Sickle Cell Trait and targeted genomic variants in Chronic Kidney Disease an African cohort

Name: Tshepiso Masekoameng

Student number: MSKTSH012

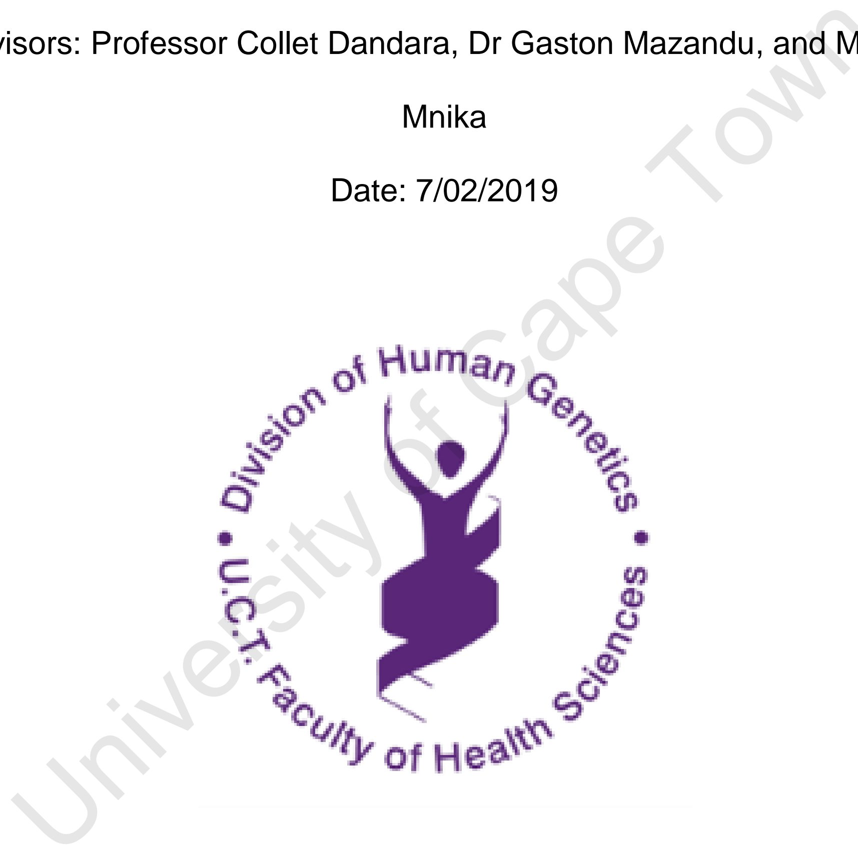
Degree: MSc (Med) Human Genetics

Supervisor: Professor Ambroise Wonkam

Co-Supervisors: Professor Collet Dandara, Dr Gaston Mazandu, and Miss Khuthala Mnika

Date: 7/02/2019

University of Cape Town



Division of Human Genetics, Department of Pathology, Faculty of Health Sciences,

University of Cape Town, Cape Town, South Africa.

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

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

University of Cape Town

Table of Contents

List of Figures 6
List of Tables 7
List of Abbreviations 8
Abstract 10
Background 10
Aims 10
Methods 10
Results 11
Conclusion and perspectives 11
CHAPTER 1: Introduction 13
CHAPTER 2: Literature review 16
2.1 Kidney functions and Kidney disease in Africa 16
2.1.1 Normal Kidney Function 16
2.1.2 Burden and aetiology of Kidney diseases in Africa 18
2.2.1 Familial clustering and Heritability of CKD 19
2.2 Genetics of kidney disease in Africans 19
MYH9 and APOL1 and kidney diseases 19
2.2.2 Sickle cell disease and Sickle cell trait and CKD 20
2.2.2.1 Epidemiology of Sickle Cell Trait 21
2.2.2.2 Genetics of SCD 23
2.2.2.3 Sickle Cell Trait and Chronic Kidney Disease (CKD) 25
2.2.2.5 Other targeted genes variants from Genome-Wide Association Studies (GWAS) SNPs that associated with CKD. 28
RATIONAL 32
Aim 32
Objectives 32
CHAPTER 3: Methods and Materials 33
3.1 PATIENTS' DATA 33
3.1.1 Ethical Approval 33
3.1.2 Participants' descriptions 33
3.2 MOLECULAR METHODS 33
3.2.1 Samples Preparation 33
3.2.2 Primers Design 34
3.2.3 Polymerase Chain Reaction (PCR) 35

3.2.4 Sickle Cell Trait Genotyping 37
3.2.5. MassArray and TAQMAN® genotyping Assays for the other targeted SNPs 37
3.2.6 Cycle Sequencing for validation of genotyping 39
3.3 Statistical Analysis 42
CHAPTER 4: RESULTS 43
4.1. Population Description and clinical phenotypes 43
4.1.1 Socio-demographic and anthropometric variables and CKD 43
4.2 Validity of the statistical model used in the analysis of the cohort 46
4.3 Description of CKD and other clinical co-morbidity 46
4.4 Description of the SCT within the CKD cases and non-CKD controls 47
4.5 Association between SCT and CKD 48
4.6 Description of SCT and other clinical co-morbidity 50
4.7 Frequencies of CKD-related genes variants across various populations 51
4.8 Association of CKD and targeted genomic variants 52
4.9 Functional Pathway Analysis 55
4.10 Influence of SCT on the APOL1 and CKD relationship. 57
CHAPTER 5: DISCUSSION 58
5.1 Originality 58
5.2 CKD and known Clinical Comorbidities 58
5.3 Association of the SCT within the CKD cases and non-CKD controls 61
5.4 Beyond association with CKD, SCT is not that benign 62
5.5 Description of possible SCT and other clinical phenotypes 63
5.6 Association of CKD and targeted genomic variants 64
5.7 Functional Pathway Analysis 65
5.8 Influence of SCT on the APOL1 and CKD relationship 66
5.9 Practical implications and research recommendation 66
5.10 Limitation 67
5.11 Conclusion and perspectives 68
REFERENCES 69
Appendices 83
Appendix 1 Reagents Used 83
PCR Reagents 83
Appendix 2: Machines Used 86
Thermocyclers 86

Centrifuge 86
Vortex 86
Appendix 3: Gel Electrophoresis Reagents 87
Appendix 4: PCR Amplification Reagents 87
Appendix 5: Sequencing Conditions and Protocol 88

Plagiarism Declaration

1. I, Tshepiso Masekoameng know that plagiarism is wrong. Plagiarism is to use another's work and to pretend that it is one's own
2. I have used the Harvard referencing style for citation and referencing. Each contribution to, and quotation in, this project from the work(s) of other people has been attributed and has been cited and referenced.
3. This project is my own work.
4. I have not allowed, and will not allow, anyone, to copy my work with the intention of passing it off as his or her own work.

Signature Signed by candidate

Tshepiso Masekoameng

Date: 11/06/2019

List of Figures

Figure 1: Diagram of the Kidney…………………………………………………….…….13

Figure 2: Graphical representation of Global burden of SCD………………………..16

Figure 3: Structure of the haemoglobin molecule……………………………………18

Figure 4: PCR Amplification of the DNA fragment containing the SCD mutation….30

Figure 5: RFLP imaging of the SCD mutation using the DdeI enzyme………… ….31

Figure 6: TaqMan SNP Genotyping plot for the cohort for rs881858………………..33

Figure 7: Sanger sequencing validation for rs881858 in patient 7, 10 and 15……..35

Figure 8: Histogram distribution of BMI and age in CKD cases and non-CKD control……………………………………………………………………………...38

Figure 9: Histogram distribution of Systolic Blood pressure (mmHg) and Diastolic Blood pressure (mmHg) in CKD cases and non-CKD controls……… ………….....39

Figure 10: Graphical representation of the validity of the model used in the analysis of the cohort………………………………………………………………………….40

Figure 11: Bar graph representation of additional clinal phenotypes that are positively associated in influencing the development of CKD…………………………………41

Figure 12: Bar graph representation of positively associated factors influencing the development of CKD………………………………………………………………43

Figure 13: Bar graph representation of clinical phenotypes influenced by the presence of SCT……………………….……………………………………………44

Figure 14: Association between targeted variants and CKD…………….……….48

Figure 15: Illustration of the NFKB1 based sub-network…………………………………..56

List of Tables

Table 1: Different stages of chronic kidney disease…………………………..11

Table 2: Genotypes and phenotypes of sickling disorders……………………19

Table 3: List of SNPs investigated in the patients and controls cohorts……24-25

Table 4: Tabulation for the primers used for SCD Diagnosis and validation of the genotyped SNPs…………………………………………………………………29

Table 5: Social-demographic variable and anthropometric variables of the CKD cases and non-CKD controls……………………………………………………..37

Table 6: Association of CKD and other clinical comorbidities …………………..41

Table 7: Association of SCT within CKD cases and non-CKD controls…..…42

Table 8: Factors influencing CKD within the cohort……………………………43

Table 9: Association of SCT and other clinical co-Morbidity……………..……44

Table 10: Allele frequencies of selected CKD-related genes variants among Cameroonian patients with CKD and data extracted from the 1000 Genome project……………………………………………………………………………45

Table 11: Tabulation of the 29 targeted SNPs investigated within the cohort……………………………………………………………………………..46

Table 12: Tabulation of the network clusters……………………………………50

Table 13: Association of *APOL1* and CKD………………………………………52

List of Abbreviations

^{o}C– degree Celsius

μL - Microliter

μM– micromolar

3' – three prime

5'– five prime

BLAST – Basic Local Alignment Search Tool

BMI – body mass index

Bp – base pairs

CKD – Chronic kidney disease

DBP – diastolic blood pressure

DNA –deoxyribonucleic acid

ddNTP – dideoxynucleotide triphosphates

dNTPs – deoxynucleotide triphosphates

*Exo*I – Exonuclease I

FastAP– Thermosensitive Alkaline Phosphatase

g – grams

GFR – glomerular filtration rate

Glu – Glutamic Acid

GWAS – Genome-wide association study

Hb –Haemoglobin

HbA – adult haemoglobin

HbAS – heterozygous for the HbS allele

HbF– foetal haemoglobin

HbS – sickle haemoglobin

HbSS – homozygous for the sickle haemoglobin allele

HWE – Hardy Weinberg Equilibrium

IDT – Integrated DNA Technologies

IQR – interquartile range

M – Molar

MAF – Minor Allele Frequency

Mg2 – Magnesium ions

$MgCl_2$ – Magnesium Chloride

mL – milliliter

mV – millivolts

N – number

NCBI – National Centre for Biotechnology

P – probability value

PCR – Polymerase Chain Reaction

SBP – systolic blood pressure

SCD – Sickle Cell Disease

SCT– Sickle Cell Trait

SNP – Single Nucleotide Polymorphism

Ta – annealing temperature

TBE – Tris-borate EDTA

Tm – melting temperature

U – units

USA – United States of America

Val – Valine

w/v – weight volume ratio

Abstract

Background

Chronic Kidney Disease (CKD), has a high and increasing burden in sub-Saharan Africa. Environmental factors that have been associated to CKD are associated with multiple co-morbidities such as hypertension, diabetes, and HIV. Some genetics factors such *APOL1* have been associated with the highest burden of CKD among population of African ancestries. Other emerging genetic factors such as Sickle Cell trait (SCT) have been investigated mostly among African Americans. Sickle Cell trait (SCT) has the highest burden in sub-Saharan Africans, because of a natural selection, attributed to its protective advantages against the severest form of Malaria, caused by *Plasmodium falciparum*. Many studies showed that SCT has an impact on the normal functioning of the kidneys among African Americans with some studies indicating significant association between SCT and CKD. However, no study has been reported from Sub-Saharan Africa, where most SCT carrier reside. Moreover, there are multiple other loci and variants in the genome that have been associated with CKD in many populations, and that are used for Polygenic Risk Score (PRS) models but have not been explored in populations living in Africa.

Aims

This project aimed to study in a sub-Saharan African cohort, the association between 1) Sickle cell trait (SCT) with Chronic Kidney disease (CKD), and 2) the association of CKD with 29 targeted single nucleotide polymorphisms (SNPs) identified in multiple Genome-Wide Association studies (GWAS).

Methods

Patients and controls: 300 Cameroonian adult participants were included: 150 CKD cases and 150 non-CKD age, sex, and comorbidities matched controls.

Molecular methods: SCT heterozygosity was determined by RFLP-PCR using the restriction enzyme *Dde*I. A total of 29 targeted SNPs was genotyped using MassArray and TaqMan techniques, followed by Sanger sequencing in a subset of samples.

Statistical Analysis: Descriptive statistics and logistic regression, and Fisher exact test were used. Functional pathway analysis: following the identification SNPs with significant association with CKD, we performed functional pathway test using the Linux programme Cytoscape.

Results

The mean age of cases was 53 years (range 46-55 years), with 43% that were female; there were no age and sex significant differences with controls. We identified, an expected, association between CKD and various co-morbidities, demographic and anthropometric variables: hypertension (p value = $5.16X10^{-9}$), HIV (p value = $2.68x10^{-9}$), diabetes (p value = $7.12X10^{-7}$), BMI (p value = $4.58X10^{-8}$) and age (p value = $4.5X10^{-8}$).

HbAS carrier status was significantly associated CKD (p value= $4.3X10^{-9}$; Odds Ratio:7.05). Only three targeted SNPs (3/29) previously associated with CKD in GWAS among African Americans, European and Asian population, were significantly associated with CKD among this group of Cameroonians (*KBTBD2 rs3750082, PTPRO rs7956634* and *LPR2* r*s4667594* with p values of 0.02335, 0.0408 and 0.0398). Genes protein-protein interactions analysis identified the two key functional pathways and one network cluster that could play a crucial role in kidney dysfunctions. Lastly, we distinguished that HbS carrier state doesn't influence the relationship between *APOL1* G1/G2 risk alleles and CKD (p value = 0.5725) in this group from sub-Saharan Africans.

Conclusion and perspectives

Our study illustrates a strong association between SCT and CKD, an important discovery that will have a major implication in preventative medicine policies and practices in both sub-Saharan African where there is a very high prevalence of SCT. The data also has global resonance, with the projected increase in the prevalence of

individual with SCT, due to migration and the improve life expectancy and genetic fitness of people living with both SCT and SCD.

We identified a relatively low proportion of (3/29) of target SNPs positively associated with CKD among this group of Cameroonians. The study illustrates that the vast majority of targeted SNPs associated with CKD in GWAS studies in multiple populations including African American, Europeans, and Asians, are not relevant for sub-Saharan Africans, indicating the urgent need to include diverse populations, specifically those living in Africa. Therefore, the data support the possible bias in currently available Polygenic Risk Score generated from GWAS data, where population from sub-Saharan Africa are largely underrepresented. The data further indicate that there is potential to discover new loci associated with CKD when investigating populations of African ancestry living in Africa.

CHAPTER 1: Introduction

Chronic kidney disease (CKD) is defined as abnormalities in kidney functions or structure, which occurs for a duration of >3 months (Sharma 2013). CKD can be categorised into five stages based on how optimally the kidney function (Table 1). The stages are determined by laboratory techniques, most commonly used is the estimated glomerular filtration rate (eGFR) calculation.

Table 1: Different stages of chronic kidney disease

CKD Stage	Description	Possible symptoms	eGFR (mL/min/1.73 m^2)
Stage 1	Minimal kidney dysfunction		90–120
Stage 2	Mild-moderate kidney dysfunction	Typically, No symptoms	60–89
Stage 3	Moderate-severe kidney dysfunction		30–59
Stage 4	Severe kidney dysfunction	Complications such as anaemia, (hypertension) and abnormal blood levels of Vitamin D, phosphorus and calcium	16–29
Stage 5 (End Stage Renal Disease (ESRD))	Complete kidney failure	Fatigue associated with anaemia, decreased appetite, nausea, vomiting, increased potassium, abnormalities in hormones related to bone health, increased phosphorus and/or decreased calcium, hypertension, and shortness of breath	15 or less

Mortality rates associated with CKD are staggeringly high: following the adjustment for gender, age, co-morbidity, prior hospitalization and race, the mortality in patients with CKD in 2009 was 56% higher than individuals without CKD. The adjusted mortality rate for Individuals with stage 4-5 CKD is estimated to be above 76% (United States Renal Data System. 2017 Annual Data Report. Available at http://www.usrds.org/adr.aspx. Accessed: January 8, 2018). Specifically, Among Africans than Caucasian population, mortality is higher, particularly in men than in women (United States Renal Data System. 2017 Annual Data Report. Available at http://www.usrds.org/adr.aspx. Accessed: January 8, 2018). A study by Schold et.al (2011) conducted within individuals that receive kidney transplants illustrated that the rate of acute rejection and graft loss is higher within African populations when compared to Caucasian populations (Schold et al. 2011). Important differences were also noted in a study conducted on 419 children aged between 1- 16 years. Their

cohort consisted of 285 Caucasian children, 67 African-Americans, 67 from a multiracial background and they illustrated that glomerular diseases are much more prevalent within African population as compared to other ethnicities explored (Wong et al. 2009). CKD appears to be more common in boys, this is due to the posterior urethral valves, which are the most common birth defect causing CKD (https://emedicine.medscape.com/article/238798-overview Accessed January 8, 2018).

Both environmental factors that are associated with multiple co-morbidities such as hypertension, diabetes, and HIV have been associated to CKD (Bruce et al. 2009; Tonelli et al. 2015) . Of note, *APOL1* has been associated with the highest burden of CKD among population of African ancestries (Friedman et al. 2011; Foster et al. 2013; Limou et al. 2014; Kruzel-Davila et al. 2016). Other emerging genetic factors such as Sickle Cell trait (SCT) have been investigated mostly among African Americans (Derebail et al. 2010; Hicks, Carl D. Langefeld, et al. 2011; Rakhi P. Naik et al. 2014).

The renal medulla is characterised as an acidic environment, with hypertonicity and extreme hypoxia (Steinberg & Embury 1986; Kiryluk et al. 2007). When blood flows through the medullary vasa recta, the hyperosmolar milieu causes dehydration of erythrocytes resulting in the erythrocyte sickling and most probable vaso-occlusion and medullary microinfarctions (Van Eps et al. 1970). Micro-infarcts that develop in the kidney's renal medulla and can lead to the inability to concentrate urine because of glomerular hyperfiltration, leading to proteinuria and glomerulosclerosis (Sears 1978; Hostetter 2003; Kiryluk et al. 2007). Sickle Cell trait (SCT) has the highest burden in sub-Saharan Africans, because of a natural selection, attributed to it protective advantage against the severest form of Malaria, caused by *Plasmodium falciparum* (Aidoo et al. 2002; Williams, Mwangi, Wambua, Alexander, et al. 2005; Williams et al. 2012). Many studies showed that SCT has an impact on the normal functioning of the kidneys among African Americans with some studies indicating significant association between SCT and CKD (Rakhi P Naik et al. 2014; Burnham-Marusich et al. 2016). However, no study has been reported from Sub-Saharan Africa, where most SCT carrier lives. Moreover, there are multiple other loci and variants in the genome that have been associated with CKD in many populations, and that are

used for Polygenic Risk Score (PRS) models (Vega & Bustamante 2018) but have not been explore in populations living in Africa.

Assessment of the populations publicly available in Genome-wide association studies (GWAS) databases have revealed a clear, and significant underrepresentation of non-European populations (Rosenberg et al. 2010). A query of the Research Portfolio Online Reporting Tools (RePORT) (Internet resource 1) yielded 2 267 NIH funded GWAS studies, of which ~14% (317 studies) concentrated of the African American population. Additionally, query of PubMed resulted in 4 942 publications, of which ~3% (148) is attributed to African American population (Peprah et al. 2015). The underrepresentation of African ancestry populations in GWAS is of extreme consequence as genetic variation have illustrated significant frequency variation across populations, including the frequency of the risk allele, causal and correlated variants along with prevalence of the disease. Significant differences have been noted within the genetics determinants of disease and their effect size when comparing the European and non-European populations (Rosenberg et al. 2010).

Moreover, GWAS conducted in European population have frequently failed replication in non-European populations (Simón-Sánchez & Singleton 2008; Turner et al. 2008; Mu et al. 2010). GWAS has discovered genetic variants linked with susceptibility for several different traits and disease conditions including asthma (*IGSF3, HPSE2, PSAP, ATG3, MKLN1, XPR1, ABI3BP, KLHL5),* cancer, obesity, hypertension *ALDH1A2, ADH7)*, blood pressure (*RSPO3, PLEKHG1, FAM46A, C6orf37, EVX1, HOXA, ULK4, SOX6, CYB5R2, SEC16B, ADCY3, GNPDA2, FTO, GALNT10, KLHL32, MIR148A, NFE2L3, MC4R, TMEM18)* and malaria *(HBB,SCO1,DDC,ATP2B4,ABO,MARVELD3)* in recent African-ancestry populations. Thus, advocating for the need to increase GWAS conducted with large African population.

This project aimed to study in a sub-Saharan African cohort, the association between 1) Sickle cell trait (SCT) with Chronic Kidney disease (CKD), and 2) the association of CKD with 29 targeted single nucleotide polymorphisms (SNPs) identified in multiple Genome-Wide Association studies (GWAS).

CHAPTER 2: Literature review

2.1 Kidney functions and Kidney disease in Africa

2.1.1 Normal Kidney Function

The kidneys are bean-shaped organs that are situated behind the peritoneum of the vertebral column (Figure 1). Humans have two kidneys on each side of their bodies. Each kidney weighs approximately 150g and measures 12cm in length. The kidneys consist of a hilus, which is a small slit that allows the renal vein, renal artery, renal nerve, and the lymphatics passage into the kidney and includes the renal pelvis. Dissection of the kidney shows two distinct regions, the cortex which is the paler inner region and the medulla which is the darker outer region. The renal pelvis opens up to the renal pyramids which houses the collecting ducts, loops of Henle and vasa recta (blood vessels) which gives the kidney the observed striation pattern appearance (Lote 2013).

