GENETIC ASSOCIATION AND GENE-GENE INTERACTION ANALYSIS OF APOL1, MYH9 AND G6PD VARIANTS IN PATIENTS WITH CHRONIC KIDNEY DISEASE

BY

PRISCILLA ABENA AKYAW

(10306073)

THIS THESIS IS SUBMITTED TO THE UNIVERSITY OF GHANA, LEGON IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE AWARD OF MPHIL MOLECULAR BIOLOGY DEGREE

DECEMBER, 2017

DECLARATION

I, PRISCILLA ABENA AKYAW, do hereby declare that with the exception of references to other people's work, which have been duly acknowledged, this thesis is the outcome of my own research conducted at the Department of Biochemistry, Cell and Molecular Biology and at the Noguchi Memorial Institute for Medical Research, University of Ghana, College of Basic and Applied Science under the supervision of Dr. Anita Ghansah and Professor Gordon Awandare. Neither all nor parts of this project have been presented for another degree elsewhere.

……. Date:……..

PRISCILLA ABENA AKYAW

(Student)

……. Date: ……

DR. ANITA GHANSAH

(Supervisor)

…….. Date: ……..

PROF. GORDON AWANDARE

(Supervisor)

DEDICATION

This research work is foremost dedicated to the Almighty God for His undeserved kindness that has been showered upon me throughout my academic life. It is also dedicated to all those who have died because they couldn't afford the cost of renal therapy.

ACKNOWLEDGEMENT

I am most grateful to the Almighty God for his wisdom, guidance and protection throughout my academic years. I want to thank my supervisors; Dr. Anita Ghansah for her patience, the help she gave to make me better understand the project and her concern about how the work was progressing as well as the corrections and advice. I am also grateful to my co-supervisor, Prof. Gordon Awandare for his guidance and support throughout the project. My profound gratitude goes to the research team from the H3Africa Kidney Disease Research Network, especially to the principal investigator, Dr. Dwomoa Adu for creating this amazing network of scientists from where I started my research career, to increase the awareness of kidney disease genetics in Africa and also to H3Africa genomics lab at the Noguchi Memorial Institute for Medical Research (NMIMR) University of Ghana, Legon (U.G) especially Ernestine Kubi, Barbara Mensah and Nana Yaa Agyemang Karikari, for assisting me whenever help was needed. I also want to thank Dr. Christian Obirikorang from the KNUST School of medicine for his help and advice especially during the troubleshooting period of the lab work.

A big thank you also goes to the West African Centre for Cell Biology of Infectious Pathogens (WACCBIP) for funding this work by providing a scholarship for research cost and school fees. I also want to thank the administrative staff of WACCBIP for helping me in diverse ways from the start to the completion of this work.

Finally, I am grateful to my family and friends especially my mother, Mary Asamoah and sister, Ama Henewaa for their encouragement and emotional support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ACR	Albuminuria-to Creatinine Ratio
ACTN4	Actinin-4
AIDS	Acquired Immune Deficiency Syndrome
AKI	Acute Kidney Injury
APOL1	Apolipoprotein 1
APS	Ammonium per Sulphate
ATP	Adenosine Tri Phosphate
BCL2	B cell Lymphoma 2
BH3	Bcl-2 Homology region 3
BP	Blood Pressure
CKD	Chronic Kidney Disease
DNA	Deoxyribonucleic Acid
EDTA	Ethylenediaminetetraacetic Acid
ESKD	End-Stage Kidney Disease
ESRD	End-Stage Renal Failure
FSGS	Focal Segmental Glomerular Sclerosis

GDP	Gross Domestic Product
GFR	Glomerular Filtration Rate
G6PD	Glucose-6-Phosphate dehydrogenase
GSH	Glutathione
GSSG	Glutathione Disulphide
GHS	Ghana Health Service
GSH-PX	Glutathione Peroxidase
GWAS	Genome Wide Association Study
HAT	Human African Trypanosomiasis
HDL	High Density Lipoprotein
H-ESKD	Hypertension –attributed End Stage Kidney Disease
HIV-AN	HIV-Associated Nephropathy
HIV	Human immunodeficiency Virus
H_2O_2	Hydrogen Peroxide
HPR	Haptoglobin-related protein
INFɣ	Interferon gamma
IPTp	Intermittent Preventive Treatment during Pregnancy
IRB	Institutional Review Board

LAMB2	Laminin-β2
LD	Linkage disequilibrium
LDR	Ligation Detection Reaction
MAD	Membrane Addressing Domain
MYH9	Human non-muscle Myosin Heavy Chain 9 gene
NADPH	Nicotinamide Adenine Dinucleotide Phosphate
NADP+	Reduced Nicotinamide Adenine Dinucleotide Phosphate
NIH	National Institute of Health
NMIMR	Noguchi Memorial Institute for Medical Research
NPHS1	Nephrin 1
NPHS2	Nephrin 2
NSAIDS	Non-Steroidal Anti-Inflammatory Drugs
OPD	Out Patients Department
PCR	Polymerase Chain Reaction
PFD	Pore Forming Domain
PLCE1	Phospholipase C Epsilon 1
RBC's	Red Blood Cells
ROS	Reactive Oxygen Species

SRA	Serum Resistance Associated binding domain
SRNS	Steroid-Resistant Nephrotic Syndrome
SNP	Single Nucleotide Polymorphism
SP	Sulphadoxine –Pyrimethamine
TBE	Tris-borate-EDTA
TEMED	Tetramethylethylenediamine
TRPC6	Transient Receptor Potential Cation channel subfamily member 6
TNF	Tumour Necrosis Factor
UTI	Urinary Tract Infections
UMOD	Uromodulin encoding gene
W.H.O	World Health Organization

ABSTRACT

Chronic Kidney Disease (CKD) is a major health crisis globally because of the high prevalence recorded in many developing and developed countries and is estimated to affect about 200 million persons worldwide. Variants in the human Apolipoprotein L1 (APOL1) and the Human non-Muscle Myosin (MYH9) genes are associated with the risk of CKD in people of West African descent, but not all individuals with these mutations develop CKD in their lifetime. There is a need to identify other genetic mutations that predisposes individuals to disease association. This work was aimed at determining the prevalence of G6PD variants A- (202) and A (376) in CKD patients of unknown aetiology and also to determine whether G6PD variants could be a risk factor by interacting with APOL1 and MYH9 gene variants in increasing or reducing the risk of developing CKD in the West African region. This was a case-control study, which utilised archived blood samples from CKD patients of unknown aetiology from Ghana and Nigeria, collected on behalf of the H3Africa Kidney disease Research Network (KDRN) in Ghana. Samples were genotyped using Restriction Fragment Length Polymorphism (RFLP) following PCR amplification of the region of the G6PD gene that carries these mutations. Whiles genotyping of APOL1 and MYH9 variants was done using Ligation Detection Reaction (LDR) assay. A total of 411 samples made up of 202 cases (CKD patients of unknown aetiology) and 209 controls (from the general population) were genotyped for 2 G6PD variants (A- and A), 4 APOL1 variants (rs73885319 (G1), rs60910145 (G1), rs71785313 (G2) and rs9622363) and 4 MYH9

variants (rs20324871, rs4821481 (E1 haplotype) and rs5750248 and rs5750250 (S1 haplotype)).

The general prevalence of G6PD A- and A observed for cases was 49% and 54.5% respectively, whiles that of the control population was 44% and 45.9% respectively. In terms of gender, the prevalence observed for the male and female cases were 43.2% and 56% for G6PD A- and 47.7% and 62.6% for G6PD A variants respectively. There was a significant difference between the prevalence observed for the females and the control group ($p = 0.001$) but not for the male subjects ($p = 0.4205$) compared to the male control group. Significantly strong associations was observed between females who had at least one risk allele (heterozygous carriers) of G6PD A- and CKD of unknown aetiology (OR 1.96, 95%CI 1.0044-3.3030, $p = 0.048$) which further increased for those with 2 risk alleles (homozygous carriers) (OR 6.8, 95%CI 2.2225-19.2411, $p = 0.001$), whilst a similar association was observed between females with 1 risk allele of G6PD A variant and CKD of unknown aetiology (OR 2.0, 95% CI 1.1206-3.6372, $p = 0.019$). Significantly strong associations was also observed with previously reported APOL1 variants rs73885319 (OR 1.3, 95%CI 1.6435-4.6118, $p = 0.02$) and rs60910145 (OR 1.3, 95%CI 1.7042-4.9177, $p = 0.004$) after controlling for age, gender and clinical site location. No significant evidence of association was observed between CKD of unknown aetiology and any of the MYH9 variants after accounting for multiple testing in the sample. Interaction analysis however revealed significant epistatic interactions between deficient G6PD A- and A genotype and heterozygous rs4821481 ($p = 0.016$), a trend towards association for rs5750250 ($p = 0.062$), and the APOL1 variants rs60910145 (0.014) and rs73885319 (0.002) which indicates a synergistic epistatic interaction

between carriers of these genotypes which could predispose them to CKD of unknown aetiology.

CHAPTER ONE

INTRODUCTION

1.1 Background

Chronic Kidney Disease (CKD) is defined as the progressive loss of kidney function over a period of time (usually for 3 months) with the cardinal symptom being the irreversible renal decline irrespective of the underlying cause of the kidney disease (Eckardt & Kasiske, 2009). Other symptoms include general tiredness, frequent urination, blood in the urine and excessive foam creation in urine. According to the National kidney foundation, two main tests are considered in order to diagnose an individual to be suffering from any form of kidney disease. The two main tests are the Albuminuria-to-Creatinine Ratio (ACR) and the Glomerular Filtration Rate (GFR) (National Kidney Foundation, 2014). There are 5 different stages of kidney disease, and the GFR is usually used as a means of distinguishing between the stages, starting from Acute Kidney Injury (AKI) to End-Stage Kidney Disease (ESKD). A glomerular filtration rate of 90 ml/min/1.73 m^2 over a 24-hour urine collection period may indicate a likelihood of kidney disease which is also sometimes known as Stage 1 or AKI. This stage can be easily reversed when detected early by a physician using recommended drugs and dietary plans. However, a glomerular filtration rate which is below 60 ml/min/1.73 m^2 indicates the onset of CKD which is from stage 2 to stage 4 and is presented as either a mild to moderately reduced kidney function or a GFR that is less than the estimated value (60 ml/min/1.73 m^2) for a period of 3 months. Finally, a glomerular filtration rate of less than 15 ml/min/1.73 m^2 represents the 5^{th} stage and is an indication of the severe form of kidney damage also referred to as End-Stage Renal Failure (Levey *et al*, 2009). At this stage, individuals will have to

either prepare to get a kidney transplant or be put on a dialysis machine till they get a kidney transplant.

Chronic Kidney Disease is a worldwide problem affecting millions of people, especially in the middle and low-income countries (Di Angelantonio *et al*., 2007). According to the 2015 global burden of disease study, CKD was ranked 27th in the list of causes of total number of deaths worldwide in 1990, but rose to 18th in 2010. This degree of movement up the list was second only to that for HIV and AIDS (Jha & Prasad, 2016). Mortality due to CKD is still on the increase and rose by 32% between 2005 and 2015 with 1.2 million deaths worldwide. Incidence and prevalence rates still keep increasing all over the world. In sub-Saharan Africa due to the extreme cases of poverty and inadequate healthcare facilities, there is a sharp increase in prevalence and incidence of non-communicable diseases like hypertension and diabetes, as well as communicable diseases like HIV, Hepatitis B, leishmaniosis and infectious glomerulonephritis all of which are known risk factors for CKD. Adults within the age bracket of 20-50 years are mainly known to develop CKD from hypertension, glomerulonephritis and unknown aetiologies (Arogundade *et al*., 2008).

The annual worldwide costs are estimated at 70 to 75 billion US dollars to maintain the renal replacement therapy of the roughly 1.1 million worldwide dialysis patients in 2001. In Ghana, renal replacement therapy (mainly haemodialysis and peritoneal dialysis) is available only in two of the teaching hospitals. The estimated cost of dialysis, initial laboratory investigations and medications is GHC 57,600 (approximately $12,500) per patient per annum. This amount is rather high for a country with a per capita income of $1,858.24 and a Gross Domestic Product (GDP) growth rate of 3.6%.

The cause of CKD is not always known, but research shows that it can be caused by other disease conditions and environmental factors such as heavy metals overexposure or, long term use of certain drugs can cause damage to the kidneys. Studies have shown that there is a genetic component to CKD (Köttgen *et al*., 2009). This was triggered by the observation that African-Americans are 3 times more likely to develop CKD and progress faster to ESRD than European-Americans even after accounting for all socio-economic factors like access to quality healthcare as well as the high poverty rate (Köttgen *et al*., 2009). The initial research associated the progression with mutations in the Human non-muscle Myosin Heavy Chain 9 gene (MYH9). This study which was conducted in the African-American population using mapping by admixture Linkage disequilibrium (LD), identified the MYH9 gene variants as the only major disease susceptibility locus for CKD and ESRD. These variants have also been associated with an increased susceptibility to Focal Segmental Glomerular Sclerosis (FSGS), HIV-Associated Nephropathy (HIV-AN) and Hypertension-attributed End Stage Kidney Disease (H-ESKD) in the African-American population. The exact mechanism by which the MYH9 variant causes kidney disease still remains unknown (Genovese *et al*., 2010; Tayo *et al*., 2013). With the onset of the 1000 genome project, later studies have also attributed the disease progression to other gene variants of the human Apolipoprotein 1 (APOL1); a gene found on the same chromosome (Chromosome 22) as MYH9 and codes for the APOL1 protein (Freedman *et al*., 2010; Genovese *et al*., 2013). Other studies have also shown that APOL1 variants are more strongly associated to the risk and progression of the disease to end stage than the MYH9 gene variants (Shay Tzur et al., 2010c). Two APOL1 nephropathy variants, G1 and G2 have been associated with high risk of non- diabetic ESKD in people of African descent (Genovese *et al*., 2010).

Research conducted shows that Africans and people from African descent who have at least one copy of the variants are protected from *Trypanosoma brucei rhodiense*, a parasite responsible for causing African sleeping sickness in the African sub-region. Inheriting 2 copies of the variants however increases ones risk of developing kidney disease (Kruzel-Davila *et al*., 2016). This selection was established to be as a result of Africans developing resistance to Human African Trypanosomiasis (HAT) (Palmer *et al*., 2015). This selection pattern confirms the high frequencies of the APOL1 variants observed across the West African region where trypanosomiasis is endemic, with Ghana and Nigeria recording the highest frequencies (30-50% for G1 and 20-30% for G2) (Kopp *et al*., 2011; Tzur *et al*., 2012). Interestingly, not all Individuals who have these mutations actually develop kidney disease (Genovese *et al*., 2010). This led to the hypothesis that in addition to the APOL1 and MYH9 gene variants, other possible gene mutations or environmental factors could also exist that can either reduce the effect of the APOL1 and MYH9 in causing kidney disease or serve as alternative risk factor for kidney disease. This is sometimes referred to as the "second hits hypothesis" or "the modifying factors hypothesis" (Freedman & Skorecki, 2014). This theory is not farfetched because there is evidence which supports the interaction between APOL1 and MYH9 gene variants and the risk of developing other genetic associated diseases like cardiovascular disease (Ito *et al*., 2014) and the progression of Familiar Haematuria (Voskarides *et al*., 2013).

Another genetic variant that is known to be predominant in the African sub-region and records high frequencies in similar geographical regions like APOL1 and MYH9 genetic variants is the Glucose-6-Phosphate dehydrogenase (G6PD) deficiency, a well-known genetic mutation that affects over 400 million people worldwide (Cappellini & Fiorelli, 2008). Glucose-6-Phosphate Dehydrogenase deficiency occurs

worldwide and is strongly observed in black males with a prevalence of 10% in the United States (Manganelli, Masullo, Passarelli, & Filosa, 2013) and 32.5% in sub-Saharan Africa (Howes *et al*., 2012). In Ghana, the incidence of G6PD deficiency was known to be 15% far back in the 1990's but current research shows that the prevalence is 19% for the coastal region and 9% for the middle belt (Carter *et al*., 2011) whereas prevalence rates recorded among Nigerian Children is known to be 16.9% (Williams *et al*., 2013a).

Over 400 G6PD variants have been identified and the polymorphisms are predominantly defined to specific geographic locations (Beutler, 1996). Research has showed that the highest prevalence rates of G6PD deficiencies are found predominantly in malaria endemic areas such as some parts of Asia, tropical Africa and the Mediterranean (Phompradit *et al*., 2011). A study done in Asia specifically in Sri Lanka recorded a high prevalence of G6PD deficiency in patients with CKD of unknown aetiology by measuring the enzyme activity levels of individuals who reported to the hospital with symptoms of renal failure and were diagnosed to have CKD of unknown aetiology (Jmkb *et al*., 2013). Like Asia, Africa shares similarities in prevalence of G6PD deficiency, malaria endemicity and the increasing prevalence and incidence of CKD. Among all the genetic variations that are found to be prevalent in these region, only G6PD deficiency has been known to lead to both acute and chronic renal failure (Jmkb *et al*., 2013).

1.2 Problem Statement

Chronic Kidney Disease (CKD) is a major health problem globally because of its high prevalence recorded in many developing and developed countries (Wouters *et al*., 2015) and is estimated to affect about 200 million persons worldwide (Freedman & Skorecki, 2014). Even though CKD is listed as a non-infectious disease according to the World Health Organization (WHO), it is gradually overtaking other infectious diseases like malaria as the leading cause of premature death in the world (Couser *et al*., 2011). Genetic variants found in APOL1 and MYH9 genes have been associated with the increased risk of CKD associated with FSGS and HIV-AN in African-Americans as compared to European-Americans (Ashley-Koch *et al*., 2011; Freedman *et al*., 2010; Gutierrez *et al*., 2016; Weckerle *et al*., 2016). Studies have also confirmed this association in some African countries with the highest frequencies from Ghana and Nigeria (Tayo *et al*., 2013), but not all individuals with the APOL1/MYH9 genetic variant actually develop CKD in their lifetime (Genovese *et al*., 2010). Currently, no study has been done to investigate the possible evidence of association between G6PD variants and the potential genetic interaction it has with the APOL1 and MYH9 genetic variants that could increase the risk of CKD in the African population (especially in countries that record the highest APOL1/MYH9 nephropathy variant frequencies like Ghana and Nigeria) as well as the role of G6PD deficiency in CKD patients of unknown aetiology in the African sub-region. All genetic association studies done on this subject have not considered patients with CKD of unknown aetiology, and as such no data exists that associates the APOL1 and MYH9 genetic variants to the risk of CKD of unknown aetiology. Knowledge of these genetic associations and other additional candidate genes that interacts with APOL1 and MYH9 or are independently associated with CKD can aid in early diagnosis of

CKD. It can also help inform physicians of the right consultation and recommended drugs to administer in order to prevent the disease at an early stage and reduce its effect even at the end-stage. This will also help scientist to develop more potent drugs that can target the disease at different stages.

1.3 Hypothesis

Genetic association and gene-gene interactions between G6PD, APOL1 and MYH9 variants increases the risk of CKD of unknown aetiology.

1.4 Aims and Objectives

1.4.1 Aim

To determine the genetic association and gene-gene interactions between G6PD, APOL1 and MYH9 variants and the risk of CKD of unknown aetiology.

1.4.2 Objectives

1. To establish the prevalence of G6PD gene variants A- and A in the general population as well as CKD cases of unknown aetiology.
2. To determine the associations between G6PD, APOL1 or MYH9 genetic variants and CKD of unknown aetiology
3. To determine the effect of any gene-gene interactions between APOL1 or MYH9 and G6PD genetic variants on the incidence of CKD of unknown aetiology

CHAPTER TWO

LITERATURE REVIEW

2.1 Functions of the Kidneys

The kidneys are a pair of bean shaped organs that are found below the liver in the peritoneal cavity and has a pair of adrenal glands called suprarenal glands sitting directly on top of them. In addition to the respiratory *system* as well as the skin, they form the main excretory organs of the body. The kidneys are mainly responsible for filtering the blood. The process of filtration involves the removal of unwanted liquid materials as well as excess salts, potassium, ions and other metabolic waste that are toxic to the body when left in the system. The Kidneys are also known for maintaining the systems normal balance of body fluids. Apart from this, they are also involved in some immune functions due to the presence of a network of macrophages and dendritic cells that are found in them (Teteris *et al*., 2011). These specialized immune cells contribute to both innate and adaptive immunity in the kidney cells by fighting against infections that can lead to kidney injury (Kurts *et al*., 2013). For instance, the kidney dendritic cells rapidly produce neutrophil-recruiting chemokines during bacterial pyelonephritis; the most prevalent form of kidney infection (Tittel *et al*., 2012). The kidneys are also very important because they secret hormones that act on other organs; for example, erythropoietin is produced in the kidneys and is necessary to stimulate the production of Red Blood Cells (RBC's), whiles renin and calcitol regulates blood pressure and plasma calcium levels respectively (Katz & Lindheimer, 1977). Kidney function is thus very essential for life and as such, a complete loss of function will result in death in a few days.

The basic functioning unit of the kidney is the nephron. They are responsible for eliminating metabolic waste from the body, regulating blood pressure and also controlling the level of electrolytes and metabolites in the blood. The kidneys comprises of millions of these nephrons all carrying out the same function at any point in time. In humans for example, the kidneys contains 800,000 to 1.5 million nephrons (Carroll & Abdel-Rahman, 2011). Each nephron has a filtrating unit called the glomerulus and a long tubule made of a dozen differentiated segments. Other components of the nephron includes the loop of Henle and the collecting duct which empties excretory waste from the nephron into the bladder as described in figure 2.1.

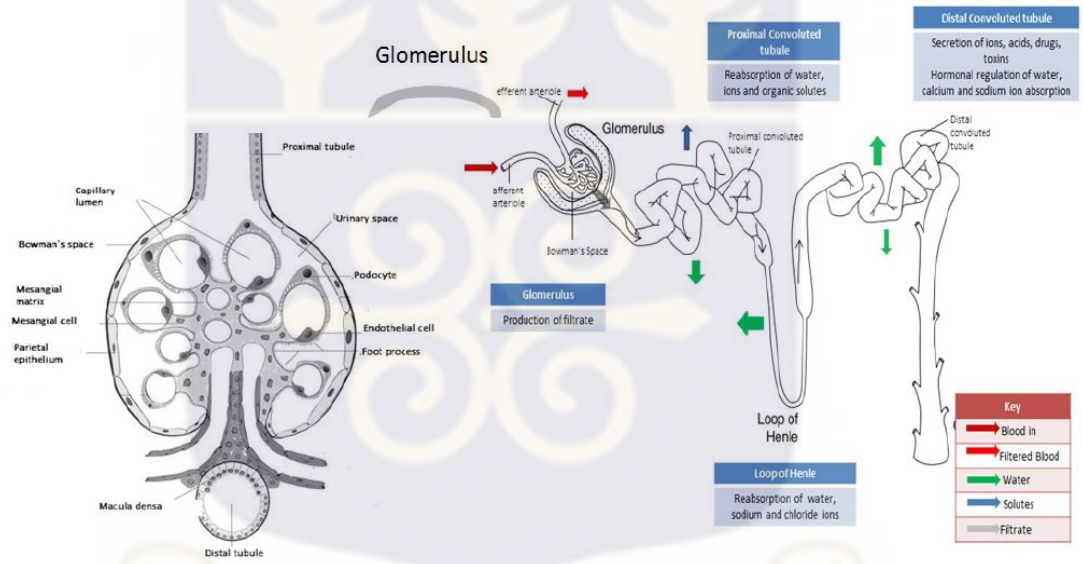

Figure 2.1 Structure of the nephron, the basic unit of the kidney. Each nephron filters a small amount of blood at a time. It is made of a sack of vessels called the glomerulus which sits inside the Bowman's capsule which is connected to the collecting ducts through two convoluted tubules (Kurts *et al.*, 2013)**.**

2.2 Kidney Disease

Kidney disease results when there is a reduction in the amount of functioning nephrons in the kidney. Ordinarily, the kidney can still function even when 90% of the nephrons are damaged but the gradual loss of function of the kidneys is what is termed as chronic kidney disease (National Kidney Foundation, 2002). Typical symptoms of kidney disease include tiredness, trouble with sleeping, experiencing dry and itchy skin, blood in the urine, swellings around the ankles and joints and foamy or excessive foam in the urine (National Kidney foundation, 2002).

According to the world kidney foundation (2014), two main tests are considered in other to diagnose an individual to be suffering from any form of kidney disease. They are the Albuminuria-to Creatinine Ratio (ACR) and the Glomerular Filtration Rate (GFR) (Levey *et al*., 2003). Albuminuria is defined as the presence of a disproportionate amount of a protein called albumin in the urine and is one of the most common markers for identifying persons with kidney or renal diseases (Turin *et al*., 2014). Persons with normal kidneys have no albumin in their urine, the ACR is therefore a means of estimating the amount of albumin that is present in the urine and helps physicians identify the extent of damage that the kidneys have incurred. The Glomerular Filtration Rate on the other hand is defined as the measure of the level of kidney function by considering factors such as the blood creatinine levels, age, body size and gender. Creatinine is a by-product of muscle metabolism and is usually found in the urine in a constant level to depict the urine concentration (Garg *et al* ., 2002).

Kidney disease occurs in distinct stages (described in Table 2.1 below), all of which can be determined by the GFR. The glomerular filtration rate is calculated based on how well the blood is filtered by the kidneys and normally takes into consideration the patients age and serum creatinine levels. A GFR which is under 90 ml/min/1.73m^2

over a 24-hour urine collection period may indicate a likelihood of kidney disease (National kidney foundation, 2002). The chronic stage of the disease is defined as either a progressive kidney damage or a GFR that is less than the estimated value for a period of 3 months (Wouters *et al*., 2015).

Table 2.1: Stages of CKD according to the National Kidney Foundation.

Stage	Description	GFR/ml/min/1.73m^2
1	Kidney damage with normal or high GFR	90>
2	Kidney damage with mildly low GFR	60-89
3	Kidney damage with moderately low GFR	30-59
4	Kidney damage with severely low GFR	15-29
5	Kidney failure	<15

Source: webmd.com (accessed on February 10, 2016). Stages 1 and 2 are usually classified together as Acute Kidney Injury which can be reverted back to its original function. Stages 3 and 4 together are also classified as Chronic Kidney Damage and cannot be reverted.

2.3 Global Distribution of Chronic Kidney Disease

Chronic kidney disease has become an important public health concern worldwide with an estimated incidence growth of 6% annually (Kasembeli *et al*., 2015). According to the 2010 global burden of disease study, CKD rose from 27th to 18th position within 2 decades. Currently, 10% of the population worldwide is affected by

CKD which results in millions of deaths each year especially in low income countries. This is mostly because of the high cost of treatment (World Kidney Disease, 2015). In Africa more than 500,000 people develop ESRD every year with a prevalence of 10.7% in Nigeria (Afolabi *et al*., 2009) and a varying prevalence in Ghana, ranging from 4% in hypertensives in the Greater-Accra region (Addo *et al*., 2009) to 46.9% in hypertensives in the general population of Ghana (Osafo *et al*., 2015).

Studies also reveal that the prevalence of CKD is higher in women than in men, even though the rate of progression to End-Stage is higher in men than in women (Goldberg & Krause, 2016). In the US, for instance, between 2007 and 2012, the US Renal Data System reported a prevalence of 15.1% in women and 12% in men (US Renal Data System, 2012). This theory, however, has been debunked by other studies across different countries where they observed a higher prevalence in men than in women (Jungers *et al*., 1996; Zhang *et al*., 2008).

2.4 Risk factors of Chronic Kidney Disease

Most people who suffer from CKD normally start without any sign of symptoms and slowly progress till they start showing symptoms long after the first instance of kidney damage has occurred. This makes it very difficult to determine the exact aetiology of some of the kidney disease cases. There are however various indicators that point to certain environmental, genetic and lifestyle factors that can easily affect ones kidneys and thus cause it to deteriorate slowly (Kazancioğlu, 2013; Freedman *et al*., 2009b; Moore & Williams, 2009). These have been grouped into two classes; they are those that cause the CKD (risk factors) (shown in table 2.2 below) and those that are associated with CKD in the absence of established causal relations (risk markers).

Table 2.2: Factors associated with an increased risk of Chronic Kidney Disease

• Hypertension • Illicit drug use • Chronic anaemia • Socio-economic status • Heavy metal consumption • Race and ethnicity • Gender/ Age	• Obesity • Hereditary/Genetic factors • Oxidative stress • infections (e.g. HIV and Hepatitis) • High protein intake/ Proteinuria • Organic solvents • Diabetes

Source: (Mitch, 2007)

2.4.1 Hypertension

Chronic Kidney Disease and hypertension are closely associated with an overlying and intermingled cause and effect relationship. There is also an overwhelming evidence that shows that Hypertension causes a decline in renal function (Fox *et al*., 2004; Ishida *et al*., 2001; Young *et al*., 2002) and also leads to end stage renal disease (Klag *et al*., 1996). High blood pressure can damage blood vessels in the kidneys, reducing their ability to work properly. When the pressure of blood flow is high, blood vessels stretch so that blood flows more easily. Eventually, this stretching scars and weakens blood vessels throughout the body, including those in the kidneys. If the kidneys' blood vessels are damaged, they may stop removing wastes and extra fluid from the body (Safian & Textor, 2001) .

2.4.2 Diabetes

Diabetes is known to be one of the leading causes of kidney disease and is also attributed to the cause of death of most type 1 and type 2 diabetic patients (Gross ML, 2004). According to the National Kidney Foundation, about 30% of all patients with type 1 and 40% of all patients with type 2 diabetes will ultimately suffer from kidney failure (National Kidney Foundation, 2014). Research suggests that there is a genetic association between type 1 and 2 diabetes and an increased susceptibility to kidney disease. However, no genetic mutations have been identified that can explain this theory (Bergrem & Leivestad, 2001). Apart from the genetic theories, environmental factors like obesity and poor eating lifestyles can also contribute to cases of diabetes associated nephropathies.

