)			20 (21.74)			
	A/A	5 (1.32)			0			

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.

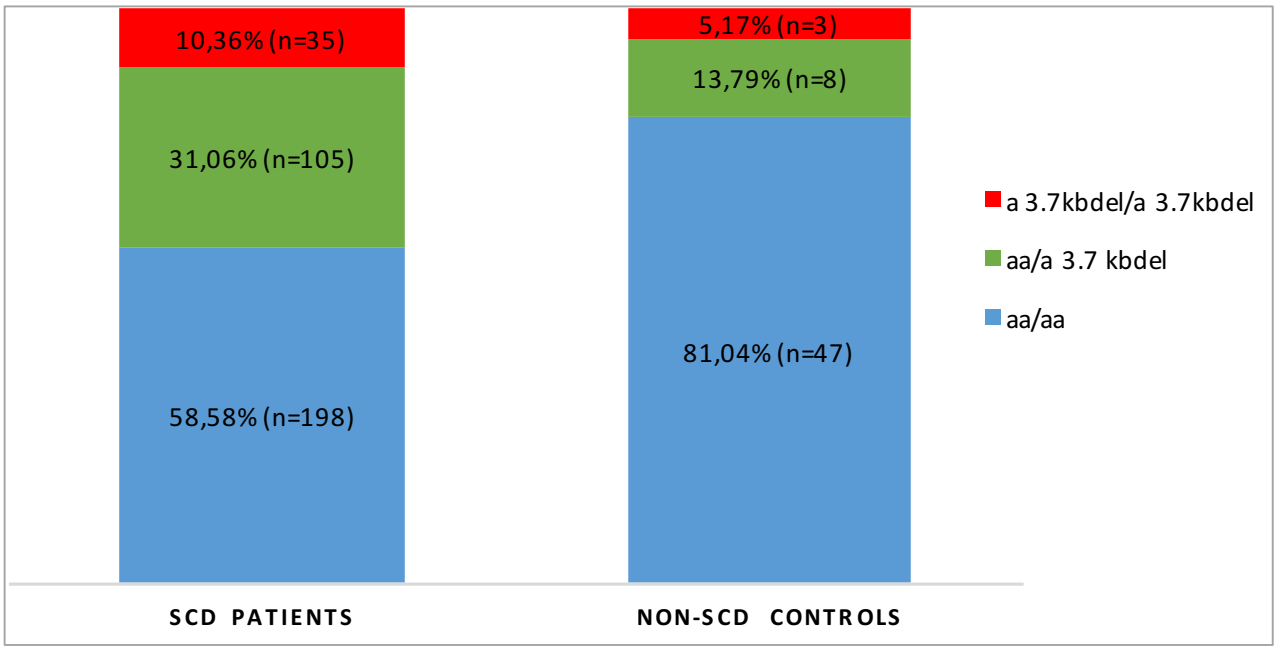


Figure 51: Co-inheritance of alpha-thalassemia among SCD patients and non-SCD controls. The SCD patients display a much higher prevalence of co-inherited 3.7 kb α-globin gene deletion than the non-SCD controls (p=0.0028)

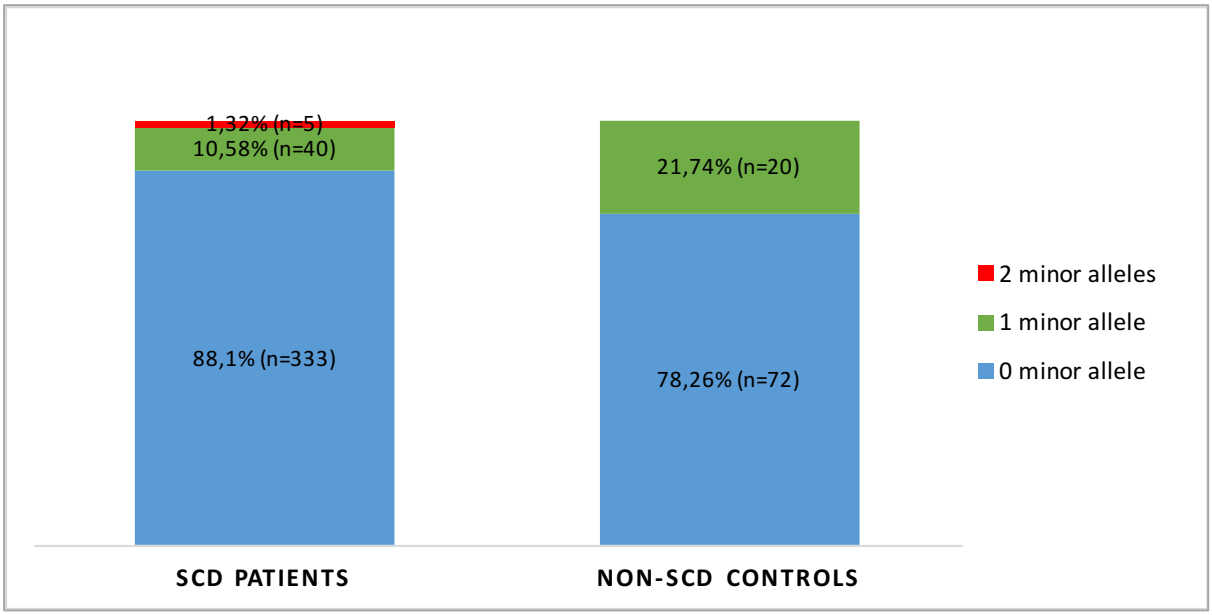


Figure 52: Prevalence of *APOL1* G1/G2 variant among SCD patients and non-SCD controls. The SCD patients display a prevalence of 2 minor alleles for the *APOL1* G1/G2 variant in comparison to the controls (p=0.0097)

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.

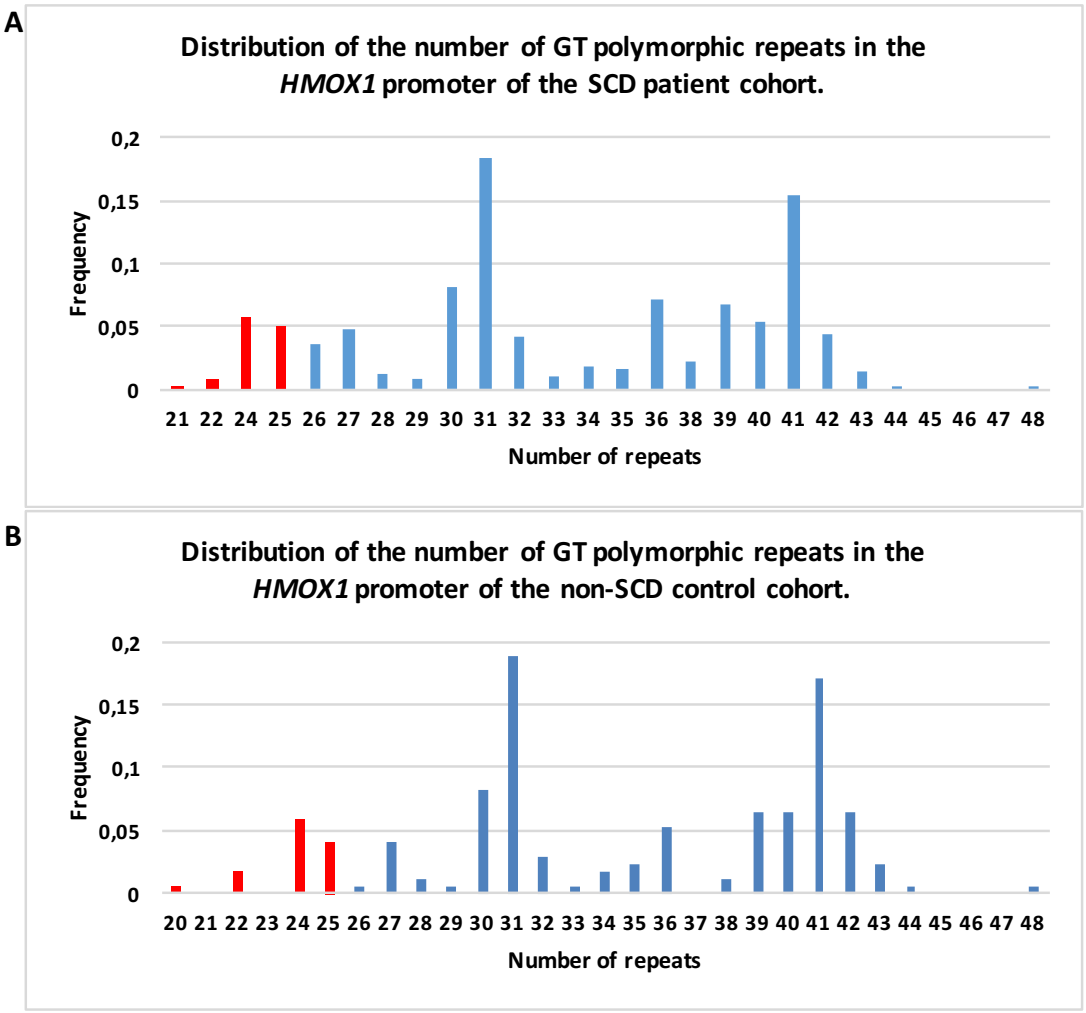


Figure 53: Histogram of distribution of *HMOX1* rs3074372 $(GT)_n$ promoter repeats in the (A) SCD patient and (B) non-SCD controls, with short repeats (</= 25) shown in red, and the long repeats (>25) shown in blue.

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.

Region	MAF						
	rs11912763	**rs16996648**	**rs1005570**	**rs16996672**	**rs743811**	**rs73885319**	**rs60910145**
SCD patient cohort (Cameroon)	**0.119**	**0.249**	**0.373**	**0.255**	**0.111**	**0.130**	**0.124**
Non-SCD control cohort (Cameroon)					**0.113**	**0.010**	**0.090**
ESKD case cohort (Cameroon)					**0.146**	**0.146**	**0.116**
ESKD control cohort (Cameroon)					**0.122**	**0.106**	**0.084**
African American (Southwest US)	0.156	0.23	0.385	0.262	0.156	0.221	0.221
Esan (Nigeria)	0.399	0.49	0.591	0.419	0.162	0.495	0.495
Yoruba (Kenya)	0.324	0.407	0.5	0.37	0.120	0.375	0.375
Europe	0.001	0.021	0.101	0.002	0.253	0	0
East Asia	0	0	0.067	0	0.664	0	0

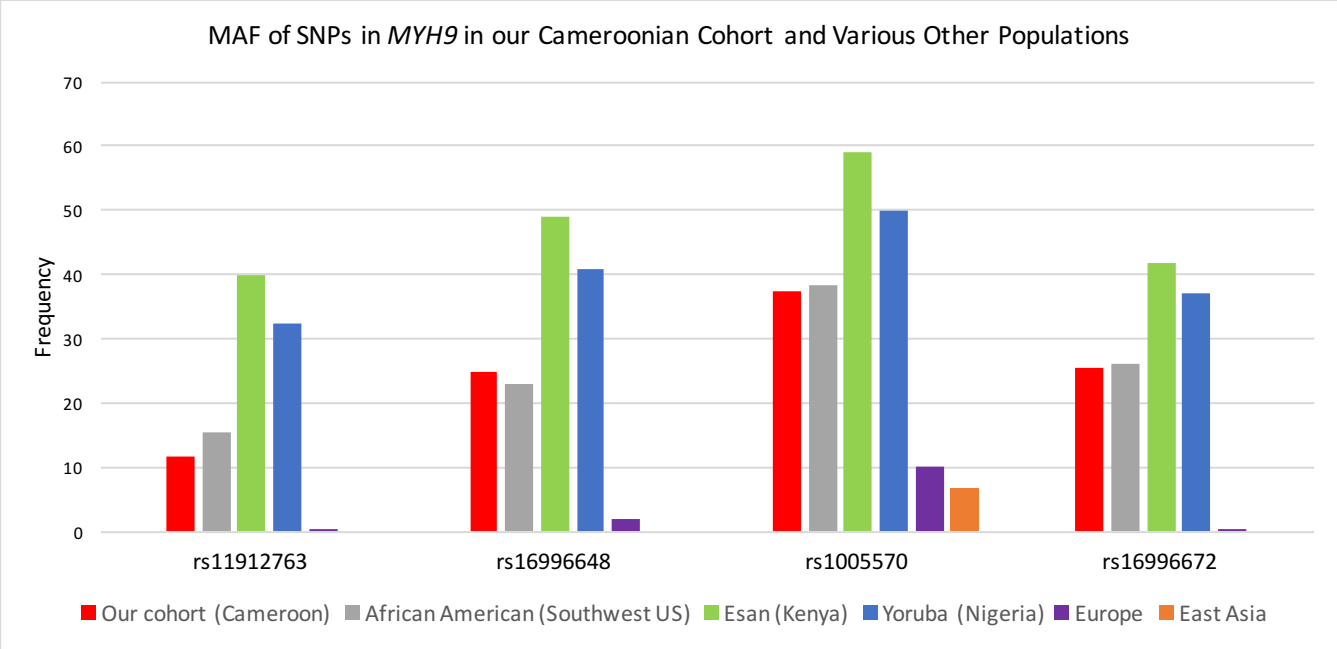


Figure 54: Comparison of MAF of *MYH9* SNPs between various populations.
Bar graph comparing the minor allele frequencies (MAF) of *MYH9* SNPs rs11912763, rs16996648, rs1005570, rs16996672 in the SCD patient cohort (Cameroon) compared to various African populations, Europe and East Asia. These MAFs were obtained from the 1000 Genome Project Data published in the Ensembl Genome browser website (http://www.ensembl.org/index.html).

The MAF of SNPs in *MYH9* (Figure 54) rs11912763 (0.118), rs16996672 (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.

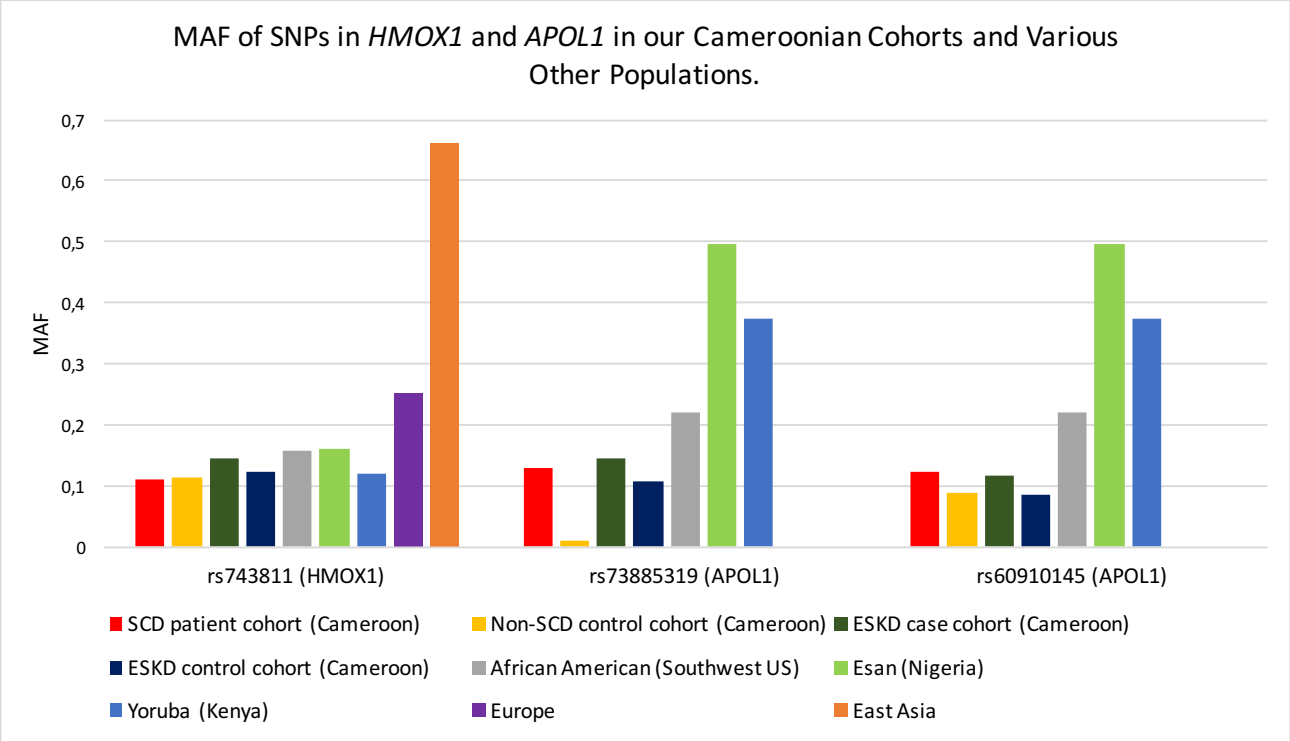


Figure 55: Comparison of MAF of *HMOX1* and *APOL1* SNPs between various populations.
Bar graph comparing the minor allele frequencies (MAF) of *HMOX1* SNP rs743811, and *APOL1* SNPs rs73885319 and rs60910145 in the SCD patient cohort (Cameroon), non-SCD controls, and ESKD case and control cohorts compared 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.