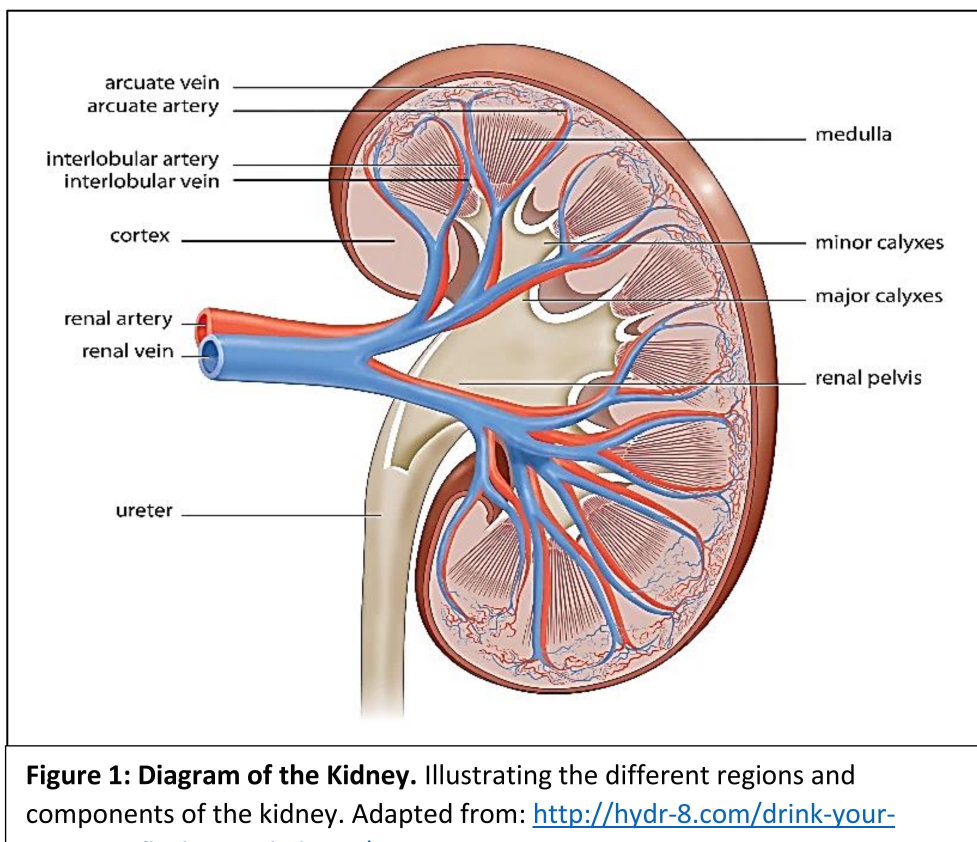


Figure 1: Diagram of the Kidney. Illustrating the different regions and components of the kidney. Adapted from: http://hydr-8.com/drink-your-

The kidneys are responsible for maintaining the environment within the body, through the regulation of water, sodium, potassium, pH, minerals, blood pressure and haemoglobin. All waste material accumulated within the body is filtered out and excreted by the kidneys as urine (Rayner et al. 2016). The kidneys are also responsible for regulating the production of erythrocytes by producing the hormone erythropoietin, regulating blood pressure by secreting the enzyme renin and regulation of calcium absorption by activating vitamin D (Shier et al. 2009).

Kidney function can be determined by measuring the levels of serum creatinine, by measuring the by-product creatinine that is produced from tearing and wearing of muscles. Serum creatinine level greater than 1.2 for females and 1.4 for males is indicative that the kidneys are not functioning optimally. Another test used to measure kidney function is the estimated Glomerular Filtration Rate (eGFR), which is a test that measures how well the kidneys are excreting waste products. The calculation is based on individual serum creatinine level, age, and sex. The normal value for eGFR is above 90. Kidney function can also be monitored by measuring Blood Urea Nitrogen (BUN), which measures the urea nitrogen produced from the breakdown on protein. Normal BUN levels are between 7–20 (NFK, 2017 Available from https://www.kidney.org/atoz/content/kidneytests).

Less invasive procedure for measuring kidney function include urinalysis, which determines the components within the urine such as protein, albuminuria, and creatinine clearance. Increased protein levels in the urine are indicative of kidney dysfunction. The kidney can also be visualized using CT scan and ultrasound to determine any structural abnormalities. The most invasive procedure would be to collect a biopsy of the kidneys to evaluate their function (NFK, 2017 Available from https://www.kidney.org/atoz/content/kidneytests).

2.1.2 Burden and aetiology of Kidney diseases in Africa

Globally, an estimated 1.2 million people died from kidney failure (which is an increase of 32% from 2005)(GBD 2015 DALYs and HALE Collaborators 2016). In 2010, an approximated 2.3 –7.1 million people with end-stage kidney disease (ESKD) died without admission to chronic dialysis (Liyanage et al. 2015). Furthermore, each year, 1.7 million people are assumed to die as a result of acute kidney injury. Overall, 5 –10 million people die from kidney disease each year (Mehta et al. 2015).

Africa is the second largest continent in the world, where it harbours a population of approximately 1 billion people (United Nations 2015). Epidemiologic studies have illustrated that the incidence of kidney diseases is higher in the developing countries than developed countries. There is projections illustrating that by 2030 over 70% of patients with ESKD are estimated to be living in low-income countries (Naicker 2003). Kidney disease, particularly glomerular disease, is more prevalent in Africa and appears to be more severe form than kidney disease found in Western countries (Naicker 2003). Studies have demonstrated that individuals of African ancestry are at a higher risk for developing CKD, with the prevalence of CKD higher in the SSA population that in North Africa, which is nearly two times higher than in the general population (Kiberd & Clase 2002; Peralta et al. 2010).

The known causes of CKD are hypertension (Rao et al. 1984; Botdorf et al. 2011; Segura & Ruilope 2011), diabetes (Villar et al. 2007; US Renal Data System 2013; Liyanage et al. 2015), Human Immunodeficiency Virus (HIV) (Fabian & Naicker 2009; Lucas et al. 2014; Liyanage et al. 2015) and Body Mass Index (BMI) (Gelber et al. 2005; Ejerblad et al. 2006; Lu et al. 2014; Herrington et al. 2017). For sub-Saharan Africa, projections for the year 2030 forecast that 18 -65 million people will have diabetes and forecasts for obesity and hypertension are similarly distressing (Wild et al. 2004; Kearney et al. 2005; Haslam 2007). Considering this projection in the context of CKD, they clearly illustrate the how the burden of CKD will drastically increase, especially in the African populations. Additionally, developing countries are also affected by poverty-related factors such as lack of water supply, infectious diseases, environmental pollutants and high concentrations of disease-transmitting vectors (Jha et al. 2013).

2.2.1 Familial clustering and Heritability of CKD

Studies have revealed that CKD aggregates in families suggesting that the disease has a genetic component, although environmental exposure is still important. A study conducted within an African-American cohort in North Carolina illustrated a fivefold increase in the risk of CKD in an individual with first- or second-degree relatives who have CKD. The risk increased to nine-fold higher when the individuals are limited to those with first-degree CKD relatives (Freedman 1999). Familial clustering within the European population is lower when compared to the African population (Steenland et al. 1990; Spray et al. 1995). The risk of CKD in first degree relatives of European dialysis patients increased to 3-fold compared with the general population (Spray et al. 1995).

Majority of kidney disease are not hereditary; however, some conditions have known inherited genetic components such as autosomal dominant polycystic kidney disease, tuberous sclerosis, Alport syndrome, and Von-Hippel Lindau Syndrome. If these disorders are present within the family, the individual has increased chances of inheriting the disease (Available from https://www.northshore.org/personalized-medicine/medical-genetics/hereditary-conditions/other-conditions/ Accessed January 2019.

2.2 Genetics of kidney disease in Africans

MYH9 and APOL1 and kidney diseases

Genetic factors involved in the development of kidney disease indicate significant predictors in the Myosin Heavy chain 9 (*MYH9*) gene within African Americans (Kao et al. 2008), Hispanic-American (Behar et al. 2010) and Europeans (O'Seaghdha et al. 2011). However, indicated no associative role of *MYH9* polymorphisms within non-diabetic CKD patients from the Nigerian and South African population (Matsha et al. 2012; Tayo et al. 2013). Although, linkage disequilibrium structure may explain the lack of association illustrated within the African population.

It is estimated that the *Apol1/MYH9* locus can explain up to 70% of the differences in ESRD rates between European and African Americans (Cavanaugh & Lanzkron 2010). The *APOL1* nephropathy risk variants are under positive selection within the African continent due to the protective mechanism that they offer against the trypanosomal infection, similarly to the natural selection of the sickle variant (Hicks, Carl D Langefeld, et al. 2011). Although a study by (Naik et al. 2015) illustrated that the association of SCT with CKD appeared to be independent of *APOL1* risk variants.

A significant association between Apolipoprotein L1 (*APOL1*) polymorphisms and non-diabetic kidney disease has been clearly demonstrated in various studies (Behar et al. 2011; Wasser et al. 2012; Fine et al. 2012; Foster et al. 2013; Genovese et al. 2013; Ulasi et al. 2013; Colares et al. 2014; Estrella et al. 2015; Kasembeli et al. 2015; Parsa et al. 2013; Geard et al. 2017), following the preliminary findings reported in African Americans (Tzur et al. 2010; Genovese et al. 2010). The risk is witnessed in the presence of two copies of the risk alleles (homozygous or compound heterozygous) when compared to none or one *APOL1* risk variant (Tzur et al. 2010; Genovese et al. 2010).

HMOX1 and Kidney diseases

Additionally, Heme oxygenase 1 (*HMOX1*) which forms the rate-limiting step in the breakdown of free haem (Tenhunen et al. 1968), with a GT-dinucleotide repeat (rs3074372), two promoter polymorphisms, and a nucleotide polymorphism (SNP) (rs743811), is associated with CKD, particularly in patient affected by sickle cell disease (Bean et al. 2012; Saraf et al. 2015).

2.2.2 Sickle cell disease and Sickle cell trait and CKD

Sickle cell disease (SCD) is a monogenic, multisystem disease, linked to progressive organ damage and acute illness (Weatherall et al. 2005). This haematological disorder results from a point mutation (T>A) on the sixth codon of the Beta-globin gene, located on the 11th chromosome (Chr11:5227002) (Sebastiani et al. 2005; Brousseau et al. 2007). The resulting mutation on the HbS gene causes polymerization and precipitation of haemoglobin when exposed to deoxygenated or dehydrated conditions. This initiates a cascade of events starting with the sickled erythrocytes,

followed by, haemolysis, inflammation, abnormal adhesion of leukocyte and platelets, and hyper-coagulation, which can result in vaso-occlusive crisis (VOC), hypoxia and eventually organ damage (Lozano et al. 2010; Bartolucci & Galacteros 2012).

Sickle Cell Trait (SCT), alternatively, is a heterozygous condition, which results from an individual inheriting a single pathogenic copy of the sickle cell gene (HbS) along with another the wildtype allele for normal haemoglobin (HbA). The compound effect of heterozygotes such as haemoglobin S/beta-thalassemia (Sβthal) and sickle cell haemoglobin C disease, exhibit a milder course when compared to the homozygous form (HbSS) (Schneider et al. 2016). Individuals with SCT exhibit a tri-modal distribution of HbS levels, based on how many α-globin chains are deleted (Brittenham et al. 1985; Tripette, Loko, et al. 2010).

2.2.2.1 Epidemiology of Sickle Cell Trait

SCT affects an estimated 300 million people worldwide. In 2012 in the USA, 7.31% of African, 0.3% Caucasian and 0.22% Asian newborns were born with SCT (Goldsmith et al. 2012; Taylor et al. 2014). However, the highest prevalence of SCD and SCT is observed in individuals of African descent residing in Africa, particularly in areas endemic for malaria (Figure 2).

Nigeria has the highest number individuals affected with SCT at a prevalence of 25–30%. The central African countries of Cameroon and Gabon have a prevalence of 19 and 20%, respectively (Ama et al. 2012; Elguero et al. 2015), while Ghana, Liberia, and Uganda have a prevalence of 10-15% each (Baffour et al. 2015; Tubman et al. 2016; Ndeezi et al. 2016). The increased prevalence of SCD and SCT in European and North American countries is mainly due to the migration of individuals from Africa to other countries. Globally, the overall migrations increased from 92.6 million in 1960 to 165.2 million in 2000 (Özden et al. 2011). However, within the same time period, the number of migrants with HbS increased from 1.6 million in 1960 to 3.6 million in 2000 (Piel et al. 2014).

The burden of the carrier state is on the rise in the European countries, where an estimated 1% of the population harbours the sickle variant. This increase is hypothesized to occur as result of the global migration patterns (Modell et al. 2007) and has resulted in the initiation of neonatal screening programmes in European hospitals, thus allowing early detection and treatment (Bain 2009).

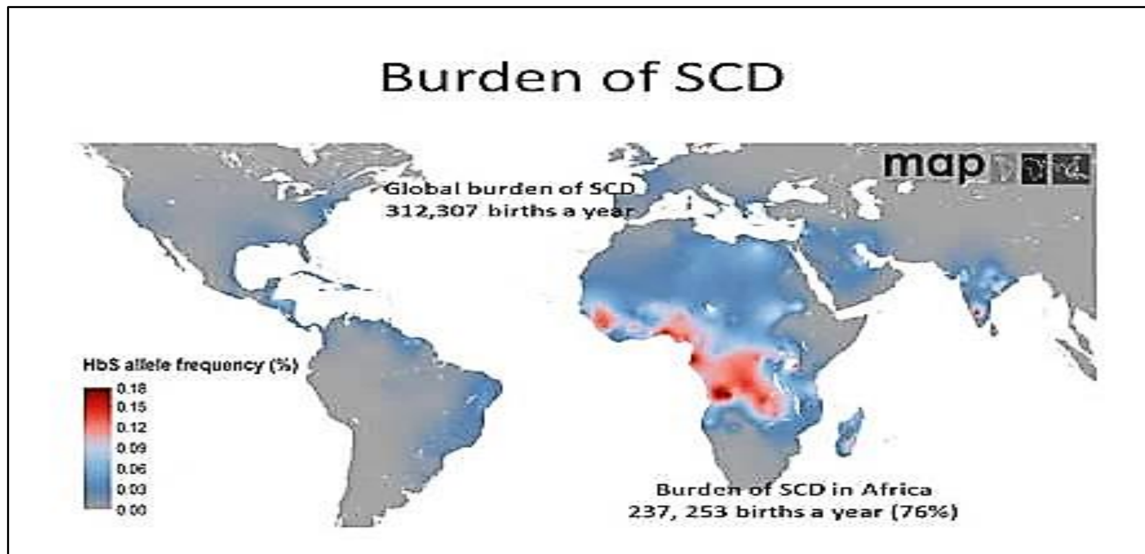


Figure 2: Graphical representation of Global burden of SCD. The highest burden is witnessed within the African continent, particularly in area endemic with malaria. Adapted from (Piel et.al 2013)

SCT erythrocytes contain approximately 40% HbS, where HbA contribute to the remaining balance (Brittenham et al. 1985; Tripette, Loko, et al. 2010). The HbA significantly reduces the possibility of polymer formation, by weakening the effect of the HbS (Brittenham et al. 1985; Tripette, Loko, et al. 2010). Haematological parameters of HbS carriers are largely normal - particularly the size and shape of the erythrocyte, Hb levels and the reduced or complete lack of intravascular haemolysis. Even as some mild abnormalities manifest, they are usually considered subclinical (Tripette, Loko, et al. 2010).

A robust association has been illustrated between the frequency of the HbS and the incidence of malarial infection (Charache et al. 1992; Williams, Mwangi, Wambua, Alexander, et al. 2005). The association was first reported by Allison in 1954, where he first suggested the protection offered by SCT against the severest form of malaria, *Plasmodium falciparum* (Allison 1954). The protection offered by SCT against malaria is thought to be responsible for the continuous presence of the trait within malaria-endemic regions of the world (Elguero et al. 2015).

Williams et.al (2005) clearly illustrated the protective effect of the SCT against malaria. They conducted a study where they inoculated 30 individuals with *Plasmodium falciparum* (15 individuals with SCT and 15 without SCT). They noted that 14 of 15 individuals without SCT developed malaria, in contrast to only 2 of 15 individuals with SCT developed malaria (Allison 1964; Wilcox et al. 1983; Williams, Mwangi, Wambua, Peto, et al. 2005; Hill et al. 2016). Correspondingly, the rate of complicated *falciparum* malaria cases requiring immediate hospitalization reduced by a staggering 90% within individuals with SCT (Williams, Mwangi, Wambua, Alexander, et al. 2005).

There are 4 possible mechanisms proposed on how SCD and SCT offer protection against malaria: 1) Reducing the growth of the parasite within the erythrocyte; 2) Reduced invasion and development of Plasmodium falciparum within the erythrocytes; 3) Stimulation of the innate immune system; and/or 4) Heightened clearance of parasitized erythrocytes by the spleen (Williams, Wambua, et al. 2005; Derebail et al. 2010; Gong et al. 2012; Taylor et al. 2013). SCT protection offered against malaria is only specific for *P. falciparum* and not the other 3 types of malaria (*Plasmodium ovale, Plasmodium malariae and Plasmodium vivax)* (Williams, Mwangi, Wambua, Alexander, et al. 2005). The protection offered by SCT will increase with age, beginning at 20% at the age of two years and increasing to a maximum of 56% at the age of 10 years (Williams, Mwangi, Wambua, Alexander, et al. 2005). It will however eventually decrease to 30% at ages above 10 years (Williams, Mwangi, Wambua, Alexander, et al. 2005).

2.2.2.2 Genetics of SCD

Haemoglobin is a tetrameric molecule consisting of two pairs of polypeptide subunits (Figure 3). The heterotetramer consists of two α-like globin genes which are located on chromosome 16, and two β-like or γ –like globin genes which are located on chromosome 11. During foetal phase of development, the major type of haemoglobin produced is HbF ($\alpha_2\gamma_2$), it is at this point that most haemoglobinopathies go undetected and undiagnosed (Bank 2006). The main haemoglobin within adult humans is a

heterotetramer termed haemoglobin A, which consists of two α-globin and two β-globin polypeptides, along with haem-group (Hardison 2012).

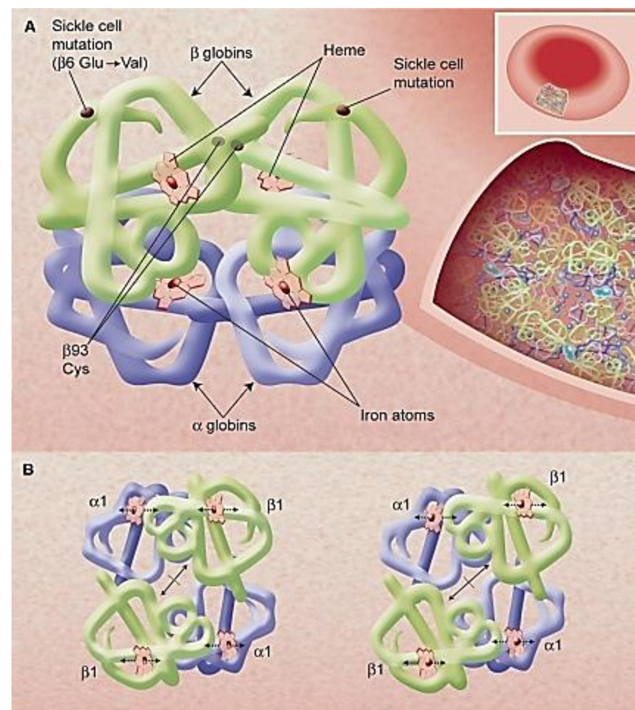


Figure 3: Structure of the haemoglobin molecule. Haemoglobin comprises of two β-globin (green) and two α-globin (purple) chains. The SCD mutation is located on the β-globin chain. Adapted from (Schechter et.al 2008)

The SCD mutation, in haemFoglobin, was first identified by Pauling and his colleagues in 1949 (Pauling et al. 1949). Their work was corroborated and followed-up by Ingram et. al in 1959, where they characterized the mutation as a single point mutation which resulted in a glutamine-to-valine substitution at the sixth residue of the β-globin peptide (Ingram 1959). The SCD mutation results from a single nucleotide base change; whereby it is a monogenic condition; however the clinical manifestation exhibited by the patients are complex, and can range from early childhood mortality to an almost undetectable condition where patients survive to late adulthood (Sebastiani et al. 2005; Bartolucci & Galacteros 2012). The phenotype is highly influenced by the other genetic variants, particularly the deletion of the α-globin chains. The relationship between the genotypic and phenotypic relationship is illustrated in Table 2.

Table 2: Genotypes and phenotypes of sickling disorders*.*

Genotype	Interacting Genes	Typical Severity	% of Hb/total Hb in a typical patient				
			HbS (%)	HbA (%)	HbF (%)	HbC (%)	HBA_2(%)
HbAA	β and β	None	-	96	2	-	2
HbSS	β^s and β^s	Severe	95	-	3	-	2
HbSC	β^s and β^c	Mild	48	-	3	47	2
$HbS\beta^0$	β^s and β^0	Severe	93	-	2	-	5
$HbS\beta^*$	β^s and β^*	Moderate	85	6	5	-	4
$HbS\beta^*$	β^s and β^*	Mild	70	23	3	-	2

*HbAA- Normal haemoglobin
*HbC- Haemoglobin C
*HbF- Foetal haemoglobin
*HbSC- Haemoglobin SC disease
*$HbS\beta^0$- Thalassemia
*HbSS- Sickled haemoglobin

Vaso-occlusive crises (VOC) is the main clinical manifestation of adult SCD and it is characterised by intense bone pain (Bartolucci & Galacteros 2012). Clinical manifestations that are considered very severe include stroke, haemolytic anaemia, cardiovascular disease, pain, acute chest syndrome, priapism, hypertension, nephropathy and pulmonary hypertension hypoxia leading to kidney dysfunctions (Lozano et al. 2010; Bartolucci & Galacteros 2012). Whether Kidney dysfunctions, observed in SCD, is also prevalent in people with sickle cell trait have been controversial in the literature, and is the focus of this dissertation.

2.2.2.3 Sickle Cell Trait and Chronic Kidney Disease (CKD)

Kidney diseases are estimated to affect 5 -18% of patient with Sickle Cell Disease/Trait (Ashley-Koch et al. 2011). Studies have illustrated that renal impairment can affect 4-20% of SCD patients (Herrick 1910; Guasch et al. 2006), while CKD can affect approximately 60% of SCD patient (Guasch et al. 2006; Saraf et al. 2014). CKD impacts all races; however, individuals of African ancestry are more affected. End Stage Renal Disease (ESRD) among African individuals in the USA is approximately 4 times that of the Caucasian population. The highest estimated eGFR is witnessed within the African population is approximately 45-59 mL/min/1.73 m^2 (Hicks, Carl D Langefeld, et al. 2011).

The environment within the kidney is characterised as an acidic environment, with hypertonicity and extreme hypoxia (Steinberg 1986; Kiryluk et al. 2007). When blood flows through the medullary vasa recta, the hyperosmolar milieu causes dehydration of erythrocytes resulting in the erythrocyte sickling and probable vaso-occlusion and medullary microinfarctions (Van Eps et al. 1970). Micro infarcts that develop in the kidney's renal medulla can lead to the inability to concentrate urine as a result of glomerular hyperfiltration, leading to proteinuria and glomerulosclerosis (Sears 1978; Hostetter 2003; Kiryluk et al. 2007).

The medullary infarcts are accompanied with papillary necrosis, which is associated with both microscopic and macroscopic haematuria (presence of blood in urine) and hyposthenuria (impaired urinary conetration) (Van Eps et al. 1970; Gupta et al. 1991; Key & Derebail 2010). In most cases, the left kidney is more involved due to its slightly larger size and higher venous pressure and as a result of compression from the left renal vein by the superior mesenteric vein and aorta (Mitchell 2007; Connes et al. 2008; Tsaras et al. 2009). In a study by (Heller et al. 1979), macroscopic haematuria occurred in both SCD and SCT individuals where it affected approximately 4% of their African American SCT patients.