2.4.3 Socio-economic Status

Research suggests that low socioeconomic status is associated with both development and progression of CKD (Safian & Textor, 2001). Socioeconomically disadvantaged populations like countries in the African and Asian continents exhibit a disproportionate burden of CKD which is usually complicated because there are inadequate funds to support proper treatment and research into the disease (Jha *et al*., 2013). Given that CKD requires very expensive drug therapy as well as access to expensive medical equipment, individuals of lower socioeconomic status have a higher risk for mortality and morbidity compared with those of higher socioeconomic status (Bello *et al*., 2008).

2.4.4 Occupational Exposures

There is increasing interest in occupational risk factors for CKD. This stems from observations of elevated CKD prevalence in low and middle income countries, particularly in rural agricultural communities located in Africa and Asia (Bello *et al*., 2008). Similarly, miners and industrial workers who are daily exposed to several hazardous chemicals are known to develop kidney disease due to the exposure to heavy metals like cadmium, chromium and lead as well as organic solvents. Cadmium is most reported to be the main cause of CKD in mining communities (Jmkb *et al*., 2013).

2.4.5 Oxidative Stress

Oxidative stress is defined as the disturbance between the systemic manifestation of reactive oxygen species and antioxidant defences (Betteridge, 2000). Research has shown the link between an excessive build-up of reactive oxygen species and its contribution to the development of diseases like diabetes mellitus and CKD (Baynes, 1991; Stenvinkel, 2003). Free radicals like super oxides and peroxides are produced daily in the body as a result of the metabolic processes like respiration and ATP production. Physiologically, free radicals are formed through the breakdown of reducing sugars, enzymatic reactions like NADPH-Oxidase. Diseases like hypertension, dyslipidaemia and diabetes mellitus are usually associated with oxidative stress which can trigger the immune system to generate an inflammatory response. This can lead to severe cases of accelerated kidney failure (Modaresi, Nafar, & Sahraei, 2015).

2.4.6 Chronic Kidney Disease of unknown aetiology

Some studies have reported cases of CKD that is not associated with any particular risk factor. CKD of unknown aetiology generally affects adults in their fifties and is typically fatal due to the fast rate of disease progression (Weaver, Fadrowski, & Jaar, 2015). Cases of CKD of unknown aetiology have been reported in Asia, several central American countries and some parts of Africa (Almaguer, Herrera, & Orantes, 2014).

2.4.7 Genetic Susceptibility

Research over the past decade has shown a trend where African Americans are 3 times more likely to develop and progress in CKD as compared to their European American counterparts given the same economic and healthcare accessibilities (US Renal System, 2009). This has led to further investigations which had gone to prove that CKD has a heritable or genetic component (Köttgen *et al*., 2009). Since then, major experiments have been conducted on a genome wide association study in order to identify possible susceptibility loci for CKD.

2.5 Genes associated with Chronic Kidney Disease

The genetic cause of many kidney diseases have been shown to be as a result of single-gene defects (Hildebrand, Meyer, & Eyre-Walker, 2010). From various

genome wide association studies, several mutations in different genes have been associated with the risk of different types of kidney disease. They include variants in the Uromodulin encoding gene (UMOD) (Köttgen *et al*., 2009), the human non-muscle myosin heavy chain 9 (MYH9) (Freedman *et al*., 2011; Kao *et al*., 2009; Jeffrey B. Kopp, Winkler, & Nelson, 2011) and the human Apolipoprotein L1 (APOL1) (Genovese *et al*., 2010; Kruzel-Davila *et al*., 2016; Thomson *et al*., 2014a). These genes are all known to be associated with hypertension associated kidney disease, Focal Segmental Glomerular Sclerosis (FSGS) and HIV associated nephropathies. Apart from these, mutations found in NPHS1(nephron), NPHS2 (podocin), LAMB2 (laminin-β2), and PLCE1(phospholipase C epsilon 1) are known to cause childhood onset Steroid-Resistant Nephrotic Syndrome (SRNS), whereas the rare mutations in dominant genes, including actinin-4 (ACTN4) and TRPC6 lead to the adult onset of the disease (Hildebrand *et al*., 2010).

2.5.1 The Non-Muscle Myosin Heavy Chain 9 (MYH9) Gene

The non-muscle myosin heavy chain 9 is a gene which in humans codes for the MYH9 protein (Simons *et al*., 1991). This protein is one part (subunit) of the myosin IIA protein. There are three forms of myosin II, called myosin IIA, myosin IIB and myosin IIC and are all involved in cell movement (cell motility), organelle transport (e.g. nucleus, mitochondria etc.), maintenance of cell shape and cytokinesis (Kopp *et al*., 2008). While some cells use more than one type of myosin II, certain blood cells such as platelets and white blood cells (leukocytes) use only myosin IIA. MYH9 gene is expressed in podocytes as well as in mesangial cells and arteriolar and peritubular capillaries (Arrondel *et al*., 2002). Each type of myosin II protein consists of two

heavy chains and four light chains (Figure 2.2 below). The heavy chains each have two main parts: a head region and a tail region. The head region interacts with actin, a protein which plays a vital role in cell movement and maintenance of cell shape. The long tail region interacts with other proteins, including but not limited to the tail regions of other myosin proteins (Kunishima *et al*., 2014). Mammalian non muscle myosin II molecules comprise three isoforms; A, B, and C. the unique heavy chains are encoded by MYH9, MYH10, and MYH14 respectively.

MYH9 gene is located on chromosome 22 on the long arm at position 12.3 and spans 110 kb of genomic sequence, with 41 exons (or 40 exons if the first noncoding exon is excluded) (Figure 2.3). The main RNA transcript is 7.5 kb in length, encoding a protein of 960 amino acid residues (Jeffrey B. Kopp, Winkler, & Nelson, 2010).

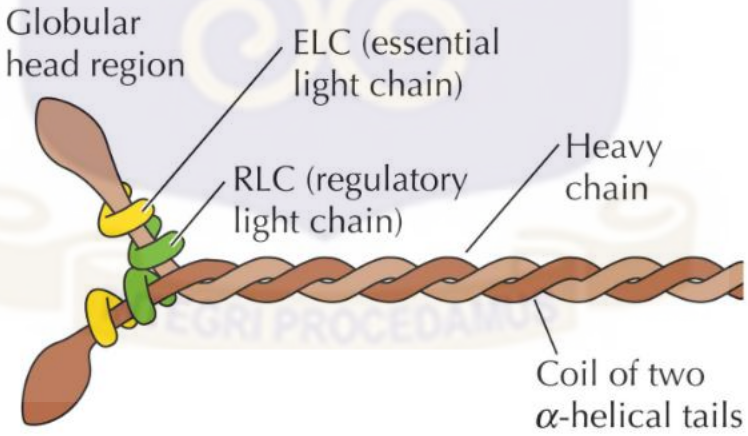

Figure 1.2: Structure of the Myosin II protein. The head region binds the filamentous actin and uses ATP hydrolysis to generate force and "walk" along filament towards the barbed end. The tail region mediates interaction with cargo molecules and or other myosin subunits (THE CELL, Third Edition, figure 11.23, 2000).

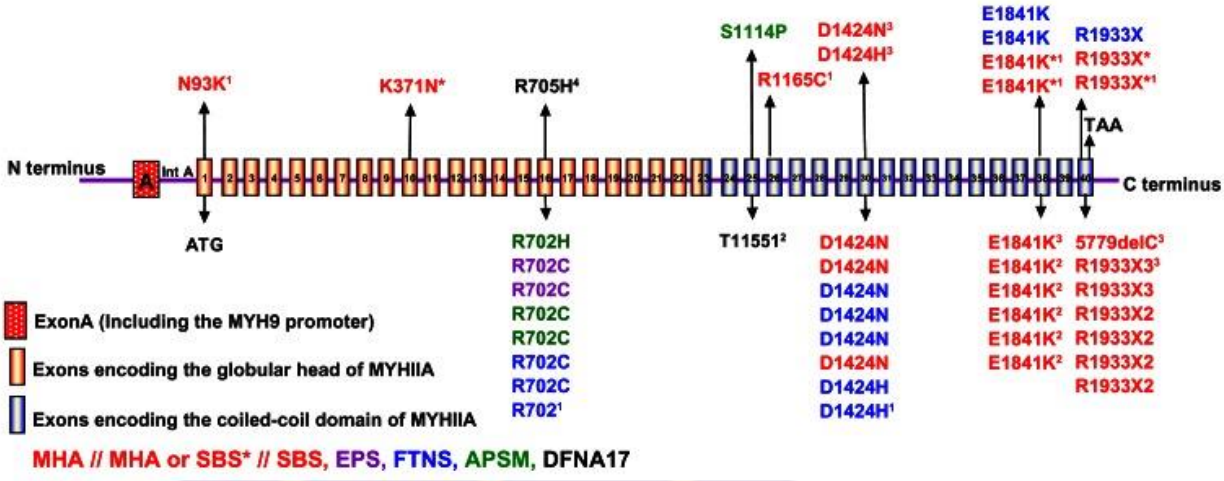

Figure 2.2: Schematic representation of the MYH9 gene Showing the various exons as well as the common mutations that occurs throughout the sequence of the gene (Agrawal, Agarwal, & Naik, 2010).

The MYH9 kidney risk variant is identified by a number of Single Nucleotide Polymorphisms (SNPs), all of which are intronic and are in linkage disequilibrium. Published studies demonstrated a strong association of genetic variants in the MYH9 gene, with ESKD in African Americans (Freedman *et al*., 2009a; Kao *et al*., 2009; Jeffrey B Kopp *et al*., 2008). These studies used mapping by admixture linkage disequilibrium (LD) in African Americans to identify the MYH9 gene, as a single major disease susceptibility locus for ESKD (Behar et al., 2010a). From these studies, four SNPs where identified which defined a risk haplotype, termed the E-1 haplotype (Behar et al., 2010a). In particular, the risk conferred by African ancestry at this genetic locus was only evident for non-diabetic ESKD aetiologies, and was strongest for FSGS, HIV-associated nephropathy (HIVAN) and hypertension associated ESKD. A subsequent study confirmed the strong association of the MYH9 gene with hypertension associated ESKD in African Americans (Freedman *et al*., 2009a). More recently, 10 more SNPs including some SNPs from the S1 haplotype have been

identified as having weaker associations with ESRD (Behar *et al*., 2010a) Despite this, intensive efforts at identifying other functional SNPs that are directly associated with the disease has proven futile (Behar *et al*., 2010b).

2.5.2 The Human Apolipoprotein (APOL1) Gene

The human APOL1 protein belongs to a six membered family of genes known as the Apolipoproteins 1-6 (APOL1-6) (Figure 2.4). The APOL1 family of genes are all located in close proximity to each other as a result of gene duplication and are expressed functionally only in humans, baboons and gorillas (Thomson *et al*., 2014). APOL1 is known to be a minor component of High Density Lipoprotein (HDL) also known as good cholesterol and is usually synthesized in many tissues including the liver, pancreas, kidneys and the brain (Figure 2.5) (Madhavan *et al*., 2011).The gene that encodes the APOL1 protein is 14,522 base pairs long and is located on chromosome 22, on the long arm at position 13.1 from 36,253,070 to 36,267,530 base pairs (Page et al., 2001).

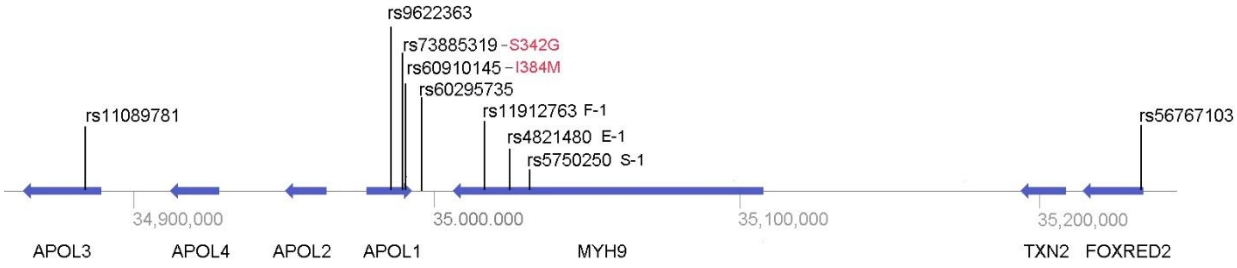

Figure 2.3: APOL1 gene structure. Locations of the G1 allele (S342G and I384M) and G2 allele (NYK388-389K) with other known gene markers including rs9622363 (Limou *et al*., 2015).

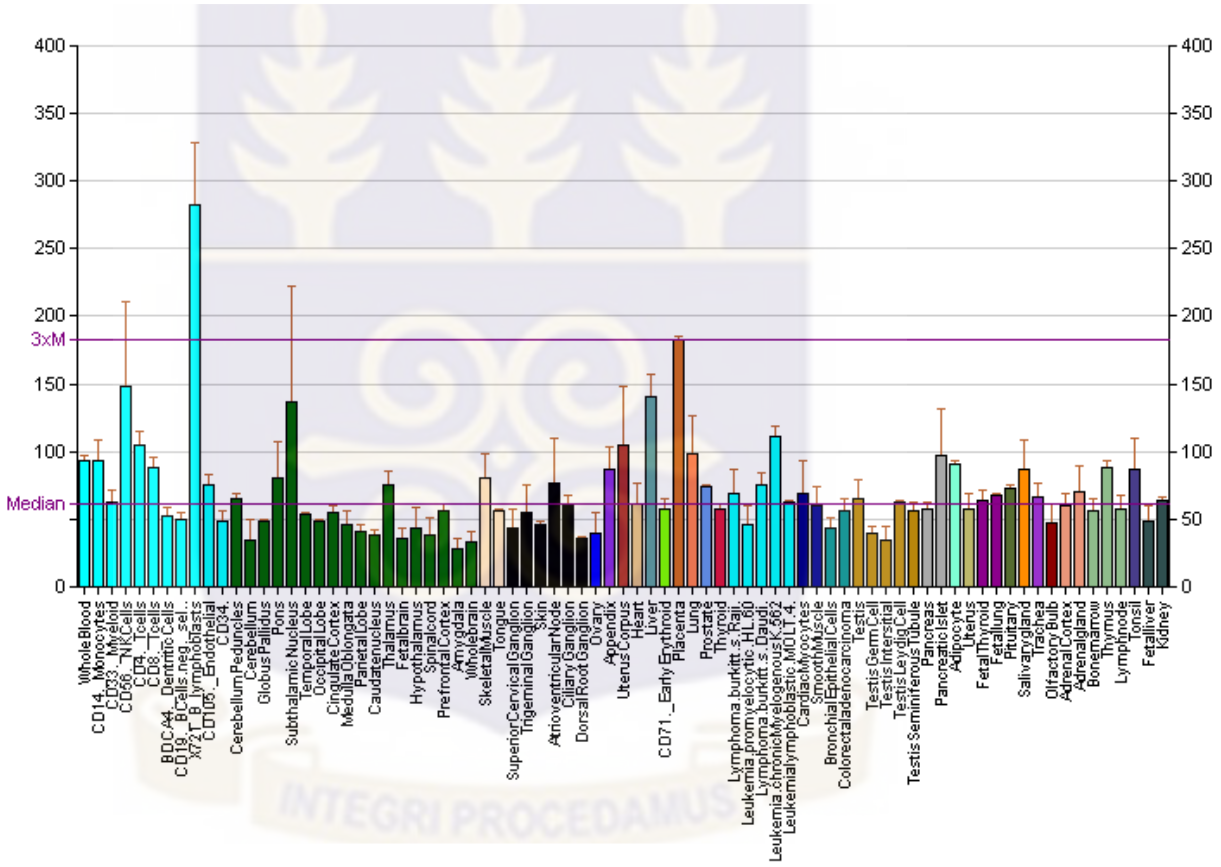

Figure 2.4: APOL1 RNA expression pattern Showing the various tissues that synthesize the APOL1 protein Adapted from Genomics Research Institute of the Novartis Research Foundation (Page *et al*., 2001).

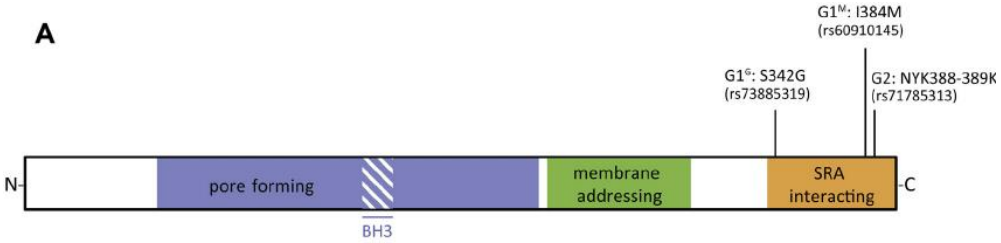

Figure 2.5: APOL1 protein domain structure Locations of the G1 allele (S342G and I384M) and G2 allele (NYK388-389K) indicated in the SRA binding domain at the C terminal (Limou *et al*., 2015).

The APOL1 protein constitutes 398 amino acids and has 5 main functional domains (Figure 2.6);

1. The S domain (also known as the secretory signal)
2. The MAD (Membrane Addressing Domain)
3. The BH3 domain which is associated with programmed cell death
4. The PFD (Pore forming domain)
5. SRA (Serum resistance –associated binding domain) which is known to confer resistance to African *Trypanosoma brucei* (Vanhamme *et al*., 2003).

The APOL gene family evolved by gene duplication in primates, but only humans, gorillas, and baboons retained a functionally expressed APOL1 gene (Limou, Dummer, Nelson, Kopp, & Winkler, 2015). The APOL1 genes are all known to play a role in innate immunity (Page *et al*., 2001; Smith & Malik, 2009) based on research which showed the upregulation of APOL genes by pro inflammatory cytokines (e.g. interferon gamma (INFɣ) and tumour necrosis factor alpha (TNFα)) and their involvement in autophagy and apoptosis (Limou *et al*., 2015; Wan *et al*., 2008; Zhaorigetu *et al*., 2008; Zhaorigetu *et al*., 2011). APOL1 also forms a component of

the High Density Lipoprotein (HDL) when secreted into the blood stream, and is known to form a complex with Apolipoprotein A1(APOA1) and the haemoglobin-binding, haptoglobin-related protein (HPR) (Zhaorigetu *et al*., 2008). APOL1 is a member of the BCL2 genes that are known to play crucial roles in autophagic cell death and has also been shown to single handedly cause autophagy when overexpressed (Wan *et al*., 2008).

Two coding variants of the apolipoprotein L1 gene (APOL1), termed G1 and G2, are strongly associated with non- diabetic forms of ESRD in populations with recent African ancestry. Kidney disorders in the APOL1 spectrum include FSGS, focal global glomerulosclerosis with interstitial and vascular changes (FGGS or hypertension-attributed nephropathy), HIV-associated nephropathy, progressive lupus nephritis, and sickle cell nephropathy (Genovese *et al*., 2010; Lipkowitz *et al*., 2013; Shay Tzur *et al*., 2010c). The APOL1 G1 and G2 renal risk alleles are located in the SRA-interacting domain. G1 is as a result of 2 non synonymous SNPs at position 342 (Ser342Gly) and position 384 (Ile384Met) which are almost in complete linkage disequilibrium whiles the G2 is an in frame deletion of a stretch nucleotides which codes for the amino acid residues at position 388(N) and position 389(Y) (Genovese *et al*., 2010).

The two APOL1 nephropathy variants G1 and G2 are in close proximity such that there is a low likelihood of recombination between the two alleles. As such, the G1 and G2 alleles are never inherited together on the same chromosome (Kruzel-Davila *et al*., 2016). These variants are commonly found in populations with recent African ancestry like African Americans and are known to be very prevalent in the sub-Saharan African region. This unique trend is attributed normally to natural selection.

Individuals with the G1 and G2 nephropathy variants of APOL1 are protected from *Trypanosoma brucei rhodesiense*, a cause of African sleeping sickness in sub-Saharan Africa. Inheriting two APOL1 variants is strongly associated with susceptibility to CKD. The mechanism by which the variants incur injury to the kidney cells leading to kidney injury is still not established (Limou *et al*., 2015).

2.6 Evolution of the Trypanolitic factor

The 2 genetic variants of the APOL1 are known to confer protection on humans and primates against the trypanosome parasites identified to cause the African sleeping sickness. The geographic distribution of the trypanosome subspecies in Africa has evolved over a 1000 years and are usually aligned with APOL1 allele frequencies (Kao *et al*., 2008). The G1 allelic variant is known to have the highest frequency in areas that have high endemicity of trypanosomiasis whereas the G2 allelic variant is more widespread across the West African population (Osafo *et al*., 2015).

APOL1 is the only secreted member of the APOL family that has acquired a signal peptide from a gain-of-function mutation occurring after the APOL1/APOL2 divergence (Smith & Malik, 2009). In the APOL1 gene, the strongest selection pattern in response to certain environmental stressors is found in the C- terminal Serum Resistance-Associated (SRA)-interacting domain which shows the likelihood that this domain plays a key role in the regulation of the gene function. The direct function of the APOL1 protein is not known, however, it is found to be complexed in high density lipoprotein 3 particles which have been identified as the main component of the trypanolytic factor in human serum (Vanhollebeke & Pays, 2006).

Research shows that the presence of the APOL1 risk variants confers protection against Human African Trypanosomiasis (HAT) disease also known as the sleeping sickness by the ability of circulating APOL1 to lyse *T. brucei* (Genovese *et al*., 2010). This occurs when the parasite internalizes the APOL1 containing trypanolytic factor through both phase and receptor mediated endocytosis (as described in Figure 2.7). The progressive acidification of the environment triggers a conformational change in the membrane addressing domain of the APOL1 protein which results in the formation of an ionic channel. The influx of ions into the parasite then causes osmotic swelling and the subsequent bursting resulting in the death of the parasite (Limou *et al*., 2015).

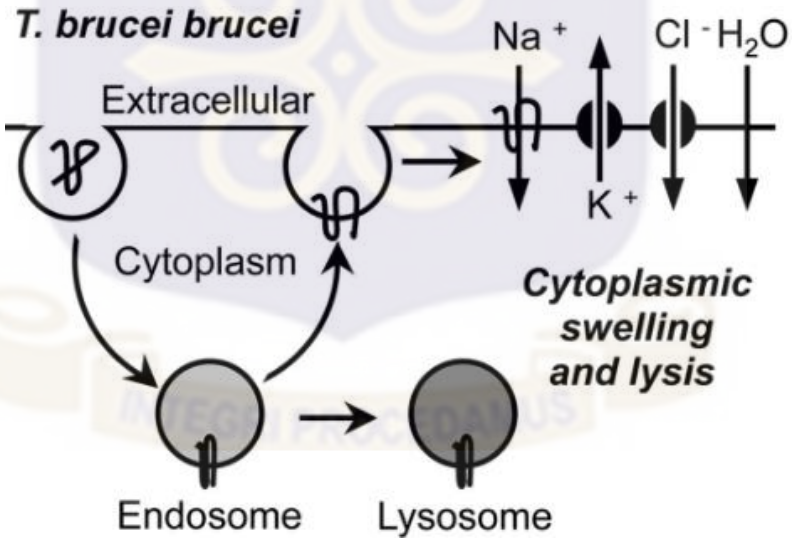

Figure 2. 6: The mechanism of the trypanolitic factor. After the APOL1 protein is endocytosed by the *Trypanosoma brucei* parasite, it encounters the acidic endosomes where it may insert into the endosomal membrane (shading indicates decreasing pH) at an acidic pH of <6. APOL1 forms an essentially nonconductive state (polypeptide backbone represented as a closed coil), but in the case of human serum-susceptible trypanosomes like the *Trypanosoma brucei*, it may generate cation selective channels if it is recycled to the plasma membrane and exposed to neutral pH conditions. Cytoplasmic swelling results from the APOL1- induced sodium influx, as well as chloride influx which causes the parasite to swell up and burst (Russell Thomson et al., 2014).

The rise to high frequency of the risk alleles in West Africa is due to the apparent adaptive advantage provided by the presence of either a G1 or G2 allele against the pathogenicity of *T.b rhodesiense*. However, carrying two copies of the APOL1 risk alleles (i.e. G1/G1, G1/G2, and G2/G2) greatly increases the risk of CKD later in life (Figure 2.8). This means that in areas where sleeping sickness is endemic, mutant APOL1 homozygotes have an added advantage and hence their population becomes increased (Kruzel-Davila *et al*., 2016).

Even though the G1 and G2 variants have the ability to restore trypanolytic effect against *T. brucei rhodesiense*, it does not show its trypanolytic effect on the specie *T. brucei gambiense* and as such, populations that have a high prevalence of the G1 and G2 protein isoforms also happens to be endemic for *T. brucei gambiense* (Limou, Nelson, Kopp, & Winkler, 2014).

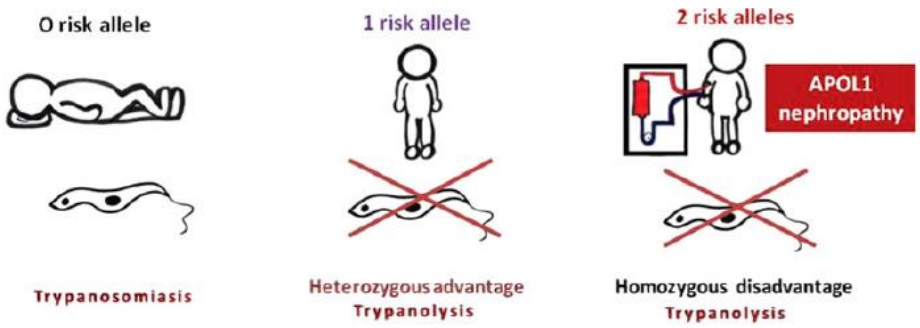

Figure 2.7: Illustration of the effects of APOL1 protein isoforms. Individuals with two APOL1 wild-type alleles or zero risk alleles (G0) are prone to *T.b. rhodesiense* infection or the African sleeping sickness (left). Individuals with only one APOL1 risk allele (G1 or G2) are protected from *T.b. rhodesiense* infection and are at a lower risk of developing APOL1 associated nephropathy (middle). Individuals with two APOL1 risk alleles, are protected from *T.b. rhodesiense* infection but are also at an increased risk of developing progressive adult CKD (right) (Kruzel-Davila et al., 2016).

2.7 Epidemiology of APOL1 and MYH9 Variants

Genetic association studies involving APOL1 and MYH9 gene variants in patients with CKD was first identified in the US mostly in African Americans. Studies now shows that African Americans are 4 times more likely to suffer from CKD compared to European Americans (US renal data system, 2009). This trend was traced back to the African continent where it has been identified that the APOL1 gene evolved to confer resistance to the African Trypanosoma parasite. Currently the gene variants are distributed mainly in the West African region mostly in Ghana, Nigeria and in Senegal (Figure 2.9) (Thomson *et al*., 2014a). Studies also show the association between MYH9 gene to increased susceptibility to CKD in some parts of South America and especially in Hispanic Americans (Behar *et al*., 2010a). Even though studies have been done in parts of Europe to identify the APOL1 and MYH9 variants in relation to CKD, only one MYH9 SNP is known to be associated with an increased risk of non-diabetic CKD in individuals of European ancestry (O'Seaghdha *et al*., 2011).

Interestingly, the East African countries record very low APOL1 variants especially in Ethiopia such that even HIV patients are less likely to die of kidney disease (Thomson *et al*., 2014).

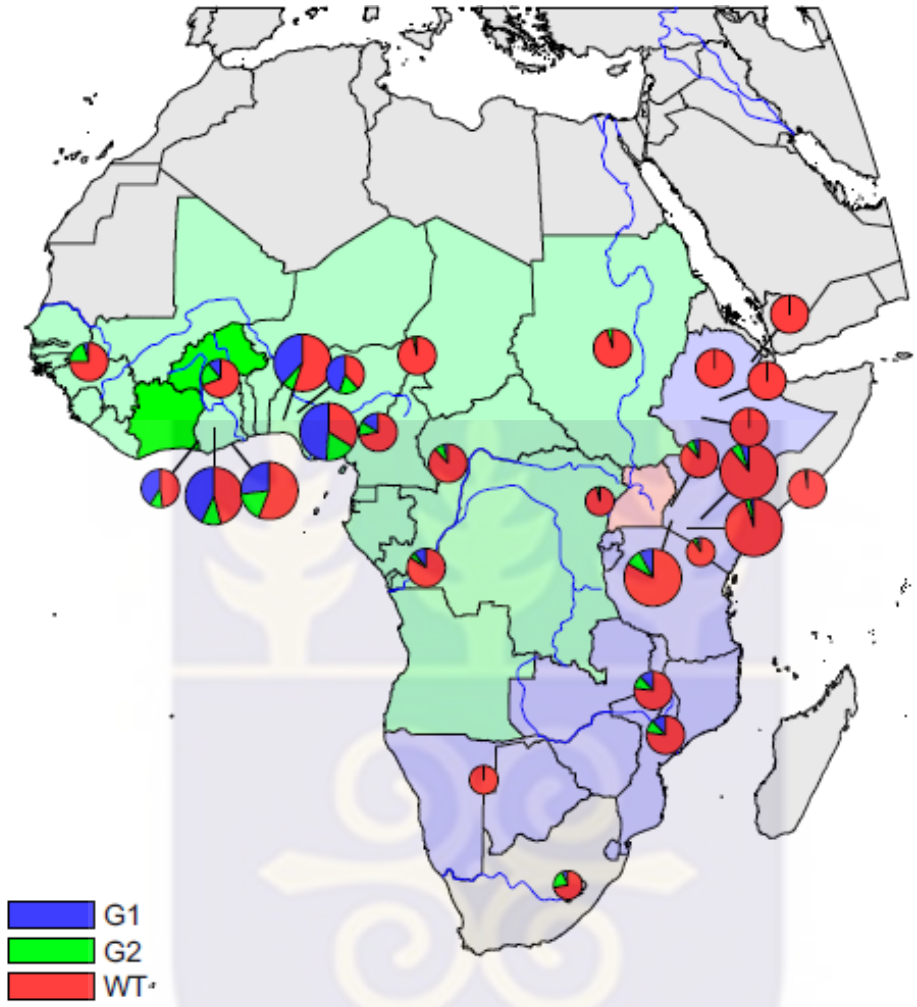

Figure 2.8: Distribution of the G1 and G2 APOL1 variants across the African continent. Allele frequencies G1 and G2 are indicated in blue and green respectively with normal wildtype indicated in red. Darker green shades represent *T. brucei* subspecies 1 and 2, light green shade represent *T. gambiense* subspecies type 1 , pink shade represents both *T. rhodesiense* and *T. gambiense* whiles purple shade represents *T. rhodesiense* subspecies (Thompson *et al*., 2014).

