

https://dspace.org/

# Genetic and molecular characterization of HIV-associated kidney disease in black South Africans

**Kasembeli, Alex Nganga**

**2016-02-22**
http://hdl.handle.net/10539/19679

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

# Genetic and molecular characterization of HIV-associated kidney disease in black South Africans

**Alex Nganga Kasembeli (Student number 593860)**

Faculty of Health Science

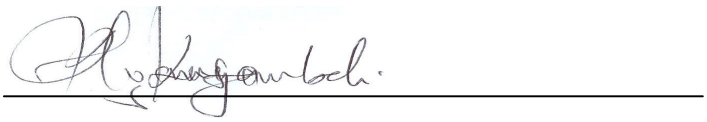
Department of Internal Medicine

A thesis submitted in fulfillment of the requirements for the degree of Doctor of Philosophy in Internal Medicine

August, 2015

# DECLARATION

I, Alex Nganga Kasembeli declare that the thesis, which I hereby submit for the degree of Doctor of Philosophy in Medicine at the University of the Witwatersrand, Johannesburg, South Africa, is my own work and has not been submitted by me or any other person for degree purposes at any other university



Alex Nganga Kasembeli (Signature).

7th day of August 2015

# ABSTRACT

Chronic kidney disease (CKD) due to HIV-associated nephropathy (HIVAN) remains an important cause of morbidity and mortality in African HIV-infected individuals. Genetic susceptibility to HIVAN has been attributed to genetic variations within *MYH9* and *APOL1* on chromosome 22. The expression of apolipoprotein L1 isoforms B and C coupled with high risk variants, play a crucial role in the development of kidney disease. The aims of this study were to determine the prevalence and examine the association of *MYH9* and/or *APOL1* high risk variants in HIVAN compared to other forms of CKD in HIV-infected individuals in a black South African population; to profile the pattern of messenger RNA *N-terminal* apolipoprotein L1 isoform expression in HIVAN, HIV-positive FSGS, other HIV-positive CKD, HIV-negative FSGS, other HIV-negative CKD, HIV-associated immune complex kidney disease (HIVICK) and in normal kidney; and to determine the expression of the *N-terminal* apolipoprotein L1 isoforms in human embryonic kidney (HEK) 293 cells.

Kidney biopsies from Charlotte Maxeke Johannesburg Academic Hospital were obtained from adults (aged ≥18 years) and selected according to histological diagnoses, clinical data were collected and patients were stratified according to age, gender and clinical presentation. Single nucleotide polymorphisms (SNPs) within and flanking *MYH9* and *APOL1* genes were selected by identifying all SNPs documented in 2011 from 1000 base pairs upstream and downstream of the two genes, and with a minor allele frequency of ≥0.05 in African populations based on the *International HapMap* and *1000 Genomes Projects*. SNPs with previous associations to kidney disease and literature SNPs were included, as was a panel of ancestry informative markers (AIMs). Genomic DNA was extracted from formalin fixed paraffin embedded (FFPE) kidney

tissues and from whole blood. SNP genotyping was performed using the Illumina BeadXpress SNP genotyping assay and the TaqMan SNP genotyping assay. N-terminal apolipoprotein L1 mRNA expression was performed using Reverse Transcription-Polymerase Chain Reaction (RT-PCR) method. Expression of N-terminal apolipoprotein L1 isoforms in HEK 293 cells was performed using PCR-based cloning and transfection techniques. Data analysis was performed using R program and Stata v11.1. Logic regression, Fisher's exact test, Kruskal Wallis test and Student's *t test* were employed in the data analysis.

A total of 228 samples [38 HIVAN, 41 HIV (+) CKD, 41 HIV (-) CKD, 54 HIV (+) Controls and 54 population controls] were successfully genotyped. Of 96 SNPs genotyped, 77 passed quality control. A total of nine *MYH9* and *APOL1* SNPs were significantly ($P< 0.05$) associated with HIVAN and not with other forms of CKD including focal segmental glomerulosclerosis (FSGS). However, the risk of *MYH9* SNPs with *P-values* < 0.05 could not be determined. Single *MYH9* E1 haplotype SNPs did not show any association, however, E1 haplotype block C-G-C showed a weaker significant factor with HIVAN (OR 3.45 *P=0.008*). The *APOL1* risk alleles [G1 (0.560), G1 (0.500) and G2 (0.342)] had the highest prevalence among HIVAN individuals. $G1^{GM}$ haplotype was the most frequent in HIVAN (0.530). 19/38 (50.0%) of the HIVAN cases were compound heterozygous for G1 and G2 risk alleles; 11/38 were homozygous for either G1 or G2 risk alleles, 6/38 were heterozygous for a single risk allele and 2/38 were homozygous for non-risk allele. Thus, 30/38 (79.0%) of HIVAN but only 1/54 (2%) of population controls carried two risk alleles. Population allele frequencies were 7.3% for G1 and 11.1% for G2. In a recessive model, individuals carrying two *APOL1* risk alleles had 89.0-fold higher odds 95% CI 17.68, 911.72; $P=1.24x10^{-14}$ for developing HIVAN. *APOL1* risk variants were not significantly associated with other forms of CKD [HIV (+) FSGS (OR=2.13 95% CI 0.03, 44.30; *P=0.48*) and

HIVICK (OR=5.60 95% CI 0.4, 86; *P=0.13*) in the HIV (+) CKD group compared to HIV (+) controls] and [Primary FSGS (OR=6.30 95% CI 0.04, 248.70; *P=0.26*) in the HIV (-) CKD group compared to population controls]. Isoform B *N-terminal apolipoprotein L1* mRNA levels were elevated in HIVAN and HIVICK while isoform C *N-terminal apolipoprotein L1* levels were elevated in normal kidney. We showed that the three isoforms are not only expressed in HEK 293 cells but the expression is variable where isoforms B and C *N-terminal apolipoprotein L1* that has a defective secretory domain and lack the secretory domain respectively, are seen to be more expressed within the HEK 293 cells than isoform A *N-terminal apolipoprotein L1*.

A combination of apolipoprotein L1 isoform B expression with C-terminal risk variants is a major contributor to the development of HIVAN and possibly other kidney diseases. There is a striking increase in the prevalence of *APOL1* risk variants among HIVAN compared to other forms of kidney disease and we estimate that HIV-infected South African blacks with two *APOL1* risk alleles not receiving anti-retroviral therapy (ART) have the greatest risk of developing HIVAN. Further studies are required to elucidate the mechanism of apolipoprotein L1-mediated kidney disease in order to develop effective therapeutic measures.

# DEDICATION

To my family for their support and encouragement: My beautiful wife, Hellen Kahaso Dena and my son, Ryan Kasembeli Nganga.

# ACKNOWLEDGEMENTS

I would like to thank God for giving me the wisdom, strength and endurance during this period and being so gracious and kind to me.

**My supervisors:**

- Dr. Raquel Duarte.
- Prof. Michèle Ramsay.
- Prof. Saraladevi Naicker.

For having faith in me, being incredible mentors, for their guidance, all-round support and shared knowledge throughout the study.

**Provision of samples:**

- Prof. Martin Hale (Department of Anatomical Pathology, University of the Witwatersrand).
- Dr. Pulane Mosiane (Department of Anatomical Pathology, University of the Witwatersrand).
- Prof. Francois Venter (HIV clinic at Charlotte Maxeke Johannesburg Academic Hospital).
- Patients (HIV clinic at Charlotte Maxeke Johannesburg Academic Hospital).
- Prof. Michèle Ramsay (Division of Human Genetics, National Health Laboratory Service and School of Pathology, Faculty of Health Sciences and The Sydney Brenner Institute for Molecular Bioscience, University of the Witwatersrand, Johannesburg, South Africa).

**Technical Laboratory assistance:**

- Dr. Caroline Dickens (Department of Internal Medicine, University of the Witwatersrand).
- Miss Therese Dix-Peek (Department of Internal Medicine, University of the Witwatersrand).
- Dr. Nadia Carstens (Division of Human Genetics, National Health Laboratory Service and School of Pathology, Faculty of Health Sciences and The Sydney Brenner Institute for Molecular Bioscience, University of the Witwatersrand).
- Miss Kiashanee Moodley (Department of Internal Medicine, University of the Witwatersrand).
- Mr. Victor David (Molecular Genetics Epidemiology Section. Frederick National Laboratory for Cancer Research, National Cancer Institute, National Institutes of Health, SAIC, NCI-Frederick, Maryland, USA).
- Mrs. Elizabeth Binns-Roemer (Molecular Genetics Epidemiology Section. Frederick National Laboratory for Cancer Research, National Cancer Institute, National Institutes of Health, SAIC, NCI-Frederick, Maryland, USA).
- Wits Non-Communicable Diseases Research Leadership group colleagues for advice and encouragement.

**HIV clinic at Charlotte Maxeke Johannesburg Academic Hospital:**

- HIV positive control participants.
- Prof. Francois Venter for granting us access/permission to the clinic.

**Postgraduate training from July, 2011 to August, 2013:**

- National Institutes of Health-funded Wits Non-Communicable Diseases Research Leadership programme.

**Funding:**

- NIH Fogarty International Center Grant 1D43TW008330-01A1 (07/01/2010 – 06/30/2015) Millennium Promise Award, NIDDK Intramural Research Programs, Non-communicable Chronic Diseases Leadership Training Program.
- Medical Research Council of South Africa, the National Research Foundation of South Africa.
- National Health Laboratory Service Research Trust.
- Division of Nephrology, University of the Witwatersrand.
- Faculty Research Committee Individual Grant, University of the Witwatersrand.
- Columbia University-South Africa Training Program for Research on AIDS-related Malignancies, National Cancer Institute, NIH (grant # 1D43CA153715-03).
- Funded in part with federal funds from the National Cancer Institute, National Institutes of Health, under contract HHSN26120080001E.

**Collaboration:**

- Dr. Cheryl Winkler for her technical advice on the genetics of the kidney disease and for her invitation to her Lab at Molecular Genetics Epidemiology Section. Frederick National Laboratory for Cancer Research, National Cancer Institute, National Institutes of Health, SAIC, NCI-Frederick, Maryland, USA.

- Dr. George Nelson (Molecular Genetics Epidemiology Section. Frederick National Laboratory for Cancer Research, National Cancer Institute, National Institutes of Health, SAIC, NCI-Frederick, Maryland, USA), for his technical advice on how to handle data from a genetics study and his guidance on how to analyze and validate genetic data.

- Dr. Agnes Fogo (Department of Pathology, Microbiology and Immunology, Division of Renal Pathology/Electron Microscopy, University of Vanderbilt Medical Center, Nashville Tennessee, USA), for reviewing histology diagnoses.

# RESEARCH OUTCOMES

## PUBLICATIONS

1. **Alex N. Kasembeli*,** Raquel Duarte, Michèle Ramsay, Pulane Mosiane, Caroline Dickens, Thérèse Dix-Peek, Sophie Limou, Efe Sezgin, George W. Nelson, Agnes B. Fogo, Stewart Goetsch, Jeffery B. Kopp, Cheryl A. Winkler, and Saraladevi Naicker **"*APOL1* risk variants strongly associated with HIV-associated Nephropathy in Black South Africans".** *(Published in the Journal of American Society of Nephrology. March, 2015).* (Appendix 1). To qualify as the first author of this paper, Alex N. Kasembeli finalised the protocol of the study, recruited patients, performed the assays, collected and analyzed data, and produced the initial draft of the paper.

2. **Alex N. Kasembeli*,** Raquel Duarte, Michèle Ramsay and Saraladevi Naicker **"African origins and Chronic Kidney Disease susceptibility in the HIV era"** *(Accepted for publication in the World Journal of Nephrology. January, 2015. In Press).* (Appendix 2). To qualify as the first author of this paper, Alex N. Kasembeli was responsible for producing the first complete draft of the review.

## CONFERENCES

1. **Alex N. Kasembeli[1*],** Raquel Duarte, Michèle Ramsay, Pulane Mosiane, Caroline Dickens, Thérèse Dix-Peek, Sophie Limou, Efe Sezgin, George W. Nelson, Agnes B. Fogo, Stewart Goetsch, Jeffery B. Kopp, Cheryl A. Winkler, and Saraladevi Naicker (2014). **"*MYH9* and *APOL1* association with HIV-associated nephropathy (HIVAN) and chronic kidney disease (CKD) in black South African population". Oral and Poster presentation** at the World Congress of Nephrology held on 13th to 18th March 2015 at Cape Town Convention Center, Cape Town, South Africa. **Accepted.**
2. **Alex N. Kasembeli[1*],** Raquel Duarte, Michèle Ramsay, Pulane Mosiane, Caroline Dickens, Thérèse Dix-Peek, Sophie Limou, Efe Sezgin, George W. Nelson, Agnes B. Fogo, Stewart Goetsch, Jeffery B. Kopp, Cheryl A. Winkler, and Saraladevi Naicker (2014). **"*APOL1* risk variants are strongly associated with HIV-Associated Nephropathy in Black South Africans". Poster presentation** at the American Society of Nephrology, Kidney Week Conference held on 11th to 16th November 2014 at Pennsylvania Conventional Center in Philadelphia, Pennsylvania, USA. **Accepted.**
3. **Alex N. Kasembeli[1*],** Raquel Duarte, Michèle Ramsay, Pulane Mosiane, Caroline Dickens, Thérèse Dix-Peek, Sophie Limou, Efe Sezgin, George W. Nelson, Agnes B. Fogo, Stewart Goetsch, Jeffery B. Kopp, Cheryl A. Winkler, and Saraladevi Naicker (2014). **"*MYH9* and *APOL1* association with HIV-associated nephropathy (HIVAN) and chronic kidney disease (CKD) in black South African population". Poster presentation** at University of the Witwatersrand, Faculty of Health Sciences Research day held on 16th to 17th September 2014. **Accepted.**

4. **Alex Kasembeli**, Raquel Duarte, Michèle Ramsay, Pulane Mosiane, Saraladevi Naicker. (2012). **"HIVAN Prevalence in Black South Africans and SNP selection within *MYH9 and APOL1* for an association study: A Preliminary Analysis". Poster presentation** at the South African Renal Congress held on 30th August to 2nd September 2012 at the CSIR convention centre in Pretoria, South Africa. **Accepted.**

# TABLE OF CONTENTS

**DECLARATION** ..... ii

**ABSTRACT** ..... iii

**DEDICATION** ..... vi

**ACKNOWLEDGEMENTS** ..... vii

**RESEARCH OUTCOMES** ..... xi

PUBLICATIONS ..... xi

CONFERENCES ..... xii

**TABLE OF CONTENTS** ..... xiv

**LIST OF FIGURES** ..... xix

**LIST OF TABLES** ..... xxi

**LIST OF ABBREVIATIONS** ..... xxiii

Chapter 1 ..... 1

**1) INTRODUCTION AND LITERATURE REVIEW** ..... 1

1.1 Chronic kidney disease ..... 1

1.2 Human Immunodeficiency Virus (HIV) ..... 2

1.2.1 Brief Overview of HIV infection ..... 4

1.2.2 HIV Epidemiology ..... 6

1.2.3 South African HIV Epidemiology ..... 7

1.3 HIV and kidney disease ..... 8

1.4 Genetic variation and disease ..... 10

1.5 Genetic susceptibility, kidney disease and HIV infection ..... 13

1.6 Overview of Chromosome 22 ..... 14

1.7 MYH9 and APOL1 Loci ..... 15

1.8 MYH9 Gene ..... 17

1.8.1 Mechanism of Myosin IIA action on podocytes ..... 17

1.8.2 MYH9 related disorders ..... 20
1.8.3 MYH9 associations with kidney disease ..... 20
1.8.4 MYH9 E1 haplotype ..... 21
1.9 APOL1 Gene ..... 23
1.9.1 APOL1 gene expression ..... 24
1.9.2 Human African Trypanosomiasis ("sleeping sickness") ..... 27
1.9.3 Trypanolytic activity of Apolipoprotein L1 ..... 28
1.9.4 Trypanolytic activity of Apolipoprotein L1 in relation to kidney disease ..... 32
1.9.5 APOL1 association with Kidney Disease ..... 34
1.10 Relationship between MYH9 and APOL1 risk variants ..... 37
1.11 Rationale for the study ..... 38
1.12 Hypothesis ..... 39
1.13 Objectives of the study ..... 40
Chapter 2 ..... 41
**2) MATERIALS AND METHODS** ..... 41
2.1 Study population ..... 41
2.2 Sample Size ..... 42
2.3 Participants ..... 43
2.3.1 Inclusion criteria ..... 45
2.3.2 Exclusion criteria ..... 45
2.3.3 Variables ..... 45
2.4 SNP Selection ..... 46
2.4.1 SNP scoring for compatibility with BeadXpress assay ..... 49
2.5 Laboratory procedures ..... 49
2.5.1 DNA extraction from Formalin Fixed Paraffin Embedded (FFPE) kidney tissues ..... 49
2.5.1.1 Purification of genomic DNA ..... 50
2.5.1.2 Purification of total RNA ..... 51

2.5.2 DNA Extraction from Human Nucleated Cells in blood samples ..... 53
2.5.3 Custom TaqMan SNP Genotyping Assay ..... 54
2.5.3.1 Materials ..... 56
2.5.3.1.1 Genomic DNA ..... 56
2.5.3.1.2 SNP genotyping assays ..... 56
2.5.3.1.3 Universal PCR Master Mix ..... 56
2.5.3.2 Procedure ..... 57
2.5.3.2.1 Master Mix preparation ..... 57
2.5.3.2.2 Reaction plate preparation ..... 57
2.5.3.2.3 Allelic discrimination plate read ..... 57
2.5.4 TaqMan Quality Control ..... 58
2.5.5 Custom Illumina® GoldenGate BeadXpress SNP Genotyping Assay ..... 58
2.5.5.1 BeadXpress Quality control ..... 61
2.5.6 Estimated Glomerular filtration rate (eGFR) ..... 63
2.6 Population stratification ..... 63
2.7 Messenger RNA (mRNA) expression profile of the N-terminal apolipoprotein L1 isoforms 64
2.7.1 Primer design for APOL1 gene N-terminal domain ..... 66
2.7.2 cDNA synthesis ..... 68
2.7.3 Polymerase Chain Reaction ..... 69
2.7.4 ImageJ ..... 70
2.8 Developing an expression system for Apolipoprotein L1 isoform expression in kidney cells 70
2.8.1 Cloning of ApoL1 N-terminal isoforms into pJET 1.2/blunt vector ..... 71
2.8.1.1 Transformation into already prepared E. coli DH5α competent cells ..... 72
2.8.1.2 Screening for positive clones ..... 73
2.8.1.3 Plasmid DNA purification ..... 74
2.8.2 Sub-cloning apolipoprotein L1 isoforms into the pDsRed2-N1 vector ..... 75
2.8.2.1 Digestion of DNA insert/isoform from pJET 1.2/blunt vector ..... 75
2.8.2.2 Fast simultaneous pDsRed2-N1 linearization and dephosphorylation ..... 76

2.8.2.3 Ligation of isoform into linearized and dephosphorylated pDsRed2-N1 vector ..... 77
2.8.2.4 Transformation into already prepared E. coli DH5α competent cells ..... 77
2.8.2.5 Screening for positive clones ..... 78
2.8.2.6 Plasmid DNA purification ..... 79
2.9 Culturing of Human Embryonic Kidney (HEK) 293 cells ..... 79
2.9.1 Determining the Cell number and viability ..... 80
2.9.2 Transfection of HEK 293 cells with APOL1 expressing constructs ..... 80
2.9.3 Immunofluorescence staining of transfected HEK 293 cells and confocal microscopy ..... 81
2.9.3.1 Cell preparation ..... 81
2.9.3.2 Cell fixation ..... 81
2.10 Statistical Analysis ..... 82
Chapter 3 ..... 84
**3) RESULTS** ..... 84
3.1 Clinical Characteristics ..... 84
3.2 SNP Selection ..... 86
3.3 SNP Genotyping Quality Control ..... 86
3.4 Frequencies and associations of MYH9 and APOL1 SNPs and APOL3 literature SNP ..... 89
3.4.1 Frequencies of MYH9 SNPs ..... 89
3.4.2 Associations of MYH9 SNPs ..... 89
3.4.3 Frequencies of the APOL1 and APOL3 SNPs ..... 93
3.4.3.1 Haplotype frequencies of the G1 and G2 risk alleles ..... 95
3.4.4 Associations of the APOL1 and APOL3 SNPs ..... 97
3.5 Combined haplotype analysis of the MYH9 and APOL1 SNPs ..... 104
3.6 Population stratification ..... 108
3.7 mRNA expression profiles of N-terminal-apolipoprotein L1 isoforms ..... 109
3.7.1 N-terminal APOL1 mRNA isoform expression patterns ..... 109
3.8 Expression of the N-terminal- apolipoprotein L1 isoforms in kidney cells ..... 112
3.8.1 Screening of clones expressing N-terminal-ApoL1 isoforms in pJET 1.2/blunt vector 112

3.8.2 Screening of clones expressing N-terminal-ApoL1 isoforms in pDsRed2-N1 vector.. 116
3.8.3 Transfected HEK 293 cells ..... 118
Chapter 4 ..... 119
**4) DISCUSSION** ..... 119
4.1 MYH9 association ..... 120
4.2 APOL1 association ..... 121
4.3 Suggestion for future study ..... 129
4.4 Proposed pathogenesis of APOL1-mediated nephropathy ..... 131
**LIMITATIONS OF THE STUDY** ..... 133
**CONCLUSIONS** ..... 134
**REFERENCES** ..... 136
**APPENDICES** ..... 150
APPENDIX 1 ..... 150
APPENDIX 2 ..... 159
APPENDIX 3 ..... 193
APPENDIX 4 ..... 195
APPENDIX 5 ..... 196
APPENDIX 6 ..... 199
APPENDIX 7 ..... 206
APPENDIX 8 ..... 209
APPENDIX 9 ..... 211
APPENDIX 10 ..... 213
APPENDIX 11 ..... 215
APPENDIX 12 ..... 217
APPENDIX 13 ..... 218
APPENDIX 14 ..... 220
APPENDIX 15 ..... 229

# LIST OF FIGURES

Figure 1.1: The Structure of HIV. Source: http://www.niaid.nih.gov/ ..... 3

Figure 1.2: Schematic representation of HIV Life Cycle. ..... 5

Figure 1.3: Map of the world showing prevalence rate of HIV in adults and children in 2012. ..... 7

Figure 1.4: Map showing estimated HIV prevalence among South Africans, 2008 ..... 8

Figure 1.5: Effect of recombination on linkage disequilibrium and haplotype blocks. ..... 12

Figure 1.6: Schematic representation of Chromosome 22 ..... 15

Figure 1.7: Region of high association with non-diabetic and non-hypertensive ESRD. ..... 16

Figure 1.8: The position of *MYH9* and *APOL1* on Chromosome 22 *(Not drawn to scale).* ..... 17

Figure 1.9: Model for actin-myosin functioning cycle on podocytes. ..... 19

Figure 1.10: Worldwide distribution of the *MYH9* haplotypes frequencies. ..... 22

Figure 1.11: *APOL1* gene structure. ..... 23

Figure 1.12: Expression of *APOL1* showing post-transcription modification and translation to its precursors. ..... 25

Figure 1.13: Amino acid sequence showing apolipoprotein L1 isoforms ..... 26

Figure 1.14: A. Distribution of Human African Trypanosomiasis (*T.b.rhodesiense*) and B. *T.b.gambiense* (W.H.O. 2010). ..... 28

Figure 1.15: Apolipoprotein L1-mediated Trypanolytic action. ..... 30

Figure 1.16: Association between *APOL1* risk variants, trypanolytic ability of apolipoprotein L1 and increased risk for kidney disease. SRA; serum resistant associated. LD; Linkage disequilibrium. Adapted from (Kronenberg, 2011). ..... 33

Figure 1.17: Worldwide frequency distribution of *APOL1* variants G1 and G2. ..... 36

Figure 1.18: The position of *MYH9* and *APOL1* risk variants on chromosome 22 showing the relationship between *APOL1* and *MYH9* genetic variants associated with common forms of kidney disease. ..... 38

Figure 2.1: Light microscopy showing histological features of HIVAN, FSGS and HIVICK. ..... 44

Figure 2.2: Summary of SNP selection process for Illumina assay compatibility ..... 48

Figure 2.3: Schematic illustration of TaqMan Allelic Discrimination Assay. ..... 55

Figure 2.4: Golden Gate BeadXpress overview and workflow assay. ..... 59

Figure 2.5: N-terminal sequence of *APOL1* isoforms A, B and C. ..... 67

Figure 3.1: Schematic diagram showing selected SNPs..... 87

Figure 3.2: Gene structure and linkage disequilibrium plot of *MYH9* and *APOL1* SNPs. ..... 88

Figure 3.3: Distribution of *APOL1* haplotypes and risk alleles. ..... 96

Figure 3.4: SNP association using logistic regression analysis..... 99

Figure 3.5: Odds ratios for the effect sizes for HIVAN..... 104

Figure 3.6: Gene structure and haploview linkage disequilibrium plot. ..... 106

Figure 3.7: Principal component analysis of all study participants along the top two eigen values. ..... 108

Figure 3.8: mRNA expression of the N-terminal-apolipoprotein L1 in kidney disease phenotypes..... 111

Figure 3.9: Colonies grown on LB-Ampincilin agar plate. Positive colonies are circled in red. ..... 112

Figure 3.10: Positive clones from pJet1.2/blunt vector of the N-terminal-apolipoprotein L1 isoforms A, B and C..... 113

Figure 3.11: Sanger sequencing to verify successful cloning of isoforms A, B and C..... 115

Figure 3.12: Colonies grown on LB-Kanamycin agar plate. Positive colonies are circled in red..... 116

Figure 3.13: PCR amplification of positive clones in *pDsred2-N1*containing the N-terminal-apolipoprotein L1 isoforms A, B and C..... 117

Figure 3.14: N-terminal apolipoprotein L1 (*ApoL1*) expression in HEK 293 cells. ..... 118

Figure 4.1: Timeline showing the discovery of genetic association to CKD in African ancestry populations..... 130

Figure 4.2: Proposed pathogenesis of *APOL1*-mediated nephropathy. ..... 132

# LIST OF TABLES

Table 1.1: Glomerular disease spectrum in HIV (Fabian *et al.,* 2007). ..... 9

Table 1.2: Summary of Worldwide distribution of E1, E2, E3, E4 and E5 haplotypes (Oleksyk *et al.*, 2010) ..... 21

Table 1.3: Allele frequencies of SNPs that delineate *MYH9* E1 and *APOL1* G1/G2 risk haplotypes. ..... 37

Table 2.1: Case study participants ..... 43

Table 2.2: TaqMan® master mix components ..... 57

Table 2.3: Thermocycling conditions for the TaqMan® assay ..... 57

Table 2.4: Thermocycling conditions for the Illumina® GoldenGate BeadXpress assay ..... 60

Table 2.5: Histology diagnoses, HIV status and number of risk alleles of samples selected for mRNA N-*ApoL1* isoform profiling ..... 65

Table 2.6: Primer sequences for N-terminal *APOL1* gene ..... 66

Table 2.7: RT-PCR master mix ..... 68

Table 2.8: Thermocycler conditions for RT-PCR reaction synthesis of cDNA ..... 68

Table 2.9: PCR master mix components ..... 69

Table 2.10: Thermocycler conditions ..... 69

Table 2.11: Blunting reaction ..... 72

Table 2.12: Ligation reaction ..... 72

Table 2.13: Colony PCR reaction ..... 73

Table 2.14: Thermocycler conditions ..... 74

Table 2.15: BglII restriction digest recombinant pJET1.2 vector ..... 76

Table 2.16: Fast simultaneous *pDsReD2-N1* linearization and dephosphorylation reaction ..... 77

Table 2.17: Ligation reaction ..... 77

Table 2.18: Colony PCR for Isoforms A, B and C reaction ..... 78

Table 2.19: Thermocycler conditions ..... 78

Table 3.1: Demographic characteristics of the study population ..... 85

Table 3.2: Logistic regression analysis showing *MYH9* E1 haplotype individual SNP association with HIVAN, HIV (+) CKD and HIV (-) CKD cases ..... 90

Table 3.3: Logistic regression analysis showing Individual SNP association of rs16996648, rs5756152, rs16996668, rs16996672, rs16996674, rs16996677 with HIVAN, HIV (+) CKD and HIV (-) CKD cases compared to their respective controls. ..... 91

Table 3.4: Logistic regression analysis showing *MYH9* SNP associations with 38 HIVAN cases and 54 HIV (+) Controls ..... 92

Table 3.5: Logistic regression analysis showing association of *APOL1* and *APOL3* SNPs with HIVAN, HIV (+) CKD and HIV (-) CKD ..... 93

Table 3.6: *APOL1* G1 and G2 genotypes in cases and controls. ..... 95

Table 3.7: Logistic regression analysis showing significant SNP Associations with 38 HIVAN cases and 54 HIV (+) controls ..... 98

Table 3.8: *APOL1* risk allele associations with HIVAN, HIV-associated FSGS, and HIV (-) FSGS (primary FSGS) ..... 101

Table 3.9: Association between *APOL1* risk alleles and various glomerular diseases ..... 103

Table 3.10: Combined haplotype association with 38 HIVAN and 54 HIV (+) Controls. ..... 107

# LIST OF ABBREVIATIONS

ABR mean AB mean for intensity R

ABT mean AB mean for theta

AD Allelic discrimination

AIDS Acquired Immune deficiency syndrome

AIMs Ancestry information markers

AP-1 Activator protein-1

*APOL1* Apolipoprotein L-1 gene

*ApoL1* Apolipoprotein L1 protein

APS Ammonium persulfate

ART Antiretroviral therapy

ASO Allele specific oligonucleotide

bp base pair

CCR5 C-C Chemokine co-receptor type 5

CD4 Cluster of differentiation 4

CD8 Cluster of differentiation 8

cDNA Complementary Deoxyribonucleic acid

CKD Chronic kidney disease

CMJAH Charlotte Maxeke Johannesburg Academic Hospital

CNS Central nervous system

CNVs Copy number variations

CRP C-reactive protein

CXCR4 C-X-C Chemokine co-receptor type 4

DNA Deoxyribonucleic acid

dNTPs  deoxynucleotide triphosphates

drFP583 *Discosoma sp.* red fluorescent protein 583

*E. coli Escherichia coli*

EDTA Ethylenediaminetetraacetic acid

eGFR estimated glomerular filtration rate

ESRD End stage renal disease

FBS Fetal Bovine Serum

FFPE Formalin Fixed paraffin embedded

FRET Forster fluorescence resonance energy transfer

FSGS Focal segmental glomerulosclerosis

GC Gen Call

GFR glomerular filtrate

GWAS Genome wide association study

HAART Highly active anti-retroviral therapy

HAT Human African Trypanosomiasis

HDL3 High density lipoprotein class 3

HEK Human embryonic kidney

HGP Human Genome Project

HIV Human Immunodeficiency virus

HIVAN HIV associated nephropathy

HIVICK HIV immune complex kidney disease

HWE Hardy Weinburg equilibrium

IDT Integrated DNA Technologies

IN Integrase

INDELs insertion and deletions

Iso A Isoform A

Iso B Isoform B

Iso C Isoform C

Kb Kilo base pair

KIM-1 kidney injury molecule 1

LB Lysogeny broth

LD Linkage disequilibrium

LDL Low density lipoprotein

LSO Locus specific oligonucleotide

LWK Luhya from Webuye, Kenya

MAF Minor Allele Frequency

MALD Mapping by admixture linkage disequilibrium

MCS Multiple cloning site

MGB Minor groove binding

MKK Maasai from Kinyawa, Kenya

MTCT Mother-to-child transmission

*MYH9* Non-muscle myosin heavy chain 9 gene

Myosin IIA Non-muscle myosin heavy chain IIA protein

NaCl Sodium chloride

NCI National Cancer Institute

NFAT nuclear factor of activated T-cells

NF-kβ nuclear factor kappa-light-chain-enhancer of activated B cells

NGAL neutrophil gelatinase-associated lipocalin

NHLS National Health Laboratory Service

NIH National Institutes of Health

NKUDIC National kidney and urologic diseases information cleaninghouse

NNRTIs Non-nucleoside reverse transcriptase inhibitors

NQF Non fluorescent quencher

NRTIs Nucleoside reverse transcriptase inhibitors

OPA Pool of primers

ORFs Open reading frames

PBS Phosphate buffered saline

PC Population control

PCR Polymerase chain reaction

Pen-strep Penicillin and streptomycin.

POLYPHEN Polymorphism phenotyping

QxMD Mayo clinic quadratic equation

RNA Ribonucleic acid

Rpm revolutions per minute

RT reverse transcriptase

RT-PCR Reverse transcription-polymerase chain reaction

SDS Sodium dodecyl sulphate

SIFT Sorting Intolerant from tolerant

SNPs Single nucleotide polymorphisms

SOC Super Optimum Culture

SP-1 specificity protein 1

SSA Sub Saharan Africa

STI Sexually transmitted infections

Tat Trans-activator of transcription

TBE Tris Borate EDTA

TE Tris-EDTA

TEMED N, N, N', N'-tetramethylethylenediamine

TMA Thrombotic microangiopathy

UMM Universal master mix

VNTRs Variable number tandem repeats

VSG Variant surface glycoproteins

WHO World Health Organization

WTCCC Wellcome trust case control consortium

YRI Yoruba from Ibadan

# Chapter 1

# 1) INTRODUCTION AND LITERATURE REVIEW

## 1.1 Chronic kidney disease

Chronic kidney disease (CKD) is an important worldwide public health problem. According to the End Stage Renal Disease Patients in 2013: A Global Perspective report, an estimated 3.2 million people were on renal replacement therapy by the end of 2013, approximately 2,522,000 people undergoing dialysis treatment (haemodialysis or peritoneal dialysis) and 678,000 people living with renal transplants; it is also estimated that CKD incidence grows by approximately 6% annually (*ESRD Patients in 2013: A Global Perspective*). Chronic kidney disease encompasses a spectrum of different pathophysiologic processes associated with abnormal renal function and a progressive decline of the glomerular filtration rate (GFR), resulting in an irreversible loss of renal function to stage V of CKD or end stage renal disease (ESRD) (<15 ml/min per 1.73$m^2$) and premature death (Kher, 2002). There are striking regional and ethnic differences in the incidence and prevalence of ESRD, such that, in the US, African Americans have the highest incidence and prevalence of ESRD (Rostand *et al.,* 1982). Diabetes and hypertension, which have been considered the two leading aetiologies of CKD together with differences in conventional clinical, social-demographic or lifestyle factors, are insufficient to account satisfactorily for the excess risk of ESRD in African Americans (Klag *et al.,* 1997; Shlush *et al.*, 2010). There is a four-fold increase in CKD among individuals of African descent and an 18-50 fold increase in kidney disease in human immunodeficiency virus (HIV) positive African individuals (Genovese *et al.*, 2010; Kopp *et al.*, 2011). Therefore, HIV plays a vital role in the development of non-diabetic and non-hypertensive kidney disease by acting as a major environmental factor in causing renal dysfunction. Over the years, Africa has also been

disproportionally burdened by CKD. The increased burden of CKD in Africa could be as a result of various communicable diseases such as leishmaniasis, schistosomiasis, infectious glomerulonephritis and importantly, human immunodeficiency virus (HIV) infection superimposed on non-communicable diseases such as hypertension and diabetes mellitus *(Idem, 2005*). These infections, coupled with genetic factors, establish CKD as a major problem in Africa. Hence ESRD is probably the most multi-factorial disease affected by both genetic and environmental processes.