A study conducted by Key et.al (2010) estimated the prevalence of SCT within 188 African-American End Stage Renal Disease (ESRD) patients on dialysis. Approximately 15% (28 of 188) of the patients were found to have SCT. Their data suggest that SCT exerts an influence on renal disease. Moreover, the patients with SCT were on dialysis for a longer duration. Their findings were supported by another study which used a cohort of 5319 patients on haemodialysis, were they noted a prevalence of 10.2% (543) patients with SCT (Derebail et al. 2014). Additionally, African-American individuals with SCT exhibit two times the prevalence of ESRD than individuals without SCT (Derebail et al. 2010).

In contrast, a case control study by Hick et.al (2011), involving a cohort of 2081 African Americans with ESRD and 1177 controls without ESRD, illustrated no difference among those with ESRD when compared to those without (Hicks, Carl D Langefeld, et al. 2011). Similar findings illustrating the association of SCT and CKD were noted within the Latino and Hispanic with a Caribbean background population (Kramer et al.

2017; Dueker et al. 2017), however, this finding could not be replicated within the Democratic Republic of Congo population (Mukendi et al. 2015).

The largest study conducted on the association of SCT and Renal disorders was published in 2014 in the Journal of the American Medical Association (JAMA) on 15 975 patients, of African American descent, of which 1248 had SCT. Their results illustrated an increased risk of renal diseases in patients with SCT**.** The SCT patients illustrated a higher risk for developing diabetic complications when compared to those with normal haemoglobin (Ajayi & Kolawole 2004; Rakhi P Naik et al. 2014). However, some literature (Bleyer et al. 2010; Ama et al. 2012), shows evidence that suggests that there is no link between SCT and DM or any CKD.

A comparative study by Aloni (2017) reported on glomerular function in children with SCT to children with SCD and healthy children. Their study indicated a significant increase in blood pressure levels when comparing children with SCT and SCD, a moderate decline in eGFR when comparing SCT and SCD patients, increase in hyperfiltration in SCT children although to a lesser extent in SCD children, in comparison to their healthy counterparts. However, in all cases, proteinuria was not detected. No significant difference was seen when comparing uric acid levels in SCT children to healthy children, but the uric levels were lower when in SCT children compared to SCD children. The difference may be due to haemolytic processes occurring in both the SCT and SCD children. The increased hyperfiltration was assumed to result from the HbS polymerisation. This data correlated to previously published literature (Al-Naama et al. 2000).

SCT has been hypothesized to play a role as a possible pathophysiological condition resulting in lower blood pressure (Aloni et al. 2017). The lower blood pressure is assumed to be due to vasodilation, as a counter measure against hypoxia. The cause of the lower blood pressure is still unknown**.**

2.2.2.5 Other targeted genes variants from Genome-Wide Association Studies (GWAS) SNPs that associated with CKD.

Other genomic variants in multiple genes have been associated with kidney functions. For instance variants in the *FGF23* gene, when associated with elevated levels of FGF-23 are independent factor for ESRD within patients with adequate kidney function (Isakova et al. 2011). Sixteen SNPs within the *SHROOM3* gene have also been associated eGFR, along with the development of albuminuria (Ellis et al. 2012). A genome-wide association study (GWAS) including over 130 000 participants found six SNPs involved in decreased GFR, located in or near, *DDX1*, *INO80, CDK12*, *SLC47A1, CASP9*, and *MPPED2* (Pattaro et al. 2012).

Several genes which have been associated with ESRD development, involve aspects of the immune system (*CCR3*, *IL1RN*, *IL4*) and RAS genes, with patients more likely to develop A2350G SNP in the angiotensin-converting enzyme (*ACE)* gene and the C573T polymorphism in the angiotensin II type 1 receptor (*AGTR1* gene) (Su et al. 2012).

Presently, the role of the polymorphisms in the apolipoprotein E (*apoE*) (Lahrach et al. 2014), endothelial nitric oxide synthase (*eNOS)* (Salimi et al. 2006; Kerkeni et al. 2009), xeroderma pigmentosum group D *(XPD*) (Radwan et al. 2015) , X-ray repair cross-complementing protein 1 (*XRCC1*) (Radwan et al. 2015), *renalase* (Rezk et al. 2014), adiponectin (*ADIPOQ*)(El-Shal et al. 2014) and C-C chemokine receptor type 2 *(CCR2)* (Sezgin et al. 2011) genes in the aetiology of CKD remains debated and further investigation in a larger African studies should be conducted to validate these results.

Of a particular note, Pattaro et al. (2016) performed a GWAS meta-analysis in 133 413 individuals from European, Asian and African ancestry. Initially, they determined the phenotype definitions by measuring the serum creatinine and establishing a diabetes diagnosis criterion. CKD was defined as eGFRcrea 60ml min^{-1} per $1.73m^2$. Genotyping was performed using a wide array of NGS platforms (Illumina and Affymetrix 500K), followed by logistic regression to adjust for sex and age using the program METAL (Pattaro et al. 2016).

Replication analysis was performed using eight independent studies, whereby the index study having an established genomic-wide significance of 5.0 X 10^{-8}. The t-test was used to determine the difference between the SNP effect on eGFRcrea in diabetes versus the non-diabetes groups. A similar analysis was performed to compare results in subjects with and without hypertension. Pathway analyses were performed using a novel program DEPICT to prioritize genes at associated loci and to test whether these genes are highly expressed in specific tissues or cell types. It was further used to test whether specific biological pathways and gene sets are enriched for genes in associated loci (Pattaro et al. 2016).

Pattaro et.al (2016)'s study confirmed 29 previously identified loci and identified 48 novel single nucleotide polymorphisms (SNPs). From the 48 SNPs, only 21 were genome-wide significant (P< 5.0 X 10^{-8}). The 48 SNPs were tested for any associations with related phenotypes and very little association was demonstrated between 24 novel SNPs with other kidney-related traits. Only one SNP (*PTPRO* rs7956634) was significant in the African Ancestry, whereas seven SNPs (*IGFBP5* rs2712184, *SKIL* rs9682041,*UNCX* rs10277115,*APB5B1* rs4014195, *PTPRO* rs7956634, *KCNQ1* rs163160, *SIPA1L3* rs11666497) showed significance in the Asian populations (Pattaro et al. 2016).

In the present study, in addition to exploring the association to SCT, we have examined SNPs in 29 loci in a cohort of patients with CKD from Cameroon. The nominated loci were selected based on the following criteria: 1) they appear to be associated with a genome-wide significance (P value <0.06), with kidney functions in all three of the previously mentioned populations i.e. African Americans, Europeans, and Asians; 2) these variants where also all confirmed in a trans-ethnic analyses and meta-analyses in which all replicated loci were included (Table 3), but yet explored among sub-Saharan Africans living in Africa.

Table 3: List of SNPs investigated in the patients and controls cohorts

Loci	dbSNP	Chromosome location	Alleles	Proven disease associations (Ensembl)
UMOD	rs4293393	16:20353266	C/T	Chronic Kidney Disease
ANXA9	rs267734	1:150979001	A/G	Chronic Kidney Disease
GCKR	Rs1260326	2;27508073	C>T	Chronic Kidney Disease
TFDP2	rs347685	3:142088295	A/C	Chronic Kidney Disease
DAB2	rs11959928	5:39397030	A/T	Chronic Kidney Disease
VEGFA	rs881858	6:43838872	A/G	Chronic Kidney Disease
ATXN2	rs653178	12:111569952	A/G	Celiac Disease
SLC22A2	rs2279463	6:160247357	C/T	Chronic Kidney Disease
TMEM60	rs6465825	7:77787122	C/T	Chronic Kidney Disease
SLC6A13	rs10774021	12:240132	C/T	Chronic Kidney Disease
BCAS1	rs17216707	20:54115823	C/T	Glomerular Filtration Rate
SDCCAG8	rs2802729	1:243338461	G/T	Glomerular Filtration Rate
LRP2	rs4667594	2:169151996	A/T	Glomerular Filtration Rate
SKIL	rs9682041	3:170374114	C/T	Glomerular Filtration Rate
UNCX	rs10277115	7:1245559	A/T	Renal Function related trait
KBTBD2	rs3750082	7:32880315	A/T	Glomerular Filtration Rate
A1CF	rs10994860	10:50885664	C/T	Glomerular Filtration Rate
KCNQ1	rs163160	11:2768725	A/G	Glomerular Filtration Rate
AP5B1	rs4014195	11:65739351	C/G	Chronic Kidney Disease
PTPRO	rs7956634	12:15168260	C/T	Glomerular Filtration Rate
NFKB1	rs228611	4:102640552	A/G	Glomerular Filtration Rate
CACNA1S	rs3850625	1:201047168	A/G	Kidney Functions
WNT7A	rs6795744	3:13865353	A/G	Glomerular Filtration Rate
SYPL2	rs12136063	1:109471548	A/G	Glomerular Filtration Rate
ETV5	rs10513801	3:186104564	G/T	Glomerular Filtration Rate

DPEP1	rs164748	16:89641884	C/G	Glomerular Filtration Rate
LRP2	rs4667594	2:169151996	A/T	Glomerular Filtration Rate
SIPA1L3	rs11666497	19:37973622	C/T	Glomerular Filtration Rate
NFATC1	rs8091180	18:79404243	A/G	Glomerular Filtration Rate

RATIONAL

Studies on the impact of SCT on CKD has been limited, particularly in African countries, where the most severe impact of CKD is seen within African countries. There have not any study reported in Cameroon. Moreover, there are multiple other loci and variants in the genome that have been associated with CKD in many populations, and that are used for Polygenic Risk Score (PRS) models but have not been explored in populations living in Africa. Ultimately, the identifications of genomic variants that could affect kidney functions will help in developing a genetic risk model, that will inform anticipatory guidance in clinical practice.

Aim

This project aims to study the association between Sickle cell trait (SCT) and 29 targeted single nucleotide polymorphisms (SNPs) identified in multiple GWAS studies with Chronic Kidney disease (CKD) in an African Cohort.

Objectives

1. To Clinically describe an existing cohort of patients with CKD of various etiology and ethnically controls
2. To Perform SCD diagnosis using Polymerase Chain Reaction (PCR) followed by restriction enzymes fragment length polymorphism (RFLP) using the enzyme *Dde1*
3. To Genotype 29 SNPs that are associated with CKD using MassArray and TaqMan techniques, and Sanger Sequencing
4. To Determining the differential frequency obtained for both SCT and the 29 SNPs by comparing CKD patients and healthy controls

CHAPTER 3: Methods and Materials

In this chapter, we will be introducing all the methodology performed for this study.

3.1 PATIENTS' DATA

3.1.1 Ethical Approval

The study was performed according to the principle of the declaration of Helsinki (Gandevia & Tovell 1964; Ebihara 2000). Ethical approval for this study was granted by the University of Cape Town, Faculty of Health Sciences Human Research Ethics Committee (HREC RE: 661/2015) and the National Ethical Committee Ministry of Public Health, Republic of Cameroon (No 033/CNE/DNM/07).

3.1.2 Participants' descriptions

We used biological samples of an existing cohort of 150 with CKD patients recruited in Yaoundé and Douala, Cameroon, with the available following information: socio-demographic data (family history, age, sex, and ethnicity), Anthropometric variables (Body Mass Index, Systolic Blood Pressure (SBP) and Diastolic Blood Pressure (DBP)), kidney biological functions and the stage of Kidney disease; and the Co-morbidity associated with CKD (High Blood pressure (HBP), Diabetes mellitus, HIV, Lupus). The control cohort comprised of age, sex, and co-morbidity (HBP, Diabetes Mellitus), and an ethnically matched group of 150 control volunteers.

3.2 MOLECULAR METHODS

3.2.1 Samples Preparation

Available DNA samples were previously extracted from the peripheral, using salting our methods, at the University of Yaoundé 1. The DNA purity and concentration were assessed using a NanoDrop® ND-1000 Spectrophotometer (Wilmington, USA). The device operated using NanoDrop® Spectrophotometry computer software. DNA

quantification is based on the Beer Lambert's law which states that the concentration of a sample is directly proportional to the absorption of light by the sample. The principle follows the formula outlined below:

$$A = \varepsilon . b . c$$

The A260: A280 ratio is used to establish the purity levels of the sample. The optimal ratio is ~1.8. Ribonucleic Nucleic Acid (RNA) and Deoxyribonucleic Acid (DNA) consist of nucleic acid which absorbs light at 260 nm. Proteins have a purity ratio of ~2.0 and have an absorb light at 280 nm. Often the A260: A230 ratio is used to determine if there is any contamination such as phenol as a result of the extraction process.

The optimal working sample concentration is 100 ng/µl. Samples with a concentration above 100 ng/µl were diluted to the optimal concentration. Stock DNA with concentrations below 100 ng/µl was aliquoted out to produce the working solution without dilution. The stock DNA was returned to the -20°C freezer whilst the working DNA samples were stored in the 4°C fridge. All experiments were conducted using the working solution.

3.2.2 Primers Design

The primers for a polymerase chain reaction (PCR) amplification for the SCD diagnosis and validation for the 29 SNPs were created using the Ensemble build 79 (Accessed 04/05/2017 and 15/08/2017). The ideal primers have a GC content between 40-60%, a length between 18 and 24 bases long and melting temperature (Tm) between 54-58°C. Deviation from the range results in primers with less specificity and stability.

The WhiteScientific IDT OligoAnalyzer 3.1 (http://eu.idtdna.com/calc/analyzer , Accessed 04/05/2017 and 15/08/2017) was used to analyze the selected primer. The primers were blasted using the PrimerBlast® (http://blast.ncbi.nlm.nih.gov/Blast.cgi) to determine any non-specific binding. The primers are described in Table 4:

Table 4: Tabulation for the primers used for SCD Diagnosis and validation of the genotyped SNPs

Sequence ID	Sequence	Length(bp)	GC content (%)	Melting Temperature (Tm) °C	Objective
rs334_*HBB*_2Cut_F	AGAAGAGCCAAGGACAGGTA	20	50	57.67	**SCD Diagnosis**
rs334_*HBB*_2CUT_R	GTCAGTGCCTATCAGAAACC	20	50	55.85	
rs10994860 *A1CF*_F	ATTGGCTCCACTCTTGAACC	20	50	57.79	**Validating SNPs Genotyping**
rs10994860 *A1CF*_R	GATTATATCTGGGCTGTAGTGACC	24	58.27	58.27	
rs881858 *VEGFA*_F	AGAAGAGGCCCAAATATGCC	20	50	57.63	
rs881858 *VEGFA*_R	AGGAACCTTAGAGAACAAGCG	21	47.62	57.63	
rs6795744 *WNT7A*_F	GGAGTATAACATGCTGAGTTCCC	23	47.83	58.59	
rs6795744 *WNT7A*_R	ACACGTGAAGAAGTGTATGGC	21	47.62	58.62	
rs7956634 *PTPRO*_F	CAAGTCTGGCTTATGATTGCC	21	47.62	57.29	
rs7956634 *PTPRO*_R	GGAATGGAAATTGAGTCTTTAGGC	24	41.67	58.10	
rs4293393 UMOD_F	AAATACAAGGTCACTGTTTCGC	22	40.91	57.52	
rs4293393 *UMOD*_R	AATTTATTTCGCCTCCAAAGGG	22	40.91	57.26	
rs4667594 *LRP2*_F	AGACTGTCCTGCCTTATTTCC	21	47.62	57.36	
rs4667594 *LRP2*_R	GGTTTGCACTGACTTACATTGG	22	45.45	58.10	

The primers were manufactured by Integrated DNA Technology (IDT), but supplied by Whitehead Scientific (Iowa, USA) in powder form and were re-suspended to 100 µM using TE buffer.

3.2.3 Polymerase Chain Reaction (PCR)

PCR is the technique used in the amplification of gene fragments, where the amplified product can be used for a wide range of downstream applications. The reaction consists of a DNA template, buffer solution, deoxyribonucleotide triphosphates (dNTPs), primers and Taq Polymerase. The region-specific primers bind to the DNA template which contains the variant of interest. The primers provide 3'hydroxyl end required by the Taq Polymerase during DNA polymerization, whilst the dNTPs act as the building blocks that are required in the synthesis of the new strand. Optimal reaction conditions are provided by the buffer where it offers balanced pH and salts (mg2+).

The cycling condition for the reaction was obtained using the Bio-Rad thermal cycler T100™ (Bio-Rad Laboratories, Hercules, CA, USA). The cycling conditions included 5 minutes of initial denaturation at 95°C, followed by 30 cycles of 30 seconds of denaturation at 95°C, 30 seconds of annealing at 59°C, 2 minutes of extension at 72°C and a final elongation step of 7 minutes at 72°C.

The amplified DNA fragments were visualized by 3% agarose electrophoresis. 10 µl of the PCR product mixed with 5 µl of loading dye (ThermoFischer) and loaded onto an agarose gel and electrophoresed for 45 minutes at 120V. A 100bp ladder (ThermoFischer) was used as the molecular weight marker as a reference lane to determine the size of the PCR product. The fragments were visualised using with 1 X SYBR safe nucleic acid stain under UV illumination on an UVItech transilluminator.

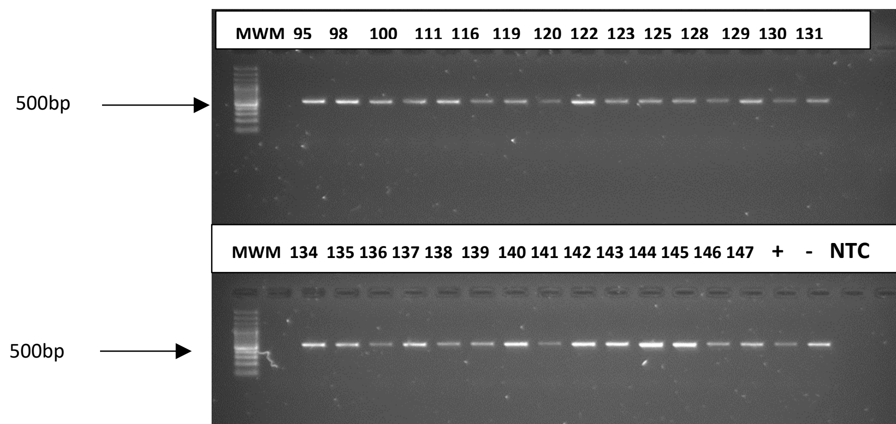


Figure 4: PCR Amplification of the DNA fragment containing the SCD mutation
The image indicates PCR fragment for SCD mutation run at 160V for 1 hour on a 3% (w/v) agarose gel. In the first lane is the molecular weight marker (MWM), the GeneRuler™ 100bp Ladder (Thermo Fisher Scientific). + - positive control, - negative control, * UND indicated the undigested control and NTC-Non-template control.

3.2.4 Sickle Cell Trait Genotyping

Following amplification, to confirm the heterozygosity of the sample the PCR product was digested using a restrictive enzyme *Dde1* enzyme (Promega, USA). The PCR product was digested overnight at 37°C and visualized in a 3% agarose gel run for 45 minutes at 120V. The fragments were visualised using with 1 X SYBR safe nucleic acid stain under UV illumination on an UVItech transilluminator

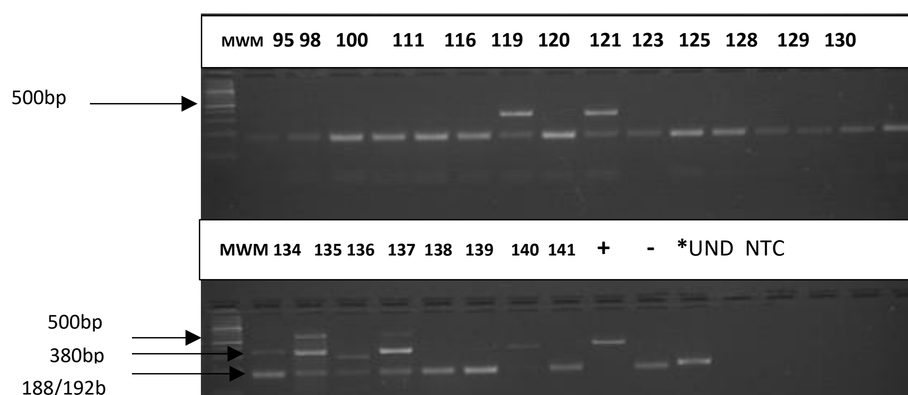


Figure 5: RFLP imaging of the SCD mutation using the DdeI enzyme
The image illustrates the SCD mutation run at 160V for 1 hour on a 3% (w/v) agarose gel. In the first lane is the molecular weight marker (MWM), the GeneRuler™ 100bp Ladder (Thermo Fisher Scientific). + - positive control, - negative control, * UND indicated the undigested control and NTC-Non-template control.

3.2.5. MassArray and TAQMAN® genotyping Assays for the other targeted SNPs

From the 29 SNPs to be genotyped, 26 SNPs were sent to a service provider (Inqaba Biotech, Pretoria, South Africa) for genotyping with the MassArray technique. They were provided with 50 µl of DNA from each of the 150 CKD cases and 150 matched controls with a concentration of 100 ng/µl. The remaining 3 SNPs were genotyped in the Human Genetics laboratory, in the University of Cape Town (UCT) using the TaqMan® technique. MassArray protocol used by Inqaba involves the use of an initial locus-specific PCR reaction, followed by single base extension using mass modified dideoxynucleotide terminators of an oligonucleotide primer which anneals immediately upstream of the polymorphic site of interest. Using MALDI-TOF mass spectrophotometry, the distinct mass primer identifies the SNP allele

TaqMan® involved specific and sensitive assays that detect the amplified products. The specificity and sensitivity are provided by targeted primers and probes in the assays, which come optimized for ideal functional performance on quantitative PCR. Detection for the Single Nucleotide Polymorphisms (SNPs) uses fluorescent signal probes which allow for easier detection of the polymorphisms

TaqMan® Assay contains Taq-polymerase, two fluorophore probes that have Minor Groove Binder (MGB) and two region-specific primer pair (forward and reverse). The probes are labelled with reporter dyes (FAM and VIC) which release fluorescence when the reporter is cleaved off from the 5' end releasing it from the non-fluorescent quencher.

The assays were received in a stock 40X concentration and were diluted to 20X using 1X TE buffer (0.3M EDTA and 7M Tris, pH 7.3) to prepare the working solution. A reaction volume of 10 µl consisted of 5 µl of 20X TaqMan® Genotyping Master Mix (Applied Biosystems, Thermo Fischer Scientific, Warrington, UK), 0.5 µl, of 20X TaqMan® Assay (Applied Biosystems, Thermo Fischer Scientific, Warrington, UK), 2.5 µl, of nuclease-free water and 2 µl of DNA. The mix was added into each well of 96-well plate (Bio-Rad Laboratories, Hercules, CA, USA). The cycling condition started with initial 10 minutes amplification at 95°C, 15 seconds denaturing at 95°C and 1-minute annealing/extension step for at 60°C using CFX96 real-time system (Bio-Rad Laboratories, Hercules, CA, USA). Analysis of the data performed using the Bio-Rad CFX96 (Bio-Rad Laboratories, Hercules, CA, USA) Manager Software (version 3.1).