2.8 Glucose-6-Phosphate Dehydrogenase (G6PD)

Glucose-6-phosphate dehydrogenase (G6PD) is a cytosolic enzyme that catalyses the hexose monophosphate shunt also known as the pentose phosphate pathway. This reaction involves the oxidation and decarboxylation of Glucose-6-phosphate to produce a five-carbon membered ribose sugar known as ribulose-5-phosphate and carbon dioxide (Figure 2.10). Studies indicates that the main function of G6PD in erythrocytes is not for glucose metabolism but for the production of the reduced form of NADPH (Berg *et al.*, 2002). G6PD catalyses the reaction by reducing Nicotinamide adenine dinucleotide phosphate (NADP+) to NADPH which is a precursor for the production of Glutathione (GSH). The production of GSH is needed to reduce the level of oxidative stress on the erythrocytes as a result the accumulation of oxidative agents which are by products of metabolism. Because red blood cells lack mitochondria, the pentose phosphate pathway is the only source of production of NADPH and hence the only defence barrier against oxidative agents (Luzzatto, 2011). Red blood cells transport oxygen from the lungs to various parts of the body tissues to enable the cells to adequately perform their functions. The highly reactive oxygen molecules that are carried by the red blood cells can sometimes produce oxygen radicals which can either decay spontaneously or can get converted to hydrogen peroxide (H_2O_2) by the enzyme superoxide dismutase. The peroxide produced is very toxic to the system since it has the tendency to damage the DNA in the cells or lead to cancer. Detoxification of H_2O_2 to H_2O is done by glutathione peroxidase (GSHPX) and catalase (Gaetani *et al.*, 1994). NADPH is therefore very crucial for removal of these reactive oxygen species since it is a structural component of catalase and is also required as a co-enzyme by glutathione reductase which regenerates GSH when it has been oxidized to GSSG by GSHPX.

Glucose-6-phosphate dehydrogenase deficiency is as an effect of genetic variations in the G6PD gene that results in the production of different protein variants with different levels of enzyme activity (Minucci *et al*., 2012). These genetic variations are associated with different kinds of clinical phenotypes most of which can be lethal whiles the others do not usually have severe consequences (Cappellini & Fiorelli, 2008).

G6PD deficiency is the most common enzymopathy in the world affecting over 400 million people worldwide. Currently there are over 300 known genetic variants which causes various degrees of enzyme deficiency (Beutler & Vulliamy, 2002).

G6PD can be found in all cells but varies in concentration from tissue to tissue (Mason *et al*., 2007). A large reserve of reductive potential is present and is more reduced in G6PD deficient erythrocytes which leads to pathophysiological features (Au *et al*., 2000). After G6PD deficiency officially described as a clinical disorder, its phenotypic expression was then noted as heterogeneous (Beutler, 1996). G6PD-deficient cells are vulnerable to oxidative damage because of the low enzyme activity or the total absence of the enzyme that will mean that the amount of NADPH and GSH that is produced is either very low or not enough to counteract the effects of the free radicals that are produced. This can lead to haemoglobin denaturation which results in haemolysis and in some severe cases can cause both liver and kidney damage (Dorgalaleh, Shahzad, & Younesi, 2013). Oxidative stress in people with G6PD deficiency can be triggered by a range of oxidative agents such as some type of infections, certain foods and drugs including anti-malarial drugs (Beutler *et al*., 2007).

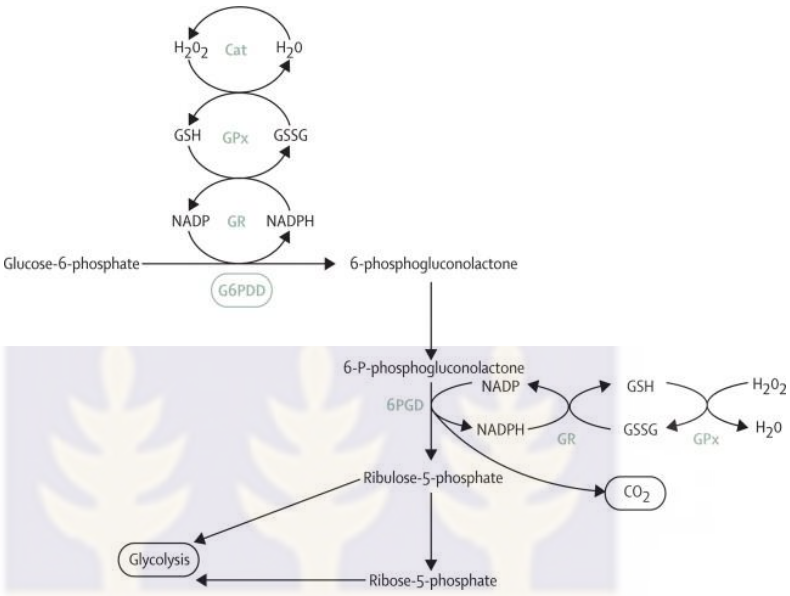

Figure 2.9: Pentose phosphate pathway. NADPH is produced through the action of the enzyme G6PD and 6-phosphogluconate dehydrogenase. It serves as a proton donor for the regeneration of reduced glutathione through NADPH, and as a ligand for catalase (Cappellini & Fiorelli, 2008).

2.8.1 Molecular Structure of G6PD

The G6PD gene which has a size of 16.2kb is encoded on the X chromosome at position q28.The monomer of the protein is made up of 514 amino acids, with a size of 59 kDa in molecular weight (Poggi *et al*., 1990). A model of the three-dimensional structure of the protein was first published in the 90's (Naylor *et al*., 1996) which was followed by the elucidation of its crystal structure in the year 2000 (Figure 2.11) (Au *et al*., 2000). The G6PD protein is an enzyme which is functional either as a tetramer or a dimer under ideal conditions. Each monomer consist of 2 main domains namely

the N terminal domain made up of amino acids at position 27 to 200, with a β-α-β domain dinucleotide binding site consisting of amino acids 38 to 44 and finally a β+α domain, made up of antiparallel nine- stranded Beta sheets (Au *et al*., 2000). The two domains are linked together by an alpha helix which contains the completely conserved eight-residue peptide which functions as a substrate binding site (Mason, 1996). Multiple sequence alignment of the 35 known amino acids sequences shows a 30% identity between the sequence of humans and other species. The protein structure at 3A (0.3nm) resolution reveals a $NADP^+$ molecule in all the subunits of the tetramer which is distant from the active site but close to the dimer interface (Au *et al*., 2000). The normal activity of the G6PD enzyme is dependent on the stability of the active quaternary structures.

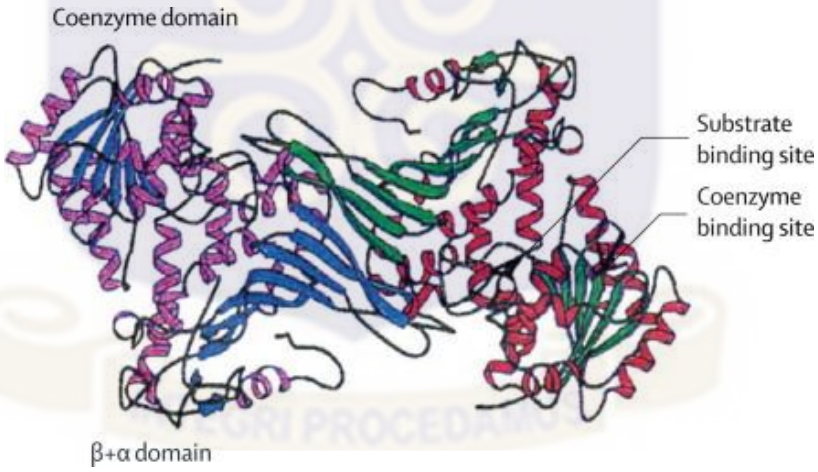

Figure 2. 10: Three dimensional model of an active G6PD dimer. Structure modified from the human G6PD model as proposed by Naylor and colleagues. This model illustrates the substrate binding site, where the mutation of the enzyme mostly affects causing the inability for the enzyme to catalyse the reaction (Capellini *et al*., 2008).

2.8.2 Glucose-6-Phosphate Dehydrogenase terminology and variant classification

The World Health Organization (W.H.O) has made recommendations for the classification of the various types of G6PD deficiency across all the different continents. The G6PD deficiency was initially characterized using biochemical tests, electrophoretic mobility and by measuring the residual enzyme activity (Daoud et al., 2008). Currently about 400 biochemical variants of G6PD have been identified and the polymorphisms are largely demarcated to specific geographical locations (Amoah *et al*., 2016). These variants were characterized by different methods such as thermostability, kinetic variables and chromatographic behaviour (Beutler, 1996).

The officially accepted abbreviation for Glucose-6-phosphate Dehydrogenase is G6PD (E.C. 1.1.1.49). The terms G6PD deficient and G6PD normal are used to assign the specific phenotypes G6PD (-) and G6PD (+). Because G6PD is X- linked, only males can express the normal (G6PD +) or deficient (G6PD -) hemizygous genotype whiles females can either be normal homozygous [G6PD (+, +)], deficient homozygous [G6PD (-,-)] or heterozygous [G6PD (+,-)] (Poggi et al., 1990). Most often females who are heterozygous for the G6PD deficiency are also called "intermediates" because their level of enzyme activity is between the normal and deficient range. As a result, the enzyme deficiency is said to be a co-dominant trait rather than a recessive trait.

According to WHO, the different types of G6PD enzyme deficiency is grouped into 5 main classes (Table 2.3) depending on the severity of the resulting phenotype that is expressed, with Class I being the most severe and Class V being the least severe (WHO, 2008).

Table 2.3: Classes of G6PD variants

Class	Clinical Manifestation	G6PD activity % of normal activity	Number of known mutants	Example	Comments
IV	None	>85	2	A, B	Normal wild type
II / III	Asymptomatic in the steady state	< 30	75	Med, A-, Orissa, Mahidol, Canton, Vanua Lava, Seattle	Most of these variants are known to be polymorphic
I	CNSHA, acute exacerbations	< 10 in most cases	61	Sunderland, Nara, Guadalajara	Never polymorphic; but same mutation can recur

2.8.3 Genetics and Molecular Basis of G6PD Deficiency

The gene that codes for the G6PD enzyme is found in the liver, brain, kidney, adrenals, skin, pancreas, nerve and muscle cells, hence the extra-erythrocytic manifestations of G6PD deficiency. The G6PD gene is located on the telomeric region of the long arm of the X chromosome (band Xq28), which is not far from the location of the gene that codes for haemophilia A, colour blindness and congenital dyskeratosis (Trask, Massa, Kenwrick, & Gitschier, 1991). The gene has 13 exons and 12 introns which spans about 20kb in total (Figure 2.12). The 5' untranslated region of the mRNA is made up of genes from exon 1 and part of exon 2 and has a promoter region which shares similarity with other housekeeping gene promoters (Toniolo et al., 1991). All mutations of the G6PD gene that results in enzyme deficiency affects the coding sequence of the gene (Vulliamy *et al*., 1991). Even though the promoter region has been extensively characterised using band shift assays, no mutations have yet been reported to be associated with the enzyme deficiency even though *in vivo* mouse models have shown that mutations of the GC

boxes can greatly affect the transcriptional activity (Philippe *et al*., 1994). It is notable that many single point mutations in the G6PD gene have been documented several times in different parts of the world, which seems to suggest that their origin is not likely to be from a common ancestor and are probably new mutations that have arisen independently (Vulliamy *et al*., 1998).

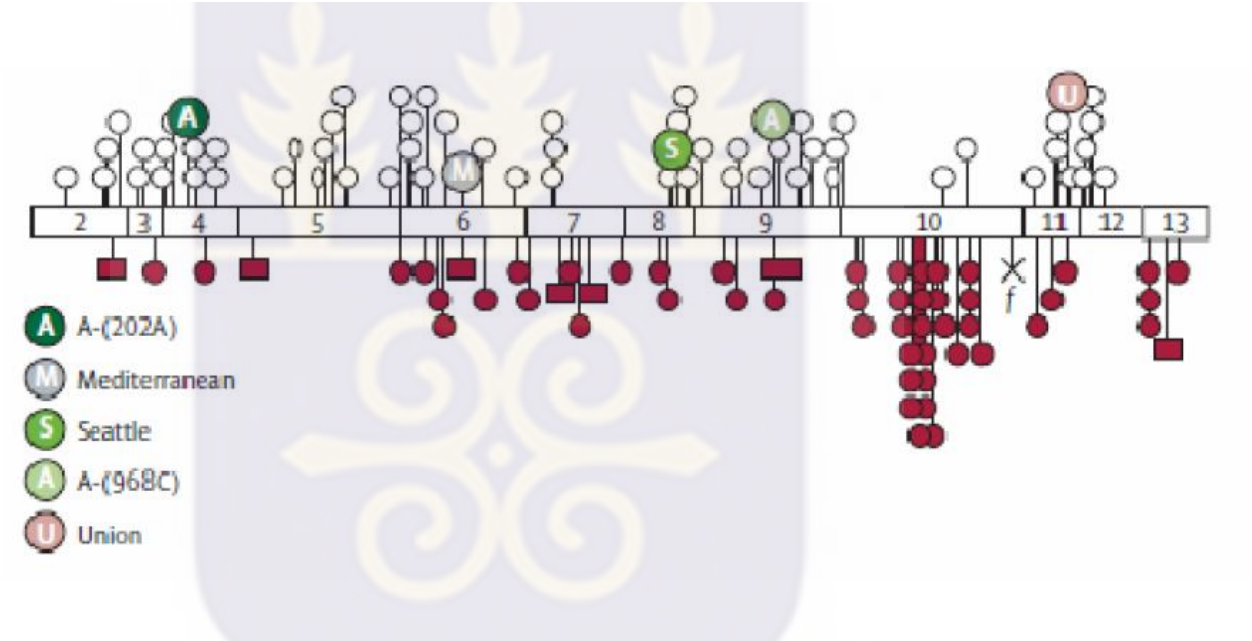

Figure 2.11: Common mutations along the G6PD gene coding sequence. Exons are showed as open numbered boxes, Open circles are mutations causing the classes II and III variants. Filled circles represent sporadic mutations giving rise to severe variants (class I) .Open ellipses are mutations causing class IV variants. Filled squares are small deletions. Cross are nonsense mutations, *f* is a splice site mutation. 202A and 968 are two sites of base substitution in G6PD A^- (Luzzato et al., 2001).

In several instances, mutations that cause G6PD deficiency act by decreasing the *in vivo* stability of the enzyme, thus significantly reducing the enzyme activity. In other cases, exchange in amino acid position can also directly affect the active site thereby reducing the function of the enzyme. This results in the disruption of the active site or

translocating the enzyme to a place it may not be needed (Manson et al., 2007). The various G6PD variants are defined based on their amino acid mutations. For instance, G6PD A^- is sometimes written as Val68Met [or 68 Val-Met] which means that the mutation has methionine residue at position 68 instead of the normal valine residue. Also G6PD Mediterranean Ser188Phe [or 188 Ser-Phe] which means that the mutation has Phenylalanine residue at position 188 instead of the normal Serine residue. A more strictly correct and unambiguous way of identifying mutations is by their base (nucleotide) change at the genomic level. For example the A^-, 202 G-A mutation means that the mutation has an Adenine at position 202 instead of the original Guanine. [Med 563 C-T] on the other hand means that the Mediterranean variant also has a Thymine substituted for Cytosine at position 563 (Manson *et al*., 2007).

Apart from the common missense mutations that are known to cause G6PD enzyme deficiency, several polymorphic sites have been identified in the introns enabling the definition of G6PD haplotypes (Maestrini *et al*., 1992; Vulliamy *et al*., 1991). These haplotypes have been used in an attempt to understand the evolutionary history of the G6PD gene especially in association with malaria selection using linkage disequilibrium (Luzzatto, 2006).

2.8.4 Epidemiology of G6PD gene variants

G6PD deficiency is the most common human enzyme defect that has resulted in 3,400 deaths in 1990 and 4,100 deaths in 2013 (Global burden of disease, 2013). It occurs with increased frequency throughout Africa, Asia, the Mediterranean and the Middle Eastern regions (Figure 2.13). In the United States, it is nown to mostly affect male

African-Americans at a higher rate than any other ethnic group with a prevalence of about 10% which is the lowest in the world. The prevalence in Africa and in Asian, the Middle East and the Mediterranean regions are reported to be over 20% (Howes *et al*., 2012). Molecular characterization data have been mostly lacking for Africa although more are becoming available most recently, for example, South-West Nigeria records a prevalence of 28.1% (May *et al*., 2000), 22.5% in Congo (Bouanga *et al*., 1998) and 12.4% among asymptomatic malaria school children living in Ghana (Amoah *et al*., 2016).

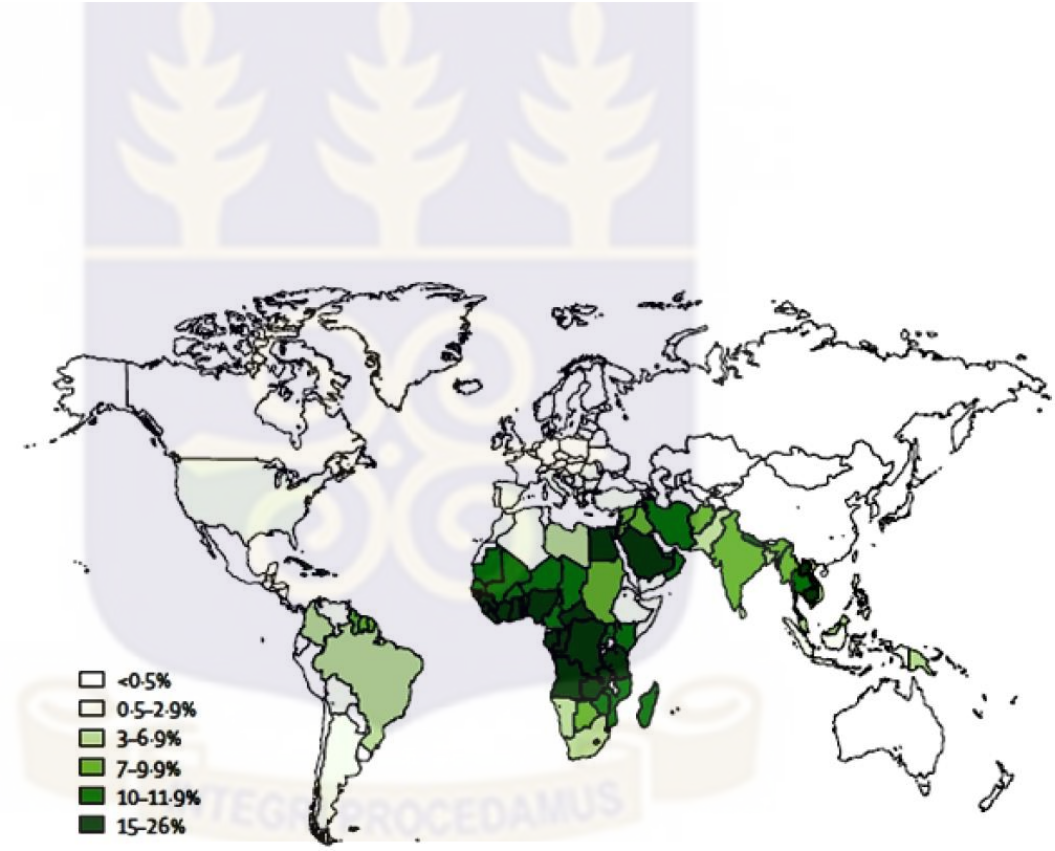

Figure 2.12: Worldwide distribution of G6PD deficiency (WHO working group, 1989).

2.9 Association between G6PD Deficiency and Malaria

The geographical areas that record high prevalence for the enzyme deficiency are also known malaria endemic regions. In recent years, it has been reported in several regions that the worldwide distribution of malaria is unusually similar to that G6PD deficient areas. It is no surprise therefore that the G6PD deficiency is widespread across the African continent. This has led to the theories that the G6PD mutations evolved to confer resistance to some cases of malaria in these regions (Amoah *et al*., 2016; Amoako *et al*., 2014; Maiga *et al*., 2014; Ouattara *et al*., 2014). Despite the fact that research now supports the claims that the G6PD gene is under a strong natural selection, evidence also suggests that the nature of protection and the resulting genotypes associated with it has yielded conflicting results most of which are skewed towards a particular gender (Howes *et al*., 2012; Luzzatto, 2011; Tishkoff, 2001).Studies also shows that these mutations are maintained in high frequency in these regions despite the hemopathologies that they cause (Tishkoff, 2001).

2.10 Clinical Manifestations of G6PD Deficiency

Most individuals who have the different forms of G6PD deficiency are usually asymptomatic and only show symptoms when exposed to the common triggers like Sulphur based drugs, certain foods and some types of infection. The effects of the enzyme deficiency is usually manifested as acute haemolysis and then gradually generates into other complications like, severe and acute anaemia, liver and kidney damage. G6PD enzyme deficiency however has not been reported to have any significant effect on the life expectancy, quality of life and the normal activity of

affected people (Cocco, 1998) unlike hemoglobinopathies (e.g. Sickle cell disease and thalassemia).The mechanism by which the known triggers initiate the process of haemolysis is not fully known, however, it is characterized by clear symptoms like fatigue, back pain, anaemia and jaundice (Edwards, 2002). Certain patients also have increased levels of unconjugated bilirubin, lactate dehydrogenase and recticulocytosis.

2.11 Drug-induced Haemolytic Anaemia in the Treatment of Malaria

In the treatment of malaria, G6PD deficiency has been found to be a major concern to individuals with the enzymopathy. Drugs like Primaquine, Sulphadoxine-Pyrimethamine (SP) which is recommended for Intermittent Preventive Treatment during Pregnancy (IPTp) and Dapsone which is used in different combination therapies for the treatment of malaria have been known to cause serious haemolytic events (Fanello *et al*., 2008; "WHO | World Malaria Report 2008," 2013). Because of this complication, the Ghana National Drugs Program (2004) institutes that pregnant women who are G6PD deficient or are allergic to sulphur-containing drugs are advised against using Sulphadoxine-Pyrimethamine for malaria treatment. As such pregnant women are normally required to be screened for G6PD deficiency before treatment for malaria using SP is initiated. In Ghana, there is a high rate of self-treatment medications which can easily be purchased in the drug stores and as such, G6PD deficient individuals can suffer cases of haemolysis without knowing the exact cause since most of them are unaware of their status.

2.12 Genetic Associations and the Risk of Disease

The widespread accessibility of high-throughput genotyping technology has gradually paved way to the age of personal genetics, which takes into consideration the use of genetic variations to predict individual susceptibility to common diseases. This notwithstanding, easy access to commercial personal genetics services, our understanding of the genetic architecture of common diseases is still very restricted and has not been able to adequately predict people who are at higher risk of certain diseases. This is partly because of the complexity of the mapping relationship between genotype and phenotype that is a consequence of epistasis (gene-gene interaction) and other phenomena such as gene-environment interaction and locus heterogeneity (Moore & Williams, 2009).

Africa is one of the most ethnically and genetically diverse regions of the world. It is thought to be the ancestral homeland of all modern humans, and is the homeland of millions of people of the recent African diaspora. Because of the central role of African populations in human history, characterizing their patterns of genetic diversity and linkage disequilibrium is crucial for reconstructing human evolution and for understanding the genetic basis of complex diseases (Tishkoff, 2001). By identifying genetic variations and associations between different genetic mutations among populations we gain a better understanding of the differences in susceptibility to disease, differential response to certain drugs and will also help in understanding the complex genetic interactions and environmental factors that produces the different phenotypes that are expressed across different geographical regions especially in the African sub region where there is high surge in the incidence of different kinds of infectious and non-infectious diseases.

The pattern of genetic variation in the African sub region is largely influenced by geographic history. Interestingly, the G6PD, APOL1 and MYH9 variants all record high frequencies in similar geographical locations and are also known to be as a result of evolutionary selection which confers protection to infectious diseases (malaria and Trypanosomiasis) whiles increasing the risk of another form of non-infectious diseases (Sickle cell anaemia, Kidney disease). Even though there is enough research that supports the theories of selection of these mutations and their association with an increased susceptibility to other diseases, no study has been done to determine the genetic associations and gene-gene interactions between these 3 genes to determine their level of interdependence on each other. In the Asian population which also shares similarities in the geographical region as well as disease prevalence with Africa, it has been reported that there is a high prevalence of G6PD deficiency in CKD patients of unknown aetiology (Jmkb *et al*., 2013) but no genetic association study was conducted to determine if these genes play a role in the increase in susceptibility to CKD especially in people who are G6PD deficient.

Even though these APOL1 and MYH9 genetic variants are located on the same chromosome and hence is possible to identify genetic association between the two, the G6PD gene is located on the X chromosome and might not always be inherited together with the APOL1 and MYH9 gene mutations. However, studies now show evidence of associations between genes located on different chromosomes. Such circumstances may be as a result of physical linkage between the different loci that lowers the rate of recombination, natural selection and genetic drift. Typical example of such genetic associations between genes from different chromosomes is the genetic association between APOL1, MYH9 genes on chromosome 22 and Sickle Cell mutations found in chromosome 11 which is known to increase the risk of Sickle cell

nephropathy (Ashley-Koch *et al*., 2011). The identification of additional genes that contributes to CKD risk in Africans could reveal strong and reliable association with non-diabetic nephropathy, specifically kidney disease of unknown aetiology. Screening tests and novel therapies derived from these genetic associations may offer new hope for slowing the worldwide epidemic of kidney disease.

CHAPTER THREE

MATERIALS AND METHODS

3.1 Study Design and Study Sites

This is a case-control study which is part of an ongoing NIH funded project (H3Africa Kidney Disease Study, NIH project number 1 U54 HG006939) to perform genotyping analysis of known disease variants in the APOL1 and MYH9 genes, as well as to perform a Genome Wide Association Study (GWAS) in a cohort of Africans with CKD of varying aetiologies, in order to establish a genetic loci associated with CKD. The design of this project was therefore in direct lineage with the Umbrella Project. Participants of this study were considered with an unmatched standardized evaluation of risk factors and clinical outcomes.

Samples were collected mainly from Ghana and Nigeria since these 2 countries have an increasing incidence and prevalence of CKD and as such served as valid study sites for the collection of samples (Friedman, 2014). Samples used were collected between the periods of 2014 to 2016 from 7 different clinical centres across the two participating countries (shown in table 3.1).

Table 3.1: Number of samples recruited across the various clinical centres

Country	Clinical site	Number of samples
Ghana	Korle- bu Teaching Hospital, Accra	100
Ghana	Kwame Nkrumah University of Science and Technilogy, Kumasi	100
Nigeria	Obafemi Awolowo University ,Ile-Ife	80
Nigeria	University of Abuja Teaching Hospital, Abuja	50
Nigeria	University of Ibadan, Ibadan	50
Nigeria	University of Nigeria, Enugu	21
Nigeria	University of Ilorin, Ilorin	10

3.2 Selection of Study participants

The sample size was calculated using the formulae for calculating sample sizes for case-control studies. This formulae is based on the prevalence of CKD in Ghana and Nigeria which are 19% and 15.3% respectively (Osafo *et al*., 2011). Based on the sample size calculation using the prevalence of 19% and an 80% power to detect an odds ratio of 3, a sample size of 136 (68 cases and 68 controls) was required for this study (detailed calculation presented in the appendix). Four hundred and eleven (411) out of the 4000 archived blood samples from CKD and non-CKD patients who were recruited for the H3Africa project were selected for this study. The 411 samples consisted of 202 CKD cases of unknown aetiology and 209 non-CKD controls who met the inclusion criteria.

3.2.1 Inclusion and Exclusion criteria

Cases were defined as patients who were clinically diagnosed to have CKD of unknown aetiology, presence of proteinuria on two occasions, absence of diabetes mellitus, hypertension associated nephropathy and Urinary Tract infections (UTI). Inclusion criteria for samples chosen as controls were healthy individuals attending Out Patients Department (OPD), persons with no history of CKD or any kidney related disease and a GFR >60ml/min/1.73m^2. Pregnant or breast-feeding women, institutionalized persons (prisoners, nursing homes etc.) and people who suffer from cancer of the kidney were all excluded from the study.

3.3 Ethical Consideration

Ethical clearance was sought from the Institutional Review Board (IRB) of the Noguchi Memorial institute for Medical Research (NMIMR) at the University of Ghana, Legon. Each volunteer signed or thumb-printed an informed written consent after the research had been explained to them in a language they understood. Protocols used for the various experiments were in line with the ethical standards of NMIMR.

3.4 Sample Preparation and DNA Isolation

5ml of venous blood was taken from each participant by an experienced phlebotomist in a standard venepuncture into EDTA vacutainer tubes and labelled with unique identification codes. Samples were then transported to the laboratory on ice and stored in the −80°C freezer.

Blood samples meant for DNA isolation were centrifuged at 3000 rpm for 15 minutes and separated into Plasma and buffy coat which were stored separately at −80°C using unique identification numbers awaiting DNA extraction.