## 1.2 Human Immunodeficiency Virus (HIV)

HIV infection was first observed clinically in the United States of America in 1981 (Gallo & Montagnier, 2003). It is a member of the retrovirus family that causes acquired immunodeficiency syndrome (AIDS) (Figure 1.1). AIDS is a human condition that causes a slow destruction of the host's immune system in turn predisposing infected individuals to fatal opportunistic infections.

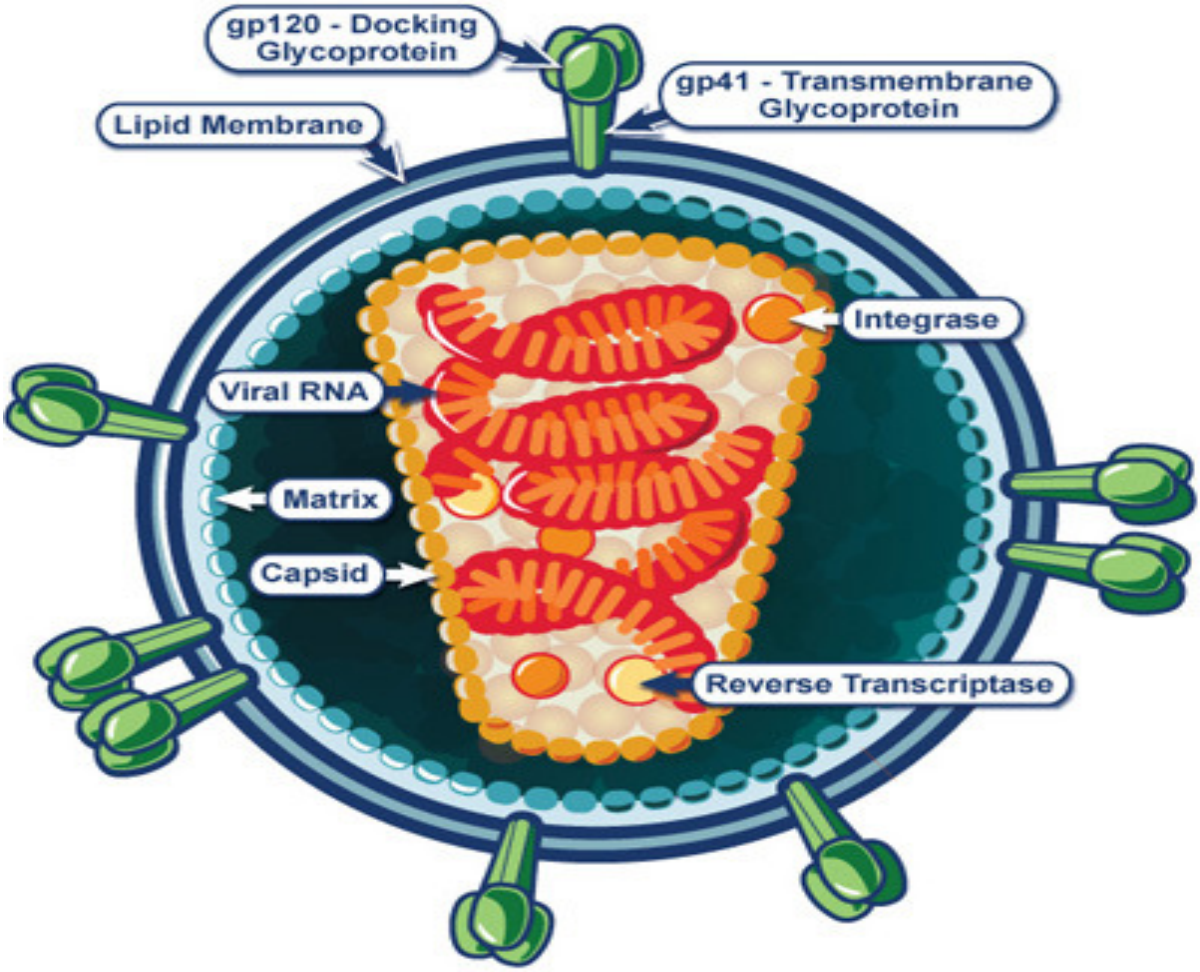



**Figure 1.1:** The Structure of HIV. Source: http://www.niaid.nih.gov/

It is spherical in shape with a diameter of 120nm which makes it to appear larger than a normal virus (Chan *et al*., 1997). It comprises of single-stranded RNA that encodes the viral protein p24. It also comprises of the enzymes; protease, reverse transcriptase, integrase and ribonuclease which are important for the development of the mature virions. The integrity of the virion is ensured by the fact that the capsid is surrounded by the viral protein p17 in the matrix. The entire membrane of the virion is surrounded by a layer of phospholipids taken from the host cell during budding off from the human cell. The virus also comprises of the env protein which is made up of glycoprotein (gp) 120, and gp 41 molecules that anchor its

structure (Chan *et al.*, 1997). This complex enables the virus to attach and fuse to target cells (CD4 cells) and initiate infection (Chan *et al.*, 1997).

### 1.2.1 Brief Overview of HIV infection

HIV infects CD4 cells, macrophages, and dendritic cells. The infection is facilitated by the presence of the virion envelope glycoproteins (gp120 and gp41) binding to the CD4 molecule receptor and chemokine co-receptors, CXCR4 and/or CCR5 on the target cells (Figure 1.2). HIV destroys the cells of the immune system mainly through direct viral killing of infected cells or by CD8 cytotoxic lymphocytes that recognize infected cells, killing them in the process. This leads to low levels of CD4 cells and subsequent high HIV viral load (copies/mL of plasma), host immunity is lost, followed by progressive exposure to opportunistic infections (Cunningham *et al.*, 2010; Douek *et al.*, 2009).

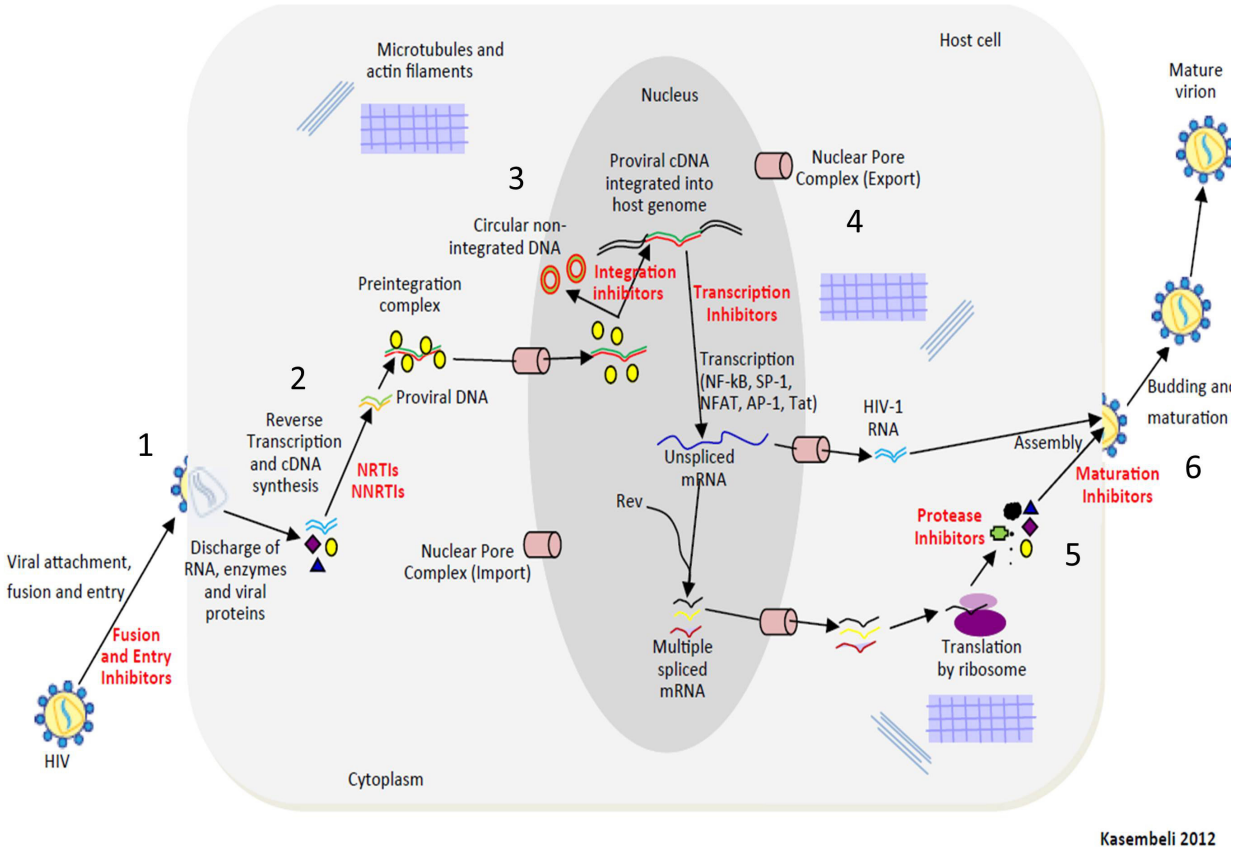



**Figure 1.2:** Schematic representation of HIV Life Cycle.
This figure shows the mechanism of HIV infection and targets for ARV therapy through steps 1-6. NRTIs: Nucleoside reverse transcriptase inhibitors. NNRTIs: Non-nucleoside reverse transcriptase inhibitors. Adapted from HIV binds to the host cell (CD4 and CD4-like molecules) via receptors and co-receptors (CCR5 and CXCR4) through **fusion (1)**. Fusion is followed by uncoating of the viral core, and deposition of components [viral RNA genome, reverse transcriptase (RT), integrase (IN), and virion regulatory proteins e.g NF-kB, SP-1, NFAT, AP-1 and Tat] into the host cell cytoplasm. Reverse transcriptase begins **assembling (2)** of DNA copies, complimentary DNA (cDNA) of the HIV genome. This is a reverse of the usual transcription in which DNA is the template for RNA. cDNA enters the host cell nucleus as a complex molecule (comprising cDNA, RT and IN). Once inside the nucleus, the cDNA integrates into the host-cell DNA forming a **proviral DNA (3)** which contains blueprint for creating virus progeny. Proviral DNA is transcribed into messenger RNA which is then transported to the cytoplasm **(4)** where it is translated into new viral components, enzymes and genomic elements. An essential viral enzyme that is synthesized during this process is known as protease (PR). PR influences viral components to associate with host cell membrane **(5)** and then **bud off** as immature virions. Protease also influences the molecular changes that ensure **maturation** into fully infectious **virions (6).** Source: Adapted from http://www.niaid.nih.gov/.

### 1.2.2 HIV Epidemiology

The joint United Nations programme on HIV/AIDS as at 2011 approximated that, "35.3 million people were living with HIV globally, including an estimated 3.4 million children of less than 15 years as at 2011 (Figure 1.3). About 2.3 million new infections including about 240,000 children less than 15 years were also reported in that year" (UNAIDS). However, there is a huge regional variation; "sub-Saharan Africa (SSA) has two thirds of all people living with HIV and nearly 75% of AIDS-related deaths and an estimated 1.9 million people were infected in 2011", UNAIDS 2011. The introduction of ART in 1995 averted 2.5 million deaths in low- and middle-income countries globally by reducing the AIDS-related mortality, mostly from SSA including more than 350,000 children and mainly women of the reproductive age," (Department of Health, 2010; Organization, 2011).

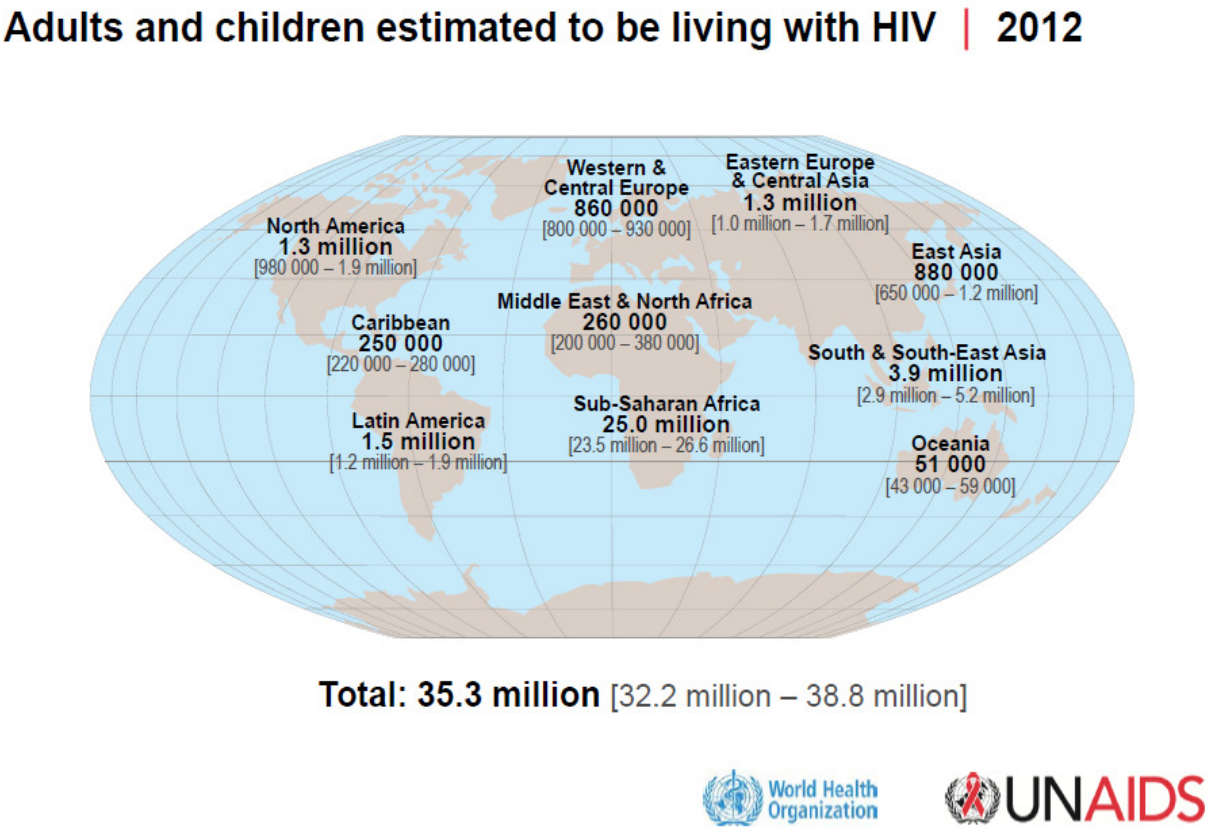



**Figure 1.3:** Map of the world showing prevalence rate of HIV in adults and children in 2012. Sub-Saharan Africa has the highest number of people (25.0 million) living with HIV. (Source: WHO/UNAIDS).

### 1.2.3 South African HIV Epidemiology

"South Africa is one of the countries most severely affected by the AIDS epidemic. It has more than 15% of the population aged 15-49 years living with HIV. UNAIDS 2011, estimated that there are 5.7 million people living with HIV in the country; more than 3 million of these are women aged 15years and above, and 280,000 are children aged 0-14 years making it the world's largest population of people living with HIV", (UNAIDS, 2011). In 2011, there were about 250,000 deaths related to AIDS. Although this number reflects the highest number of deaths that a country has lost over the past, it is approximately 100,000 fewer deaths than in 2001, demonstrating the effectiveness of the treatment scale-up (UNAIDS, 2011). The

prevalence of HIV was about 17.3% among the general population as at 2008; however, this varied a lot by region with the highest rate of 38.7% in KwaZulu Natal as shown by Figure 1.4.

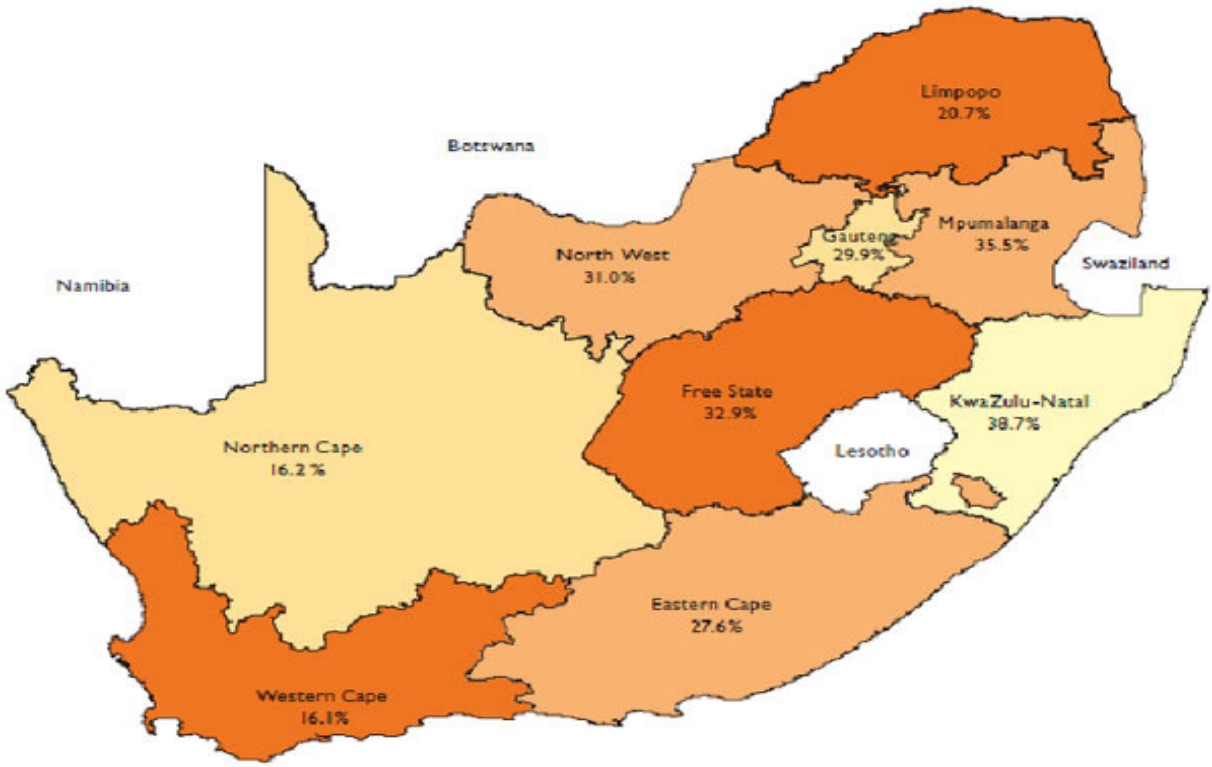



**Figure 1.4:** Map showing estimated HIV prevalence among South Africans, 2008
**Source:** The South African National HIV Survey 2008, Human Sciences Research Council (**www.hsrcpress.ac.za**). This is the current published data in South Africa.

## 1.3 HIV and kidney disease

AIDS-associated nephropathy was discovered in AIDS patients in New York and Miami in 1984. This name later changed to HIV-associated nephropathy (HIVAN) after asymptomatic HIV-infected individuals showed similar clinical and histological features (Rao *et al.,* 1984). Since then, kidney disease has been a well-known complication of HIV infection. HIV causes clinical manifestations with a broad spectrum, ranging from an asymptomatic state to severe immunodeficiency. Even though HIV prevalence and incidence rates have changed since the

roll-out of highly active antiretroviral therapy (HAART), kidney disease remains associated with considerable morbidity and mortality in HIV-infected individuals (Michelle M Estrella, 2009). Kidney disease occurs in approximately 10% of HIV-positive patients with a GFR < 60 mL/min. About 50-60% of these cases are considered classic HIVAN (Hays T, 2012; Kalayjian, 2010). This is a clinicopathological condition characterized by proteinuria and rapidly developing renal failure and histologically, by a collapsing variant of focal segmental glomerulosclerosis (FSGS) with microcystic tubular dilatation and mild interstitial inflammation (D'Agati & Appel, 1997; Han *et al.*, 2006; Rao *et al.*, 1984; Wearne *et al.*, 2012). Risk factors for HIVAN are older age, lower CD4 counts, high viral load, co-morbidity (such as diabetes mellitus, hypertension and hepatitis C co-infection) (Ray, 2012a) and is present predominantly in people of African ancestry. A number of other renal histology has been reported in HIV infected patients (D'Agati & Appel, 1997). These include HIV related FSGS, HIV immune complex kidney disease (HIVICK) and other histologies as shown in Table 1.1.

**Table 1.1**: Glomerular disease spectrum in HIV (Fabian *et al.,* 2007).

| **Glomerular pattern** | **Subtypes** |
|---|---|
| Classic HIVAN | Sometimes described as a mixed variant of HIV FSGS together with proliferative glomerulonephritis |
| HIVICK (mostly with hepatitis B or C co-infection) | Mesangial proliferative<br>Membranoproliferative (type I and III)<br>Lupus-like<br>Crescentic IgA<br>Membranous<br>Exudative-proliferative |
| HIV Thrombotic microangiopathy<br>Co-morbid disease | <br>Diabetic nephropathy<br>Hypertensive nephrosclerosis<br>Auto-immune disease such as lupus nephritis |
| Various glomerulonephropathies (heterogenous group with different aetiologies) | <br>Minimal change<br>Immunotactoid<br>Amyloidosis |

HIV-associated nephropathy (HIVAN), HIV immune complex kidney (HIVICK) disease and other HIV-associated histologies).

However, the mechanism by which HIV induces glomerular injury leading to the pathologic syndrome of HIVAN is not well understood, hence a number of theories have been postulated on how HIV causes renal injury (Gerntholtz *et al.*, 2006; Ray, 2012b). There have been several major mechanisms described on how HIV infects renal cells: First, a direct viral infection of podocytes, renal parenchyma cells, especially the visceral epithelial cells of the glomerulus, and the tubular epithelial cells and as a result, this elicits cytopathic effects including proliferation and apoptosis. Secondly, HIV activates the innate immune system (likely infecting the lymphocytes and macrophages) that enter the kidney, resulting in the release of inflammatory lymphokines or cytokines which promote injury and fibrosis (Gerntholtz *et al.*, 2006; Kopp & Winkler, 2003). However, there have been studies that demonstrated that two main HIV co-receptors, *CCR5* and *CXCR4,* are absent on intrinsic renal cells in HIV associated renal disease, but are expressed in circulating and infiltrating leukocytes at sites of tubulo-interstitial inflammation (Eitner *et al.*, 1998; Eitner *et al.,* 2000).

## 1.4 Genetic variation and disease

Genetic variation is the occurrence of the genetic difference both within and among individuals and populations (Conrad *et al.,* 2011). It is brought about by demographic and evolutionary factors including mutation, which is a permanent change in the chemical structure of the gene, and differences in allele frequency brought about by selection, population size changes, migration and admixture, as well as by chance referred to as genetic drift. This variation is important for the survival of a population in the face of dynamic environmental changes (Ramsay, 2012). The most abundant type of genetic variation found in the human genome is the single nucleotide polymorphism (SNP). To be described as polymorphic, the rare allele should occur at a frequency of at least 1% of the population. Some polymorphic variants have been

implicated in diseases, for instance the sickle cell variant in Central and West African populations (Ko & Urban, 2013). Many SNPs have been associated with different diseases through Genome Wide Associated Studies (GWAS) and the information stored in databases that are easily accessible (NCBI GWAS catalogue http://www.genome.gov/26525384). Genome Wide Associated Studies such as the Wellcome Trust Case Control Consortium (WTCCC), helped uncover the associations between individual SNPs and disorders with a heritable component and identified variants that could be used to determine an individual's risk of developing a particular disease (*GWAS, Wellcone-Trust, 2007*). Single nucleotide polymorphism allele frequencies differ among populations, and many studies have demonstrated that the greatest variation occurs among the African populations who have been investigated the least (Kottgen, 2010; Rosset *et al.*, 2011).

Single nucleotide polymorphisms can also be used as "sign posts" (in linkage disequilibrium) in mapping other variants that are functionally involved in disease phenotypes and by doing so, are considered as disease markers. Linkage disequilibrium (LD) is the non-random association of certain alleles at two or more loci that descend from single, ancestral chromosomes. Single nucleotide polymorphisms that are adjacent to each other and are highly correlated (in high LD) allow scientists to map locations of mutations that cause diseases (Reich *et al.*, 2001). Other genetic variants are insertions-deletions (INDELs), variable number tandem repeats (VNTRs) and copy number variants (CNVs), which have also been shown to play an important role in disease.

Apart from genetic variation, chromosome translocations also play a role in disease development (Rosset *et al.*, 2011). Genetic recombination also affects the level of association between different SNPs and the extent to which groups of the neighbouring SNPs are preserved in the

same combination on parental chromosome. Genetic recombination breaks the allelic relationship such that some descendants could harbor the wild type at one site and the mutant on the other as illustrated in Figure 1.5. The resulting combination of alleles in a given population and inherited together is then referred to as haplotypes (Rosset *et al.,* 2011). If the variants are close to each other, there is an increase in frequency of the resulting haplotypes. The LD and haplotype blocks capture the extent of association of specific SNPs in haplotypes.

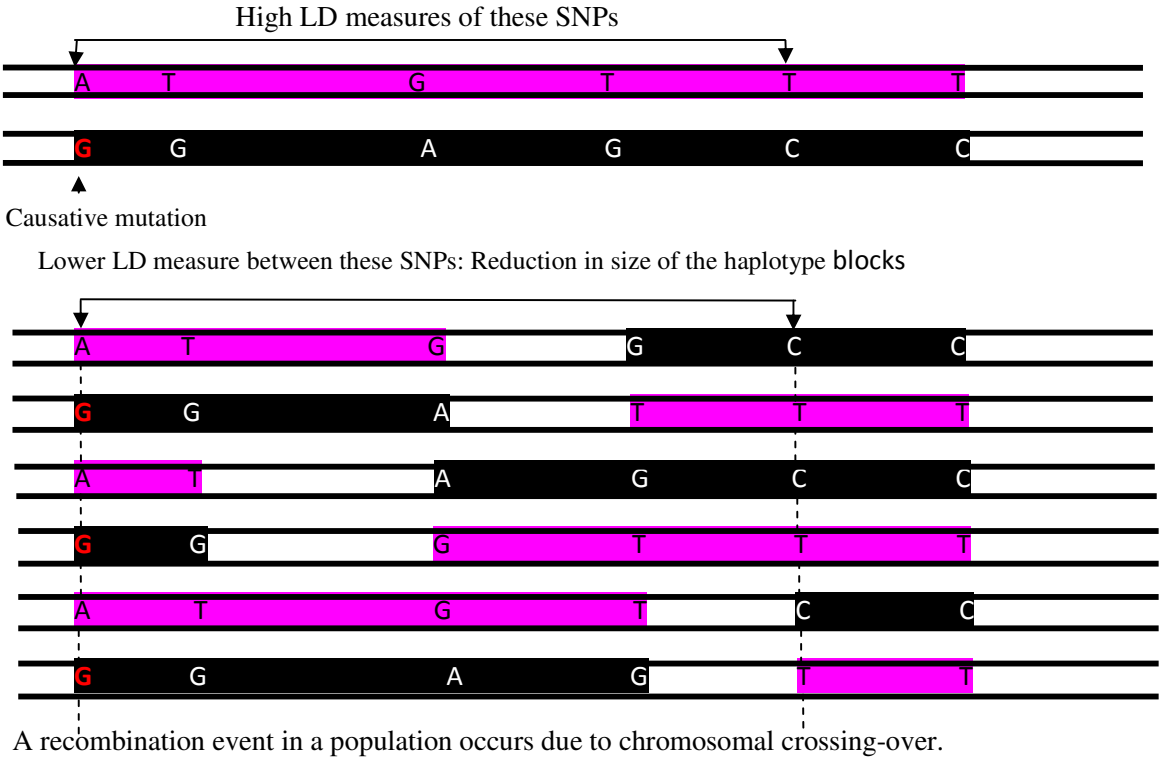



**Figure 1.5:** Effect of recombination on linkage disequilibrium and haplotype blocks.
When a mutation occurs in a population, it becomes highly associated (in high LD) with neighbouring SNPs which are inherited together as haplotype blocks. If in any case, such a mutation is a causative mutation or of high disease risk, then SNPs in the region can be used as markers of its presence. However, as the mutation spreads within a population, recombination occurs as it passes through the generations and as a result the LD breaks down reducing the size of the initial haplotype block, in the process creating new haplotypes where two initial haplotypes with six SNPs give rise to six different haplotypes in the population. Purple and black represents haplotype blocks from different chromosomes. Permission obtained from Nature Publishing Group © Rosset, S. *et al. Nat. Rev. Nephrol*. 7, 313–326 (2011). (Rosset *et al.,* 2011). Appendix 15 A

## 1.5 Genetic susceptibility, kidney disease and HIV infection

Studies of associations between host susceptibility genes and HIV have concentrated on the host's immune defense mechanism and HIV life cycle genes (Dean *et al.*, 1996; Huang *et al.,* 1996). Due to this, the key viral proteins that determine HIV pathogenesis or drug resistance have been identified (Kiryluk *et al.,* 2007). In addition, GWAS have successfully determined genetic variants that influence the outcome of HIV infection. Despite the fact that there is a high extent of treatment for diabetes and hypertension especially in the African set up, the incidence and prevalence rates of ESRD continue to increase (Satko *et al.,* 2005). HIVAN is an important cause of renal dysfunction in Africa and is a distinct pathological complication of HIV infection. Different ethnic distribution and inter-individual variability of HIVAN with other forms of ESRD have suggested that host genetic susceptibility is a major contributing factor to development of ESRD (Kiryluk *et al.*, 2007). In addition, according to the United States Renal Data System 2013 (USRDS 2013) report, there is an 18-50 fold increased risk of developing kidney disease among HIV-positive individuals of African descent compared to other non-African individuals. In general, these individuals have a poorer prognosis than individuals from the European population. HIVAN is rare among people of European and Asian descent, suggesting that genetic factors play a vital role in susceptibility (NIDDK, 2012). Furthermore, not all African populations seem to have the same predisposition to HIVAN. This is evidence for the genetic diversity of African populations and therefore demonstrate inter-population heterogeneity with regard to susceptibility to HIVAN (Nochy *et al.,* 1993). Behar et al., (2011), reported the absence of HIVAN and the *APOL1* risk variants among HIV-infected individuals of Ethiopian ancestry. This finding supports a Western rather than Pan-African ancestry risk for ESRD, and can readily explain the lack of HIVAN among individuals of

Ethiopian ancestry. This notion was further illustrated in three studies of South Africans where HIVAN was the commonest histopathological entity observed among HIV-positive individuals (Gerntholtz *et al.,* 2006; Han *et al.*, 2006; Wearne *et al.,* 2012). However, other HIV-positive South African blacks also presented with immune-complex glomerulopathy suggesting inter-individual variation in classic HIVAN susceptibility within the population (Gerntholtz *et al.*, 2006; Wearne *et al.,* 2012). Racial differences, inter-individual and human genetic variation strongly suggest the existence of HIVAN or kidney disease-specific genetic susceptibility factors that are triggered or unmasked in the presence of HIV infection hence postulated that genetic susceptibility to kidney disease has been attributed to genetic variations within non-muscle myosin heavy 9 chain (*MYH9)* and apolipoprotein L1 (*APOL1)* genes on chromosome 22 (Genovese *et al.,* 2010; Kopp *et al.,* 2011; Kopp *et al.*, 2008; Tzur *et al.*, 2010; Winkler *et al.,* 2010).

## 1.6 Overview of Chromosome 22

Chromosome 22 is the second smallest human chromosome comprising 51.3 million base pairs (Dunham *et al.,* 1999). Dunham *et al.* (1999), described chromosome 22 to have 507 protein coding regions, 390 non-coding gene regions, 307 pseudogene counts and 800,598 variations (Figure 1.6). It was the first human chromosome to be fully sequenced during the Human Genome Project (HGP) in December 1999. The short arm (22p) contains a number of tandem repeat structures including genes that encode the structural ribonucleic acids (RNAs) of the ribosomes while the long arm (22q) contains protein coding genes (Dunham *et al.*, 1999).

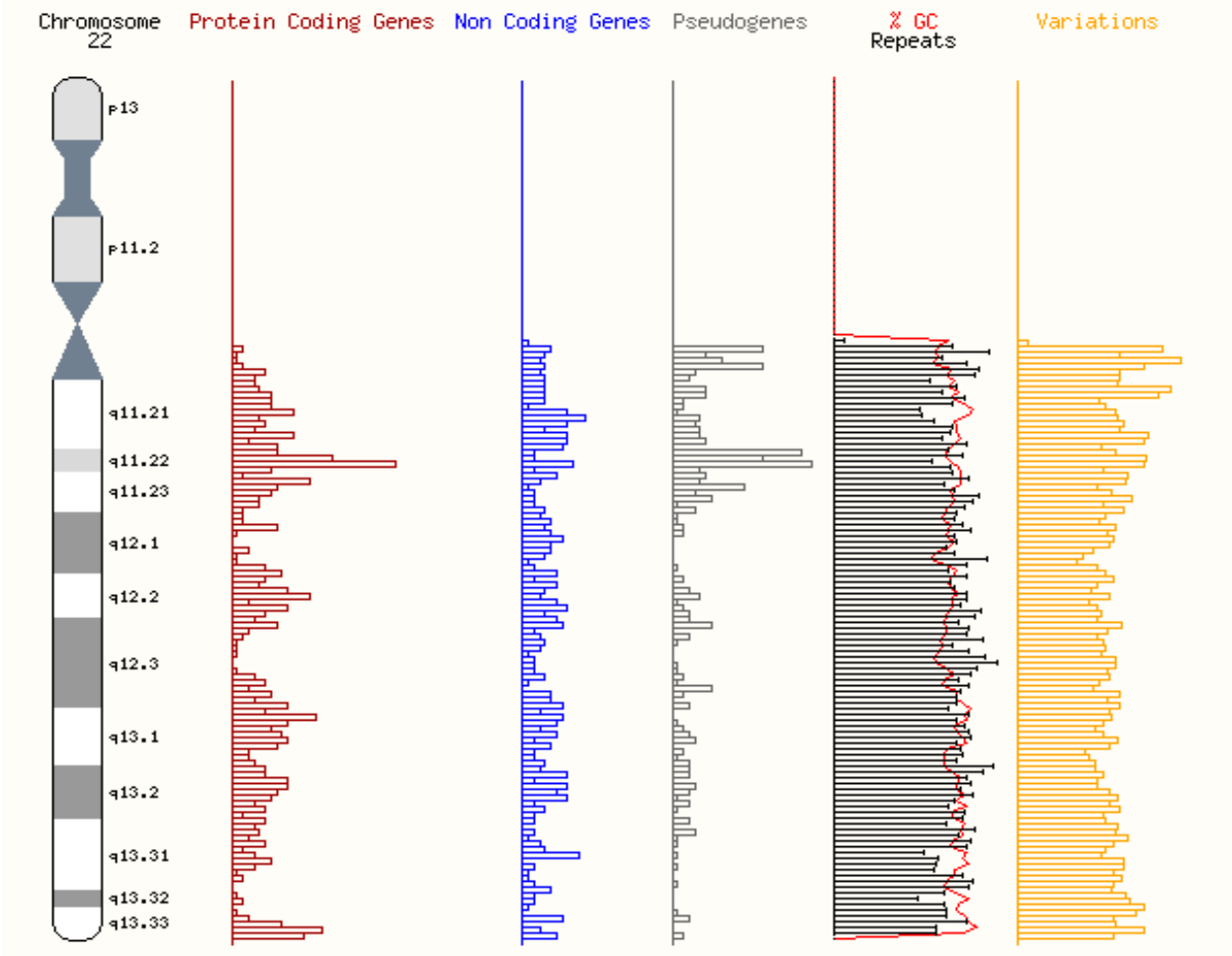



**Figure 1.6:** Schematic representation of Chromosome 22
It comprises of 507 protein coding regions, 390 non-coding gene regions, 307 pseudogene counts and 800,598 variations. Source; Ensembl (**http://www.ensembl.org/Homosapiens/Location/Chromosome?chr=22;r=22:1-51304566**) Accessed on December, 2013.