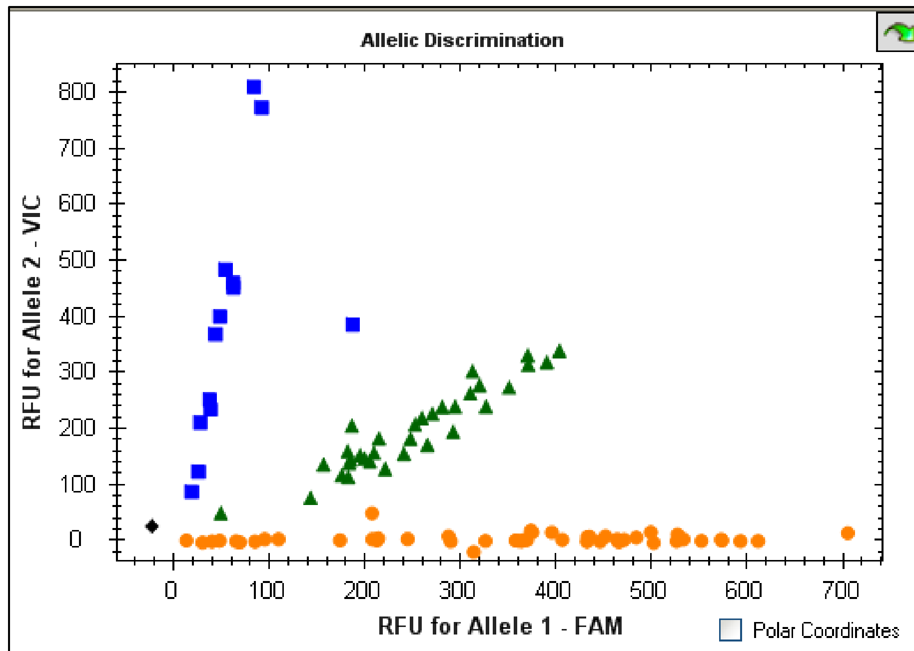


Figure 6: TaqMan SNP Genotyping plot for the cohort for rs881858. The TaqMan allelic discrimination plot was produced using the Bio-Rad CFX96 (Bio-Rad laboratories, Hercules, CA, USA). The Y-axis corresponds to the (RFU) for the VIC 5'-dye label while the X-axis corresponds to the (RFU) for FAM 5'-dye. The FAM and VIC 5'dye corresponds to the C and A alleles, respectively. The blue squares represent the homozygous for the CC allele, the green triangles represent the heterozygous AC allele and orange circles represent the homozygous AA allele. The analysis was preformed using the Bio-Rad CFX96 (Bio-Rad laboratories, Hercules, CA, USA) Manager software version 3.1.

3.2.6 Cycle Sequencing for validation of genotyping

A subset of the cohort (6 SNPs in 10% of the cohort) initially genotyped either by Mass Assay or the TaqMan, was validated through Sanger sequencing and illustrated in figure 7. Sanger sequencing is the gold standard technique for genotype validation, as it allows visualization of the individual base in the sequence (Sanger et al., 1977). A total of 10 µl reaction requires 1 µl of BigDye® Terminator v3.1 Cycle Sequencing mix (Life Technologies), 2 µl BigDye® Terminator v3.1 dilution buffer, 1 µl of primer (reverse or forward primer), 4 µl distilled Sabax water, 2 µl of the PCR product. The reaction mix contains four dye-labelled dideoxynucleotides (ddNTPs) which have an absent hydroxyl group, resulting in the termination of DNA synthesis. The sequencing product was precipitated by adding 50 µl of absolute ethanol and 2 µl of sodium acetate. Followed by, vortexing and overnight incubation at -20°, ensuring DNA precipitation and high yield of the DNA sequences.

The sequences were vortexed, and the supernatant was discarded. The sequences were re-suspended in 50 µl of 70% ethanol and mixed by vortexing. This was followed by centrifugation for 10 minutes at 10 000g and the supernatant was discarded, and the sequences were allowed to air-dry. This pellet was re-suspending 10µl HiDi Formamide.

The sequencing products have different fragments lengths, which are fluorescently labelled at the 3'end. DNA is separated by size (1bp) using capillary electrophoresis and fluorescently tagged ddNTPs using the 3130*xl* Genetic Analyzer ABI Prism (Applied Biosystems).

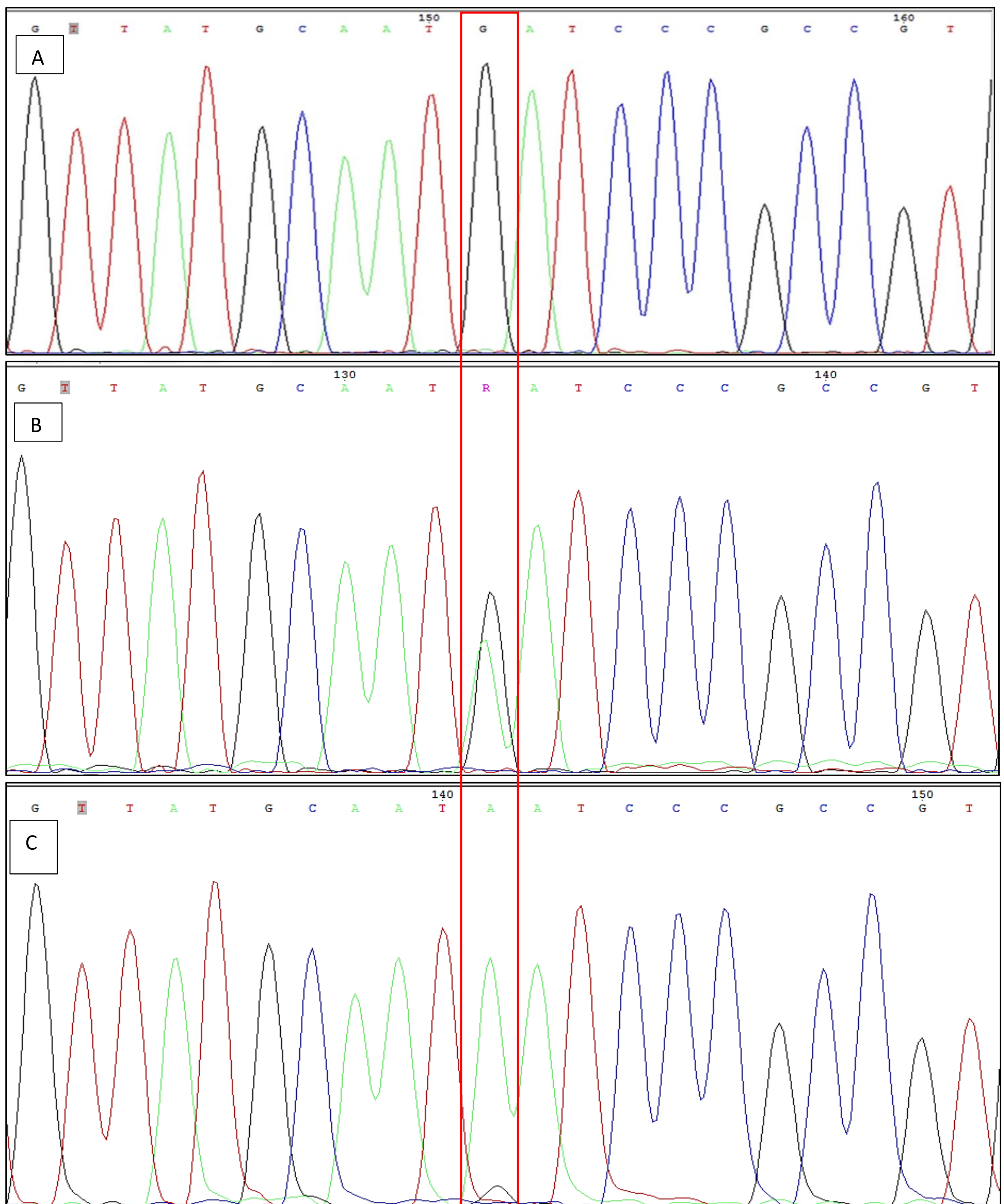


Figure 7: Sanger sequencing validation for rs881858 in patient 7, 10 and 15. Electrogram produced using ABI Prism 3130*xl* Genetic Analyser (Applied Biosystems). Analysis was performed using the software program Chromas (Version 2.6.6) for analysis of the genotype for rs881858. (A) is the GG wildtype genotype for patient 7, (B) is the GA heterozygous genotype for patient 10 and (C) is the AA mutant genotype for patient 15

3.3 Statistical Analysis

Statistical analysis of the cohort was performed using STATA® SE-64 software program (version 14.0.370 for Windows). Descriptive statistics were obtained for the sociodemographic data attained during recruitment. Plink (Appendix 1)(Purcell et al. 2007) was used in the analysis of the results obtained for the SCT diagnosis and also the association of the targeted SNPs with CKD. Logistic regression was employed to examine the genetic association between SCT and CKD (p-value < 0.05) and any association between CKD and the targeted SNPs (p-value < 0.05), while Fischer's exact test was used to examine any association with SCT and other clinical phenotypes (p-value < 0.05).

Following the identification of any associated SNPs, we performed functional pathway test using the Linux programme Cytoscape. Cytoscape is an open source platform used for visualization and complex networks analysis such as networks of genes or interacting proteins.

Lastly, using already available data on *APOL1* (Geard et.al 2017) which illustrated no association with CKD in the sample cohort, we analysed the influence of SCT on the *APOL1*-CKD relationship using ODDs ratio analysis.

CHAPTER 4: RESULTS

In this chapter, we will be presenting the results obtained from our study.

4.1. Population Description and clinical phenotypes

4.1.1 Socio-demographic and anthropometric variables and CKD

Table 5 describes the socio-demographic and anthropometric variables of the CKD patients and non-CKD controls indicating most of the participants were adults with a mean age of 53 years (range: 46-55) and matched as designed with non-CKD controls with a mean age of 54 years (range: 46-61). As designed, the cohort consisted of equal proportions of females and males with a p-value of 0.771.

Both the cases and control illustrated a normal distribution of BMI values, systolic and diastolic blood pressure with p-values of 0.00, 0.002 and 0.10 respectively. The diastolic and systolic blood pressure appears as above normal in the patients' group, due to the associated co-morbidity, as increased BP values are expected within a CKD cohort.

Table 5: Social-demographic and anthropometric variables and CKD

Variables		CKD Cases %(n)	Non-CKD controls %(n)	P-value
Age(years)		53	54	0.459
Gender	Female %(Number)	28.6 (43)	30 (45)	0.771
	Male %(Number)	71.3 (107)	70 (105)	
BMI (kg/m^2)		24	29	**0.00**
Systolic blood pressure(mmHg)		148	142	**0.002**
Diastolic blood pressure(mmHg)		85	83	0.10
Total Observation		150	150	

Figure 8 illustrates the distribution of BMI and Age in CKD cases and control while figure 9 illustrates the distribution of the Systolic blood pressure and diastolic blood pressure. The BMI, age, systolic and diastolic blood pressure showed normal distribution for both patients and control.

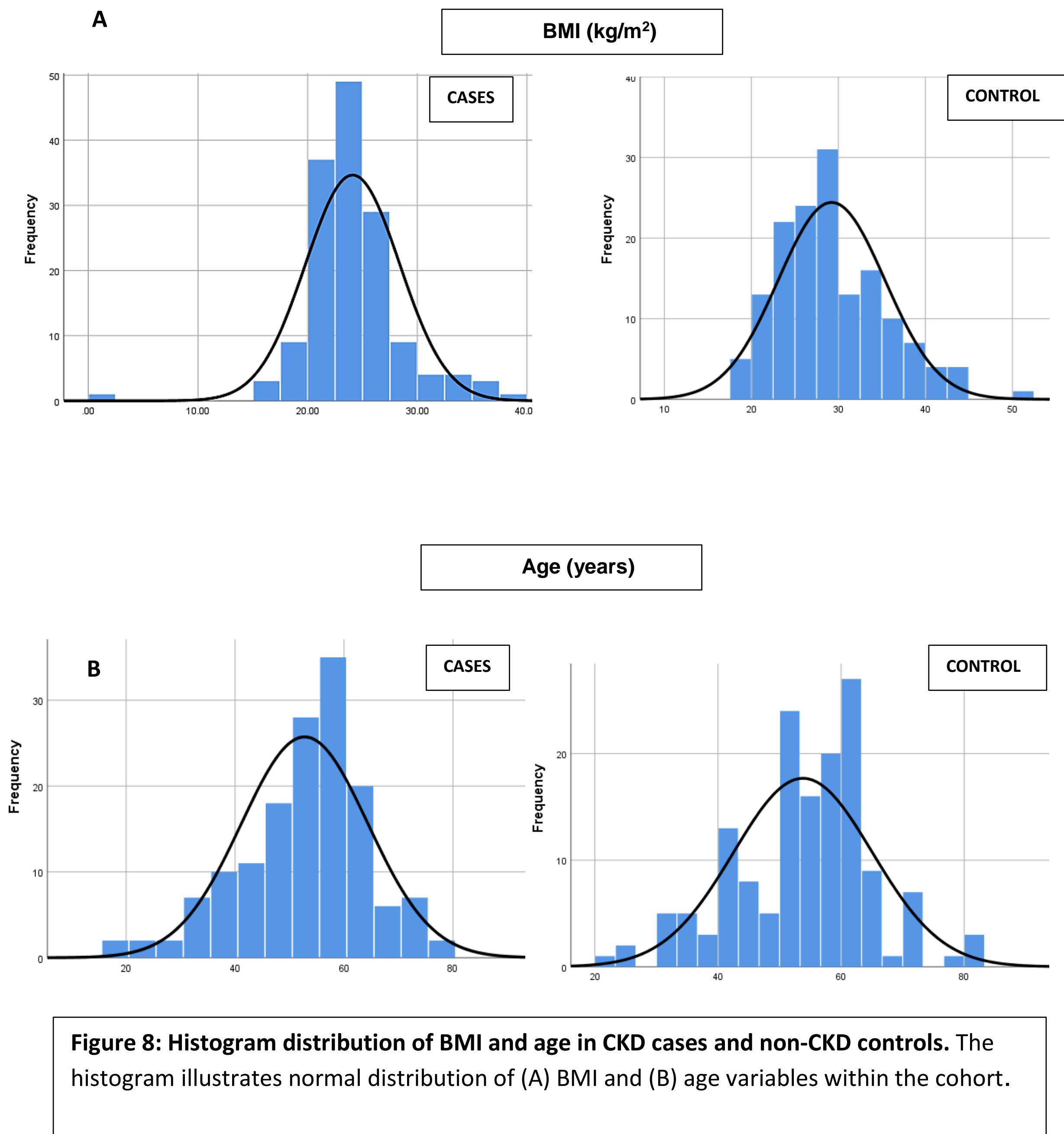


Figure 8: Histogram distribution of BMI and age in CKD cases and non-CKD controls. The histogram illustrates normal distribution of (A) BMI and (B) age variables within the cohort.

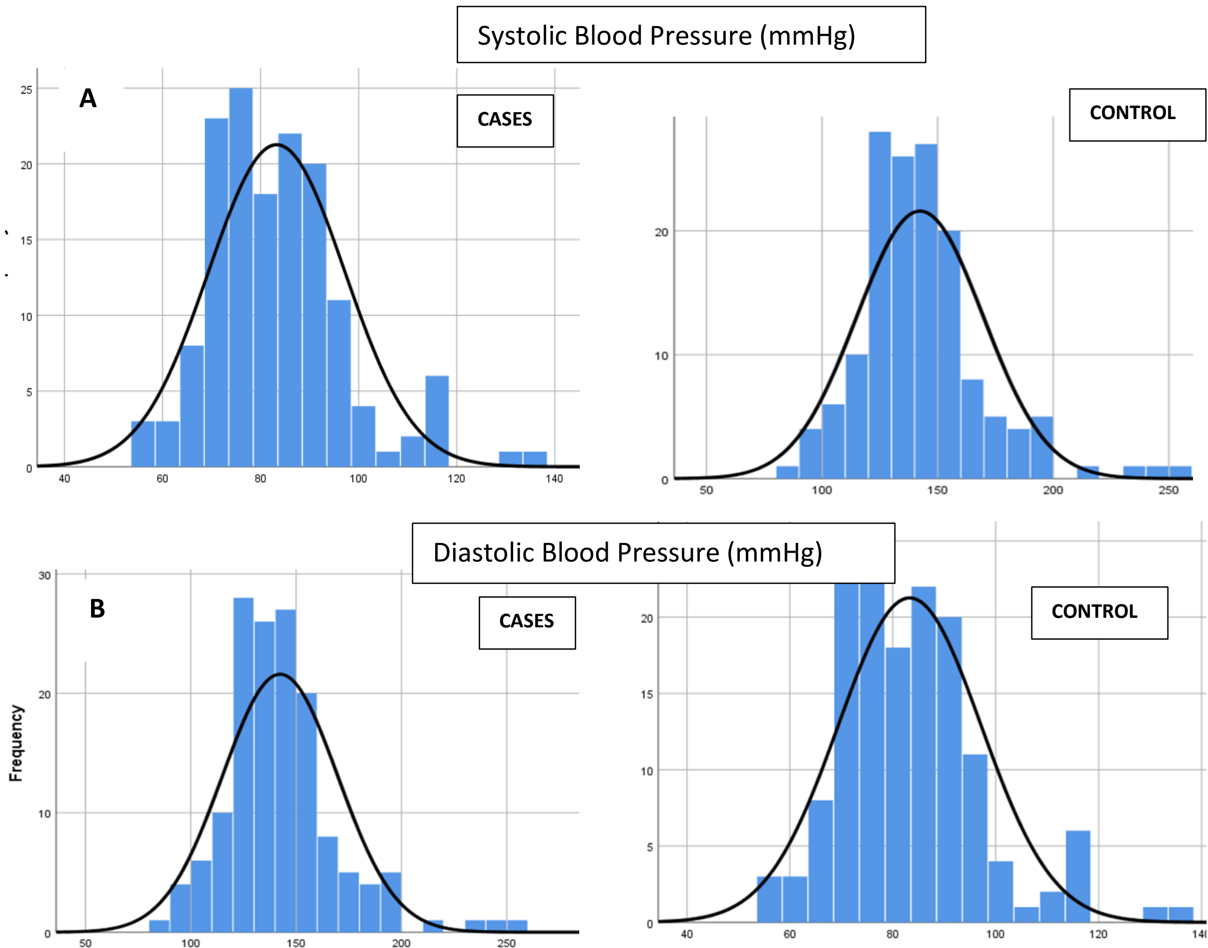


Figure 9: Histogram distribution of Systolic Blood pressure (mmHg) and Diastolic Blood pressure (mmHg) in CKD cases and non-CKD controls. The histogram illustrates normal distribution of (A) Systolic Blood pressure (mmHg) and (B) Diastolic Blood pressure (mmHg) variables within the cohort.

4.2 Validity of the statistical model used in the analysis of the cohort

The model used in the analysis of the cohort indicates True Positive Rate (TPR) of approximately 1 and a precision rate of approximately 1 (Figure 10), both illustrating that the model used is dependable, rigid and accurate. The regression model has a chi-square of 281.697 with a p-value of $1.60X10^{-52}$ and the model has achieved accuracy and precision of 0.91 and 0.89 respectively.

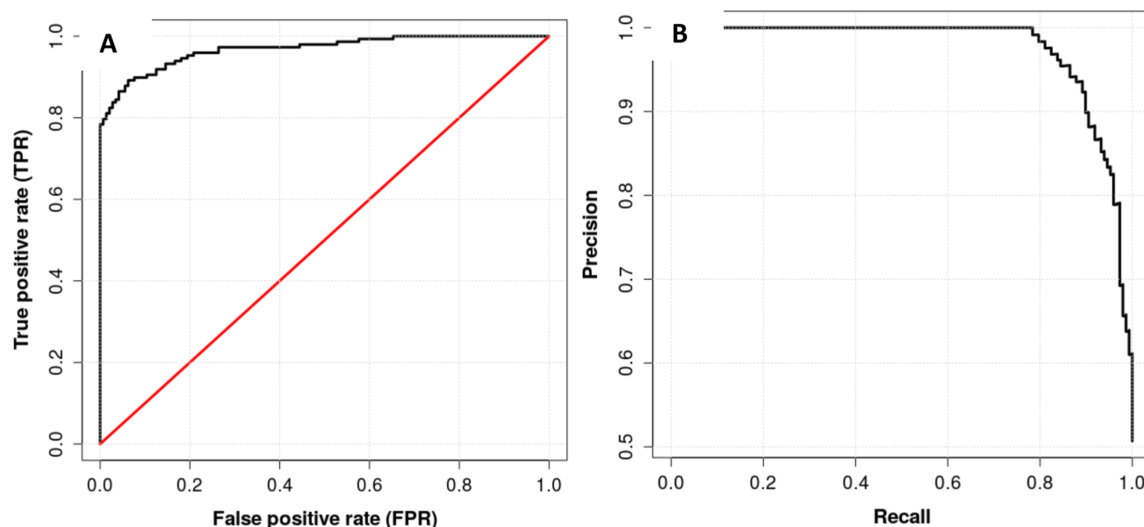


Figure 10: Graphical representation of the validity of the model used in the analysis of the cohort. The TPR is approximately 1, while the precision and recall are approximately 1, indicating that the results obtained from the model are reliable.

4.3 Description of CKD and other clinical co-morbidity

The logistic regression model confirmed an association between CKD and diabetes, HIV and hypertension at p values of $7.12X10^{-7}$, $2.68X10^{-9}$, $5.16X10^{-9}$ respectively (Table 6). An association between CKD and hypertension (Table 6) was established with a log ODDs ratio of 8.09226 indicating that a hypertensive individual has a probability of 1 of developing CKD when compared to healthy individuals. The study also revealed an association of CKD and HIV (Table 6) with a log ODDs ratio of 8.7045 indicating that the carrier has a probability 1 of developing CKD. Diabetes was associated with CKD (Table 6) at a log ODDs ratio of 6.5663 indicating that the individual has a probability 1 of developing CKD.

The associated phenotypes were represented in bar graphs in figure 11.