Genomic DNA was extracted from buffy coat samples isolated from the whole blood using QIAamp DNA midi kit (Qiagen) according to the manufacturer's instructions with slight modifications. The Qiagen DNA extraction kit is based on the silica-membrane technology. The DNA has a high affinity for the silica membrane and as such binds to the membrane whiles other nucleases and cellular debris are washed off. The DNA is then eluted using a ready-to-use elution buffer. This procedure produces high-purity genomic DNA suitable for most molecular biology and clinical research applications, such as restriction digestion, ligation, labelling, amplification, and radioactive or fluorescent sequencing (Ausubel *et al*., 1991). Two (2) ml of buffy coat was pipetted into a 15 ml centrifuge tube (falcon tube) and 200 ul of proteinase K solution and 2 ml of Lysis buffer (Buffer AL) were added to the blood sample. The mixture was then gently mixed by inversion for 1 minute and vortexed to completely mix all the contents. It was then incubated in waterbath at 70°C for 10 minutes for total lysis and degradation of cells and protein in the mixture. The sample was removed from the waterbath and 2 ml of absolute ethanol of molecular grade was added to the resulting solution and mixed by pulse-vortexing for 1 minute. The sample was pipetted into a clean and sterile spin column and centrifuged at 3500rpm for 3 minutes. The filtrate after centrifuging was discarded into 10% bleach solution whiles the 2mls of AW1 wash buffer was pipetted into the spin column and centrifuged at 3500rpm for 6 minutes. The washing step was repeated with AW2 wash buffer and centrifuged at 3500rpm for 25 minutes. After the washing step was completed, the spin column was placed into a fresh and sterile collection tube and

300ul of elution buffer was added and incubated on ice for 5 minutes before centrifuging at 3500rpm for 6 minutes. Eluted DNA was pipetted into a clean and sterile Eppendorf tube and stored at 4°C for downstream molecular work.

3.5 DNA Quality Control Analysis

3.5.1 Agarose Gel Electrophoresis

To determine the quality of the extracted DNA, agarose gel electrophoresis was employed. The agarose gel electrophoresis is a method used in biochemistry, molecular biology, genetics, and clinical chemistry to separate a mixed population of macromolecules such as DNA in a matrix of agarose. The DNA is separated by applying an electric field to move the charged molecules through an agarose matrix, and the biomolecules are separated by size in the agarose gel matrix. In the case of genomic DNA, they will move towards the positive end because the DNA itself is negatively charged (Sambrook & Russel 2001).

To prepare a 1% agarose gel, 4 grams of Agarose powder (Cleaver scientific Ltd., U.K) was weighed into a graduated beaker containing 400 ml of 1X Sodium Borate (SB) buffer. The mixture was heated till all the agarose powder completely dissolved and allowed to stand on the bench to cool for about 3 minutes. After cooling, 4 ul of Ethidium Bromide (Sigma, USA) was added to the solution and swirled gently till it completely mixed with the solution. The agarose was casted by gently pouring the melted agarose into the gel tank with inserted combs and allowed to cool at room temperature till solidified. After solidification, gel combs were removed from the gel 5 ul of DNA was mixed with 5 ul of loading dye (bromophenol blue) and loaded unto

the wells in the agarose gel immersed in 1X SB buffer and run at 240 V for 10 minutes. The resolved DNA was then visualized using a UV-trans illuminator.

3.5.2 DNA Quantitation using Qubit Fluorimeter

To determine the quantity of DNA extracted from blood samples, the qubit fluorimeter was used. The Qubit fluorimeter employs the use of a fluorescent dye such as Pico green to determine the concentration of nucleic acids and proteins in some cases in a sample. The dye has a very low fluorescence until they bind with their target, in this case the DNA. Once the dye is bound, they become intensely fluorescent. The difference in bound and unbound dye is so wide such that, they record almost no fluorescence until bound to DNA. Upon binding, the dye intercalates between the bases of the DNA which then assumes a more rigid shape and becomes highly fluorescent. The dye can bind to DNA within seconds of being added to the DNA and reaches equilibrium under 5 minutes which can be measured at wavelength of 260 nm (McKnight et al., 2007). The concentration of DNA is then calculated as the amount of double stranded DNA that is bound to the fluorescent dye in 1ul of buffer (ng/µl).

The Qubit fluorimetric assay was performed using the Quant-iT dsDNA HS assay as described by Hamza *et al*., (2009) with slight modifications. Standards and assay buffers were prepared according to the manufacturer's protocol (Table 3.2). One hundred and ninety-nine (199) ul of assay buffer is pipetted into a clean and sterile PCR tube and labelled appropriately. One (1) ul of the DNA sample is added to the PCR tube and gently vortexed to completely mix the solution. The PCR tubes are then incubated in the dark for 5 minutes and the absorbance read using the Qubit Fluorimeter.

3.6 Amplification of G6PD

The coding regions of the G6PD gene encompassing A and A- variants were amplified by Polymerase Chain Reaction (PCR) using oligonucleotide primer sets (Table 3.3) appropriate for the amplification of the G6PD exons 4 and 5 using already established protocols and cycling conditions (Bouanga *et al*., 1998; Daud *et al*., 2008) with slight modifications. A 25µl amplification reaction mixture for the G6PD conventional PCR contained 4µl of template DNA, 2.5µl of 10X buffer, 0.3µl of 25mM $MgCl_2$, 0.12µl of Taq polymerase enzyme (New England Bio labs, USA), 0.3µl of each primer and 0.25µl of a 10mm DNTPS's (New England Bio labs, USA) as well as 6.83ul of Sterilized Double distilled water (Amresco, U.S.A) (Table 3.4).

After 10 min of incubation at 94°C, a total of 30 amplification cycles were run for G6PD variants A and A- primer sets under the following PCR conditions: For G6PD A, denaturation at 94°C for 45s, annealing for 61°C for 45s, elongation at 72°C for 45s and a final elongation at 72°C for 7 min (Figure 3.1). For G6PD variant A-, the PCR conditions was follows: denaturation at 94°C for 45s, annealing for 65°C for 45s, elongation at 72°C for 45s, and a final elongation at 72° for 7min (Figure 3.2). PCR products obtained was resolved on a 2% agarose gel to confirm a band size of 273bp for variant A and 202bp for variant A- to indicate a successful amplification of the G6PD gene.

Table 3.2: Volumes of reagents for preparing buffers for standards for reaction assay

	Standard Assay tubes	User sample assay tubes
Volume of working solution	190 µl	199µl
Volume of standard(from kit) solution	10 µl	0
Volume of user solution	0	1 µl
Total volume in each assay tube	200 µl	200 µl

Table 3.3: Primer sets utilized for amplification of the encoding regions of G6PD gene for variants A and A-

rs number	Name	Forward Primer	Reverse Primer
rs1050829	G6PD A	ACCCCAGAGGAGAAGCTCA	CGTGAATGTTCTTGGTGACG
rs1050828	G6PD A-	TGACCTGGCCAAGAAGAAGA	GAAGGGCTCACTCTGTTTGC

Table 3.4: Reagents with concentrations and the volumes used in preparing Master Mix for amplifying G6PD SNPs.

Reagents in reaction mix	Initial concentration	Volume/ µl	Final concentration in 15µl reaction mix
Sterilized Double Distilled Water		6.83	
PCR Buffer	10x	1.5	1X
$MgCl_2$	50mM	0.3	0.5mM
DNTP	10mM	0.25	0.17mM
Forward Primer	10µM	1.0	0.1nM
Reverse Primer	10µM	1.0	0.1nM
Taq Polymerase	5U/µl	0.12	1U
DNA		4	

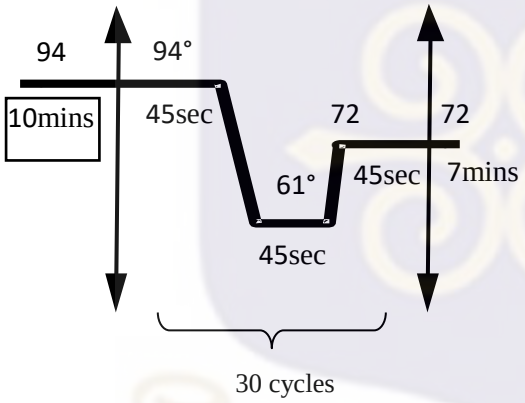

Figure 3.1: Cycling conditions for G6PD A showing the various cycling temperatures at each stage of the PCR

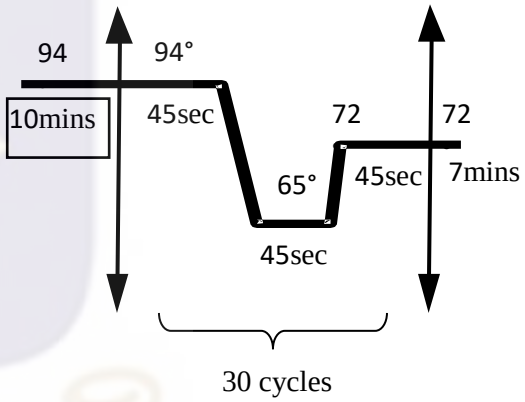

Figure 3.2: Cycling conditions for G6PD A$^-$ showing the various cycling temperatures at each stage of the PCR

3.7 Restriction Fragment Length Polymorphism analysis for G6PD SNPs

After successful amplification of the coding regions of the G6PD gene encompassing A and A- variants, the amplicons were screened for the G6PD deficiency mutations using restriction enzymes appropriate for such mutations according to the manufacturer's protocol with slight modifications. The restriction endonuclease *FokI* recognizes the sequence of nucleotides that make up A376G mutation leading to the *G6PD* A variant whiles *NlaIII* recognizes the sequence of nucleotides that make up the G202A variant (Table 3.5). Both restriction enzymes were obtained from the New England Bio labs, U.S.A. For a 25 ul reaction mixture, 5 ul of PCR products was mixed with 5units of restriction enzyme, 2.5ul of 10X enzyme buffer and 17 ul of sterile nuclease free distilled water. The mixture was incubated at 37°C for 1 hour after which the reaction was terminated through heat inactivation at 65°C for 20 min in a waterbath. The resulting fragments were resolved on a 3% agarose gel precast with ethidium Bromide and visualized using a UV-trans illuminator. For G6PD variant A, successful digestion of the PCR product to produce 2 band sizes at 155bp and 117bp indicated the presence of the SNP. For G6PD variant A-, successful digestion of the PCR product to produce 2 band sizes at 115bp and 69bp indicated the presence of the SNP.

3.8 Amplification of APOL1 and MYH9 gene

The coding regions of the APOL1 and MYH9 genes were amplified by a multiplex PCR using oligonucleotide primer sets designed to specifically amplify the section of the MYH9 and APOL1 gene containing the SNPs (Table 3.6). A 25 ul amplification

reaction mixture for the APOL1 and MYH9 multiplex PCR contained 4 ul of template DNA, 2.5 µl of 10X buffer, 0.3 µl of 50mM $MgCl_2$, 0.12 µl of Taq polymerase enzyme (New England Biolabs, USA), 0.3 µl of each primer and 0.25 µl of a 0.5 mm DNTPS's as well as 6.83 µl of Sterilized Double distilled water (Amresco, U.S.A) (Table 3.7). After 5 min of incubation at 95°C, a total of 40 amplification cycles were run for APOL1 and MYH9 primer sets under the following PCR conditions: denaturation at 95°C for 30s, annealing for 65°C for 20s, elongation at 72°C for 1min and a final elongation at 72°C for 7 min (Figure 3.3).

Table 3.5: Restriction enzymes for digestion of G6PD PCR amplicons

Enzyme	Sequence	Recognition Site
NlaIII[a]	CATG	5'…..CATG▼…..3' 3'…..▲GTAC…..5'
FokI[b]	GGATG	5'….GGATG(N)$_9$▼….3' 3'….CCTAC(N)$_{13}$▲….5'

[a] restriction endonuclease from an E Coli strain that carries the *NlaIII* gene from *Neisseria lactamica* (NRCC 2118).

[b] restriction endonuclease from an E Coli strain that carries the *FokI* gene from *Flavobacterium okeanokoites* (IFO 12536).Source: Manufacturer, New England Bio labs, U.S.A. http://www.neb.com

Table 3.6: Primer sets used for the PCR amplification of the various SNPs in APOL1, MYH9 and G6PD genes

rs number	Gene	Forward Primer	Reverse Primer
rs73885319	APOL1	GCCAGAGCCAATCTTCAGTC	GACTTTGCCCCCTCATGTAA
rs60910145	APOL1	GGCCCCTGTAAGCTTCTTTC	GCCCTGTGGTCACAGTTCTT
rs9622363	APOL1	CCTGGGTGAGAGAGCAAGAC	AAGTCAGCTGCCACCAAAAC
rs2032487	MYH9	TGGGTGAGAAAGGATTCCAG	GCCACACAGAACAGAAAGCA
rs11912763	MYH9	GAGTCACTGAACCCCGAGAC	GCAGAGCGAGGAGAAGAAGA
rs4821481	MYH9	CCCACAGTGACCAACACTTG	GCTGCTGCTTCCTGTCTCTT
rs5750248	MYH9	GTGCTAAGGTCATCCCCAAA	TCCGTAAGCCAGGTTTTCAC
rs5750250	MYH9	GGACATCAGAATCCCCTGAA	CGGATGTGCACTTTTCATCA

Table 3.7: Reagents with concentrations and the volumes used in preparing Master Mix for amplifying APOL1 and MYH9 SNPs.

Reagents in reaction mix	Initial concentration	Volume/ µl	Final concentration in 15 µl reaction mix
Sterilized Double Distilled Water		6.83	
PCR Buffer	10x	1.5	1X
$MgCl_2$	50mM	0.3	0.5mM
DNTP	10mM	0.25	0.17mM
Forward Primer	10µM	1.0	0.1nM
Reverse Primer	10µM	1.0	0.1nM
Taq Polymerase	5U/µl	0.12	1U
DNA		4	

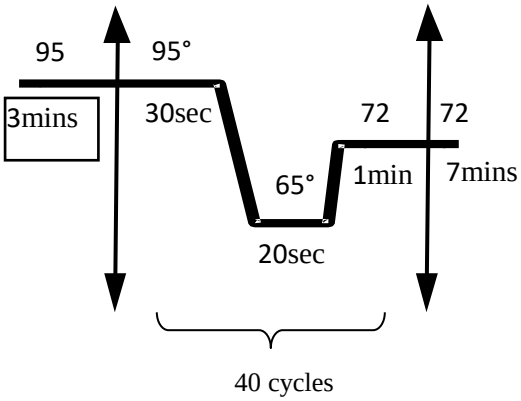

Figure 3.3: Cycling conditions for APOL1 and MYH9 SNPs showing the various cycling temperatures at each stage of the PCR

3.9 Ligation Detection Reaction (LDR)

Genotyping of samples to identify APOL1 and MYH9 gene variants was done using the Ligation Detection Reaction (LDR), also known as Ligase Chain Reaction (LCR) as described by Fitness (2011) with slight modifications. The Ligation Detection Reaction is a post PCR step which provides a well-designed way for the multiplex genotyping of different SNPs, micro deletions and insertions found in the gene. It utilizes the ability of the enzyme ligase to preferentially connect adjacent oligonucleotides hybridized to target DNA in which there is a perfect complementation at the nick junction using 3 probes, 2 of which are allele specific and one common probe for both alleles (Figure 3.4).

After performing the multiplex PCR for the APOL1 and MYH9 SNPs, a Uni PCR step was set up using conditions in figure 3.5 to increase the number of PCR amplicons from the previous step. Ten (10) µl of protease is then added to the Uni PCR product and run at 50°C for 20min and 75°C for 15 min. This step ensures that all the enzymes used for the PCR reaction are completely degraded before the LDR assay begins. The reagents used for the LDR reaction are listed in table 3.8. The LDR assay was then run using conditions in figure 3.6 and reagents in table 3.9. The LDR products were then run on a polyacrylamide gel to resolve the various amplicons to identity the SNP's present in the sample.

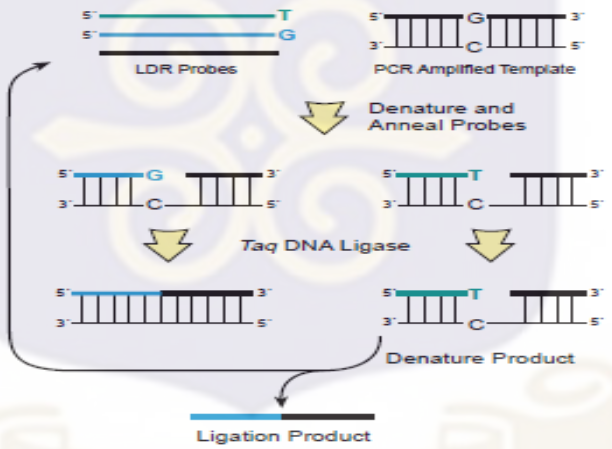

Figure 3.4: LDR Typing of a G-to-T SNP. The allelic probe corresponding to the wildtype allele is shown in blue and has a G at its 3' end. The common probe is shown in bold black. In this example, the template has been amplifies from an individual homozygous for the wildtype G allele. Thus only the wildtype allelic probe is ligated to the common probe, generating single type of ligation product. The thermal cycling is repeated for a few times in order to generate sufficient ligation product for detection.

Table 3.8: Reagents with concentrations and the volumes used in preparing Master Mix for Uni PCR for APOL1 and MYH9 SNPs.

Reagents Reaction mix	Initial concentration	Volume/ µl	Final concentration in 15 µl reaction mix
Sterilized Double Distilled Water		6.73	
PCR Buffer	10x	1.0	1X
$MgCl_2$	50mM	0.8	0.5mM
DNTP	10mM	0.8	0.17mM
Uni Forward Primer	10µM	0.5	0.1nM
Uni Reverse Primer	10µM	0.5	0.1nM
Taq Polymerase	5U/µl	0.12	1U
PCR template		4	

Table 3.9: List of reagents with concentrations and the volumes used in preparing Master Mix for LDR assay of APOL1 and MYH9 SNPs.

Reagents reaction mix	Initial concentration	Volume/ µl
APB Buffer	10x	2
AmX L1	5µM	1
AmX L2	100µm	1
AmY L1	5µM	1
AmY L2	100µM	1
Marker 1 L1 mix	50µM	1
Marker 1 L2	100µM	1
Marker 2 L1 mix	50µM	1
Marker 2 L2	100µM	1
Ligase	5U/µl	0.1
Uni PCR Template		10ul

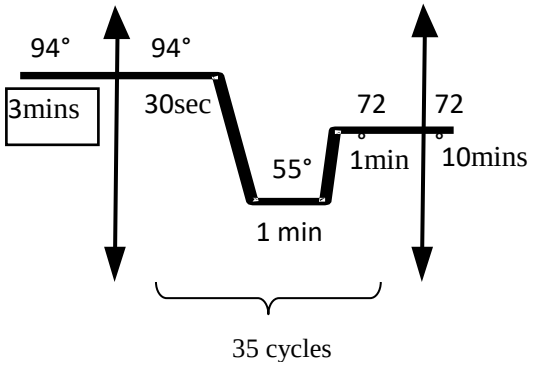

Figure 3.5: Cycling conditions for Uni PCR for APOL1 and MYH9 PCR amplicons.

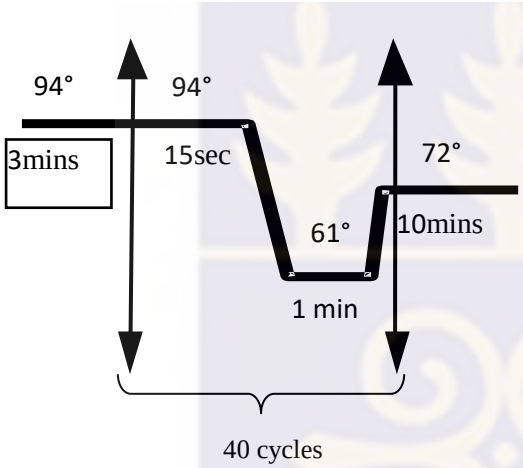

Figure 3.6: Cycling conditions for LDR assay of APOL1 and MYH9 SNPs

Two main quality control methods were employed for the genotyping assays used in this study. For the RFLP assay, negative controls which involved the replacing the template DNA or PCR product with distilled water which shows no band when run on an agarose gel was used. This was used to confirm that the reaction was not contaminated hence, the band sizes observed are solely from the template DNA. Positive controls were also run using a sample that is known to contain the SNPs being studied. This helped in ensuring that the reaction was working perfectly. For the LDR assay, the amelogenin gene was amplified and genotyped together with the

APOL1 and MYH9 genes as a marker to ensure that the results obtained was a true reflection of the samples.

Amelogenin is a gene found on both the X and Y chromosome and codes for a group of closely related proteins that plays a role in the development of tooth enamel. Differences in the X and Y chromosome versions of the gene enable it to be used in determining the sex of human or animal samples. The X chromosome form (i.e. AMELX) has a 8bp deletion in intron 1 which is not present in the Y chromosome form (i.e. AMELY), because of this, differences between the 2 variants of the gene can be detected using PCR followed by gel electrophoresis and is usually used as a quality control for genotyping human samples (Yutaka & Osamu, 1991).

After obtaining the genotyped data, the quality control method for determining if generated genotype data was the true reflection of the general population was to determine if the results for the control group obeyed the Hardy-Weinberg equilibrium.

The Hardy-Weinberg equilibrium is based on a principle which states that the genetic variation in a population will remain constant from one generation to the next in the absence of any other evolutionary influences or factors. These factors include random mating, the assumption that the population is infinitely large, genetic drift, non-overlapping generations, and natural selection (Moonesinghe *et al*., 2010). Because most of these factors generally occurs in nature, the Hardy- Weinberg equilibrium rarely applies in reality. It therefore assumes the ideal state where all these factors are believed to apply. Deviations from the Hardy-Weinberg equilibrium will depend in which of the factors were violated. For instance, in a population where there is the practice of inbreeding, this population will not obey the Hardy-Weinberg principle,

since the absence of random mating leads to the rise in high frequency in homozygosity for all genes.

3.11 Polyacrylamide Gel Electrophoresis

The polyacrylamide gel electrophoresis is an analytical method used to separate components of proteins or PCR amplicons based on their size. The technique is based on the principle that charged molecules migrate in an electric field towards an electrode with the opposite charge. In this case, because the amplified DNA fragments have an overall negative charge, they will migrate towards the positive charge and will be separated into bands based on their respective sizes.

Reagents used to prepare the gel includes 40% mixture of acrylamide (acrylic amide) and bis-acrylamide (*N, N′*-Methylenebisacrylamide) in the ratio 19:1. To prepare the 10% Polyacrylamide gel, 12.5ml of 40X Acugel (19:1 acrylamide, bisacrylamide) (Amresco, USA) was mixed with 5ml of 10X Tris-Borate-EDTA (TBE) buffer, 500ul of freshly prepared APS (Amresco, USA), 50ul of TEMED (Amresco, USA) and 32.5 ml of Distilled water in a clean and sterile beaker and swirled gently to mix. The mixture was then dispensed into the acrylamide glass plate set up from the top and the gel comb inserted into the solution in between the glass plates. The setup was then left to solidify before loading LDR products into the wells. The gel was then placed in the acrylamide gel tank containing 1X TBE buffer avoiding the formation of bubbles at the bottom of the gel plate as this could interfere with the gel resolution. The gel was run at 250V for 90 minutes after which the resolved products are viewed on the UV-trans illuminator.

3.12 Data Management and Statistical Analysis

All associated summary statistics including the mean and standard deviations were reported as mean ± SD with the statistical difference tested at 5% significance level. The prevalence of the G6PD variants were calculated by using a proportion of number of cases to the total number of samples used for the study.

The chi-square test was used to determine association between the observed genotypes and disease outcome in a crude analysis (without controlling for possible confounders). The extent of association (whether or not the association reduced risk/had a protective effect or increased risk of disease) was determined and presented as an odds ratio (OR). Subsequently, a Logistic regression analysis using STATA (version 13.0), was also conducted to confirm the associations observed. Six different models of disease association (additive, dominant, recessive, genotype, test of trends and the heterozygous) were tested in the logistic regression, adjusting for age, clinical site location and gender. The G6PD regression models were however run independently since they were stratified according to gender (G6PD is an X-linked gene).

The additive model of association was defined as the risk conferred by an allele in an increasing fold (i.e. X1 for heterozygotes and 2X for homozygotes). The recessive model of association was defined as the effect of only the recessive allele on disease outcome. The Dominant model of association was defined as one that examines the association observed in only the dominant allele (i.e. the effect of a single dominant allele on disease outcome). The genotype model of association determined the effect each genotype on disease outcome. The heterozygous model of association determined the effect of only the heterozygous genotype on disease outcome.

To determine the genetic interactions between G6PD, APOL1 and MYH9 risk variants, a logistic regression analysis was computed with the G6PD variants and the individual APOL1 and MYH9 genotypes as covariates. The interactions were tested using the standard centred cross-product of the APOL1 or MYH9 genotypes and the G6PD A- and A genotypes.

CHAPTER FOUR

RESULTS

4.1 Characteristics of Study Population

Four hundred and eleven (411) participants consisting of 202 cases of CKD of unknown aetiology and 209 healthy controls were selected for the study. The 202 cases consisted of 111 (54.95%) males and 91 (45.05%) females, whose age ranged from 9 to 74 years. There was no significant difference (p= 0.1625) between the number of male and female participants. The age of the males ranged from 13 to 72 years, with the mean age and standard deviation of 41±15 whiles that of the females was 9-74 with the mean age and standard deviation of 40±16. The mean age of the males and females did not differ significantly (p= 0.5140), (Table 4.1).

Table 4.1: Age and Sex distribution of case and control participants

	Case	Control	P-value
Count	202	209	0.6855
Male	111	95	0.2535
Female	91	114	0.0885
Mean ±SD	41±15	41±15	
Age range	9-76	17-76	

4.2 Gel Electrograms

4.2.1 Agarose Gel Electrophoresis

Representative gel electrograms for agarose gel electrophoresis for both PCR and RFLP genotyping reactions have been shown in figures 4.3 to 4.6. The sizes of the PCR amplicons for G6PD A- was 203bp whiles G6PD A was 272bp. Following the RFLP, the fragment size for G6PD A- was 123bp and 69bp whiles that for G6PD A was 155bp and 117bp. The fragment sizes of the variants after performing the RFLP assay have been presented in table 4.2.

Table 4.2: G6PD A deficiency distribution among participants from various clinical sites

G6PD Variant	Exon	Nucleotide change	PCR uncut	Mutant digested fragments
A^-	4	202 G ➡ A	203bp	123bp,69bp
A	5	376 A ➡ G	272bp	155bp,117bp

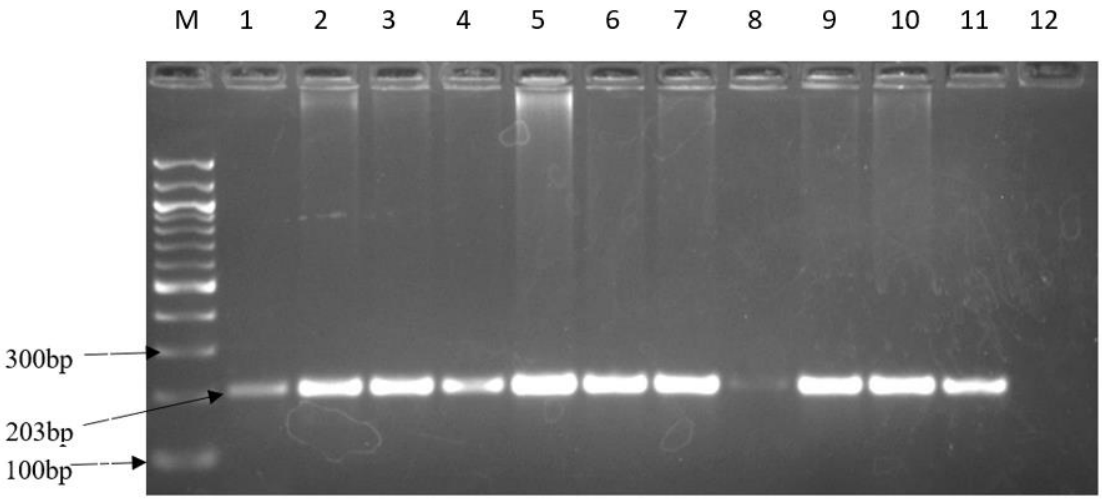

Figure 4.3: Representative gel showing PCR amplicons for G6PD A-. After successful amplification of the G6PD gene region has the G6PD A- variant, a band size of 203bp was observed. Lane M= 100bp molecular marker, 1-10 = samples, 11= positive control, 12= negative control.

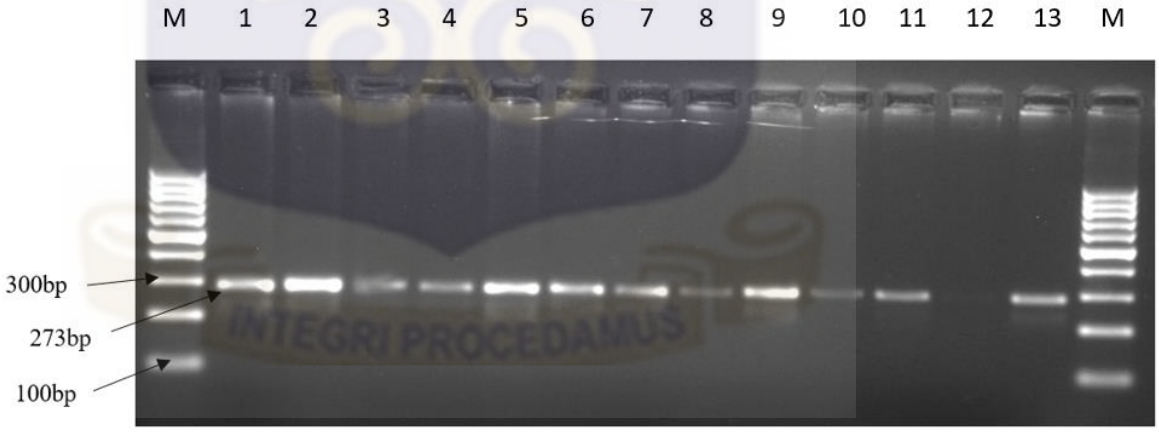

Figure 4.4: Representative gel showing PCR amplicons for G6PD A: After successful amplification of the G6PD gene region has the G6PD A variant, a band size of 273bp was observed. Lane M= 100bp molecular marker, 1-11 = samples, 12= negative control, 13= positive control.