## 1.7 *MYH9* and *APOL1* Loci

The epidemiological and clinical observations of non-diabetic and non-hypertensive kidney disease among the African population prompted researchers to conduct population-based genetic studies that exploit LD in genome-wide screens (Kopp *et al.*, 2008; Rosset *et al.,* 2011; Shlush *et al.*, 2010). Mapping by admixture linkage disequilibrium (MALD) studies specifically for the ESRD susceptibility loci among those of African American decent led to the identification of association with a specific genomic region on chromosome 22q12 that contains more than 21

genes including *MYH9* and *APOL1* (Figure 1.7). MALD is basically the mapping of genetic variants involved in disease by taking advantage of long haplotypes blocks generated by gene flow in recently admixed populations (M. W. Smith & O'Brien, 2005).

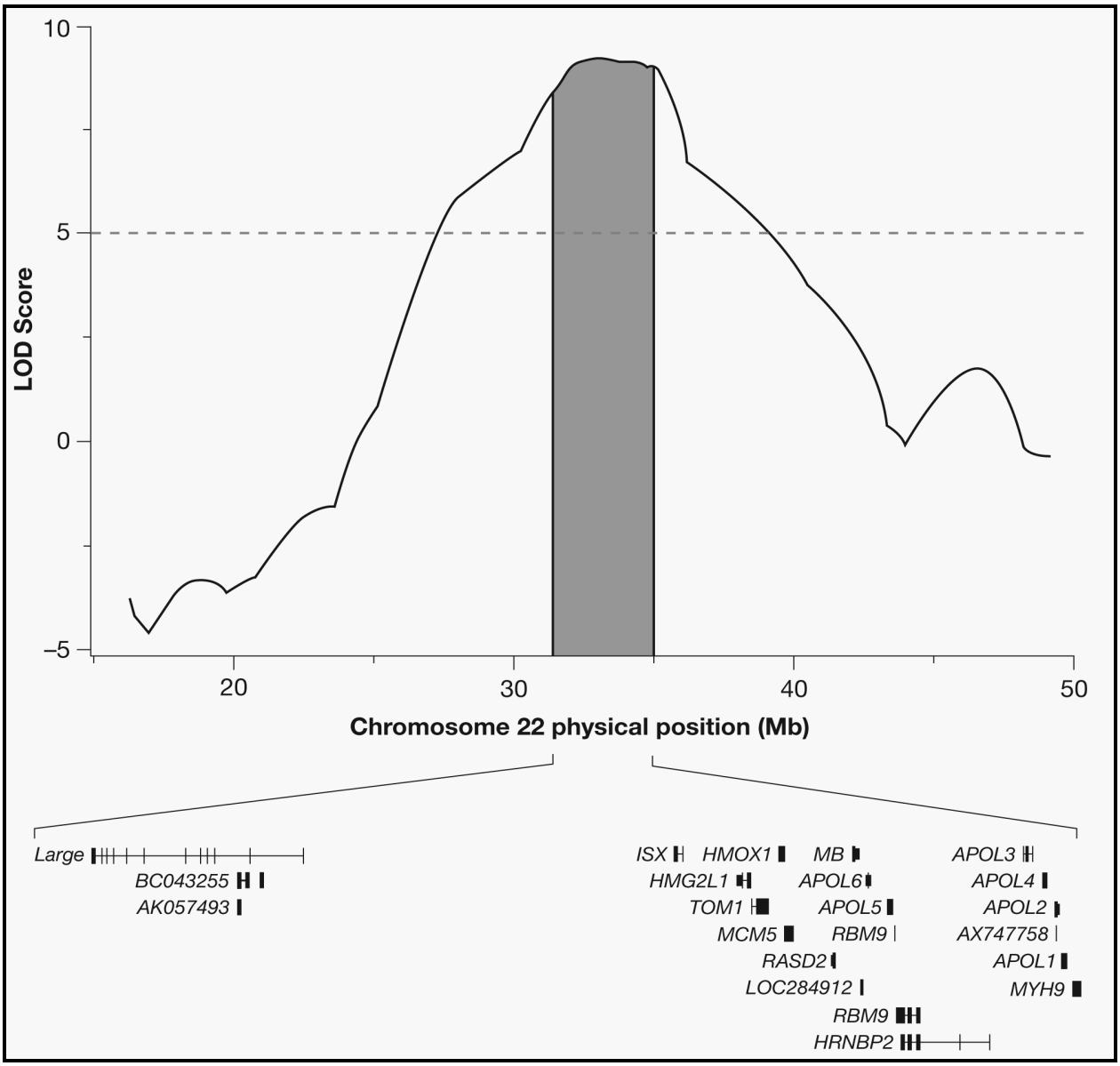



**Figure 1.7:** Region of high association with non-diabetic and hypertensive ESRD.
Logarithm of the odds (LOD). The peak LOD score was 9.31, and the 95% CI spanned from 31.388650 to 35.039798 Mb, covering 22 genes. Permission obtained from Nature Publishing Group © Rosset, S. *et al*. *Nat. Rev. Nephrol*. 7, 313–326 (2011) (Rosset *et al.*, 2011). Appendix 15 B

*MYH9* and *APOL1* are adjacent genes on chromosome 22 in a region that shows maps of recent selection in the certain African populations (Genovese *et al.,* 2010; Kopp *et al.,* 2011). *APOL1* is approximately 20 kilo bases (kb) upstream of *MYH9* (O'Seaghdha *et al.,* 2011). It is transcribed from the forward strand while *MYH9* is transcribed from the reverse strand (Figure 1.8).

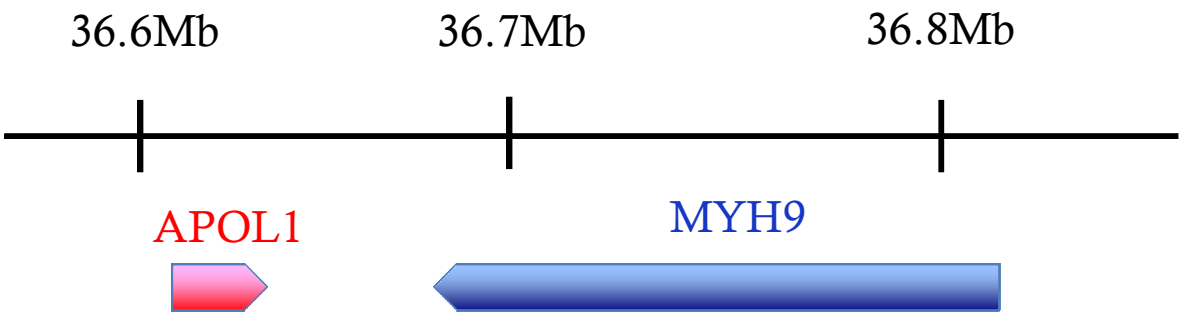



**Figure 1.8:** The position of *MYH9* and *APOL1* on Chromosome 22 *(Not drawn to scale).*

## 1.8 *MYH9* Gene

*MYH9* is a 107 kb gene located on the long arm of chromosome 22 (22q13.1). It is located from base pair 36,677,327 to base pair 36,784,063 (Human GRCh 37) and comprises 41 exons. The *MYH9* gene encodes a 1960 amino acid residue non-muscle myosin heavy chain IIA protein (Myosin IIA). It is expressed in the podocytes and is a widely distributed cellular motor protein that is essential for cytoskeleton rearrangement, cell motility, division, and cell-cell adhesion (Winkler *et al.,* 2010). It consists of two heavy chains and four light chains. The heavy chain consists of a head region and a tail region.

### 1.8.1 Mechanism of Myosin IIA action on podocytes

Myosin IIA head region, interacts with actin, an important protein that is involved in cell movement and structure. In the podocyte, the actin-myosin complex is linked to several important protein complexes, including the slit diaphragm and podocalyxin on the apical domain and alpha3/beta1 integrin and dystroglycan complex on the basal domain (Winkler *et al.,* 2010).

*MYH9* exons 1 to 19 encode the myosin globular head containing an actin binding site, and the ATP binding site while exons 20 to 41 encode the coiled tail domain. The actin-myosin complex is composed of an actin filament and multimeric myosin assembly (non-muscle myosin IIA and IIB molecules) and constitutes muscle fibers (Winkler *et al.,* 2010). ATP binds the myosin head and becomes hydrolyzed by ATPase enzyme to ADP and Pi. This "corks" the myosin head to a high energy state enabling its release from the actin filament. It then binds to another part of actin when the site becomes available, generating a power stroke that pushes the actin filament and simultaneously ADP and Pi are released (Winkler *et al.,* 2010). This is known as the actin-myosin functional cycle (Figure 1.9). This actin-myosin complex maintains cell tension and cyto-architecture, and facilitates contraction and migration (Faul *et al.*, 2007; Winkler *et al.*, 2010).

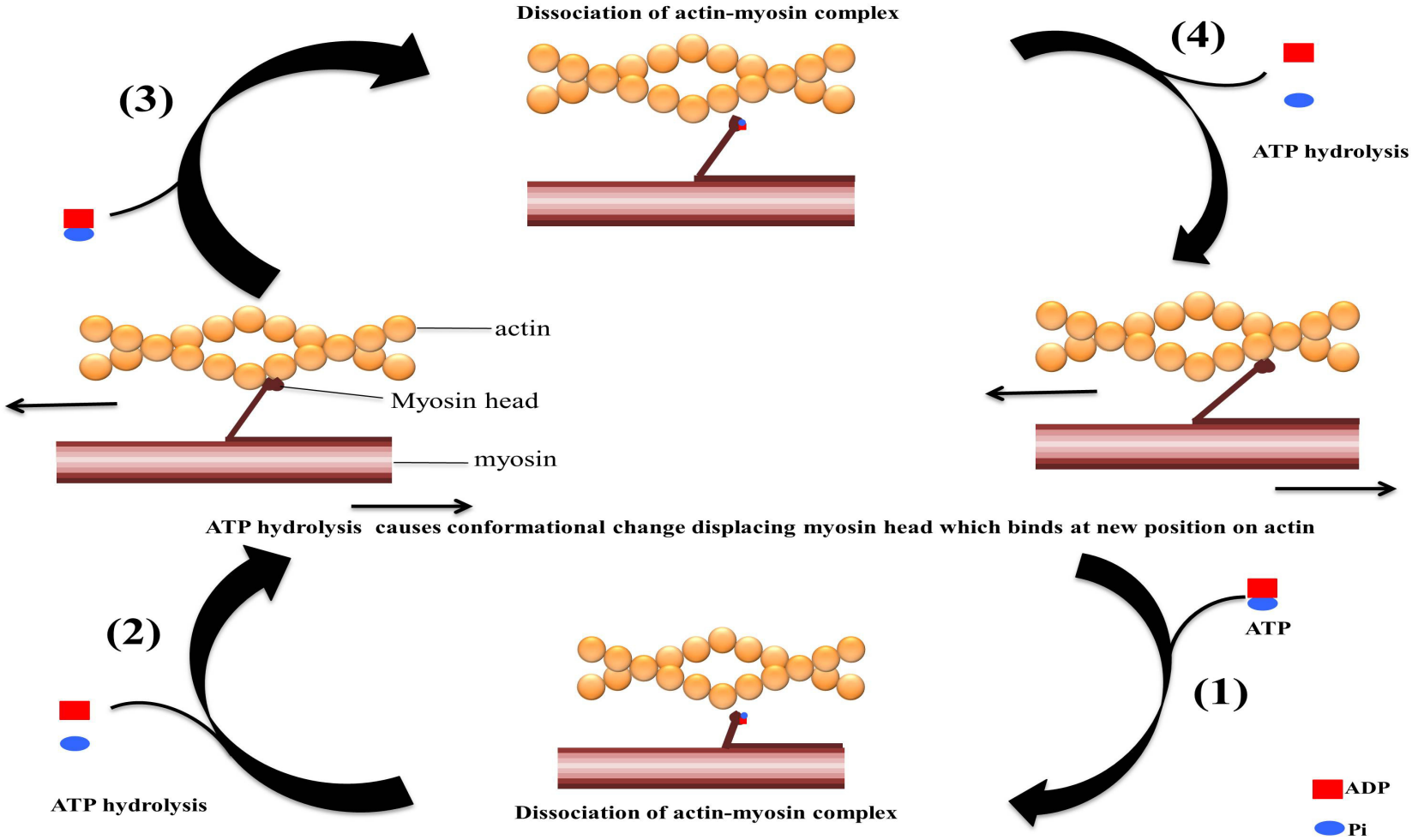



**Figure 1.9:** Model for actin-myosin functioning cycle on podocytes.
**(1):** The binding of ATP dissociates **myosin** from **actin**. **(2)**: ATP hydrolysis then induces a conformational change that displaces the myosin head group. This is followed by binding of the myosin head to a new position on the actin filament and release of ADP and $P_i$. The binding of the myosin head at new position drives actin filament sliding. **(3)** and **(4)**: The cycle continues as the movement takes place. Arrows indicate the movement of actin-myosin complex filament. Adapted from (Winkler et al., 2010).

### 1.8.2 *MYH9* related disorders

Many mutations in the *MYH9* gene lead to the production of a non-functional protein. A non-functional myosin-9 protein cannot properly interact with other subunits to form myosin IIA. There are more than 45 dominant mutations in the *MYH9* gene that have been found to cause a *MYH9*-related disorder which is characterized by four over-lapping syndromes, described as the *MYH9*-related platelet disorder spectrum: May-Hegglin anomaly, Flechtner syndrome, Epstein syndrome, and Sebastian syndrome (De Rocco *et al.,* 2013; Landi *et al.,* 2012; Savoia *et al.,* 2010). These *MYH9*-related syndromes include congenital giant platelets with thrombocytopenia (bleeding problems) with hearing impairment, renal impairment, and cataracts as someone matures into adult-hood. Although the pathological mutations span exons 1 through 41, kidney involvement is noted for missense mutations in exons 16, 25, and 30 (Winkler *et al.,* 2010). Mutations that are located near the head of the myosin protein tend to lead to a more severe disorder than mutations that are located toward the tail of the protein. Recurring mutations involving the amino acid arginine at position 702 in the protein tend to result in many problems, including a marked decrease in the number of platelets, early-onset renal disease, and hearing loss in infancy (De Rocco *et al.,* 2013).

### 1.8.3 *MYH9* associations with kidney disease

*MYH9* was identified by MALD as a major susceptibility locus for HIVAN, FSGS and ESRD. *MYH9* provides a genetic explanation since *MYH9* alleles and the haplotype frequencies known to strongly associate with kidney disease are frequently found among individuals of African descent (allele frequencies of approximately 69%) and are infrequent in Europeans (≤ 4%) (Kopp *et al.,* 2008; Oleksyk *et al.,* 2010) (Table 1.2). Hence, the effect is easiest to detect in

populations of African descent where high frequencies and strong effects of the risk alleles provide excellent statistical power even with small sample sizes (Winkler *et al.*, 2010).

*MYH9* SNPs were found to significantly associate with HIVAN and FSGS, any of the three E1 haplotype SNPs, rs4821480, rs2032487, and rs4821481 (intron 23) and rs3752462 (intron 13) and have a higher effect than any single SNP for association with kidney disease in African American populations (Kopp *et al.*, 2008). There are other *MYH9* tagging SNPs known to significantly associate with kidney disease and they are categorized into two main groups; Protective-E2 haplotype and Neutral-E3, E4 and E5 haplotypes (Figure 1.10). These occur at different frequencies worldwide, with allele frequencies of the E2 haplotype being the highest among the European population at 62% and that of neutral haplotypes being the highest among the Oceania population with 98% (Table 1.2) (Oleksyk *et al.,* 2010).

**Table 1.2**: Summary of Worldwide distribution of E1, E2, E3, E4 and E5 haplotypes (Oleksyk *et al.*, 2010)

| **Regions** | **Haplotypes** | | |
|---|---|---|---|
| | Risk Haplotype (E1) | Protective Haplotype (E2) | Neutral Haplotypes (E3,E4 and E5) |
| African | 69% | 5% | 26% |
| European | 4% | 62% | 34% |
| Middle-east | 18% | 52% | 30% |
| East Asia | 0% | 20% | 80% |
| South-central Asia | 5% | 43% | 52% |
| European America | 0% | 48% | 52% |
| Oceania | 0% | 2% | 98% |

Frequencies of *MYH9* E haplotypes.

### 1.8.4 *MYH9* E1 haplotype

This *MYH9* E1 haplotype spans 14.9kb, extending across two haplotype blocks that span introns 12 to 23. All the *MYH9* SNPs that are strongly associated with kidney disease fall within this extended block (Kopp *et al.,* 2008). The *MYH9* E1 haplotype explains most burden of major forms of kidney disease in African Americans; for instance, the attributable risks are 100% for HIVAN and 70% for FSGS respectively and a significant percentage for hypertensive

nephrosclerosis (Freedman *et al.,* 2009b; Kopp *et al.,* 2008; Winkler *et al.,* 2010). The association of *MYH9* risk alleles with HIVAN is of particular concern in sub-Saharan Africa where risk alleles are at high frequency (Figure 1.10) and more than 25 million adults and children are living with HIV (Oleksyk *et al.,* 2010).

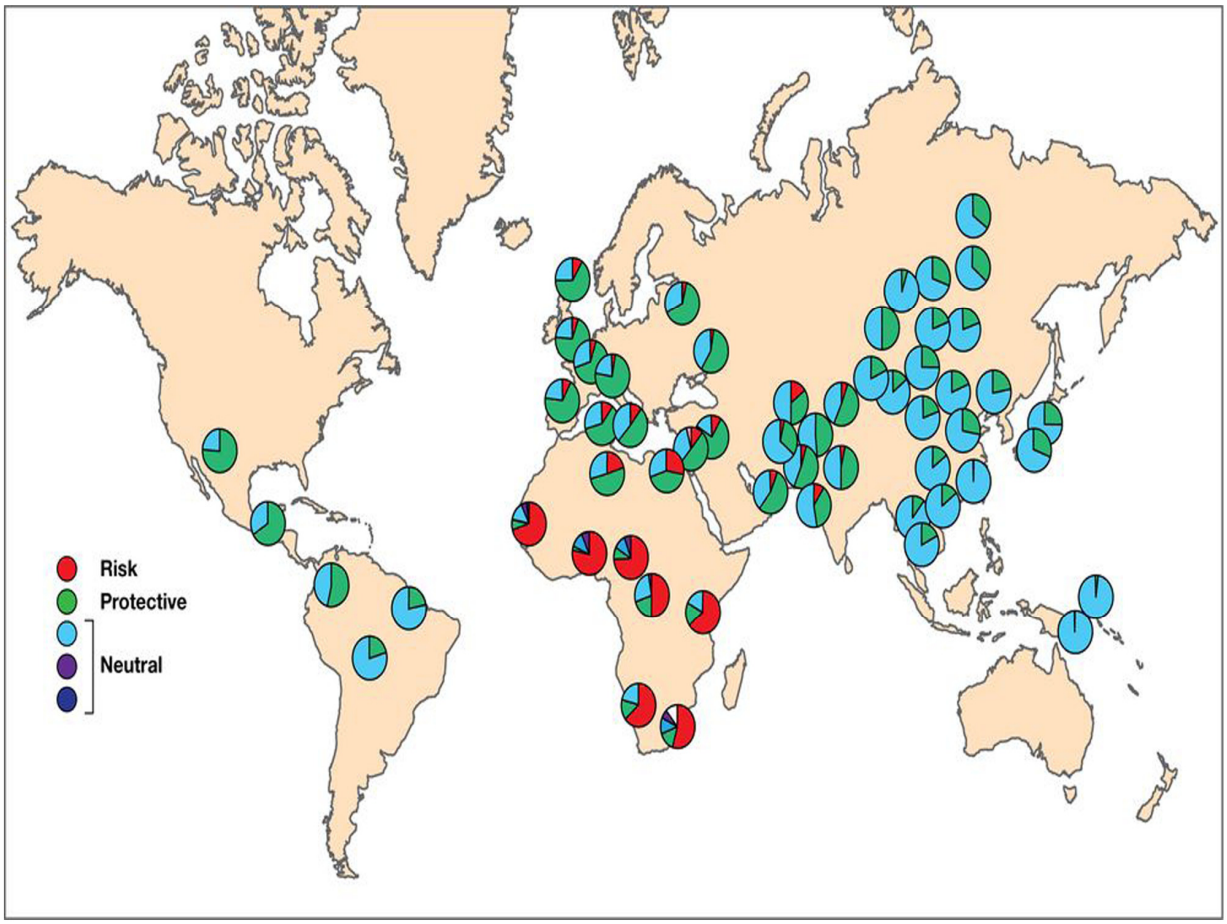



**Figure 1.10:** Worldwide distribution of the *MYH9* haplotypes frequencies.
(Human Genome Diversity Panel-HGDP). The haplotype proportions represented by the pie charts can also be found in Table 1.2 The figure was taken from an open-Access freely available online license article and therefore no permission was required (Oleksyk *et al.*, 2010).

## 1.9 *APOL1* Gene

The *APOL1* gene is a member of the *APOL* gene family which is composed of six genes (*APOL1, APOL2, APOL3, APOL4, APOL5 and APOL6*) in humans, grouped within 619kb on human chromosome 22. Among the six genes, the best studied family member is the *APOL1* gene which encodes a 43kD trypanolytic apolipoprotein L1 protein, a minor apoprotein component of high-density lipoprotein class 3 (HDL3) family which is synthesized in the liver and many other tissues, including pancreas, kidney, and brain (E. E. Smith & Malik, 2009). *APOL1* is a 14.0kb gene located on the long arm of chromosome 22 (22q12.3). More precisely, the *APOL1* gene is located from base pair 36,649,056 to base pair 36,663,576 (Human GRCh 37). It comprises seven exons and six introns (Figure 1.11). It has five functional domains: a secretory domain (SD), pore forming domain (PFD), B-cell lymphoma 2 homology domain 3 (BH3), membrane addressing domain (MAD), leucine zipper domain (LZ) and serum resistant-associated interacting domain (SRA), listed from the N-terminal to C-terminal respectively (Figure 1.12).

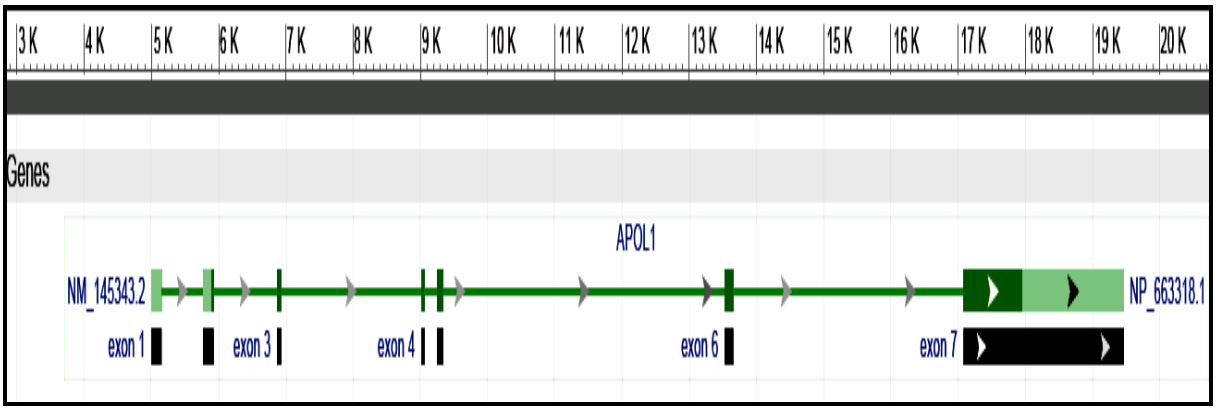



**Figure 1.11:** *APOL1* gene structure.
It is transcribed on the forward strand of chromosome 22. It contains seven exons and six introns. The SRA binding domain, the driving force of the *APOL1* activity is situated in exon seven where the *APOL1* risk variants (G1 and G2) are located. Source: PubMed (http://www.ncbi.nlm.nih.gov/gene/8542). Accessed December 2014

### 1.9.1 *APOL1* gene expression

The *APOL1* gene undergoes alternative splicing during transcription giving rise to four different messenger ribonucleic acids (mRNAs) splice variants (splice variants 1, 2, 3 and 4) which code for apolipoprotein L1 isoforms (isoform A, B and C) (Figure 1.12). Splice variants 1 (NCBI Reference Sequence: NM_003661.3) and 3 (NCBI Reference Sequence: NM_001136540.1) translate to the predominant/principal isoform A (398 amino acid protein) which contains exons 1 and 3-7, splice variant 2 (NCBI Reference Sequence: NM_145343.2) translates to isoform B (414 amino acid protein) and contains exons 1-7. The B isoform has 3 sub-isoforms, namely B1, B2 and B3. Splice variant 4 (NCBI Reference Sequence: NM_001136541.1) translates to isoform C, a 380 amino acid protein that contains exons 1, 3 and 5-7 as shown in Figures 1.12 and 1.13. The different proteins differ in the N-terminus coding regions where isoform A has the secretory/signal peptide cleavage site. Splice variant 2 has an alternate in-frame exon in the 5' coding region and uses an upstream in-frame start codon compared to variant 1. The coded protein, isoform B has a longer N-terminus compared to isoform A; as a result the signal sequence cleavage site is too far to promote cleavage. Splice variant 3 uses a different splice site in the 5' un-translated region (UTR), compared to variant 1. It encodes the same protein as splice variant 1 (isoform A). Splice variant 4 lacks an alternate in-frame 5'coding segment, compared to variant 1 resulting in a protein (isoform C) which has a shorter N-terminus as compared to isoform A and therefore lacks the signal sequence cleavage site. This means exon 4 is lacking in isoform C. (Figure 1.13). (http://www.ncbi.nlm.nih.gov/gene/8542).

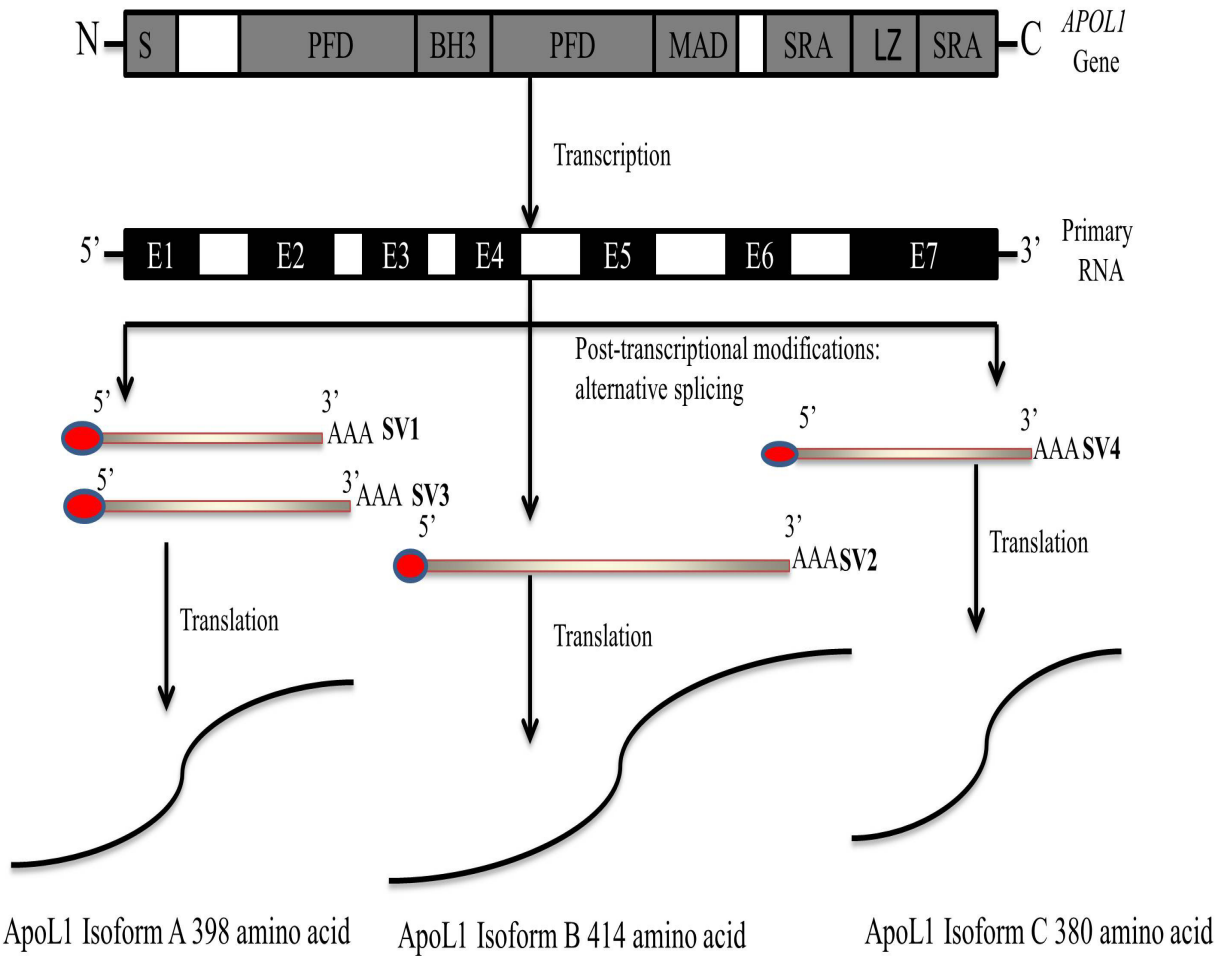



**Figure 1.12:** Expression of *APOL1* showing post-transcription modification and translation to its precursors. Secretory domain (SD), pore forming domain (PFD), B-cell lymphoma 2 homology domain 3 (BH3), membrane addressing domain (MAD), Leucine zipper (LZ) domain and serum resistant-associated (SRA) interacting domain. Splice variant (SV), Exon (E). *ApoL1*; Apolipoprotein L1protein.

The predominant *ApoL1* protein is unique among *APOL* gene family proteins because of the N-terminal signal peptide that enables it to be secreted outside the cell. Other *APOL* gene family proteins do not have this signal peptide and have predicted localizations inside the cell (Page *et al*., 2001). The *APOL1* encoded protein (*ApoL1*) is also known to lyse pathogenic *Trypanosoma* (protozoa that causes Human African Trypanosomiasis) during blood stream infections. Most other primates lack *APOL1* gene and do not have trypanolytic activity, except for certain species of baboons, mandrills, and mangabeys (Lugli *et al.,* 2004).

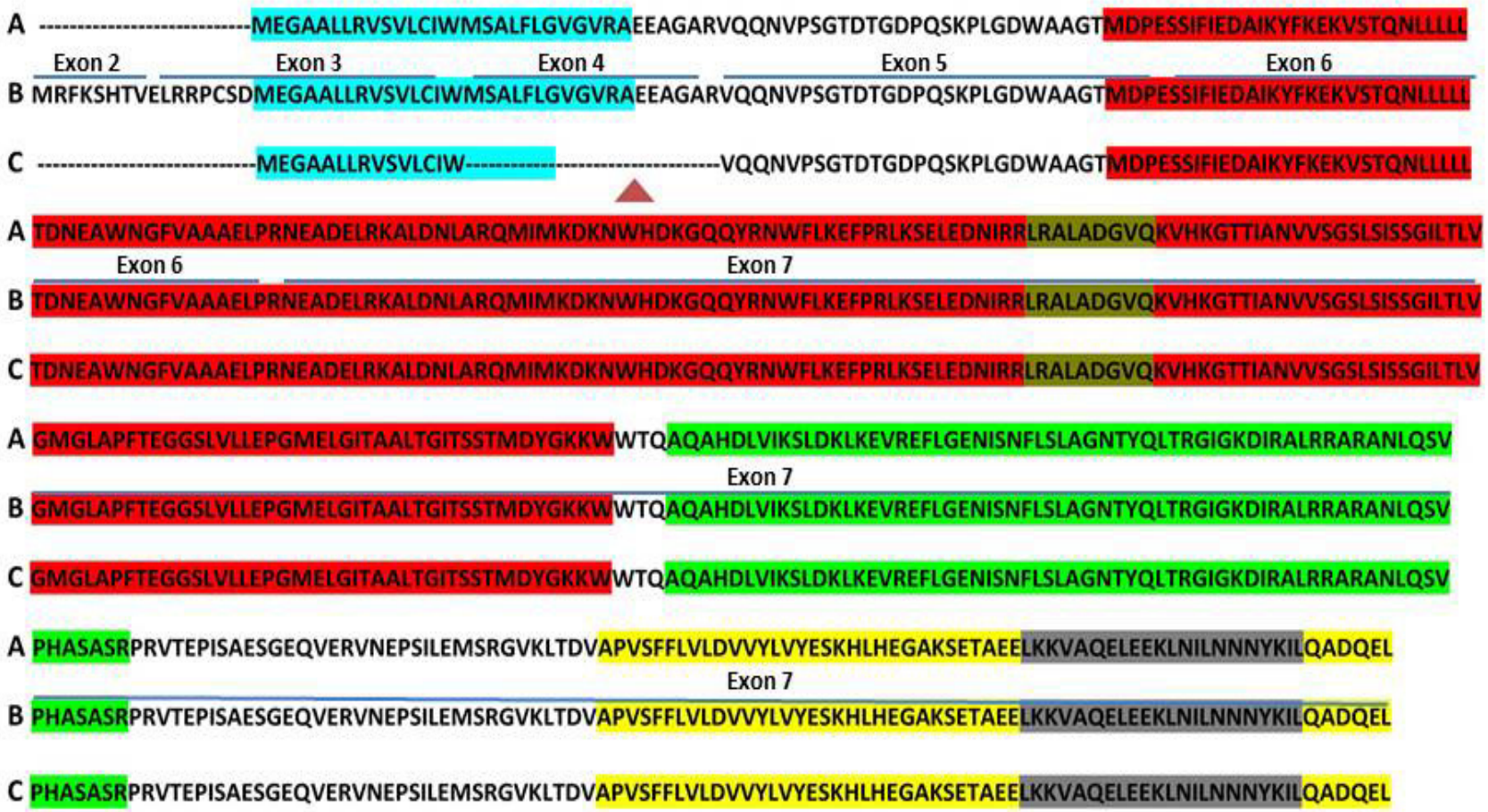



**Figure 1.13:** Amino acid sequence showing apolipoprotein L1 isoforms

**A** (NCBI Reference Sequence: NP_001130012.1), **B** (NCBI Reference Sequence: NP_663318.1) and **C** (NCBI Reference Sequence: NP_001130013.1). Secretory peptide (blue), pore-forming domain (red), BH3 domain (dark yellow), membrane-addressing domain (green), serum resistant associated domain (yellow) and Leucine zipper domain (grey). The secretory peptide is responsible for the secretion of apolipoprotein L1 outside the cell, BH3 domain is responsible for apoptosis, MAD responsible for the release of apolipoprotein L1 from the HDL particles into the lysosome, SRA, responsible for the binding of trypanosomiasis serum associated protein onto apolipoprotein L1 for its activity to be effected and Leucine zipper domain for the binding of DNA. Red caret indicates predicted signal sequence cleavage site present in A, but absent in C. The site in B may be too far from the N terminus to promote cleavage. Exons 2-7 of *APOL1* encode the amino acid sequences shown above. All exons are translated only in apolipoprotein L1 isoform B. Amino acid residues shown in between the exons, are encoded by nucleotides from neighbouring exons.