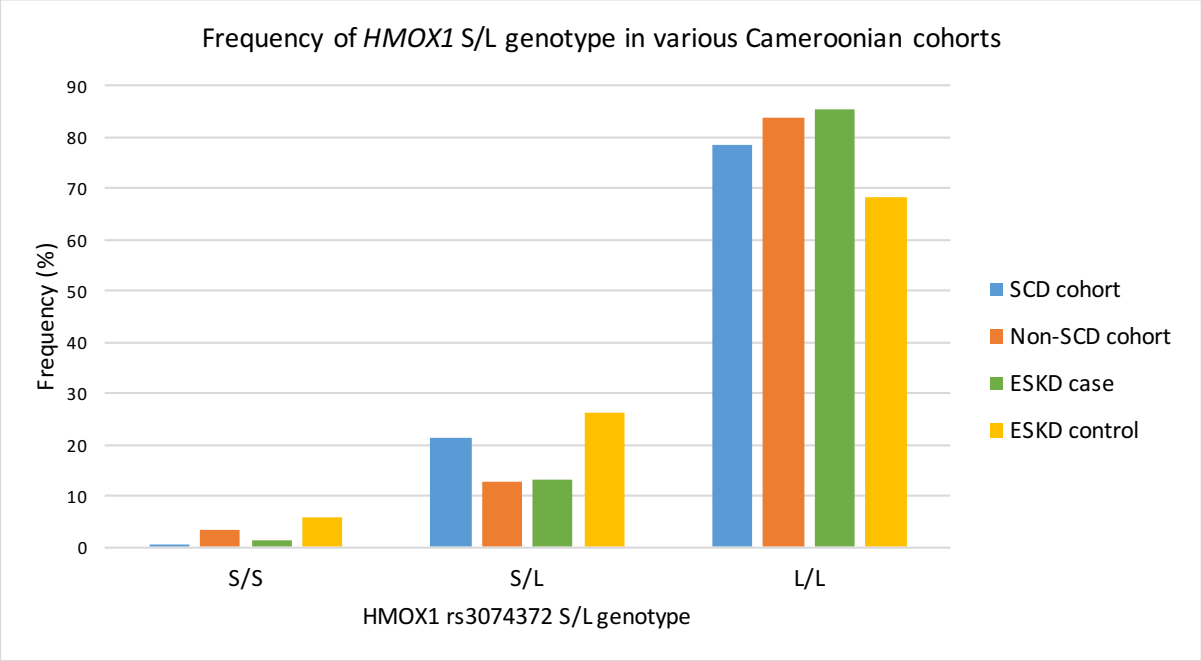


Figure 56: Comparison of genotype frequencies of *HMOX1* rs3074372 between our Cameroonian cohorts.

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.

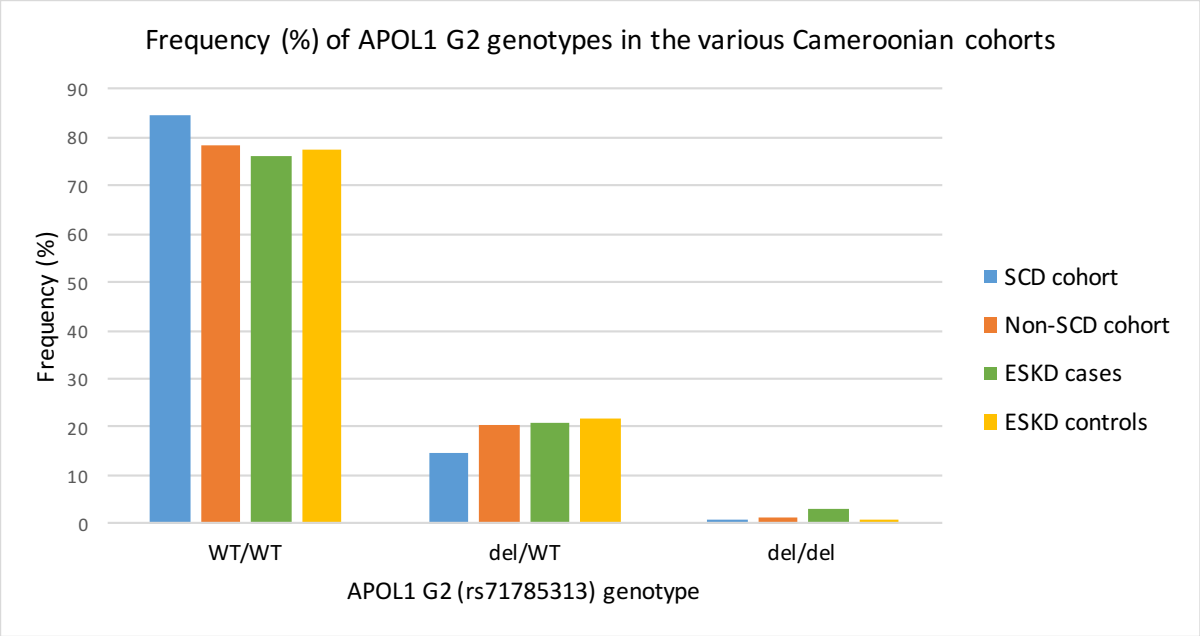


Figure 57: Comparison of genotype frequencies of *APOL1* G2 (rs71785313) between our Cameroonian cohorts.

Bar graph comparing the genotype frequencies for the 6 bp deletion (del) in *APOL1*, constituting the G2 variant (rs71785313), 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 ($p=0.2243$).

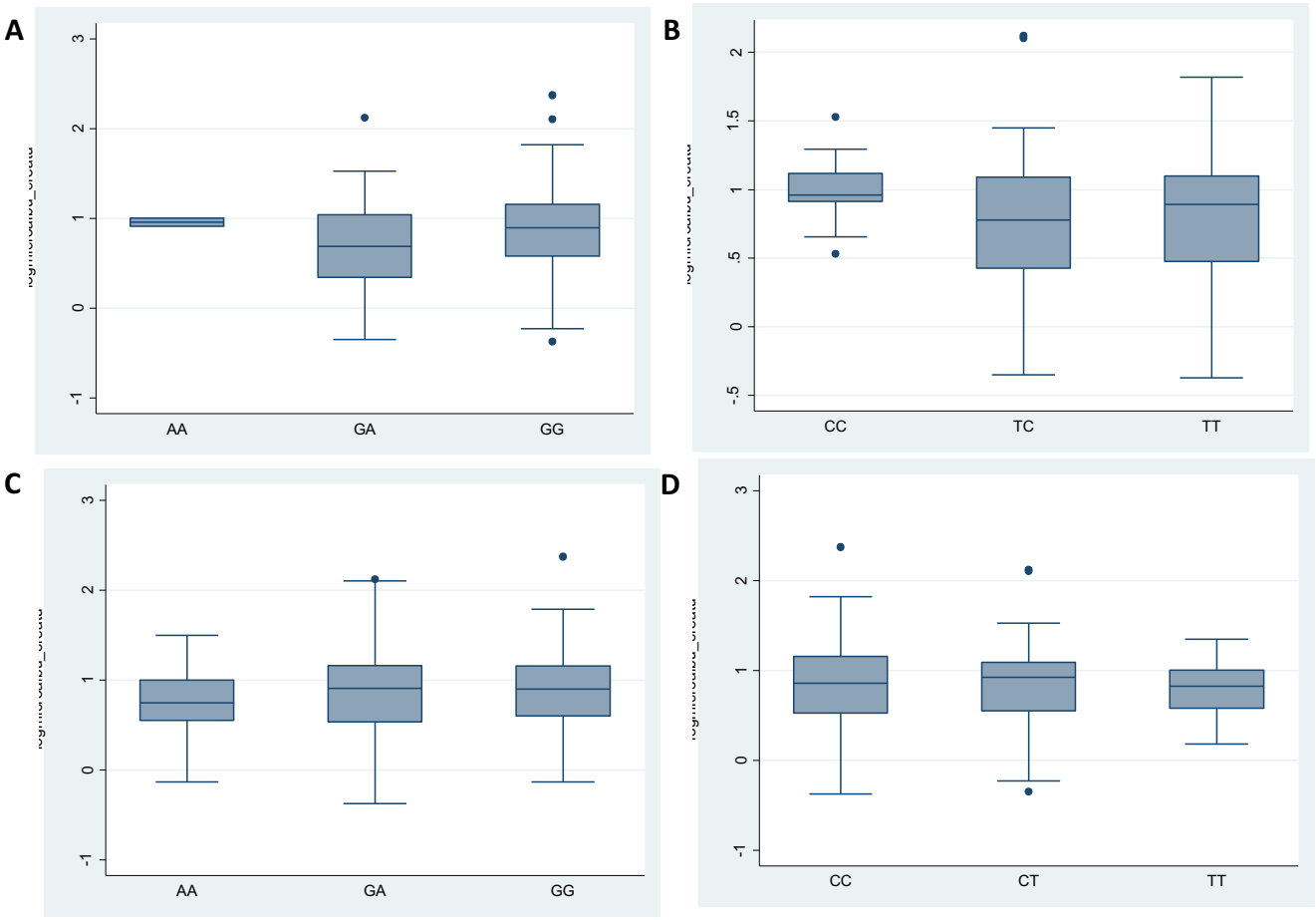


Figure 58: Box and whisker plots showing the association of log_{10}ACR values with each of the four target SNPs in *MYH9* in the general cohort.

Box and whisker plots illustrating the distribution of log ACR values conditioned on the SNP genotypes for the entire patient cohort. Log ACR values, conditioned on (A) the rs11912763 genotypes, (B) the rs16996648 genotypes, (C) the rs1005570 genotypes, and (D) the rs16996672 genotypes. 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 * ($p<0.05$).

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 ($p\cong0.05$) are highlighted in bold.

Polymorphism	Crude Albuminuria	Macro-albuminuria	Log(ACR)	eGFR (CKD-EPI)	eGFR (MDRD)	eGFR (CG)
***APOL1* G1/G2**	0.9239 (recessive)	**0.08537** (recessive)	**0.01829** (recessive) **0.05982** (codominant)	**0.07283** (recessive)	0.2434 (recessive)	0.1069 (recessive)
***HMOX1* – rs3074372**	0.4095 (codominant)	0.9218 (codominant)	0.8142 (codominant)	0.8521 (codominant)	0.6920 (codominant)	0.6946 (codominant)
***HMOX1* – rs743811**	**0.07372** (codominant) **0.08013** (recessive)	0.1549 (codominant) **0.0629** (recessive)	0.8455 (codominant)	0.2081 (codominant) 0.1169 (recessive)	**0.09865** (codominant) 0.1364 (recessive	0.1836 (codominant) 0.1391 (recessive)
3.7del α-globin gene genotypes	0.1371 (recessive)	**0.03441** (codominant) **0.051535** (dominant) **0.009707** (overdominant)	0.8098 (recessive)	0.4516 (codominant) 0.2318 (recessive)	0.7236 (codominant) 0.4320 (recessive)	0.6123 (codominant) 0.3292 (recessive)

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 overdominant model (p=0.009707). No significant association was observed with eGFR values for all three equations utilized.

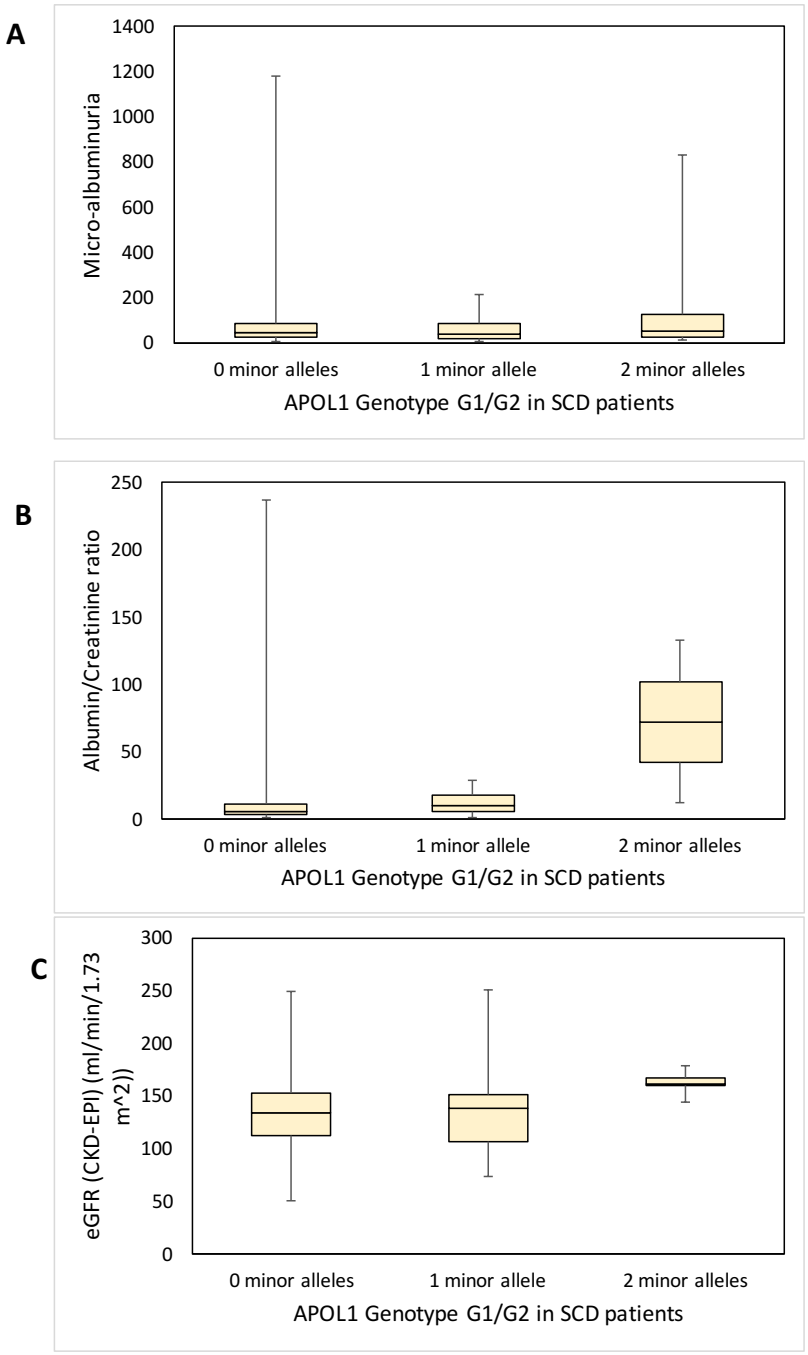


Figure 59: Box and whisker plots showing the association of *APOL1* G1/G2 with crude albuminuria, ACR, and eGFR values in the SCD patient cohort.

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 * ($p<0.05$).