Table 6: Association of CKD and other clinical phenotypes

Variables	ODDs ratio	P-value
Diabetes	6.56663	**$7.12X10^{-7}$**
HIV	8.7045	**$2.68X10^{-9}$**
Hypertension	8.09226	**$5.16X10^{-9}$**

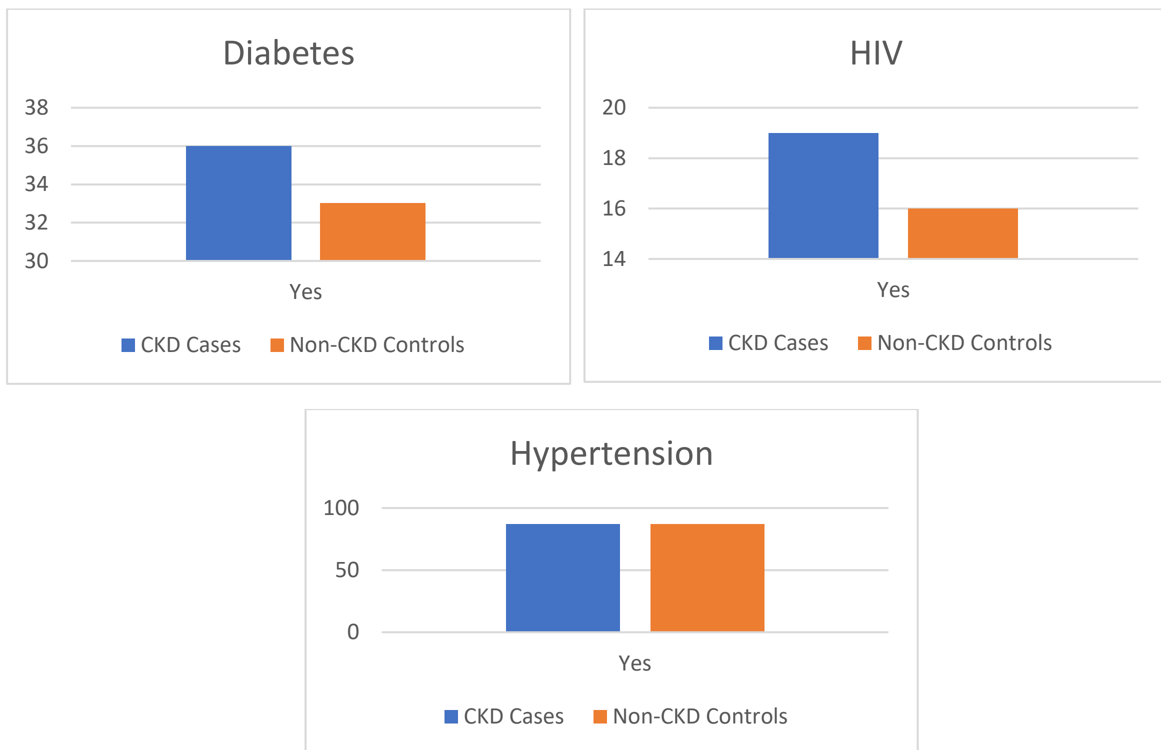


Figure 11: Bar graph representation of additional clinical phenotypes that are positively associated in influencing the development of CKD. A) Diabetes B) HIV and C) Hypertension

4.4 Description of the SCT within the CKD cases and non-CKD controls

Table 7 illustrates a significantly higher frequency of AS individuals within the CKD cases (n=49) when compared to the non-CKD controls (n=10) with an ODDs ratio of 7.05 and p-value of $4.371X10^{-9}$. This indicates that carriers have a 7.05 times chances of developing CKD when compared to controls individuals.

Table 7: Association of SCT within CKD cases and non-CKD controls.

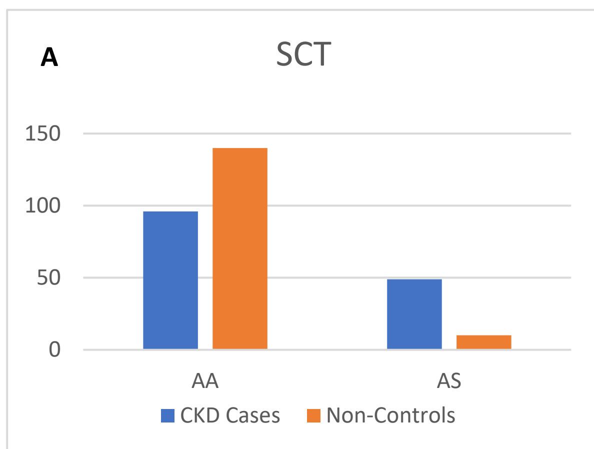
		CASES %(n)		CONTROLS %(n)		ODDs Ratio	P Value
		AA	**AS**	**AA**	**AS**	**-**	**-**
Frequencies		66 (96)	34 (49)	93 (140)	7 (10)	7.04	$4.371X10^{-9}$
Age (mean)		53	53	54	54	1	0.459
Gender	**Female**	25 (24)	97 (32)	32 (44)	10 (1)	NA	
	Male	75 (72)	3 (1)	68 (95)	90 (9)	NA	
Family History	**yes**	9 (9)	10 (5)	7 (10)	100 (10)	NA	
	No	91 (87)	90 (44)	93 (128)	0 (0)	NA	
Systolic Blood Pressure(mmHg)		149	149	142	145	1	0.02
Diastolic Blood Pressure(mmHg)		85	85	83	84	1	0.10
BMI (kg/m^2)		24	24	29	29	1	0.00

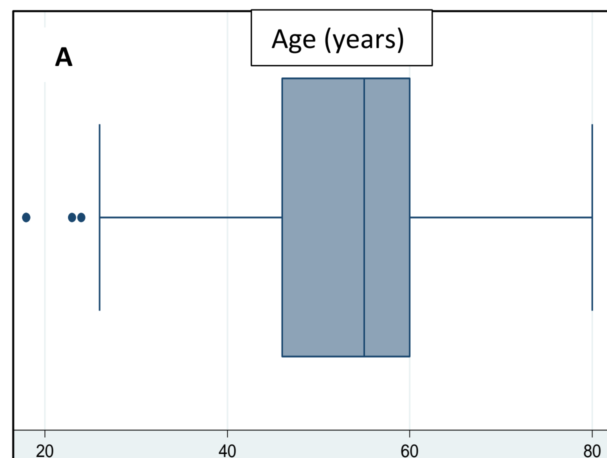
4.5 Association between SCT and CKD

Logistic regression (Table 8) was performed on the cohort which illustrated an association between SCT, BMI, Age and CKD at p values of 0.000625, $4.58X10^{-8}$ and 0.016227 respectively. The association between SCT and CKD achieved a log ODDs ratio of 2.40223 indicating that a carrier has a 0.92 probability of developing CKD as compared to healthy individuals. The association between BMI and CKD illustrated a log ODDs ratio of 0.32816 indicating that the carrier has 0.58 probability to develop CKD when compared to healthy individuals. Age was also identified as a variable that influences the development of CKD at a log ODDs ratio of 0.0621 indicating that the carrier has a 0.52 probability of developing CKD. All the factors that were associated are illustrated in figure 12.

Table 8: Factors influencing CKD within the cohort

Variables	Z value	P value
SCT	3.421	**0.000625**
Age	2.404	**0.016227**
Sex	0.425	0.671162
BMI (kg/m2)	5.467	**$4.58X10^{-8}$**
Systolic Blood Pressure (mmHg)	-0.517	0.604984
Diastolic Blood Pressure (mmHg)	-1.238	0.215817





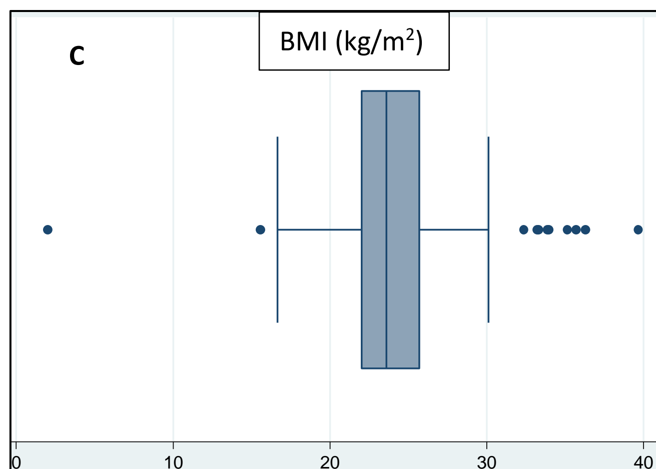


Figure 12: Graph representation of positively associated factors influencing the development of CKD. A) Bar graph- SCT B) Box and Whisker- Age and C) Box and Whisker -BMI

4.6 Description of SCT and other clinical co-morbidity

Analysis of the cohort revealed the association of SCT and other clinical co-morbidity; particularly, hypertension and Human Immunodeficiency Virus (HIV); the only exception was diabetes which was not influenced by the SCT, illustrated in Table 8 and figure 13. An association between SCT and HIV (Table 9) was established with a highly significant p-value of 0.000323 and ODDs ratio of 6.586044 indicating that the carrier could have 6.59 times chances of increased risk of attaining HIV when compared to healthy individuals. Further analysis of the cohort revealed an association between SCT and hypertension (Table 9) with a p-value of 0.04639 and ODDs ratio of 1.831774 indicating that the carrier has 1.83 times more risk of hypertension as compared to healthy individuals. Diabetes was the only variable not influenced by the SCT (Table 9) with a p-value of 0.7203 and an ODDs ratio of 1.129938 indicating no evidence that SCT influences diabetes state.

Table 9: Association of SCT and other clinical co-Morbidity

Variables	ODDs ratio	P-value
Diabetes	1.129938	0.7203
HIV	6.586044	**0.000323**
Hypertension	1.831774	**0.04639**

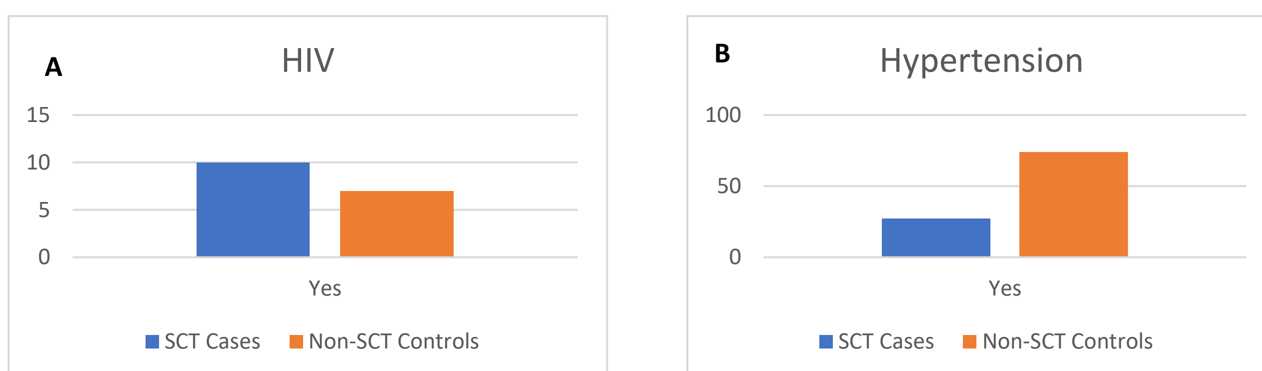


Figure 13: Bar graph representation of clinical phenotypes influenced by the presence of SCT. A) HIV B) Hypertension

4.7 Frequencies of CKD-related genes variants across various populations

Table 10: Allele frequencies of selected CKD-related genes variants among Cameroonian Controls and data extracted from the 1000 Genome project.

Gene	dbSNP	location	Allele	Cameroon	African	East Asian	European	Cameroon vs African P values	Cameroon vs East Asia P values	Cameroon vs Europe P values
UMOD	Rs4293393	16:20353266	A>G	0.22	0.193	0.084	0.201	**0.0394**	**0.00001**	0.0903
ANXA9	rs267734*	1:150979001	T>C	0.009	0.069	0.026	0.194	**0.0035**	0.1432	**0.0000**
GCKR	rs1260326	2:27508073	C>T	0.005	0.094	0.481	0.411	0.7792	**0.00001**	**0.00001**
TFDP2	rs347685	3:142088295	A>C	0.26	0.264	0.289	0.278	0.9244	0.3744	0.5765
DAB2	rs11959928	5:39397030	A>G	0.28	0.287	0.174	0.451	0.0762	**0.00001**	**0.0011**
SLC22A2	rs2279463	6:43838872	A>G	0.19	0.186	0.082	0.128	**0.0374**	**0.00001**	**0.00001**
TMEM60	rs6465825*	6:160247357	C>T	0.44	0.493	0.780	0.574	0.2882	**0.00001**	**0.0002**
SLC6A13	Rs10774021	12:240132	C>T	0.40	0.45	0.800	0.0185	**0.0001**	**0.00001**	**0.00001**
BCAS1	rs17216707*	20:54115823	C>T	0.02	0.035	0.043	0.227	**0.0185**	**0.0001**	**0.00001**
LRP2	rs4667594	2:169151996	A>T	0.37	0.521	0.865	0.488	**0.0364**	**0.0118**	**0.00001**
SKIL	rs9682041	3:170374114	C>T	0.26	0.290	0.060	0.121	**0.0000**	**0.00001**	**0.00001**
UNCX	rs10277115	7:1245559	A>T	0.21	0.116	0.308	0.754	0.0640	**0.00001**	**0.00001**
KBTBD2	rs3750082	7:32880315	A>T	0.37	0.330	0.656	0.635	0.5725	**0.00001**	**0.00001**
A1CF	rs10994860	10:50885664	C>T	0.28	0.222	0.047	0.202	0.1577	**0.00001**	**0.00001**
KCNQ1	rs163160*	11:2768725	A>G	0.05	0.062	0.208	0.179	0.6550	**0.00001**	0.2277
AP5B1	rs4014195	11:65739351	C>G	0.28	0.263	0.237	0.351	0.3945	**0.00001**	**0.00001**
PTPRO	rs7956634	12:15168260	C>T	0.45	0.550	0.273	0.829	**0.0248**	0.7481	**0.0042**
NFKB1	rs228611	4:102640552	A>G	0.22	0.188	0.486	0.482	**0.0000**	**0.00001**	**0.00001**
CACNA1S	rs3850625	1:201047168	A>G	1	0.995	0.960	0.882	0.3726	**0.00001**	**0.00001**
WNT7A	rs6795744	3:13865353	A>G	0.24	0.231	0.081	0.158	0.2274	**0.0008**	**0.00001**
SYPL2	rs12136063	1:109471548	A>G	0.28	0.269	0.957	0.689	0.7273	**0.00001**	0.0140
ETV5	rs10513801*	3:186104564	G>T	0.004	0.012	0.063	0.138	0.7086	**0.00001**	**0.00001**
DPEP1	rs164748	16:89641884	C>G	0.08	0.064	0.016	0.477	0.2086	**0.0001**	0.1831
SIPA1L3	rs11666497	19:37973622	C>T	0.06	0.061	0.110	0.195	0.2574	**0.00001**	**0.00001**
NFATC1	rs8091180	18:79404243	A>G	0.12	0.125	0.784	0.597	0.9736	0.5448	0.1831
*IGFBP5***	Rs2712184	2:217682779	A>C	0.46	0.442	0.533	0.572	0.5921	**0.00001**	**0.00001**
VEGFA	rs881858	:43838872	A>G	0.28	0.656	0.773	0.695	0.3373	**0.0004**	**0.00001**
ATXN	rs653178*	12:111569952	A>G	0.04	0.981	0.997	0.534	**0.0000**	**0.00001**	**0.00001**
SDCCAG8	rs2802729	1:243338461	A>C	0.46	0.490	0.319	0.445	0.3356	0.0969	**0.0002**

Significant P values are bolded. dbSNP ID; single nucleotide polymorphism database identification; NA, not applicable.* indicates monomorphic SNPS

The differential frequencies across populations of the SNPs investigated are presented in Table 10. A total of 10/29 SNPs (34%) were differentially distributed among Cameroonian individuals compared to African controls from the 1000 genomes project. Only 2/9 SNPs illustrated a significant or borderline association with CKD. Up to 83 % of the variants studied (24/29) were differentially frequent when comparing Cameroonian versus East Asian individuals with CKD (Table 10). Additionally, comparison with European American population illustrated a significant difference in allele in 22/29 (75.8 %) SNPs.

4.8 Association of CKD and targeted genomic variants

Five SNPs (*ANXA9 rs267734, BCAS1 rs17216707,KCQN1 rs163160,CACNA1S rs3850625, ETV5 rs10513801* and *ATNX2 rs653178*) were monomorphic and further excluded from analysis, however, the model illustrated an association between CKD *KBTBD2 rs3750082,PTPRO rs7956634* and *LPR2* r*s4667594* with p values of 0.02335, 0.0408 and 0.0398and the 3 targeted SNPs, respectively.

Additionally, two SNPs DAB2 *rs11959928* and UNCX *rs10277115* illustrated a borderline association with CKD with p values of 0.08937 and 0.06895 respectively (Table 11).

Table 11: Tabulation of the 29 targeted SNPs investigated within the cohort.

Loci	dbSNP	Chromosome location	Alleles	Minor Allele Frequency (MAF)	P-Values	Chi-Square	ODDs Ratio
UMOD	rs4293393	16:20353266	C/T	0.25	0.5461	0.3643	1.132
ANXA9	rs267734	1:150979001	A/G	0.006	0.7585	0.09451	0.6485
GCKR	Rs1260326	2;27508073	C>T	0.40	0.398	4.227	1.969
TFDP2	rs347685	3:142088295	A/C	0.010	0.8107	0.05737	0.9534
DAB2	rs11959928	5:39397030	A/T	0.26	**0.08937**	2.886	1.379
VEGFA	rs881858	6:43838872	A/G	0.34	0.6999	0.1486	1.077
ATXN2	rs653178	12:111569952	A/G	0.24	0.1469	2.105	0.3848
SLC22A2	rs2279463	6:160247357	C/T	0.45	0.2209	1.498	1.297

TMEM60	rs6465825	7:77787122	C/T	0.01	0.9299	0.007728	0.9847
SLC6A13	rs10774021	12:240132	C/T	0.31	0.6625	0.1905	1.095
BCAS1	rs17216707	20:54115823	C/T	0.24	0.2409	1.375	0.4514
SDCCAG8	rs2802729	1:243338461	G/T	0.13	0.5167	0.4206	0.8847
LRP2	rs4667594	2:169151996	A/T	0.28	0.333	0.9371	0.7867
SKIL	rs9682041	3:170374114	C/T	0.23	0.1056	2.619	0.7138
UNCX	rs10277115	7:1245559	A/T	0.05	**0.06895**	3.308	0.6591
KBTBD2	rs3750082	7:32880315	A/T	0.25	**0.02335**	**5.142**	**0.6337**
A1CF	rs10994860	10:50885664	C/T	0.45	0.3036	1.058	0.8171
KCNQ1	rs163160	11:2768725	A/G	0.21	0.3813	0.7666	0.6912
AP5B1	rs4014195	11:65739351	C/G	1	0.5458	0.3648	0.8545
PTPRO	rs7956634	12:15168260	C/T	0.22	**0.0408**	**4.184**	**1.444**
NFKB1	rs228611	4:102640552	A/G	0.26	0.8068	0.05984	0.9504
CACNA1S	rs3850625	1:201047168	A/G	0.004	NA	NA	NA
WNT7A	rs6795744	3:13865353	A/G	0.08	0.6032	0.2702	0.8989
SYPL2	rs12136063	1:109471548	A/G	0.06	0.4873	0.4825	0.8711
ETV5	rs10513801	3:186104564	G/T	0.11	0.979	0.000693	1.038
DPEP1	rs164748	16:89641884	C/G	0.40	0.6516	0.2039	1.155
LRP2	rs1260326	2:169151996	A/T	0.30	**0.0398**	**4.277**	**1.969**
SIPA1L3	rs11666497	19:37973622	C/T	0.01	1	NA	NA
NFATC1	rs8091180	18:79404243	A/G	0.44	0.7399	0.1102	0.9152

Significant P values are bolded. dbSNP ID; single nucleotide polymorphism database identification; NA, Not applicable.

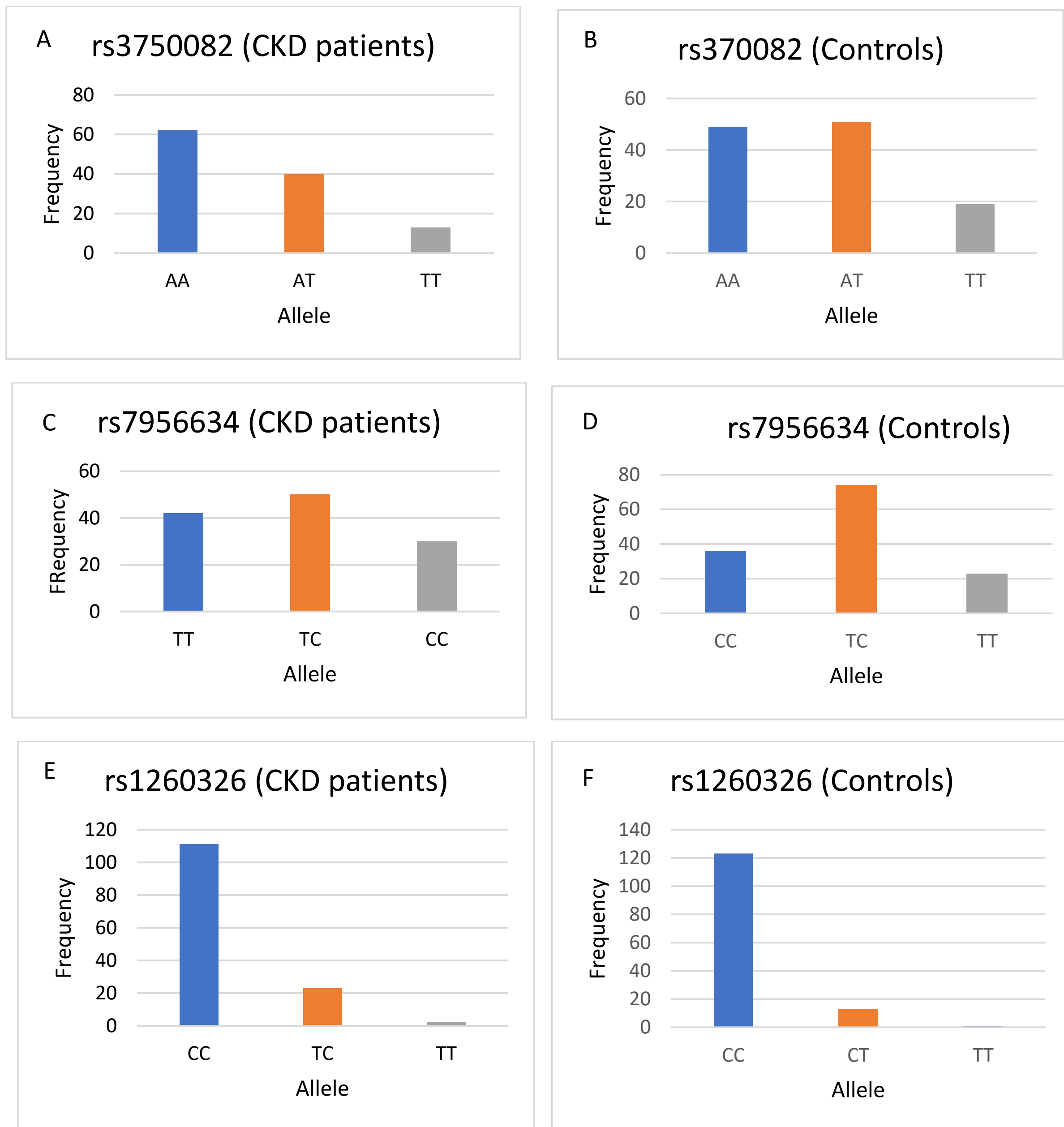


Figure 14: Association between targeted variants and CKD (CKD Cases: panels A, C, E respectively) vs Control (panels B, D, F, respectively). Associated SNPs were *KBTBD2* rs3750082,*PTPRO* rs7956634 and *LPR2* rs466759 4 at p-values of 0.02335, 0.0408 and 0.0398, respectively.