### 1.9.2 Human African Trypanosomiasis ("sleeping sickness")

Human African Trypanosomiasis (HAT) also known as sleeping sickness, is a parasitic disease of humans caused by protozoa of the species *Trypanosoma brucei rhodensiense* in East Africa and *Trypanosoma brucei gambiense* in West Africa (Figure 1.14). "The two protozoan species are morphologically indistinguishable, but have different epidemiological features", (WHO 2010). "Several species of Tsetse flies (*Hematophagous glossina)* mostly *Glossina palpalis*, are the vectors of these related diseases, and are responsible for cyclical transmission of the parasitic protozoan between numerous vertebrate hosts" (WHO 2010) and the central nervous system (CNS) is affected in both forms of sleeping sickness (Enanga *et al.*, 2002; Trenaman *et al.,* 2013).

"The West African form of trypanosomiasis is the most common form of HAT accounting for about 95% of reported cases and causes a chronic infection. In contrast, the East African form of trypanosomiasis is less common, accounting for approximately 5%, and is characterized by rapid and acute development of the disease", (Centers for Disease Control and prevention). In both, untreated patients die within weeks or months of infection.

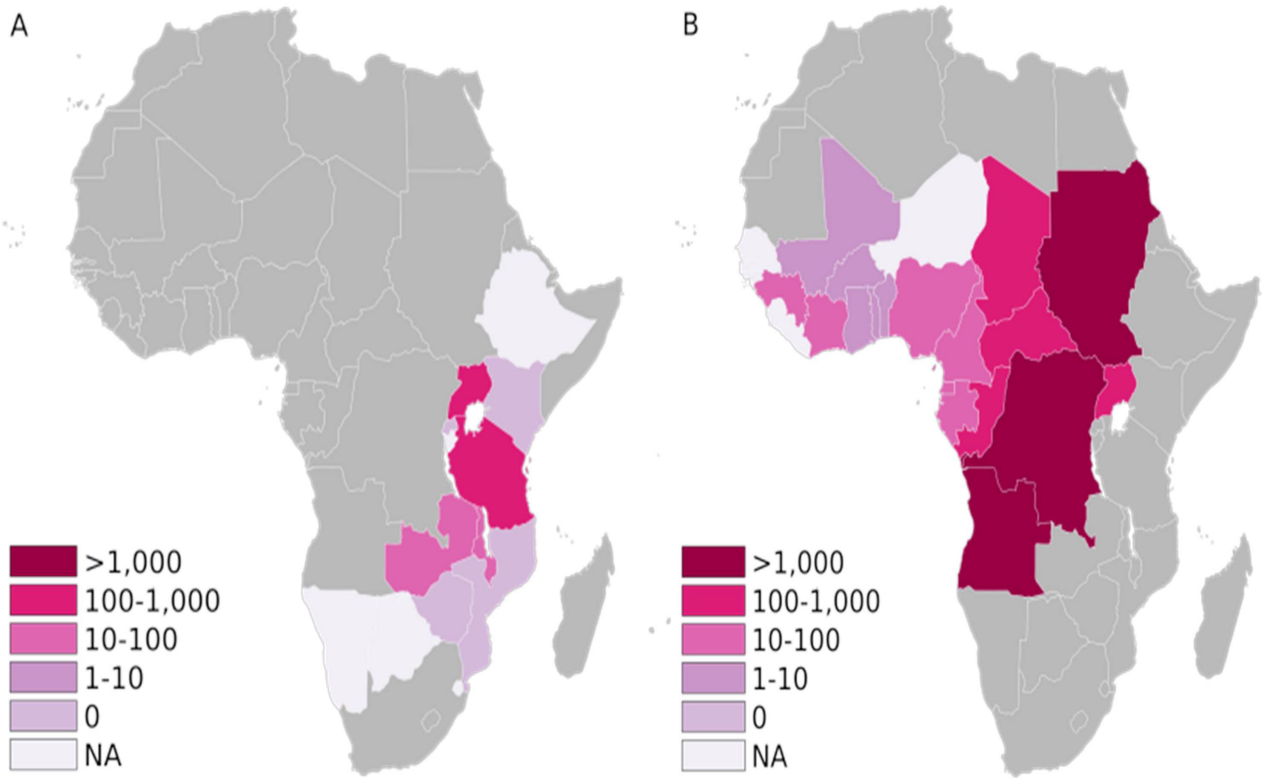



**Figure 1.14:** A. Distribution of Human African Trypanosomiasis (*T.b.rhodesiense*) and B. *T.b.gambiense* (W.H.O. 2010).

HAT epidemics were more rampant in the past. However, according to the most recent WHO estimates, "60 million people are at risk for HAT today with approximately 500,000 people currently infected and the disease discontinuously spread over 9 million square kilometers and affects populations across 37 sub-Saharan countries" **(Figure 1.14).**

### 1.9.3 Trypanolytic activity of Apolipoprotein L1

A complex co-evolution has occurred between the Trypanosome species and primates/humans that has led to the emergence of trypanosome-specific defence and counter-defence measures leading to the first line of defence in humans facilitated by the trypanolytic apolipoprotein-L1 (Capewell *et al.,* 2015). Trypanolytic activity of apolipoprotein L1 is carried out through a class of HDL particles known as Trypanolytic factor (TLF). There are two main sub-classes of

trypanolytic factors: TLF1 in the HDL3 which is the densest HDL sub-fraction and a large and dense TLF2 that is not considered to be a normal HDL particle (Gillett & Owen, 1991; Hajduk *et al.,* 1989; Wheeler, 2010). These two classes of HDL particles share several protein components including apolipoprotein L1, apolipoprotein A1 and haptoglobin related protein (Hpr) indicating that one of these is responsible for TLF activity (Tomlinson *et al.,* 1997). However, it has been shown previously that apolipoprotein A1 lacked trypanolytic activity (Owen *et al.,* 1992; Tomlinson *et al.*, 1997) and therefore, apolipoprotein L1 and Hpr are thought to be responsible for trypanolysis. Hpr-induced lysis is proposed to occur by oxidative damage to the lysosomal membrane caused by hydroxyl radicals while apolipoprotein L1-induced lysis is by the autophagic process (Bishop *et al.,* 2001; A. B. Smith *et al.*, 1995).

Once infected via the vector (tsetse fly of the *Glossina* species), the *Trypanosoma* subspecies (*T.b rhodesience* or *T.b gambiense*) in the blood stream endocytose the apolipoprotein L1-containing HDL3 particles as a means of obtaining both lipids and iron. This process is facilitated by the presence of a trypanosome surface receptor that recognizes Hpr and HDL-receptor mediated endocytosis (Figure 1.15).

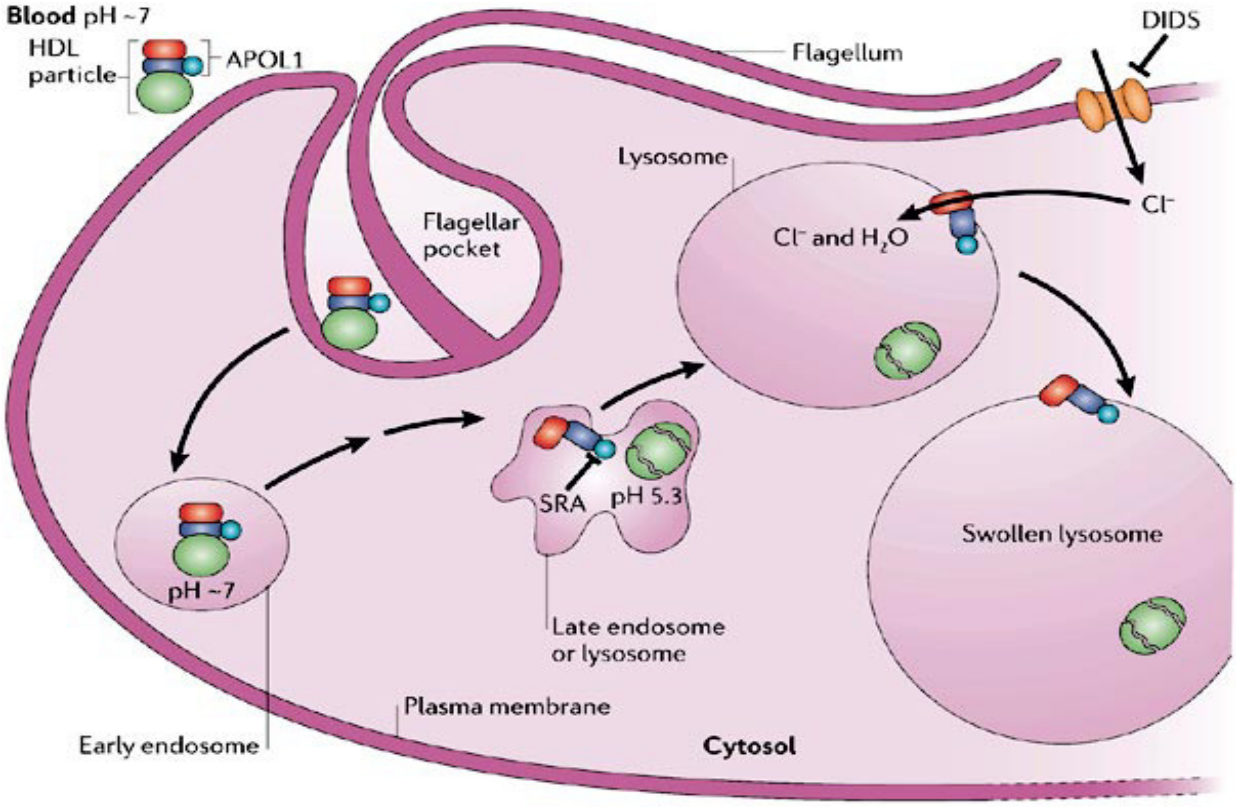



**Figure 1.15:** Apolipoprotein L1-mediated Trypanolytic action.
Apolipoprotein L1contains a pore-forming domain (red), a membrane-addressing domain (purple) and a serum resistance associated interacting domain (blue) associated with HDL particles (green) circulating in the blood. *APOL1*-HDL3 complex is engulfed by trypanosomes in the bloodstream through HDL-receptor-mediated endocytosis. Once internalized, HDL particles are taken through the endocytic pathway. Cl-: Chloride ions. DIDS: 4, 4-diisothiocyano-2, 2-disulphonate stilbene. Permission obtained from Nature Publishing Group © Pays, E. *et al*. *Nature Reviews Microbiology* **4**, 477-486 (June 2006). (Pays *et al.*, 2006; Vanhollebeke & Pays, 2010). Appendix 15 C

The apolipoprotein L1-containing HDL3 particles then follow the endocytic pathway (from early endosomes with neutral pH~7 to late endosomes with acidic pH~5). Once inside the late endosomes, the *T.b rhodesience* or *T.b gambiense's* serum resistant associated (SRA) protein interacts with the apolipoprotein L1 C-terminal SRA-binding domain in order to neutralize the effect of apolipoprotein L1 and prevent its lytic activity. SRA is a truncated variant surface glycoprotein (VSG). Due to this pH change, conformational changes in the apolipoprotein L1 MAD domain occur resulting in the opening of a salt-bridge-linked on the lysosomal membrane, releasing apolipoprotein L1 from the HDL3 particles into the lysosome (Perez-Morga *et al.*,

2005). Once inside the lysosome, the apolipoprotein L1 utilizes the PFD to create anion-specific pores that enables the movement of chloride ions ($Cl^-$) from the cytoplasm into the lysosome, creating an anion imbalance resulting in the influx of 4, 4-diisothiocyano-2, 2-disulphonate stilbene (DIDS) in the cytoplasm to reinstate an osmotic balance destabilized by the efflux of $Cl^-$ into the lysosomes. DIDS is an anion channel inhibitor and *Perez Morga et al., 2005* showed that, addition of DIDS to the medium led to a substantial delay in lysis by normal human serum. This creates an osmotic pressure that leads to the movement of water into the lysosome resulting in uncontrolled osmotic swelling of the lysosome which increases the intracellular pressure resulting in damage to the Trypanosome's membrane, leading to lysis and damage of plasma membrane, triggering the autophagic response (Perez-Morga *et al.*, 2005; E. E. Smith & Malik, 2009). In addition, there has been some evidence suggesting that trypanolysis involves the swelling of the entire trypanosome (Thomson & Finkelsteil, 2015).

Apolipoprotein L1 is also believed to initiate programmed cell death (PCD) through apoptosis of host cells due to the presence of the BH3 protein domain **(Figures 1.12 and 1.13)**. Some studies have shown that one part of the apolipoprotein L1 is shaped like B-Cell lymphoma- 2 (Bcl-2), a protein known to regulate apoptosis (Happo *et al.*, 2012; E. E. Smith & Malik, 2009; Wan *et al.*, 2008). BH3 proteins function as upstream activators of PCD markers [Cysteine Aspartyl-specific proteases (caspases), Bcl2, Fas and Bax] responding to stimuli such as DNA damage, cytokine withdrawal, or cell detachment (Strasser, 2005). It is believed that apolipoprotein L1 uses this pathway to achieve its apoptotic activity. Cell damaging events initiates the mitochondrial (intrinsic) apoptotic signalling pathway. This is the pathway that pro-apoptotic members of the Bcl-2 family get activated and translocate to the mitochondria and in the process neutralize anti-

apoptotic proteins in turn promoting apoptosis (Happo *et al.,* 2012; Rodriguez & Lazebnik, 1999; Strasser, 2005; Thome & Tschopp, 2001).

### 1.9.4 Trypanolytic activity of Apolipoprotein L1 in relation to kidney disease

The trypanolytic activity of *APOL1* could be the driving force of development of kidney disease as hypothesized in Figures 1.15 and 1.16. Since apolipoprotein L1 is a circulating protein, component of HDL3, and the function of HDL3 is to protect low density lipoprotein (LDL) from oxidation, it is postulated although unclear, that apolipoprotein L1 plays a role in this anti-oxidation process. *APOL1* G1 and G2 non-synonymous variants described in detail below (1.9.5) are located on the C-terminal domain in the last exon (exon 7) (residues 342-398) of the apolipoprotein L1, which is the binding site for the trypanosomal SRA protein (Vanhamme *et al.,* 2003). The fact that the two pathogenic variants occur at this domain, and being the only disease-associated *APOL1* variants that lyse sleeping sickness disease-causing *Trypanosoma brucei rhodesiense* (Kronenberg, 2011), could be an indication for the pathogenesis of kidney disease. The high prevalence of these variants occurring over a period of time in individuals of African ancestry is mainly attributed to natural selection as a result of Trypanosomiasis endemic in Africa years ago hence providing an evolutionary advantage. The evolution of these *APOL1* variants was a critical survival factor, which on the other hand may have contributed to the high incidence and prevalence of kidney disease in the African population (Genovese *et al.,* 2010).

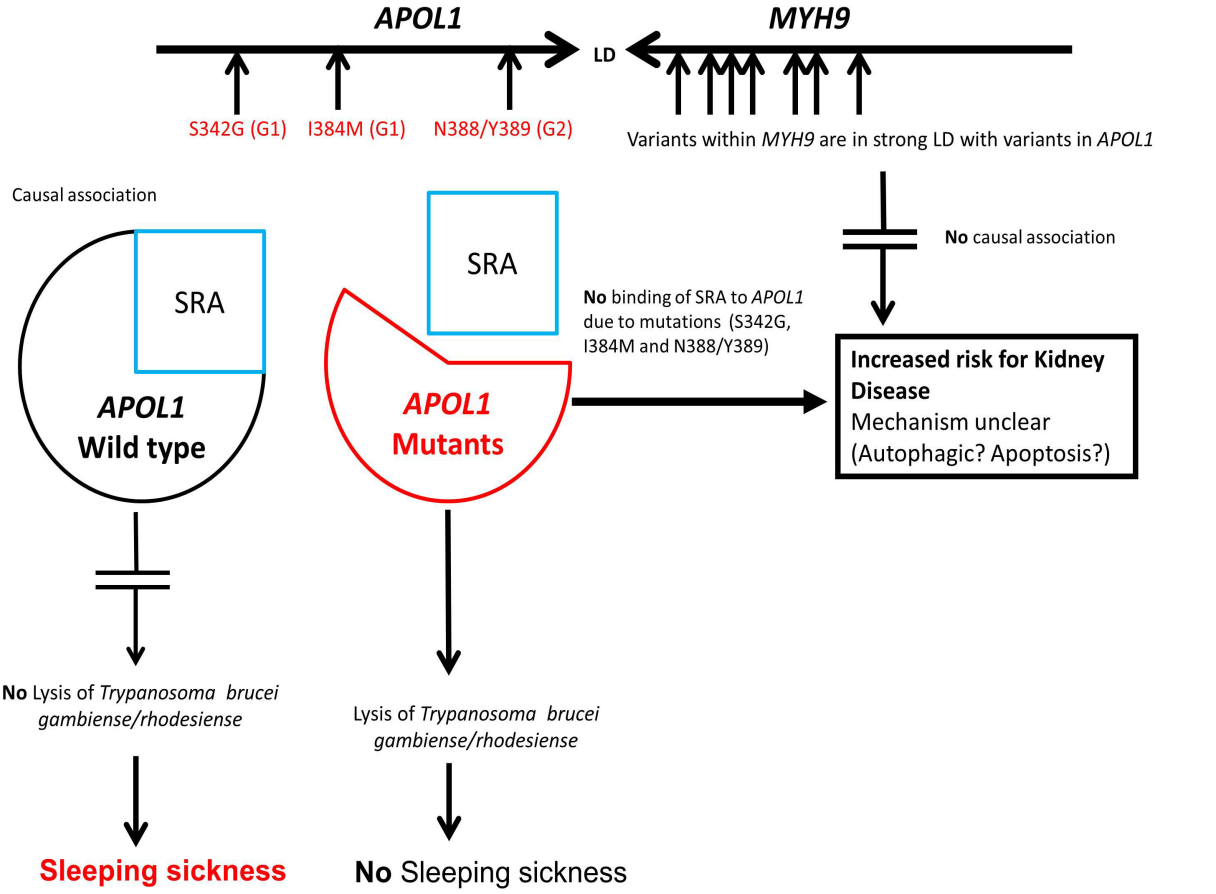



**Figure 1.16:** Association between *APOL1* risk variants, trypanolytic ability of apolipoprotein L1 and increased risk for kidney disease. SRA; serum resistant associated. LD; Linkage disequilibrium. Adapted from (Kronenberg, 2011).

In case of the *APOL1* gene wild type alleles, the apolipoprotein L1 interacts with the SRA expressed by the parasites (Figure 1.16) resulting in the inactivation of apolipoprotein L1 (no longer able to lyse the parasites), hence the parasites manifests themselves and in turn causing sleeping sickness (Kronenberg, 2011). However, the non-synonymous variants are located exactly at the site where SRA protein binds apolipoprotein L1 and therefore, in the case of the mutations at positions 342, 384 and 388-389 (presence of G1 and G2), the SRA protein becomes

incapable of binding to apolipoprotein L1, allowing apolipoprotein L1 to lyse the parasite through ionic pore formation in endosomal membranes of the parasite.

It could be through this process that a positive selection for these variants occurred in Africans since they had a major survival advantage by being resistant to *T. brucei rhodesiense* hence against sleeping sickness. However, these variants are associated with an increased risk for non-diabetic kidney disease for which the molecular mechanisms have not yet been established (Kronenberg, 2011).

### 1.9.5 *APOL1* association with Kidney Disease

Intensive efforts and analysis of the *MYH9* gene were performed in order to find a causative functional variant within the African ancestry population that predisposes to kidney disease (Behar *et al.*, 2010; Freedman *et al.*, 2009a; Winkler *et al.*, 2010). However, even after extending association studies to earlier stages of related kidney disease phenotypes and to groups with admixed African population (Tzur *et al.,* 2010; Winkler *et al.*, 2010), functional mutations were not identified in the *MYH9* gene. This could be an explanation behind the absence of *APOL1* risk variants coinciding with the absence of HIVAN in spite of high frequency of *MYH9* risk haplotype in the Ethiopian population (Behar *et al.,* 2011). This led a group of scientists (Tzur *et al.*, 2010) to re-examine the region surrounding the *MYH9* gene and they found novel missense mutations with causative functional effects in the neighbouring *APOL1* gene, which are more significantly associated with kidney disease than all previously associated *MYH9* SNPs. These missense mutations are I384M-rs60910145 (G1), S342G-rs73885319 (G1) and a 6 base pair deletion-rs71785313 (-\TTATAA) - N388_Y389del referred to as G2. These two risk alleles are mutually exclusive and in negative LD. These coding variants (G1 and G2) are suggested to be causally related to CKD and provide an explanation for selection of *APOL1*-associated CKD risk

polymorphisms as a protective measure against Trypanosomiasis. Studies by Tzur *et al.,* (2010), Genovese *et al.,* (2010) and Kopp *et al.,* (2011) (Genovese *et al.*, 2010; Kopp *et al.*, 2011; Tzur *et al.*, 2010), revealed evidence of a strong association within a 10kb region in the exon 7 of the *APOL1* gene. The strong signal was found to be at non-synonymous coding variants, rs73885319 and rs60910145 which were in perfect linkage disequilibrium ($r^2$ = 1.0). Controlling for the effects of these two variants exposed a second strong *APOL1* signal, 12 base pairs from I384M. This signal is a 6-base pair deletion represented by rs71785313 which removes two amino acid residues (Asparagine-N and Tyrosine-Y). The odds ratio (OR) of association for carrying at least one risk (G1-G2) allele was 10.5 (95% confidence interval [CI] 6.0 to 18.4) in the study by Genovese *et al.,* (2010) (Genovese *et al.*, 2010). Controlling for both G1 and G2 did not result in significant association with *MYH9.* However, controlling for *MYH9* variants maintained significant *APOL1* signal at G1 and G2. Kopp *et al.,* (2011), in a larger FSGS and HIVAN cohort confirmed the association but this time, a greater association was observed in HIVAN (OR=29.2 CI 13.1 to 68.5 $P=6 \times 10^{-22}$) compared to FSGS (OR=16.9 CI 11.0 to 26.5 $P=1.3 \times 10^{-48}$) (Kopp *et al.*, 2011). The authors reported that 50% of African Americans with two *APOL1* risk alleles would develop HIVAN if not on antiretroviral therapy. The observed mode of inheritance of the *APOL1* risk variants was fully recessive in both FSGS and HIVAN cohorts. Therefore, the originally described associations between variants within the *MYH9* gene alone had to be revised since the association finding was thought to be caused by the high LD between the variants in the *MYH9* and the *APOL1* genes (Kronenberg, 2011). These three variants within the *APOL1* gene were found to have a very high frequency among African populations (Figure 1.17). G1 and G2 allele frequencies are shown in Table 1.3.

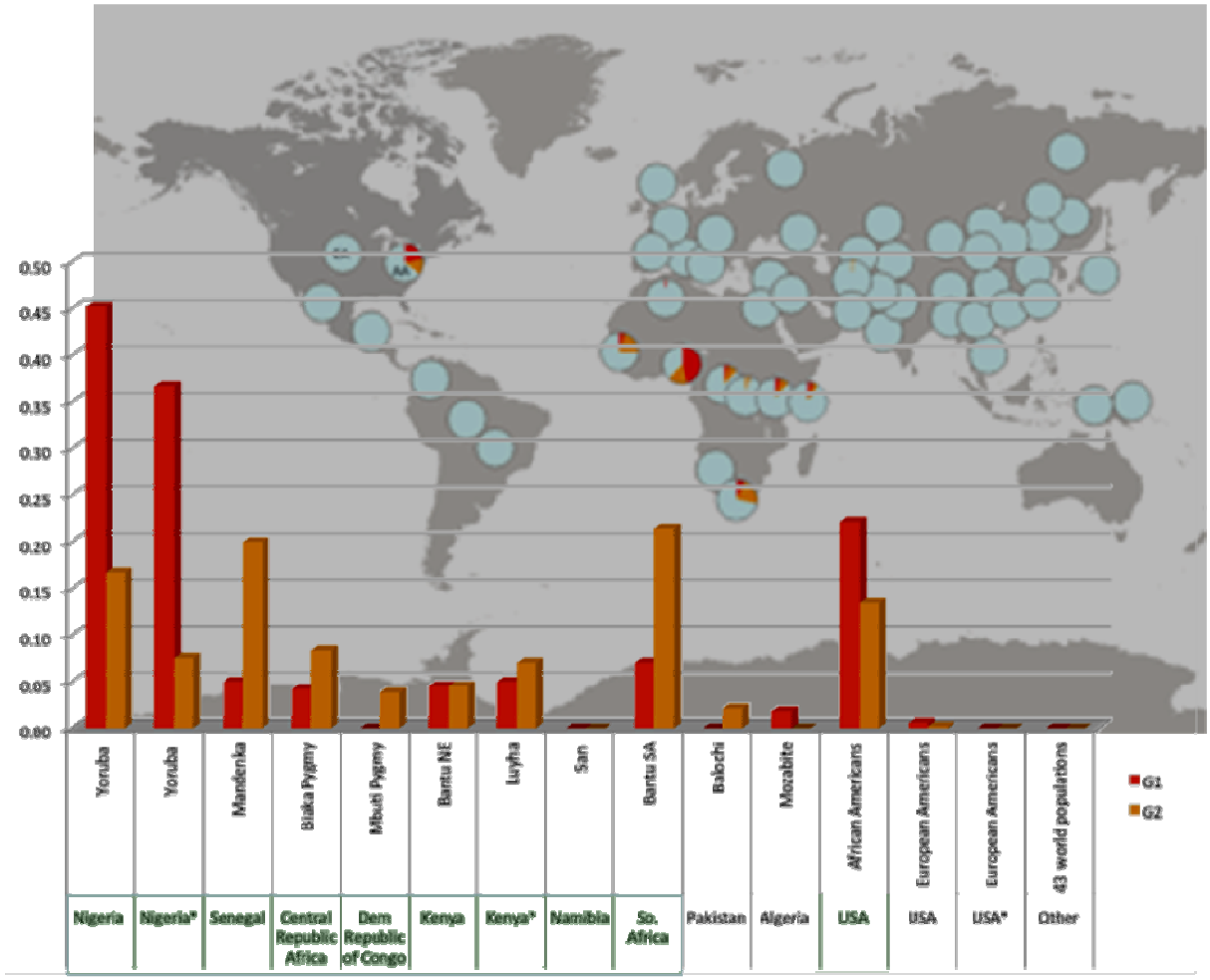



**Figure 1.17:** Worldwide frequency distribution of *APOL1* variants G1 and G2.
Data obtained from: Centre d'Etude du Polymorphisme Humain (CEPH), Human Genome Diversity Project (HGDP), Luhya from Webuye, Kenya (International HapMap Project), African Americans and European American controls from the National Institutes of Health (NIH) FSGS study cohort. The pie charts represents allele frequencies of G1 (red), G2 (orange), and wild-type alleles (light blue) in each population overlaid upon a world map (Kopp *et al.*, 2011). Permission obtained from the American Society of Nephrology publishing group (Journal of the American Society of Nephrology). Appendix 15 D.

**Table 1.3**: Allele frequencies of SNPs that delineate *MYH9* E1 and *APOL1* G1/G2 risk haplotypes.

| Genes | SNPs | Alleles | Allele Frequencies | | | | |
|---|---|---|---|---|---|---|---|
| | | | YRI | LWK | MKK | EUR | SA |
| ***MYH9*(E1haplotype)** | rs4821480 | G/T | 0.238 | 0.321 | 0 | 0.086 | ND |
| | rs2032487 | C/T | 0.238 | 0.306 | 0 | 0.075 | 0.3138 |
| | rs4821481 | C/T | 0.273 | 0.358 | 0.433 | 0.075 | 0.3298 |
| | rs3752462 | T/C | 0.224 | 0.227 | 0.314 | 0.374 | 0.1862 |
| ***APOL1*(G1/G2)** | rs73885319 | A/G | 0.442 | 0.037 | ND | 0 | 0.04787 |
| | rs60910145 | G/T | ND | ND | ND | 0 | ND |
| | rs71785313 | Indel | 0.13 | 0.25 | 0.25 | 0 | ND |

These are frequencies in the general population among the 3 African populations compared to a European population and a South African black population. Indel; Insertion/deletion, ND; no data meaning the SNP has not been typed in HapMap or 1000 Genomes project nor was typed in the SA study, 0; Not observed as polymorphic in this population, YRI; Yoruba in Ibadan, Nigeria, LWK; Luhya in Webuye, Kenya and MKK; Maasai in Kinyawa, Kenya, EUR; Europeans (http://www.ensembl.org and http://spsmart.cesga.es/search.php/dataSet=engines), SA; South Africa (May *et al*., 2013).

## 1.10 Relationship between *MYH9* and *APOL1* risk variants

Initial studies in early genome-wide mapping, identified African ancestral genetic variants centered at the genomic region containing *MYH9*, encoding non-muscle myosin heavy chain IIA in *MYH9* that were strongly associated with kidney disease (Freedman *et al.,* 2009b; Kopp *et al.,* 2008; Lipkowitz *et al.,* 2013; Winkler *et al.*, 2010). Further studies however, showed that these SNPs were only "spectators" for *APOL1* SNPs with stronger association with HIVAN and FSGS (Kopp *et al.,* 2008; Kronenberg, 2011; Rosset *et al.*, 2011; Tzur *et al.,* 2010). *APOL1* gene is located approximately 20.0kb upstream of *MYH9,* encoding apolipoprotein L1 (Figure 1.18). However, accumulating evidence now indicates that both genes may independently contribute to kidney disease susceptibility, and that although *APOL1* variation probably accounts for a majority of susceptibility to HIVAN and FSGS, it does not explain a significant concept of disease (Hays & Wyatt, 2012; Wyatt *et al.,* 2012). The *APOL1* gene was also strongly associated with non-diabetic CKD, with statistical associations being stronger than that of *MYH9* (Genovese *et al.,* 2010). This high association was caused by two *APOL1* risk alleles, G1 (rs73885319 and rs60910145) consisting of two missense variants (S342G and I384M); and G2 (rs71785313), a 6

base pair (bp) inframe deletion (N388del:Y389del). It is therefore hypothesized and postulated that, *MYH9* risk variants are in high LD with *APOL1* risk variants but the causal effect is driven by the *APOL1* risk variants.

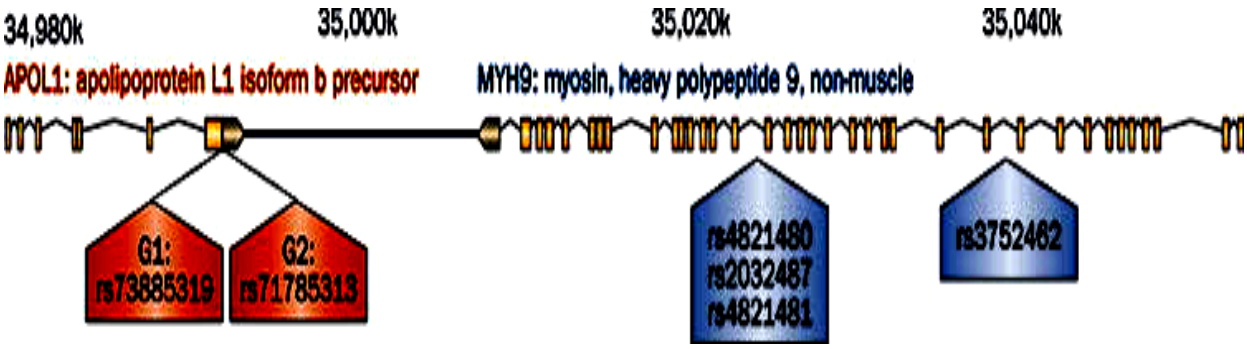



**Figure 1.18:** The position of *MYH9* and *APOL1* risk variants on chromosome 22 showing the relationship between *APOL1* and *MYH9* genetic variants associated with common forms of kidney disease.
Permission obtained from Nature Publishing Group © Drawz, P. E. & Sedor, J. R. *Nat. Rev. Nephrol.* 7, 458–468 (2011) (Drawz & Sedor, 2011). Appendix 15 E.