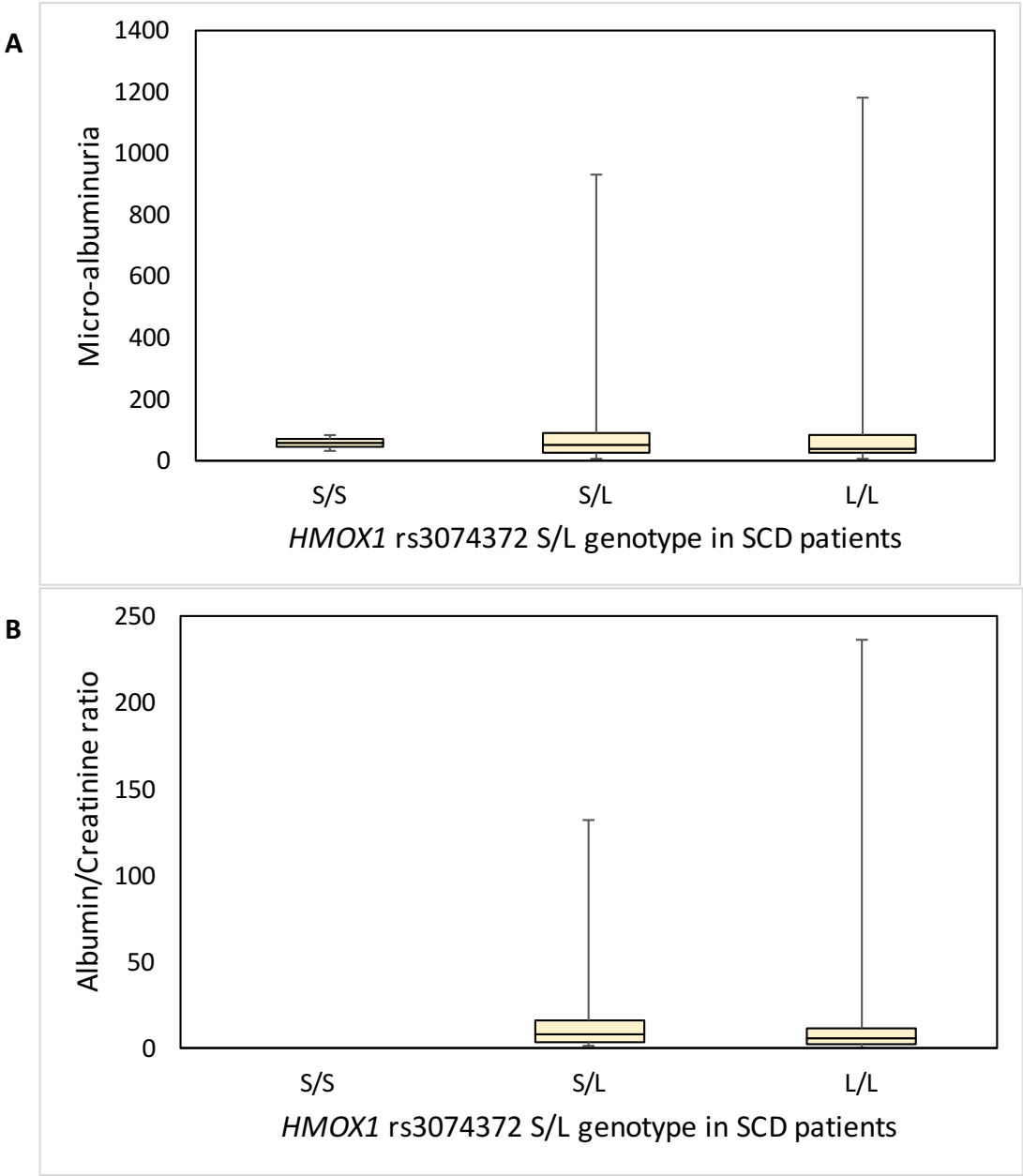


Figure 60: Box and whisker plots showing the association of *HMOX1* S/L Genotypes with crude albuminuria and ACR values in the SCD patient cohort.

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 * ($p<0.05$).

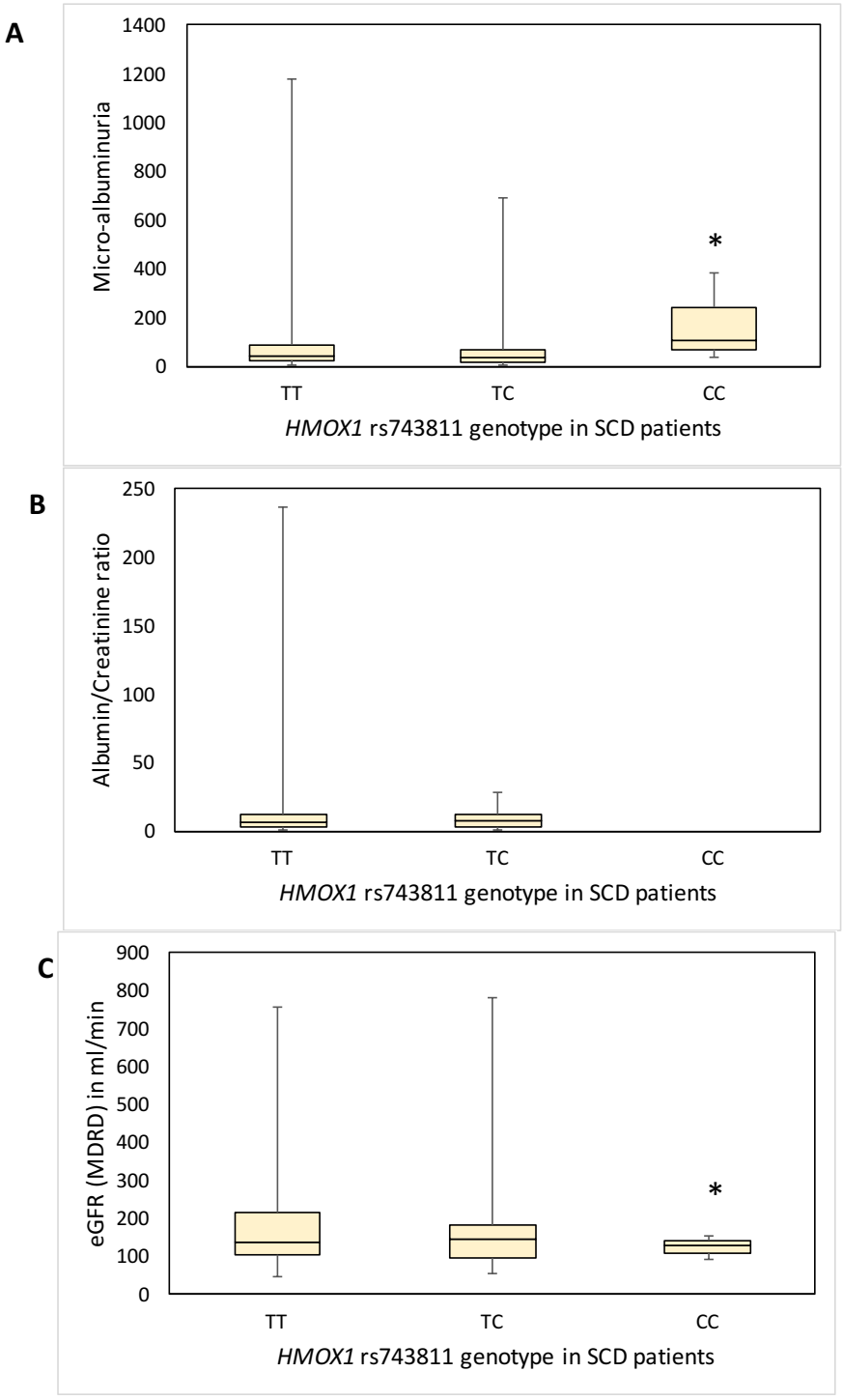


Figure 61: Box and whisker plots showing the association of *HMOX1* rs743811 Genotypes with crude albuminuria and ACR values in the SCD patient cohort.

Box and whisker plots illustrating the distribution of (A) crude albuminuria and (B) ACR values conditioned on the *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 ($p\cong0.05$) 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)

Characteristic	Cases (n=167)	Controls (n=162)	p-value
Mean age* (SD)	51.65 (12.60)	53.01 (11.67)	0.31
Age groups			0.749
≤40 years, n (%)	29 (17.4)	26 (16.0)	
> 40 years, n (%)	138 (82.6)	136 (84.0)	
Sex (F), n (%)	49 (29.3)	54 (33.3)	0.435
Ethnic Group			0.149
Bantu, n (%)	77 (46.1)	92 (56.8)	
Semi-Bantu, n (%)	72 (43.1)	57 (35.2)	
Sudanese, n (%)	18 (10.8)	13 (8.0)	
Blood pressure			
Mean SBP (SD)	147.36 (22.35)	140.92 (28.43)	0.023
Mean DBP (SD)	85.46 (12.53)	82.84 (13.86)	0.073
SES			0.036
Low, n (%)	5 (3.0)	1 (0.6)	
Middle, n (%)	48 (28.7)	65 (40.1)	
High, n (%)	114 (68.3)	96 (59.3)	
Family history of CKD, yes n (%)	19 (11.4)	12 (7.4)	0.218

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

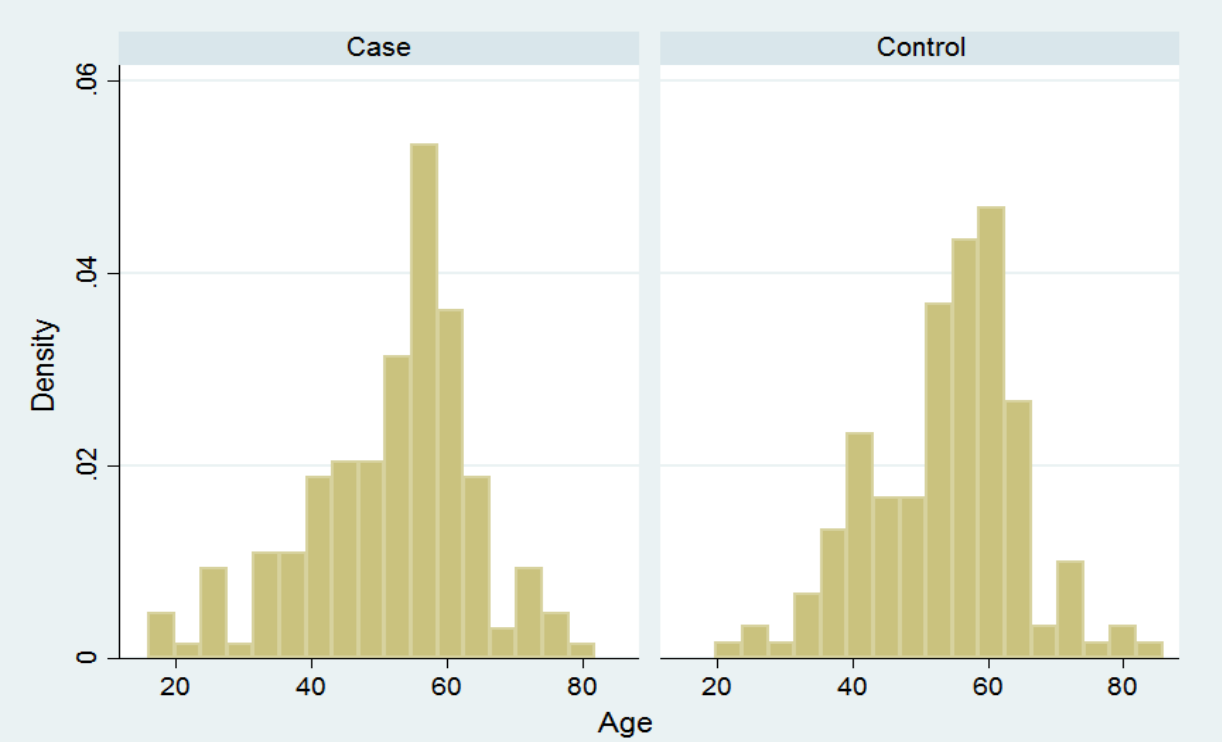


Figure 62: Histogram distribution of age in the ESKD cohort. The age distribution for the ESKD case (left, n=164) and control (right, n=155) patients from the general Cameroonian population.

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 ($p \cong 0.05$) are highlighted in bold.

Variables		**ESKD Cases**			**ESKD Controls**			
		Freq (%)	**HWE (p-value)**	**Number of observations**	**Freq (%)**	**HWE (p-value)**	**Number of observations**	**Association test p-value (model)**
HMOX1 – rs3074372 (S/L)	S/S	2 (1.45)	0.190	138	6 (5.77)	0.13	104	**0.002** (codominant) **0.0005** (dominant) **0.063** (recessive) **0.005** (overdominant)
	S/L	18 (13.04)			27 (25.96)			
	L/L	118 (85.51)			71 (68.27)			
HMOX1 – rs743811 (T/C)	TT	107 (74.31)	0.19	144	102 (77.86)	0.39	131	0.63 (codominant) 0.34 (dominant) 0.68 (recessive) 0.41 (overdominant)
	TC	32 (22.22)			26 (19.85)			
	CC	5 (3.47)			3 (2.29)			
APOL1 – rs73885319 (A/G)	AA	115 (76.16)	**0.0017**	151	113 (79.58)	0.60	142	
	AG	28 (18.54)			28 (19.72)			
	GG	8 (5.30)			1 (0.70)			
APOL1 – rs60910145 (T/G)	TT	119 (80.95)	**0.0011**	147	110 (83.97))	0.93	131	
	TG	22 (14.97)			20 (15.27)			
	GG	6 (4.08)			1 (0.76)			
APOL1 – rs71785313 (WT/del)	WT/ WT	105 (76.09)	0.27	138	78 (77.23)	0.69	101	
	Del/ WT	29 (21.01)			22 (21.78)			
	Del/del	4 (2.90)			1 (0.99)			
APOL1 – G1/G2	G/G							0.90 (codominant) 0.67 (dominant) 0.99 (recessive) 0.64 (overdominant)
	G/A							
	A/A							

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/1.73m^2 and >140 ml/min/1.73m^2 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 macro-albuminuria 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, in SCD.

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 m^2 among Jamaican adults with SCD (Asnani et al., 2016), that is much higher the normal expected levels of 0.75 ml/min/1.73 m^2 (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 *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 *APOL1*, and not *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 *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 *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 *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, *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 macro-albuminuria also approached significance in a recessive model (p=0.085). These results indicate that variants in *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 *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 *HMOX1*

When erythrocytes are lysed, extracellular hemoglobin is released and easily oxidized from ferrous (Fe^{2+}) to ferric (Fe^{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 *HMOX1* (Saraf et al., 2015). *HMOX1* rs743811 was associated with chronic kidney disease stage, and longer *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 *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 ($\leq$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 *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 *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 *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 *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 *HMOX1* (rs743811) with albuminuria and eGFR is in line with previous findings of *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 α-GLOBIN GENE DELETION AND SCD NEPHROPATHY

The 3.7kB α-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 *APOL1* G1/G2 and *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 *MYH9* SNPs and the *HMOX1* dinucleotide promoter polymorphism (rs3074372) may be unsuitable to detect any early stage of renal impairment in a Cameroonian SCD population. *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

Abo-Zenah H, Moharram M, El Nahas AM, 2009. Cardiorenal risk prevalence in sickle cell hemoglobinopathy. Nephron Clin. Pract. 112, c98--c106.

Adams Gaye T, Snieder Harold, McKie Virgil C, Clair Betsy, Brambilla Donald, Adams Robert J, Kutlar Ferdane, Kutlar Abdullah, 2003. Genetic risk factors for cerebrovascular disease in children with sickle cell disease: design of a case-control association study and genomewide screen. BMC Med. Genet. 4, 6. doi:10.1186/1471-2350-4-6

Adams RJ, Kutlar A, McKie V, Carl E, Nichols FT, Liu JC, McKie K, Clary A, 1994. Alpha thalassemia and stroke risk in sickle cell anemia. Am. J. Hematol. 45, 279–282.

Adams RJ, McKie VC, Hsu L, Files B, Vichinsky E, Pegelow C, Abboud M, Gallagher D, Kutlar A, Nichols FT, Bonds DR, Brambilla D, 1998. Prevention of a first stroke by transfusions in children with sickle cell anemia and abnormal results on transcranial Doppler ultrasonography. N. Engl. J. Med. 339, 5–11.

Adams RJ, Ohene-Frempong K, Wang W, 2001. Sickle cell and the brain. Hematology Am. Soc. Hematol. Educ. Program 31–46. doi:10.1182/asheducation-2001.1.31

Adedoyin OT, Adesiyun OO, Adegboye OA, Bello OA, Fatoye OP, 2012. Sickle cell Nephropathy in children seen in an African Hospital - Case Report. Niger. Postgrad. Med. J. 19, 119–122.

Alhashimi Dunia, Fedorowicz Zbys, Alhashimi Fatima, Dastgiri Saeed, 2010. Blood transfusions for treating acute chest syndrome in people with sickle cell disease. Cochrane Libr.

Aliyu ZY, Kato GJ, Taylor J 6th, Babadoko A, Mamman AI, Gordeuk VR, Gladwin MT, 2008. Sickle cell disease and pulmonary hypertension in Africa: a global perspective and review of epidemiology, pathophysiology, and management. Am. J. Hematol. 83, 63–70.