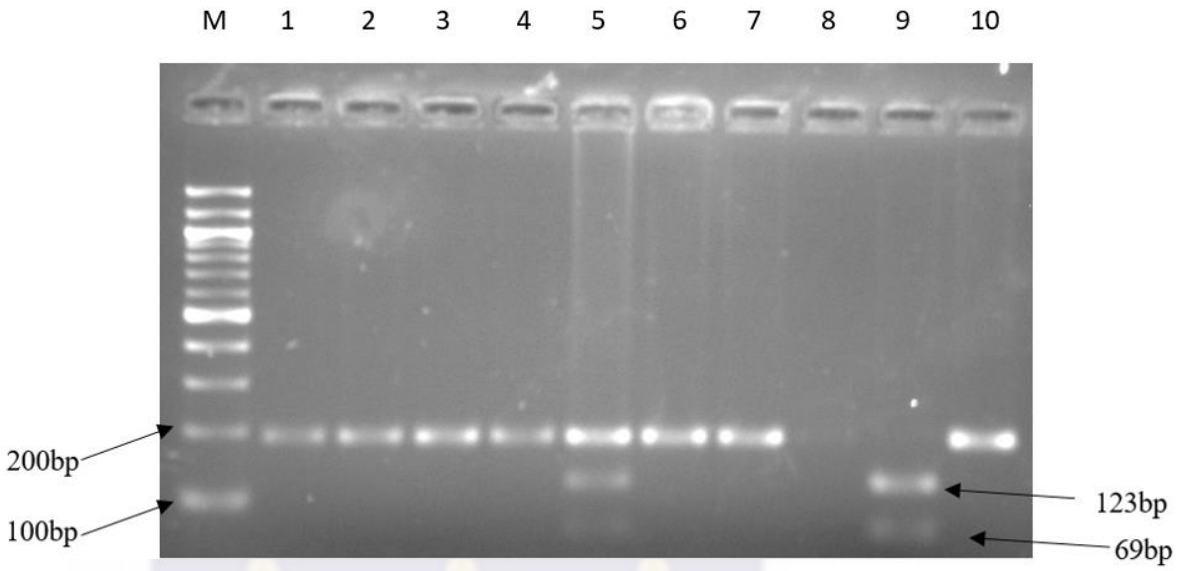

Figure 4.5: Representative gel showing RFLP for G6PD A- : *Nla III* digested 203bp amplified region of G6PD A- to produce 123bp and 69bp band sizes. Lane M= 100bp molecular marker, 1-8 = samples, 9= positive control, 10= uncut product (negative control).

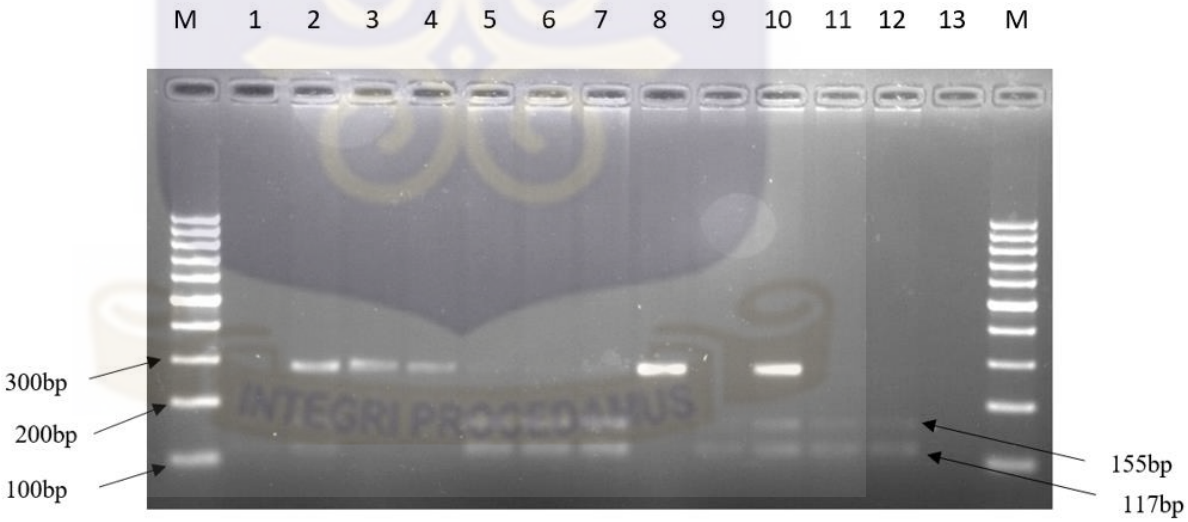

Figure 4.6: Representative gel showing RFLP for G6PD A: *Fok I* digests 273bp amplified region of G6PD A to produce 155bp and 117bp band sizes. Lane M= 100bp molecular marker, 1-11 = samples, 12= positive control, 13= empty well

4.2.2 Polyacrylamide Gel Electrophoresis

Representative gels for Polyacrylamide gel electrophoresis for LDR genotyping reactions have been shown in figures 4.5 to 4.9. The fragment sizes of the variants following genotyping have been presented in table 4.3. To detect the genotype of each individual, the gel images were analysed for homozygosity or heterozygosity of each SNP. Homozygous alleles showed 2 bands for each SNP whiles heterozygous SNPs showed only one band for the particular SNP. Each allele is differentiated however by size and colour.

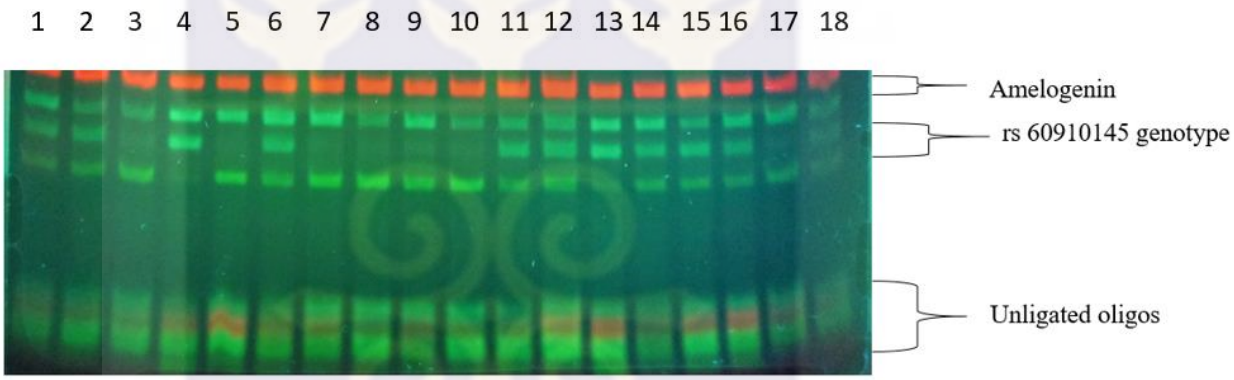

Figure 4.5: Representative gel showing LDR genotyping of APOL1 rs60910145: 10% Acrylamide gel electrophoresis showing genotypes for variant rs60910145. Lanes 1-18 = samples .19 = negative control.

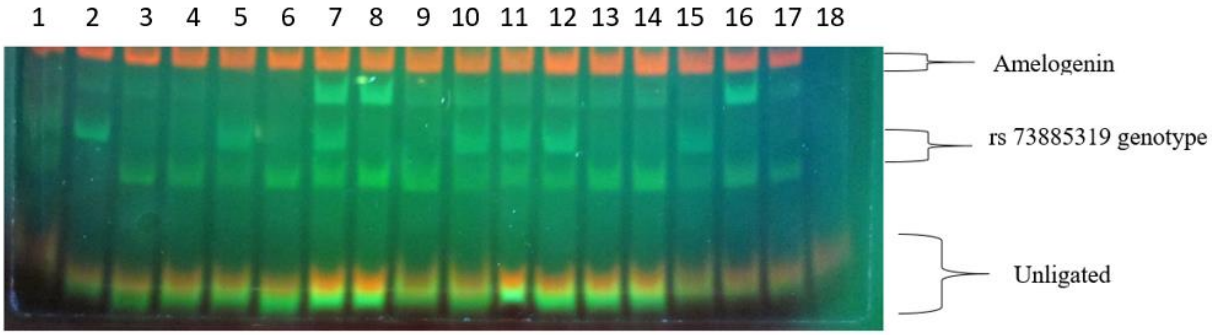

Figure 4.6: Representative gel showing LDR genotyping of rs73885319: 10% Acrylamide gel electrophoresis showing genotyping for variant rs73885319. Lanes 1-17 = samples .18 = negative control.

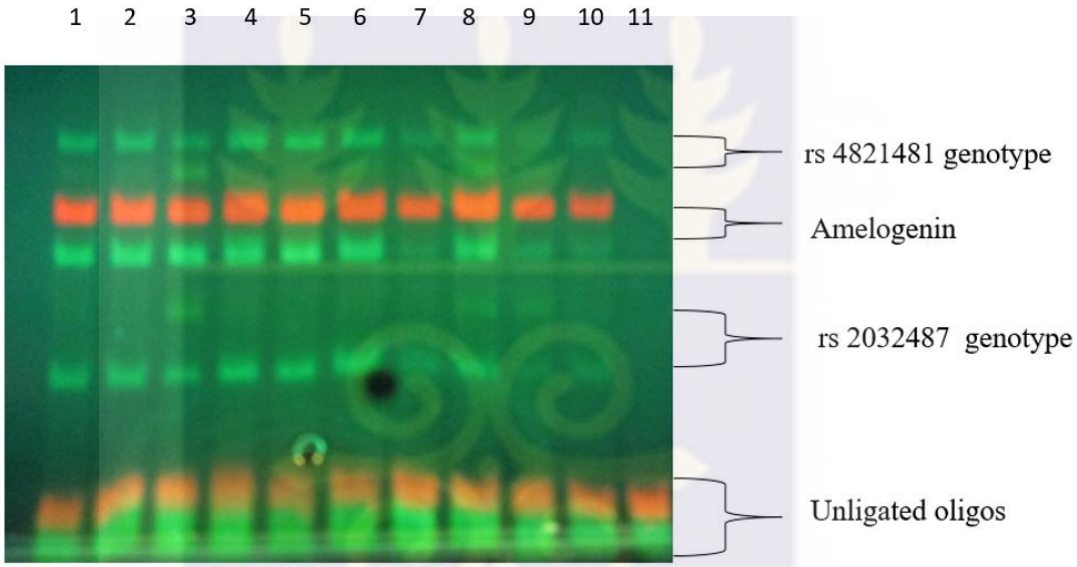

Figure 4.7: LDR genotyping of rs4821481: 10% Acrylamide gel electrophoresis showing genotyping for variant rs4821481 and rs2032487. Lanes 1-10 = samples .11 = negative control.

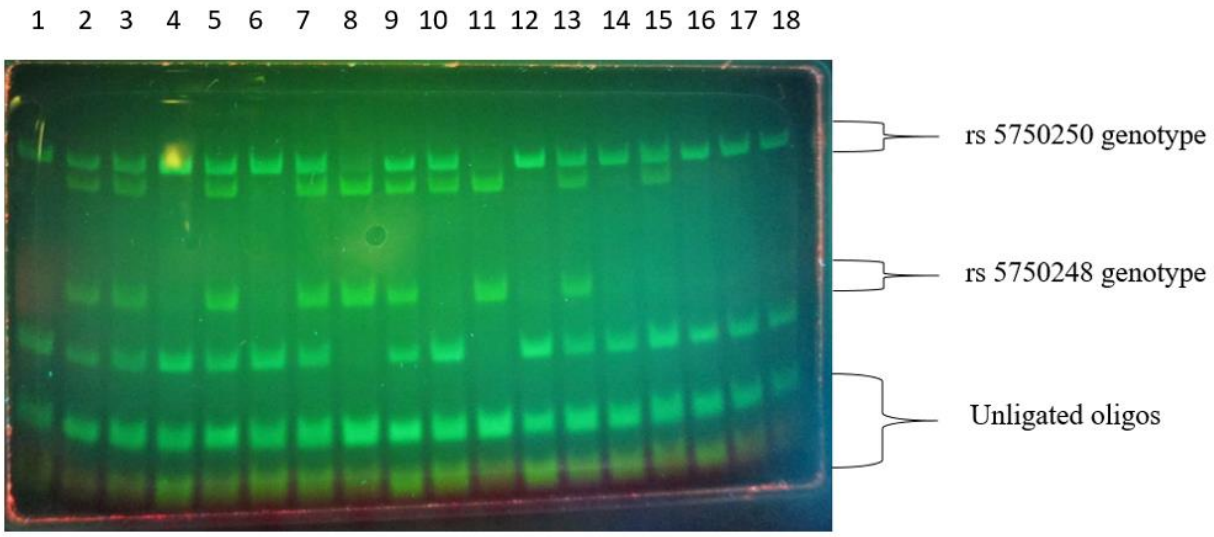

Figure 4.8: LDR genotyping of rs5750250 and rs5750248: 10% Acrylamide gel electrophoresis showing genotyping for variant rs5750250 and rs5750248. Lanes 1-18 = samples.

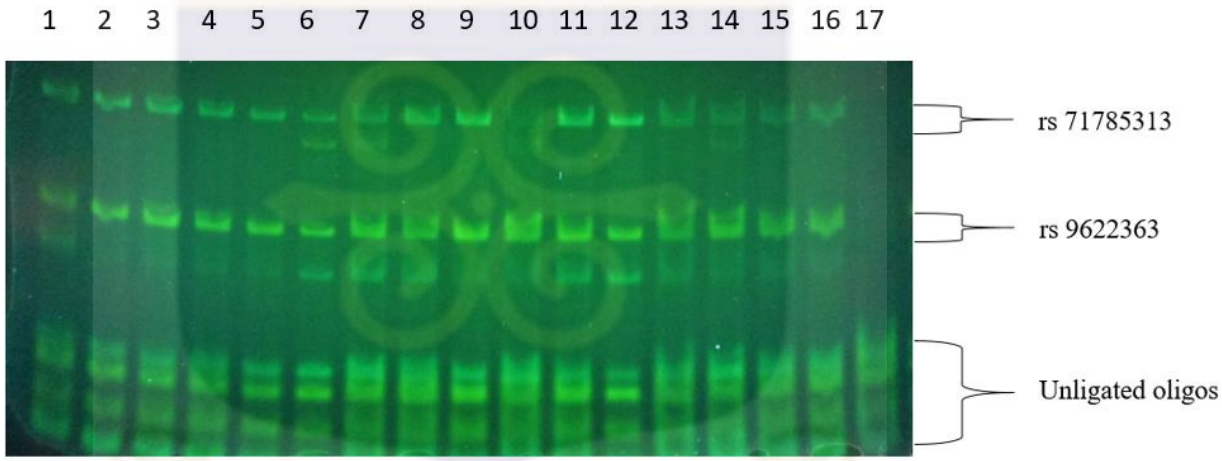

Figure 4.9: LDR genotyping of rs71785313 and rs9622363: 10% Acrylamide gel electrophoresis showing genotyping for variant rs5750250 and rs5750248. Lanes 1-16 = samples 17= negative control.

4.3 Molecular Characterization of G6PD mutations among the Study Participants

Following the successful PCR amplification of the section of the G6PD gene that codes for the variants A- and A, RFLP genotyping assay for the G6PD variants A and A^- was performed among the study participants. It was observed that for G6PD A^-, 99 persons representing 49.01% of the cases were positive for variant whiles 110 representing 54.46% of the cases were positive for variant A. We further grouped them according to gender and also observed that for variant A^-, the genotype revealed that females recorded 25 (27.47%) homozygotes whilst males recorded 48 (43.24%) hemizygotes. For variant A, females recorded 23(25.2%) homozygous cases whiles males recorded 58 (52.25%) hemizygous cases. There were no male heterozygous cases for both variants since the gene is X-linked and as such males cannot be heterozygous. (Table 4.4 and 4.5).

Table 4.4: G6PD A$^-$ variant distribution among Case participants

Sex	Full deficient	Partial deficient	Non- deficient	Total
Male	48	-	63	111
Female	25	26	40	91
Total	73	26	103	202

Table 4.5: G6PD A variant distribution among Case participants

Sex	Full deficient	Partial deficient	Non- deficient	Total
Male	58	-	53	111
Female	23	34	34	91
Total	81	34	87	202

Table 4.6: G6PD A$^-$ variant distribution among participants from various clinical sites

Clinical Site	Homozygous deficient (%)	Heterozygous deficient (%)	Wild type (%)	Total
Korle-Bu	25(24.5)	25(37.3)	67(27.7)	117
Komfo Anokye	12(11.8)	4(5.9)	26(10.7)	42
Abuja	15(14.7)	7(10.5)	26(10.7)	48
Ile-Ife	10(9.8)	5(7.5)	22(9.1)	37
Ilorin	16(15.7)	10(14.9)	60(24.8)	86
Ibadan	14(13.7)	11(16.4)	29(12.0)	54
Enugu	10(9.8)	5(7.5)	12(5.0)	27
Total (100%)	102(100)	67(100)	242(100)	411

Grouping the samples according to clinical sites, it was observed that participants from Korle-Bu teaching hospital had the highest prevalence of G6PD deficiency for both A^- and A variants (24.5% and 24.4% respectively) whiles those from Ile-Ife had the lowest prevalence rates for both G6PD variants A^- and A (9.8% and 6.6% respectively). Generally, the average prevalence of the sample population from Ghana was higher than that from Nigeria (Table 4.6 and 4.7). Because the various study sites were from Ghana and Nigeria, the prevalence of the G6PD variant A^- and A were also calculated based on the country of origin according to gender (Table 4.8 and 4.9). Similar to the prevalence recorded according to clinical site location, the prevalence of participants from Ghana was higher for all the genotypes in both G6PD A^- (females (A/A) = 73.8%, males (A/A) = 68.4%) and G6PD A (females (G/G) = 59.0%, males (G/G) = 84.2%) compared to those from Nigeria (G6PD A^- : females (A/A) = 31.1%, males (A/A) = 48.7% . G6PD A: females (G/G) = 32.8%, males (G/G) = 53.95%).

The General prevalence of G6PD variants A^- was 28% (female population) and 34.74% (male population) whiles that of G6PD A was 38.5% (female population) and 42.1% (male population). There was no significant difference ($p< 0.0896$) in the prevalence between G6PD variants A^- and A. The prevalence of G6PD variants among the cases was however 56.04% (female population) and 43.24% (male population) for variant A^- then 62.64% (female population) and 47.75% (male population) for variant A respectively. The prevalence of G6PD A^- and A have been summarized on tables 4.10 and 4.11

Table 4.7: G6PD A variant genotype distribution among participants from various clinical sites

Clinical site	Homozygous deficient (%)	Heterozygous deficient (%)	Wild type (%)	Total
Korle-Bu	33(24.4)	16(29.1)	52(29.2)	101
Komfo Anokye	12(8.8)	7(12.7)	23(12.9)	42
Abuja	20(14.8)	4(7.3)	24(13.5)	48
Ile-Ife	9(6.6)	3(5.5)	13(7.3)	25
Ilorin	29(21.5)	10(18.2)	35(19.7)	74
Ibadan	20(14.8)	10(18.2)	21(11.8)	51
Enugu	12(8.8)	5(9.1)	10(5.6)	27
Total (100%)	135(100)	55(100)	178(100)	411

Table 4.8: Prevalence of G6PD A$^-$ among the study population from Ghana and Nigeria

	Prevalence in cases (%)				Prevalence in controls (%)			
Country	F(A/A)	F(G/A)	M(A/A)	P-value	F(A/A)	F(G/A)	M(A/A)	P-value
Ghana	73.8	8.2	68.4	0.0714	4.8	29.3	32.7	0.0034
Nigeria	31.1	21.3	48.7	0.6680	6.8	17.8	34.9	0.2369

*P-value for an overall effect. F(A/A) = homozygous females, F(G/A) = heterozygous females, M(A/A) = homozygous males

Table 4.9: Prevalence of G6PD A among the general population Ghana and Nigeria

	Prevalence in cases (%)				Prevalence in controls (%)			
Country	F(G/G)	F(A/G)	M(G/G)	P-value	F(G/G)	F(A/G)	M(G/G)	P-value
Ghana	59	26.2	84.2	0.2304	29.2	9.7	48.1	0.3785
Nigeria	32.8	24.6	53.95	0.9429	17.8	20.5	58.1	0.0647

*P-value for an overall effect. F(G/G) = homozygous females, F(A/G) = heterozygous females, M(G/G) = homozygous males

Table 4.10: General prevalence of G6PD genotypes among CKD cases

G6PD variant	Sample type	Number	Prevalence (%)
G6PD A$^-$	Female Hetero/Homozygous	51	56.04
	Male hemizygous	48	43.24
G6PD A	Female Heterozygous	57	62.64
	Male hemizygous	53	47.75

Table 4.11: General prevalence of G6PD genotypes among General Population control samples

G6PD variant	Sample type	Number	Prevalence (%)
G6PD A$^-$	Female Hetero/homozygous	59	28.0
	Male hemizygous	33	34.74
G6PD A	Female Hetero/homozygous	45	38.5
	Male hemizygous	51	42.1

4.4 Molecular Characterization of G6PD, APOL1 and MYH9 SNPs among Study Participants

4.4.1 DNA Quality: Genotyping Performance

Table 4.12 below represents information on the alleles and minor allele frequency of the polymorphisms in the study population as well as results for Hardy-Weinberg equilibrium for the various SNPs to determine whether variants were evenly distributed.

Table 4.12: Minor allele frequencies (MAF), and p-value for comparing genotype distributions among participants.

SNP	Gene	Allele	MAF	HWE P Value	Allele Frequency	
					Cases	Controls
rs73885319	APOL1	A/G	G(0.37)	0.658	G(0.40)	G(0.26)
rs60910145	APOL1	G/T	T(0.66)	0.101	T(0.63)	T(0.69)
rs71785313	APOL1	D/I	N/A	0.714	N/A	N/A
rs2032487	MYH9	T/C	C(0.79)	0.282	C(0.79)	C(0.79)
rs4821481	MYH9	T/C	C(0.76)	0.123	C(0.75)	C(0.77)
rs5750248	MYH9	C/T	T(0.70)	0.057	T(0.68)	T(0.71)
rs5750250	MYH9	A/G	G(0.67)	0.158	G(0.67)	G(0.69)
rs1050828	G6PD A-(202)	G/A	A(0.35)	7.62E-10	A(0.44)	A(0.25)
rs1050829	G6PD A(376)	A/G	G(0.43)	0.000	G(0.47)	G(0.40)

*HWE, Hardy-Weinberg Equilibrium (HWE p-values are from exact tests). MAF = Minor Allelic Frequency

4.4.2 Association between G6PD, APOL1 and G6PD gene variants and CKD of Unknown Aetiology

Test for association for all SNPs was performed using STATA (version 13.0). The results for the association studies are summarized in tables 4.13 to 4.16. As shown in table 4.13, only 2 SNPs, (rs73885319 and rs60910145) which happened to be the G1 variants of APOL1 showed a strong association indicating a 30% and 40% increase in risk of disease and was statistically significant (OR 1.3, 95% CI 1.033- 1.7282, p = 0.027 and OR, 1.4 95% CI 1.0987-1.6958, p = 0.005 for rs7885319 and rs60910145 respectively) for the additive and recessive models. The dominant model however showed a 20% reduced risk of disease but was not statistically significant (OR 0.8, 95% CI 0.3743-1.5891, p = 0.481 and OR 0.8, 95% CI 0.3417-1.5123, p = 0.384 for rs7885319 and rs60910145).

The test for association using the genotype model (table 2 in the appendix) revealed that the association that was observed in the test of trends was actually from the heterozygous genotype for both SNPs, with a more increased association than observed (OR 1.97, 95% CI 1.2726- 3.0442, p = 0.002 for rs73885319 and OR, 1.86 95% CI 1.2083-2.8683, p = 0.005 for rs60910145) for the test of trends. For rs9622363 however, the heterozygous model showed a significant association which was not observed in the additive and recessive models (OR 0.5, 95% CI 0.2530-0.9447, p = 0.033) even though this association was identified to exhibit a 50% reduced risk of disease. Nonetheless, it had an increased risk of disease association in the test of trends model but was not significant (OR, 1.5 95% CI 0.8404 -2.566, p = 0.177).

Test for association between MYH9 variants and CKD of unknown aetiology did not reveal any significant associations between all the SNPs in all models of association (Table 4.14).

Table 4.13: Logistic regression analysis showing association between APOL1 variants and CKD of unknown aetiology in different models

Gene	SNP	Risk Allele	Model	Unadjusted		Adjusted	
				OR(95%CI)	P-value	OR(95%CI)	P-value
APOL1	rs73885319	G	Additive	**1.3(1.8906-3.1239)**	**0.012**	**1.3(1.644-4.6118)**	**0.02**
			Dominant	0.8(0.3873-1.6006)	0.509	0.8(0.374-1.5891)	0.481
			Recessive	1.3(0.6248-2.5818)	0.509	1.3(0.629-2.6719)	0.481
			Test of trends	**1.3(1.033-1.7282)**	**0.027**	**1.4(1.049-1.7737)**	**0.02**
			Heterozygous	**1.7(1.1525-2.6289)**	**0.008**	**1.8(1.186-2.7559)**	**0.006**
			Genotype(A/G)	**1.9(1.2726-3.0442)**	**0.002**	**2.0(1.321-3.2213)**	**0.001**
			Genotype (G/G)	1.86(0.880-3.934)	0.104	1.9(0.907-4.1839)	0.087
APOL1	rs60910145	G	Additive	**1.3(1.8905-3.1899)**	**0.014**	**1.3(1.704-4.9177)**	**0.004**
			Dominant	0.8(0.3584-1.5522)	0.433	0.8(0.3417-1.512)	0.384
			Recessive	1.3(0.6442-2.7903)	0.433	1.4(0.6612-2.926)	0.384
			Test of trends	**1.4(1.0987-1.6958)**	**0.005**	**1.4(1.105-1.7196)**	**0.004**
			Heterozygous	**1.7(1.1077-2.5349)**	**0.015**	**1.7(1.111-2.5812)**	**0.014**
			Genotype (G/T)	**1.86(1.208-2.8683)**	**0.005**	**1.9(1.222-2.9495)**	**0.004**
			Genotype (G/G)	1.8(0.8604-3.9658)	0.115	1.9(0.891-4.2069)	0.095
APOL1	rs71785313	Del	Additive				
			Dominant	0.8(0.7-1.1)	0.248	0.9(0.718-1.0989)	0.275
			Recessive			0.9(0.718-1.0989)	0.275
			Test of trends	0.9(0.7158-1.0901)	0.248	0.9(0.718-1.0989)	0.275
			Heterozygous	0.8(0.7-1.1)	0.248	0.9(0.718-1.0989)	0.275
			Genotype (Ins/Del)	0.7(0.3668-1.2953)	0.248	0.7(0.3696-1.327)	0.275
APOL1	rs9622363	G	Additive	1.4(3.4467-5.6708)	0.177	1.3(2.6329-9.390)	0.177
			Dominant	**0.5(0.2669-1.0382)**	**0.064**	**0.5(0.261-1.0273)**	**0.06**
			Recessive	**1.8(0.9639-3.7466)**	**0.064**	**1.9(0.973-3.8252)**	**0.06**
			Test of trends	1.5(0.8404-2.566)	0.177	1.5(0.846-2.6091)	0.168
			Heterozygous	**0.5(0.2530-0.9447)**	**0.033**	**0.5(0.247-0.9339)**	**0.031**
			Genotype (A/G)	0.34(0.042-2.7059)	0.307	0.3(0.0409-2.695)	0.302
			Genotype (G/G)	0.7(0.0782-5.6850)	0.711	0.7(0.077-5.7522)	0.712

Values shaded in red indicate statistically significant association. Values shaded in blue indicate a trend of association (not statistically significant)

Table 4.14: Logistic regression analysis for additive, dominant and recessive models showing association between MYH9 variants and CKD of unknown aetiology

Gene	SNP	Risk Allele	Model	Unadjusted		Adjusted	
				OR(95%CI)	P-value	OR(95%CI)	P-value
MYH9	rs2032487	C	Additive	0.9(0.6848-1.3987)	0.906	0.6(0.6921-1.4293)	0.977
			Dominant	0.6(0.1967-1.9657)	0.419	0.7(0.2058-2.1254)	0.488
			Recessive	1.0(0.50871-5.0828)	0.851	1.0(0.4705-4.8592)	0.488
			Test of Trends	0.9(0.8186-1.2162)	0.982	0.9(0.8094-1.2107)	0.922
			Heterozygous	0.9(0.5781-1.3704)	0.597	0.9(0.5727-1.3757)	0.594
			Genotype (T/T)	1.6(0.4867-4.9714)	0.456	1.45(0.4489-4.7421)	0.53
			Genotype (T/C)	0.91(0.5902-1.4111)	0.681	0.9(0.5824-1.4111)	0.664
MYH9	rs4821481	C	Additive	0.9(0.6549-1.3053)	0.656	0.6(0.6661-1.3423)	0.754
			Dominant	0.6(0.1967-1.9657)	0.419	0.7(0.2058-2.1253)	0.488
			Recessive	0.9(0.5087-5.0828)	0.827	0.9(0.4705-4.8592)	0.488
			Test of Trends	1.1(0.8131-1.4834)	0.541	1.1(0.7919-1.4604)	0.642
			Heterozygous	0.9(0.6414-1.4899)	0.916	0.9(0.6281-1.4787)	0.866
			Genotype (T/T)	1.6(0.5033-5.1643)	0.421	1.5(0.4617-4.9032)	0.498
			Genotype (T/C)	1.0(0.6569-1.5417)	0.976	0.99(0.6406-1.5241)	0.957
MYH9	rs5750248	T	Additive	0.9(1.5158-2.1399)	0.328	2.0(1.2418-3.2379)	0.122
			Test of trends	0.9(0.7539-1.0991)	0.328	0.9(0.7404-1.0866)	0.267
			Heterozygous	1.0(0.6966-1.5768)	0.822	1.0(0.6883-1.5815)	0.841
			Genotype (T/T)	0.58(0.2331-1.4290)	0.577	0.5(0.2036-1.3001)	0.16
			Genotype (C/T)	0.65(0.2617-1.5984)	0.345	0.6(0.2309-1.4682)	0.252
MYH9	rs5750250	G	Additive	0.7(1.4849-2.1042)	0.142	2.0(1.2104-3.1705)	0.122
			Dominant	1.4(0.9132-2.0828)	0.126	1.4(0.9397-2.1929)	0.094
			Recessive	0.7(0.4801-1.0949)	0.126	0.7(0.4560-1.0641)	0.094
			Test of trends	0.8(0.5525-1.0891)	0.142	0.8(0.5271-1.0566)	0.099
			Heterozygous	1.3(0.8576-1.9430)	0.221	1.3(0.8636-1.9967)	0.203
			Genotype G/G)	0.67(0.2917-1.5636)	0.359	0.6(0.2613-1.4520)	0.268
			Genotype A/G)	0.92(0.4006-2.1224)	0.849	0.9(0.3714-2.0319)	0.746