4.9 Functional Pathway Analysis

Following the use of the Linux programme Cytoscape, two functional pathways were identified which are the Advanced glycation end products (AGE or RAGE) signaling pathway in diabetic complication and the regulation of the receptor-mediated endocytosis. The AGEs comprise of a complex group of compounds which are a result from the oxidation and non-enzymatic glycation of nucleic acids, protein, and lipids, because of a pathologic condition such as hyperglycaemia or naturally through aging. The receptors corresponding to the AGE pathway are known as RAGE (Yamagishi 2011). The receptor-mediated endocytosis is a mechanism in which molecules are engulfed into a cell. The mechanism is based on the specificity of the receptor-ligand interaction. Receptors on the membrane of the target tissue will specifically bind to ligands on the exterior of the cell, resulting in an endocytotic process.

Analysis of the protein network pathways using the Gene Ontology (GO), we identified 8 homo-candidate gene clusters (Table 12). The total number of candidate genes that are key protein is 1. The disease-associated key protein is cluster 2 with the NFKB1 gene, P19838 (Figure 15). We consider the cluster as a key genetic cluster which comprises of 4619 genes and key genes in key cluster interacting with the key candidate are 181. The main processes involved in the target sets are: cellular response to mechanical stimulus (GO:0071260), T cell receptor signalling pathway (GO:0050852), NF-kappaB signalling (GO:0038061), positive regulation of NF-KappaB transcription factor activity (GO:0051092) and the Fc-epsilon receptor signalling pathway (GO:0038095). We identified enriched Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway in which key genes are implicated, we managed to identify 22 pathways of interest.

Table 12: Tabulation of the network clusters

Cluster	Genes
Cluster 0	*SCL6A13,ATXN2,DPEP1, A1CF, SCL22A2*
Cluster 1	*DAB2, TMEM60*
Cluster 2	*UNCX, BCAS1, LRP2, WNT7A, SYPL2, NFKB1, VEGFA, UMOD, SKIL, ETV5, AP5BI, ANXA9*
Cluster 3	*NFKB1*
Cluster 4	*TFDP2, KBTBD2*
Cluster 5	*CACNA1S, KCNQ1*
Cluster 6	*SDCCAG8,NFATC1,PTPRO*
Cluster 7	*SIPA1L3*

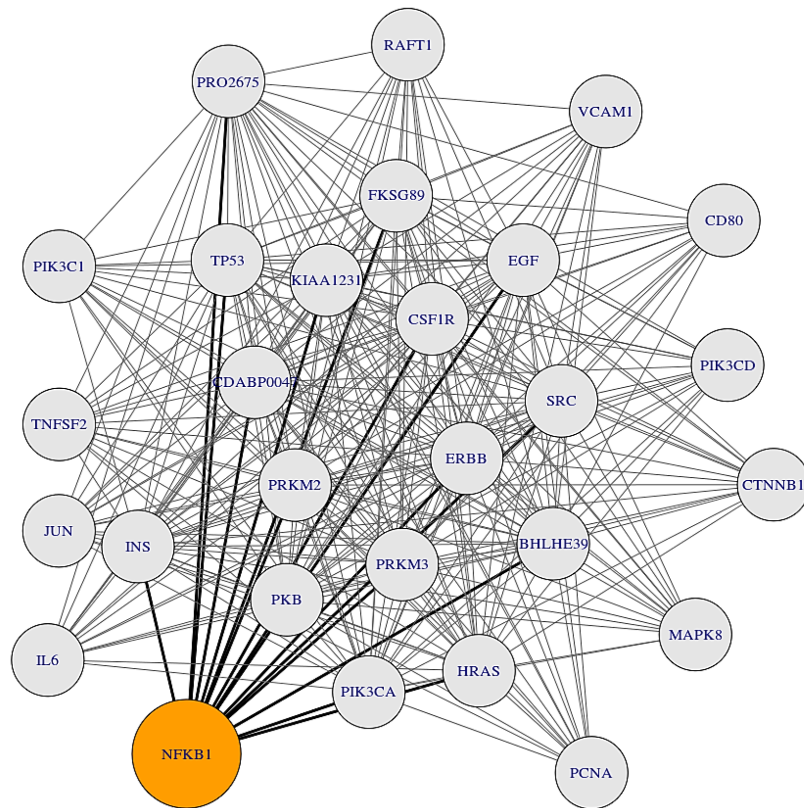


Figure 14: Illustration of the NFKB1 based sub-network. The network clearly indicates the major role of the NFKB1 in the protein-protein sub-network.

4.10 Influence of SCT on the APOL1 and CKD relationship.

APOL1 G1/G2 was previously genotype in the cohort by another investigator (Geard et al. 2017), and the was not significant association with CKD. We investigated whether HbS carrier state can influence the association between *APOL1* and CKD (Table 13): we did not find a significant association (p-value = 0.5725; Odds ratio = 1.433), indicating that there is no evidence that SCT influences the relationship between *APOL1* and CKD.

Table 13: Association of APOL1 G1/G2 and CKD, conditioned to SCT

Gene	P value	ODDs Ratio
APOL1	0.5725	1.433

CHAPTER 5: DISCUSSION

In this chapter, we will be analysing all the results obtained from the study and their potential biological impact in the cohort, moreover within the population.

5.1 Originality

Both SCT and CKD affect a great proportion of individuals within the African population, however, studies on the impact of SCT on CKD have been limited, particularly in Africa, where most carriers live. With our current study, we aimed to elucidate the missing link between the impact of SCT on CKD. Additionally, we investigated 29 SNPs in 29 genes (Table 10), that may affect the development of CKD, in a group of patients from Cameroon.

The nominated loci were selected from a recent study by Pattaro et.al (2016), based on the significant correlation with CKD among multiple populations i.e. African Americans, Asians, and Europeans, and that have also achieved a genome-wide significance. To the best of our knowledge, our study will be the first to look at the identified loci and SCT in a cohort of CKD patients living in Africa. Of particular note, we closely matched the socio-demographic and clinical co-morbidity between patients and controls (Table 5).

5.2 CKD and known Clinical Comorbidities

An expected, association between CKD and hypertension was found (Table 9). Hypertension occurs within 85-95% of patients with CKD (Rao et al. 2008; Botdorf et al. 2011; Segura & Ruilope 2011). Approximately, one billion people worldwide have hypertension and that number is expected to increase to 1.56 billion people by 2025 (Kearney et al. 2005). Hypertension can lead to CKD because of the negative effect that increased BP has on kidney vasculature. Chronic hypertension can result in impairing glomerular filtration and high intraglomerular pressure (Yoshioka et al. 1987; Keane et al. 1999). The damage caused to the glomeruli results in increased protein

filtration, causing abnormally high amounts of protein in the urine known as microalbuminuria (presence of the protein albumin within the urine) and proteinuria (occurs when the protein-to-creatinine ratio is above 200 mg/g and often develops with the progression of CKD) (Yoshioka et al. 1987; Keane et al. 1999).

The study also revealed an expected association of CKD and HIV (Table 9). HIV is a huge burden within Africa, with approximately 24.7 million people within sub-Saharan Africa affected (Joint United Nations Programme on HIV/AIDS 2013; Stanifer et al. 2014; World Kidney Day 2016; GBD 2015 DALYs and HALE Collaborators 2016). Approximately, 3.4 million South Africans living with HIV are receiving Antiretroviral Therapy (ART) (Islam et al. 2012) and the numbers are expected to increase over time. Kidney disease is an acknowledged complication of HIV infection affecting 3.5 – 48.5% of individuals (Fabian & Naicker 2009; Lucas et al. 2014). Kidney diseases have moved from the 27th to the 18th most important global cause of death in the last twenty years (Liyanage et al. 2015). The link between CKD and HIV was first reported in New York in 1984 (Rao et al. 1984). Kidney diseases associated to HIV can be caused by the HIV virus as it enters the kidney or alternatively by the side effects of the medication which can be lactic acidosis, crystals, interstitial nephritis and electrolyte abnormalities (Available at https://www.davita.com/education/kidney-disease/risk-factors/hiv-aids-and-chronic-kidney-disease Accessed 24/08/2019). Risk factors that can lead to CKD are African ancestry, diabetes, CD4 count <200 cells/μL, unsuppressed viral load, family history of CKD, hypertension, and HCV infection (Available at https://www.hiv.va.gov/provider/manual-primary-care/renal-disease.asp Accessed 24/08/2018).

Diabetes was associated with CKD (Table 9). Globally, an estimated 171 million people are affected with diabetes (Wild et al. 2004) with a projected growth reaching up to 642 million people by 2040 (International Diabetes Federation (IDF) et al. 2015). Additionally, an estimated 39 330 000 of individuals with diabetes present with moderate to severe CKD (Coresh et al. 2003; Middleton et al. 2006; American Diabetes & American Diabetes Association 2011). Diabetes has been illustrated as the leading cause of End-Stage Kidney Disease (ESKD) and this trend expected to grow over time especially in developing countries (Villar et al. 2007; US Renal Data System 2013; Liyanage et al. 2015). The development of CKD from diabetes results

from the increased sugar levels within the bloodstream damaging the small blood vessels throughout the body including the small vessels within the kidney (Available at: https://www.davita.com/education/kidney-disease/risk-factors/diabetes-is-the-leading-cause-of-chronic-kidney-disease Accessed 24/08/2018).

In the present study, and as expected, we also found an association between BMI and CKD. Our results correlate with studies from some authors in many different populations (Gelber et al. 2005; Ejerblad et al. 2006; Lu et al. 2014; Herrington et al. 2017). Increasing BMI was also associated with CKD in sickle cell disease in African populations (Asnani et al. 2016; Geard et al. 2017). However, no association between BMI and CKD was found in a study in Taiwan (Chang et al. 2018), and a Caucasian population from USA (Brown et al. 2012). The link between BMI and CKD can be hypothesized to result from obesity and increased weight being predisposing factors for developing various chronic diseases (Pinkowish 1998). Arguably, BMI is not a reliable factor for assessing risk in CKD, since the CKD patients usually suffer from fluid retention and BMI fails to account for it (Navaneethan et al. 2016).

Age, as expected, was also identified as a variable that influences the development of CKD. It has been previously established that eGFR declines the older an individual becomes (Davies & Shock 1950), even in SCD (Asnani et al. 2016; Geard et al. 2017). Within the Chinese population, it was noted that the females aged between 18-39 years have a 7.4% prevalence of CKD, while females aged 60-69 have 18% and females aged 70 have a 24.2% prevalence of CKD (Zhang et al. 2012). Similar results were illustrated in USA (Coresh et al. 2007), Canada and Europe (Zhang & Rothenbacher 2008; Arora et al. 2013). However, the risk of developing CKD when aging may also be due to the normal age-associated decline in kidney function which is not attributed to other unknown factors (Zhou et al. 2008).

5.3 Association of the SCT within the CKD cases and non-CKD controls

The association between SCT and renal dysfunctions has been highly debated with a several publications showing contradicting data (Rakhi P. Naik et al. 2014). Few publications found the association, particularly with regards to hyposthenia and haematuria (Gupta et al. 1991). The present study has revealed a strong association between SCT trait (HbAS individuals) with CKD cases; this is a novel finding on a population living in sub-Saharan Africa, where SCT is highly prevalent and could reach a prevalence of more than 25% in country such Nigeria and DRC (Ezenwosu et al. 2015; Burnham-Marusich et al. 2016).

These findings could have a major implications in public health and clinical practice and in health policy, there are potentially nearly hundreds of millions of individuals with SCT in sub-Saharan Africa, that could benefit from counseling for CKD, including life style preventive measures , to avoid developing exacerbating co-morbidity such HBP, diabetes or HIV. Our findings correlate with a few available literature from African American studies (Key & Derebail 2010; Derebail et al. 2014) and within Latino and Hispanic individuals with a Caribbean background (Dueker et al. 2017; Kramer et al. 2017). However, the SCT and CKD was not found in populations living in the Democratic Republic of the Congo population (Mukendi et al. 2015), that could be attributed to the difference in the size, sociodemographic and with the present Cameroonian cohort. The development of CKD within SCT individuals was hypothesized to be caused by sickled erythrocytes resulting in vaso-occlusion and medullary microinfarctions within the kidneys. Resulting in inhibition of urine concentration whilst leading to proteinuria, glomerulosclerosis and glomerular hyperfiltration (Van Eps et al. 1970; Sears 1978; Hostetter 2003; Kiryluk et al. 2007).

Discovery of association between HbAS and CKD can be used in the therapeutic management of CKD. Individuals already known to suffer from CKD, especially individuals with African ancestry, need to genotype for SCT for appropriate counselling.

5.4 Beyond association with CKD, SCT is not that benign

The present results support other emerging data that indicate that SCT is not a benign condition as previously thought. Multiple disorders have now been associated with SCT, including rhabdomyolysis (Harmon et al. 2012), vaso-occlusive crisis (Tripette, Connes, et al. 2010) and medullary carcinoma (Alvarez et al. 2015). Individual fitness most likely impacts the progression of rhabdomyolysis. Improving exercise conditions, oxidative stress and inflammation may reduce the rapid progression of the disease (Chirico et al. 2012; Aufradet et al. 2010). The acidic environment and decreased venous partial pressure of oxygen (PO_2) that develops within the kidney may be contributors to the pathophysiology of rhabdomyolysis. Debatable, the presence of the sickled erythrocyte may exert a strain to vessels within the striated muscle resulting muscular deterioration (Harmon et al. 2012).Over the duration of 2000- 2010, there has been an increase in African American participants for the National Collegiate Athletic Association (NCAA) college football league. An approximated 22 892 of African American athlete have joined the league within that period. Assuming that the prevalence of SCT is ~7% this can indicate that about 1 602 of the enrolled participants may be in risk of developing rhabdomyolysis (Eichner 2010). These results illustrate the probability of how many individuals may be SCT carriers and advocates for better understanding of the disease, thus allowing the development of guidelines in the detection, treatment and ultimately the prevention of rhabdomyolysis.

Medullary carcinoma has been noted in individuals of all ages but predominantly in younger patients with SCT or SCD at the age below 40 (Davis et al. 1995; Coogan et al. 1998; Warren et al. 1999; Watanabe et al. 2007). The association between SCT and medullary carcinoma is a strong robust association, without much literature to the contrary. Individuals with SCT are highly at risk for the development of medullary carcinoma, thus early prevention through early screening for SCT is essential. There is a great need for a better understanding of the genetics and molecular factors influencing the development of medullary carcinoma, enabling the development of promising therapies.

The mechanism of vaso occlusion (VOC) in SCT patients, is a more complexed process than anticipated and may comprise multiple receptor-ligand pairs expressed on at least three cell types (Frenette 2002) such as leukocytes. The intensity of the VOC depends on how many alpha-globin genes deleted (Steinberg 1986; Piel et al. 2013). In most cases, HbAS take up 40% of the globin within erythrocyte, any exposure to adverse conditions such as dehydration and hypothermic environment can lead to the sickling of erythrocyte, although to a lesser extent than in HbSS (Brittenham et al. 1985; Tripette, Loko, et al. 2010).

5.5 Description of possible SCT and other clinical phenotypes

We also found an association between SCT and a few clinical conditions that should be interpreted with caution, as they could be hugely biased by our cohort selection confounded by multiple co-morbidities. However, the data suggest that these findings could deserve additional investigations, in the general populations. For example, an association between SCT and HIV (Table 8) was found. Nevertheless, several case-reports that indicate that SCD (although not SCT) slows the progression of HIV and hypothesized that this association can elucidate new ideas in HIV treatment (Nouraie et al. 2012; Obaro 2012; Owusu et al. 2015). Several credible explanations were established for this observation, including reduced HIV viral replication because of the insistent up-regulation of inflammation, immunologic changes and iron metabolism common in SCD individuals. Researchers illustrated that the white blood cells within SCD patients have a resistance mechanism against HIV-1 infection because of how the iron is metabolized in SCD patients (Kumari et al. 2016). However, there is no publicly available data on the association between SCT and HIV.

Further analysis of the cohort revealed an association between SCT and hypertension, that could be due to the selection bias of this group of patients that is not representative of the general populations (Table 8). No data was publicly available on the association of SCT and hypertension.

Diabetes was not associated with SCT in the present study (Table 8). Our data correlates with studies by (Bleyer et al. 2010; Ama et al. 2012; Rakhi P Naik et al.

2014) which show evidence that suggest that there is no link between SCT and DM, but contradicts studies by (Ajayi & Kolawole 2004; Goldsmith et al. 2012) were SCT patients illustrated higher risk for diabetic complications when compared to those with normal haemoglobin. A recent 2018 study indicated SCT as a possible risk for the development of T2D-related complications. Their cohort numbers, particularly the individuals which are SCT carriers, were greater than our cohort, which might explain the contradicting data (Skinner et al. 2018).

5.6 Association of CKD and targeted genomic variants

This is the first study to explore multiple SNPs associated with CKD and consistently replicated in many world populations, in an African Cohort. We found that five SNPs (*ANXA9 rs267734, BCAS1 rs17216707,KCQN1 rs163160,CACNA1S rs3850625, ETV5 rs10513801,* and *ATXN2 rs653178*) were monomorphic, indicating that only one allele occurs at the locus in the population (Elston et al. 2017). This is an important finding as it indicates that African Americans data, used in the previous study and that shown association with genome-wide significance with these SNPs, are not proxy for African living in Africa. Indeed, African ancestry in self-declared African American can range from 1- 99% with an average of 70% (Campbell & Tishkoff 2008; Gomez et al. 2014). Moreover, up to 10 SNPs also show a significant difference between our control group and African data extracted from the 1000 genome project, demonstrating the highest genetic diversity found in Africans living in Africa.

The data supports and emphasizes as much as possible using controls from the same population background, because in the same African country even regional difference might be massive (Lambert & Tishkoff 2009). In addition, that major difference in MAF when comparing these SNPs with data from European (75.8%) and Asians (83%), support that risk model developed for CKD, will be largely uninformative in populations of African ancestry. It is also a call to improve diversity in GWAS studies globally, which are still overpowered by studies in a population of African ancestry (Peprah et al. 2015), and in Africa in particular, thus improving and attaining a polygenic risk score that can be universally used (Vega & Bustamante 2018).

Surprisingly, only 3/29 SNPs Gene- *KBTBD2-rs3750082,PTPRO-rs7956634* and *LRP2-*r*s4667594* were associated with CKD at a p-values of 0.02335, 0.0408 and 0.0398, respectively (Table 10). These results support the need to perform GWAS studies for CKD in populations living in Africa, indicating that there is a potential to discover novel loci not found in other world populations. SNP *rs3750082* is in the Kelch Repeat and BTB Domain Containing 2 (KBTBD2) gene, SNP *rs7956634* is found within the Protein Tyrosine Phosphatase Receptor Type O (PTPRO) gene while SNP r*s4667594* is in the low-density lipoprotein-related protein 2 (LRP2) gene. All three SNPs are intronic variants associated with glomerular filtration rate within the kidneys, eluding to their potential to be pathogenic especially if they occur on the intron-exon junction or regulator site of the gene. These SNPs have only been cited once in a study conducted by Pattaro et.al 2016 and colleagues where they indicated *KBTBD2 rs3750082* (p-value $2.52X10^{-7}$)*, PTPRO rs7956634* (p-value $2.46X10^{-9}$) and *LRP2 rs4667594* (p-value $2.37X10^{-7}$) as novel SNPs associated with eGFR within individuals of European ancestry. Only SNP *PTPRO rs7956634* was associated with other kidney-related traits within the African population (Pattaro et al. 2016).

Two SNPs *DAB2 rs11959928* and *UNCX rs10277115* illustrated a borderline association with CKD with p values of 0.08937 and 0.06895. SNP *DAB2 rs11959928* is an intergenic variant located in the Clathrin Adaptor Protein, where it encodes for the mitogen-responsive phosphoprotein. SNP *UNCX rs10277115* is a regulatory region variant in the UNX homeobox that plays in maintenance and differential of elements of the axial skeleton.

5.7 Functional Pathway Analysis

When computing the association of the variants investigated with relevant biological pathways, two pathways were identified when using the programme Cytoscape the Advanced glycation end products (AGE or RAGE), that are signalling pathway in diabetic complication and the regulation of the receptor-mediated endocytosis. The AGE/RAGE signalling prompts the activation of various intracellular signal pathways such as protein kinase C, Nicotinamide Adenine Dinucleotide Phosphate (NADPH)

oxidase and Mitogen-Activated Protein Kinase (MAPKs) which results in NF-kappaβ activity. Additionally, the RAGE pathway induces the Janus Kinase/Signal Transducer and Activator of Transcription (JAK-STAT) mediated, hypoxia-mediated induction and the PI3K-Atk dependent pathways which in turn contribute to apoptosis and cell proliferation. The increased destruction of erythrocyte will inevitably result in increased intracellular material and sheared erythrocyte which can affect the function of kidneys resulting in kidney dysfunction (Yamagishi 2011).

The receptor-mediated endocytosis plays a crucial role in the proximal tubule epithelial cells (PTECs), with regards to the reabsorption and metabolism of protein, carrier-bound vitamins, conservation of nutrients and various substances in the glomerular filtrates. Disruption of this process can result in the loss of essential substances and development of proteinuria, which is a hallmark of kidney dysfunction.

Analysis of the protein network pathways illustrated 8 homo-candidate gene clusters. The disease-associated key protein is cluster 2 with the NFKB1 gene, P19838. The Nuclear Factor Kappa β Subunit 1 (NFKB1) gene has been identified as an important by-product of the functional pathway AGE or RAGE signalling pathway in diabetic complication.

5.8 Influence of SCT on the APOL1 and CKD relationship

In this cohort, a separate investigator did not previously find any association between APOL1 and CDK (Geard et al, 2017). We investigated whether the result might be different if conditioned with the significantly influential HbAS. However, we did not find any significant association. This could be due to the small sample size, under power by the low frequency of APOL1 G1/G2 alleles (14%) in the studied population, as we previously postulated (Geard et al, 2017).