Aloni Michel Ntetani, Ngiyulu Ren?? Makwala, Gini-Ehungu Jean Lambert, Nsibu C??lestin Ndosimao, Ekila Mathilde Bothale, Lepira Fran??ois Bompeka, Nseka Nazaire Mangani, 2014. Renal function in children suffering from sickle cell disease: Challenge of early detection in highly resource-scarce settings. PLoS One 9, 1–5. doi:10.1371/journal.pone.0096561

Alvarez Ofelia, Lopez-Mitnik Gabriela, Zilleruelo Gaston, 2008. Short-term follow-up of patients with sickle cell disease and albuminuria. Pediatr. Blood Cancer 50, 1236–1239. doi:10.1002/pbc.21520

Alvarez Ofelia, Montane Brenda, Lopez Gabriela, Wilkinson James, Miller Tracie, 2006. Early blood transfusions protect against microalbuminuria in children with sickle cell disease. Pediatr. Blood Cancer 47, 71–76. doi:10.1002/pbc.20645

Aoki Regina Y, Saad Sara TO, 1995. Enalapril reduces the albuminuria of patients with sickle cell disease. Am. J. Med. 98, 432–435. doi:10.1016/S0002-9343(99)80341-6

Arlet JB, Ribeil JA, Chatellier G, Eladari D, Seigneux S De, Souberbielle JC, Friedlander G, de Montalembert M, Pouchot J, Prie D, Courbebaisse M, 2012. Determination of the best method to estimate glomerular filtration rate from serum creatinine in adult patients with sickle cell disease: a prospective observational cohort study. BMC Nephrol. 13, 83.

Arrondel Christelle, Vodovar Nicolas, Knebelmann Bertrand, Grünfeld Jean-Pierre, Gubler Marie-Claire, Antignac Corinne, Heidet Laurence, 2002. Expression of the nonmuscle myosin heavy chain IIA in the human kidney and screening for MYH9 mutations in Epstein and Fechtner syndromes. J. Am. Soc. Nephrol. 13, 65–74.

Ashley-Koch Allison E, Okocha Emmanuel C, Garrett Melanie E, Soldano Karen, De Castro Laura M, Jonassaint Jude C, Orringer Eugene P, Eckman James R, Telen Marilyn J, 2011. MYH9 and APOL1 are both associated with sickle cell disease nephropathy. Br. J. Haematol. 155, 386–394. doi:10.1111/j.1365-2141.2011.08832.x

Asnani Monika R, Lynch O'Neil, Reid Marvin E, 2013. Determining Glomerular Filtration Rate in Homozygous Sickle Cell Disease: Utility of Serum Creatinine Based Estimating Equations. PLoS One 8. doi:10.1371/journal.pone.0069922

Asnani Monika R, Reid Marvin E, 2015. Renal function in adult Jamaicans with homozygous sickle cell disease. Hematology 20, 422–428.

Asnani Monika, Reid Marvin, 2015. Cystatin C: A useful marker of glomerulopathy in sickle cell disease? Blood Cells, Mol. Dis. 54, 65–70. doi:10.1016/j.bcmd.2014.07.018

Asnani Monika, Serjeant Graham, Royal-Thomas Tamika, Reid Marvin, 2016. Predictors of renal function progression in adults with homozygous sickle cell disease. Br. J. Haematol. 173, 461–468. doi:10.1111/bjh.13967

Asnani MR, Fraser RA, Reid ME, 2011. Higher rates of hemolysis are not associated with albuminuria in Jamaicans with sickle cell disease. PLoS One 6, e18863.

Ataga Kenneth I, Derebail Vimal K, Archer David R, 2014. The glomerulopathy of sickle cell disease. Am. J. Hematol. 89, 907–914. doi:10.1002/ajh.23762

Atz AM, Wessel DL, 1997. Inhaled nitric oxide in sickle cell disease with acute chest syndrome. Anesthesiology 87, 988–990.

Aygun Banu, Mortier Nicole A, Smeltzer Matthew P, Hankins Jane S, Ware Russell E, 2011. Glomerular hyperfiltration and albuminuria in children with sickle cell anemia. Pediatr. Nephrol. 26, 1285–1290. doi:10.1007/s00467-011-1857-2

Aygun Banu, Mortier Nicole A, Smeltzer Matthew P, Shulkin Barry L, Hankins Jane S, Ware Russell E, 2013. Hydroxyurea treatment decreases glomerular hyperfiltration in children with sickle cell anemia. Am. J. Hematol. 88, 116–119. doi:10.1002/ajh.23365

Bains SK, Foresti R, Howard J, Atwal S, Green CJ, Motterlini R, 2010. Human Sickle Cell Blood Modulates Endothelial Heme Oxygenase Activity: Effects on Vascular Adhesion and Reactivity. Arterioscler. Thromb. Vasc. Biol. 30, 305–312. doi:10.1161/ATVBAHA.109.196360

Ballas Samir K, Kesen Muge R, Goldberg Morton F, Lutty Gerard A, Dampier Carlton, Osunkwo Ifeyinwa, Wang Winfred C, Hoppe Carolyn, Hagar Ward, Darbari Deepika S, 2012. Beyond the definitions of the phenotypic complications of sickle cell disease: an update on management. Sci. World J. 2012.

Barros Fabiana B, Lima Carmen SP, Santos Allan O, Mazo-Ruiz Mariana FC, Lima Mariana CL, Etchebehere Elba CSC, Costa Fernando F, Saad Sara TO, Camargo Edwaldo E, Ramos Celso D, 2006. 51Cr-EDTA measurements of the glomerular filtration rate in patients with sickle cell anaemia and minor renal damage. Nucl. Med. Commun. 27, 959–62. doi:10.1097/01.mnm.0000243373.03636.6e

Bartolucci Pablo, Habibi | Anoosha, Stehlé Thomas, Gaetana |, Liberto Di, Marie |, Rakotoson Georgine, Gellen-Dautremer Justine, Sylvain Loric |, Moutereau Stéphane, Sahali Dil, Wagner-Ballon Orianne, Remy Philippe, Lang Philippe, Grimbert Philippe, Audureau Etienne, Godeau Bertrand, Galacteros Frédéric, Audard Vincent, Hospitalier Groupe, Mondor- Henri, Chenevier Albert, 2015. Six Months of Hydroxyurea Reduces Albuminuria in Patients with Sickle Cell Disease. Public Heal. Unit EA J Am Soc Nephrol 27. doi:10.1681/ASN.2014111126

Bartolucci P, Galacteros F, 2012. Clinical management of adult sickle-cell disease. Curr. Opin.

Hematol. 19, 149–155.
Bauer Daniel E, Kamran SC, Lessard S, Xu J, Fujiwara Y, Lin C, Shao Z, Canver MC, Smith EC, Pinello L, Sabo PJ, Vierstra J, Voit RA, Yuan GC, Porteus MH, Stamatoyannopoulos JA, Lettre G, Orkin SH, 2013. An Erythroid Enhancer of BCL11A Subject to Genetic Variation Determines Fetal Hemoglobin Level. Science (80-.). 342, 253–257. doi:10.1126/science.1242088
Bean CJ, Boulet SL, Ellingsen D, Pyle ME, Barron-Casella E, Casella JF, Payne B, Driggers J, Trau H, Yang G, Jones K, Ofori-Acquah SF, Hooper WC, DeBaun MR, 2012. Heme oxygenase-1 gene promoter polymorphism is associated with reduced incidence of acute chest syndrome among children with sickle cell disease. Blood 120, 3822–3828. doi:10.1182/blood-2011-06-361642
Becker Amy M, Goldberg Jordan H, Henson Michael, Ahn Chul, Tong Liyue, Baum Michel, Buchanan George R, 2014. Blood pressure abnormalities in children with sickle cell anemia. Pediatr. Blood Cancer 61, 518–522.
Becton LJ, Kalpatthi R V, Rackoff E, Disco D, Orak JK, Jackson SM, Shatat IF, 2010. Prevalence and clinical correlates of microalbuminuria in children with sickle cell disease. Pediatr. Nephrol. 25, 1505–1511.
Behar DM, Rosset S, Tzur S, Selig S, Yudkovsky G, Bercovici S, Kopp JB, Winkler CA, Nelson GW, Wasser WG, Skorecki K, 2010. African ancestry allelic variation at the MYH9 gene contributes to increased susceptibility to non-diabetic end-stage kidney disease in Hispanic Americans. Hum. Mol. Genet. 19, 1816–1827.
Behar Doron M, Kedem Eynat, Rosset Saharon, Haileselassie Yonas, Tzur Shay, Kra-Oz Zipi, Wasser Walter G, Shenhar Yotam, Shahar Eduardo, Hassoun Gamal, others, 2011. Absence of APOL1 risk variants protects against HIV-associated nephropathy in the Ethiopian population. Am. J. Nephrol. 34, 452–459.
Belcher John D, 2003. Transgenic sickle mice have vascular inflammation. Blood 101, 3953–3959. doi:10.1182/blood-2002-10-3313
Belcher John D, Mahaseth Hemachandra, Welch Thomas E, Otterbein Leo E, Hebbel Robert P, Vercellotti Gregory M, 2006. Heme oxygenase-1 is a modulator of inflammation and vaso-occlusion in transgenic sickle mice. J. Clin. Invest. 116, 808–816. doi:10.1172/JCI26857
Bernaudin F, Socie G, Kuentz M, Chevret S, Duval M, Bertrand Y, Vannier JP, Yakouben K, Thuret I, Bordigoni P, Fischer A, Lutz P, Stephan JL, Dhedin N, Plouvier E, Margueritte G, Bories D, Verlhac S, Esperou H, Coic L, Vernant JP, Gluckman E, SFGM-TC, 2007. Long-term results of related myeloablative stem-cell transplantation to cure sickle cell disease. Blood 110, 2749–2756.
Bitoungui Valentina J Ngo, Pule Gift D, Hanchard Neil, Ngogang Jeanne, Wonkam Ambroise, 2015. Beta-globin gene haplotypes among cameroonians and review of the global distribution: is there a case for a single sickle mutation origin in Africa? Omi. a J. Integr. Biol. 19, 171–179.
Bleyer Anthony J, Reddy Sri V, Sujata Leon, Russell Gregory B, Akinnifesi Damilola, Hire Donald, Shihabi Zak, Knovich Mary Ann, Daeihagh Pirouz, Calles Jorge, Freedman Barry I, 2010. Sickle cell trait and development of microvascular complications in diabetes mellitus. Clin. J. Am. Soc. Nephrol. 5, 1015–20. doi:10.2215/CJN.08841209
Bolarinwa RA, Akinlade KS, Kuti Ma O, Olawale OO, Akinola NO, 2012. Renal disease in adult Nigerians with sickle cell anemia: a report of prevalence, clinical features and risk factors. Saudi J. kidney Dis. Transplant. 23, 171–175.

Canver Matthew C, Smith Elenoe C, Sher Falak, Pinello Luca, Sanjana Neville E, Shalem Ophir, Chen Diane D, Schupp Patrick G, Vinjamur Divya S, Garcia Sara P, Luc Sidinh, Kurita Ryo, Nakamura Yukio, Fujiwara Yuko, Maeda Takahiro, Yuan Guo-Cheng, Zhang Feng, Orkin Stuart H, Bauer Daniel E, 2015. BCL11A enhancer dissection by Cas9-mediated in situ saturating mutagenesis. Nature 527, 192–7. doi:10.1038/nature15521

Castro O, Brambilla DJ, Thorington B, Reindorf CA, Scott RB, Gillette P, Vera JC, Levy PS, 1994. The acute chest syndrome in sickle cell disease: incidence and risk factors. The Cooperative Study of Sickle Cell Disease. Blood 84, 643–649.

Charache S, Terrin ML, Moore RD, Dover GJ, Barton FB, Eckert S V, McMahon RP, Bonds DR, 1995. Effect of hydroxyurea on the frequency of painful crises in sickle cell anemia. Investigators of the Multicenter Study of Hydroxyurea in Sickle Cell Anemia. N. Engl. J. Med. 332, 1317–1322.

Chintagari Narendranath Reddy, Nguyen Julia, Belcher John D, Vercellotti Gregory M, Alayash Abdu I, 2015. Haptoglobin attenuates hemoglobin-induced heme oxygenase-1 in renal proximal tubule cells and kidneys of a mouse model of sickle cell disease. Blood Cells, Mol. Dis. 54, 302–306. doi:10.1016/j.bcmd.2014.12.001

Cockcroft Donald W, Gault M Henry, 1976. Prediction of creatinine clearance from serum creatinine. Nephron 16, 31–41.

Cox Sharon E, Makani Julie, Newton Charles R, Prentice Andrew M, Kirkham Fenella J, 2013. Hematological and genetic predictors of daytime hemoglobin saturation in tanzanian children with and without sickle cell anemia. ISRN Hematol. 2013.

Day TG, Drasar ER, Fulford T, Sharpe CC, Thein SL, 2012. Association between hemolysis and albuminuria in adults with sickle cell anemia. Haematologica 97, 201–205.

Dharnidharka VR, Dabbagh S, Atiyeh B, Simpson P, Sarnaik S, 1998. Prevalence of microalbuminuria in children with sickle cell disease. Pediatr. Nephrol. Berlin Ger. 12, 475–478.

Diagne I, Ndiaye O, Moreira C, Signate-Sy H, Camara B, Diouf S, Diack-Mbaye A, Ba M, Sarr M, Sow D, Fall M, 2000. Les syndromes drépanocytaires majeurs en pédiatrie à Dakar (Sénégal). Arch. Pédiatrie 7, 16–24.

Diallo Dapa, Tchernia Gil, 2002. Sickle cell disease in Africa. Curr. Opin. Hematol. 9, 111–116. doi:10.1097/00062752-200203000-00005

Diop S, Thiam D, Cisse M, Toure-Fall a O, Fall K, Diakhate L, 1999. New results in clinical severity of homozygous sickle cell anemia, in Dakar, Senegal. Hematol. Cell Ther. 41, 217–221.

Earley CJ, Kittner SJ, Feeser BR, Gardner J, Epstein a, Wozniak Ma, Wityk R, Stern BJ, Price TR, Macko RF, Johnson C, Sloan Ma, Buchholz D, 1998. Stroke in children and sickle-cell disease: Baltimore-Washington Cooperative Young Stroke Study. Neurology 51, 169–176. doi:10.1212/WNL.51.1.169

Eke Christopher Bismarck, Okafor Henrietta Uche, Ibe Bede Chidozie, 2012. Prevalence and correlates of microalbuminuria in children with sickle cell anaemia: Experience in a tertiary health facility in Enugu, Nigeria. Int. J. Nephrol. doi:10.1155/2012/240173

Eknoyan G, Levin NW, 2002. K/DOQI clinical practice guidelines for chronic kidney disease: Evaluation, classification, and stratification-Foreword. Am. J. Kidney Dis. 39, S14–S266.

Enosolease ME, Ejele OA, Awodu OA, 2005. The influence of foetal haemoglobin on the frequency of vaso-occlusive crisis in sickle cell anaemia patients. Niger. Postgrad. Med. J. 12, 102–105.