## 1.11 Rationale for the study

The burden of renal disease associated with HIV infection among black individuals is on the increase (Estrella *et al.*, 2010; Post & Holt, 2009). According to the ESRD Patients in 2013: A Global Perspective report, an estimated 3.2 million people are on renal replacement therapy with an increased prevalence in individuals of African ancestry. This is mainly attributed to the fact that, other than environmental and socio-economic factors, these individuals also carry 1 or 2 risk alleles of the *APOL1* gene which codes for an apolipoprotein L1. Mutants of this protein have been shown to have trypanolytic factors and therefore it protects against Trypanosomiasis but at the same time predisposes an HIV infected individual of African ancestry to CKD. Further studies are therefore needed in order to understand the genetic and molecular mechanism of apolipoprotein L1 and its role in inducing HIV-associated kidney disease. The dramatic reduction in the incidence of opportunistic infections as a result of HAART has resulted in a shift of health concern toward chronic morbidities, including renal dysfunction, however, the burden

of renal disease associated with HIV infection is increasing exponentially (D'Agati & Appel, 1997; de Silva *et al.,* 2007; Wyatt & Klotman, 2007). The high prevalence of HIVAN and the impact on morbidity and mortality drives the need to come up with definitive indicators of disease progression and prognosis. There has been an increase in the number of patients progressing to ESRD, the majority of who are working-age adults, breadwinners and parents. Countries cannot afford to lose people in this age group especially in the developing world where it is not possible to offer renal replacement therapy to all those in need. Data from this study will enable us to determine the impact of genetic and molecular factors in the pathogenesis of HIVAN and other forms of HIV CKD and determine their potential role in terms of prediction or susceptibility to developing kidney disease in the presence of HIV infection. In addition, a better understanding of the genetic predisposition to kidney disease in a South African population will assist in alerting clinicians to kidney disease susceptibility in HIV-positive patients and the need to increase surveillance and possibly adjust treatment. At present, little is known about the biology of apolipoprotein L1 and its role in kidney disease although it is postulated that since it has the BH3 domain, it may use a similar mechanism described earlier in trypanolysis to cause kidney disease. This is the first study of its kind in Africa to functionally characterize apolipoprotein L1-related kidney disease.

## 1.12 Hypothesis

Genetic susceptibility to kidney disease in the context of HIV infection is influenced by variations in the *MYH9* and *APOL1* genes. mRNA isoforms B and C, having a defective secretory domain and lacking N-terminal secretory peptides respectively are highly expressed in HIVAN and the expression of high density lipoprotein-binding apolipoprotein L1 isoforms B and

C coupled with C-terminal risk alleles (G1/G2) play a vital role in induction and development of HIV-associated kidney disease in individuals of African ancestry.

## 1.13 Objectives of the study

1. To determine the prevalence of high risk genetic variants in *MYH9* and *APOL1* genes in HIVAN patients compared to patients with other forms of CKD.

2. To examine the association of these *MYH9* and *APOL1* high risk genetic variants with HIVAN patients compared to patients with other forms of CKD.

3. To profile the pattern of apolipoprotein L1 messenger RNA N-terminal isoform expression in HIVAN, HIV-positive FSGS and other CKD, HIV-negative FSGS and other CKD, HIV-associated immune complex kidney disease (HIVICK) and in normal kidney using a semi-quantitative reverse transcription polymerase chain reaction method.

4. To determine the expression patterns of the N-terminal apolipoprotein L1 isoforms in Human embryonic kidney (HEK) 293 cells by developing isoform constructs using PCR-based cloning technique and transfecting them in HEK 293 cells and using confocal microscopy to assess their expression levels.

# Chapter 2

# 2) MATERIALS AND METHODS

## 2.1 Study population

The Department of Anatomical Pathology, University of the Witwatersrand, Charlotte Maxeke Johannesburg Academic Hospital (CMJAH) has archived human pathology tissues since 1940s and kidney biopsy tissues have been archived since 1970s. For our study, archived kidney biopsies of adults (aged ≥18years) were retrieved according to their histological diagnosis and HIV status as shown in Table 2.1. Standard diagnostic criteria were employed in the renal histological diagnoses. The histology diagnoses were also confirmed by an independent pathologist (Prof. Agnes Fogo from the University of Vanderbilt Medical Center, Department of Pathology, Microbiology and Immunology, Division of Renal Pathology/Electron Microscopy, Nashville, Tennessee, USA). HIVAN was defined by "the presence of glomerular capillary collapse and glomerular visceral epithelial cell proliferation affecting at least one glomerulus, together with microcystic tubular dilatation and interstitial inflammation" (Figure 2.1A). HIV-associated immune complex kidney disease (HIVICK) (Figure 2.1C) was defined by "the presence of glomerular endocapillary cell proliferation together with an increased deposition of at least one immunoglobulin"; hence HIV (+) IgA nephropathy was included. Other histopathology categories included the following: HIV (+) and HIV (-) FSGS (Figure 2.1B), other glomerulonephritis (membranoproliferative glomerulonephritis and membranoproliferative glomerulonephritis consistent with C3 glomerulopathy, membranous glomerulonephritis, HIV (-) IgA nephropathy and lupus nephritis) and other kidney disease (thrombotic microangiopathy, nodular glomerulosclerosis, minimal change disease, arterionephrosclerosis and benign

nephrosclerosis, oxalosis and global glomerulosclerosis) Table 2.1. The viral load and CD4 data was obtained from the patient records.

We collected blood samples from 54 individual volunteers of each of 2 control groups: HIV (+) individuals without clinical evidence of CKD [HIV (+) Controls] from the HIV clinic at Charlotte Maxeke Johannesburg Academic Hospital and a population control (PC) group from the National Health Laboratory Service (NHLS) who were all "healthy" individuals who gave consent to be enrolled in population genetic studies as control individuals. All HIV (+) case individuals with CKD were ART-naïve. However, 75% of the HIV (+) control group were on ART. This study was approved by the University of the Witwatersrand Human Research Ethics Committee (medical) with approval numbers M10745 and M111185 (Appendix 3) and prospectively participating patients and controls gave their consent.

## 2.2 Sample Size

The sample size was based on statistical analysis using effect size and frequency estimates for Yoruba from Ibadan (YRI) and African Americans from previous studies (Kao et al., 2008; Kopp et al., 2011; Kopp et al., 2008; Winkler et al., 2010). Assuming the general population risk allele frequency for the combined two *APOL1* risk alleles of at least 15% , at least 2.3% of the general population are expected to be homozygous or compound heterozygotes for the two *APOL1* risk alleles; and that among HIVAN cases 30% are risk allele homozyogtes (a conservative estimate), then a sample of 30 cases in the study group and 50 population controls would provide adequate power to detect a significant difference between cases and controls with an alpha = 5%, have 1-beta (power) of 97% with a two tailed test.

## 2.3 Participants

The study included 39 biopsies from HIV-positive individuals with HIVAN, 40 HIV-positive individuals with CKD other than HIVAN [HIV (+) CKD] and 41 HIV-negative individuals with CKD [HIV (-) CKD] as cases (Table 2.1). The study also included 54 controls from each group; HIV-positive individuals without CKD [HIV (+) Controls] and Population controls (PC) of black ethnicity.

**Table 2.1:** Case study participants

| Renal Pathology | Subtype | N |
|---|---|---|
| **HIVAN** | | 39 |
| **HIV (+) CKD** | | |
| | HIVICK | 13 |
| | FSGS | 13 |
| | Benign nephrolosclerosis | 1 |
| | Membranous glomerulonephritis | 8 |
| | Minimal change disease | 1 |
| | Membranoproliferative glomerulonephritis | 1 |
| | Membranoproliferative glomerulonephritis consistent with C3 glomerulopathy | 1 |
| | Global glomerulosclerosis | 1 |
| | Thrombotic microangiopathy | 1 |
| **HIV (-) CKD** | | |
| | FSGS | 9 |
| | Membranoproliferative glomerulonephritis | 2 |
| | Membranous glomerulonephritis | 12 |
| | IgA nephropathy | 1 |
| | Minimal change disease | 5 |
| | Lupus nephritis | 3 |
| | Global glomerulosclerosis | 2 |
| | Benign nephrosclerosis | 1 |
| | Thrombotic microangiopathy | 2 |
| | Oxalosis | 1 |
| | Severe ANS with secondary FSGS | 1 |
| | Nodular Glomerulosclerosis | 1 |
| | Proliferative glomerulonephropathy | 1 |
| **Total** | | **120** |

HIVAN; HIV associated nephropathy. HIVICK; HIV immune complex kidney disease. FSGS; Focal segmental glomerulosclerosis. ANS; arterial nephrosclerosis.

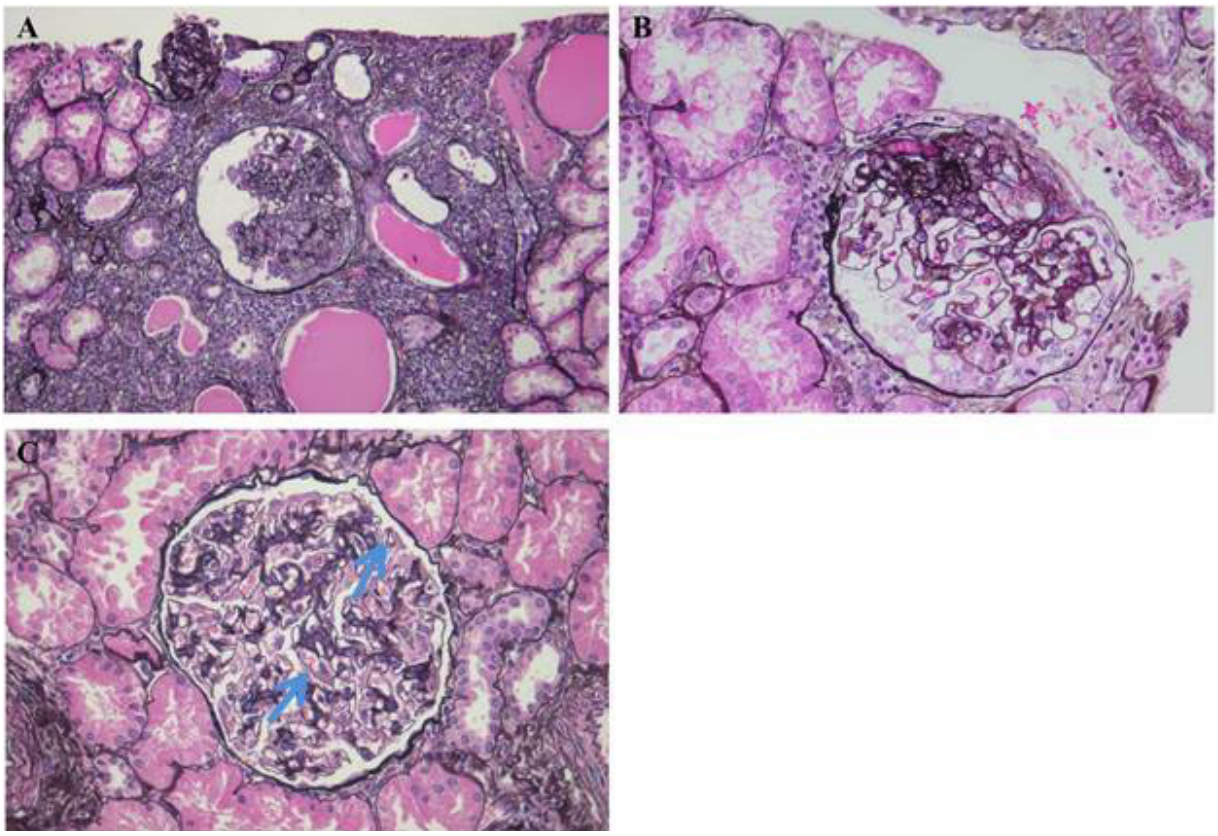



**Figure 2.1:** Light microscopy showing histological features of HIVAN, FSGS and HIVICK..
Collapsing glomerulopathy with microcytic dilatation of tubules and interstitial inflammation, characteristic of HIVAN (Silver methanamine X200); B. FSGS showing increased cellularity in cellular variant caused by endocapillary hypercellularity (inside the glomerular basement membrane) usually with foam cells (Silver methanamine X 400) (Thomas et al., 2006); C. Mild mesangial expansion and sub-epithelial immune complex deposits with a basement membrane reaction (blue arrows), characteristic of HIVICK (Silver methanamine X200). Photo micrograph courtesy of Department of Anatomical Pathology, University of the Witwatersrand).

### 2.3.1 Inclusion criteria

All study cases and controls had to be black South Africans of 18 years and above with no evidence of diabetes. Other inclusion criteria were specific for each group:

I. Group 1: HIV-positive, histologically showing collapsing glomerulopathy with microcystic tubular dilatation and mild interstitial inflammation (Figure 2.1A). [HIVAN].

II. Group 2; HIV-positive, histologically showing "non-HIVAN" histology such as FSGS (Figure 2.1B) and HIVICK (Figure 2.1 C). [HIV (+) CKD]. Table 2.1.

III. Group 3; HIV negative, with CKD of various histologies. [HIV (-) CKD]. Table 2.1.

IV. Group 4; HIV positive; with normal renal function tests (serum creatinine<100µmol/L and eGFR> 60mL/min/1.73m$^2$). [HIV (+) Controls].

V. Group 5; Population control; [PC].

### 2.3.2 Exclusion criteria

These included patients who were younger than 18 years; a history of diabetes; and non-black non-South Africans.

### 2.3.3 Variables

These included: age, gender, serum creatinine, viral load, CD4, eGFR and HIV status.

## 2.4 SNP Selection

Gene co-ordinates and intron/exon structure of each gene (*MYH9* and *APOL1*) was retrieved using the web based ENSEMBL genome browser (http://www.ensembl.org/index.html). A 1000 bps upstream and downstream for each gene was included and the longest transcript of each gene was used (Transcript: *MYH9-001*, Transcript ID: ENST 00000216181, Length of Transcript: 7501 bps and Transcript: *APOL1-003*, Transcript ID: ENST 00000319136, Length of Transcript: 3000 bps). The total number of SNPs present in and around each gene was then generated. The ENGINES SPSmart Program (http://spsmart.cesga.es/engines.php) was then used to determine the number of SNPs present in both International HapMap and 1000 Genome projects focusing on the three African populations; YRI, LWK and MKK (Table 1.3) as proxies for the black South African population since none of the black South African community had been sequenced nor genotyped in detail at the time the project began. SNPs with minor allele frequency (MAF) of equal to or greater than 5% in one of the three African populations mentioned above were selected. In addition, using the ENSEMBL genome browser, we used sorting intolerant from tolerant (SIFT) and polymorphism phenotyping (POLYPHEN) options to perform functional analysis in order to predict and map the potentially functional *MYH9* SNPs. Both SIFT and POLYPHEN use the homology of the sequence to predict missense mutations that could affect protein function and hence, potentially alter phenotype (Ng & Henikoff, 2001, 2002, 2003; Sim *et al.,* 2012). This has been applied to variant databases and was able to predict mutations involved in disease (Ng & Henikoff, 2002). In order to conform to the 96plex Illumina BeadXpress SNP genotyping system (Illumina, Inc.9885 Town Centre Drive San Diego, CA 92121 USA), we included previously published associated *MYH9* and *APOL1* SNPs including *APOL3* SNP rs11089781, as literature SNPs, having been previously reported to be associated with

kidney disease (Tzur *et al.,* 2010). Ancestry informative markers (AIMs) obtained from the department of Human Genetics, University of the Witwatersrand that had been previously used in population genetics studies and minor allele frequencies established were also included to assess whether our sample contained individuals with significant European admixture. These AIMs were obtained from HapMap (release 24). Ten of these SNPs were identified by Lao et al., 2006 as informative of geographic population structure and genetic ancestry while eight of those SNPs identified by Schlebusch et al., 2012 discriminate among southern African population (Lao *et al.,* 2006, Lombard *et al.,* 2012, Williams *et al.,* 2015 and Schlebusch *et al.,* 2012). We could not include tagSNPs as this could go beyond our budget of 96plex Illumina BeadXpress and also our candidate genes are not big and therefore the above selection process was best suited for our study. Figure 2.2 shows an illustration of the SNP selection process.

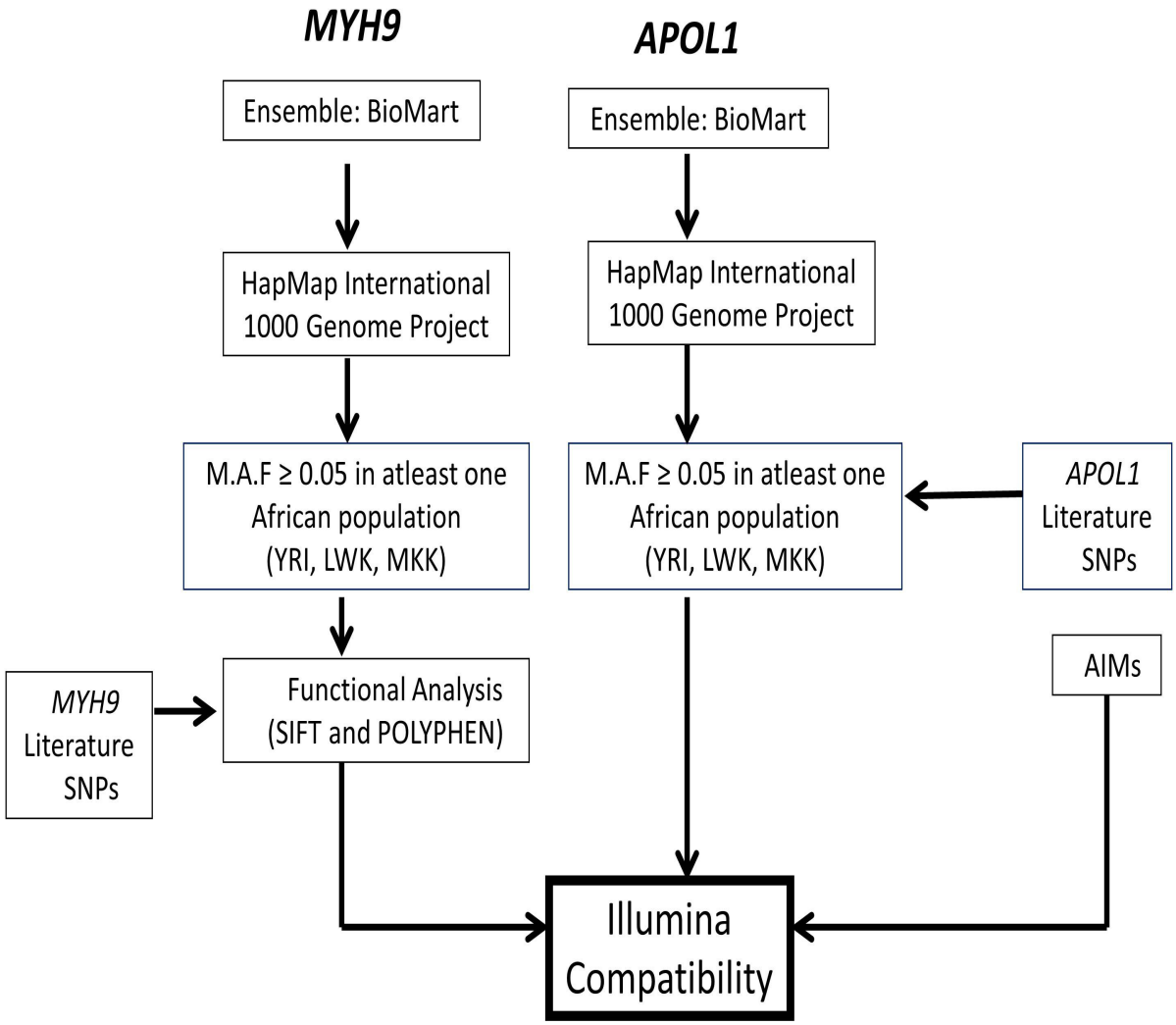



**Figure 2.2:** Summary of SNP selection process for Illumina assay compatibility
SIFT; sorting intolerant from tolerant. POLYPHEN; polymorphism phenotyping. AIMs; ancestry informative markers. M.A.F; minor allele frequency. YRI; Yoruba from Ibadan. LWK; Luhya from Western Kenya. MKK; Maasai from Kinyawa Kenya. Literature SNP were SNPs previously associated with kidney disease.

### 2.4.1 SNP scoring for compatibility with BeadXpress assay

The list of selected SNPs was sent to Illumina for scoring and assay design. All of the selected SNPs passed the scoring filter with a notable exception of the *APOL1* G1 (rs73885319 and rs60910145) and insertion/deletion variant (rs71785313). Custom TaqMan genotyping assay (ABI Foster city, California, USA) was used to genotype rs73885319, rs60910145 and rs71785313 as designed by Genovese *et al.* (Genovese *et al.*, 2010) and used by Kopp *et al.* (Kopp *et al.*, 2011).

## 2.5 Laboratory procedures

### 2.5.1 DNA extraction from Formalin Fixed Paraffin Embedded (FFPE) kidney tissues

Genomic DNA and RNA from kidney tissues were extracted using All Prep QIA amp DNA/RNA FFPE tissue kit (Qiagen Inc., Chatsworth, CA) according to the manufacturer's instructions. This kit is specifically designed for purifying DNA and RNA from FFPE tissue sections. The kit uses special lysis conditions to release DNA/RNA from tissue sections and to overcome inhibitory effects caused by formalin cross linking of nucleic acids. The kit also uses QiaAmp MinElute spin columns for purification of high quality DNA/RNA in small volumes (Appendix 4). This was done at the University of the Witwatersrand, Medical school (CMJAH) Molecular Biology Laboratory in the Department of Internal Medicine.

Excess paraffin was trimmed off the sample tissue using a scalpel. Immediately the sections were placed in 1.5ml microcentrifuge tubes and 320µl of deparafinization solution (Qiagen Inc., Chatsworth, CA) added to each tube containing one 20µm section of the sample. The lids were closed and the tubes vortexed vigorously for 10 seconds. This was followed by brief centrifugation at room temperature to bring the tissue to the bottom of the tube. The tubes were

then incubated at 56°C for 3 minutes and allowed to cool at room temperature. The tubes were then centrifuged at full speed (14000rpm) for 2 minutes. The supernatant was removed by pipetting without disturbing the pellet and any residual deparafinization solution was carefully removed. The lids were kept open and the tubes incubated for 37°C for 10 minutes to dry the pellet. The pellet was resuspended in 150µl of PKD (Proteinase K digest) buffer followed by the addition of 10µl Proteinase K solution and mixed by vortexing. Incubation at 56°C for 15 minutes followed and the tubes were placed on ice for 3 minutes to precipitate nucleic acids. The tubes were then centrifuged at 14000 rpm for 15 minutes. The supernatant was carefully removed and added to a different tube for RNA purification leaving the DNA containing pellet.

#### 2.5.1.1 Purification of genomic DNA

The DNA containing pellets were resuspended in 180µl of ATL buffer (Tissue lysis buffer) followed by 40µl of Proteinase K solution and mixed thoroughly by vortexing. The tubes were then incubated at 56°C for 1 hour followed by incubation at 90°C for 2 hours without agitating the tubes. There was an addition of 200µl of AL buffer added and tubes mixed thoroughly by vortexing followed by the addition of 200µl of absolute (96-100%) ethanol and mixed thoroughly again by vortexing. A white precipitate forms after addition of Buffer AL and ethanol and should be mixed immediately and thoroughly at this step in order to yield a homogenous solution. The entire sample was transferred to the QiaAmp minElute spin columns (supplied by the kit) placed in 2ml collection tubes, the lids closed and centrifuged for 1 minute at 10000 rpm. The collection tubes were then discarded with the flow through. If any of the samples did not completely pass through the membrane after centrifugation, a further centrifugation was done again at a higher speed until the QiaAmp minElute spin column was empty. QiaAmp minElute spin columns were put in new 2ml collection tubes (supplied by the kit) and 700µl of AW1

buffer (Wash buffer 1) added to the spin columns and centrifuged at 10000 rpm for 15 seconds to wash the spin column membranes and the flow through discarded. The collection tubes were put back in the QiaAmp minElute spin columns and 700µl of AW2 buffer (Wash buffer 2) was added to the spin columns and centrifuged at 10000 rpm for 15 seconds to wash the spin column membranes and the flow through discarded. The collection tubes were put back in the QiaAmp minElute spin columns and 700µl of absolute ethanol (96-100%) was added to the spin columns and centrifuged at 10000 rpm for 15 seconds and the collection tubes containing the flow through discarded. New 2ml collection tubes were placed to the QiaAmp minElute spin columns, the lids were opened and centrifuged at full speed (14000rpm) for 5 minutes. The collection tubes with the flow through were then discarded. The QiaAmp minElute spin columns were placed in new appropriately 1.5 ml microcentrifuge tubes, dated and labeled with sample ID followed by 50µl ATE buffer (Elution buffer) directly to the spin column membranes. The lids were gently closed, and samples incubated for 1 minute at room temperature. The DNA was finally eluted by centrifuging the tubes at full speed (14000 rpm) for 1 minute.

#### 2.5.1.2 Purification of total RNA

The RNA containing supernatants were incubated at 80°C for 15 minutes. This incubation step partially reverses formaldehyde modification of nucleic acids resulting in maximum RNA yields. The tubes were briefly centrifuged to remove drops from the inside of the lid and 320 µl of Buffer RLT was added to adjust binding conditions, and mixed by vortexing. 720 µl of absolute ethanol (96–100%) was added and mixed well by vortexing. 700 µl of the sample was transferred to an RNeasy MinElute spin column placed in a 2 ml collection tube. The lids were gently closed and centrifuged at 10,000 rpm for 15 seconds and the flow-through discarded. The collection tubes were put back in the RNeasy minElute spin columns and this step repeated until the entire

sample was passed through the RNeasy MinElute spin column. The collection tubes were re-used for the next step. 350 µl Buffer FRN was then added to the RNeasy MinElute spin column. The collection tubes were re-used for the next step. The lids were gently closed, and centrifuged at10, 000 rpm for 15 seconds. The flow-through was then discarded. The collection tubes were re-used for the next step. 10 µl of DNase I stock solution was added to 70 µl Buffer RDD, mixed gently by inverting the tubes, and centrifuged briefly to collect residual liquid from the sides of the tubes. 80 µl of the DNase I incubation mix was added directly to the RNeasy MinElute spin column membrane, and placed on the benchtop (20-30°C) for 15 minutes. 500 µl of Buffer FRN was added to the RNeasy MinElute spin column, the lids closed gently and centrifuged for 15 seconds at 10,000 rpm. The flow-through was saved for use in the next step since it contained RNA including small RNAs. The RNeasy MinElute spin column were placed in new 2 ml collection tubes and the flow-through from the previous step applied to the spin column. The lids were gently closed and centrifuged for 15 seconds at 10,000 rpm. The flow through was then discarded. The collection tubes were re-used for the next step. 500 µl Buffer RPE was then added to the RNeasy MinElute spin column. The lids were gently closed and centrifuged for 15 seconds at 10,000rpm to wash the spin column membrane. The flow through was then discarded. The collection tubes were re-used for the next step. This was followed by 500 µl Buffer RPE onto the RNeasy MinElute spin column. The lids were gently closed and centrifuged for 15 seconds at 10,000rpm to wash the spin column membrane. The flow through was then discarded. RNeasy MinElute spin columns were then placed in new 2ml collection tubes, the lids opened, and centrifuged at full speed (14,000 rpm) for 5 minutes. The flow through was then discarded.

To avoid damage to their lids, the spin columns were placed into the centrifuge with at least one empty position between columns and orientated the lids so that they point in a direction opposite

to the rotation of the rotor. The RNeasy MinElute spin columns were then placed in new 1.5 ml collection tubes and 30 µl of RNase-free water was added directly to the spin column membrane. The lids were gently closed, and incubated for 1 minute at room temperature (15-25°C) followed by centrifugation at full speed (14,000 rpm) for 1 minute to elute the RNA.

The DNA and RNA concentrations and purity were checked by spectrophotometry using the NanoDrop (Thermo Scientific, 3411 Silverside Road Bancroft Buliding Wilmington, DE 19810 USA).

The eluted genomic DNA and RNA were stored at -70$^{o}$C until they were required for the purposes of this study.

### 2.5.2 DNA Extraction from Human Nucleated Cells in blood samples

Salting out method was used for high yield DNA extraction from whole blood. This was done at the University of the Witwatersrand, Medical School (CMJAH) Molecular Biology Laboratory in the Department of Internal Medicine. The procedure was modified from Nasiri *et al.*(Nasiri *et al.*, 2005 & Miller *et al.,* 1988) and is described in detail in Appendix 5.

Lysis buffer dissolved the cell membrane while Proteinase K digests proteins, while chelating agents like 10% Sodium dodecyl sulphate (SDS) bind to bivalent cations of nuclease cofactors and thereby prevent DNA degradation by these enzymes. Finally, DNA was precipitated using absolute (96-100%) ethanol and dissolved in 1X Tris-EDTA (TE) buffer (Nasiri *et al.*, 2005).

The DNA concentration and purity were checked by spectrophotometry using the NanoDrop (Thermo Scientific, 3411 Silverside Road Bancroft Buliding Wilmington, DE 19810 USA). The eluted genomic DNA was stored at -70$^{o}$C until it was required for the purposes of this study.

### 2.5.3 Custom TaqMan SNP Genotyping Assay

All genomic DNA samples were genotyped for SNPs rs73885319, rs60910145 and rs71785313 using the TaqMan® minor groove-binding (MGB) based allelic discrimination (AD) assay (Applied Biosystems). TaqMan® AD Assay is a multiplexed, robust and highly reproducible assay. It utilizes the 5' to 3' Taq DNA polymerase exonuclease activity (Livak, 1999). A unique pair of fluorescent dye detectors is used for each sample in an AD assay, for instance, two allele-specific TaqMan® probes labelled with a different 5' fluorophore dye (either VIC or FAM dye) and a 3' quencher dye together with a primer pair to target a SNP site (Figure 2.3). The fluorophore's signal using Förster fluorescence resonance energy transfer (FRET) is eliminated by the close proximity of the dyes. In order to improve the hybridization process, the DNA double helix is bound onto the MGB molecule hence stabilizing the MGB-probe/template complex. Detection is determined by the exonuclease cleavage of a 5' allele-specific dye label that generates a permanent assay signal. All the MGB probes include a non-fluorescent quencher (NQF). This eliminates the background fluorescence providing a greater signal-to-noise ratio for hence improving assay sensitivity (Kamau et al., 2012). One fluorescent dye detector represents a match to the wild type (allele 1) and the other fluorescent dye detector represents a match to the mutation (allele 2). The presence of two fluorescent dye detectors (primer/probe) pairs in each reaction allows genotyping of the two variants at that particular SNP site. The AD assay therefore classifies the unknown samples as; Homozygotes (samples having only allele 1 or allele 2) and Heterozygotes (samples having both allele 1 and allele 2).

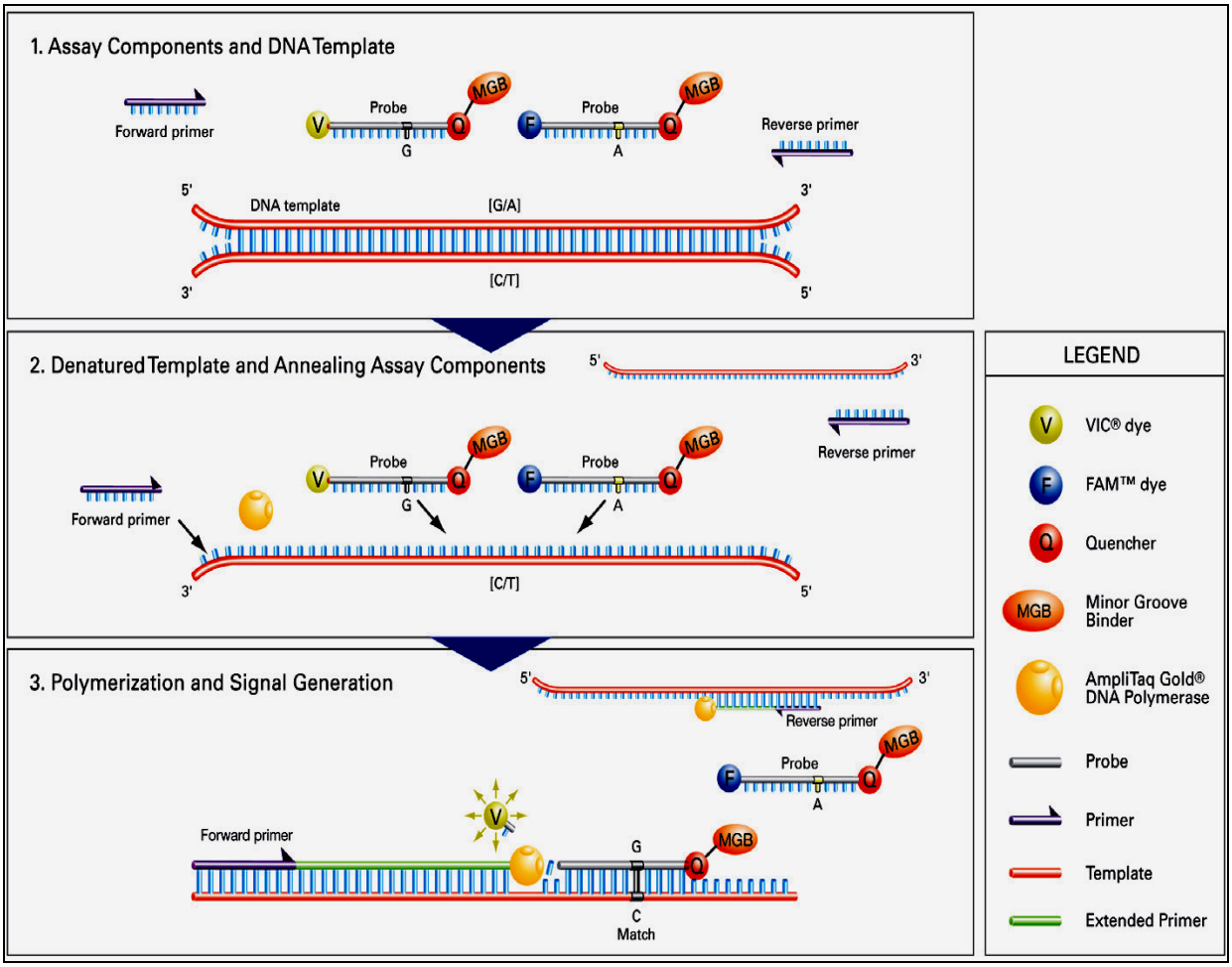



**Figure 2.3:** Schematic illustration of TaqMan Allelic Discrimination Assay.
Achieved by the selective annealing of TaqMan MGB probes. Source: www.Appliedbiosystems.com

#### 2.5.3.1 Materials

##### 2.5.3.1.1 Genomic DNA

Each DNA sample was diluted (normalized) to 20ng/µl and 11.25µl per well was aliquoted into 96-well reaction plates. The following controls were also included on each plate:

- Known genotype: this included three controls with known genotypes
- No template control: this did not contain DNA template hence fluorescence levels were due to the uncleaved probe

##### 2.5.3.1.2 SNP genotyping assays

The SNPs (rs73885319, rs60910145 and rs71785313) genotyping assay were supplied as 40X assay and contained:

- 5' VIC® dye and a 3' quencher (tetramethylrhodamine) labelled to a TaqMan® probe to detect the allele 1 sequence
- 5' FAM® (6-carboxyl-fluorescein) dye and a 3' quencher (tetramethylrhodamine) labelled to a TaqMan® probe to detect the allele 2 sequence
- Forward and reverse primers to amplify the sequences of interest that represents the polymorphisms

The SNP genotyping assays were diluted to 20X using 1X TE buffer before preparing the Master Mix.