Table 4.15: Logistic regression analysis for additive, dominant and recessive models showing association between G6PD variants and CKD of unknown aetiology in the female subjects

Gene	SNP	Risk Allele	Model	Unadjusted		Adjusted	
				OR(95%CI)	P-value	OR(95%CI)	P-value
G6PD A-	1050828	A	Additive	**6.3(2.1314-19.0673)**	**0.001**	**6.8(2.2458-20.7838)**	**0.001**
			Dominant	**1.8(1.0044-3.3030)**	**0.048**	**1.96(1.067-3.6046)**	**0.03**
			Recessive	**6.5(2.2225-19.2411)**	**0.001**	**6.8(2.2893-20.4858)**	**0.001**
			Test of trends	**8.2(2.860-24.2589)**	**0.0001**	**17.6(3.0477-101.2814)**	**0.001**
			Genotype (G/A)	**0.14(0.0435-0.4695)**	**0.001**	**0.1(0.04329-0.4863)**	**0.002**
			Genotype (A/A)	**8.7(3.1099-24.6190)**	**0.0001**	**19.0(3.4155-106.4739)**	**0.001**
G6PD A	1050829	G	Additive	**2.8(1.3335-6.2372)**	**0.007**	**3.0(1.3807-6.6341)**	**0.006**
			Dominant	**1.96(1.1033-3.5143)**	**0.022**	**2.0(1.1206-3.6372)**	**0.019**
			Recessive	1.0(0.5326-2.0073)	0.921	1.0(0.5245-2.0053)	0.941
			Test of trends	1.3(0.7794-2.0661)	0.338	1.9(0.5424-6.5494)	0.319
			Genotype (A/G)	2.0(0.8463-4.9435)	0.112	2.1(0.8712-5.2116)	0.097
			Genotype (G/G)	1.8(0.9979-3.1671)	0.051	3.0(0.8032-11.4521)	0.102

* Columns shaded in red indicate statistically significant association for the model

Table 4.16: Logistic regression analysis for additive, dominant and recessive models showing association between G6PD variants and CKD of unknown aetiology in the male subjects

Gene	SNP	Risk Allele	Model	Unadjusted		Adjusted	
				OR(95%CI)	P-value	OR(95%CI)	P-value
G6PD A-	1050828	A	Additive	1.29(0.7067-2.3550)	0.407	1.1(0.6154-2.1295)	0.669
			Dominant	1.3(0.7067-2.3550)	0.407	1.14(0.6154-2.1295)	0.669
			Recessive	1.29(0.7067-2.355)	0.407	1.14(0.6154-2.1295)	0.669
			Test of trends	2.7(1.2889-5.6458)	0.008	11.1(2.9015-42.2014)	0.0001
			Genotype	2.4(1.4747-3.826)	0.0001	10.3(2.9869-35.8087)	0.0001
G6PD A	1050829	G	Additive	1.0(0.5747-1.8470)	0.92	0.88(0.4845-1.6343)	0.707
			Dominant	1.0(0.5747-1.8470)	0.92	0.88(0.4845-1.6343)	0.707
			Recessive	1.0(0.5747-1.8470)	0.920	0.88(0.4845-1.6343)	0.707
			Test of trends	2.1(1.3809-3.1659)	0.0001	9.5(2.7897-32.0609)	0.0001
			Genotype	2.1(1.3809-3.1659)	0.0001	9.5(2.7896-32.0609)	0.0001

Values shaded in red indicate statistically significant association. Values shaded in blue indicate a trend of association (not statistically significant)

Tables 4.15 and 4.16 show a summary of the logistic regression analysis performed for male and female participants with adjusted and unadjusted odds ratio for additive, dominant and recessive models for the G6PD variants A and A^-. In the female population there was evidence of a strong association of both G6PD variant A^- in the additive, recessive and dominant models which indicated an increase in risk of disease for individuals with at least one risk allele (i.e. heterozygous individuals). The dominant model recorded a 96% increased risk of disease (OR 1.96, 95% CI 1.067-3.6046, p = 0.03) and this increased further for recessive model (OR 6.8, 95% CI 2.2893-20.4858, p = 0.001) the recessive model however showed a wider confidence interval which indicates that the samples used for this analysis were small.

G6PD A showed a similar trend of association with CKD unknown aetiology in the female population with the additive (OR 3.0, 95% CI 1.3807-6.6341, p = 0.006) and the dominant model (OR 2.0, 95% CI 1.1206-3.6372, p = 0.019) showing significant associations. The recessive model however was not statistically significant (OR 1.0, 95% CI 0.5245-2.0053, P=0.941). For the male population, the association observed for both variants were not statistically significant (Table 4.14). All other models that were run for the variants showed a similar pattern (see Appendix for tables of other models).

4.4.4 Test for gene-gene interactions between G6PD and APOL1 genotype

The test for genetic interactions between the variants G6PD A^- , G6PD A and APOL1 genotypes are summarized in table 4.17 and 4.18. Statistically significant epistatic interactions were observed between deficient G6PD A- genotype and APOL1 genotypes (heterozygous rs73885319 females (OR 12.6, 95% CI 2.5-62.4, p = 0.002) and males (OR 2.7, 95% CI 1.0-6.8, p = 0.04), heterozygous rs60910145 females (OR 9.2, 95% CI 1.9-44.4, p = 0.006) whiles epistatic interactions between wild type G6PD A- and heterozygous rs60910145 males (OR 2.3, 95% CI 1.0-5.2, p = 0.04) was also observed to be statistically significant. These observations confirmed the association that was detected in the G6PD A- variants and APOL1 G1 variants. A trend towards association was however observed for interactions between wildtype G6PD A- genotype (G/G) and heterozygous rs73885319 males (OR 2.2, 95% CI 1.0-5.1, p = 0.058) and heterozygous rs60910145 (G/T) males (OR 3.8, 95% CI 0.9-15.5, p = 0.062) but were all not statistically significant. Generally, more significant interactions were observed for homozygous mutant G6PD A- and heterozygous APOL1 genotype in females than in males. In the case of G6PD A however, all statistically significant interactions were observed in females and happened to be between heterozygous G6PD A genotype and heterozygous APOL1 genotype. There were observable epistatic interactions however between G6PD A- and both rs73885319 or rs60910145 SNPs which were not statistically significant.

Table 4.17: Test for genetic interactions between G6PD A- with APOL1 genotypes among Participants in terms of gender

			Interaction type	Female		Male	
SNP	Gene	Allele	G6PD*APOL1	OR(95% CI)	P-value	OR(95% CI)	P-value
rs73885319	APOL1	A	wt*A/G	1.2(0.5-2.6)	0.722	**2.2(1.0-5.1)**	**0.058**
			def*A/A	4.1(0.8-21.7)	0.098	1.5(0.6-3.9)	0.414
			def*A/G	**12.6(2.5-62.4)**	**0.002**	**2.7(1.0-6.8)**	**0.04**
			def*G/G	1		0.2(0.01-2.3)	0.202
			het*A/G	2.5(0.7-8.8)	0.147		
			het*G/G	7.2(0.8-67.9)	0.087		
rs60910145	APOL1	T	wt*G/G	1.6(0.4-6.0)	0.499	**3.8(0.9-15.5)**	**0.062**
			wt*G/T	0.9(0.4-2.2)	0.966	**2.3(1.0-5.2)**	**0.046**
			def*T/T	3.9(0.8-20.0)	0.103	1.7(0.7-4.3)	0.254
			def*G/G	1		0.5(0.07-3.5)	0.498
			def*G/T	**9.2(1.9-44.4)**	**0.006**	**2.5(0.9-6.1)**	**0.053**
			het*T/T	0.4(0.1-1.13)	0.085		
			het*G/T	2.9(0.8-10.48)	0.1		
rs71785313	APOL1		wt*Ins/Del	0.4(0.1-1.2)	0.118	0.6(0.1-3.9)	0.677
			def*Ins/Del	1.3(0.1-11.9)	0.837	0.6(0.2-2.2)	0.463
			het*Ins/Del	0.5(0.09-3.1)	0.475		
Rs9622363	APOL1	A	wt*G/A	0.8(0.08-8.5)	0.888	1	
			wt*G/G	0.8(0.07-8.8)	0.858	2.7(0.5-16.7)	0.272
			def*A/A	1		1	
			def*G/A	1.9(0.2-24.9)	0.603	0.8(0.2-3.6)	0.774
			def*G/G	1		1	
			het*A/A	1			

Table 4.18: Test for genetic interactions between G6PD A with APOL1 genotypes among participants in terms of gender.

			Interaction type	Female		Male	
SNP	Gene	Allele	G6PD*APOL1	OR(95% CI)	P-value	OR(95% CI)	P-value
rs73885319	APOL1	A	wt*A/G	1.6(0.7-3.7)	0.264	1.7(0.7-4.2)	0.271
			wt*G/G	**4.7(0.9-26.8)**	**0.075**	1.2(0.3-6.1)	0.789
			het*A/A	**3.5(0.9-13.1)**	**0.061**		
			het*A/G	**7.5(2.1-26.6)**	**0.002**		
			het*G/G	1.2(0.2-6.7)	0.85		
			def*A/A	0.8(0.2-2.6)	0.713	0.8(0.3-2.0)	0.623
			def*A/G	2.4(0.9-7.1)	0.097	2.0(0.7-5.7)	0.165
			def*G/G	1		0.7(0.2-3.3)	0.668
rs60910145	APOL1	T	wt*G/G	1.2(0.2-6.4)	0.789	2.3(0.4-13.5)	0.336
			wt*G/T	1.2(0.6-2.9)	0.564	2.0(0.8-4.9)	0.121
			het*T/T	2.7(0.8-9.0)	0.096		
			het*G/G	1.8(0.3-11.2)	0.514		
			het*G/T	**4.8(1.4-16.9)**	**0.014**		
			def*T/T	0.5(0.2-1.6)	0.269	1.0(0.4-2.4)	0.994
			def*G/G	1		1.4(0.4-5.7)	0.614
			def*G/T	2.6(0.9-7.9)	0.087	2.0(0.8-5.2)	0.167
rs71785313	APOL1		wt*Ins/Del	**0.3(0.09-0.8)**	**0.03**	2.4(0.3-22.0)	0.427
			het*Ins	1.3(0.4-4.5)	0.57		
			het*Ins/Del	1			
			def*Ins	0.9(0.3-2.6)	0.846	0.7(0.3-1.8)	0.464
			def*Ins/Del	0.7(0.1-4.1)	0.675	0.4(0.09-1.3)	0.121
Rs9622363	APOL1	A	wt*G/A	3.1(0.2-56.3)	0.438	0.9(0.2-3.4)	0.88
			wt*G/G	3.9(0.2-75.9)	0.363		0.374
			het*A/A	1			
			het*G/A	6.1(0.3-118.3)	0.235		
			het*G/G	1			
			def*A/A	1		1	
			def*G/A	1.9(0.1-36.9)	0.658	0.5(0.2-1.9)	0.362
			def*G/G	1		1	

*figures in red indicate statistically significant interactions combination of genotypes whiles figures in blue indicate interactions that are not statistically significant but show a trend towards association.

4.4.4.1 Test for gene-gene interactions between G6PD and MYH9 genotype

The test for genetic interactions between the variants G6PD A^- , G6PD A and MYH9 genotypes are summarized in tables 4.19 and 4.20. Interaction analysis between G6PD A- and MYH9 genotypes revealed statistically significant positive epistatic interactions between mutant G6PD A- genotype (A/A) and both homozygous rs4821481 (C/C) (OR 13.0, 95% CI 1.6-104.9, p = 0.016) and heterozygous rs4821481 (C/T) (OR 5.5, 95% CI 1.1-27.1, p = 0.036) females. Although there was no observable association between MYH9 SNPs and CKD of unknown aetiology, there was a statistically significant epistatic interactions observed for heterozygous (OR 5.5, 95% CI 1.1-27.1, P =0.036) and homozygous (OR 13.0, 95% CI 1.6-104.9, P=0.016) females with the rs4821481 and mutant G6PD A- genotype. Interactions analysis also revealed that epistatic interactions between heterozygous rs5750248 and G6PD A genotype (OR 7.8, 95% CI 1.1-54.7, P = 0.03). The caveat in these analysis however is that they have a very wide confidence interval, an indication of small sample size utilized for the analysis.

Table 4.19: Test for genetic interactions between G6PD A- and MYH9 genotypes among participants in terms of gender.

			Interaction type	Female		Male	
SNP	Gene	Allele	G6PD*APOL1	OR(95% CI)	P-value	OR(95% CI)	P-value
rs4821481	MYH9	C	wt*C/T	0.9(0.4-1.9)	0.689	1.3(0.6-2.8)	0.553
			def*C/C	**13.0(1.6-104.9)**	**0.016**	1.1(0.5-2.4)	0.805
			def*C/T	**5.5(1.1-27.1)**	**0.036**	1.4(0.5-3.6)	0.531
			def*T/T	1.8(0.2-18.9)	0.622	1	
			het*C/C	1.4(0.5-3.9)	0.509		
			het*C/T	0.7(0.2-1.9)	0.505		
			het*T/T	0.6(0.04-10.2)	0.726		
rs5750250	MYH9	A	wt*G/A	1.0(0.2-5.1)	0.986	0.6(0.1-3.5)	0.586
			def*A/A	1.5(0.2-14.3)	0.729	0.8(0.06-13.4)	0.919
			def*G/A	**7.5(0.9-61.7)**	**0.062**	0.6(0.09-3.3)	0.514
			def*G/G	1		0.4(0.07-2.2)	0.286
			het*A/A	0.6(0.02-14.3)	0.737		
			het*G/A	0.7(0.1-4.0)	0.712		
			het*G/G	1.4(0.2-8.6)	0.749		
rs5750248	MYH9	C	wt*C/T	0.6(0.1-3.8)	0.599	0.96(0.2-5.8)	0.969
			wt*T/T	0.6(0.09-3.9)	0.608	0.5(0.08-2.6)	0.374
			def*C/C	1.0(0.09-12.0)	0.949	1	
			def*C/T	4.4(0.5-42.3)	0.2	0.7(0.1-4.1)	0.652
			def*T/T	1		0.6(0.09-3.5)	0.541
			het*C/C	0.4(0.01-10.9)	0.591		
			het*C/T	0.4(0.05-2.4)	0.291		
			het*T/T	1.4(0.2-10.8)	0.728		

* Figures in red indicate statistically significant interactions combination of genotypes whiles figures in blue indicate interactions that are not statistically significant but show a trend towards association.

Table 4.20: Test for genetic interactions between G6PD A with MYH9 genotypes among participants in terms of gender.

			INTERACTION TYPE	Female		Male	
SNP	Gene	Allele	G6PD*APOL1	OR(95% CI)	P-value	OR(95% CI)	P-value
rs4821481	MYH9	C	wt*C/T	0.7(0.3-1.5)	0.352	1.7(0.7-4.1)	0.269
			het*C/C	1.4(0.5-3.7)	0.542		
			het*C/T	**4.4(0.9-21.6)**	**0.068**		
			def*C/C	1.5(0.5-4.1)	0.44	1.0(0.5-2.3)	0.861
			def*C/T	0.7(0.2-2.1)	0.502	1.0(0.4-2.3)	0.954
			def*T/T	0.9(0.1-6.2)	0.948	2.1(0.2-19.7)	0.526
rs5750250	MYH9	A	wt*G/A	3.5(0.6-19.9)	0.162	0.9(0.2-5.6)	0.94
			wt*G/G	2.8(0.5-17.0)	0.248	0.3(0.06-1.8)	0.205
			het*A/A	1			
			het*G/A	**10.4(1.6-68.6)**	**0.015**		
			het*G/G	5.2(0.7-37.6)	0.103		
			def*A/A	9.2(0.6-145.7)	0.114	1.0(0.07-16.1)	0.982
			def*G/A	2.7(0.4-16.9)	0.279	0.5(0.09-2.9)	0.444
			def*G/G	**7.3(0.9-54.4)**	**0.052**	0.4(0.06-2.0)	0.245
rs5750248	MYH9	C	wt*C/T	2.2(0.4-13.3)	0.394	0.4(0.04-3.6)	0.402
			wt*T/T	3.2(0.5-19.9)	0.212	0.2(0.016-1.4)	0.092
			het*C/C	1			
			het*C/T	**7.8(1.1-54.7**	**0.037**		
			het*T/T	5.5(0.7-40.8)	0.095		
			def*C/C	9.7(0.6-154.4)	0.106	0.5(0.02-10.9)	0.665
			def*C/T	2.0(0.3-13.5)	0.467	0.2(0.2-2.1)	0.187
			def*T/T	5.4(0.7-39.7)	0.1	0.2(0.019-1.59)	0.122

*figures in red indicate statistically significant interactions combination of genotypes whiles figures in blue indicate interactions that are not statistically significant but show a trend towards association.

CHAPTER FIVE

DISCUSSION

The main aim of this work was to determine if G6PD variants A^- and A, the most common variants in West Africa had a role to play in increasing the risk of CKD especially of CKD with unknown aetiology. This aim was in accordance with identifying potential additional factors that could contribute to disease susceptibility as stated by Freeman *et al*, (2012). This work also sought to genotype only patients with CKD of unknown aetiology for the first time in order to report on their genetic predisposition to CKD unlike previous studies by other researchers who associated genetic factors to CKD of known aetiology especially for non-diabetic forms of CKD, Hypertension associated nephropathy and HIV associated nephropathy (Freedman et al., 2009b; Genovese et al., 2010; Shay Tzur et al., 2010a). This study also sought to determine if the genetic association reported in African Americans and some parts of Africa in non-diabetic CKD patients could also apply to patients with CKD of unknown aetiology in the major renal clinical centres in Ghana and Nigeria. Finally, this study was aimed at determining any epistatic interactions between G6PD, APOL1 and MYH9 gene variants and the risk of CKD of unknown aetiology. This work was restricted to just these variants that have been reported to have association with non-diabetic forms of CKD (Tayo *et al*., 2013).

Only the 2 G6PD variants were significantly out of Hardy-Weinberg equilibrium (p = 0.000 for G6PD A and p = 7.62E-10 for G6PD A-) and hence violated the Hardy-Weinberg principle whiles the APOL1 and MYH9 variants were all in Hardy-Weinberg equilibrium. The first reason to account for the violation of the Hardy-Weinberg principle by the G6PD SNPs was to assume that there was some errors

during the genotyping process. This reason however was nullified due to the quality control checks that was strictly adhered to during the genotyping assays (i.e. the use of Positive and negative controls as well as the Amelogenin marker to ensure the results obtained was the true reflection of the genotype from the population).

The violations of the Hardy-Weinberg equilibrium can therefore be attributed to the small sample size that was used for this study which could have caused a random change in the allele frequency. Also, violation of the Hardy-Weinberg principle by the G6PD SNPs could be likely due to the selective advantage of these mutations since they are known to be in high frequencies in the sample population. The G6PD variants are known to be in high frequency because of the selective advantage it confers against Malaria especially in the sample population under study (Howes et al., 2012; Ruwende et al., 1998; Tsegaye, Golassa, Mamo, & Erko, 2014).

Glucose-6-Phosphate Dehydrogenase is one of the oldest and commonest enzymopathies that occurs in humans. Previous studies have suggested that the high rise in frequency of the mutant gene is because of the selective advantage it had on the malaria parasite (Howes et al, 2013; Nkhoma et al., 2009; Orimadegun & Sodeinde, 2011). This is confirmed by the fact that there is an overlap in geographical regions between areas that have high prevalence of the G6PD mutations and where malaria is endemic, similar to the evolution of the APOL1 gene variants and the resistance to African human trypanosomiasis.

The prevalence of G6PD A^- in the study population was found to be 33.3% and 28.5% for the Ghana and Nigeria study sites respectively, whiles that of G6PD A was found to be 44.1% and 45.7% (Table 4.8 and 4.9). This is in conformity with the >25% estimated by the WHO (Howes *et al*., 2016). This happens to be higher than any of

the previously reported prevalence in relation to asymptomatic malaria resistance in children and pregnant women previously reported in Ghana and Nigeria (Williams *et al*., 2013; Amoah *et al*., 2016). It can however be argued that those previous reports were skewed to only pregnant women and children under 5 years. This notwithstanding, there was no significant difference between the prevalence observed in this study compared to what has been reported in current literature (p-value = 0.5710 for G6PD A- and 0.1005 for G6PD A). For CKD patients however, the prevalence recorded was significantly higher (49% and 53% for G6PD variants A^- and A respectively. P-value = 0.0001) and confirms earlier reports by Jayasekara *et al* 2013 about the increased prevalence of G6PD deficiency in CKD patients with unknown aetiology. This study did not look at the personal characteristic of the study subjects in order to predict the likely cause of the triggers that could have led to their cases of haemolysis which could have been the initiator of an acute kidney failure leading to CKD.

Similar to Ghana and Nigeria, there is little or no public health education about the effects of G6PD and its known triggers. Because there is a high prevalence of G6PD deficient carriers as well as a high rate of the abuse of over the counter drugs, many individuals who do not know their G6PD status could be damaging their kidneys unknowingly and as such, there should be an immediate cause to raise public awareness of the effects of this genetic defect and how to avoid known triggers.

In terms of gender, the prevalence of G6PD deficiency compared between case and control population was significantly different (p-value = 0.0001) and was also observed to be generally higher in women compared to that of the male case and control subjects. This however does not agree with previously reported literature (Zhang *et al*., 2008) where the prevalence was reported to be higher in males than in

females. Nonetheless, this was expected because most of the time, studies that record prevalence of G6PD variants use the enzyme activity test, which does not give the true reflection of the genotype of the population. For instance, heterozygous females may appear to be negative for an enzyme activity test which will lead to a misrepresentation of the true prevalence of the deficiency. There have been several instances where false negatives have been reported and were corrected after the genotyping was done to confirm the observed enzyme activity. This was confirmed as it was also observed that majority of female patients (55.5% of mutant females) were heterozygous for the G6PD deficiency (and may have been missed or misrepresented by previous studies to be negative or having a normal enzyme activity.

On the contrary, observed genotype does not always correlate with enzyme activity. This means that not all individuals who had the mutant genotype actually have lower enzyme activity. It will be prudent to determine the enzyme activity of samples participants to observe this trend in other to determine whether that can also increase or reduce the risk of disease.

Analysis of genetic association data has evolved from the use of simple statistical tools such as the simple chi-squared tests for analysing multivariate (logistic regression) tests (Moonesinghe et al., 2010). In some cases, the use of the simple chi-squared test is enough if the data is not powered enough for the multivariate analysis (i.e. in the case where sample size is too small or most samples fail to pass the quality control checks). This notwithstanding, the most complete analysis of the association data is the logistic regression with key epidemiological covariates and known genetic factors that may influence the observed associations (Kao et al., 2008).

In this study, we observed for the first time a significant association between G6PD, MYH9 and APOL1 gene variants and the incidence of CKD of unknown aetiology. Even though the association between APOL1 and MYH9 has been reported for known aetiologies of CKD like FSGS, hypertension attributed CKD and HIV associated CKD, the association observed with the G6PD variants however is the first to be observed for the sub-Saharan African population with direct emphasis on Ghana and Nigeria. Previous reports made by Jayasekara *et al* in Sri Lanka confirmed that it is possible for individuals who develop acute renal injury as a result of G6PD deficiency to progress to CKD if diagnosis is not done earlier and treatment is not immediate (Jmkb *et al*., 2013). Because the sample population used for this study was from two different sites, it was prudent to present the logistic regression models adjusting for age, sex and clinical site location since any of these factors could have caused the observed odds to shift. This was confirmed when the odds ratio was seen to have increased moderately in some cases which shows that these factors could also play a role in the association that was observed.

The APOL1 gene encodes the apolipoprotein L-1, whose known functions include autophagic cell death (Wan *et al*., 2008), lipid metabolism, cellular senescence and other vascular and biological activities (Monajemi & al, 2002; Vanhollebeke & Pays, 2006). Out of the family of 6 closely spaced and related APOL genes, only APOL1 is known to have evolved to develop a signal peptide, which enables it to be classified as both a circulating as well as an intracellular protein (Vanhollebeke & Pays, 2006). It is this feature that enables it to play a crucial role in the lytic activity of the human serum to many species of trypanosome because the mutations that occurred are all present in the signal peptide. Studies suggest that, even though the G1 and G2 mutations are both equally instrumental in the protection against the Trypanosoma

parasite, it appears that the G1 records a higher frequency than the G2 and as such, the absence of a statistically significant association between G2 and CKD is likely due to its 'low power to detection' as a result of the low frequencies recorded. This notwithstanding, the high frequencies of these variants confirms the theory that they do confer a selective advantage to the African population (Genovese *et al*., 2010). This can also explain why they did not obey the Hardy-Weinberg principle in this study aside the use of a small sample size.

The minor allelic frequencies (MAF) recorded for G1 (rs60910145 and rs73885319) were 0.66 and 0.37 respectively. Similar to previous reports that studied the association between APOL1 variants and non-diabetic forms of CKD, this present study also observed significant associations for CKD cases of unknown aetiology for both G1 variants in all models but not for rs9622363 and rs73885319 which is consistent with what has been reported for non-diabetic forms of CKD (Kopp *et al*., 2011; Shay Tzur *et al*., 2010a). This is the first time the association has been reported in CKD of unknown aetiology and goes to buttress the fact that APOL1 variants do play a critical role in increasing the risk of Africans in developing CKD with or without a known aetiology. Significant association was however observed for rs9622363 using the heterozygous model (OR 0.5, 95%CI 0.2530-0.9447, p = 0.033) and showed a 50% decrease in risk of disease for carriers of this genotype. Individuals who were heterozygous for variants rs60910145, rs73885319 and rs9622363 had a significantly higher risk of disease compared to individuals who were homozygous for the variants. Recessive models of association however did not reveal any significant association, which seems to suggest that the recessive model may not be the best model to describe the association between these variants and CKD of unknown aetiology even though the recessive model best fits FCGS associated CKD

and hypertension attributed CKD from previous reports (Genovese *et al*., 2010). This however agrees with earlier reports by Freedman *et al* about the likelihood that individuals with a single risk allele could have an increased risk of disease compared to individuals with two risk alleles (Friedman *et al*., 2011). This trend could be attributed to the other possible genetic or environmental factors that are present in the population under discussion. Earlier reports have confirmed this theory especially in relation to HIV and hypertension associated forms of CKD (Bostrom *et al*., 2012; Divers *et al*., 2014) and could likely be a similar pattern for CKD cases of unknown aetiology.

There was an increase in the Odds ratio of rs9622363 for after adjusting for age, gender and clinic location. The odds ratio for rs73885319 and rs60910145 however did not change after adjusting for those factors. This shows that for the rs9622363 SNP, factors like age and gender or the location of the clinical site contributed to the increase in risk of disease association compared to the other SNPs.

Variant rs71785313 did not have any statistically significant associations even after adjusting for parameters such as age and gender. This observation is consistent with reported literature for most non-diabetic forms of CKD and as such could likely be as a result of the low frequencies that are found in the sub-Saharan African region when compared to the other variants (Genovese *et al*., 2010).

Gene-gene interaction analysis revealed that variants rs60910145 and rs73885319 recorded significant epistatic interactions between the deficient G6PD variants which lead to the subsequent increase in the risk of disease (OR 12.6, 95%CI 2.5-62.4, p = 0.002 for rs73885319, OR 9.2, 95%CI 1.9-44.4, p = 0.006 for rs60910145).This type of epistatic interaction is known as the polymeric gene interaction, where the

individual alleles have a different effect but produces an enhanced effect when they are together. However the caveat here is that, the confidence intervals were broad, indicative of the small sample sizes used in the analysis. A larger sample size will be required in order to confirm these numbers that were observed and confirm the interactions observed. Interestingly, interactions between heterozygous rs71785313 variant and the wild type G6PD A variant was observed for the female population but was not statistically significant.

Even though no significant association was observed between G6PD A^-, G6PD A and CKD unknown aetiology in the male study subjects, significant epistatic interactions were observed with the deficient G6PD A- variants and heterozygous APOL1 variants rs73885319 (OR 2.7 95%CI 1.0-6.8, p = 0.04) and rs60910145 (OR 2.3 95% CI 1.0-5.2, p = 0.046) which revealed that individuals with these genotypes have 30% risk of disease. There was a trend towards association was also observed for rs60910145 and deficient G6PD A- (OR 2.5 95% CI 0.9-6.1, p = 0.053). This gives an indication of a likely synergistic epistatic interaction between these variants that possibly leads to an increase in risk of disease for individuals with those SNP combinations which could explain earlier observations of no significant associations between CKD of unknown etiology and this variant. This shows that, for males who are G6PD deficient, presence of at least one risk variant of rs60910145 makes them susceptible to kidney damage.

Interactions was also observed between the heterozygous G6PD genotype and other APOL1 variants but were all not statistically significant. It will be good to replicate this analysis using a larger sample size to see if the interactions observed here could become significant.

The MYH9 gene product, myosin-IIA, is localized to the podocyte foot process and is responsible for moving actin filaments in cells. The proteins comprising the filtration slit barrier actively regulate actin dynamics to maintain normal cell structure. Mutations affecting other podocyte proteins rearrange the actin cytoskeleton, ultimately disrupting the filtration barrier and causing renal disease.

Unlike previous studies where association was found between MYH9 gene variants and CKD (Freedman *et al*., 2009a; Jeffrey B Kopp *et al*., 2008), This study did not record any statistically significant association between CKD of unknown aetiology and the MYH9 genetic variants, even after adjusting for factors such as age, gender and clinical location. Significant associations was however observed during the gene-gene interaction analysis for all the MYH9 SNPs even though most of them were not statistically significant.