Fatunde OJ, Scott-Emuakpor AB, 1992. Foetal haemoglobin in Nigerian children with sickle

cell anaemia. Effect on haematological parameters and clinical severity. Trop. Geogr. Med. 44, 264–266.
Flanagan JM, Sheehan V, Linder H, Howard TA, Wang YD, Hoppe CC, Aygun B, Adams RJ, Neale GA, Ware RE, 2013. Genetic mapping and exome sequencing identify 2 mutations associated with stroke protection in pediatric patients with sickle cell anemia. Blood 121, 3237–3245.
Flanagan Jonathan M, Frohlich Denise M, Howard Thad a, Schultz William H, Driscoll Catherine, Nagasubramanian Ramamoorthy, Mortier Nicole a, Kimble Amy C, Aygun Banu, Adams Robert J, Helms Ronald W, Ware Russel E, 2011. Genetic predictors for sroke in children with sickle cell anemia. Blood J. Am. Soc. Hematol. 117, 6681–6684. doi:10.1182/blood-2011-01-332205.The
Flint Jonathan, Harding Rosalind M, Boyce Anthony J, Clegg John B, 1998. The population genetics of the haemoglobinopathies. Baillieres. Clin. Haematol. 11, 1–51.
Frackman Susan, Kobs Gary, Simpson Dan, Storts Doug, 1998. Betaine and DMSO: enhancing agents for PCR. Promega notes 65, 27–29.
Freedman Barry I, Hicks Pamela J, Bostrom Meredith A, Comeau Mary E, Divers Jasmin, Bleyer Anthony J, Kopp Jeffrey B, Winkler Cheryl A, Nelson George W, Langefeld Carl D, Bowden Donald W, 2009. Non-muscle myosin heavy chain 9 gene MYH9 associations in African Americans with clinically diagnosed type 2 diabetes mellitus-associated ESRD. Nephrol. Dial. Transplant. 24, 3366–3371. doi:10.1093/ndt/gfp316
Freedman Barry I, Hicks Pamela J, Bostrom Meredith A, Cunningham Mary E, Liu Yongmei, Divers Jasmin, Kopp Jeffrey B, Winkler Cheryl A, Nelson George W, Langefeld Carl D, Bowden Donald W, 2009. Polymorphisms in the non-muscle myosin heavy chain 9 gene (MYH9) are strongly associated with end-stage renal disease historically attributed to hypertension in African Americans. Kidney Int. 75, 736–745. doi:10.1038/ki.2008.701
Gabriel Abram, Przybylski J, 2010. Sickle-cell anemia: a look at global haplotype distribution.
Genovese Giulio, Friedman David J, Ross Michael D, Lecordier Laurence, Uzureau Pierrick, Freedman Barry I, Bowden Donald W, Langefeld Carl D, Oleksyk Taras K, Uscinski Knob Andrea L, Bernhardy Andrea J, Hicks Pamela J, Nelson George W, Vanhollebeke Benoit, Winkler Cheryl A, Kopp Jeffrey B, Pays Etienne, Pollak Martin R, 2010a. Association of Trypanolytic ApoL1 Variants with Kidney Disease in African Americans. Science (80-.). 329, 841–845. doi:10.1126/science.1193032
Genovese Giulio, Tonna Stephen J, Knob Andrea U, Appel Gerald B, Katz Avi, Bernhardy Andrea J, Needham Alexander W, Lazarus Ross, Pollak Martin R, 2010b. A risk allele for focal segmental glomerulosclerosis in African Americans is located within a region containing APOL1 and MYH9. Kidney Int. 78, 698–704.
Ghosh Samit, Tan Fang, Yu Tianwei, Li Yuhua, Adisa Olufolake, Mosunjac Mario, Ofori-Acquah Solomon F, 2011. Global Gene Expression Profiling of Endothelium Exposed to Heme Reveals an Organ-Specific Induction of Cytoprotective Enzymes in Sickle Cell Disease. PLoS One 6. doi:10.1371/journal.pone.0018399
Gladwin Mark T, Sachdev Vandana, 2012. Cardiovascular abnormalities in sickle cell disease. J. Am. Coll. Cardiol. 59, 1123–1133. doi:10.1016/j.jacc.2011.10.900
Gladwin MT, Sachdev V, Jison ML, Shizukuda Y, Plehn JF, Minter K, Brown B, Coles WA, Nichols JS, Ernst I, Hunter LA, Blackwelder WC, Schechter AN, Rodgers GP, Castro O, Ognibene FP, 2004. Pulmonary hypertension as a risk factor for death in patients with sickle cell disease. N. Engl. J. Med. 350, 886–895.
Gosmanova Elvira O, Zaidi Sahar, Wan Jim Y, Adams-Graves Patricia E, 2014. Prevalence and

progression of chronic kidney disease in adult patients with sickle cell disease. J. Investig. Med. 62, 804–7. doi:10.1097/01.JIM.0000446836.75352.72
Guasch A, Cua M, Mitch WE, 1996. Early detection and the course of glomerular injury in patients with sickle cell anemia. Kidney Int. 49, 786–791.
Guasch A, Cua M, You W, Mitch WE, 1997. Sickle cell anemia causes a distinct pattern of glomerular dysfunction. Kidney Int. 51, 826–833. doi:10.1038/ki.1997.116
Guasch a, Zayas CF, Eckman JR, Muralidharan K, Zhang W, Elsas LJ, 1999. Evidence that microdeletions in the alpha globin gene protect against the development of sickle cell glomerulopathy in humans. J. Am. Soc. Nephrol. 10, 1014–1019.
Guasch Antonio, Navarrete Jose, Nass Kaleed, Zayas Carlos F, 2006. Glomerular involvement in adults with sickle cell hemoglobinopathies: Prevalence and clinical correlates of progressive renal failure. J. Am. Soc. Nephrol. 17, 2228–2235. doi:10.1681/ASN.2002010084
Gurdasani Deepti, Carstensen Tommy, Tekola-Ayele Fasil, Pagani Luca, Tachmazidou Ioanna, Hatzikotoulas Konstantinos, Karthikeyan Savita, Iles Louise, Pollard Martin O, Choudhury Ananyo, 2015. The African Genome Variation Project shapes medical genetics in Africa. Nature 517, 327–332.
Haymann JP, Stankovic K, Levy P, Avellino V, Tharaux PL, Letavernier E, Grateau G, Baud L, Girot R, Lionnet F, 2010. Glomerular hyperfiltration in adult sickle cell anemia: a frequent hemolysis associated feature. Clin. J. Am. Soc. Nephrol. 5, 756–761.
Hirai Hisao, Kubo Hiroshi, Yamaya Mutsuo, Nakayama Katsutoshi, Numasaki Muneo, Kobayashi Seiichi, Suzuki Satoshi, 2003. Brief report Microsatellite polymorphism in heme oxygenase-1 gene promoter is associated with susceptibility to oxidant-induced apoptosis in lymphoblastoid cell lines. Apoptosis 102, 2002–2004. doi:10.1182/blood-2002-12-3733.BLOOD
Hirschberg Raimund, 2010. Glomerular hyperfiltration in sickle cell disease. Clin. J. Am. Soc. Nephrol. 5, 748–749. doi:10.2215/CJN.01340210
Hsu LL, Miller ST, Wright E, Kutlar A, McKie V, Wang W, Pegelow CH, Driscoll C, Hurlet A, Woods G, Elsas L, Embury S, Adams RJ, (STOP) Stroke Prevention Trial, the Cooperative Study of Sickle Cell Disease (CSSCD), 2003. Alpha Thalassemia is associated with decreased risk of abnormal transcranial Doppler ultrasonography in children with sickle cell anemia. J. Pediatr. Hematol. Oncol. 25, 622–628.
Imuetinyan Blessing Abhulimen-Iyoha, Okoeguale Michael Ibadin, Egberue Gabriel Ofovwe, 2011. Microalbuminuria in Children with Sickle Cell Anemia. Saudi J. Kidney Dis. Transplant. 2, 733–738. doi:10.5144/0256-4947.2012.243
Iwalokun BA, Iwalokun SO, Hodonu SO, Aina OA, Agomo PU, 2012. Evaluation of microalbuminuria in relation to asymptomatic bacteruria in Nigerian patients with sickle cell anemia. Saudi J. Kidney Dis. Transpl. 23, 1320–1330. doi:10.4103/1319-2442.103589
Jison Maria L, Munson Peter J, Barb Jennifer J, Suffredini Anthony F, Talwar Shefali, Logun Carolea, Raghavachari Nalini, Beigel John H, Shelhamer James H, Danner Robert L, Gladwin Mark T, 2004. Blood mononuclear cell gene expression profiles characterize the oxidant, hemolytic, and inflammatory stress of sickle cell disease. Blood 104, 270–80. doi:10.1182/blood-2003-08-2760
Kanakiriya Sharan KR, Croatt Anthony J, Haggard Jill J, Ingelfinger Julie R, Tang Shiow-shih, Alam Jawed, Nath Karl A, 2003. Heme: a novel inducer of MCP-1 through HO-dependent and HO-independent mechanisms. Am. J. Physiol. - Ren. Physiol. 284, F546–

F554. doi:10.1152/ajprenal.00298.2002
Kao WH, Klag MJ, Meoni LA, Reich D, Berthier-Schaad Y, Li M, Coresh J, Patterson N, Tandon A, Powe NR, Fink NE, Sadler JH, Weir MR, Abboud HE, Adler SG, Divers J, Iyengar SK, Freedman BI, Kimmel PL, Knowler WC, Kohn OF, Kramp K, Leehey DJ, Nicholas SB, Pahl M V, Schelling JR, Sedor JR, Thornley-Brown D, Winkler CA, Smith MW, Parekh RS, of Nephropathy Family Investigation, Group Diabetes Research, 2008. MYH9 is associated with nondiabetic end-stage renal disease in African Americans. Nat. Genet. 40, 1185–1192.
Kasembeli AN, Duarte R, Ramsay M, Mosiane P, Dickens C, Dix-Peek T, Limou S, Sezgin E, Nelson GW, Fogo AB, Goetsch S, Kopp JB, Winkler CA, Naicker S, 2015a. APOL1 Risk Variants Are Strongly Associated with HIV-Associated Nephropathy in Black South Africans. J. Am. Soc. Nephrol.
Kasembeli AN, Duarte R, Ramsay M, Naicker S, 2015b. African origins and chronic kidney disease susceptibility in the human immunodeficiency virus era. World J. Nephrol. 4, 295–306.
Kaul DK, Hebbel RP, 2000. Hypexia/reoxygenation causes inflammatory response in transgenic sickle mice but not in normal mice. J. Clin. Invest. 106, 411–420. doi:10.1172/JCI9225
Kimpara T, Takeda a, Watanabe K, Itoyama Y, Ikawa S, Watanabe M, Arai H, Sasaki H, Higuchi S, Okita N, Takase S, Saito H, Takahashi K, Shibahara S, 1997. Microsatellite polymorphism in the human heme oxygenase-1 gene promoter and its application in association studies with Alzheimer and Parkinson disease. Hum. Genet. 100, 145–7.
Kopp JB, Smith MW, Nelson GW, Johnson RC, Freedman BI, Bowden DW, Oleksyk T, McKenzie LM, Kajiyama H, Ahuja TS, Berns JS, Briggs W, Cho ME, Dart RA, Kimmel PL, Korbet SM, Michel DM, Mokrzycki MH, Schelling JR, Simon E, Trachtman H, Vlahov D, Winkler CA, 2008. MYH9 is a major-effect risk gene for focal segmental glomerulosclerosis. Nat. Genet. 40, 1175–1184.
Kopp Jeffrey B, Nelson George W, Sampath Karmini, Johnson Randall C, Genovese Giulio, An Ping, Friedman David, Briggs William, Dart Richard, Korbet Stephen, Mokrzycki Michele H, Kimmel Paul L, Limou Sophie, Ahuja Tejinder S, Berns Jeffrey S, Fryc Justyna, Simon Eric E, Smith Michael C, Trachtman Howard, Michel Donna M, Schelling Jeffrey R, Vlahov David, Pollak Martin, Winkler Cheryl A, 2011. APOL1 genetic variants in focal segmental glomerulosclerosis and HIV-associated nephropathy. J. Am. Soc. Nephrol. 22, 2129–37. doi:10.1681/ASN.2011040388
Lamarre Yann, Romana Marc, Lemonne Nathalie, Hardy-Dessources Marie Dominique, Tarer Vanessa, Mougenel Danielle, Waltz Xavier, Tressi??res Beno??t, Lalanne-Mistrih Marie Laure, Etienne-Julan Maryse, Connes Philippe, 2014. Alpha thalassemia protects sickle cell anemia patients from macro-albuminuria through its effects on red blood cell rheological properties. Clin. Hemorheol. Microcirc. 57, 63–72. doi:10.3233/CH-131772
Laurin Louis Philippe, Nachman Patrick H, Desai Payal C, Ataga Kenneth I, Derebail Vimal K, 2014. Hydroxyurea is associated with lower prevalence of albuminuria in adults with sickle cell disease. Nephrol. Dial. Transplant. 29, 1211–1218. doi:10.1093/ndt/gft295
Lettre G, Sankaran VG, Bezerra MA, Araujo AS, Uda M, Sanna S, Cao A, Schlessinger D, Costa FF, Hirschhorn JN, Orkin SH, 2008. DNA polymorphisms at the BCL11A, HBS1L-MYB, and beta-globin loci associate with fetal hemoglobin levels and pain crises in sickle cell disease. Proc. Natl. Acad. Sci. U. S. A. 105, 11869–11874.
Levey Andrew S, Becker Cassandra, Inker Lesley A, 2013. Glomerular Filtration Rate and

Albuminuria for Detection and Staging of Acute and Chronic Kidney Disease in Adults: A Systematic Review. JAMA 313, 837–846. doi:10.1001/jama.2015.0602.Glomerular

Levey Andrew S, Bosch JP, Lewis JB, Greene T, Rogers N, Roth D, 1999. A more accurate method to estimate glomerular filtration rate from serum creatinine: a new prediction equation. Modification of Diet in Renal Disease Study Group. Ann. Intern. Med. 130, 461–470.

Levey Andrew S, Coresh Josef, Balk Ethan, Kausz Annamaria T, Levin Adeera, Steffes Michael W, Hogg Ronald J, Perrone Ronald D, Lau Joseph, Eknoyan Garabed, 2003. National Kidney Foundation practice guidelines for chronic kidney disease: evaluation, classification, and stratification. Ann. Intern. Med. 139, 137–147.

Levey Andrew S, Eckardt Kai Uwe, Tsukamoto Yusuke, Levin Adeera, Coresh Josef, Rossert Jerome, De Zeeuw Dick, Hostetter Thomas H, Lameire Norbert, Eknoyan Garabed, Willis Kerry, 2005. Definition and classification of chronic kidney disease: A position statement from Kidney Disease: Improving Global Outcomes (KDIGO)z. Kidney Int. 67, 2089–2100. doi:10.1111/j.1523-1755.2005.00365.x

Levey Andrew S, Stevens Lesley A, Schmid Christopher H, Zhang Yaping, Castro Alejandro F, Feldman Harold I, Kusek John W, Eggers Paul, Van Lente Frederick, Greene Tom, Coresh Josef, (CKD-EPI) Chronic Kidney Disease Epidemiology Collaboration, 2009. A New Equation to Estimate Glomerular Filtration Rate 150, 604–612.

Limou Sophie, Nelson George W, Kopp Jeffrey B, Winkler Cheryl A, 2014. APOL1 Kidney Risk Alleles: Population Genetics and Disease Associations. Adv. Chronic Kidney Dis. 21, 426–433. doi:10.1053/j.ackd.2014.06.005

Lindeman Robert D, Tobin Jordan, Shock Nathan W, 1985. Longitudinal studies on the rate of decline in renal function with age. J. Am. Geriatr. Soc. 33, 278–285.

Maines Mahin D, 1997. THE HEME OXYGENASE SYSTEM: A Regulator of Second Messenger Gases. Annu. Rev. Pharmacol. Toxicol 37, 517–54. doi:10.1146/annurev.pharmtox.37.1.517

Makani J, Williams TN, Marsh K, 2007. Sickle cell disease in Africa: burden and research priorities. Ann. Trop. Med. Parasitol. 101, 3–14. doi:10.1179/136485907X154638

Mawanda M, Ssenkusu JM, Odiit A, Kiguli S, Muyingo A, Ndugwa C, 2011. Micro-albuminuria in Ugandan children with sickle cell anaemia: a cross-sectional study. Ann. Trop. Paediatr. 31, 115–121. doi:10.1179/1465328111Y.0000000013

McBurney PG, Hanevold CD, Hernandez CM, Waller JL, McKie KM, 2002. Risk factors for microalbuminuria in children with sickle cell anemia. J Pediatr Hematol Oncol 24, 473–477.

Mckie Kathleen T, Hanevold Coral D, Hernandez Caterina, Waller Jennifer L, Ortiz Luis, Mckie Kathleen M, 2007. Prevalence , Prevention , and Treatment of Cell Disease 30912, 140–144.