##### 2.5.3.1.3 Universal PCR Master Mix

This is a 2X single reaction buffer that contains (AmpliTaq Gold™ DNA polymerase, reaction buffer and dNTPs) for the TaqMan® assay. It facilitates the synthesis of a new DNA strand from the dNTPs using single-stranded DNA after denaturation. The annealed primers are used for initiation of synthesis.

#### 2.5.3.2 Procedure

##### 2.5.3.2.1 Master Mix preparation

The total master mix for each assay, was prepared in a sterile centrifuge tube, mixed well by vortexing gently and centrifuged to settle the components in the tube. The volume of each component needed for per well of the 96-well plate reaction plate are shown in Table 2.2.

**Table 2.2:** TaqMan® master mix components

| **PCR Components** | **Volume per well (µl)** |
|---|---|
| (2X) TaqMan® Universal PCR Master Mix | 12.50 |
| 20X SNP Genotyping Assay Mix | 1.25 |
| **TOTAL** | **13.75** |

A template for a single sample used to calculate the reaction volumes. In our calculations we scaled-up master mix reactions accordingly.

##### 2.5.3.2.2 Reaction plate preparation

For the SNP assays, 13.75µl of the master mix was pipetted into each well of the 96-well plate. This was followed by 11.25µl of DNA to make a total reaction volume of 25.0 µl. All wells were inspected for uniformity of volume and the plate sealed with an adhesive cover. It was vortexed to mix the reactions and centrifuged briefly to bring down the contents and eliminate air bubbles. The plates were then placed in the Applied Biosystems 7500 Fast Real-Time PCR System specifying a TaqMan® AD assay. The thermocycling conditions are shown in Table 2.3 below;

**Table 2.3:** Thermocycling conditions for the TaqMan® assay

| **AmpliTaq Gold™ Enzyme activation** | **(40 cycles)** | |
|---|---|---|
| HOLD | Denature | Anneal/Extend |
| 10 minutes at 95°C | 15 seconds at 92°C | 1 minute at 60°C |

##### 2.5.3.2.3 Allelic discrimination plate read

Sample IDs were entered in the template on the machine computer and reaction volume per well (25.0µl) specified. Applied Biosystems 7500 Fast Real-Time PCR System Sequence Detection System Software measured the intensity of the fluorescent signal. The plotted fluorescent signals

were used to call the genotypes for each sample. An increase in signal of either FAM® or VIC® indicated homozygous for the FAM®-specific or VIC®-specific allele while an increase in both signals indicated heterozygous for both FAM® and VIC®.

### 2.5.4 TaqMan Quality Control

The DNA samples tested for SNPs rs73885319, rs60910145 and rs71785313 were tested in duplicate and sequencing of selected samples was used for confirmation of the genotypes. Further validation of the genotypes was conducted in Dr. Cheryl Winkler's lab at NIH, NCI Molecular genetics unit in Frederick, Maryland, USA (our collaborator). The genotype frequencies were also tested for deviation from Hardy Weinberg Equilibrium (HWE).

### 2.5.5 Custom Illumina® GoldenGate BeadXpress SNP Genotyping Assay

Illumina® is a flexible pre-optimized assay that provides the ability to design custom panels for high-throughput, two-color laser detection bioassay system that allows you to scan large numbers of VeraCode assays in a wide range of multiplex levels. This assay targets specific SNPs in genomic DNA and it is based on sequence-sequence extension and ligation of hybridized oligonucleotides distinguished by their shared primers sites (Fan *et al.*, 2006). This is illustrated in Figure 2.4.

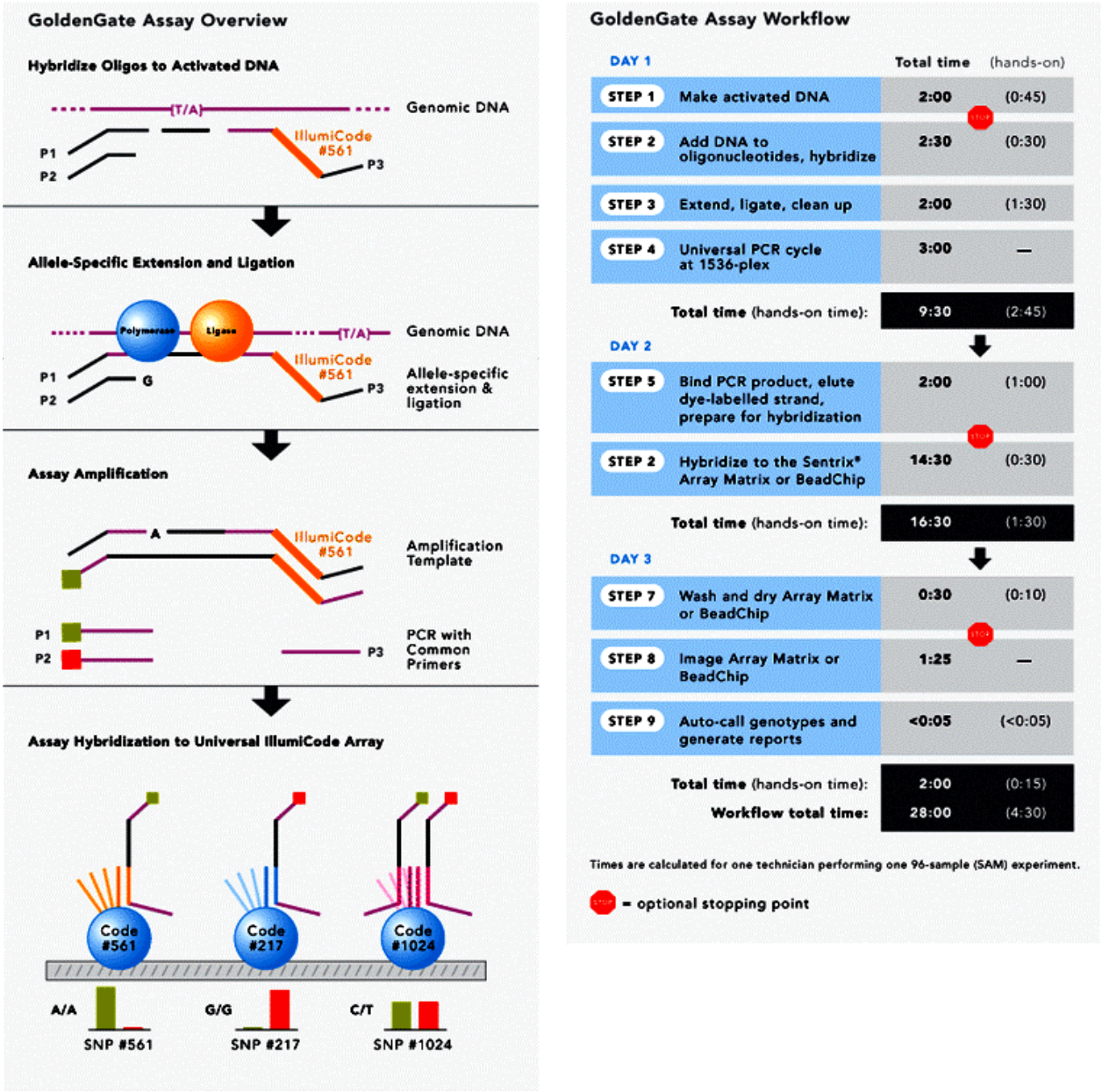



**Figure 2.4:** Golden Gate BeadXpress overview and workflow assay.
This assay uses allele-specific extension and ligation between annealed locus-specific oligonucleotides (LSOs) and allele-specific oligonucleotides (ASOs) for genotype calling, enabled by an enzymatic discrimination step. Subsequent PCR amplification step is used to amplify the appropriate product using "built in" primers in the ASO and LSO sequence sites. The products are then hybridized and read out on an array of universal-capture illumicodes (Fan *et al.,* 2006).

All genomic DNA samples were normalized to a concentration of 50ng/µl and genotyped for 96 SNPs that passed SNP scoring using the BeadXpress SNP Genotyping system (Illumina, Inc.9885 Town Centre Drive San Diego, CA 92121 USA) according to the manufacturer's instructions. Each normalized genomic DNA (5µl) sample was hybridized simultaneously to a pool of primers (OPA) attached to beads in solution and after allele specific primer extension and ligation, followed by the addition of a set of fluorescently labeled universal primers (Cy3 and Cy5 labeled P1 and P2 respectively), PCR was carried out using conditions shown in Table 2.4, generating multiple labeled amplicons representing different SNPs.

**Table 2.4:** Thermocycling conditions for the Illumina® GoldenGate BeadXpress assay

| | Temperature | Time |
|---|---|---|
| **Denaturation** | 37°C | 10 minutes |
| | 95°C | 3 minutes |
| **PCR (34 cycles)** | 95°C | 35 seconds |
| | 56°C | 35 seconds |
| | 72°C | 2 minutes |
| **Final Extension** | 72°C | 10 minutes |
| **Hold** | 4°C | 5 minutes |

PCR amplicons containing strands with fluorescent signals were isolated and hybridized to VeraCode beads (rod-shaped glass microbeads measuring ~240 x 28µm). These VeraCode beads are inscribed digitally with novel barcodes, with their surface available for attachment of an oligonucleotide complimentary to the specific sequence on the beads. After hybridization, the VeraCode beads were washed and scanned on the BeadXpress Reader calling alleles for each genotype based on the identity of the VeraCode bead and the fluorescence label.

Appendix 6 contains the detailed procedure for Illumina® GoldenGate Assay.

#### 2.5.5.1 BeadXpress Quality control

This was performed using the BeadStudio software package used to cluster the fluorescence data and generate genotypes for each sample and each SNP. Assay performance was first assessed using the inbuilt sample-dependent, sample independent and contamination controls. These assessed the overall performance of the samples, reagents, equipment, and BeadChips. The inbuilt controls used were:

I. *Allele Specific Extension* tests the extension efficiency of properly matched versus mismatched allele specific oligos (ASO). Failures in this control could indicate either a processing failure at the Make ASE or Make MEL steps or poor DNA sample quality. Samples with a theta between 0.2 – 0.8 were excluded.

II. *PCR Uniformity* tests the PCR amplification efficiency for high AT and high GC regions of DNA. Low signal intensities could indicate either a problem with the PCR amplification process or low sample quality. Samples with a theta between 0.2 – 0.8 were excluded. *Extension Gap* tests the efficiency of the 15 base extensions from the 3' end of the ASO to the 5' end of the locus specific oligos (LSO). Processing the Make MEL step or poor DNA sample quality could impact the behaviour of these controls. Samples with R < 1000 were excluded.

III. *First Hybridization* tests for the specificity of the annealing ASO with different melting temperatures (Tm) to the different DNA loci. Observed effects to signal intensity or clustering could indicate either a general processing failure at Make ASE step or poor DNA sample quality. Samples with R < 1000 were excluded.

IV. *Second Hybridization* tests the specificity of synthetic oligo targets to probes present on the beads. Issues related to hybridization of samples to BeadChips or VeraCode beads could lead to unexpected behaviour of these controls. Samples with R < 1000 were excluded.

Contamination dashboard was viewed to make sure that one colour was displayed meaning there was no contamination of the samples.

GC 10% which is a measure of the fluorescence intensity and distance of the call to the centre of the cluster was performed and GC below 0.15 was considered poor and therefore samples falling in that category were excluded.

All samples with a low call rate of less than 0.95 and low 10% GC were excluded. Samples with poor genotyping performance were highlighted before evaluating the quality of SNP clusters. The GenCall (GC) score which is a quality metric ranging from 0-1 was calculated for each genotype. The sample data points that appeared to be further from the center of the cluster had low GenCall score and were therefore excluded. In order to identify SNPs that need to be manually edited or zeroed (excluded), we evaluated newly reclustered SNPs using the metrics listed in the "SNP Table". These metrics provide the overall performance information for each SNP. Each SNP was evaluated based on the angle, dispersion, and overlap of clusters and intensities. In order to find SNPs that might have needed editing or exclusion, we started by determining hard cut offs and grey zones. We sorted data one column at a time, exploring values at the extremes of the ranges. Hard cut offs levels below and above including the grey zones were defined to contain loci with 80-90% successful and were improved by manual editing. Otherwise, all the other SNPs were zeroed and excluded from the final analytical data set.

Cluster separation was used to evaluate individual SNPs for overlapping clusters. SNPs with overlapping clusters were zeroed and excluded from the project while SNPs which were well separated, were manually edited. The AB mean for intensity (R) - ABR mean, was used to identify all SNPs with intensities too low for genotypes to be called reliably. All those with low intensities were zeroed. AB mean for theta – ABT mean, ranging from 0-0.2 and 1-0.8 were evaluated in order to identify SNPs where the heterozygote cluster shifted towards the homozygote cluster. SNPs falling between the two clusters were not reliably separated, were zeroed and excluded from the analysis. A call rate and call frequency of 95% was used on all the SNPs that passed quality control. During manual editing, any changes made were consistent with principles of population genetics in order to prevent introduction of subjective bias into the data set and make sure that we remained with data of the highest quality. We also excluded SNPs that did not respect the Hardy-Weinberg equilibrium at $p<10^{-3}$. The final genotyping data were then exported from Bead Studio software using the "Final Report Wizard" for further downstream analysis.

### 2.5.6 Estimated Glomerular filtration rate (eGFR)

This was calculated using the simplified Modification of Diet in Renal Disease (MDRD) equation: (eGFR (ml/min/1.73 $m^2$) = 186× (serum creatinine) $^{-1.154}$×age$^{-0.203}$ (×0.742 if female) (×1.210 if black) (Foundation, 2002; Levey *et al.,* 1999).

## 2.6 Population stratification

Population substructure (stratification) was assessed based on a total of 42 SNPs including the AIMs SNPs and the selected chromosome 22 SNPs using the Principal Components Analysis (PCA) module of *Eigensoft* software (Price *et al.,* 2006). For these analyses, chromosome 22

SNPs were limited to $r^2 < 0.05$ with each other, and *APOL1* risk allele SNPs were excluded from the population stratification analyses in order to prove that the strong association seen with *APOL1* SNPs is not as a result of population substructure but due to the *APOL1* high risk alleles..

## 2.7 Messenger RNA (mRNA) expression profile of the N-terminal apolipoprotein L1 isoforms

Reverse Transcription-Polymerase Chain Reaction (RT-PCR) was used to profile the expression patterns of mRNA N-terminal apolipoprotein L1 isoforms. RT-PCR is a sensitive method for detection of transcript expression in total RNA extracted from limited tissue samples such as biopsy material (Marone *et al.,* 2001). This method was adopted over quantitative RT-PCR due to the difficulty in designing a PCR strategy that could accommodate the small size difference and the extent of overlap between the N-terminals. We profiled specific apolipoprotein L1 mRNA transcripts/isoforms of various kidney disease types as represented in (Table 2.5). To ensure standardization of results all the samples were processed at the same time under the same conditions.

**Table 2.5**: Histology diagnoses, HIV status and number of risk alleles of samples selected for mRNA N-*ApoL1* isoform profiling

| | Diagnosis | | HIV status | Genotypes | No. of risk alleles |
|---|---|---|---|---|---|
| **1** | HIVAN | **CKD Cases** | HIV positive | G1/G1 | 2 |
| **2** | HIVAN | | HIV positive | G1/G2 | 2 |
| **3** | HIVAN | | HIV positive | G0/G1 | 1 |
| **4** | HIVAN | | HIV positive | G2/G2 | 2 |
| **5** | HIVAN | | HIV positive | G0/G0 | 0 |
| **6** | HIVAN | | HIV positive | G0/G0 | 0 |
| **1** | FSGS | | HIV positive | G2/G2 | 2 |
| **2** | FSGS | | HIV positive | G0/G2 | 1 |
| **3** | FSGS | | HIV positive | G0/G0 | 0 |
| **4** | FSGS | | HIV positive | G0/G0 | 0 |
| **5** | FSGS | | HIV positive | G0/G0 | 0 |
| **6** | FSGS | | HIV positive | G0/G0 | 0 |
| **1** | HIVICK | | HIV positive | G1/G2 | 2 |
| **2** | HIVICK | | HIV positive | G0/G0 | 0 |
| **3** | HIVICK | | HIV positive | G0/G1 | 1 |
| **4** | HIVICK | | HIV positive | G0/G0 | 0 |
| **5** | HIVICK | | HIV positive | G0/G0 | 0 |
| **6** | HIVICK | | HIV positive | G0/G2 | 1 |
| **1** | Membranous glomerulonephritis | | HIV positive | G0/G2 | 1 |
| **2** | Global glomerulosclerosis | | HIV positive | G1/G1 | 2 |
| **3** | Membranous glomerulonephritis | | HIV positive | G0/G0 | 0 |
| **4** | Membranous glomerulonephritis | | HIV positive | G0/G0 | 0 |
| **5** | Thrombotic microangiopathy | | HIV positive | G0/G0 | 0 |
| **6** | Minimal change | | HIV positive | G0/G0 | 0 |
| **1** | FSGS | | HIV negative | G1/G2 | 2 |
| **2** | FSGS | | HIV negative | G0/G0 | 0 |
| **3** | FSGS | | HIV negative | G0/G2 | 1 |
| **4** | FSGS | | HIV negative | G0/G1 | 1 |
| **5** | FSGS | | HIV negative | G0/G0 | 0 |
| **6** | FSGS | | HIV negative | G0/G0 | 0 |
| **1** | Minimal change | | HIV negative | G0/G0 | 0 |
| **2** | Benign Nephrolesclerosis | | HIV negative | G0/G0 | 0 |
| **3** | Membranous glomerulonephritis | | HIV negative | G0/G1 | 1 |
| **4** | Nodular glomerulosclerosis | | HIV negative | G0/G0 | 0 |
| **5** | Membranous glomerulonephritis | | HIV negative | G0/G2 | 1 |
| **6** | Minimal change | | HIV negative | G0/G0 | 0 |
| **1** | Control 1 (Normal kidney) | **Controls** | HIV negative | G0/G0 | 0 |
| **2** | Control 2 (Normal kidney) | | HIV negative | G0/G0 | 0 |
| **3** | 1740 (Normal kidney) | | HIV negative | G0/G0 | 0 |
| **4** | 2389 (Normal kidney) | | HIV negative | G0/G0 | 0 |
| **5** | 8279 (Normal kidney) | | HIV negative | G0/G0 | 0 |

HIVAN; HIV associated nephropathy. HIVICK; HIV immune complex kidney disease. FSGS; Focal segmental glomerulosclerosis.

The controls used in this experiment were from nephrectomy kidney samples archived at Department of Anatomical Pathology, University of Witwatersrand. The normal histology from these kidney tissues was confirmed by a pathologist.

### 2.7.1 Primer design for *APOL1* gene N-terminal domain

*APOL1* gene primers were designed specifically for the *APOL1* gene transcripts that hybridize to homologous sequence in the N-terminal region of the gene (Figures 1.13 and 2.5). Messenger RNA sequences for the apolipoprotein L1 isoforms (A, B and C) were retrieved from NCBI website (http://www.ncbi.nlm.nih.gov/nuccore/NG_023228.1?report=genbank&from=5001&to=19461) and primers were designed at the N-terminal region of the gene using PubMed primer-blast tool (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) synthesized by Integrated DNA Technologies (IDT) (Whitehead Scientific, Cape Town). GAPDH primers were also designed and used as a reference control gene. The sequences of the primers are shown in Table 2.6**.**

**Table 2.6**: Primer sequences for N-terminal *APOL1* gene

| **Primer** | **Sequence** | **Melting temperature ($^{o}$C)** |
|---|---|---|
| *APOL1* Forward Isoforms A/C | 5'-GGAGCTGCTTTGCTGAGAGT-3' | 59.90 |
| *APOL1* Reverse Isoforms A/C/B | 5'-AGTCACCGAGGGGCTTACTT-3' | 60.13 |
| *APOL1* Forward Isoform B | 5'-ATGAGATTCAAAAGCCACACTG-3' | 59.24 |
| GAPDH Forward | 5'-TGCACCACCAACTGCTTAG-3' | 57.98 |
| GAPDH Reverse | 5'-GGATGCAGGGATGATGTTC-3' | 55.73 |

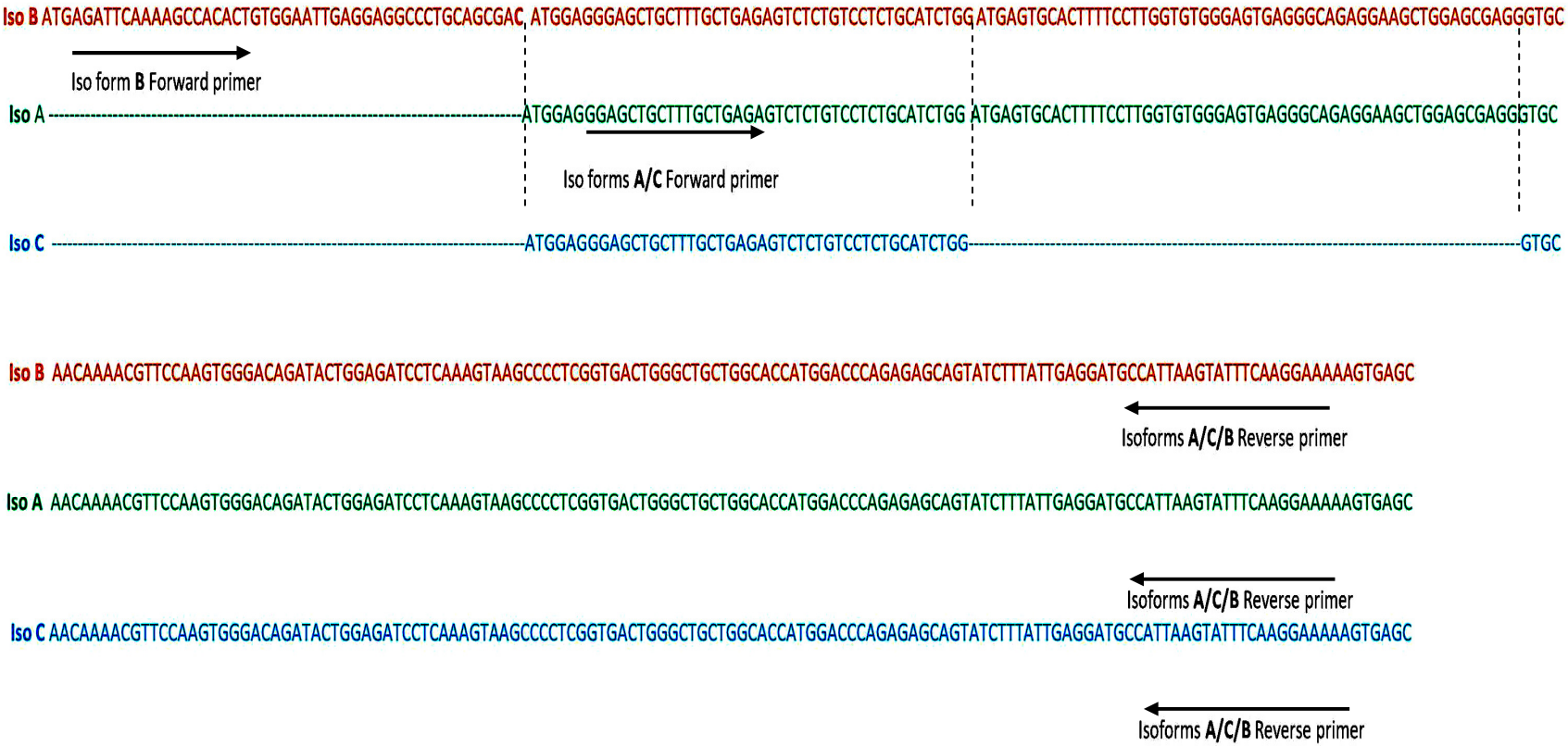



**Figure 2.5**: N-terminal sequence of *APOL1* isoforms A, B and C.
Isoform B, NCBI Reference Sequence: NM_145343 (Red), has an alternate in-frame exon in the 5' coding region and uses an upstream in-frame start codon compared to Isoform A, NCBI Reference Sequence: NM_003661 (Green) hence Isoform B has a longer N-terminus compared to isoform A. Isoform C, NCBI Reference Sequence: NM_001136541 (Blue) lacks an alternate in-frame 5'coding segment, compared to Isoform A hence resulting in a protein that has a shorter N-terminus as compared to isoforms A and B. Source:
(http://www.ncbi.nlm.nih.gov/nuccore/NG_023228.1?report=genbank&from=5001&to=19461). Accessed in July 2014.

### 2.7.2 cDNA synthesis

To measure the mRNA expression of N-terminal apolipoprotein L1 isoforms in kidney tissues, a total of 6 RNA samples previously isolated according to manufacturer's instructions [The All Prep QIA amp DNA/RNA FFPE tissue kit (Qiagen Inc., Chatsworth, CA) as described in 2.5.1.2 and Appendix 4] from FFPE kidney biopsies of each kidney disease type [HIVAN, HIV (+) FSGS, HIVICK, HIV (-) FSGS, HIV (+) other kidney diseases, HIV (-) other kidney diseases and normal kidney] (Table 2.5) were used as templates for cDNA synthesis using random primers according to the template seen in Table 2.7. We used a high capacity cDNA reverse transcription kit (ABI Biosystems Foster city, CA 94404).

**Table 2.7:** RT-PCR master mix

| **RT-PCR Components (Applied Biosystems)** | **Volume per reaction (µl)** |
|---|---|
| (10X) RT Buffer | 2.0 |
| 100µM dNTP mix | 0.8 |
| (10X) Random primers | 2.0 |
| Reverse Transcriptase (50U/µl) | 1.0 |
| RNAse free water | 13.2 |
| RNA template (100ng) | 1.0 |
| **TOTAL** | **20.0** |

This template is for a single sample and was scaled up for the total number of samples in the experiment.

The reaction tubes containing the RT-PCR reaction components were then placed in the thermocycler, MJ mini™ Gradient Thermocycler (BioRad laboratories, 1000 Alfred Noble drive Hercules, CA 94547, USA) under the following conditions (Table 2.8) in order to synthesize cDNA.

**Table 2.8:** Thermocycler conditions for RT-PCR reaction synthesis of cDNA

| **Temperature** | **Time** |
|---|---|
| 25°C | 10 minutes |
| 37°C | 2 hours |
| 85°C | 5 minutes |
| 4°C | Hold (∞) |

### 2.7.3 Polymerase Chain Reaction

Polymerase chain reaction (PCR) was performed using *APOL1* gene-specific primers designed as described in Table 2.6 and the reactions set up according to the template represented in Table 2.9.

**Table 2.9**: PCR master mix components

| **PCR Components** | **Volume per reaction (µl)** |
|---|---|
| (2X) KAPA 2G robust HotStart Readymix | 6.250 |
| *APOL1* F primer (10µM) | 0.625 |
| *APOL1* R primer (10µM) | 0.625 |
| RNAse free water | 3.000 |
| cDNA (20ng) | 2.000 |
| **TOTAL** | **12.5** |

This template is for a single sample and was scaled up for the total number of samples in the experiment.

The reaction tubes containing the PCR components were then placed in the thermocycler, MJ mini™ Gradient Thermocycler (BioRad laboratories, 1000 Alfred Noble drive Hercules, CA 94547, USA) under the following conditions (Table 2.10) for PCR to take place.

**Table 2.10:** Thermocycler conditions

| | **Temperature** | **Time** |
|---|---|---|
| **Denaturation** | 95°C | 3 minutes |
| **PCR (40 cycles)** | 95°C | 15 seconds |
| | 62°C | 15 seconds |
| | 72°C | 15 seconds |
| **Final Extension** | 72°C | 1 minute |
| **Hold** | 4°C | ∞ |

10µl of the amplified PCR products were run on 5% acrylamide gels (set up as described in Appendix 7 A) and the Gene Ruler Low Range DNA Ladder, Ready-to-Use 25-700 bp (ThermoScientific™, Bishop Meadow Road, Loughborough, Leicestershire, LE11 5RG, United Kingdom) used as the molecular weight marker. The gels were then stained in ethidium bromide solution (3µl of ethidium bromide in 100mL 1X TBE solution) for 5 minutes followed by viewing under ultra violet (UV) light using Gel Doc™ EZ Imaging system (BioRad laboratories,

1000 Alfred Noble drive Hercules, CA 94547, USA). Images of the gels were captured and the different isoforms identified based on their different molecular weights. ImageJ v1.48 (http://rsb.info.nih.gov/ij/docs/menus/analyze.html#gels) was used to determine the intensity of bands and a student t test was used to determine the correlation between N-*ApoL1* mRNA expression in cases and controls.

### 2.7.4 ImageJ

ImageJ is a public domain Java image processing program developed by the National Institutes of Health Image. It can be used as either an online or downloadable application. This tool was used to compare the density or intensity of bands captured from the acrylamide gels. The file images were opened using ***File>Open*** in ImageJ. The **Rectangular Selections** tool was chosen from the ImageJ toolbar and a rectangle of an area of 0.032 drawn next to each band to establish the background of the individual bands. Measurements were taken using ***Analyze>Measure*** measurements.

The same size rectangle was drawn around each PCR fragment band and band intensities measurements of each band recorded using ***Analyze>Measure***. The area of the rectangle was retained throughout the measurements of all the bands on acrylamide gels. The background measurement values were subtracted from the band measurement values to give the actual band intensity values of each band. These values were then used in the subsequent analysis.

## 2.8 Developing an expression system for Apolipoprotein L1 isoform expression in kidney cells

Proteins are transported to appropriate destinations in the cell or out of the cell depending on the sequence of the N-terminal domain. Proteins can be targeted to organelles, different

intracellular membranes, the plasma membrane, or for secretion outside of the cell through specific targeting signals or signal peptides. To elucidate the effect of the N-terminal domain on apolipoprotein L1 isoform trafficking, a series of fluorescently labeled *ApoL1* N-terminal isoform (A, B and C) constructs were created to be utilized in confocal microscopy. A two-step PCR-based cloning strategy was adopted to create fluorescently tagged *ApoL1* N-terminal isoform constructs.

### 2.8.1 Cloning of *ApoL1* N-terminal isoforms into *pJET 1.2/blunt* vector

*pJET1.2/blunt* is a linearized cloning vector that accommodates inserts from 6bp to 10kb. This vector also contains a lethal gene that is disrupted by ligation of a DNA insert into the multiple cloning site (MCS), hence only cells with recombinant plasmids are able to propagate for easier positive screening of colonies and eliminate expensive blue/white screening.

Isoform A, B and C products previously obtained from the mRNA isoform expression analysis in **2.7** were ligated into the general donor plasmid, *pJET1.2/blunt* expression vector (2974bp) according to manufacturer's instructions (ThermoScientific$^{TM}$, USA). The plasmid map and features of the *pJET 1.2/blunt* cloning vector is shown in Appendix 8. Since the KAPA 2G robust HotStart Ready-mix used in RT-PCR of *APOL1* mRNA isoform expression gave rise to PCR products with 3'-dA overhangs, the sticky-end cloning protocol of the *pJET 1.2/blunt* cloning kit (ThermoScientific$^{TM}$, USA) was used. A blunting reaction was set up according to the reaction conditions below (Table 2.11).

**Table 2.11:** Blunting reaction.

| Components | Volume per reaction (µl) |
|---|---|
| 2X Reaction Buffer | 10.0 |
| Non-purified PCR product/sticky-end fragment | 1.0 |
| Water, nuclease-free water | Up to 17.0 |
| DNA blunting Enzyme | 1.0 |
| **TOTAL** | **18.0** |

A template for a single sample used to calculate the reaction volumes. The reaction volume was scaled up for the total number of blunting reactions required.

The above mixture was vortexed and centrifuged briefly followed by incubation at 70$^{o}$C for 5 minutes and chilled on ice thereafter.

The following ligation reaction (Table 2.12) was added to the blunting reaction mixture above in order to ligate the PCR products onto the *pJET1.2/blunt* vector. This brought the total volume to 20.0µl.

**Table 2.12:** Ligation reaction

| Components | Volume per reaction (µl) |
|---|---|
| *pJET 1.2/blunt* Cloning Vector (50ng/µl) | 1.0 |
| T4 DNA Ligase (5U/µl) | 1.0 |
| **TOTAL** | **20.0** |

The reaction volume for one reaction is represented.

The contents of the tubes were then vortexed and centrifuged briefly to collect drops on the top and sides of the tubes. The tubes containing the ligation mixture were then incubated at room temperature for 15 minutes and then transformed into *Escherichia coli* (*E. coli)* DH5α competent cells.

The DH5α competent cells were prepared according to the protocol described in Appendix 9.

#### 2.8.1.1 Transformation into already prepared E. coli DH5α competent cells

50 µl of freshly prepared competent cells were thawed on ice and placed in chilled 1.5ml microcentrifuge tubes and 4 µl of ligation mixtures added, mixed gently and incubated on ice for 30 minutes. The tubes were then heat shocked by shifting the chilled mixture to 42°C water bath

for 2 minutes, and then chilled on ice for further 2 minutes. To this 950 µl of Super optimum culture (SOC) medium (Appendix 10 A) was added and incubated at 37°C in a shaking incubator at 150 rpm for 1 hour 30 minutes in order for the bacteria to recover and express the antibiotic marker encoded by the plasmid. The culture was then centrifuged at 3,000 rpm for 3 minutes and ~700 µl of supernatant discarded and the pellet dissolved in the remaining ~300 µl supernatant. 50 µl, 100 µl and 150 µl of the transformed cells were plated up and spread on Luria Burtani (LB) agar (Appendix 10 B) containing 100µg/ml Ampicillin (Amp100) and incubated overnight at 37°C.

#### 2.8.1.2 Screening for positive clones

After an overnight incubation, colonies were picked and streaked on LB agar plates with Amp100 and screened for recombinants (vectors with N-*APOL1* isoform inserts) by colony PCR using the reaction template below (Table 2.13). The plates were then incubated at 37°C overnight.

**Table 2.13:** Colony PCR reaction

| **Components** | **Volume per reaction (µl)** |
|---|---|
| (2X) KAPA 2G robust HotStart Readymix | 10.0 |
| pJET 1.2 Forward sequencing primer (10µM) | 0.4 |
| pJET 1.2 Reverse sequencing primer (10µM) | 0.4 |
| Water, nuclease-free | 9.2 |
| **TOTAL** | **20.0** |

A single reaction is represented.

The contents in the tubes were then mixed and 20µl aliquoted in labelled PCR tubes on ice. Colonies streaked on the plates above were re-suspended in corresponding labelled PCR tubes and placed in a thermocycler, MJ mini™ Gradient Thermocycler (BioRad laboratories, 1000 Alfred Noble drive Hercules, CA 94547, USA) under the following conditions (Table 2.14), in order to amplify the region containing the inserts/isoforms.