For interaction analysis between MYH9 and G6PD A^- genotypes, significant interactions were observed between G6PD A- deficient and both heterozygous and homozygous genotypes of the rs4821481 ([def*C/C] OR13.0 95% CI 1.6-104.9, P= 0.016; [def*C/T] OR 5.5 95% CI 1.1-27.1, p = 0.036). This is an indication of a synergistic epistatic interaction between G6PD A- and at least one risk allele of rs4821481. The wide confidence interval observed for both however shows that the samples size went into this analysis was small and would have to be confirmed using a larger sample size. It was also observed that there was an increasing trend of association was observed for heterozygous rs5750250 and G6PD A^- deficient genotype, although not statistically significant (OR 7.5, 95% CI 0.9-61.7, p = 0.062).

In the interaction analysis for G6PD A variants, it was observed that heterozygous G6PD A genotype and heterozygous rs4821481 showed an increasing trend towards

association (OR 4.4, 95% CI 0.9-21.6, p = 0.068) whiles heterozygous rs5750250 and deficient G6PD A (OR 10.4, 95%CI 1.6-68.6, p = 0.015) as well as wildtype rs57050250 and deficient G6PD A (OR 7.3, 95% CI 0.9-54.4, p = 0.052). The interactions between heterozygous G6PD A genotype and rs5750248 (OR 7.8, 95%CI 1.1-54.7, p = 0.037) showed significant interactions that resulted in increased risk of disease. This seems to suggest that for individuals who have the MYH9 mutations especially at least one risk allele of either rs4821481, rs5750250 or rs4821481, and genetic factors like G6PD A- or A could be the likely cause of the increased association with the disease possibly as a result of the synergistic epistatic interactions occurring between these genotypes.

Because research have shown that rs5750248 and rs5750250 which belong to the S1 haplotype, and rs4821481 and rs2032487 which also belongs to the E 1 haplotype are known to be in complete linkage disequilibrium and are associated with CKD, most often than not, they are mostly inherited together (Behar et al., 2010a). This indicates the presence of an observed association in one of these SNPs could mean an association for both. Finding significant associations for rs482481 and rs5750250 after interacting with the G6PD variants in this study could therefore imply that the other SNPs that belong to these haplotypes could also be associated with CKD of unknown aetiology. This study however did not look at these other SNP because the effect of the small sample size.

Since all MYH9 risk alleles do not actually develop kidney disease, environmental exposures or interactions with other genes are clearly necessary to initiate MYH9-associated nephropathy in genetically susceptible hosts. This was confirmed in this study when some SNPs of MYH9 showed significant association after interaction analysis with some G6PD genotypes. Further interaction analysis between G6PD and

other MYH9 SNPs can be evaluated to determine whether the interactions observed here applies to any other MYH9 SNPs.

Studies have shown that G6PD deficiency causes haemolytic anaemia, usually after individuals who have the deficiency are exposed to the known triggers such as certain types of anti-malarial or antibacterial drugs, infections like viral hepatitis, certain foods like fava beans as well as some chemicals and herbal medications. Literature also shows that the deficiency was only associated to acute renal failure until scientist in Sri Lanka observed that, patients who reported to the hospital with CKD of unknown aetiology had very low levels of enzyme activity. This led to the conclusion that, the high prevalence of G6PD may be an indication that G6PD deficiency could have played a role in the pathogenesis of the disease (Jmkb *et al*., 2013). Since then, other studies have shown that G6PD deficiency anaemia is associated with variable degree of acute renal injury during haemolytic episodes which can persist if not detected at the early stages.

In this study, it was observed that there was a higher prevalence of G6PD A- in both males (34.7%) and females (28%). As well as G6PD A males (42.1%) and females (38.5%). The prevalence recorded in this study however appeared to be higher that what has been previously reported for some parts of Nigeria (Orimadegun & Sodeinde, 2011; Williams *et al*., 2013b) and Ghana (Amoah *et al*., 2016; Amoako *et al*., 2014). It can be argued however that, most of these studies only considered an aspect of the whole population (i.e. pregnant women and children under 5 years). Also, most of these studies only did the enzyme activity test to determine enzyme deficiency. This test however does not give the true representation of the deficiency because it has a high rate of false positive and negatives. There are instances where individuals have shown to be negative for the enzyme test but turn out to be positive

during the genotype test. Also, the enzyme activity does not detect heterozygous deficiency in females. As a result of this, most females may be either regarded as negative or positive which may not be a true reflection of their G6PD status. Prevalence of G6PD A- (43.2% for males and 56% for females) and A (47.7% for males and 62.6% for females) in CKD cases on unknown aetiology was however higher than what was reported for the control samples. Generally, the prevalence recorded for females appeared to be higher for both G6PD variants and were also statistically different (p-value = 0.0001 for A- and 0.0007 for A) whiles no statistical difference was observed for the male participants (p-value = 0.2128 for G6PD A- and 0.4205 for G6PD A).

Results from this study showed very strong evidence of an increased risk in CKD of unknown aetiology with both variants A- and A. Because G6PD is an X linked gene, association analysis was done according to gender. Females who had two risk alleles had the greatest risk for CKD with odds ratio ranging from 2 to 6.8 using the addictive model. Meaning that females who are homozygous for the G6PD variant are at more risk of disease compared to females without the mutation. Males recorded odds ratio from 1.3 to 1.4 for both variants. Association with males however did not appear to be very significant, a situation which can be attributed to the small sample size.

Even though subsequent *in vitro* and *in vivo* studies have reported the link between acute renal failure and CKD, this is the first time a genetic association study has revealed that individuals with G6PD deficiency are at an increased risk of developing CKD that is not associated with hypertension, HIV or FSGS.

In all the models of association that was run for G6PD variant A- and A, it appears that the disease association is observed more in individuals with 2 risk alleles which indicates that the recessive nature of the disease. This confirms why the odds of homozygous females (who have 2copies of the SNP) are at a higher risk than hemizygous males (who have 1 copy of the SNP).

Interaction analysis was also able to identify a strong association between the deficient genotype of G6PD with heterozygous APOL1 whiles an increasing trend was observed for interactions between G6PD A- or G6PD A and MYH9 variants. This could likely explain the reason as to why individuals with 2 risk alleles for only APOL1 and MYH9 do not actually develop CKD in their lifetime.

CHAPTER SIX

CONCLUSION AND RECOMMENDATION

6.0 Conclusion

Results from this study shows that G6PD variant A (A376G) and G6PD variant A- (G202A) have a high prevalence in the populations under study. Also for the first time, this study has revealed that both G6PD variants A and A- as well as APOL1 risk variants (rs73885319 and rs60910145) are associated with CKD of unknown aetiology. Using the logistic regression analysis, genetic interactions between APOL1, MYH9 and G6PD risk variants A and A- increases the risk of developing CKD of unknown aetiology in the population under study, and as such, G6PD variants A and A- is a possible risk factor for the development of CKD of unknown aetiology. This study also showed that having two or at least one APOL1 and MYH9 risk variants is necessary, but not sufficient for the development (or progression) of kidney disease of unknown etiology. Environmental modifiers such as excessive abuse of over the counter drugs have been identified to be the main cause of the triggers of G6PD deficiency and as such should be closely monitored with public awareness so as to reduce the incidence of G6PD associated renal injury.

Finally we can conclude that G6PD deficiency may play a critical role in the pathogenesis of CKD of unknown aetiology.

6.1 Recommendation

For future studies, larger sample size from both Ghana and Nigeria, as well as other sub-Saharan African countries should be included in order to confirm the role G6PD could play in the pathogenesis of CKD especially of unknown aetiology. Studies

should also be conducted to determine the direct mechanisms by which G6PD contributes to kidney failure.

As a result of the high prevalence of G6PD deficiency that was observed in the study population, it should be made mandatory for all clinics and hospitals to check the status of their patients before prescribing any drug to them, especially for kidney patients and people who require treatment for malaria as some of these drugs are known triggers that could cause severe episodes of haemolysis leading to the onset of kidney damage.

Public health awareness should be raised in order to educate people about G6PD deficiency as well as the APOL1 and the MYH9 variants, in other to create the public awareness. The public should also be educated on the potential dangers of abusing over the counter drugs and the abuse of the known triggers of the deficiency. This will help in reducing the onset of acute kidney failure and by extension, to CKD.

Further studies will have to be conducted to elucidate the direct mechanism by which APOL1 and MYH9 variants could contribute to kidney disease in CKD of unknown aetiology as has been done for FSGS and HIV associated nephropathies (Freedman *et al*., 2009b). These results may improve understanding of the pathogenesis and mechanisms underlying the progression of this diverse spectrum of non-diabetic kidney diseases and assist in explaining why not all individuals possessing two APOL1 or MYH9 risk variants ultimately develop kidney disease.

Finally, Individuals should also be encouraged to pick up a healthy lifestyle, in terms of eating healthy, taking alcohol in moderation and avoiding sedentary lifestyles as well as exercising regularly to prevent damaging their kidneys.

REFERENCES

Addo, J., Smeeth, L., Leon, D. A., Opie, L., Seedat, Y., Seedat, Y., … Calhoun, H. (2009). Hypertensive Target Organ Damage in Ghanaian Civil Servants with Hypertension. *PLoS ONE*, *4*(8), e6672. http://doi.org/10.1371/journal.pone.0006672

Almaguer, M., Herrera, R., & Orantes, C. M. (2014). Chronic kidney disease of unknown etiology in agricultural communities. *MEDICC Review*, *16*(2), 9–15. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/24878644

Amoah, L. E., Opong, A., Ayanful-Torgby, R., Abankwa, J., & Acquah, F. K. (2016). Prevalence of G6PD deficiency and Plasmodium falciparum parasites in asymptomatic school children living in southern Ghana. *Malaria Journal*, *15*(1), 388. http://doi.org/10.1186/s12936-016-1440-1

Amoako, N., Asante, K. P., Adjei, G., Awandare, G. A., Bimi, L., & Owusu-Agyei, S. (2014). Associations between red cell polymorphisms and plasmodium falciparum infection in the middle belt of Ghana. *PLoS ONE*, *9*(12), 1–15. http://doi.org/10.1371/journal.pone.0112868

Arogundade, F. A., Barsoum, R. S. (2008). CKD prevention in Sub-Saharan Africa: a call for governmental, nongovernmental, and community support. *American Journal of Kidney Diseases*, *51*(3), 515–523.

Arrondel, C., Vodovar, N., Knebelmann, B., Grünfeld, J.-P., Gubler, M.-C., Antignac, C., & Heidet, L. (2002). Expression of the nonmuscle myosin heavy chain IIA in the human kidney and screening for MYH9 mutations in Epstein and Fechtner syndromes. *Journal of the American Society of Nephrology : JASN*, *13*(1), 65–74. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11752022

Ashley-Koch, A. E., Okocha, E. C., Garrett, M. E., Soldano, K., de Castro, L. M., Jonassaint, J. C., … Telen, M. J. (2011). MYH9 and APOL1 are both associated with sickle cell disease nephropathy. *British Journal of Haematology*, *155*(3), 386–394. http://doi.org/10.1111/j.1365-2141.2011.08832.x

Au, S. W. N., Gover, S., Lam, V. M. S., & Adams, M. J. (2000). Human glucose-6-phosphate dehydrogenase: The crystal structure reveals a structural NADP+ molecule and provides insights into enzyme deficiency. *Structure*, *8*(3), 293–303. http://doi.org/10.1016/S0969-2126(00)00104-0

Baynes, J. W. (1991). Role of oxidative stress in development of complications in diabetes. *Diabetes*, *40*(4), 405–12. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/2010041

Behar, D. M., Rosset, S., Tzur, S., Selig, S., Yudkovsky, G., Bercovici, S., … Skorecki, K. (2010a). 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*, *19*(9), 1816–1827. http://doi.org/10.1093/hmg/ddq040

Behar, D. M., Rosset, S., Tzur, S., Selig, S., Yudkovsky, G., Bercovici, S., … Skorecki, K. (2010b). 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*, *19*(9), 1816–1827. http://doi.org/10.1093/hmg/ddq040

Bello, A. K., Peters, J., Rigby, J., Rahman, A. A., & El Nahas, M. (2008). Socioeconomic status and chronic kidney disease at presentation to a renal service in the United Kingdom. *Clinical Journal of the American Society of*

Nephrology : CJASN, *3*(5), 1316–23. http://doi.org/10.2215/CJN.00680208

Berg, J. M., Tymoczko, J. L., & Stryer, L. (2002). Biochemistry.

Bergrem, H., & Leivestad, T. (2001). Diabetic nephropathy and end-stage renal failure: the Norwegian story. *Advances in Renal Replacement Therapy*, *8*(1), 4–12. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11172323

Betteridge, D. J. (2000). What is oxidative stress? *Metabolism: Clinical and Experimental*, *49*(2 Suppl 1), 3–8. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/10693912

Beutler, E. (1996). A series of new screening procedures for pyruvate kinase deficiency, glucose-6-phosphate dehydrogenase deficiency, and glutathione reductase deficiency. *Blood*, (28), 553–62.

Beutler, E., Duparc, S., & G6PD Deficiency Working Group. (2007). Glucose-6-phosphate dehydrogenase deficiency and antimalarial drug development. *The American Journal of Tropical Medicine and Hygiene*, *77*(4), 779–89. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/17978087

Beutler, E., & Vulliamy, T. J. (2002). Hematologically important mutations: glucose-6-phosphate dehydrogenase. *Blood Cells, Molecules & Diseases*, *28*(2), 93–103. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12064901

Bostrom, M. A., Kao, W. H. L., Li, M., Abboud, H. E., Adler, S. G., Iyengar, S. K., … Family Investigation of Nephropathy and Diabetes (FIND) Research Group. (2012). Genetic association and gene-gene interaction analyses in African American dialysis patients with nondiabetic nephropathy. *American Journal of Kidney Diseases : The Official Journal of the National Kidney Foundation*, *59*(2), 210–21. http://doi.org/10.1053/j.ajkd.2011.09.020

Bouanga, J. C., Mouele, R., Prehu, C., Wajcman, H., Feingold, J., & Galacteros, F. (1998). Glucose-6-phosphate dehydrogenase deficiency and homozygous sickle cell disease in congo. *Human Heredity*, *48*(4), 192–197. http://doi.org/10.1159/000022801

Cappellini, M. D., & Fiorelli, G. (2008). Glucose-6-phosphate dehydrogenase defi ciency. *Seminar 64 Www.thelancet.com*, *371*.

Carroll, R. G., & Abdel-Rahman, A. A. (2011). Regulation of acid-base balance. In *xPharm: The Comprehensive Pharmacology Reference*. http://doi.org/10.1016/B978-008055232-3.60298-3

Carter, N., Pamba, A., Duparc, S., & Waitumbi, J. N. (2011). Frequency of glucose-6-phosphate dehydrogenase deficiency in malaria patients from six African countries enrolled in two randomized anti-malarial clinical trials. *Malaria Journal*, *10*(1), 241. http://doi.org/10.1186/1475-2875-10-241

Cocco, P. (1998). Occupational Lead Exposure and Screening of Glucose-6-Phosphate Dehydrogenase Polymorphism: Useful Prevention or Nonvoluntary Discrimination? - Journals - NCBI, *71*(2), 148–150. 3. Retrieved from https://www.ncbi.nlm.nih.gov/labs/articles/9553792/

Couser, W. G., Remuzzi, G., Mendis, S., & Tonelli, M. (2011). The contribution of chronic kidney disease to the global burden of major noncommunicable diseases. *Kidney International*, *80*(12), 1258–1270. http://doi.org/10.1038/ki.2011.368

Daoud, B. Ben, Mosbehi, I., Préhu, C., Chaouachi, D., Hafsia, R., & Abbes, S. (2008). Molecular characterization of erythrocyte glucose-6-phosphate dehydrogenase deficiency in Tunisia Caractérisation moléculaire du déficit en glucose-6-phosphate déshydrogénasse érythrocytaire en Tunisie. *Pathologie Biologie*, *56*,

260–267. http://doi.org/10.1016/j.patbio.2007.08.009

Di Angelantonio, E., Danesh, J., Eiriksdottir, G., Gudnason, V., & Naidoo, D. (2007). Renal Function and Risk of Coronary Heart Disease in General Populations: New Prospective Study and Systematic Review. *PLoS Medicine*, *4*(9), e270. http://doi.org/10.1371/journal.pmed.0040270

Divers, J., Palmer, N. D., Lu, L., Langefeld, C. D., Rocco, M. V., Hicks, P. J., … Freedman, B. I. (2014). Gene-gene interactions in APOL1-associated nephropathy. *Nephrology Dialysis Transplantation*, *29*(3), 587–594. http://doi.org/10.1093/ndt/gft423

Dorgalaleh, A., Shahzad, M. S., & Younesi, M. R. (2013). Evaluation of liver and kidney function in favism patients, *27*(1), 17–22.

Eckardt, K.-U., & Kasiske, B. L. (2009). KDIGO clinical practice guideline for the diagnosis, evaluation, prevention, and treatment of Chronic Kidney Disease-Mineral and Bone Disorder (CKD-MBD). *Kidney International*, *76*(113), S1–S2. http://doi.org/10.1038/ki.2009.188

Edwards, C. Q. (2002). Anemia and the liver. Hepatobiliary manifestations of anemia. *Clinics in Liver Disease*, *6*(4), 891–907, viii. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12516198

Fanello, C. I., Karema, C., Avellino, P., Bancone, G., Uwimana, A., Lee, S. J., … Modiano, D. (2008). High Risk of Severe Anaemia after Chlorproguanil-Dapsone+Artesunate Antimalarial Treatment in Patients with G6PD (A-) Deficiency. *PLoS ONE*, *3*(12), e4031. http://doi.org/10.1371/journal.pone.0004031

Fox, C. S., Larson, M. G., Leip, E. P., Culleton, B., Wilson, P. W. F., & Levy, D.

(2004). Predictors of New-Onset Kidney Disease in a Community-Based Population. *JAMA*, *291*(7), 844. http://doi.org/10.1001/jama.291.7.844

Freedman, B. I., Hicks, P. J., Bostrom, M. A., Cunningham, M. E., Liu, Y., Divers, J., … Bowden, D. W. (2009a). Polymorphisms in the non-muscle myosin heavy chain 9 gene (MYH9) are strongly associated with end-stage renal disease historically attributed to hypertension in African Americans. *Kidney International*, *75*(7), 736–745. http://doi.org/10.1038/ki.2008.701

Freedman, B. I., Hicks, P. J., Bostrom, M. A., Cunningham, M. E., Liu, Y., Divers, J., … Bowden, D. W. (2009b). Polymorphisms in the non-muscle myosin heavy chain 9 gene (MYH9) are strongly associated with end-stage renal disease historically attributed to hypertension in African Americans. *Kidney International*, *75*(7), 736–45. http://doi.org/10.1038/ki.2008.701

Freedman, B. I., Kopp, J. B., Langefeld, C. D., Genovese, G., Friedman, D. J., Nelson, G. W., … Pollak, M. R. (2010). The apolipoprotein L1 (APOL1) gene and nondiabetic nephropathy in African Americans. *Journal of the American Society of Nephrology : JASN*, *21*(9), 1422–1426. http://doi.org/10.1681/ASN.2010070730

Freedman, B. I., Langefeld, C. D., Lu, L., Divers, J., Comeau, M. E., Kopp, J. B., … Bowden, D. W. (2011). Differential effects of MYH9 and APOL1 risk variants on FRMD3 association with diabetic ESRD in African Americans. *PLoS Genetics*, *7*(6), 1–8. http://doi.org/10.1371/journal.pgen.1002150

Freedman, B. I., & Skorecki, K. (2014). Gene???gene and gene???environment interactions in apolipoprotein L1 gene-associated nephropathy. *Clinical Journal of the American Society of Nephrology*, *9*(11).

http://doi.org/10.2215/CJN.01330214

Friedman, D. J., Kozlitina, J., Genovese, G., Jog, P., & Pollak, M. R. (2011). Population-Based Risk Assessment of APOL1 on Renal Disease. *Journal of the American Society of Nephrology*. http://doi.org/10.1681/ASN.2011050519

Gaetani, G..,Kirkman, H.N., Mangrini, R., Ferraris, A. M. (1994). No Title. *Blood*, *84*, 325–330.

Garg, A. X., Kiberd, B. A., Clark, W. F., Haynes, R. B., & Clase, C. M. (2002). Albuminuria and renal insufficiency prevalence guides population screening: Results from the NHANES III. *Kidney International*, *61*(6), 2165–2175. http://doi.org/10.1046/j.1523-1755.2002.00356.x

Genovese, G., Friedman, D. J., & Pollak, M. R. (2013). APOL1 variants and kidney disease in people of recent African ancestry. *Nature Reviews Nephrology*, *9*(4). http://doi.org/10.1038/nrneph.2013.34

Genovese, G., Friedman, D. J., Ross, M. D., Lecordier, L., Uzureau, P., Freedman, B. I., & Bowden, D. W. (2010). Association of Trypanolytic ApoL1 Variants with Kidney Disease in African-Americans. *Science. August*, *13*(3295993), 841–845. http://doi.org/10.1126/science.1193032

Gross ML, A. K. (2004). Progression of renal disease: new insight into risk factors and pathomechanisms. *Curr Opin Nephrol Hypertens.*, *13*, 307–312.

Gutierrez, O. M., Judd, S. E., Irvin, M. R., Zhi, D., Limdi, N., Palmer, N. D., … Freedman, B. I. (2016). APOL1 nephropathy risk variants are associated with altered high-density lipoprotein profiles in African Americans. *Nephrology Dialysis Transplantation*, *31*(4), 602–608. http://doi.org/10.1093/ndt/gfv229

Hildebrand, F., Meyer, A., & Eyre-Walker, A. (2010). Evidence of Selection upon

Genomic GC-Content in Bacteria. *PLoS Genetics*, *6*(9), e1001107. http://doi.org/10.1371/journal.pgen.1001107

Howes, R. E., Battle, K. E., Satyagraha, A. W., Baird, J. K., & Hay, S. I. (2013). G6PD Deficiency. Global Distribution, Genetic Variants and Primaquine Therapy. *Advances in Parasitology*, *81*. http://doi.org/10.1016/B978-0-12-407826-0.00004-7

Howes, R. E., Piel, F. B., Patil, A. P., Nyangiri, O. A., Gething, P. W., Dewi, M., … Hay, S. I. (2012). G6PD Deficiency Prevalence and Estimates of Affected Populations in Malaria Endemic Countries: A Geostatistical Model-Based Map. *PLoS Medicine*, *9*(11). http://doi.org/10.1371/journal.pmed.1001339

Ishida, K., Ishida, H., Narita, M., Sairenchi, T., Saito, Y., Fukutomi, H., … Koyama, A. (2001). Factors affecting renal function in 119 985 adults over three years. *QJM*, *94*(10).

Ito, K., Bick, A. G., Flannick, J., Friedman, D. J., Genovese, G., Parfenov, M. G., … Seidman, C. (2014). Increased burden of cardiovascular disease in carriers of APOL1 genetic variants. *Circulation Research*, *114*(5). http://doi.org/10.1161/CIRCRESAHA.114.302347

Jha, V., Garcia-Garcia, G., Iseki, K., Li, Z., Naicker, S., Plattner, B., … Yang, C.-W. (2013). Chronic kidney disease: global dimension and perspectives. *The Lancet*, *382*(9888), 260–272. http://doi.org/10.1016/S0140-6736(13)60687-X

Jha, V., & Prasad, N. (2016). CKD and Infectious Diseases in Asia Pacific: Challenges and Opportunities. *American Journal of Kidney Diseases*, *68*(1), 148–160. http://doi.org/10.1053/j.ajkd.2016.01.017

Jmkb, J., Dm, D., Mdn, G., Sivakanesan, R., & Dmts, D. (2013). Research Article

Prevalence of G6Pd Deficiency in Patients With Chronic Kidney Disease of Unknown Origin in North Central Region of Sri Lanka : Case Control Study, *4*(July 2016), 455–458.

Jungers, P., Chauveau, P., Descamps-Latscha, B., Labrunie, M., Giraud, E., Man, N. K., … Jacobs, C. (1996). Age and gender-related incidence of chronic renal failure in a French urban area: a prospective epidemiologic study. *Nephrology, Dialysis, Transplantation : Official Publication of the European Dialysis and Transplant Association - European Renal Association*, *11*(8), 1542–6. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8856208

Kao, W. H. L., Klag, M. J., Meoni, L. A., Reich, D., Berthier-schaad, Y., Li, M., … L, P. (2009). NIH Public Access. *October*, *40*(10), 1185–1192. http://doi.org/10.1038/ng.232.A

Kasembeli, A. N., Duarte, R., Ramsay, M., Mosiane, P., Dickens, C., Dix-Peek, T., … Naicker, S. (2015). APOL1 Risk Variants Are Strongly Associated with HIV-Associated Nephropathy in Black South Africans. *Journal of the American Society of Nephrology : JASN*, (January 2016), ASN.2014050469-. http://doi.org/10.1681/ASN.2014050469

Katz, A. I., & Lindheimer, M. D. (1977). Actions of Hormones on the Kidney. *Annual Review of Physiology*, *39*(1), 97–133. http://doi.org/10.1146/annurev.ph.39.030177.000525

Kazancioğlu, R. (2013). Risk factors for chronic kidney disease: an update. *Kidney International Supplements*, *3*(4), 368–371. http://doi.org/10.1038/kisup.2013.79

Klag, M. J., Whelton, P. K., Randall, B. L., Neaton, J. D., Brancati, F. L., Ford, C. E., … Stamler, J. (1996). Blood Pressure and End-Stage Renal Disease in Men. *New*

England Journal of Medicine, *334*(1), 13–18. http://doi.org/10.1056/NEJM199601043340103

Kopp, J. B., Nelson, G. W., Sampath, K., Johnson, R. C., Genovese, G., An, P., … Winkler, C. A. (2011). APOL1 Genetic Variants in Focal Segmental Glomerulosclerosis and HIV-Associated Nephropathy. *Journal of the American Society of Nephrology*, *22*(11), 2129–2137. http://doi.org/10.1681/ASN.2011040388

Kopp, J. B., Smith, M. W., Nelson, G. W., Johnson, R. C., Freedman, B. I., Bowden, D. W., … Winkler, C. A. (2008). MYH9 is a major-effect risk gene for focal segmental glomerulosclerosis. *Nature Genetics*, *40*(10), 1175–1184. http://doi.org/10.1038/ng.226

Kopp, J. B., Winkler, C. a., & Nelson, G. W. (2011). MYH9 Genetic Variants Associated With Glomerular Disease: What Is the Role of Genetic Testing. *Semin Nephrol*, *30*(4), 409–417. http://doi.org/10.1016/j.semnephrol.2010.06.007.MYH9

Kopp, J. B., Winkler, C. A., & Nelson, G. W. (2010). MYH9 Genetic Variants Associated With Glomerular Disease: What Is the Role for Genetic Testing? *Seminars in Nephrology*, *30*(4), 409–417. http://doi.org/10.1016/j.semnephrol.2010.06.007

Köttgen, A., Glazer, N. L., Dehghan, A., Hwang, S.-J., Katz, R., Li, M., … Fox, C. S. (2009). Multiple loci associated with indices of renal function and chronic kidney disease. *Nature Genetics*, *41*(6), 712–717. http://doi.org/10.1038/ng.377

Kruzel-Davila, E., Wasser, W. G., Aviram, S., & Skorecki, K. (2016). APOL1 nephropathy: From gene to mechanisms of kidney injury. *Nephrology Dialysis*

Transplantation, *31*(3), 349–358. http://doi.org/10.1093/ndt/gfu391

Kunishima, S., Kitamura, K., Matsumoto, T., Sekine, T., & Saito, H. (2014). Somatic mosaicism in *MYH9* disorders: the need to carefully evaluate apparently healthy parents. *British Journal of Haematology*, *165*(6), 885–887. http://doi.org/10.1111/bjh.12797

Kurts, C., Panzer, U., Anders, H.-J., & Rees, A. J. (2013). The immune system and kidney disease: basic concepts and clinical implications. *Nature Reviews Immunology*, *13*(10), 738–753. http://doi.org/10.1038/nri3523

Levey A.S., Coresh J., Balk E., Kausz A.T., Levin A., Steffes M.W., Hogg R.J., Perrone R.D., L. J. and E. G. (2003). National Kidney Foundation practice guidelines for chronic kidney disease: evaluation, classification, and stratification. *Ann Intern Med*.

Limou, S., Dummer, P. D., Nelson, G. W., Kopp, J. B., & Winkler, C. A. (2015). APOL1 toxin, innate immunity, and kidney injury. *Kidney International*, *88*(1), 28–34. http://doi.org/10.1038/ki.2015.109

Limou, S., Nelson, G. W., Kopp, J. B., & Winkler, C. A. (2014). APOL1 Kidney Risk Alleles: Population Genetics and Disease Associations. *Advances in Chronic Kidney Disease*. http://doi.org/10.1053/j.ackd.2014.06.005

Linda Kao, W., Klag, M. J., Meoni, L. A., Reich, D., Berthier-Schaad, Y., Li, M., … Genet Author, N. (2008). A genome-wide admixture scan identifies MYH9 as a candidate locus associated with non-diabetic end stage renal disease in African Americans. *Nat Genet*, *40*(10), 1185–1192. http://doi.org/10.1038/ng.232

Lipkowitz, M. S., Freedman, B. I., Langefeld, C. D., Comeau, M. E., Bowden, D. W., Linda Kao, W. H., … SK Investigators. (2013). Apolipoprotein L1 gene variants

associate with hypertension-attributed nephropathy and the rate of kidney function decline in African Americans. *Kidney International*, *83*(1), 114–120. http://doi.org/10.1038/ki.2012.263

Luzzatto, L. (2006). Glucose 6-phosphate deficiency: from genotype to phenotype. *Hematology*, *2*, 63–68.