McPherson Yee Marianne, Jabbar Shameem F, Osunkwo Ifeyinwa, Clement Lisa, Lane Peter A, Eckman James R, Guasch Antonio, 2011. Chronic Kidney Disease and Albuminuria in Children with Sickle Cell Disease. Clin. J. Am. Soc. Nephrol. 6, 2628–2633. doi:10.2215/CJN.01600211

Morita T, Kourembanas S, 1995. Endothelial cell expression of vasoconstrictors and growth factors is regulated by smooth muscle cell-derived carbon monoxide. J. Clin. Invest. 96, 2676–2682. doi:10.1172/JCI118334

Muller-Eberhard U, Javid J, Liem HH, Hanstein A, Hanna M, 1968. Plasma Concentrations of Hemopexin, Haptoglobin and Heme in Patients with Various Hemolytic Diseases. Blood

32, 811–816.
Naik Rakhi P, Derebail Vimal K, Grams Morgan E, Franceschini Nora, Auer Paul L, Peloso Gina M, Young Bessie a, Lettre Guillaume, Peralta Carmen a, Katz Ronit, Hyacinth Hyacinth I, Quarells Rakale C, Grove Megan L, Bick Alexander G, Fontanillas Pierre, Rich Stephen S, Smith Joshua D, Boerwinkle Eric, Rosamond Wayne D, Ito Kaoru, Lanzkron Sophie, Coresh Josef, Correa Adolfo, Sarto Gloria E, Key Nigel S, Jacobs David R, Kathiresan Sekar, Bibbins-Domingo Kirsten, Kshirsagar Abhijit V., Wilson James G, Reiner Alexander P, 2014. Association of Sickle Cell Trait With Chronic Kidney Disease and Albuminuria in African Americans. Jama 312, 2115. doi:10.1001/jama.2014.15063
Nath KA, Grande JP, Haggard JJ, Croatt AJ, Katusic ZS, Solovey A, Hebbel RP, 2001. Oxidative stress and induction of heme oxygenase-1 in the kidney in sickle cell disease. Am. J. Pathol. 158, 893–903. doi:10.1016/S0002-9440(10)64037-0
Nath Ka, Katusic ZS, 2012. Vasculature and Kidney Complications in Sickle Cell Disease. J. Am. Soc. Nephrol. 23, 781–784. doi:10.1681/ASN.2011101019
Nath Karl A, 2014. Heme oxygenase-1 and acute kidney injury. Curr. Opin. Nephrol. Hypertens. 23, 17–24. doi:10.1097/01.mnh.0000437613.88158.d3
Naumann HN, Diggs LW, Barreras L, Williams BJ, 1971. Plasma hemoglobin and hemoglobin fractions in sickle cell crisis. Am. J. Clin. Pathol. 56, 137–147.
Nebor Danitza, Broquere C??dric, Brudey Karine, Mougenel Danielle, Tarer Vanessa, Connes Philippe, Elion Jacques, Romana Marc, 2010. Alpha-thalassemia is associated with a decreased occurrence and a delayed age-at-onset of albuminuria in sickle cell anemia patients. Blood Cells, Mol. Dis. 45, 154–158. doi:10.1016/j.bcmd.2010.06.003
Nelson GW, Freedman BI, Bowden DW, Langefeld CD, An P, Hicks PJ, Bostrom MA, Johnson RC, Kopp JB, Winkler CA, 2010. Dense mapping of MYH9 localizes the strongest kidney disease associations to the region of introns 13 to 15. Hum. Mol. Genet. 19, 1805–1815.
Nielsen Marianne Jensby, Petersen Steen Vang, Jacobsen Christian, Oxvig Claus, Rees David, M??ller Holger Jon, Moestrup S??ren Kragh, 2006. Haptoglobin-related protein is a high-affinity hemoglobin-binding plasma protein. Blood 108, 2846–2849. doi:10.1182/blood-2006-05-022327
Norris KC, 2006. Baseline Predictors of Renal Disease Progression in the African American Study of Hypertension and Kidney Disease. J. Am. Soc. Nephrol. 17, 2928–2936. doi:10.1681/ASN.2005101101
Novelli Enrico M, Hildesheim Mariana, Rosano Caterina, Vanderpool Rebecca, Simon Marc, Kato Gregory J, Gladwin Mark T, 2014. Elevated pulse pressure is associated with hemolysis, proteinuria and chronic kidney disease in sickle cell disease. PLoS One 9, 1–14. doi:10.1371/journal.pone.0114309
Ohene-Frempong K, Weiner SJ, Sleeper La, Miller ST, Embury S, Moohr JW, Wethers DL, Pegelow CH, Gill FM, 1998. Cerebrovascular accidents in sickle cell disease: rates and risk factors. Blood 91, 288–294.
Oli JM, Watkins PJ, Wild B, Adegoke OJ, 2004. Albuminuria in Afro-Caribbeans with Type 2 diabetes mellitus: is the sickle cell trait a risk factor? Diabet. Med. 21, 483–486.
Ono Koh, Goto Yoichi, Takagi Shuichi, Baba Shunroku, Tago Naomi, Nonogi Hiroshi, Iwai Naoharu, 2004. A promoter variant of the heme oxygenase-1 gene may reduce the incidence of ischemic heart disease in Japanese. Atherosclerosis 173, 315–319. doi:10.1016/j.atherosclerosis.2003.11.021
Osei-Yeboah CT, Rodrigues O, 2011. Renal Status of Children with Sickle Cell Disease in

Accra, Ghana. Ghana Med. J. 45, 155–160.
Otterbein Leo E, Mantell Lin L, Choi Augustine MK, 1999. Carbon monoxide provides protection against hyperoxic lung injury. Am. J. Physiol. - Lung Cell. Mol. Physiol. 276, 688–694.
Pakasa NM, Sumaili EK, 2012. Pathological peculiarities of chronic kidney disease in patient from sub-Saharan Africa. Review of data from the Democratic Republic of the Congo. Ann. Pathol. 32, 40–52.
Palmer Nicholette D, Ng Maggie CY, Hicks Pamela J, Mudgal Poorva, Langefeld Carl D, Freedman Barry I, Bowden Donald W, 2014. Evaluation of candidate nephropathy susceptibility genes in a genome-wide association study of African American diabetic kidney disease. PLoS One 9, 1–6. doi:10.1371/journal.pone.0088273
Parsa Afshin, Kao WH Linda, Xie Dawei, Astor Brad C, Li Man, Hsu Chi-yuan, Feldman Harold I, Parekh Rulan S, Kusek John W, Greene Tom H, Fink Jeffrey C, Anderson Amanda H, Choi Michael J, Wright Jackson T, Lash James P, Freedman Barry I, Ojo Akinlolu, Winkler Cheryl A, Raj Dominic S, Kopp Jeffrey B, He Jiang, Jensvold Nancy G, Tao Kaixiang, Lipkowitz Michael S, Appel Lawrence J, 2013. *APOL1* Risk Variants, Race, and Progression of Chronic Kidney Disease. N. Engl. J. Med. 369, 2183–2196. doi:10.1056/NEJMoa1310345
Pham Phuong-Thu T, Pham Phuong-Chi T, Wilkinson Alan H, Lew Susie Q, 2000. Renal abnormalities in sickle cell disease. Kidney Int. 57, 1–8.
Piel Frédéric B, Patil Anand P, Howes Rosalind E, Nyangiri Oscar A, Gething Peter W, Dewi Mewahyu, Temperley William H, Williams Thomas N, Weatherall David J, Hay Simon I, 2013. Global epidemiology of sickle haemoglobin in neonates: a contemporary geostatistical model-based map and population estimates. Lancet 381, 142–151.
Platt Oprah S, Thorington Bruce D, Brambilla Donald J, Milner Paul F, Rosse Wendell F, Vichinsky Elliott, Kinney TR, 1991. Pain in Sickle Cell Disease Rates and Risk Factors. N. Engl. J. Med. 325, 11–16.
Platt Orah S, Brambilla Donald J, Rosse Wendell F, Milner Paul F, Castro Oswaldo, Steinberg Martin H, Klug Panpit P, 1994. Mortality in Sickle Cell Disease. N. Engl. J. Med. 330, 1639–16344.
Platt OS, Orkin SH, Dover G, Beardsley GP, Miller B, Nathan DG, 1984. Hydroxyurea enhances fetal hemoglobin production in sickle cell anemia. J. Clin. Invest. 74, 652–656.
Powars Darleen R, Elliott-Mills Donna D, Chan Linda, Niland Joyce, Hiti Alan L, Opas Lawrence M, Johnson Cage, 1991. Chronic renal failure in sickle cell disease: risk factors, clinical course, and mortality. Ann. Intern. Med. 115, 614–620.
Powars DR, Chan L, Schroeder WA, 1990. Beta S-gene-cluster haplotypes in sickle cell anemia: clinical implications. Am. J. Pediatr. Hematol. Oncol. 12, 367–374.
Pule Gift Dineo, Bitoungui Valentina Josiane Ngo, Chemegni Bernard Chetcha, Kengne Andre Pascal, Antonarakis Stylianos, Wonkam Ambroise, 2015. Association between Variants at BCL11A Erythroid-Specific Enhancer and Fetal Hemoglobin Levels among Sickle Cell Disease Patients in Cameroon: Implications for Future Therapeutic Interventions. Omi. a J. Integr. Biol.
Pule Gift, Wonkam Ambroise, 2014. Treatment for sickle cell disease in Africa: should we invest in haematopoietic stem cell transplantation? Pan Afr. Med. J. 18.
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.

doi:10.1086/519795
Quimby Kim R, Moe Stephen, Sealy Ian, Nicholls Christopher, Hambleton Ian R, Landis R Clive, 2014. Clinical findings associated with homozygous sickle cell disease in the Barbadian population--do we need a national SCD registry? BMC Res. Notes 7, 102. doi:10.1186/1756-0500-7-102
Ranque Brigitte, Menet Aymeric, Diop Ibrahima Bara, Thiam Marie Michèle, Diallo Dapa, Diop Saliou, Diagne Ibrahima, Sanogo Ibrahima, Kingue Samuel, Chelo David, Wamba Guillaume, Diarra Mamadou, Anzouan Jean Baptiste, N'Guetta Roland, Diakite Cheick Oumar, Traore Youssouf, Legueun Gaëlle, Deme-Ly Indou, Belinga Suzanne, Boidy Kouakou, Kamara Ismael, Tharaux Pierre Louis, Jouven Xavier, 2014. Early renal damage in patients with sickle cell disease in sub-Saharan Africa: A multinational, prospective, cross-sectional study. Lancet Haematol. 1, e64–e73. doi:10.1016/S2352-3026(14)00007-6
Rees David C, Williams Thomas N, Gladwin Mark T, 2010. Sickle-cell disease. Lancet 376, 2018–2031.
Rees William A, Yager Thomas D, Korte John, Hippel Peter H Von, 1993. Betaine can eliminate the base pair composition dependence of DNA melting. Biochemistry 32, 137–144.
Rosenberg Avi Z, Naicker Saraladevi, Winkler Cheryl A, Kopp Jeffrey B, 2015. HIV-associated nephropathies: epidemiology, pathology, mechanisms and treatment. Nat. Rev. Nephrol. 11, 150–160.
Rumaney Maryam Bibi, Ngo Bitoungui Valentina Josiane, Vorster Anna Alvera, Ramesar Raj, Kengne Andre Pascal, Ngogang Jeanne, Wonkam Ambroise, 2014. The co-inheritance of alpha-thalassemia and sickle cell anemia is associated with better hematological indices and lower consultations rate in Cameroonian patients and could improve their survival. PLoS One 9, 1–10. doi:10.1371/journal.pone.0100516
Saraf Santosh L, Zhang Xu, Kanias Tamir, Lash James P, Molokie Robert E, Oza Bharvi, Lai Catherine, Rowe Julie H, Gowhari Michel, Hassan Johara, Desimone Joseph, Machado Roberto F, Gladwin Mark T, Little Jane A, Gordeuk Victor R, 2014. Haemoglobinuria is associated with chronic kidney disease and its progression in patients with sickle cell anaemia. Br. J. Haematol. 164, 729–739. doi:10.1111/bjh.12690
Saraf SL, Zhang X, Shah B, Kanias T, Gudehithlu KP, Kittles R, Machado RF, Arruda JAL, Gladwin MT, Singh AK, Gordeuk VR, 2015. Genetic variants and cell-free hemoglobin processing in sickle cell nephropathy. Haematologica 100, 1275–1284. doi:10.3324/haematol.2015.124875
Sasongko Teguh H, Nagalla Srikanth, Ballas Samir K, 2013. Angiotensin-converting enzyme (ACE) inhibitors for proteinuria and microalbuminuria in people with sickle cell disease. doi:10.1002/14651858.CD009191.pub3.www.cochranelibrary.com
Schechter Alan N, 2008. Hemoglobin research and the origins of molecular medicine. Blood 112, 3927–3938. doi:10.1182/blood-2008-04-078188
Scheinman Jon I, 2009. Sickle cell disease and the kidney. Nat. Clin. Pract. Nephrol. 5, 78–88.
Schwartz George J, Munoz Alvaro, Scheider Michael F, Mak Robert H, Kaskel Frederick, Warady Bradley A, Furth Susan L, 2009. New Equations to Estimate GFR in Children with CKD. J Am Soc Nephrol. 20, 629–637. doi:10.1681/ASN.2008030287
Sebastiani Paola, Ramoni Marco F, Nolan Vikki, Baldwin Clinton T, Steinberg Martin H, 2005. Genetic dissection and prognostic modeling of overt stroke in sickle cell anemia. Nat. Genet. 37, 435–440.

Selistre Luciano, Rabilloud Muriel, Cochat Pierre, de Souza Vandr??a, Iwaz Jean, Lemoine Sandrine, Beyerle Fran??oise, Poli-de-Figueiredo Carlos E, Dubourg Laurence, 2016. Comparison of the Schwartz and CKD-EPI Equations for Estimating Glomerular Filtration Rate in Children, Adolescents, and Adults: A Retrospective Cross-Sectional Study. PLoS Med. 13, 1–18. doi:10.1371/journal.pmed.1001979

Stehlé Thomas, Bartolucci Pablo, Bouanane Mohamed, Galacteros Frederic, Dudreuilh Caroline, Grimbert Philippe, Deux Jean-François, Audard Vincent, 2016. Reversible kidney iron accumulation in a patient with sickle cell disease treated with hydroxyurea. Am. J. Hematol. 1–29. doi:10.1002/ajh.24544

Steinberg Martin H, 2005. Predicting clinical severity in sickle cell anaemia. Br. J. Haematol. 129, 465–481. doi:10.1111/j.1365-2141.2005.05411.x

Steinberg MH, Barton F, Castro O, Pegelow CH, Ballas SK, Kutlar A, Orringer E, Bellevue R, Olivieri N, Eckman J, Varma M, Ramirez G, Adler B, Smith W, Carlos T, Ataga K, DeCastro L, Bigelow C, Saunthararajah Y, Telfer M, Vichinsky E, Claster S, Shurin S, Bridges K, Waclawiw M, Bonds D, Terrin M, 2003. Effect of hydroxyurea on mortality and morbidity in adult sickle cell anemia: risks and benefits up to 9 years of treatment. Jama 289, 1645–1651.

Steinberg MH, Lu ZH, Nagel RL, Venkataramani S, Milner PF, Huey L, Safaya S, Rieder RF, 1998. Hematological effects of atypical and Cameroon beta-globin gene haplotypes in adult sickle cell anemia. Am. J. Hematol. 59, 121–126.

Stockman JA, Nigro MA, Mishkin MM, Oski FA, 1972. Occlusion of large cerebral vessels in sickle-cell anemia. N. Engl. J. Med. 287, 846–849.

Stuart Marie J, Nagel Ronald L, Jefferson Thomas, 2004. Sickle-cell disease. Lancet 364, 1343–1360. doi:10.1016/S0140-6736(04)17192-4

Sundaram Nambirajan, Bennett Michael, Wilhelm Jamie, Kim Mi Ok, Atweh George, Devarajan Prasad, Malik Punam, 2011. Biomarkers for early detection of sickle nephropathy. Am. J. Hematol. 86, 559–566. doi:10.1002/ajh.22045

Taha Hevidar, Skrzypek Klaudia, Guevara Ibeth, Nigisch Anneliese, Mustafa Stefan, Grochot-Przeczek Anna, Ferdek Pawel, Was Halina, Kotlinowski Jerzy, Kozakowska Magdalena, Balcerczyk Aneta, Muchova Lucie, Vitek Libor, Weigel Guenter, Dulak Jozef, Jozkowicz Alicja, 2010. Role of heme oxygenase-1 in human endothelial cells: Lesson from the promoter allelic variants. Arterioscler. Thromb. Vasc. Biol. 30, 1634–1641. doi:10.1161/ATVBAHA.110.207316

Taylor Steve M, Parobek Christian M, Fairhurst Rick M, 2012. Haemoglobinopathies and the clinical epidemiology of malaria: a systematic review and meta-analysis. Lancet Infect. Dis. 12, 457–468.

Tayo Bamidele O, Kramer Holly, Salako Babatunde L, Gottesman Omri, McKenzie Colin A, Ogunniyi Adesola, Bottinger Erwin P, Cooper Richard S, 2013. Genetic variation in APOL1 and MYH9 genes is associated with chronic kidney disease among Nigerians. Int. Urol. Nephrol. 45, 485–494. doi:10.1007/s11255-012-0263-4

Tenhunen R, Marver HS, Schmid R, 1969. Microsomal Heme Oxygenase: Characterization of the enzyme. J. Biol. Chem. 244, 6388–6394.

Tenhunen R, Marver HS, Schmid R, 1968. The enzymatic conversion of heme to bilirubin by microsomal heme oxygenase. Proc. Natl. Acad. Sci. U. S. A. 61, 748–755. doi:10.1073/pnas.61.2.748

Thein Swee Lay, Menzel Stephan, 2009. Discovering the genetics underlying foetal haemoglobin production in adults. Br. J. Haematol. 145, 455–467.