**Table 2.14:** Thermocycler conditions

| | Temperature | Time |
|---|---|---|
| **Denaturation** | 95°C | 3 minutes |
| **PCR (25 cycles)** | 94°C | 30 seconds |
| | 60°C | 30 seconds |
| | 72°C | 30 seconds |
| **Final Extension** | 72°C | 1 minute |
| **Hold** | 4°C | ∞ |

10µl of the colony PCR reactions were then analyzed on 2.5% agarose gel (set up as described in Appendix 7 B) stained with ethidium bromide [3µl of ethidium in 2.5% agarose solution (2.5g of agarose in 100mL of 1X TBE solution)] followed by viewing under ultra violet (UV) light using Gel Doc™ EZ Imaging system (BioRad laboratories, 1000 Alfred Noble drive Hercules, CA 94547, USA). The images were captured and the isoform insert identified based on their different molecular weights.

Once positively identified, two of each clone were incubated in LB broth (Appendix 10 C) containing Amp100 at 37°C in an incubator and agitated at 100 rpm on a shaker (Orbital shaker incubator, LM-510) for bacterial propagation.

#### 2.8.1.3 Plasmid DNA purification

DNA plasmid purification was performed using The Zyppy™ Plasmid Miniprep Kit (Zymo Research 17062 Murphy Ave., Irvine, CA 92614, and U.S.A.) This kit employs a pellet-free modified alkaline lysis method that has an advantage of bypassing the centrifugation and re-suspension steps common to classic plasmid preparation procedures. This entails addition of the uniquely formulated 7X Lysis Buffer directly to the bacterial culture, followed by a neutralization step and then plasmid purification using *Fast-Spin* column technology. The coloured buffers included in the kit helps in the visualization and identification of complete bacterial cell lysis and neutralization. The resulting plasmid DNA is of the high quality and is

well suited for use in downstream applications such as transfections, bacterial transformations, restriction endonuclease digestions, DNA ligations, PCR, transcriptions and DNA sequencing. A detailed protocol is described in Appendix 11.

### 2.8.2 Sub-cloning apolipoprotein L1 isoforms into the *pDsRed2-N1* vector

*pDsRed2-N1* (4680bp) encodes *Discosoma sp.* red fluorescent protein 2 (DsRed2) derived from the *Discosoma sp.* red fluorescent protein 583 (drFP583). It has a series of non-synonymous base-pair mutations which play an important role in the high expression in mammalian cells. It also contains six amino acid substitutions: V105A, I161T, and S197A, that result in rapid appearance of red fluorescence in transfected cell lines and R2A, K5E, and K9T, that prevent the protein from aggregating. Due to these characteristics, this protein is well tolerated by mammalian cell cultures. These transfected mammalian cell cultures show no obvious signs of reduced viability and are similar in morphology and in the manner they grow as non-transfected cell cultures. The plasmid map and features of the *pDsRed2-N1* cloning vector is shown in Appendix 12.

#### 2.8.2.1 Digestion of DNA insert/isoform from *pJET 1.2/blunt* vector

FastDigest® BglII (Fermentas, Life Sciences, USA) was found to be the most suitable restriction enzyme to use to restrict the isoform inserts from the recombinant *pJET1.2/blunt* vector since we could not find any BglII restriction sites within the isoform inserts using *RestrictionMapper version3.0* program and the vector has two BglII sites that enables the double-digest of the insert from the vector resulting in 3'dA overhang sticky-ends. BglII is a type II restriction endonuclease enzyme isolated from certain species of *Bacillus globigii* that recognizes

5'…A^GATCT…3' sites and cuts best at 37°C in 5-30 minutes. The restriction digest was performed according to the following reaction below in Table 2.15:

**Table 2.15:** BglII restriction digest recombinant pJET1.2 vector

| **Components** | **Volume per reaction (µl)** |
|---|---|
| 10X FastDigest® Buffer | 2.0 |
| Plasmid DNA | 2.0 (1µg) |
| Water, nuclease-free | 15.0 |
| FastDigest® Enzyme | 1.0 |
| **TOTAL** | **20.0** |

Reaction conditions for a single digest are represented and were scaled up accordingly.

The reaction tubes were mixed and incubated at room temperature for 20 minutes to ensure complete digestion of the insert of interest from the *pJET 1.2* cloning vector. The resulting inserts have 3' dA overhang sticky-ends and following purification, were used in ligation reactions to linearized and dephosphorylated *pDsRed2-N1* vector.

#### 2.8.2.2 Fast simultaneous *pDsRed2-N1* linearization and dephosphorylation

The circularized plasmid, *pDsRED2-N1* was linearized and dephosphorylated in preparation for cloning of the inserts derived from the restriction digests (2.8.2.1). Linearization of the *pDsRED2-N1* was carried out using the restriction enzyme FastDigest® BglII (Fermentas, Life Sciences, USA). BglII restriction was followed by dephosphorylation of the vector by the addition of the FastAP alkaline phosphatase (Fermentas, Life Sciences, USA). This reaction was carried out simultaneously in the same reaction tube according to the reaction outlined in Table 2.16 below:

**Table 2.16:** Fast simultaneous *pDsReD2-N1* linearization and dephosphorylation reaction

| Components | Volume per reaction (µl) |
|---|---|
| 10X FastDigest® Buffer | 2.0 |
| Plasmid DNA (Empty pDsRED2-N1) | 1.0 |
| FastDigest® Enzyme (BglII) | 1.0 |
| FastAP® Alkaline phosphatase (5U/µl) | 1.0 |
| Water, nuclease-free | 15.0 |
| **TOTAL** | **20.0** |

A template for a single sample used to calculate the reaction volumes. In our calculations we scaled-up our master mix according to the required reactions

The contents in the tubes were mixed thoroughly and incubated at 37°C for 10 minutes followed by 65°C for 15 minutes in order to de-activate alkaline phosphatase.

#### 2.8.2.3 Ligation of isoform into linearized and dephosphorylated *pDsRed2-N1* vector

The digested DNA insert/isoforms above were ligated into linearized and dephosphorylated *pDsReD2-N1* vector according to the following reaction (Table 2.17):

**Table 2.17:** Ligation reaction

| Components | Volume per reaction |
|---|---|
| Linearized and dephosphorylated *pDsReD2-N1* vector | 20 ng |
| DNA insert-isoforms (A/B/C) | 4:1 ratio over vector |
| 10X T4 DNA Ligase Buffer | 2.0 µl |
| T4 DNA Ligase | 1.0 µl |
| Water, nuclease-free | Up to 20.0µl |
| **TOTAL** | **20.0µl** |

A template for a single sample used to calculate the reaction volumes. In our calculations we scaled-up our master mix.

The reaction mixtures containing all the three isoforms were then incubated at room temperature for 10 minutes and 5µl of each ligation reaction (containing Isoforms A, B and C) was used for transformation into *E. coli* DH5α competent cells.

#### 2.8.2.4 Transformation into already prepared E. coli DH5α competent cells

The transformations were carried out as described previously (2.8.1.1). 50 µl, 100 µl and 150 µl of the transformed cells were plated up and spread on LB agar (Appendix 10 B) plates containing 50µg/ml of the selectable marker kanamycin (Kan50) and incubated overnight at 37°C.

#### 2.8.2.5 Screening for positive clones

After an overnight incubation, colonies were picked up and streaked on LB plates with Kan50 and screened for recombinant vectors as described previously (2.8.1.2) but this time using *APOL1*-specific primers as seen in the table below (Table 2.18). The plates were then incubated at 37°C overnight, for multiplication.

**Table 2.18:** Colony PCR for Isoforms A, B and C reaction

| **Components** | **Volume per reaction (µl)** |
|---|---|
| (2X) KAPA 2G robust HotStart Readymix | 6.250 |
| *APOL1* Iso A/C/B Forward primer | 0.625 |
| *APOL1* Iso A/C/B Reverse primer | 0.625 |
| Water, nuclease-free | 5.0 |
| **TOTAL** | **12.5** |

A template for a single sample used to calculate the reaction volumes. In our calculations we scaled-up our master mix.

The contents in the tubes were then mixed well and 12.5µl aliquoted in labelled PCR tubes on ice. The colonies streaked on plates were re-suspended in corresponding labelled PCR tubes and placed in thermocycler, MJ mini™ Gradient Thermocycler (BioRad laboratories, 1000 Alfred Noble drive Hercules, CA 94547, USA) under the following conditions (Table 2.19) in order to amplify the region containing the inserts/isoforms.

**Table 2.19:** Thermocycler conditions

| | **Temperature** | **Time** |
|---|---|---|
| **Denaturation** | 95°C | 3 minutes |
| **PCR (35 cycles)** | 95°C | 15 seconds |
| | 60°C | 15 seconds |
| | 72°C | 15 seconds |
| **Final Extension** | 72°C | 1 minute |
| **Hold** | 4°C | ∞ |

The colony PCR reactions were then analyzed on 2.5% agarose gel as described earlier (2.8.1.2) and images were taken of the different insert/isoforms which were identified based on their different molecular weights.

Once positively identified, each clone containing the required insert was incubated in LB broth that contains Kan50 at 37°C in a shaking incubator (Orbital shaker incubator, LM-510) at 100 rpm overnight for effective growth.

#### 2.8.2.6 Plasmid DNA purification

This was performed as described earlier (2.8.1.3) and Appendix 11.

The recombinant *pDsRed2-N1*plasmids containing the N-terminal region of the different isoforms A, B and C were used for transfection of human embryonic kidney cells 293 (HEK 293) cells.

## 2.9 Culturing of Human Embryonic Kidney (HEK) 293 cells

Human Embryonic Kidney (HEK) 293 cells were thawed and propagated in Dullbecco's modified Eagle medium (DMEM) supplemented with 10% Foetal bovine serum (FBS) and 1 % penicillin and streptomycin (pen-strep).

The cells were then incubated and maintained at 37°C with 5% $CO_2$ until 60-80 % confluent (~5 Days). On the 5$^{th}$ day, with the cells 80% confluent, the growth medium was discarded, cells rinsed with 1X phosphate buffer saline (PBS), followed by 2ml of Trypsin EDTA and a brief incubation at 37$^{o}$C in order to detach the cells from the surface of the flask. 2ml of fresh medium was then added and mixed up and down on the surface of the flask until all the cells have detached and lumps dispersed. 2ml of the cells were then taken out in a separate tube to be used in cell counting and seeding. 8ml of fresh medium was then added to the remaining 2ml in the flask and placed in the incubator at 37$^{o}$C.

### 2.9.1 Determining the Cell number and viability

Cell number and viability is important in standardization of culture conditions and performance of accurate experiments. This was done using a hemacytometer and trypan blue staining. A ratio of 1:1 of the cell suspension to trypan blue was mixed in a tube and left to stand for 5 minutes before loading to the hemacytometer. Using a sterile pipette, the trypan blue stained cell suspension was transferred to the edge of the hemacytometer counting chamber by holding the tip of the pipette under the coverslip and dispensing one drop of the suspension. The suspension was then drawn under the coverslip by capillary action. The cells were allowed to settle for a few minutes before counting and the excess liquid blotted off. The slide was viewed under the microscope, Olympus U-TV 0.5 C-3 (Olympus Corporation) at 40X magnification. The slide was positioned to view the large central area of the grid which is bordered by a set of three parallel lines. Non-viable cells normally take up the dye while live cells are impermeable to dye. A hand-held counter was used to count viable cells in each of the four corners and central squares (quadrants). Cells touching the middle line of the triple lines on the bottom or right side of the square were not counted. In order to determine the number of viable cells per milliliter (cells/ml), the following formula was used:

**Cells/ml= Number of cell counts/number of the quadrants X 10,000 X dilution factor.** Cells were seeded at 125,000 cells per well on cover slips in 3 well plates for the three isoforms (A, B and C) and placed at 37$^{o}$C in preparation for transfection the following day.

### 2.9.2 Transfection of HEK 293 cells with *APOL1* expressing constructs

On the day of transfection, the growth medium was removed from cells and replaced with 2ml of complete growth medium. For each well of cells to be transfected, we diluted 0.5 µg of plasmid DNA expressing different N-terminal isoforms A, B and C in 150 µl of Opti-MEM® I Reduced

Serum Media without serum. For each well of cells, we added 6.0µl of Lipofectamine 2000® Reagent into the above diluted Opti-MEM®: DNA solution mixed gently and incubated for 5 minutes at room temperature to form DNA- Lipofectamine 2000® Reagent complexes. After 5 minutes incubation, 275µl of the DNA- Lipofectamine 2000® Reagent complexes were added to each well containing and mixed gently by rocking the plate back and forth. The cells were incubated at 37$^{o}$C in a $CO_2$ incubator for 72 hours post-transfection before assaying for transgene expression. The media was changed after the 6 hours.

### 2.9.3 Immunofluorescence staining of transfected HEK 293 cells and confocal microscopy

#### 2.9.3.1 Cell preparation

The cells were removed from the incubator and the DMEM media discarded and the cells attached on the coverslips were washed three times in PBS, pH 7.2-7.4. Excess solution was rinsed off gently from the coverslips using a paper towel.

#### 2.9.3.2 Cell fixation

The coverslips were then placed on appropriately labelled slides with the side containing the cells facing up. The coverslips were then fixed in chilled methanol at -20$^{o}$C for 10 minutes. They were then washed three times in phosphate buffer saline (PBS), pH 7.2-7.4. Excess solution was rinsed off gently from the coverslips using a paper towel. DAPI (4', 6-diamidino-2-phenylindole) solution (1:1000) was added onto the slides in order to stain them followed by a three-time wash in phosphate buffer saline (PBS), pH 7.2-7.4 and distilled water. Excess solution was rinsed off gently from the coverslips using a paper towel. The coverslips were then mounted with Dako fluorescence (An Agilent Technologies Company, Dako North America, Inc. 6392 Via Real

Carpinteria, CA 93013) mounting medium and inverted onto new and dry glass slides. They were then incubated at room temperature in the dark for 1 hour and inspected under the confocal microscope (Carl Zeiss cell observer CD). The images were then captured in accordance with the manufacturer's manual of instructions.

## 2.10 Statistical Analysis

The Kruskal Wallis test was used to test for the association of the clinical (serum creatinine, eGFR, viral load and CD4) data among the groups using Stata v11.1. The quality controlled genotypic data and clinical outcome files were merged using *R program version 2.15.2*. Tests for single SNP associations with the three kidney diseases [HIVAN, HIV (+) CKD, and HIV (-) CKD] were first performed using logistic regression analysis adjusting for age and gender. We then fitted logistic regression models in which the highly associated SNPs [*APOL1* (rs73885319, rs60910145, rs71785313, rs7284919, rs136147), *MYH9* (rs4821480, rs4821481, rs3752462, rs16996648, rs16996668, rs16996672, rs16996674, rs16996677, rs5750252) and *APOL3* (rs11089781)] with HIVAN. The predictor variable values were coded as follows: equal to the number of copies of the minor allele (0, 1, 2) in the additive model, presence of at least one copy of the minor allele (1 or 2) in the dominant model or presence of the two copies of the minor allele in the recessive model.

In order to determine possible associations between the HIVAN and various configurations of multiple SNPs and also validate the association of *MYH9* E1 haplotype SNPs (rs4821480, rs4821481, rs3752462) with HIVAN since this has been shown to be associated with kidney diseases in other studies (Freedman et al., 2009a; Kopp et al., 2008; Tayo et al., 2013; Winkler et al., 2010), we performed haplotype association analysis of the highly associated SNPs using

logistic regression models adjusting for age and gender and also correcting for multiple comparisons.

We used the software Haploview (Barrett et al., 2005) to compute the estimates of linkage disequilibrium for each pair of SNPs by the standard $r^2$ method and to determine the haplotype blocks, regions with no evidence of historical recombination events, but significant levels of LD. The haplotype blocks were defined by the four-gamete test as implemented in Haploview.

We also performed Fisher's exact test to determine the effect size of the G1 (rs73885319 and rs60910145) and G2 (rs71785313) alleles with HIVAN, HIV (+) CKD and HIV (-) CKD, where we stratified the analysis by comparing individuals with two risk alleles (G1 or G2) *versus* zero or one risk alleles using the recessive model. A similar analysis was performed comparing the haplotypes.

Fisher's exact test and Logistic regression were used to test the effect of *APOL1* risk alleles on the kidney disease phenotypes, adjusting for age, sex and the first two eigen values obtained from a PCA of population substructure. We also tested recessive models (comparing individuals with two risk alleles, i.e. G1/G1, G1/G2, or G2/G2, against all others), but also tested other models (recessive; and 1 vs. 0, 2 vs. 0, 2 vs. 1, and 2 vs. 1 or 0 risk alleles). Comparisons were made between HIV (+) cases and HIV (+) controls and between HIV (-) cases and Population Controls, comparing the distribution of two risk alleles (explanatory exposure) to 0 or 2 risk alleles (no exposure). To test for homogeneity of the odds ratios between the South African and African American FSGS populations, a Woolf test was used.

The Student's *t test* was used to determine the level of significance between cases and controls in the N-terminal mRNA *ApoL1* isoform expression patterns. *P*-values less than 0.05 were considered significant.

# Chapter 3

# 3) RESULTS

## 3.1 Clinical Characteristics

This study was composed of 228 adult individuals of black African ancestry from Johannesburg, South Africa. The clinical characteristics are summarized in Table 3.1. The mean age for HIVAN group was 34.6 ± 8.3 years, HIV (+) CKD group was 36.1 ± 9.3 years, HIV (-) CKD was 36.4 years, HIV (+) Control group was 38.8 ± 7.1 years and among the population control group was 38.5 ± 8.6 years. The presence of glomerular capillary collapse and glomerular visceral epithelial cell proliferation affecting at least one glomerulus, together with microcystic tubular dilatation and interstitial inflammation differentiated HIVAN from HIV (+) FSGS.

Serum creatinine levels were observed to be highest in the HIVAN group with a median of 534.0µmol/L and, as expected, lowest (in the normal range) among the HIV positive with no CKD at median of 66.0 µmol/L. Elevated serum creatinine was also observed among the HIV (+) CKD (median of 233.0µmol/l) and HIV (-) CKD (median of 155.0µmol/l). Median estimated glomerular filtration rate (eGFR) was observed to be lowest in HIVAN (11.9mL/min/1.73m$^2$) with HIV (+) CKD and HIV (-) CKD groups having 33.5 and 56.6mL/min/1.73m$^2$ respectively. All the CKD patients had eGFR<60.0mL/min/1.73m$^2$. However, among HIV (+) Controls, eGFR was 121.6mL/min/1.73m$^2$. The association of the clinical (serum creatinine and eGFR) data among groups was statistically significant ($p<0.0001$) when comparing HIV (+) CKD and HIV (+) controls. The population control group was assumed to be "healthy" individuals since they lacked measures of kidney function. They however, gave consent to be enrolled for population genetic studies as control individuals. There was no significant difference by age (34.6, 36.1,

36.4 and 38.8 years respectively) and gender at biopsy or study enrollment (*P >0.05*) between the CKD groups and HIV (+) control group (Table 3.1). HIVAN showed the highest median viral load (Log10 copies/ml) levels (5.1 log copies/ml) compared to HIV (+) CKD (4.7 log copies/ml). Undetectable median levels were observed in HIV (+) controls. HIVAN also showed the lowest median CD4 counts HIVAN (92 cells/mm$^3$) compared to HIV (+) CKD (244.5 cells/mm$^3$) and HIV (+) controls (371.0 cells/mm$^3$); Table 3.1.

**Table 3.1:** Demographic characteristics of the study population

| **Characteristic** | **HIVAN[1] (N=39)** | **HIV (+) CKD[2] (N=40)** | **HIV (+) Controls[3] (N=54)** | **P value[4]** | **HIV (-) CKD[5] (N=41)** | **Population controls[6] (N=54)** |
|---|---|---|---|---|---|---|
| **Sex, N (%)** | | | | | | |
| Male | 16.0 (41.0) | 18.0 (42.9) | 15.0 (28.0) | 0.069 | 24.0 (58.1) | 24·0 (44.4) |
| Female | 23.0 (59.0) | 22.0 (57.1) | 39.0 (72.0) | 0.726 | 17.0 (41.9) | 30·0 (55.6) |
| **Age, Mean (SD)** | 34.6 (8.3) | 36.1 (9.3) | 38.8 (7.1) | 0.623 | 36.4 (11.0) | 38.5 (8.6) |
| **Viral Load (Log copies/ml)** [a] Median (IQR) | 5.1 (4.5-5.3) | 4.7 (3.8-4.9) | UD[8] | 0.001 | - | - |
| **CD4 (cells/mm$^3$)**[b] Median (IQR) | 92.0 (76-195) | 244.5 (71.5-351.5) | 371.0 (215.0-545.0) | 0.001 | - | - |
| **eGFR (ml/min/1.73m$^2$)** Median (IQR) | 11.9 (7.5-22.6) | 33.5 (15.9-68.8) | 121.6 (103.3-146.9) | 0.001 | 56.6 (8.3-101.7) | NA |
| **Serum Creatinine (µmol /L)** Median (IQR) | 534.0 (262.0-799.0) | 233.0 (106.0-388.0) | 66.0 (55.0-77.0) | 0.001 | 155.0 (91.0-778.0) | NA |

[1]HIV-associated nephropathy (collapsing FSGS); [2]Chronic kidney disease (see Table 3 for histological diagnosis); [3]HIV positive controls with no kidney disease; [4]Kruskal Wallis test comparing HIV positive CKD cases to HIV positive controls. [5]HIV negative chronic kidney disease (see Table 3); [6]Population controls were from NHLS; NA; Not available. [7] eGFR calculated according to the modified MDRD formula. [8]UD; undetectable. [a]HIVAN (N=30), HIV (+) CKD (N=27), HIV (+) Controls (N=45). [b]HIVAN (N=26), HIV (+) CKD (N=32), HIV (+) Controls (N=41). Permission obtained from the American Society of Nephrology publishing group (Journal of the American Society of Nephrology). Appendix 15 F.

## 3.2 SNP Selection

The SNP selection process resulted in a total of 97 SNPs from non-synonymous coding, synonymous coding intron and splice site, 5' and 3'UTRs, upstream, downstream near start of transcription sites of the *MYH9* and *APOL1* genes including the AIMs (Figure 3.1). However, after, subjecting the SNPs to standard quality control procedures described in Chapter 2, 20 SNPs did not pass quality control procedures and therefore 77 SNPs (17 *APOL1* SNPs, 41 *MYH9* SNPs, 1 *APOL3* SNP and 18 AIMs) were eligible for analysis. The full SNP list is found in Table A13.1 (Appendix 13). The gene structure of the 59 *(APOL1*, *APOL3* and *MYH9)* SNPs (Chromosome 22) is shown in Figure 3.2.

## 3.3 SNP Genotyping Quality Control

A total of 228 samples [39 HIVAN, 40 HIV (+) CKD and 41 HIV (-) CKD cases and 54 HIV (+) controls and 54 Population control (PC)] were genotyped (Table 2.1). After the initial assay performance was assessed using the inbuilt sample-dependent, sample independent and contamination controls (Allele Specific Extension, PCR Uniformity, Extension Gap, First Hybridization and Second Hybridization) and a scatter plot of the call rate as a function of the sample number against a 10% GC score, 4 samples that did not meet the quality assurance requirements were excluded. Therefore, a total of 224samples [38 HIVAN, 39 HIV (+) CKD and 39 HIV (-) CKD cases and 54 HIV (+) controls and 54 PC)] were eligible for analysis.

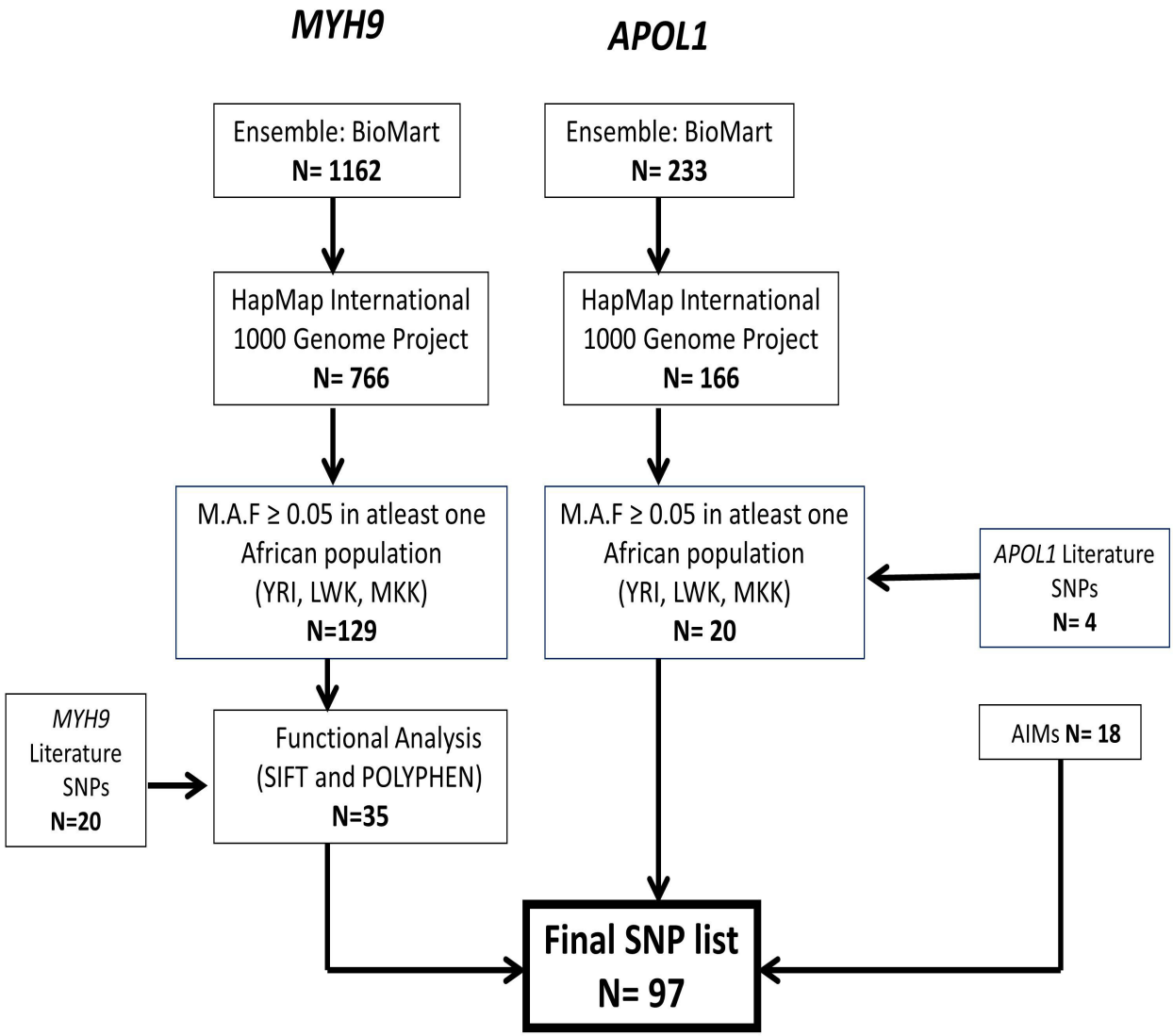



**Figure 3.1:** Schematic diagram showing selected SNPs
SIFT; sorting intolerant from tolerant. POLYPHEN; polymorphism phenotyping. AIMs; ancestry informative markers. YRI; Yoruba from Ibadan, Nigeria. LWK; Luhya from Kenya. MKK; Maasai from Kenya. Literature SNP were SNPs previously associated with kidney disease.

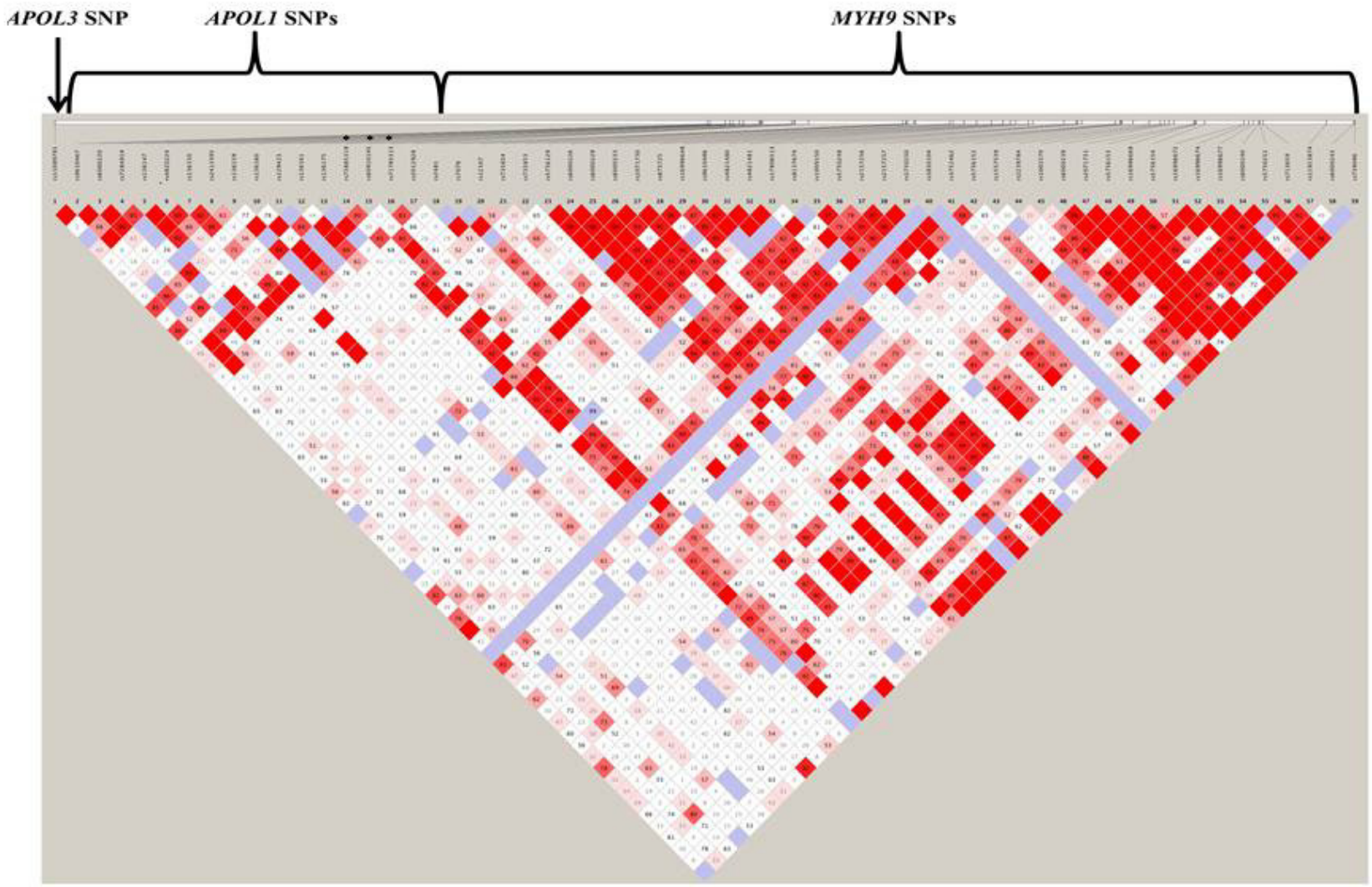



**Figure 3.2:** Gene structure and linkage disequilibrium plot of *MYH9* and *APOL1* SNPs.
59 *APOL1, APOL3* and *MYH9* SNPs. Inter SNP D' values are displayed on the plot. SNPs marked with asterix are the G1/G2 risk alleles which represent independent strong associations.

## 3.4 Frequencies and associations of *MYH9* and *APOL1* SNPs and *APOL3* literature SNP

### 3.4.1 Frequencies of *MYH9* SNPs

The frequencies of the alleles of the individual 41 *MYH9* SNPs differed across CKD cases [HIVAN, HIV (+) CKD and HIV (-) CKD) and controls (HIV (+) controls and PC group] (Appendix 14). Risk alleles had higher frequencies in HIVAN cases when compared to HIV (+) controls than other CKD case groups [HIV (+) CKD and HIV (-) CKD] when compared to HIV (+) controls and population controls respectively. These included *MYH9* E1 risk haplotype SNPs as shown in Table 3.2. More details on allele frequencies are found in Tables A14.1, A14.2 and A14.3.

### 3.4.2 Associations of *MYH9* SNPs

We compared HIVAN and HIV (+) CKD cases with the HIV (+) controls, and the HIV (-) CKD cases with the PC group. The *MYH9* individual SNP association with different forms of kidney disease revealed no significant association with HIVAN, HIV (+) CKD or HIV (-) CKD groups, using logistic regression adjusted for age and gender except for rs16996648, rs5756152, rs16996668, rs16996672, rs16996674 and rs16996677 with higher odds ratios as shown in Table 3.3. These SNPs are however not part of the risk haplotypes that have been shown to predispose to kidney disease and therefore considered to have no possible effect on kidney disease. *MYH9* E1 haplotype also showed no association with any of the case groups (Table 3.2). Since this observation was a unique *MYH9* association with the HIVAN in our population, our focus was narrowed down to HIVAN in subsequent analysis. We performed the covariate-adjusted case/control analyses considering the three genetic modes of associations (dominant, recessive

and additive) in HIVAN in order to find out whether there could be an association with other genetic modes, and we still did not find significant association even after correcting for multiple comparisons (Table 3.4). More details on the associations of the individual *MYH9* SNPs are found in Appendix 14. We also validated the association of *MYH9* E1 haplotype SNPs known to have been associated with kidney disease in other studies (Freedman *et al.,* 2009a; Kopp et al., 2008; Tayo *et al.*, 2013; Winkler *et al.,* 2010) by including them in our combined haplotype analysis (Figure 3.6 and Table 3.10).