Luzzatto, L. (2011). G6PD deficiency and malaria selection. *Heredity*, *108*(456). http://doi.org/10.1038/hdy.2011.90

Madhavan, S. M., O'Toole, J. F., Konieczkowski, M., Ganesan, S., Bruggeman, L. A., & Sedor, J. R. (2011). APOL1 Localization in Normal Kidney and Nondiabetic Kidney Disease. *Journal of the American Society of Nephrology*, *22*(11), 2119–2128. http://doi.org/10.1681/ASN.2011010069

Maestrini, E., Rivella, S., Tribioli, C., Rocchi, M., Camerino, G., Santachiara-Benerecetti, S., … Toniolo, D. (1992). Identification of novel RFLPs in the vicinity of CpG islands in Xq28: application to the analysis of the pattern of X chromosome inactivation. *Am J Hum Genet*, *50*(1), 156–163.

Maiga, B., Dolo, A., Campino, S., Sepulveda, N., Corran, P., Rockett, K. a, … Clark, T. G. (2014). Glucose-6-phosphate dehydrogenase polymorphisms and susceptibility to mild malaria in Dogon and Fulani, Mali. *Malaria Journal*, *13*(1), 270. http://doi.org/10.1186/1475-2875-13-270

Manganelli, G., Masullo, U., Passarelli, S., & Filosa, S. (2013). Glucose-6-phosphate dehydrogenase deficiency: disadvantages and possible benefits. *Cardiovascular & Hematological Disorders Drug Targets*, *13*(1), 73–82. http://doi.org/10.2174/1871529X11313010008

Mason, P.J., Bautisa, J.M., Gilsanz, F. (2007). G6PD deficiency: the genotype-

phenotype association. *Blood Cells, Molecules, and Diseases*, *21*, 267–283.

Mason, P. J. (1996). New insights into G6PD deficiency. *British Journal of Haematology*, *94*, 585–91.

May, J., Meyer, C. G., Grossterlinden, L., Ademowo, O. G., Mockenhaupt, F. P., Olumese, P. E., … Bienzle, U. (2000). Red cell glucose-6-phosphate dehydrogenase status and pyruvate kinase activity in a Nigerian population. *Tropical Medicine and International Health*, *5*(2), 119–123. http://doi.org/10.1046/j.1365-3156.2000.00529.x

McKnight, R. E., Gleason, A. B., Keyes, J. A., & Sahabi, S. (2007). Binding mode and affinity studies of DNA-binding agents using topoisomerase I DNA unwinding assay. *Bioorganic and Medicinal Chemistry Letters*, *17*(4), 1013–1017. http://doi.org/10.1016/j.bmcl.2006.11.038

Minucci, A., Moradkhani, K., Hwang, M. J., Zuppi, C., Giardina, B., & Capoluongo, E. (2012). Glucose-6-phosphate dehydrogenase (G6PD) mutations database: Review of the "old" and update of the new mutations. *Blood Cells, Molecules, and Diseases*, *48*(3), 154–165. http://doi.org/10.1016/j.bcmd.2012.01.001

MO Afolabi MBBS, MPH, FMCFM, FWACP, EA Abioye-Kuteyi MBBS, F., FWACP, FRACGP, FA Arogundade MBBS, FMCP, F. & I. B. M., & FMCGP. (2009). Prevalence of chronic kidney disease in a Nigerian family practice populatione. *South African Family Practice*, *51*(2), 132–137. http://doi.org/10.1080/20786204.2009.10873828

Modaresi, A., Nafar, M., & Sahraei, Z. (2015). Oxidative stress in chronic kidney disease. *Iranian Journal of Kidney Diseases*, *9*(3), 165–79. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/25957419

Monajemi, & al. (2002). The Apolipoprotein L Gene Cluster Has Emerged Recently inEvolution and Is Expressed in Human Vascular Tissue. *GENOMICS*, *79*(4). http://doi.org/10.1006

Moonesinghe, R., Yesupriya, A., Chang, M. h., Dowling, N. F., Khoury, M. J., & Scott, A. J. (2010). A Hardy-Weinberg Equilibrium Test for Analyzing Population Genetic Surveys With Complex Sample Designs. *American Journal of Epidemiology*, *171*(8), 932–941. http://doi.org/10.1093/aje/kwq002

Moore, J. H., & Williams, S. M. (2009). Epistasis and Its Implications for Personal Genetics. *American Journal of Human Genetics*, *85*(3), 309–320. http://doi.org/10.1016/j.ajhg.2009.08.006

National Kidney Foundation. (2002). *K/DOQI clinical practice guidelines for chronic kidney disease: evaluation, classification, and stratification*. *American journal of kidney diseases : the official journal of the National Kidney Foundation* (Vol. 39). Retrieved from http://linkinghub.elsevier.com/retrieve/pii/S0272638602700905%5Cnhttp://www.ncbi.nlm.nih.gov/pubmed/11904577

National Kidney Foundation. (2014). *How kidneys work*. Retrieved from htttp: www. kidney.org/kidneydisease/howkidneyswrk

Naylor, C. E., Rowland, P., Basak, A. K., Gover, S., Mason, P. J., Bautista, J. M., … Adams, M. J. (1996). Glucose 6-phosphate dehydrogenase mutations causing enzyme deficiency in a model of the tertiary structure of the human enzyme. *Blood*, *87*(7), 2974–82. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8639919

Nkhoma, E. T., Poole, C., Vannappagari, V., Hall, S. A., & Beutler, E. (2009). The

global prevalence of glucose-6-phosphate dehydrogenase deficiency: A systematic review and meta-analysis. *Blood Cells, Molecules, and Diseases*, *42*(3), 267–278. http://doi.org/10.1016/j.bcmd.2008.12.005

O'Seaghdha, C. M., Parekh, R. S., Hwang, S., Li, M., Ko, A., Coresh, J., … Kao, W. H. L. (2011). The MYH9 / APOL1 region and chronic kidney disease in European-Americans. *Human Molecular Genetics*. http://doi.org/10.1093/hmg/ddr118

Orimadegun, A. E., & Sodeinde, O. (2011). Glucose-6-phosphate dehydrogenase status and severity of malarial anaemia in Nigerian children. *Journal of Infection in Developing Countries*, *5*(11), 792–8. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/22112733

Osafo, C., Mate-Kole, M., Affram, K., & Adu, D. (2011). Prevalence of Chronic Kidney Disease in Hypertensive Patients in Ghana. *Renal Failure*, *33*(4), 388–392. http://doi.org/10.3109/0886022X.2011.565140

Osafo, C., Raji, Y. R., Burke, D., Tayo, B. O., Tiffin, N., Moxey-Mims, M. M., … Parekh, R. S. (2015). Human heredity and health (H3) in africa kidney disease research network: A focus on methods in sub-Saharan Africa. *Clinical Journal of the American Society of Nephrology*, *10*(12), 2279–2287. http://doi.org/10.2215/CJN.11951214

Ouattara, A. K., Bisseye, C., Bazie, B. V. J. T. E., Diarra, B., Compaore, T. R., Djigma, F., … Simpore, J. (2014). Glucose-6-phosphate dehydrogenase (G6PD) deficiency is associated with asymptomatic malaria in a rural community in Burkina Faso. *Asian Pacific Journal of Tropical Biomedicine*, *4*(8), 655–8. http://doi.org/10.12980/APJTB.4.2014APJTB-2014-0100

Page, N. M., Butlin, D. J., Lomthaisong, K., & Lowry, P. J. (2001). The human apolipoprotein L gene cluster: identification, classification, and sites of distribution. *Genomics*, *74*, 71–78. http://doi.org/10.1006/geno.2001.6534

Palmer, N. D., Ng, M. C. Y., Langefeld, C. D., Divers, J., Lea, J. P., Okusa, M. D., … Freedman, B. I. (2015). Lack of Association of the APOL1 G3 Haplotype in African Americans with ESRD. *Journal of the American Society of Nephrology*, *26*(5), 1021–1025. http://doi.org/10.1681/ASN.2014050444

Philippe, M., Larondelle, Y., Lemaigre, F., Mariamé, B., Delhez, H., Mason, P., … Rousseau, G. G. (1994). Promoter function of the human glucose-6-phosphate dehydrogenase gene depends on two GC boxes that are cell specifically controlled. *European Journal of Biochemistry*, *226*(2), 377–84. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/8001555

Phompradit, P., Kuesap, J., Chaijaroenkul, W., Rueangweerayut, R., Hongkaew, Y., Yamnuan, R., & Na-Bangchang, K. (2011). Prevalence and distribution of glucose-6- phosphate dehydrogenase (G6PD) variants in Thai and Burmese populations in malaria endemic areas of Thailand. *Malaria Journal*, *10*. http://doi.org/10.1186/1475-2875-10-368

Poggi, V., Town, M., Foulkes, N. S., Luzzatto, L., & Luzzattot, L. (1990). Identification of a single base change in a new human mutant glucose-6-phosphate dehydrogenase gene by polymerase-chain- reaction amplification of the entire coding region from genomic DNA. *The Biochemical Journal*, *271*(1), 157–160.

Ruwende, C., Hill, A., Campino, S., Sepulveda, N., Corran, P., Rockett, K. A., … Williams, S. (1998). Glucose-6-phosphate dehydrogenase deficiency and

malaria. *Journal of Molecular Medicine*, *76*(8), 581–588. http://doi.org/10.1007/s001090050253

Safian, R. D., & Textor, S. C. (2001). Renal-Artery Stenosis. *New England Journal of Medicine*, *344*(6), 431–442. http://doi.org/10.1056/NEJM200102083440607

Simons, M., Wang, M., McBride, O. W., Kawamoto, S., Yamakawa, K., Gdula, D., … Weir, L. (1991). Human nonmuscle myosin heavy chains are encoded by two genes located on different chromosomes. *Circulation Research*, *69*(2), 530–9. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1860190

Smith, E. E., & Malik, H. S. (2009). The apolipoprotein L family of programmed cell death and immunity genes rapidly evolved in primates at discrete sites of host-pathogen interactions. *Genome Research*, *19*(5), 850–858. http://doi.org/10.1101/gr.085647.108

Stenvinkel, P. (2003). Anaemia and inflammation: what are the implications for the nephrologist? *Nephrol Dial Transplant*, *18*, 17–22. http://doi.org/10.1093/ndt/gfg1086

Tayo, B. O., Kramer, H., Salako, B. L., Gottesman, O., McKenzie, C. a., Ogunniyi, A., … Cooper, R. S. (2013). Genetic variation in APOL1 and MYH9 genes is associated with chronic kidney disease among Nigerians. *International Urology and Nephrology*, *45*(2), 485–494. http://doi.org/10.1007/s11255-012-0263-4

Teteris, S. A., Engel, D. R., & Kurts, C. (2011). Homeostatic and pathogenic role of renal dendritic cells. *Kidney International*, *80*(2), 139–145. http://doi.org/10.1038/ki.2011.129

Thomson, R., Genovese, G., Canon, C., Kovacsics, D., Higgins, M. K., Carrington, M., … Raper, J. (2014a). Evolution of the primate trypanolytic factor APOL1.

Proceedings of the National Academy of Sciences of the United States of America, *111*(20), E2130-9. http://doi.org/10.1073/pnas.1400699111

Thomson, R., Genovese, G., Canon, C., Kovacsics, D., Higgins, M. K., Carrington, M., … Raper, J. (2014b). Evolution of the primate trypanolytic factor APOL1. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(20), E2130-9. http://doi.org/10.1073/pnas.1400699111

Tishkoff, S. A. (2001). Haplotype Diversity and Linkage Disequilibrium at Human G6PD: Recent Origin of Alleles That Confer Malarial Resistance. *Science*, *293*(5529), 455–462. http://doi.org/10.1126/science.1061573

Tittel, A. P., Heuser, C., Ohliger, C., Llanto, C., Yona, S., Hämmerling, G. J., … Kurts, C. (2012). Functionally relevant neutrophilia in CD11c diphtheria toxin receptor transgenic mice. *Nature Methods*, *9*(4), 385–390. http://doi.org/10.1038/nmeth.1905

Toniolo, D., Filippi, M., Dono, R., Lettieri, T., & Martini, G. (1991). The CpG island in the 5' region of the G6PD gene of man and mouse. *Gene*, *102*(2), 197–203. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1874446

Trask, B. J., Massa, H., Kenwrick, S., & Gitschier, J. (1991). Mapping of human chromosome Xq28 by two-color fluorescence in situ hybridization of DNA sequences to interphase cell nuclei. *American Journal of Human Genetics*, *48*(1), 1–15. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/1985451

Tsegaye, A., Golassa, L., Mamo, H., & Erko, B. (2014). Glucose-6-phosphate dehydrogenase deficiency among malaria suspects attending Gambella hospital, southwest Ethiopia. *Malaria Journal*. http://doi.org/10.1186/1475-2875-13-438

Turin, T. C., Ahmed, S. B., Tonelli, M., Manns, B., Ravani, P., James, M., …

Hemmelgarn, B. (2014). Kidney function, albuminuria and life expectancy. *Canadian Journal of Kidney Health and Disease*, *1*, 33. http://doi.org/10.1186/s40697-014-0033-6

Tzur, S., Rosset, S., Shemer, R., Yudkovsky, G., Selig, S., Tarekegn, A., … Skorecki, K. (2010a). Missense mutations in the APOL1 gene are highly associated with end stage kidney disease risk previously attributed to the MYH9 gene. *Human Genetics*, *128*(3), 345–350. http://doi.org/10.1007/s00439-010-0861-0

Tzur, S., Rosset, S., Shemer, R., Yudkovsky, G., Selig, S., Tarekegn, A., … Skorecki, K. (2010b). Missense mutations in the APOL1 gene are highly associated with end stage kidney disease risk previously attributed to the MYH9 gene. *Human Genetics*, *128*(3), 345–350. http://doi.org/10.1007/s00439-010-0861-0

Tzur, S., Rosset, S., Shemer, R., Yudkovsky, G., Selig, S., Tarekegn, A., … Skorecki, K. (2010c). Preliminary Report: Missense mutations in the APOL gene family are associated with end stage kidney disease risk previously attributed to the MYH9 gene. *Hum Genet.*, *128*(3), 345–50.

Tzur, S., Rosset, S., Skorecki, K., & Wasser, W. G. (2012). APOL1 allelic variants are associated with lower age of dialysis initiation and thereby increased dialysis vintage in African and Hispanic Americans with non-diabetic end-stage kidney disease. *Nephrology Dialysis Transplantation*, *27*(4), 1498–1505. http://doi.org/10.1093/ndt/gfr796

Vanhamme, L., Paturiaux-Hanocq, F., Poelvoorde, P., Nolan, D. P., Lins, L., Van Den Abbeele, J., … Pays, E. (2003). Apolipoprotein L-I is the trypanosome lytic factor of human serum. *Nature*, *422*(6927), 83–87. http://doi.org/10.1038/nature01461

Vanhollebeke, B., & Pays, E. (2006). The function of apolipoproteins L. *Cellular and Molecular Life Sciences*, *63*(17), 1937–1944. http://doi.org/10.1007/s00018-006-6091-x

Voskarides, K., Demosthenous, P., Papazachariou, L., Arsali, M., Athanasiou, Y., Zavros, M., … Deltas, C. (2013). Epistatic Role of the MYH9/APOL1 Region on Familial Hematuria Genes. *PLoS ONE*, *8*(3), 1–7. http://doi.org/10.1371/journal.pone.0057925

Vulliamy, T. J., Kaeda, J. S., Ait-Chafa, D., Mangerini, R., Roper, D., Barbot, J., … Mason, P. J. (1998). Clinical and haematological consequences of recurrent G6PD mutations and a single new mutation causing chronic nonspherocytic haemolytic anaemia. *British Journal of Haematology*, *101*(4), 670–5. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9674740

Vulliamy, T. J., Othman, a, Town, M., Nathwani, a, Falusi, a G., Mason, P. J., & Luzzatto, L. (1991). Polymorphic sites in the African population detected by sequence analysis of the glucose-6-phosphate dehydrogenase gene outline the evolution of the variants A and A-. *Proceedings of the National Academy of Sciences of the United States of America*, *88*(19), 8568–8571. http://doi.org/10.1073/pnas.88.19.8568

Wan, G., Zhaorigetu, S., Liu, Z., Kaini, R., Jiang, Z., & Hu, C. A. A. (2008). Apolipoprotein L1, a novel Bcl-2 homology domain 3-only lipid-binding protein, induces autophagic cell death. *Journal of Biological Chemistry*, *283*(31), 21540–21549. http://doi.org/10.1074/jbc.M800214200

Weaver, V. M., Fadrowski, J. J., & Jaar, B. G. (2015). Global dimensions of chronic kidney disease of unknown etiology (CKDu): a modern era environmental and/or

occupational nephropathy? *BMC Nephrology*, *16*, 145. http://doi.org/10.1186/s12882-015-0105-6

Weckerle, A., Snipes, J. A., Cheng, D., Gebre, A. K., Reisz, J. A., Murea, M., … Ma, L. (2016). Characterization of circulating APOL1 protein complexes in African Americans. *Journal of Lipid Research*, *57*(1), 120–30. http://doi.org/10.1194/jlr.M063453

WHO | World Malaria Report 2008. (2013). *WHO*.

Williams, O., Gbadero, D., Edowhorhu, G., Brearley, A., Slusher, T., & Lund, T. C. (2013a). Glucose-6-Phosphate Dehydrogenase Deficiency in Nigerian Children. *PLoS ONE*, *8*(7). http://doi.org/10.1371/journal.pone.0068800

Williams, O., Gbadero, D., Edowhorhu, G., Brearley, A., Slusher, T., & Lund, T. C. (2013b). Glucose-6-phosphate dehydrogenase deficiency in Nigerian children. *PloS One*, *8*(7), e68800. http://doi.org/10.1371/journal.pone.0068800

Wouters, O. J., O 'donoghue, D. J., Ritchie, J., Kanavos, P. G., & Narva, A. S. (2015). Early chronic kidney disease: diagnosis, management and models of care. *Nat Rev Nephrol*, *11*(8), 491–5. http://doi.org/10.1038/nrneph.2015.85

Young, J. H., Klag, M. J., Muntner, P., Whyte, J. L., Pahor, M., & Coresh, J. (2002). Blood pressure and decline in kidney function: findings from the Systolic Hypertension in the Elderly Program (SHEP). *Journal of the American Society of Nephrology : JASN*, *13*(11), 2776–82. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/12397049

Yutaka, Nakahori Osamu, T. Y. N. (1991). A human X-Y homologous region encodes "amelogenin." *Genomics*, *9*(2), 264–269. http://doi.org/10.1016/0888-7543(91)90251-9

Zhang, L., Zhang, P., Wang, F., Zuo, L., Zhou, Y., Shi, Y., … Wang, H. (2008). Prevalence and Factors Associated With CKD: A Population Study From Beijing. *American Journal of Kidney Diseases*, *51*(3), 373–384. http://doi.org/10.1053/j.ajkd.2007.11.009

Zhaorigetu, S., Wan, G., Kaini, R., Jiang, Z., & Hu, C. A. A. (2008). ApoL1, a BH3-only lipid-binding protein, induces autophagic cell death. *Autophagy*, *4*(8), 1079–1082. http://doi.org/10.1074/jbc.M800214200.www.landesbioscience.com

Zhaorigetu, S., Yang, Z., Toma, I., McCaffrey, T. A., & Hu, C. A. A. (2011). Apolipoprotein L6, induced in atherosclerotic lesions, promotes apoptosis and blocks beclin 1-dependent autophagy in atherosclerotic cells. *Journal of Biological Chemistry*, *286*(31), 27389–27398. http://doi.org/10.1074/jbc.M110.210245

APPENDIX

CALCULATING THE SAMPLE SIZE FOR THE STUDY

The equation for calculating sample size for case-control studies is defined as

$$n = \left(\frac{r+1}{r}\right)\frac{(\bar{p})(1-\bar{p})(Z_{\beta}+Z_{\alpha/2})^2}{(p_1-p_2)^2}$$

N is Sample size for case

P_1 is the prevalence of case (19%)

P_2 is the prevalence of the control

r is the ratio of case to control (1)

$Z_{\alpha/2}$ is the type II error at a desired power

Z_{β} is the significant level

Therefore,

For 80% power, $Z_{\beta} = 0.84$

For 0.05 significance level, $Z_{\alpha} = 1.96$

For and equal ratio of case to control, r = 1

To calculate for prevalence of control group,

$$OR = \frac{P_2(1-P_1)}{P_1(1-P_2)}, \qquad 3 = \frac{P_2(1-0.19)}{0.19(1-P_2)}$$

$$0.57(1-P_2) = 0.81\,P_2, \qquad 0.57-0.57 = 0.81\,P_2,$$

$$P_2 = \frac{0.57}{(0.81+0.57)}$$

$$P_2 = 0.41$$

$$\bar{P} = \frac{(P_2 + rP_1)}{1+ r}, \qquad \bar{P} = \frac{(0.41+0.19)}{1 + 1}$$

$$\bar{P} = 0.3$$

$$N = \frac{1+1}{1} \frac{(0.3)(1-0.3)(0.84+1.96)^2}{(0.19-0.41)^2}$$

$$N = 68$$

Table1. Test for association between the various SNPS with participants using test of trends model.

SNP	Gene	Alleles	Case	Control	χ2(p-value)
rs73885319	APOL1	A/G	153(57.74)	62(43.97)	**9.8799(0.007)**
		G/G	28(10.57)	12(8.51)	
		A/A	84(31.70)	67(47.52)	
rs60910145	APOL1	G/T	141(53.21)	57(40.3)	**8.7012(0.013)**
		T/T	97(36.60)	73(51.77)	
		G/G	27(10.19)	11(7.8)	
rs71785313	APOL1	D/I	52(19.62)	17(26.15)	1.3464(0.246)
		I/I	52(19.62)	17(26.15)	
rs2032487	MYH9	T/C	88(33.59)	50(36.23)	0.8358(0.658)
		C/C	162(61.83)	84(60.87)	
		T/T	12(4.58)	4(2.90)	
rs4821481	MYH9	T/C	103(39.31)	55(39.86)	0.6665(0.717)
		C/C	147(56.11)	79(57.25)	
		T/T	12(4.58)	4(2.90)	
rs5750248	MYH9	C/T	130(48.69)	67(47.52)	1.4986(0.473)
		T/T	116(43.45)	67(47.52)	
		C/C	21(7.87)	7(4.96)	
rs5750250	MYH9	A/G	142(53.18)	66(46.81)	2.3784(0.304)
		G/G	104(38.95)	66(46.81)	
		A/A	21(7.87)	9(6.38)	
Rs9622363	APOL1	A/G	176(66.17)	52(80.00)	4.7615(0.092)
		G/G	80(30.08)	12(18.46)	
		A/A	10(3.76)	1(1.54)	

*columns shaded in red indicate statistically significant association for the model

Table 2. Test for association between the various SNPS with participants using genotype model

SNP	Gene	Alleles	Case	Control	χ2(p-value)
rs73885319	APOL1	G/G	28(25.00)	12(15.19)	2.6927(0.101)
		A/G	**153(64.56)**	**62(48.06)**	**9.3779(0.002)**
rs60910145	APOL1	G/G	27(21.77)	11(13.10)	2.5262(0.112)
		G/T	**141(59.24)**	**57(43.85)**	**8.0197(0.005)**
rs71785313	APOL1	Ins/Del	52(19.62)	17(26.15)	1.3464(0.246)
rs2032487	MYH9	T/T	12(6.90)	4(4.55)	0.5634(0.453)
		T/C	88(35.20)	50(37.31)	0.1693(0.681)
rs4821481	MYH9	T/T	12(7.55)	4(4.82)	0.6572(0.418)
		T/C	103(41.20)	55(41.04)	0.0009(0.976)
rs5750248	MYH9	T/T	116(84.67)	67(90.54)	1.4380(0.23)
		C/T	130(86.09)	67(90.54)	0.9017(0.342)
rs5750250	MYH9	G/G	104(83.20)	66(88.0)	0.8471(0.357)
		A/G	142(87.12)	66(88.0)	0.0364(0.849)
Rs9622363	MYH9	G/G	80(88.89)	12(92.31)	0.1392(0.709)
		A/G	176(94.62)	52(98.11)	1.1439(0.285)

*columns shaded in red indicate statistically significant association for the model

Table 3. Test for association between the various SNPS with participants using heterozygous model

SNP	Gene	Allele	Case	Control	$\chi 2$(p-value)
rs73885319	APOL1	A/G G/G*A/A	153(57.74) 112(42.26)	62(43.97) 79(56.03)	**6.9988(0.008)**
rs60910145	APOL1	G/T T/T*G/G	141(53.21) 124(46.79)	57(40.3) 84(59.57)	**6.0181(0.014)**
rs71785313	APOL1	D/I I/I	52(19.62) 52(19.62)	17(26.15) 17(26.15)	0.246
rs2032487	MYH9	T/C C/C*T/T	88(33.59) 174(66.41)	50(36.23) 88(63.77)	0.2797(0.597)
rs4821481	MYH9	T/C C/C*T/T	103(39.31) 159(60.69)	55(39.86) 83(60.14)	0.0111(0.916)
rs5750248	MYH9	C/T T/T*C/C	130(48.69) 137(51.31)	67(47.52) 74(52.48)	0.0507(0.822)
rs5750250	MYH9	A/G G/G*A/A	142(53.18) 125(46.82)	66(46.81) 75(53.19)	1.5006(0.221)
Rs9622363	MYH9	A/G G/G*A/A	176(66.17) 90(33.83)	52(80.00) 13(20.00)	**4.6643(0.031)**

* Columns shaded in red indicate statistically significant association for the model

Table 4. Test for association between G6PD SNPS with female participants using genotype model

SNP	Gene	Allele	Case	Control	$\chi 2$(p-value)
rs1050828	G6PD A-(202)	G/A	25(21.37)	20(41.67)	**7.07(0.008)**
		A/A	92(37.86)	28(22.76)	**8.4448(0.004)**
rs1050829	G6PD A(376)	A/G	40(28.37)	11(17.74)	2.5853(0.108)
		G/G	101(44.30)	51(38.93)	0.9186(0.322)

* Columns shaded in red indicate statistically significant association for the model

Table 5. Test for association between G6PD SNPS with male participants using genotype model

SNP	Gene	Alleles	Case	Control	$\chi 2$(p-value)
rs1050828	G6PD A-(202)	A/A	92(37.86)	28(22.76)	**8.4448(0.004)**
rs1050829	G6PD A(376)	G/G	101(44.30)	51(38.93)	0.9186(0.322)

* Columns shaded in red indicate statistically significant association for the model

Table 6. Test for association between G6PD SNPS with female participants using test of trends model

SNP	Gene	Allele	Case	Control	χ2(p-value)
rs1050828	G6PD A-(202)	G/A	**25(19.23)**	**20(26.67)**	**14.4652(0.001)**
		A/A	**35(26.92)**	**4(5.33)**	
		G/G	**70(53.85)**	**51(68.0)**	
rs1050829	G6PD A(376)	A/G	40(30.77)	11(14.67)	7.5847(0.023)
		G/G	32(24.62)	18(24.0)	
		A/A	58(44.62)	46(61.33)	

* Columns shaded in red indicate statistically significant association for the model

Table 7. Test for association between G6PD SNPS with male participants using test of trends model

SNP	Gene	Allele	Case	Control	χ2(p-value)
rs1050828	G6PD A-(202)	G/A	57(41.30)	24(35.29)	0.6897(0.406)
		A/A	81(58.70)	44(64.71)	
		G/G			
rs1050829	G6PD A(376)	A/G	69(50.0)	33(49.25)	
		G/G	69(50.0)	34(50.75)	0.0100(0.920)
		A/A			

Table 8. Expected band sizes for the various APOL1 and MYH9 genotypes from the polyacrylamide gel electrophoresis

SNP	Allele	Band size/bp
rs60910145	G	62
	T	54
rs73885319	G	58
	A	50
rs71785313	Ins	90
	Del	82
rs2032487	T	58
	C	50
rs4821481	C	90
	T	82
rs5750248	C	58
	T	50
rs5750250	G	89
	A	81
rs9622363	G	62
	A	54
Amelogenin X	X	74
Amelogenin Y	Y	66

SOLUTION PREPARATION, SOURCE OF REAGENTS AND EQUIPMENT

1X TBE buffer

To prepare the working solution of 1X TBE buffer, 20 ml of the 50X TBE buffer (Thermo Scientific) was measured and diluted with distilled water to a total volume of 1 L. The pH of this solution was 8.0

1X SB buffer

To prepare the working solution of 1X SB buffer, 20 ml of the 50X SB buffer (Thermo Scientific) was measured and diluted with distilled water to a total volume of 1 L. The pH of this solution was 8.0

Agarose Gel (1%)

To prepare 1% agarose solution, 1 g of agarose was weighed and dissolved in 100 mL 1X SB buffer.

Ethidium bromide solution (10 mg/ ml)

To make 10 mg/ mL ethidium bromide solution, 0.1 g ethidium bromide tablet was weighed and dissolved in 10 mL 1X SB buffer

Source of Reagents and materials used

1. 5X PCR buffer (1nM $MgCl_2$), dNTPs (Inqaba biotech, Pretoria, South Africa),

2. Primers and Probes (Integrated DNA technologies (IDT), Leuven Belgium)

3. Vortex (Mini vortex, VMR Scientific Products, Atlanta, U.S.A)

4. Glass beads (Sigma-Aldrich, Darmstadt, Germany)

5. Thermal Cycler (Applied Biosystems, California, U.S.A)

6. Camera (Sony, Tokyo, Japan).

7. Agarose (AGTC Bioproducts Ltd, UK)

8. 100bp DNA ladder (Promega, Madison WI, USA)

9. Electrophoresis set up (CBS Scientific CA, USA)

10. Ethidium bromide (Promega, Madison WI, USA)

11. Benchtop 2UV Trans illuminator gel photography system (Upland, CA, USA)

12. Centrifuge (Eppendorf, Hamburg, Germany)

13. Sensoquest Labcycler PCR Machine (SensoQuest, Gottingen, Germany)

14. Nuclease free water (Hyclone Lab Inc., South Logan, Utah, USA)

15. 5U Taq Polymerase (Amresco. Fountain Park. Ohio, U.S.A)