Thompson J, Reid M, Hambleton I, GR Serjeant, 2007. Albuminuria and renal function in homozygous sickle cell disease: Observations from a cohort study. Arch. Intern. Med. 167, 701–708. doi:10.1001/archinte.167.7.701

Tracz MJ, Alam J, Nath KA, 2007. Physiology and Pathophysiology of Heme: Implications for Kidney Disease. J. Am. Soc. Nephrol. 18, 414–420. doi:10.1681/ASN.2006080894

Tzur Shay, Rosset Saharon, Shemer Revital, Yudkovsky Guennady, Selig Sara, Tarekegn Ayele, Bekele Endashaw, Bradman Neil, Wasser Walter G, Behar Doron M, Skorecki Karl, 2010. Missense mutations in the APOL1 gene are highly associated with end stage kidney disease risk previously attributed to the MYH9 gene. Hum. Genet. 128, 345–350. doi:10.1007/s00439-010-0861-0

Ulasi Ifeoma I, Tzur Shay, Wasser Walter G, Shemer Revital, Kruzel Etty, Feigin Elena, Ijoma Chinwuba K, Onodugo Obinna D, Okoye Julius U, Arodiwe Ejikeme B, others, 2013. High population frequencies of APOL1 risk variants are associated with increased prevalence of non-diabetic chronic kidney disease in the Igbo people from south-eastern Nigeria. Nephron Clin. Pract. 123, 123–128.

Van Xong Hoang, Vanhamme Luc, Chamekh Mustapha, Chimfwembe Chibeka Evelyn, Van Den Abbeele Jan, Pays Annette, Van Melrvenne Nestor, Hamers Raymond, De Baetselier Patrick, Pays Etienne, 1998. A VSG expression site-associated gene confers resistance to human serum in Trypanosoma rhodesiense. Cell 95, 839–846. doi:10.1016/S0092-8674(00)81706-7

Vanhamme Luc, Paturiaux-Hanocq Françoise, Poelvoorde Philippe, Nolan Derek P, Lins Laurence, Abbeele Jan Van Den, Pays Annette, Tebabi Patricia, Xong Huang Van, Jacquet Alain, Moguilevsky Nicole, Dieu Marc, Kane John P, Baetselier Patrick De, Brasseur Robert, Pays Etienne, 2003. Apolipoprotein L-1 is the trypanosome lytic factor of human serum. Nature 422, 83–87. doi:10.1038/nature01457.1.

Vazquez Benjamin, Shah Binal, Zhang Xu, Lash James P, Gordeuk Victor R, Saraf Santosh L, 2014. Hyperfiltration is associated with the development of microalbuminuria in patients with sickle cell anemia. Am. J. Hematol. 89, 1156–1157. doi:10.1002/ajh.23817

Vichinsky Elliott P, Earles Ann, Johnson Robert A, Hoag M Silvija, Williams Amber, Lubin Betram, 1990. Alloimmunization in Sickle Cell Anemia and Transfusion of Racially Unmatched Blood. N. Engl. J. Med. 322, 1617–1621.

Ware Russell E, Rees Renee C, Sarnaik Sharada A, Iyer Rathi V., Alvarez Ofelia A, Casella James F, Shulkin Barry L, Shalaby-Rana Eglal, Strife C Frederic, Miller John H, Lane Peter A, Wang Winfred C, Miller Scott T, 2010. Renal Function in Infants with Sickle Cell Anemia: Baseline Data from the BABY HUG Trial. J. Pediatr. 156, 66–70.e1. doi:10.1016/j.jpeds.2009.06.060

Ware Russell E, Schultz William H, Yovetich Nancy, Mortier Nicole A, Alvarez Ofelia, Hilliard Lee, Iyer Rathi V, Miller Scott T, Rogers Zora R, Scott J Paul, 2011. Stroke With Transfusions Changing to Hydroxyurea (SWiTCH): a phase III randomized clinical trial for treatment of children with sickle cell anemia, stroke, and iron overload. Pediatr. Blood Cancer 57, 1011–1017.

Weatherall DJ, Clegg JB, 2001. Inherited haemoglobin disorders: An increasing global health problem. Bull. World Health Organ. 79, 704–712. doi:10.1590/S0042-96862001000800005

Wigfall Delbert R, Ware Russell E, Burchinal Margaret R, Kinney Thomas R, Foreman John W, 2000. Prevalence and Clinical Correlates of Glomerulopathy Children With Sickle Cell Disease. J. Pediatr. 136, 749–753. doi:10.1067/rnpd.2000.105998

Witte Elsbeth C, Heerspink Hiddo J Lambers, de Zeeuw Dick, Bakker Stephan JL, de Jong Paul E, Gansevoort Ronald, 2009. First morning voids are more reliable than spot urine samples to assess microalbuminuria. J. Am. Soc. Nephrol. 20, 436–443.

Wonkam Ambroise, Ngo Bitoungui Valentina J, Vorster Anna a, Ramesar Raj, Cooper Richard S, Tayo Bamidele, Lettre Guillaume, Ngogang Jeanne, 2014a. Association of variants at BCL11A and HBS1L-MYB with hemoglobin F and hospitalization rates among sickle cell patients in Cameroon. PLoS One 9. doi:10.1371/journal.pone.0092506

Wonkam Ambroise, 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

Xie Dawei, Joffe Marshall M, Brunelli Steven M, Beck Gerald, Chertow Glenn M, Fink Jeffrey C, Greene Tom, Hsu Chi-yuan, Kusek John W, Landis Richard, others, 2008. A comparison of change in measured and estimated glomerular filtration rate in patients with nondiabetic kidney disease. Clin. J. Am. Soc. Nephrol. 3, 1332–1338.

Xu Jian, Peng Cong, Sankaran Vijay G, Shao Zhen, Esrick Erica B, Chong Bryan G, Ippolito Gregory C, Fujiwara Yuko, Ebert Benjamin L, Tucker Philip W, Orkin Stuart H, 2011. Correction of Sickle Cell Disease in Adult Mice by Interference with Fetal Hemoglobin Silencing. Science (80-.). 334, 993–996. doi:10.1126/science.1211053

Xu Jian, Sankaran Vijay G, Ni Min, Menne Tobias F, Puram Rishi V., Kim Woojin, Orkin Stuart H, 2010. Transcriptional silencing of γ-globin by BCL11A involves long-range interactions and cooperation with SOX6. Genes Dev. 24, 783–789. doi:10.1101/gad.1897310

Yetunde A, Anyaegbu CC, 2001. Profile of the Nigerian sickle cell anaemia patients above 30 years of age. Cent. Afr. J. Med. 47, 108–111.

Yoshida T, Kikuchi G, 1979. Purification and properties of heme oxygenase from rat liver microsomes. J Biol Chem 254, 4487–4491.

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)

C_2H_6SO

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)

Lot no.: 0000160698

Agarose Gel Electrophoresis

This was used for the resolution of DNA fragments as the negatively charged DNA phosphate (PO_4^-) 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)

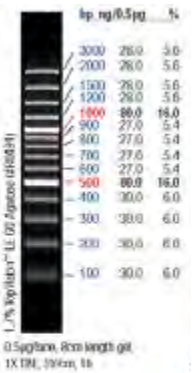
Made up to 100 mL and distilled dH_2O made the total volume up to 1 L.

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 (*Exo*I)

Lot no.: 00254712

20U/μL

Cycle sequencing

70% Ethanol (EtOH)

350 mL absolute ethanol (100%) (Kimix) mixed with 150 mL of autoclaved H_2O.

Sodium Acetate (3 M)

246.1 g sodium acetate in 1 L autoclaved dH_2O

BigDye® Terminator (v3.1)

Austin TX, USA

Cycle Sequencing Kit

Lot no: 1106093

SNaPSHOT® Sequencing

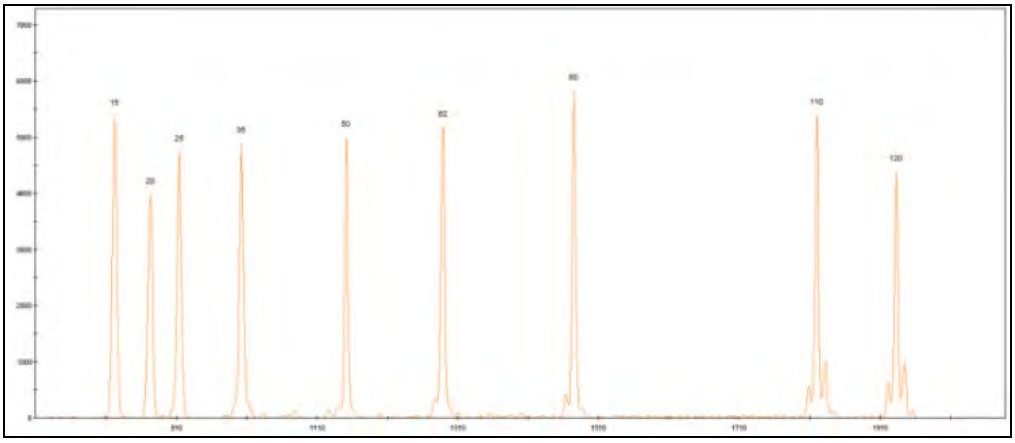
SNaPshot® Sequencing Mix

Thermo Fisher Scientific Inc.

GeneScan™ 120 Liz™ Size standard

Applied Biosystems (California, USA)

Lot no.: 1506053



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

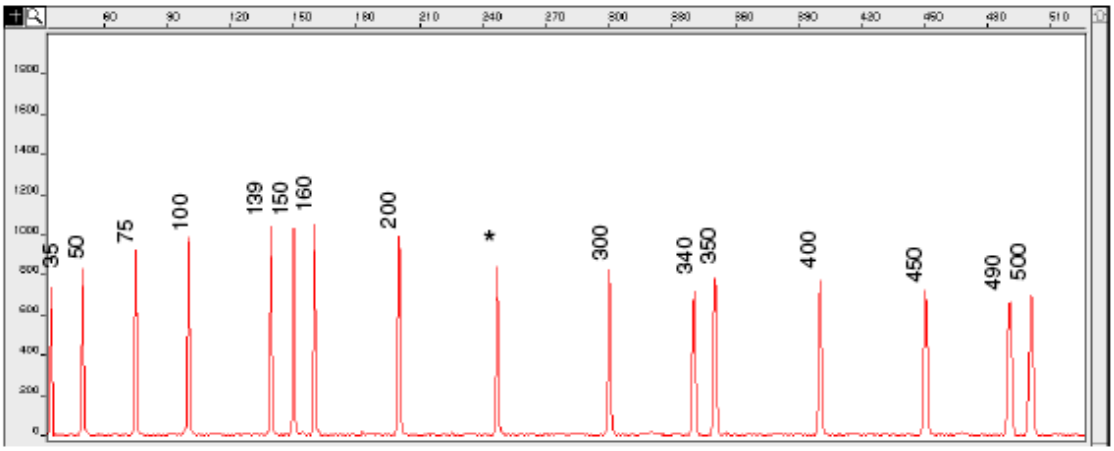
KAPA2G Fast HotStart ReadyMix

KAPA Biosystems (Wilmington, USA)

GeneScan™ 500 ROX™ Size standard

Applied Biosystems (Warrington, UK)

Lot no.: 1511345



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^{TM} Real-Time system (Bio-Rad Laboratories, Hercules, CA, USA)

Bio-Rad CFX96^{TM} 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.

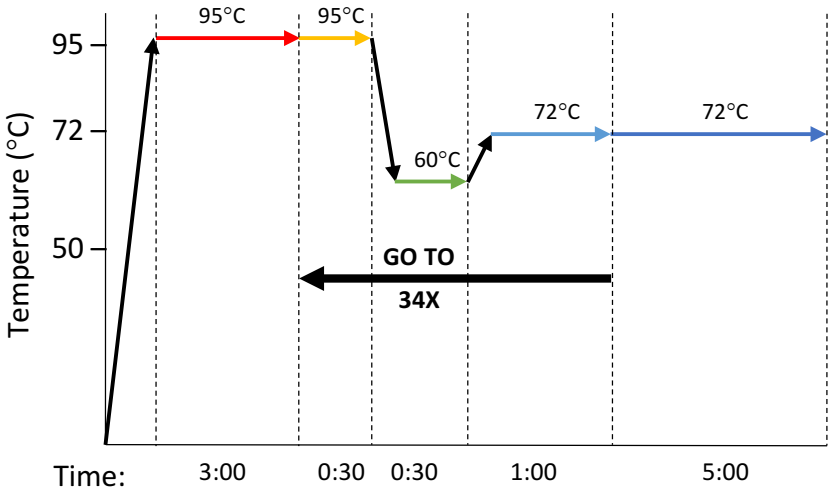
Standard PCR reaction:

REAGENTS	**STOCK CONCENTRATION**	**REQUIRED FINAL CONCENTRATION**	**VOLUME (μL)**
Sabax distilled H_2O			16.9
GoTaq® Buffer	5X	1X	5
dNTPs	5 mM	200 μM	1
Forward Primer	20 μM	10 pM	0.5
Reverse Primer	20 μM	10 pM	0.5
GoTaq® DNA Polymerase	5 U/μL	0.02 U/μL	0.1
Template DNA	100ng/μL	100 ng	1
		Total:	25

PCR CONSTITUENTS AND CYCLING PROCEDURES FOR THE AMPLIFICATION OF RS16996648

REAGENTS	STOCK CONCENTRATION	REQUIRED FINAL CONCENTRATION	VOLUME (µL)
Sabax distilled H_2O			16.9
GoTaq® Buffer	5X	1X	5
dNTPs	5 mM	200 µM	1
Forward Primer	20 µM	10 pM	0.5
Reverse Primer	20 µM	10 pM	0.5
GoTaq® DNA Polymerase	5 U/µL	0.02 U/µL	0.1
Template DNA	100ng/µL	100 ng	1
		Total:	25

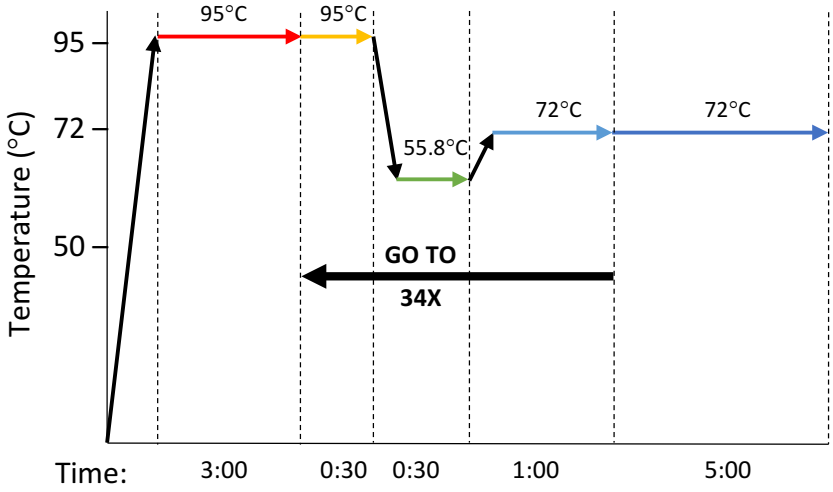
CYCLING PARAMETERS:



PCR CONSTITUENTS AND CYCLING PROCEDURES FOR THE AMPLIFICATION OF RS5750248

REAGENTS	STOCK CONCENTRATION	REQUIRED FINAL CONCENTRATION	VOLUME (µL)
Sabax distilled H_2O			14.9
GoTaq® Buffer	5X	1X	5
dNTPs	5 mM	200 µM	1
DMSO	100%	8%	2
Forward Primer	20 µM	10 pM	0.5
Reverse Primer	20 µM	10 pM	0.5
GoTaq® DNA Polymerase	5 U/µL	0.02 U/µL	0.1
Template DNA	100ng/µL	100 ng	1
		Total:	25

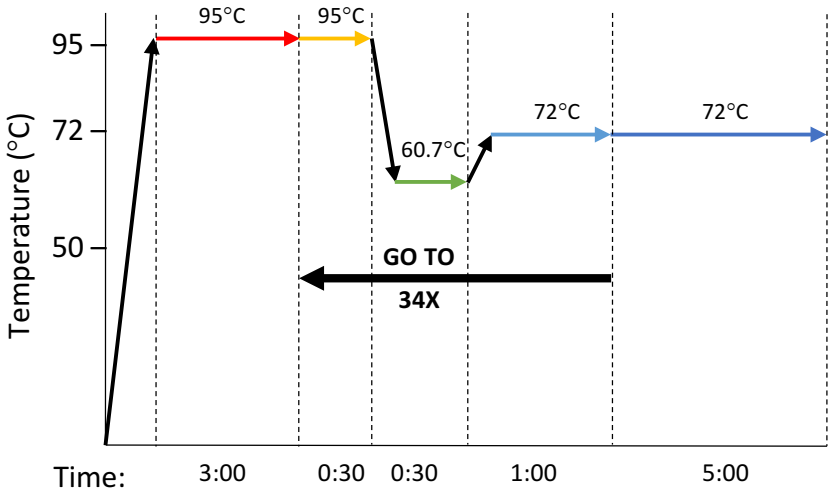
CYCLING PARAMETERS:



PCR CONSTITUENTS AND CYCLING PROCEDURES FOR THE AMPLIFICATION OF MULTIPLEX A (RS1005570, RS16996672, RS73885319)

REAGENTS	STOCK CONCENTRATION	REQUIRED FINAL CONCENTRATION	VOLUME (µL)
Sabax distilled H_2O			14.4
GoTaq® Buffer	5X	1X	5
dNTPs	5 mM	200 µM	1.5
Forward Primer (rs1005570)	20 µM	10 pM	0.5
Reverse Primer (rs1005570)	20 µM	10 pM	0.5
Forward Primer (rs16996672)	20 µM	10 pM	0.5
Reverse Primer (rs16996672)	20 µM	10 pM	0.5
Forward Primer (rs73885319)	20 µM	10 pM	0.5
Reverse Primer (rs73885319)	20 µM	10 pM	0.5
GoTaq® DNA Polymerase	5 U/µL	0.02 U/µL	0.1
Template DNA	100ng/µL	100 ng	1
		Total:	25

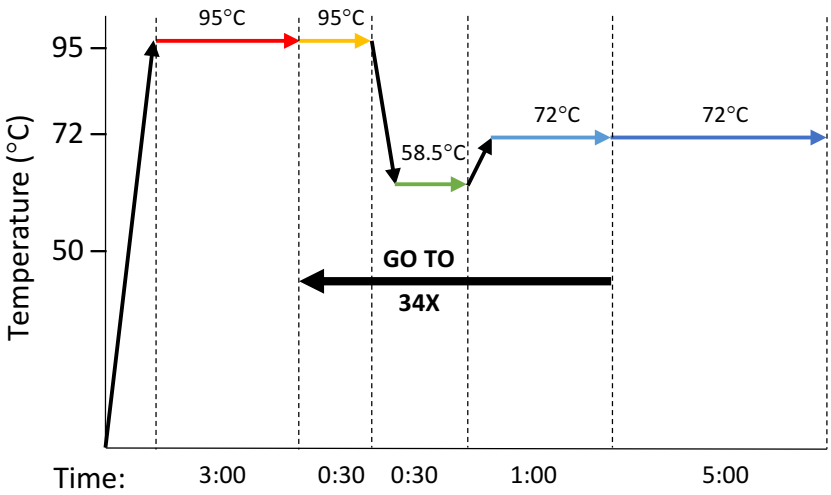
CYCLING PARAMETERS:



PCR CONSTITUENTS AND CYCLING PROCEDURES FOR THE AMPLIFICATION OF MULTIPLEX B (RS1557529, RS11912763)

REAGENTS	STOCK CONCENTRATION	REQUIRED FINAL CONCENTRATION	VOLUME (µL)
Sabax distilled H_2O			15.4
GoTaq® Buffer	5X	1X	5
dNTPs	5 mM	200 µM	1.5
Forward Primer (rs1557529)	20 µM	10 pM	0.5
Reverse Primer (rs1557529)	20 µM	10 pM	0.5
Forward Primer (rs11912763)	20 µM	10 pM	0.5
Reverse Primer (rs11912763)	20 µM	10 pM	0.5
GoTaq® DNA Polymerase	5 U/µL	0.02 U/µL	0.1
Template DNA	100ng/µL	100 ng	1
		Total:	25

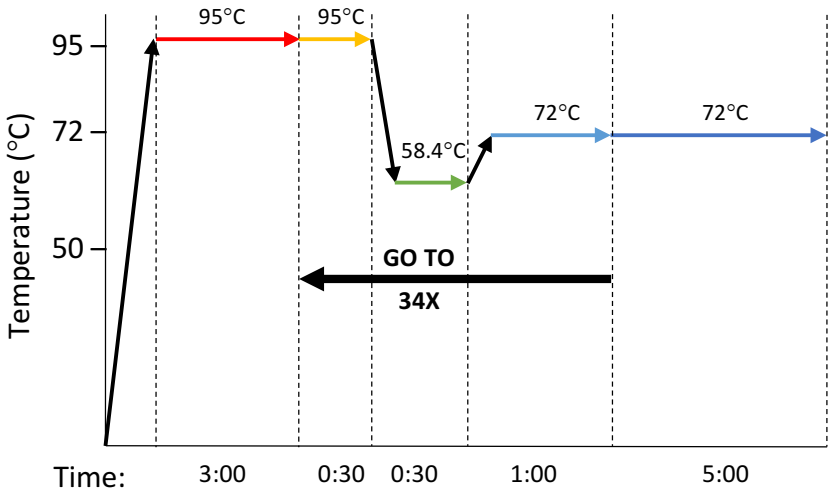
CYCLING PARAMETERS:



PCR CONSTITUENTS AND CYCLING PROCEDURES FOR THE AMPLIFICATION OF MULTIPLEX C (RS5756129, RS8141189)

REAGENTS	STOCK CONCENTRATION	REQUIRED FINAL CONCENTRATION	VOLUME (µL)
Sabax distilled H_2O			14.9
GoTaq® Buffer	5X	1X	5
dNTPs	5 Mm	200 µM	1.5
Betaine	5M	0.1M	0.5
Forward Primer (rs5756129)	20 µM	10 pM	0.5
Reverse Primer (rs5756129)	20 µM	10 pM	0.5
Forward Primer (rs8141189)	20 µM	10 pM	0.5
Reverse Primer (rs8141189)	20 µM	10 pM	0.5
GoTaq® DNA Polymerase	5 U/µL	0.02 U/µL	0.1
Template DNA	100 ng/µL	100 ng	1
		Total:	25

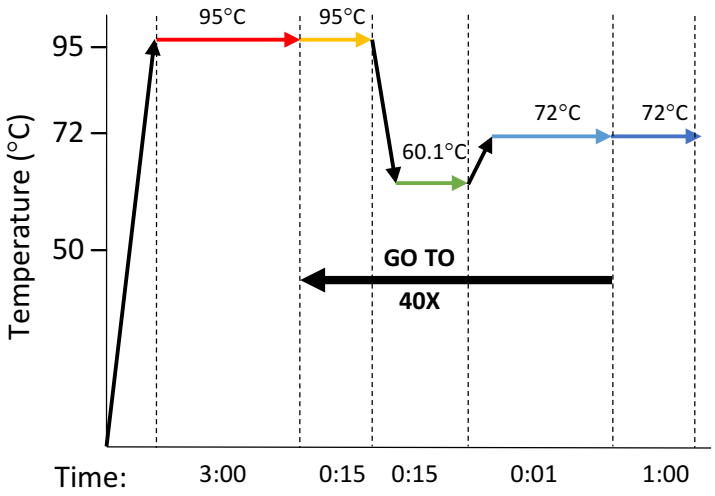
CYCLING PARAMETERS:



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

REAGENTS	STOCK CONCENTRATION	REQUIRED FINAL CONCENTRATION	VOLUME (µL)
Sabax distilled H_2O			8.7
KAPA2G Fast HotStart ReadyMix	2X	1X	10
Forward primer	20 µM	10 pM	0.4
Reverse primer	20 µM	10 pM	2.4
Template DNA	10 ng/µL	5 ng	0.5
		Total:	20

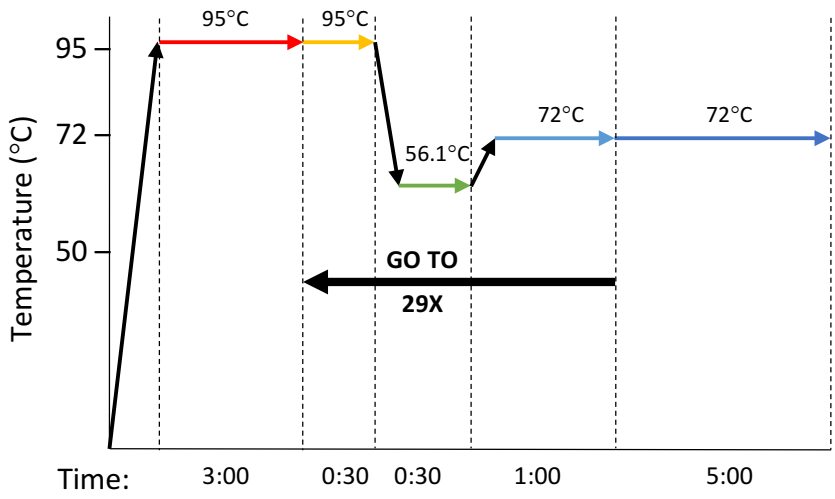
Cycling conditions:



PCR CONSTITUENTS AND CYCLING PROCEDURES FOR THE AMPLIFICATION OF rs73885319 FOR DIRECT CYCLE SEQUENCING

REAGENTS	STOCK CONCENTRATION	REQUIRED FINAL CONCENTRATION	VOLUME (µL)
Sabax distilled H_2O			11.9
GoTaq® Buffer	5X	1X	5
dNTPs	5 mM	200 µM	1
Forward Primer (rs73885319)	20 µM	10 pM	0.5
Reverse Primer (rs73885319)	20 µM	10 pM	0.5
GoTaq® DNA Polymerase	5 U/µL	0.02 U/µL	0.1
Template DNA	10 ng/µL	50 ng	5
		Total:	25

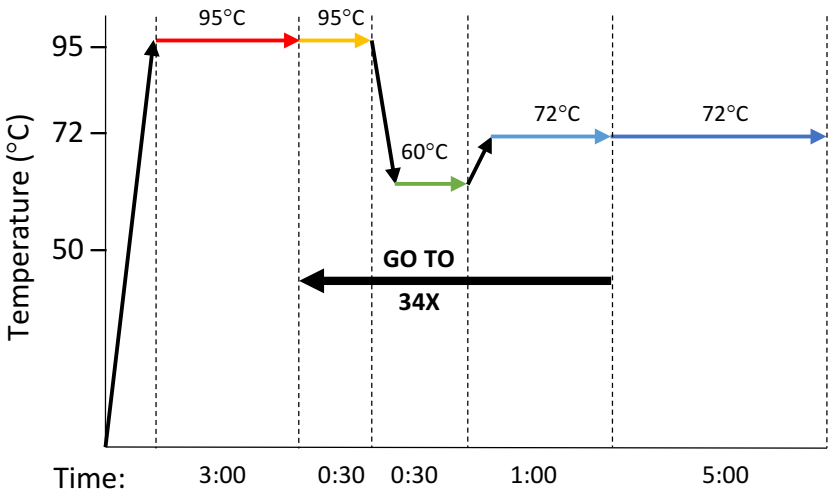
Cycling procedures:



PCR CONSTITUENTS AND CYCLING PROCEDURES FOR THE AMPLIFICATION OF rs60910145 FOR DIRECT CYCLE SEQUENCING

REAGENTS	STOCK CONCENTRATION	REQUIRED FINAL CONCENTRATION	VOLUME (µL)
Sabax distilled H_2O			11.9
GoTaq® Buffer	5X	1X	5
dNTPs	5 mM	200 µM	1
Forward Primer (rs60910145)	20 µM	10 pM	0.5
Reverse Primer (rs60910145)	20 µM	10 pM	0.5
GoTaq® DNA Polymerase	5 U/µL	0.02 U/µL	0.1
Template DNA	10 ng/µL	50 ng	5
		Total:	25

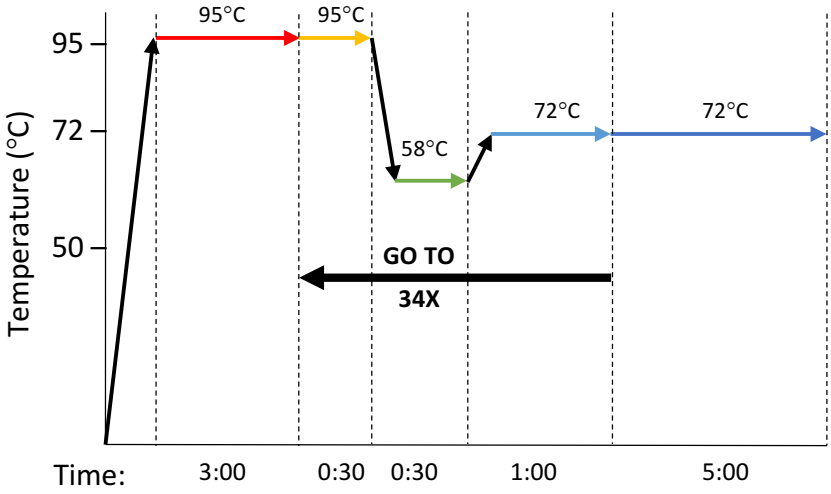
Cycling procedures:



PCR CONSTITUENTS AND CYCLING PROCEDURES FOR THE AMPLIFICATION OF rs743811 FOR DIRECT CYCLE SEQUENCING

REAGENTS	STOCK CONCENTRATION	REQUIRED FINAL CONCENTRATION	VOLUME (µL)
Sabax distilled H_2O			11.9
GoTaq® Buffer	5X	1X	5
dNTPs	5 mM	200 µM	1
Forward Primer (rs743811)	20 µM	10 pM	0.5
Reverse Primer (rs743811)	20 µM	10 pM	0.5
GoTaq® DNA Polymerase	5 U/µL	0.02 U/µL	0.1
Template DNA	10 ng/µL	50 ng	5
		Total:	25

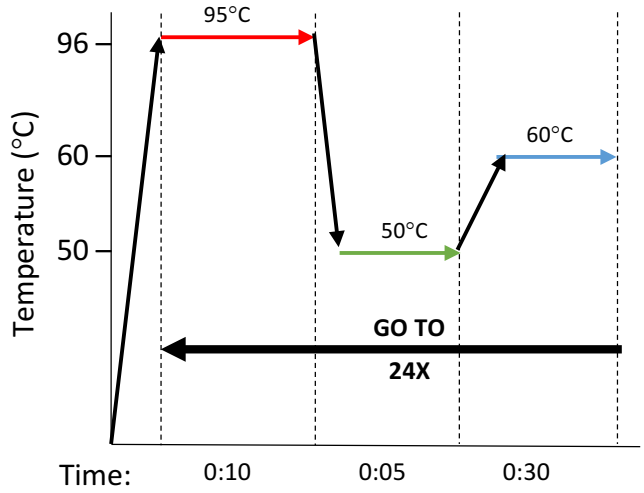
Cycling procedures:



APPENDIX 4: SNAPSHOT® REACTION CONSTITUENTS, PROTOCOL AND CAPILLARY ELECTROPHORESIS

REAGENTS	STOCK CONCENTRATION	REQUIRED FINAL CONCENTRATION	VOLUME (µL)
Sabax distilled H_2O			6
SNaPshot® Primer	20 µM	2 µM	1
SNaPshot® mix			0.5
PCR Product	100ng	10ng	2.5
		Total:	10

CYCLING PARAMETERS:

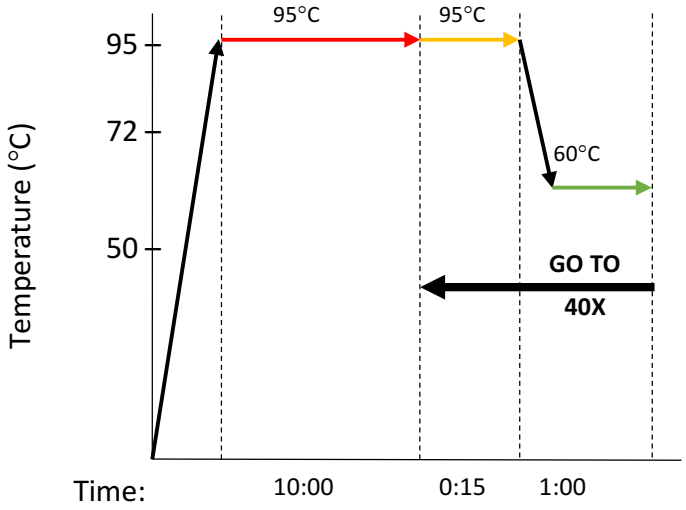


APPENDIX 5: TAQMAN® CONSTITUENTS AND PROTOCOL

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

REAGENTS	STOCK CONCENTRATION	REQUIRED FINAL CONCENTRATION	VOLUME (µL)
TaqMan® Universal PCR MasterMix	2X	1X	5
TaqMan® Assay Working stock	20X	1X	0.5
Sabax distilled H_2O			2.5
Template DNA	10 ng/µL	20 ng	2
		Total:	20

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).