5.9 Practical implications and research recommendation

Most individuals that have SCT are asymptomatic and in most cases, they don’t even know their carrier status. The global burden of SCT, already very high in Africa is projected to increase, since most SCT individuals have a normal life expectancy and the presence of the SCT is advantageous in malaria-endemic areas particularly in

African countries, and also due to the global migration. SCT prevalence is also influenced by the increased in the life expectancy of the SCD patients who are now living to reproductive age in affluent countries such as USA due to improvement in SCD treatments and care. The study illustrated a clear association between SCT and CKD, that will have major implication in preventative medicine policies and practices in both sub-Saharan African and globally. Screening and identification of individuals with SCT will therefore be more actionable than previously argue, as it can help in the modification of lifestyle and, in some case, treatment and therapies offered to already symptomatic patients.

Our study highlighted unexpected associations between SCT with HIV and hypertension. Despite the selection bias and our limited sample size, the data indicate the need for further research in the general population, in investigating how these disorders influence each other, allowing the mechanism to be fully understood..

In addition, we identified a relatively low proportion of (3/29) SNPs previously associated with CKD in GWAS studies among African American, Europeans, and Asians, indicating the urgent need to include diverse populations, specifically those in sun-Saharan African, so that the Polygenic risk score generated from GWAS data could be universally relevant. The data further indicate that there is potential to discovers new loci associated with CKD when investigating populations of African ancestry living in Africa.

5.10 Limitation

The recruitment was hospital-based, this could have allowed for bias enrolment of sicker patients which may influence the outcome of our studies. Future studies should aim to recruit more clinically stable individuals.

5.11 Conclusion and perspectives

Our study illustrates a strong association between SCT and CKD, an important discovery that will have a major implications in medicine policies and practices in sub-Saharan African where there is a high prevalence of SCT. The data also has a global resonance, with the projected increase in the prevalence of individuals with SCT, due to migration and the improve life expectancy and genetic fitness of people living with both SCT and SCD.

We identified a relatively low proportion of (3/29) of target SNPs positively associated with CKD among this group of Cameroonians. The study illustrates that the vast majority of targeted SNPs associated with CKD in GWAS studies in multiple populations including African American, Europeans, and Asians, are not relevant for sub-Saharan Africans, indicating the urgent need to include diverse populations, specifically those living in Africa, so that the Polygenic risk scores generated from GWAS data could be universally applicable. The data further indicate that there is potential to discover new loci associated with CKD when investigating populations of African ancestry living in Africa.

REFERENCES

Aidoo, M. et al., 2002. Protective effects of the sickle cell gene against malaria morbidity and mortality For personal use . Only reproduce with permission from The Lancet Publishing Group . *Lancet*, 359, pp.1311–1312.

Ajayi, A.A. & Kolawole, B.A., 2004. Sickle cell trait and gender influence type 2 diabetic complications in African patients. *European Journal of Internal Medicine 15*, 15, pp.312–315.

Al-Naama, L.M., Al-Sadoon, E.A. & Al-Sadoon, T.A., 2000. Levels of Uric Acid, Urea and Creatinine in Iraqi Children with Sickle Cell Disease. *Journal of the Pakistan Medical Association*, 50(3), pp.98–102.

Allison, A., 1964. Polymorphism and Natural Selection in Human Populations. *Cold Spring Harb Symp Quant Biol*, 29, pp.137–149.

Allison, A.C., 1954. PROTECTION AFFORDED BY SICKLE- CELL TRAIT AGAINST SUBTERTIAN. *British Medical Journal*, 1(4857), pp.290–294.

Aloni, M.N. et al., 2017. Congolese children with sickle cell trait may exhibit glomerular hyperfiltration: A case control study. *Journal of Clinical Laboratory Analysis*, (August 2016), pp.1–7. Available at: http://doi.wiley.com/10.1002/jcla.22143.

Alvarez, O., Rodriguez, M.M. & Jordan, L., 2015. Renal Medullary Carcinoma and Sickle Cell Trait : A Systematic Review. , (May 2014), pp.1694–1699.

Ama, V. et al., 2012. Short Report : Care Delivery Would sickle cell trait influence the metabolic control in sub-Saharan individuals with type 2 diabetes ? *DIABETICMedicine*, 29, pp.334–337.

American Diabetes, A. & American Diabetes Association, 2011. Standards of Medical Care in Diabetes--2014. *Diabetes Care*, 37(Supplement_1), pp.S11-61. Available at: http://care.diabetesjournals.org/cgi/doi/10.2337/dc14-S014.

Arora, P. et al., 2013. Prevalence estimates of chronic kidney disease in Canada: results of a nationally representative survey. *Canadian Medical Association.Journal*, 185(9), pp.E417--23. Available at: http://sfx.usask.ca/usask?url%7B_%7Dver=Z39.88-2004%7B&%7Drft%7B_%7Dval%7B_%7Dfmt=info:ofi/fmt:kev:mtx:journal%7B&%7Dgenre=article%7B&%7Dsid=ProQ:ProQ:nursing%7B&%7Datitle=Prevalence+estimates+of+chronic+kidney+disease+in+Canada:+results+of+a+nationally+r.

Ashley-Koch, A.E. et al., 2011. MYH9 and APOL1 are both associated with sickle cell disease nephropathy. *British Journal of Haematology*, 155(3), pp.386–394.

Asnani, M.R. et al., 2016. Interventions for patients and caregivers to improve knowledge of sickle cell disease and recognition of its related complications. *Cochrane Database of Systematic Reviews*.

Aufradet, E. et al., 2010. Habitual Physical Activity and Endothelial Activation in Sickle Cell Trait Carriers. *Medicine and Science in Sports and Exercise*, 42(11),

pp.1987–1994.

Baffour, S.A. et al., 2015. Prevalence of hemoglobin S trait among blood donors : a cross - sectional study. *BMC Research Notes*, 8(583), pp.4–9.

Bain, B.J., 2009. Neonatal/newborn haemoglobinopathy screening in Europe and Africa. *Journal of Clinical Pathology*, 62(1), pp.53–56.

Bank, A., 2006. Regulation of human fetal hemoglobin: New players, new complexities. *Blood*, 107(2), pp.435–443.

Bartolucci, P. & Galacteros, F., 2012. Clinical management of adult sickle-cell disease. *Curr Opin Hematol*, 19(1), pp.149–155.

Bean, C.J. et al., 2012. Heme oxygenase-1 gene promoter polymorphism is associated with reduced incidence of acute chest syndrome among children with sickle cell disease. *Blood*.

Behar, D.M. et al., 2011. Absence of APOL1 Risk Variants Protects against HIV-Associated Nephropathy in the Ethiopian Population. *American Journal of Nephrology*.

Behar, D.M. et al., 2010. African ancestry allelic variation at the MYH9 gene contributes to increased susceptibility to non-diabetic end-stage kidney disease in Hispanic Americans. *Human Molecular Genetics*.

Bleyer, A.J. et al., 2010. Sickle Cell Trait and Development of Microvascular Complications in Diabetes Mellitus. , pp.1015–1020.

Botdorf, J., Chaudhary, K. & Whaley-Connell, A., 2011. Hypertension in Cardiovascular and Kidney Disease. *Cardiorenal Medicine*, 1(3), pp.183–192. Available at: https://www.karger.com/Article/FullText/329927.

Brittenham, G.M., Schechter, A.N. & Constance, T.N., 1985. Hemoglobin Hemolytic. *Bloodjournal*, 65(1), pp.183–189.

Brousseau, D. et al., 2007. Treatment for Pediatric Sickle Cell Pain Crisis. *The Journal of Pediatrics*, 150(6), pp.623–626.

Brown, R.N.K.L. et al., 2012. Body mass index has no effect on rate of progression of chronic kidney disease in non-diabetic subjects. *Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association*, 27(7), pp.2776–80. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22442391.

Bruce, M.A. et al., 2009. Social Environmental Stressors, Psychological Factors, and Kidney Disease. *Journal of Investigative Medicine*.

Burnham-Marusich, A.R. et al., 2016. Prevalence of sickle cell trait and reliability of self-reported status among expectant parents in Nigeria: Implications for targeted newborn screening. *Public Health Genomics*.

Campbell, M.C. & Tishkoff, S.A., 2008. African Genetic Diversity: Implications for Human Demographic History, Modern Human Origins, and Complex Disease Mapping. *Annual Review of Genomics and Human Genetics*.

Cavanaugh, K.L. & Lanzkron, S., 2010. Time to Recognize an Overlooked Trait. *Journal of the American Society of Nephrology*, 21(3), pp.385–386. Available at: http://www.jasn.org/cgi/doi/10.1681/ASN.2010010070.

Chang, T.-J. et al., 2018. Relationship between body mass index and renal function deterioration among the Taiwanese chronic kidney disease population. *Scientific Reports*, 8(1).

Charache, B.S. et al., 1992. Hydroxyurea: Effects on Hemoglobin F Production in Patients With Sickle Cell Anemia. *Bloodjournal*, 79(1), pp.2555–2565.

Chirico, E.N. et al., 2012. Exercise training blunts oxidative stress in sickle cell trait carriers. , (33), pp.1445–1453.

Colares, V.S. et al., 2014. MYH9 and APOL1 gene polymorphisms and the risk of CKD in patients with lupus nephritis from an admixture population. *PLoS ONE*.

Connes, P., Morrison, E. & Hue, O., 2008. Physiological Responses of Sickle Cell Trait Carriers during Exercise. *Sports Medicine*, 38(January 2017), pp.931–946.

Coogan, C.L. et al., 1998. Renal medullary carcinoma in patients with sickle cell trait. *Urology*, 51(6), pp.1049–1050.

Coresh, J. et al., 2003. Prevalence of chronic kidney disease and decreased kidney function in the adult US population: Third National Health and Nutrition Examination Survey. *American Journal of Kidney Diseases: The Official Journal of the National Kidney Foundation*, 41(1), pp.1–12. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12500213%5Cnhttp://www.uptodate.com/contents/assessment-of-kidney-function/abstract/2.

Coresh, J. et al., 2007. Prevalence of chronic kidney disease in the United States. *JAMA : the journal of the American Medical Association*, 298(17), pp.2038–47. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17986697.

Davies, D.F. & Shock, N.W., 1950. Age changes in glomerular filtration rate, effective renal plasma flow, and tubular excretory capacity in adult males. *The Journal of clinical investigation*, 29(5), pp.496–507.

Davis, C.J., Mostofi, F.K. & Sesterhenn, I.A., 1995. Renal Medullary Carcinoma. *The American Journal of Surgical Pathology*, 19(1), pp.1–11.

Derebail, V.K. et al., 2010. High Prevalence of Sickle Cell Trait in African Americans with ESRD. *J Am Soc Nephrol Mar;*, 21(3), pp.413–417. Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2831865/?report=printable.

Derebail, V.K. et al., 2014. Sickle Trait in African-American Hemodialysis Patients and Higher Erythropoiesis-Stimulating Agent Dose. *J Am Soc Nephrol*, 25, pp.819–826.

Dueker, N.D. et al., 2017. Sickle cell trait and renal function in HISPANICS IN THE UNITED STATES: THE NORTHERN MANHATTAN STUDY. , 27(August), pp.11–14.

Ebihara, A., 2000. World medical association declaration of Helsinki. *Japanese Pharmacology and Therapeutics*.

Eichner, E.R., 2010. Sickle Cell Trait in Sports. *Current Sports Medicine Reports*, 9(6), pp.347–351.

Ejerblad, E. et al., 2006. Obesity and risk for chronic renal failure. *Journal of the American Society of Nephrology : JASN*, 17(6), pp.1695–702. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16641153.

El-Shal, A.S., Zidan, H.E. & Rashad, N.M., 2014. Adiponectin gene polymorphisms in Egyptian type 2 diabetes mellitus patients with and without diabetic nephropathy. *Molecular Biology Reports*.

Elguero, E. et al., 2015. Malaria continues to select for sickle cell trait in Central Africa. *PNAS*, 112(22), pp.7051–7054.

Ellis, J.W. et al., 2012. Validated SNPs for eGFR and their associations with albuminuria. *Human Molecular Genetics*.

Elston, R.C., Satagopan, J. & Sun, S., 2017. Statistical genetic terminology. In *Methods in Molecular Biology*. pp. 1–9.

Van Eps, L.W.S. et al., 1970. NATURE OF CONCENTRATING DEFECT IN SICKLE-CELL NEPHROPATHY. Microradioangiographic Studies. *The Lancet*, 295(7644), pp.450–452.

Estrella, M.M. et al., 2015. The association between APOL1 risk alleles and longitudinal kidney function differs by HIV viral suppression status. *Clinical Infectious Diseases*.

Ezenwosu, O.U. et al., 2015. Knowledge and awareness of personal sickle cell genotype among parents of children with sickle cell disease in southeast Nigeria. *Journal of Community Genetics*.

Fabian, J. & Naicker, S., 2009. HIV and kidney disease in sub-Saharan Africa. *Nature Reviews Nephrology*, 5(10), pp.591–598.

Fine, D.M. et al., 2012. APOL1 Risk Variants Predict Histopathology and Progression to ESRD in HIV-Related Kidney Disease. *Journal of the American Society of Nephrology*.

Foster, M.C. et al., 2013. APOL1 Variants Associate with Increased Risk of CKD among African Americans. *Journal of the American Society of Nephrology*.

Freedman, B.I., 1999. Familial aggregation of end-stage renal failure: Aetiological implications. *Nephrology Dialysis Transplantation*.

Frenette, P.S., 2002. Sickle cell vaso-occlusion: multistep and multicellular paradigm. *Current opinion in hematology*, 9, pp.101–106.

Friedman, D.J. et al., 2011. Population-Based Risk Assessment of APOL1 on Renal Disease. *Journal of the American Society of Nephrology*.

Gandevia, B. & Tovell, A., 1964. DECLARATION OF HELSINKI. *The Medical journal of Australia*.

GBD 2015 DALYs and HALE Collaborators, 2016. Global, regional, and national disability-adjusted life-years (DALYs) for 315 diseases and injuries and healthy

life expectancy (HALE), 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. *The Lancet VO - 388*, (10053), p.1603. Available at: https://ezp.lib.unimelb.edu.au/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edsgao&AN=edsgcl.470763692&site=eds-live&scope=site.

Geard, A. et al., 2017. Clinical and genetic predictors of renal dysfunctions in sickle cell anaemia in Cameroon. , (May).

Gelber, R.P. et al., 2005. Association between body mass index and CKD in apparently healthy men. *American journal of kidney diseases : the official journal of the National Kidney Foundation*, 46(5), pp.871–80. Available at: http://www.ncbi.nlm.nih.gov/pubmed/16253727.

Genovese, G. et al., 2010. Association of trypanolytic ApoL1 variants with kidney disease in African Americans. *Science*.

Genovese, G., Friedman, D.J. & Pollak, M.R., 2013. APOL1 variants and kidney disease in people of recent African ancestry. *Nature Reviews Nephrology*.

Goldsmith, J.C. et al., 2012. Framing the research agenda for sickle cell trait: building on the current understanding of clinical events and their potential implications. *Am J Hematol*, 87(3), pp.340–346.

Gomez, F., Hirbo, J. & Tishkoff, S.A., 2014. Genetic variation and adaptation in Africa: Implications for human evolution and disease. *Cold Spring Harbor Perspectives in Biology*.

Gong, L. et al., 2012. Evidence for both innate and acquired mechanisms of protection from Plasmodium falciparum in children with sickle cell trait. *Bloodjournal*, 119(16), pp.3808–3815.

Guasch, A. et al., 2006. Glomerular involvement in adults with sickle cell hemoglobinopathies: Prevalence and clinical correlates of progressive renal failure. *Journal of the American Society of Nephrology : JASN*, 17(8), pp.2228–35. Available at: http://www.jasn.org/cgi/doi/10.1681/ASN.2002010084%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/16837635.

Gupta, A.K. et al., 1991. Effects of a-Thalassemia and Sickle Polymerization Tendency on the Urine-concentrating Defect of Individuals with Sickle Cell Trait. *j Clin. Invest*, 88, pp.1963–1968.

Hardison, R.C., 2012. Evolution of hemoglobin and its genes. *Cold Spring Harbor Perspectives in Medicine*, 2(12), pp.1–18.

Harmon et al., 2012. Sickle cell trait associated with a RR of death of 37 times in national collegiate athletic association football athletes : a database with 2 million athlete-years as the denominator Correspondence to. *Br J Sports Med*, 46, pp.325–330.

Harmon, K.G., Drezner, J.A. & Casa, D.J., 2012. To screen or not to screen for sickle cell trait in American football? *British Journal of Sports Medicine*, 46(3), p.158. Available at: http://bjsm.bmj.com/cgi/doi/10.1136/bjsports-2012-090999%5Cnpapers2://publication/doi/10.1136/bjsports-2012-090999.

Haslam, D., 2007. Obesity: A medical history. *Obesity Reviews*.

Heller, P. et al., 1979. Clinical implications of sickle-cell trait and glucose-6-phosphate dehydrogenase deficiency in hospitalized black male patients. *The New England journal of medicine*, 300(18), pp.1001–5. Available at: http://www.ncbi.nlm.nih.gov/pubmed/431593.

Herrick, J.B., 1910. Peculiar elongated and sickle-shaped red blood corpuscles in a case of severe anemia. *Archives of Internal Medicine*, VI(5), pp.517–521.

Herrington, W.G. et al., 2017. Body-mass index and risk of advanced chronic kidney disease: Prospective analyses from a primary care cohort of 1.4 million adults in England. *PLoS ONE*, 12(3).

Hicks, P.J., Langefeld, C.D., et al., 2011. Sickle cell trait is not independently associated with susceptibility to end-stage renal disease in African Americans. *Kidney International*, 80(12), pp.1339–1343. Available at: http://dx.doi.org/10.1038/ki.2011.286.

Hicks, P.J., Langefeld, C.D., et al., 2011. susceptibility to end-stage renal disease in African Americans. *Kidney Int*, 80(12), pp.1339–1343.

Hill, O.T. et al., 2016. Sickle Cell Trait, Rhabdomyolysis, and Mortality among U.S. Army Soldiers. , pp.435–442.

Hostetter, T.H., 2003. Hyperfiltration and glomerulosclerosis. *Seminars in Nephrology*, 23(2), pp.194–199.

Ingram, V.M., 1959. Abnormal human haemoglobins. III the chemical difference between normal and sickle cell haemoglobins. *Biochimica et Biophysica Acta*, 36(2), pp.402–411. Available at: http://dx.doi.org/10.1016/0006-3002(59)90183-0.

International Diabetes Federation (IDF) et al., 2015. *IDF Diabetes Atlas 7th edition*, Available at: http://www.diabetesatlas.org/.

Isakova, T. et al., 2011. Fibroblast growth factor 23 and risks of mortality and end-stage renal disease in patients with chronic kidney disease. *JAMA - Journal of the American Medical Association*.

Islam, F.M. et al., 2012. Relative risk of renal disease among people living with HIV: a systematic review and meta-analysis. *BMC public health*, 12, p.234.

Jha, V. et al., 2013. Chronic kidney disease: Global dimension and perspectives. *The Lancet*.

Joint United Nations Programme on HIV/AIDS, 2013. *Global Report|UNAIDS report on the global AIDS epidemic 2013*, Available at: http://www.unaids.org/en/resources/campaigns/globalreport2013.

Kao, W.H.L. et al., 2008. MYH9 is associated with nondiabetic end-stage renal disease in African Americans. *Nature Genetics*.

Kasembeli, A.N. et al., 2015. APOL1 Risk Variants Are Strongly Associated with HIV-Associated Nephropathy in Black South Africans. *Journal of the American*

Society of Nephrology.

Keane, W.F., Eknoyan, G. & Willis, K., 1999. Proteinuria, albuminuria, risk, assessment, detection, elimination (PARADE): A position paper of the National Kidney Foundation. *American Journal of Kidney Diseases*, 33(5), pp.1004–1010.

Kearney, P.M. et al., 2005. Global burden of hypertension: Analysis of worldwide data. *Lancet*, 365(9455), pp.217–223.

Kerkeni, M. et al., 2009. Endothelial nitric oxide synthetase, methylenetetrahydrofolate reductase polymorphisms, and cardiovascular complications in Tunisian patients with nondiabetic renal disease. *Clinical Biochemistry*.

Key, N.S. & Derebail, V.K., 2010. Sickle-Cell Trait: Novel Clinical Significance. *Hematology*, 2010(1), pp.418–422. Available at: http://www.asheducationbook.org/cgi/doi/10.1182/asheducation-2010.1.418.

Kiberd, B.A. & Clase, C.M., 2002. Cumulative risk for developing end-stage renal disease in the US population. *Journal of the American Society of Nephrology : JASN*, 13(6), pp.1635–1644. Available at: papers://5aecfcca-9729-4def-92fe-c46e5cd7cc81/Paper/p90424.

Kiryluk, K. et al., 2007. Sickle cell trait and gross hematuria. *Kidney International*, 71(7), pp.706–710. Available at: http://dx.doi.org/10.1038/sj.ki.5002060.

Kramer, H.J. et al., 2017. African Ancestry-Specific Alleles and Kidney Disease Risk in Hispanics/Latinos. *Journal of the American Society of Nephrology : JASN*, 28(3), pp.915–922. Available at: http://www.ncbi.nlm.nih.gov/pubmed/27650483%0Ahttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5328161.

Kruzel-Davila, E. et al., 2016. APOL1 nephropathy: From gene to mechanisms of kidney injury. *Nephrology Dialysis Transplantation*.

Kumari, N. et al., 2016. Increased iron export by ferroportin induces restriction of HIV-1 infection in sickle cell disease. *Blood advances*.

Lahrach, H. et al., 2014. Association of apolipoprotein E gene polymorphism with end-stage renal disease and hyperlipidemia in patients on long-term hemodialysis. *Renal Failure*.

Lambert, C.A. & Tishkoff, S.A., 2009. Genetic structure in African populations: Implications for human demographic history. In *Cold Spring Harbor Symposia on Quantitative Biology*.

Limou, S. et al., 2014. APOL1 Kidney Risk Alleles: Population Genetics and Disease Associations. *Advances in Chronic Kidney Disease*.

Liyanage, T. et al., 2015. Worldwide access to treatment for end-stage kidney disease: A systematic review. *The Lancet*, 385(9981), pp.1975–1982.

Lote, C.J., 2013. *Principles of renal physiology*,

Lozano, R. et al., 2010. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010 : a systematic analysis for the Global Burden of Disease Study 2010. *Lancet*, 380, pp.2095–2128.

Lu, J.L. et al., 2014. Association of Body Mass Index with Outcomes in Patients with CKD. *Journal of the American Society of Nephrology*, 25(9), pp.2088–2096. Available at: http://www.jasn.org/cgi/doi/10.1681/ASN.2013070754.

Lucas, G.M. et al., 2014. Clinical practice guideline for the management of chronic kidney disease in patients infected with HIV. *Clin Infect Dis*, 59(9), pp.e96--138.

Matsha, T.E. et al., 2012. Polymorphisms in the Non-Muscle Myosin Heavy Chain Gene (MYH9) Are Associated with Lower Glomerular Filtration Rate in Mixed Ancestry Diabetic Subjects from South Africa. *PLoS ONE*.

Mehta, R.L. et al., 2015. International Society of Nephrology's 0by25 initiative for acute kidney injury (zero preventable deaths by 2025): A human rights case for nephrology. *The Lancet*.

Middleton, R.J. et al., 2006. The unrecognized prevalence of chronic kidney disease in diabetes. *Nephrology, dialysis, transplantation : official publication of the European Dialysis and Transplant Association - European Renal Association*, 21(1), pp.88–92.

Mitchell, B.L., 2007. Sickle Cell Trait and Sudden Death- Bringing it Home. *JOURNAL OF THE NATIONAL MEDICAL ASSOCIATION*, 99(3), pp.300–305.

Modell, B. et al., 2007. Epidemiology of haemoglobin disorders in Europe: An overview. *Scandinavian Journal of Clinical and Laboratory Investigation*, 67(1), pp.39–69.

Mu, J. et al., 2010. Plasmodium falciparum genome-wide scans for positive selection, recombination hot spots and resistance to antimalarial drugs. *Nature Genetics*.

Mukendi, K. et al., 2015. Sickle cell trait is not associated with chronic kidney disease in adult Congolese patients : a clinic-based , cross-sectional study. , 26(3), pp.125–129.

Naicker, S., 2003. End-stage renal disease in sub-Saharan and South Africa. *Kidney International*.

Naik, R.P. et al., 2014. Association of sickle cell trait with chronic kidney disease and Albuminuria in African Americans. *JAMA - Journal of the American Medical Association*.

Naik, R.P. et al., 2014. Association of Sickle Cell Trait With Chronic Kidney Disease and Albuminuria in African Americans. *JAMA*, 21287(20), pp.2115–2125.

Naik, R.P. et al., 2015. Association of Sickle Cell Trait With Chronic Kidney Disease and Albuminuria in African Americans. *JAMA*, 312(20), pp.2115–2125.

Navaneethan, S.D. et al., 2016. Body mass index and causes of death in chronic kidney disease. *Kidney International*, 89(3), pp.675–682.

Ndeezi, G. et al., 2016. Burden of sickle cell trait and disease in the Uganda Sickle Surveillance Study (US3): a cross-sectional study. *The Lancet Global Health*, 4(3), pp.e195–e200. Available at: http://dx.doi.org/10.1016/S2214-109X(15)00288-0.

Nouraie, M., Nekhai, S. & Gordeuk, V.R., 2012. Sickle cell disease is associated with decreased HIV but higher HBV and HCV comorbidities in US hospital discharge records: A cross-sectional study. *Sexually Transmitted Infections*, 88(7), pp.528–533.

O'Seaghdha, C.M. et al., 2011. The MYH9/APOL1 region and chronic kidney disease in European-Americans. *Human Molecular Genetics*.

Obaro, S., 2012. Does sickle cell disease protect against HIV/AIDS? *Sexually Transmitted Infections*, 88(7), p.533.

Owusu, E.D.A. et al., 2015. The Interaction Between Sickle Cell Disease and HIV Infection: A Systematic Review. *Clinical Infectious Diseases*, 60(4), pp.612–626. Available at: https://academic.oup.com/cid/article-lookup/doi/10.1093/cid/ciu832.

Özden, Ç. et al., 2011. Where on earth is everybody? The evolution of global bilateral migration 1960-2000. *World Bank Economic Review*, 25(1), pp.12–56.

Parsa, A. et al., 2013. APOL1 Risk Variants, Race, and Progression of Chronic Kidney Disease . *New England Journal of Medicine*.

Pattaro, C. et al., 2016. Genetic associations at 53 loci highlight cell types and biological pathways relevant for kidney function. *Nature communications*, 7, p.10023. Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4735748&tool=pmcentrez&rendertype=abstract.

Pattaro, C. et al., 2012. Genome-wide association and functional follow-up reveals new loci for kidney function. *PLoS Genetics*.

Pauling, L. et al., 1949. Sickle Cell Anemia, a Molecular Disease. *Science*, 110(2865), pp.543–548. Available at: http://www.sciencemag.org/content/110/2865/543.

Peprah, E. et al., 2015. Genome-wide association studies in Africans and African Americans: Expanding the framework of the genomics of human traits and disease. *Public Health Genomics*.

Peralta, C.A. et al., 2010. The association of african ancestry and elevated creatinine in the coronary artery risk development in young adults (CARDIA) study. *American Journal of Nephrology*, 31(3), pp.202–208.

Piel, F.B. et al., 2013. Global Burden of Sickle Cell Anaemia in Children under Five, 2010-2050: Modelling Based on Demographics, Excess Mortality, and Interventions. *PLoS Medicine*, 10(7), pp.1–14.

Piel, F.B. et al., 2014. Global migration and the changing distribution of sickle haemoglobin: A quantitative study of temporal trends between 1960 and 2000. *The Lancet Global Health*, 2(2).

Pinkowish, M., 1998. Obesity -- a chronic disease. *Patient Care for the Nurse Practitioner*, 1(8), pp.33–43. Available at: http://search.ebscohost.com/login.aspx?direct=true&db=cin20&AN=107302406&site=ehost-live.

Purcell, S. et al., 2007. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *The American Journal of Human Genetics*, 81(3), pp.559–575. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0002929707613524.

Radwan, W.M., Elbarbary, H.S. & Alsheikh, N.M., 2015. DNA repair genes XPD and XRCC1 polymorphisms and risk of end-stage renal disease in Egyptian population. *Renal Failure*, 37(1), pp.122–128.

Rao, M. V. et al., 2008. Hypertension and CKD: Kidney Early Evaluation Program (KEEP) and National Health and Nutrition Examination Survey (NHANES), 1999-2004. *American Journal of Kidney Diseases*, 51(4 SUPPL. 2), pp.1999–2004.

Rao, T.K. et al., 1984. Associated focal and segmental glomerulosclerosis in the acquired immunodeficiency syndrome. *The New England journal of medicine*, 310(11), pp.669–673.

Rayner, H., Milford, D. & Thomas, M., 2016. *Understanding kidney diseases*,

Rezk, N.A. et al., 2014. Renalase Gene Polymorphism and Epinephrine Level in Chronic Kidney Disease. *Applied Biochemistry and Biotechnology*.

Rosenberg, N.A. et al., 2010. Genome-wide association studies in diverse populations. *Nature Reviews Genetics*.

Salimi, S. et al., 2006. Endothelial nitric oxide synthase gene intron4 VNTR polymorphism in patients with coronary artery disease in Iran. *Indian Journal of Medical Research*.

Saraf, S.L. et al., 2015. Genetic variants and cell-free hemoglobin processing in sickle cell nephropathy. *Haematologica*.

Saraf, S.L. et al., 2014. Haemoglobinuria is associated with chronic kidney disease and its progression in patients with sickle cell anaemia. *British Journal of Haematology*, 164(5), pp.729–739.

Schneider, B.R. et al., 2016. Bl(x)d. *The Journal of The American Society of Hematology*, 48(5), pp.629–638.

Schold, J.D. et al., 2011. The relative risk of overall graft loss and acute rejection among African American renal transplant recipients is attenuated with advancing age. *Clinical Transplantation*.

Sears, D.A., 1978. The Morbidity of Sickle Cell Trait. *Am J Med*, 64(6), pp.1021–1036.

Sebastiani, P. et al., 2005. NIH Public Access. , 37(4), pp.435–440.

Segura, J. & Ruilope, L.M., 2011. Hypertension in Moderate-to-Severe Nondiabetic

CKD Patients. *Advances in Chronic Kidney Disease*, 18(1), pp.23–27.

Sezgin, I. et al., 2011. CCR2 polymorphism in chronic renal failure patients requiring long-term hemodialysis. *Internal medicine (Tokyo, Japan)*.

Sharma, K., 2013. Chapter 1: Definition and classification of CKD. *Kidney International Supplements*, 3(1), pp.19–62. Available at: http://linkinghub.elsevier.com/retrieve/pii/S2157171615311011.

Shier, D., Butler, J. & Lewis, R., 2009. Hole's essentials of Human Anatomy & Physiology. *McGraw-Hill Higher Education*, 10, pp.147–151.

Simón-Sánchez, J. & Singleton, A., 2008. Genome-wide association studies in neurological disorders. *The Lancet Neurology*.

Skinner, S.C. et al., 2018. Increased prevalence of type 2 diabetes-related complications in combined type 2 diabetes and sickle cell trait. *Diabetes Care*.

Spray, B.J. et al., 1995. Familial risk, age at onset, and cause of end-stage renal disease in white Americans. *Journal of the American Society of Nephrology : JASN*.

Stanifer, J.W. et al., 2014. The epidemiology of chronic kidney disease in sub-Saharan Africa: A systematic review and meta-analysis. *The Lancet Global Health*, 2(3), pp.174–181.

Steenland, N.K. et al., 1990. Occupational and other exposures associated with male end-stage renal disease: A case/control study. *American Journal of Public Health*.

Steinberg, H., 1986. alpha- thalassemia in Blacks: Genetic and Clinical Aspects and Interaction with the Sickle Haemoglobin Gene. *The Journal of The American Society of Hematology*, 68(5), pp.985–990.

Steinberg, M.H. & Embury, S.H., 1986. Alpha-thalassemia in blacks: genetic and clinical aspects and interactions with the sickle hemoglobin gene. *Blood*, 68(5), pp.985–90.

Su, S.L. et al., 2012. Gene polymorphisms of angiotensin-converting enzyme and angiotensin II Type I receptor among chronic kidney disease patients in a Chinese population. *JRAAS - Journal of the Renin-Angiotensin-Aldosterone System*.

Taylor, C., Kavanagh, P. & Zuckerman, B., 2014. Sickle cell trait--neglected opportunities in the era of genomic medicine. *Jama*, 311(15), pp.1495–6. Available at: http://jama.jamanetwork.com/article.aspx?articleid=1841474.

Taylor, S.M., Cerami, C. & Fairhurst, R.M., 2013. Hemoglobinopathies : Slicing the Gordian Knot of Plasmodium falciparum Malaria Pathogenesis. *PLoS ONE*, 9(5), pp.1–10.

Tayo, B.O. et al., 2013. Genetic variation in APOL1 and MYH9 genes is associated with chronic kidney disease among Nigerians. *International Urology and Nephrology*.

Tenhunen, R., Marver, H.S. & Schmid, R., 1968. The enzymatic conversion of heme to bilirubin by microsomal heme oxygenase. *Proceedings of the National Academy of Sciences*.

Tonelli, M. et al., 2015. Comorbidity as a driver of adverse outcomes in people with chronic kidney disease. *Kidney International*.

Tripette, J., Loko, G., et al., 2010. Effects of hydration and dehydration on blood rheology in sickle cell trait carriers during exercise. *Am J Physiol Heart Circ Physiol*, 299, pp.908–914.

Tripette, J., Connes, P., et al., 2010. Red blood cell deformability and aggregation, cell adhesion molecules, oxidative stress and nitric oxide markers after a short term, submaximal, exercise in sickle cell trait carriers. *Clinical hemorheology and microcirculation*, 45(June), pp.39–52.

Tsaras, G., Owusu-ansah, A. & Boateng, O., 2009. Complications Associated with Sickle Cell Trait : A Brief. *AJM*, 122(6), pp.507–512. Available at: http://dx.doi.org/10.1016/j.amjmed.2008.12.020.

Tubman, V.N. et al., 2016. Newborn Screening for Sickle Cell Disease in Liberia : A Pilot Study. *Pediatr Blood Cancer*, 63(August 2015), pp.671–676.

Turner, S.T. et al., 2008. Genomic association analysis suggests chromosome 12 locus influencing antihypertensive response to thiazide diuretic. *Hypertension*.

Tzur, S. et al., 2010. Missense mutations in the APOL1 gene are highly associated with end stage kidney disease risk previously attributed to the MYH9 gene. *Human Genetics*.

Ulasi, I.I. et al., 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 - Clinical Practice*.

United Nations, D. for E. and S.A., 2015. *World population prospects; The 2015 Revision, Key Findings and Advance Tables*,

US Renal Data System, 2013. USRDS 2013 Annual Data Report: Atlas of Chronic Kidney Disease and End-Stage Renal Disease in the United States. *National Institutes of Health National Institute of Diabetes and Digestive and Kidney Diseases*, pp.1–26.

Vega, F.M.D. La & Bustamante, C.D., 2018. Polygenic risk scores: a biased prediction? *Genome Medicine 2018 10:1*, 10(1), p.100. Available at: https://genomemedicine.biomedcentral.com/articles/10.1186/s13073-018-0610-x.

Villar, E., Chang, S.H. & McDonald, S.P., 2007. Incidences, treatments, outcomes, and sex effect on survival in patients with end-stage renal disease by diabetes status in Australia and New Zealand (1991 2005). *Diabetes care*, 30(12), pp.3070–6.

Warren, K.E., Gidvani-diaz, V. & Duval-arnould, B., 1999. Renal Medullary Carcinoma in an Adolescent With Sickle Cell Trait. , 103(2).

Wasser, W.G. et al., 2012. Population genetics of chronic kidney disease: The evolving story of APOL1. *Journal of Nephrology*.

Watanabe, I.C. et al., 2007. Renal medullary carcinoma : report of seven cases from Brazil. *Modern Pathology*, 20, pp.914–920.

Weatherall, D. et al., 2005. A case for developing North-South partnerships for research in sickle cell disease. , 105(3), pp.921–923.

Wilcox, M., Bjorkman, A. & Brohult, J., 1983. Falciparum malaria and B-thalassaemia trat in Norther. *Ann Trop Med*, 77(4), pp.335–347.

Wild, S. et al., 2004. Global Prevalence of Diabetes: Estimates for the year 2000 and projections for 2030. *Diabetes Care*, 27(5), pp.1047–1053.

Williams, T.N., Wambua, S., et al., 2005. Both heterozygous and homozygous ␣. *English*, 106(1), pp.368–371.

Williams, T.N., Mwangi, T.W., Wambua, S., Peto, T.E.A., et al., 2005. Negative epistasis between the malaria-protective effects of alpha+-thalassemia and the sickle cell trait. *Nature genetics*, 37(11), pp.1253–7. Available at: http://www.nature.com/doifinder/10.1038/ng1660%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/16227994%5Cnhttp://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3521056.

Williams, T.N. et al., 2012. Negative epistasis between the malaria-protective effects of thalassemia and the sickle cell trait Europe PMC Funders Author Manuscripts. *Nat Genet*, 37(11), pp.1253–1257.

Williams, T.N., Mwangi, T.W., Wambua, S., Alexander, N.D., et al., 2005. Sickle Cell Trait and the Risk of Plasmodium falciparum Malaria and Other Childhood Diseases. *J Infect Dis.*, 192(1), pp.178–186.

Wong, C.S. et al., 2009. Association of proteinuria with race, cause of chronic kidney disease, and glomerular filtration rate in the chronic kidney disease in children study. *Clinical Journal of the American Society of Nephrology*.

World Kidney Day, 2016. World Kidney Day. *Kidney Health for All*.

Yamagishi, S. ichi, 2011. Role of advanced glycation end products (AGEs) and receptor for AGEs (RAGE) in vascular damage in diabetes. *Experimental Gerontology*.

Yoshioka, T. et al., 1987. Role of abnormally high transmural pressure in the permselectivity defect of glomerular capillary wall: A study in early passive Heymann nephritis. *Circulation Research*, 61(4), pp.531–538.

Zhang, L. et al., 2012. Prevalence of chronic kidney disease in China: A cross-sectional survey. *The Lancet*, 379(9818), pp.815–822.

Zhang, Q.-L. & Rothenbacher, D., 2008. Prevalence of chronic kidney disease in population-based studies: Systematic review. *BMC Public Health*, 8(1), p.117. Available at: http://bmcpublichealth.biomedcentral.com/articles/10.1186/1471-2458-8-117.

Zhou, X.J. et al., 2008. The aging kidney. *Kidney International*, 74(6), pp.710–720.

Appendices

Appendix 1 Reagents Used

PCR Reagents

GOTAQ® TAQ POLYMERASE

PROMEGA

MADISON, WISCONSIN

USA

5X COLORLESS GOTAQ® REACTION BUFFER

LOT NUMBER: 0000 1173 14

PROMEGA

MADISON, WISCONSIN

USA

DNTP MIXTURE

LOT NUMBER: BH4601C

TAKARA BIO INC

NOJIHIGASHI, KUSATSU, SHIGA PREFACTURE

JAPAN

PRIMERS

INTEGRATED DNA TECHNOLOGIES

CORALVILLE, IOWA

USA

DISTRIBUTED BY: WHITEHEAD SCIENTIFIC

CAPE TOWN, WESTERN CAPE

SOUTH AFRICA

GEL ELECTROPHORESIS REAGENTS

6X DNA LOADING DYE

LOT NUMBER: 0011 6849

THERMO SCIENTIFIC

THERMOFISCHER SCIENTIFIC

WALTHAM, MASSACHUSETTS

USA

GENERULER™ 100BP DNA LADDER

INVITROGEN

THERMOFISHER SCIENTIFIC

WALTHAM, MASSACHUSETTS

USA

SEAKEM® LE AGAROSE

LOT NUMBER: 0000 4364 71

LONZA

BASEL

SWITZERLAND

DISTRIBUTED BY: WHITEHEAD SCIENTIFIC

CAPE TOWN, WESTERN CAPE

SOUTH AFRICA

PCR CLEAN-UP REAGENTS

FASTAP THERMOSENSITIVE ALKALINE PHOSPHATASE

LOT NUMBER: 0024 7986

THERMO SCIENTIFIC

THERMO FISCHER SCIENTIFIC

WALTHAM, MASSACHUSETTS

USA

EXONUCLEASE 1

LOT NUMBER 0025 4712

THERMO SCIENTIFIC

THERMO FISCHER SCIENTIFIC

WALTHAM, MASSACHUSETTS

USA

SEQUENCING REAGENTS

BIGDYE® TERMINATOR V1.1, V3.1 5X SEQUNCING BUFFER

LOT NUMBER 1103 200

APPLIED BIOSYSTEMS

THERMO FISCHER SCIENTIFIC

WALTHAM, MASSACHUSETTS

USA

BIGDYE® TERMINATOR V3.1 CYCLE SEQUENCING RR-100

LOT NUMBER: 1201 096

APPLIED BIOSYSTEMS

THERMO FISCHER SCIENTIFIC

WALTHAM, MASSACHUSETTS

USA

Appendix 2: Machines Used

Thermocyclers
BioRad T100™ Thermal Cycler

Applied Biosystems 2720 Thermal Cycler

Applied Biosystems GeneAmp® PCR System 9700

Power Pack

DS 5000XT DC Power Supply

Hoefer Scientific Instruments

Heating Block

HyBaid Touchdown Thermocycler

Scientific Group

Centrifuge
Eppendorf centrifuge 5415D

Vortex
Labnet Vortex Mixer

Vortex Mixer Model VM-1000

Digisystems Lab Instruments Inc

Appendix 3: Gel Electrophoresis Reagents

<u>SeaKem®LE Agarose</u>

Lonza (ME, USA)

1% (w/v) Agarose gel (50 ml or 100 ml)

0.5 g or 1 g agarose

50 ml or 100 ml 1x TBE buffer

Dissolve agarose powder by microwave heating

Allow to cool, add 3 µl or 6 ul of Etbr and pour into a gel electrophoresis casting tray with a well-comb

Appendix 4: PCR Amplification Reagents

Table 12 Reagents used to amplify SCD mutation

Reagent	**Stock**	**Required**	**v=1 (µl)**
dH2O	-	-	11.9
Taq Buffer	5X	1X	5
dNTPs	5mM	200µM	1
Forward Primer	20µM	0.04µM	0.5
Reverse Primer	20µM	0.04µM	0.5
Taq	5U/µl	0.02µM	0.1
DNA	-	100ng	1
Total	**-**	**-**	**25**

Table 1.2 Protocol used amplification.

	Temperature (°C)	**Time**	**Cycles**
Initial Denaturation	95	3 Minutes	-
Denaturation	94	30 Seconds	35X
Primer Annealing	58.1	30 Seconds	
Elongation	72	1 Minute	
Final Elongation	72	10 Minutes	-

Table 3: PCR clean-up protocol

Reagent	**V=1(µl)**
FastAP	1.0
Exo1	0.1
PCR Product	8.9
Total	**10.0**

Samples were incubated for 1 hour at 37°C followed by 15 minutes 72°C

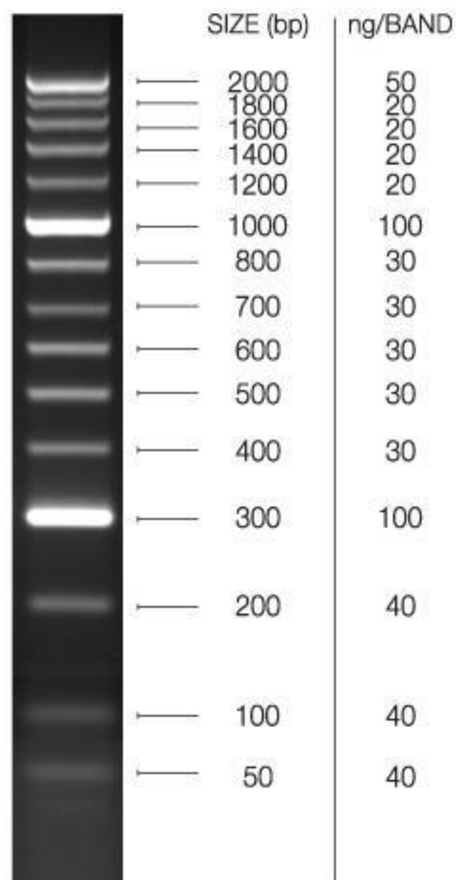
Appendix 5: Sequencing Conditions and Protocol

Table 4: Reagents used for direct cycle (Sanger) Sequencing

Reagents	**Stock Concentration**	**Required Concentration**	**v=1(µl)**
Sabax Water	-	-	4
Dilution Buffer	5x	-	2
Terminator mix	-	-	1
Primer	20 µM	0.4µM	0.5
Cleaned PCR Product	-	-	2.5
Total	**-**	**-**	**10**

Table 5: Sequencing protocol

	Temperature (°C)	**Time**	**Cycles**
Initial Denaturation	95	3 Minutes	-
Denaturation	94	30 Seconds	35X
Primer Annealing	58.1	30 Seconds	
Elongation	72	1 Minute	
Final Elongation	72	10 Minutes	-



2.5 ul of GeneRuler 100bp plus DNA Ladder was loaded onto the gels Fermentas (Life Sciences) USA