**Table 3.2**: Logistic regression analysis showing *MYH9* E1 haplotype individual SNP association with HIVAN, HIV (+) CKD and HIV (-) CKD cases

| | | | | **Risk allele frequency** | | | ***p-values*** | |
|---|---|---|---|---|---|---|---|---|
| | Gene | SNPs | Risk Allele | Cases | Controls | Odds Ratio | Unadjusted | Adjusted |
| **HIVAN** | *MYH9* | rs4821480 | G | 0.6667 | 0.4583 | 0.68 | 1.0000 | 0.9450 |
| | *MYH9* | rs4821481 | C | 0.6364 | 0.4490 | 0.67 | 1.0000 | 0.9547 |
| | *MYH9* | rs3752462 | T | 0.9677 | 0.9387 | 1.94 | 1.0000 | 0.4329 |
| **HIV (+) CKD** | *MYH9* | rs4821480 | G | 0.6756 | 0.4583 | 0.25 | 0.3129 | 0.0948 |
| | *MYH9* | rs4821481 | C | 0.6757 | 0.4489 | 0.24 | 0.3102 | 0.0907 |
| | *MYH9* | rs3752462 | T | 0.9722 | 0.9388 | 2.26 | 0.6341 | 0.6184 |
| **HIV (-) CKD** | *MYH9* | rs4821480 | G | 0.6206 | 0.5208 | 0.69 | 0.7383 | 1.0000 |
| | *MYH9* | rs4821481 | C | 0.6000 | 0.5106 | 0.59 | 0.5216 | 1.0000 |
| | *MYH9* | rs3752462 | T | 0.9677 | 0.9574 | 1.33 | 1.0000 | 1.0000 |

Allele frequencies and association obtained from linear regression in the dominant model. Adjusted for age and gender. SNPs: Single nucleotide polymorphisms

**Table 3.3**: Logistic regression analysis showing Individual SNP association of rs16996648, rs5756152, rs16996668, rs16996672, rs16996674, rs16996677 with HIVAN, HIV (+) CKD and HIV (-) CKD cases compared to their respective controls.

| | | | **Risk allele frequency** | | | | ***p-values*** | |
|---|---|---|---|---|---|---|---|---|
| | Gene | SNPs | Risk Allele | Cases | Controls | Odds Ratio | Unadjusted | Adjusted |
| **HIVAN** | *MYH9* | rs16996648 | C | 0.4706 | 0.2128 | 3.81 | 0.0099 | 0.0056 |
| | *MYH9* | rs5756152 | A | 0.5000 | 0.1875 | 5.34 | 0.0017 | 0.0027 |
| | *MYH9* | rs16996668 | C | 0.4571 | 0.1735 | 5.15 | 0.0006 | 0.0011 |
| | *MYH9* | rs16996672 | T | 0.4722 | 0.2083 | 3.67 | 0.0070 | 0.0102 |
| | *MYH9* | rs16996674 | T | 0.4722 | 0.1383 | 6.78 | $9.40x10^{-05}$ | 0.0003 |
| | *MYH9* | rs16996677 | A | 0.7333 | 0.3878 | 4.26 | 0.0050 | 0.0084 |
| **HIV (+) CKD** | *MYH9* | rs16996648 | C | 0.0294 | 0.0000 | 1.06 | 1.0000 | 0.9760 |
| | *MYH9* | rs5756152 | A | 0.3529 | 0.3542 | 0.99 | 1.0000 | 0.8118 |
| | *MYH9* | rs16996668 | C | 0.0541 | 0.0204 | 1.40 | 0.5011 | 0.9062 |
| | *MYH9* | rs16996672 | T | 0.4865 | 0.4167 | 1.32 | 0.6603 | 0.9485 |
| | *MYH9* | rs16996674 | T | 0.4054 | 0.2766 | 1.77 | 0.2489 | 0.6225 |
| | *MYH9* | rs16996677 | A | 0.4412 | 0.3878 | 1.24 | 0.6559 | 0.8585 |
| **HIV (-) CKD** | *MYH9* | rs16996648 | C | 0.0385 | 0.1111 | 0.47 | 0.2051 | 1.0000 |
| | *MYH9* | rs5756152 | A | 0.4286 | 0.3830 | 1.21 | 0.8084 | 1.0000 |
| | *MYH9* | rs16996668 | C | 0.0968 | 0.0208 | 0.95 | 1.0000 | 1.0000 |
| | *MYH9* | rs16996672 | T | 0.4000 | 0.3958 | 1.02 | 1.0000 | 1.0000 |
| | *MYH9* | rs16996674 | T | 0.3333 | 0.3261 | 1.03 | 1.0000 | 1.0000 |
| | *MYH9* | rs16996677 | A | 0.4333 | 0.4043 | 1.13 | 0.8169 | 1.0000 |

Allele frequencies and association obtained from linear regression in the dominant model. Adjusted for age and gender. SNPs: Single nucleotide polymorphisms

**Table 3.4**: Logistic regression analysis showing *MYH9* SNP associations with 38 HIVAN cases and 54 HIV (+) Controls

| | | | Risk Allele Frequencies | | | | *P Values* | | |
|---|---|---|---|---|---|---|---|---|---|
| **Gene** | **SNP** | **Risk Allele** | **Cases** | **Controls** | **Association Mode** | **OR (95% CI)** | **Unadjusted** | **Adjusted** | **Corrected*** |
| *MYH9* | rs16996648 | C | 0.4706 | 0.2128 | Dominant | 3.69(1.32-10.30) | 0.0068 | 0.0127 | 0.1059 |
| | | | | | Recessive | Inf (0.00-Inf) | 0.9981 | 0.9980 | 1.0000 |
| | | | | | Additive | 4.36 (1.78-10.68) | 0.0077 | 0.0013 | 0.0069 |
| *MYH9* | rs16996668 | C | 0.4571 | 0.1735 | Dominant | 6.28 (2.17-18.20) | 0.0007 | 0.0007 | 0.0039 |
| | | | | | Recessive | 11.77 (1.22-113.30) | 0.0233 | 0.0328 | 0.0589 |
| | | | | | Additive | 4.93 (2.01-12.12) | 0.0003 | 0.0005 | 0.0019 |
| *MYH9* | rs16996672 | T | 0.4722 | 0.2083 | Dominant | 4.40 (1.57-12.36) | 0.0064 | 0.0050 | 0.0379 |
| | | | | | Recessive | Inf (0.00-Inf) | 0.9980 | 0.9979 | 1.0000 |
| | | | | | Additive | 4.84 (1.95-12.01) | 0.0006 | 0.0007 | 0.0029 |
| *MYH9* | rs16996674 | T | 0.4722 | 0.1383 | Dominant | 9.17 (2.98-28.18) | 0.0001 | 0.0001 | 0.0009 |
| | | | | | Recessive | Inf (0.00-Inf) | 0.9980 | 0.9979 | 1.0000 |
| | | | | | Additive | 8.34 (2.95-23.55) | $4.03x10^{-05}$ | $6.27x10^{-05}$ | 0.0009 |
| *MYH9* | rs16996677 | A | 0.7333 | 0.3878 | Dominant | 6.18 (2.04-18.79) | 0.0033 | 0.0013 | 0.0079 |
| | | | | | Recessive | Inf (0.00-Inf) | 0.9980 | 0.9978 | 1.0000 |
| | | | | | Additive | 6.42 (2.34-17.42) | 0.0003 | 0.0003 | 0.0019 |
| *MYH9* | rs5756152 | A | 0.50 | 0.1875 | Dominant | 5.46 (1.86-16.03) | 0.0011 | 0.0020 | 0.0139 |
| | | | | | Recessive | 14.28 (1.54-132.00) | 0.0104 | 0.0191 | 0.0269 |
| | | | | | Additive | 4.56 (1.89-10.98) | 0.0003 | 0.0007 | 0.0039 |
| ***MYH9*** | **rs4821480** | **G** | **0.6667** | **0.4583** | **Dominant** | **0.41 (0.15-1.12)** | **0.0330** | **0.0832** | **0.5574** |
| | | | | | **Recessive** | **2.10 (0.07-60.99)** | **0.8369** | **0.6651** | **1.0000** |
| | | | | | **Additive** | **0.50 (0.20-1.25)** | **0.0598** | **0.1375** | **0.7263** |
| ***MYH9*** | **rs4821481** | **C** | **0.6364** | **0.4490** | **Dominant** | **0.48 (0.18-1.27)** | **0.0487** | **0.1376** | **0.7463** |
| | | | | | **Recessive** | **2.13 (0.07-62.78)** | **0.8254** | **0.6608** | **1.0000** |
| | | | | | **Additive** | **0.60 (0.22-1.39)** | **0.0821** | **0.2081** | **0.8691** |
| ***MYH9*** | **rs3752462** | **T** | **0.9677** | **0.9387** | **Dominant** | **0.20 (0.05-0.74)** | **0.06423** | **0.0160** | **0.1399** |
| | | | | | **Recessive** | **0.18 (0.02-1.98)** | **0.5149** | **0.1596** | **0.5435** |
| | | | | | **Additive** | **0.27(0.10-0.79)** | **0.0864** | **0.0166** | **0.1359** |

Adjusted for age and gender, SNP: Single nucleotide polymorphism, OR: Odds ratio, CI: Confidence Interval, NE: Not Estimated, **P values* corrected for multiple comparisons

### 3.4.3 Frequencies of the *APOL1* and *APOL3* SNPs

The allele frequencies of each of the *APOL1* 17 SNPs including G1-rs73885319, G1-rs60910145 and G2-rs71785313 SNPs and 1 literature SNP of the *APOL3* region, differed in cases [HIVAN, HIV (+) CKD and HIV (-) CKD] and controls [HIV (+) Controls and PC groups] as shown in Appendix 14. We compared HIVAN and HIV (+) CKD cases with the HIV (+) controls, and the HIV (-) CKD cases with the PC group. As would be expected, the risk allele was more frequent in cases than in controls for example, in four SNPs of the *APOL1* region (G1-rs73885319, G1-rs60910145, G2-rs71785313, rs136147) with a notable difference in rs7284919 and one literature SNP, rs11089781 of the *APOL3* region with HIVAN (Table 3.5). The frequencies of the risk allele in HIV (+) CKD and HIV (-) CKD in correspondence with their respective control groups, was not significantly different (*P>0.05*) as shown in Tables A14.2 and A14.3 respectively in Appendix 14 and in Table 3.5.

**Table 3.5**: Logistic regression analysis showing association of *APOL1* and *APOL3* SNPs with HIVAN, HIV (+) CKD and HIV (-) CKD

| | | | **Risk allele frequency** | | | | ***p-values*** | |
|---|---|---|---|---|---|---|---|---|
| | **Gene** | **SNPs** | **Risk Allele** | **Cases** | **Controls** | **Odds Ratio** | **Unadjusted** | **Adjusted** |
| **HIVAN** | *APOL1* | rs73885319 | G | 0.5556 | 0.0370 | 98.56 | $1.61x10^{-14}$ | $7.203x10^{-07}$ |
| | *APOL1* | rs60910145 | G | 0.4872 | 0.0370 | 46.08 | $4.31x10^{-13}$ | $2.832x10^{-07}$ |
| | *APOL1* | rs71785313 | D | 0.3421 | 0.1667 | 2.08 | 0.0060 | 0.0067 |
| | *APOL1* | rs7284919 | C | 0.0143 | 0.0938 | 0.14 | 0.0446 | 0.0336 |
| | *APOL1* | rs136147 | T | 0.6806 | 0.3370 | 29.39 | $6.39x10^{-06}$ | 0.0006 |
| | *APOL3* | rs11089781 | A | 0.4286 | 0.1170 | 6.08 | 0.0004 | 0.0023 |
| **HIV (+) CKD** | *APOL1* | rs73885319 | G | 0.0110 | 0.0370 | 2.77 | 0.0493 | 0.1252 |
| | *APOL1* | rs60910145 | G | 0.1163 | 0.3704 | 2.62 | 0.0341 | 0.1346 |
| | *APOL1* | rs71785313 | D | 0.1744 | 0.1667 | 0.97 | 0.8865 | 0.9282 |
| | *APOL1* | rs7284919 | C | 0.0278 | 0.000 | 1.66 | 0.4300 | 0.6279 |
| | *APOL1* | rs136147 | T | 0.4412 | 0.5116 | 0.76 | 0.6472 | 0.6708 |
| | *APOL3* | rs11089781 | A | 0.1351 | 0.2340 | 0.52 | 0.2793 | 0.0435 |
| **HIV (-) CKD** | *APOL1* | rs73885319 | G | 0.0909 | 0.0648 | 1.41 | 0.4943 | 0.5289 |
| | *APOL1* | rs60910145 | G | 0.0652 | 0.0556 | 1.15 | 0.7743 | 0.8204 |
| | *APOL1* | rs71785313 | D | 0.0889 | 0.1111 | 0.75 | 0.6054 | 0.5584 |
| | *APOL1* | rs7284919 | C | 0.0183 | 0.0958 | 2.03 | 0.2630 | 1.0000 |
| | *APOL1* | rs136147 | T | 0.6250 | 0.5833 | 1.19 | 0.8025 | 1.0000 |
| | *APOL3* | rs11089781 | A | 0.1379 | 0.3043 | 0.37 | 0.1641 | 1.0000 |

Allele frequencies and association obtained from linear regression in the dominant model. Adjusted for age and gender. SNPs: Single nucleotide polymorphisms

Since G1 and G2 risk alleles had been shown to be highly associated with kidney disease in other studies, we chose to focus on the same risk alleles and determine their frequencies as well as their impact on kidney disease in our black South African population.

*APOL1* G1 and G2 genotypes were determined for 96.7% of CKD cases [HIVAN, N=38; HIV (+) CKD, N=39; and HIV-negative (-) CKD, N=39] and 100% of controls [HIV (+) controls, N=54 and population controls, N=54] (Table 3.6). Genotype distributions did not differ significantly from Hardy-Weinberg equilibrium expectations in HIV (+) and population control groups ($P > 0.05$). SNP and haplotype frequencies are shown in Figure 3.3. The genotype G1/G2 was most frequent in HIVAN and half (19/38; 50.0%) of the HIVAN cases had both G1 and G2 risk alleles and 30/38 (79.0%) had a total of two risk alleles (Table 3.6). The prevalence of the risk alleles of each of the SNPs differed in CKD cases [HIVAN, HIV (+) CKD and HIV (-) CKD] and controls [HIV (+) controls and PC]. The risk allele was more frequent in HIVAN cases than controls. However, the frequencies of G1 and G2 risk alleles were highest in HIVAN than in other forms of kidney diseases and controls. G1 (rs73885319), G1 (rs60910145) and G2 (rs71785313) allele frequencies were 56.2%, 50.0% and 34.2% respectively in HIVAN (Figure 3.3). The two control groups [HIV (+) Controls and PC] were not significantly different to one another as was the case for the HIV (+) CKD and HIV (-) CKD groups.

**Table 3.6**: *APOL1* G1 and G2 genotypes in cases and controls.

| | **HIV (+) cases and controls** | | | **HIV (-) cases and controls** | |
|---|---|---|---|---|---|
| | **HIVAN[T]** | **CKD*** | **HIV (+) Controls** | **CKD[P]** | **Population Controls** |
| **Genotype** | **N (%)** | **N (%)** | **N (%)** | **N (%)** | **N (%)** |
| **0 risk allele** | | | | | |
| G0/G0 | **2 (5.3)** | **22 (56.4)** | **34 (63.0)** | **25 (64.1)** | **36 (66.7)** |
| **1 risk allele** | **6 (15.8)** | **11 (28.2)** | **18 (33.3)** | **13 (33.3)** | **17 (31.5)** |
| G0/G1 | 5 (13.2) | 4 (10.3) | 4 (7.4) | 6 (15.4) | 7 (13.0) |
| G0/G2 | 1 (2.6) | 7 (17.9) | 14 (25.9) | 7 (17.9) | 10 (18.5) |
| **2 risk alleles** | **30 (78.9)** | **6 (15.4)** | **2 (3.7)** | **1 (2.6)** | **1 (1.9)** |
| G1/G1 | 8 (21.0) | 2 (5.1) | 0 (0.0) | 0 (0.0) | 0 (0.0) |
| G1/G2 | 19 (50.0) | 2 (5.1) | 0 (0.0) | 1 (2.6) | 0 (0.0) |
| G2/G2 | 3 (7.9) | 2 (5.1) | 2 (3.7) | 0 (0.0) | 1 (1.9) |
| **Total** | **38** | **39** | **54** | **39** | **54** |

The distribution of *APOL1* genotypes and fractions with 1 or 2 risk alleles for the study groups. [T]One patient did not pass genotyping. *The single HIV (+) CKD subject carrying the $G1^{+M}$ (A-G-I) haplotype is excluded from the table and one patient did not pass genotyping. [P]Two patients did not pass genotyping. (Permission obtained from the American Society of Nephrology publishing group). Appendix 15 F

*Refer to Figure 3.3 for haplotype notation.

#### 3.4.3.1 Haplotype frequencies of the G1 and G2 risk alleles

The five G1 and G2 haplotypes are shown in Figure 3.3. The common haplotype (G0) containing non risk alleles for G1 and the 6-basepair insertion/deletion (A-T-I) was most common in other forms of CKD and controls [HIV (+) CKD (71.8%), HIV (-) CKD (83.7%), HIV (+) Controls (79.6%) and PC (82.4%)]. However, it was least frequent in HIVAN (13.7%). The haplotype containing two risk alleles at G1 and the non-risk allele at G2 ($G1^{GM}$), was most frequent in HIVAN (52.6%) followed by haplotype containing non risk alleles at G1 and 6 bp deletion at G2 (33.7%). The $G1^{G+}$ haplotype was rare and observed in the heterozygous state only in six individuals [three with HIVAN, two in HIV (-) CKD and one in PC] as shown in Figure 3.3 with haplotype frequencies of 3.8%, 1.5% and 0.9% respectively (Figure 3.3). This haplotype accounted for less than 1% of the *APOL1* haplotypes in the combined control groups. There was

one individual with HIV (+) CKD with FSGS who had a unique haplotype $G1^{+M}$ (A-G-I). This haplotype did not feature in any of the other groups. This is the first time such a combination is reported. This novel haplotype was confirmed by Sanger sequencing of the PCR product and was observed as a compound heterozygote ($G1^{+M}/G1^{GM}$) in a single HIV-positive FSGS case. It has however been found that (Kopp *et al.,* 2011) the effect size of the $G1^{GM}$ and $G1^{G+}$ haplotypes are equivalent and so the renal risk of $G1^{+M}$ could not be determined hence we excluded this individual from further analysis in this study.

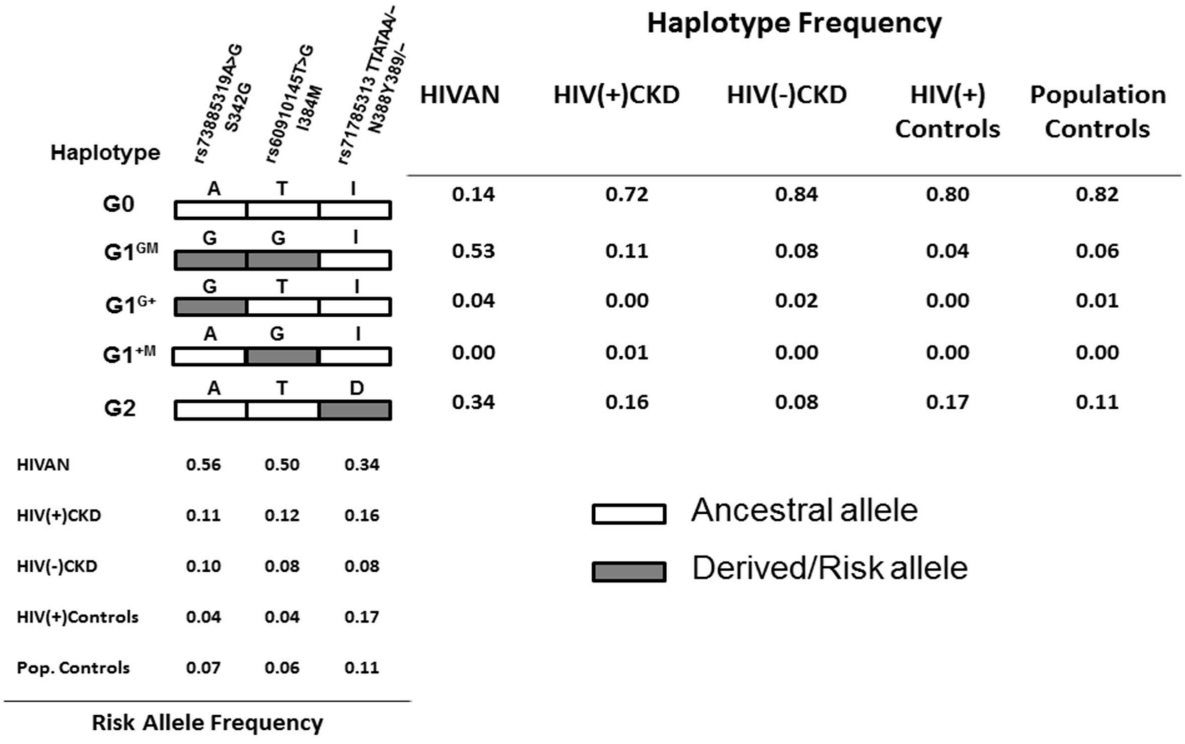



**Figure 3.3**: Distribution of *APOL1* haplotypes and risk alleles.
Five *APOL1* haplotypes were observed. The ancestral haplotype, denoted as G0, has three ancestral alleles; the $G1^{GM}$ haplotype has two missense alleles; the $G1^{G+}$ haplotype has one missense risk allele at rs73885319; the $G1^{+M}$ has one missense variant at rs60910145; the G2 haplotype has the 6 bp deletion risk allele at rs71785313. The risk allele and haplotype frequencies are shown for HIVAN, HIV (+) CKD and HIV (-) CKD cases and HIV (+) controls and population controls (PC). (Permission obtained from the American Society of Nephrology publishing group). Appendix 15 F

### 3.4.4 Associations of the *APOL1* and *APOL3* SNPs

In the association analysis, we compared the HIVAN and the HIV (+) CKD cases with HIV (+) controls and the HIV (-) CKD case with the population control (PC) group. We found 5/17 *APOL1* individual SNPs and 1 *APOL3* individual SNP were not significantly associated with different forms of kidney disease except HIVAN. The strongest signal was for the two-locus allele, rs73885319 G1, rs60910145 G1 ($P=1.61x10^{-14}$ and $P=4.31x10^{-13}$ respectively) in the last exon of *APOL1* (Figure 3.4). SNP rs71785313 G2 risk haplotype on the same exon also showed a significant association with HIVAN (Table 3.5). These SNPs had the highest association with HIVAN and no association with HIV (+) CKD or HIV (-) CKD using the logistic regression model adjusting for age and gender (Table 3.5). The covariate-adjusted case/control analyses considering the three genotypic modes of associations (dominant, recessive and additive) within the HIVAN group, showed significant genetic associations with individual SNPs for the two G1 (rs73885319 and rs60910145) and one G2 (rs71785313) variant allele under dominant and additive models of associations even after correcting for multiple comparisons (Table 3.7). The reason why we could not detect an association using logistic regression in a recessive model could be because half of the individuals (19/38) within the HIVAN group were compound heterozygotes (Table 3.6). Variants rs73885319 and rs60910145 had larger effects under the dominant mode, with the strongest effect being with rs73885319 (OR=123.10 95% CI 20.57-737.00 $P=5.99X10^{-10}$). This was followed by rs60910145 (OR= 58.25 95% CI 13.43-252.60 $P=2.92X\ 10^{-09}$) as shown in Table 3.7. These two variants are in almost perfect LD (D'=1.00, $r^2$=0.80) as shown in Figure 3.4A and therefore if adjusted for either one, the association of the other disappears. G2 association with HIVAN was OR=2.64 95% CI 1.05-6.65 $P=0.004$.

**Table 3.7**: Logistic regression analysis showing significant SNP Associations with 38 HIVAN cases and 54 HIV (+) controls

| | | | Risk Allele Frequencies | | | | *P Values* | | |
|---|---|---|---|---|---|---|---|---|---|
| **Gene** | **SNP** | **Risk Allele** | **Cases** | **Controls** | **Association Mode** | **OR (95% CI)** | **Unadjusted** | **Adjusted** | **Corrected*** |
| ***APOL1*** | rs73885319 | G | 0.5556 | 0.03704 | Dominant | 123.10 (20.57-737.00) | $5.99x10^{-10}$ | $1.35x10^{-07}$ | 0.0009 |
| | | | | | Recessive | Inf (0.00-Inf) | 0.9981 | 0.9979 | 1.0000 |
| | | | | | Additive | 98.56 (16.04-605.60) | $6.10x10^{-09}$ | $7.20x10^{-07}$ | 0.0009 |
| ***APOL1*** | rs60910145 | G | 0.4872 | 0.03704 | Dominant | 58.25 (13.43-252.60) | $2.915x10^{-09}$ | $5.61x10^{-08}$ | 0.0009 |
| | | | | | Recessive | Inf (0.00-Inf) | 0.9981 | 0.9980 | 1.0000 |
| | | | | | Additive | 46.08 (10.68-198.90) | $2.26x10^{-08}$ | $2.83x10^{-07}$ | 0.0009 |
| ***APOL1*** | rs71785313 | D | 0.3421 | 0.1667 | Dominant | 2.64 (1.05-6.65) | 0.0038 | 0.0399 | 0.2897 |
| | | | | | Recessive | 1.49 (0.22-10.02) | 0.3933 | 0.6861 | 1.0000 |
| | | | | | Additive | 2.08 (0.95-4.54) | 0.0076 | 0.0667 | 0.4665 |
| ***APOL1*** | rs7284919 | C | 0.0143 | 0.0938 | Dominant | 0.11 (0.01-1.00) | 0.0564 | 0.0500 | 0.3596 |
| | | | | | Recessive | NE | NE | NE | NE |
| | | | | | Additive | 0.11 (0.01-1.00) | 0.0564 | 0.0501 | 0.3586 |
| ***APOL1*** | rs136147 | T | 0.6806 | 0.3370 | Dominant | 20.73 (3.08-113.00) | 0.0004 | 0.0005 | 0.0029 |
| | | | | | Recessive | 3.05 (0.99-9.36) | 0.0148 | 0.0509 | 0.1039 |
| | | | | | Additive | 3.98 (1.79-8.85) | 0.0002 | 0.0007 | 0.0029 |
| ***APOL3*** | rs11089781 | A | 0.4286 | 0.1170 | Dominant | 6.43 (2.13-19.43) | 0.0002 | 0.0010 | 0.0060 |
| | | | | | Recessive | Inf (0.00-Inf) | 0.9981 | 0.9981 | 1.0000 |
| | | | | | Additive | 5.80 (2.11-15.96) | $9.71x10_{-05}$ | 0.0007 | 0.0029 |

Adjusted for age and gender, SNP: Single nucleotide polymorphism, OR: Odds ratio, CI: Confidence Interval, NE: Not Estimated, **P values* corrected for multiple comparisons

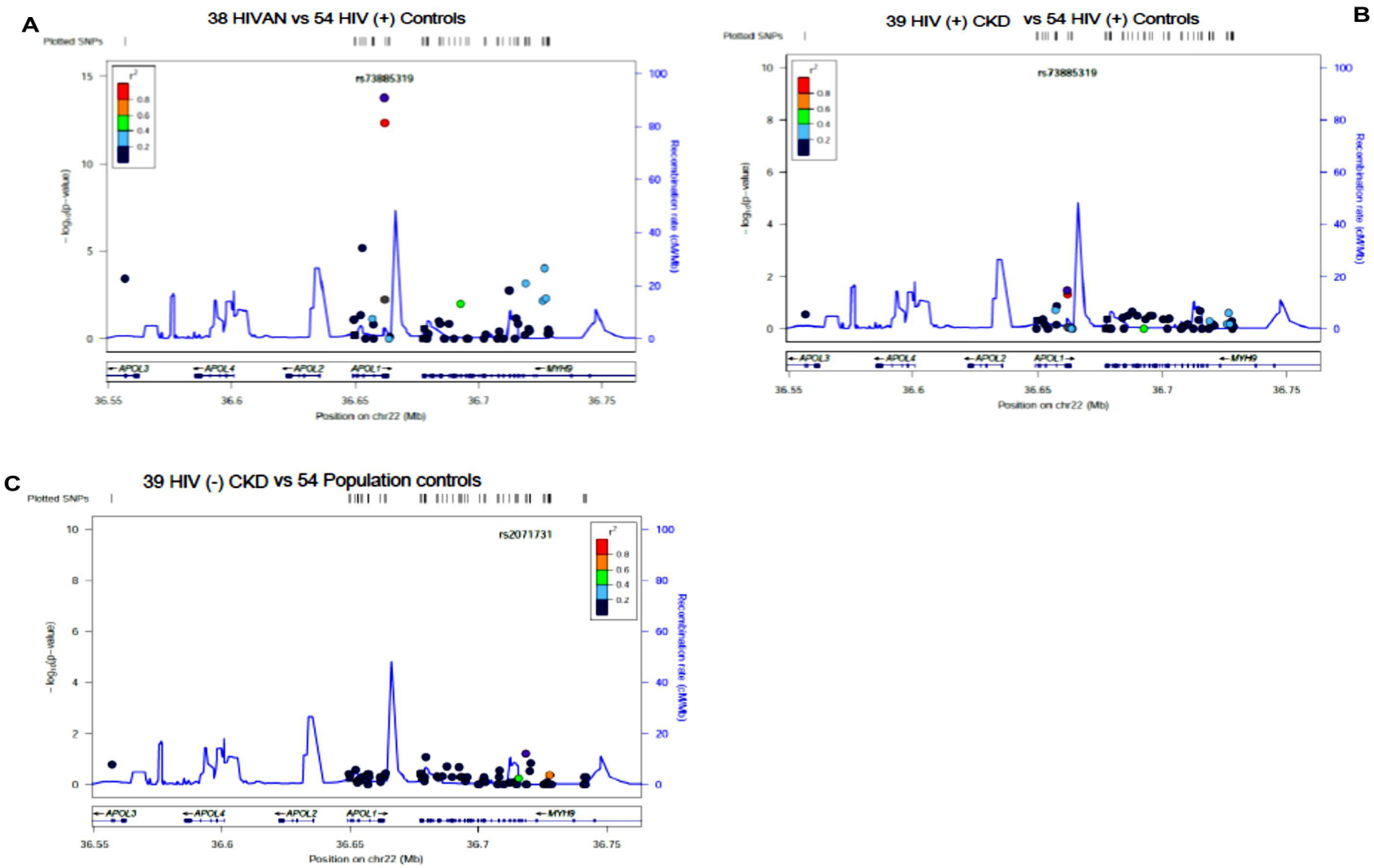



**Figure 3.4:** SNP association using logistic regression analysis.

**A:** SNP association between HIVAN cases vs HIV (+) Control **B:** SNP Association between HIV (+) CKD vs HIV (+) Controls. **C:** SNP Association between HIV (-) CKD cases vs Population Control. On the x and y –axes, genomic position and –log of the unadjusted *P values* are shown respectively.

The effect size and statistical significance for carriage of 1 or 2 *APOL1* risk variants by comparing the distribution of 2 vs. 1 or 0 copies of *APOL1* G1 and G2 risk alleles in cases and controls in a recessive model (Table 3.8) was assessed using Fisher's exact test by stratifying the analysis with the other allele adjusting for age and gender. There was no significant association with HIV (+) CKD and the HIV (-) CKD case groups. There was a marginally significant association when carrying one copy of either allele with HIVAN and because of the complete negative LD, an individual cannot carry more than two copies of G1 and/or G2. The association of *APOL1* risk alleles with HIVAN fitted best a recessive model using Fisher's exact test, consistent with other studies. Carriage of 2 copies of *APOL1* renal risk alleles was a strong predictor of HIVAN (OR=89.1 95% CI 17.68, 911.72 $P=1.24x10^{-14}$). This effect size was so strong that carriage of even 1 renal risk allele increased the risk of HIVAN with a considerable statistical support (OR 5.49 95% CI 0.87, 61.14 *P=0.05*) (Table 3.8). The results for HIVAN remained highly significant after adjusting for age, gender, and genetic ancestry (Table 3.8).

**Table 3.8**: *APOL1* risk allele associations with HIVAN, HIV-associated FSGS, and HIV (-) FSGS (primary FSGS)

| | **HIV (+) CKD Case Groups (N=51) vs. HIV (+) Controls (N=54)** | | | | | **Primary FSGS (N=9) vs. Population Controls (N=54)** | | **All FSGS (N=22) vs. All controls (N=108)** | |
|---|---|---|---|---|---|---|---|---|---|
| | **HIVAN N=38** | | | **HIV(+) FSGS N= 13** | | **Primary FSGS N=9** | | **HIV(+) FSGS (N=13) and primary FSGS (N=9)** | |
| **No. of risk alleles** | OR (95%CI) | $P_{FET}$ | $P_{adj*}$ | OR (95%CI) | $P_{FET}$ | OR (95%CI) | $P_{FET}$ | OR (95%CI) | $P_{FET}$ |
| **1 vs. 0** | 5.49 (0.87, 61.14) | 0.05 | 0.13 | 0.63 (0.10, 2.98) | 0.74 | 1.3 (0.18, 7.42) | 1.00 | 0.86 (0.25, 2.64) | 1.00 |
| **2 vs. 0** | 200.87 (27.62, 3119.17) | $1.5x10^{-14}$ | $1.9x10^{-3}$ | 1.85 (0.03, 39.57) | 0.53 | 6.70 (0.08, 577.00) | 0.26 | 3.27 (0.25, 31.46) | 0.22 |
| **2 vs. 1** | 40.50 (7.09, 446.94) | $8.7x10^{-08}$ | $1.8x10^{-4}$ | 2.82 (0.04, 73.20) | 0.44 | 5.0 (0.05, 467.90) | 0.34 | 3.74 (0.26, 40.67) | 0.20 |
| **2 vs. 1 or 0 Recessive Model** | **89.1 (17.68, 911.72)** | $1.2x10^{-14}$ | $1.5x10^{-5}$ | 2.13 (0.03, 44.30) | 0.48 | 6.30 (0.04, 248.70) | 0.26 | 3.45 (0.27, 32.23) | 0.20 |

Shown are odds ratios (OR) and confidence intervals (CI) for association between *APOL1* risk alleles and particular glomerular diseases. Analyses were adjusted for age, sex, and ancestry. $P_{FET:}$ P values from Fisher's exact test. Permission obtained from the American Society of Nephrology publishing group (Journal of the American Society of Nephrology. (Permission obtained from the American Society of Nephrology publishing group). Appendix 15 F

We also showed the distribution of *APOL1* G1 and G2 risk alleles among the different histologically diagnosed kidney phenotypes HIV (+) and HIV (-) CKD cases in Table 3.9. Among the HIV (+) CKD, there was no significant association between *APOL1* risk alleles and "HIV (+) FSGS (OR=2.13 95% CI 0.03, 44.30 *P=0.48*) and HIVICK (OR=5.60 95% CI 0.4, 86 *P=0.13*) compared to HIV (+) controls". Among the HIV (-) CKD, there was also no significant association for "primary FSGS (OR=6.30 95% CI 0.04, 248.70 *P=0.26*) compared to population controls" (Table 3.8). We combined HIV (+) FSGS (N=13) and primary FSGS (N=9) and compared this to all controls (N=108) to increase statistical power, and still no significant association with *APOL1* G1 and G2 alleles in a recessive model (OR=3.45 95% CI 0.27, 32.23 *P=0.20*) was observed (Table 3.8). We used a Woolf test for homogeneity to determine if the odds ratios for *APOL1* association with primary FSGS and HIVAN from our study were significantly different to odds ratios from those reported in Kopp et al (Kopp et al., 2011) since the point odds ratios differed between the two studies. The odds ratios in our study were higher for HIVAN and lower for FSGS compared to the ratios observed in African Americans. However, the differences were not statistically significant (*P=0.56* and *0.21* for FSGS and HIVAN respectively). "We also compared odds ratios for HIV (+) FSGS seen in this study to data from Fine et al (Fine et al., 2012), for African Americans with HIV (+) FSGS compared to a population control group from Kopp et al.(Kopp *et al*., 2011) (OR = 24.5 95% CI 9.7-65.1 *P= 4.3X10*$^{-14}$)". Woolf test gave a *P* value of 0.067 indicating that there was no significant association. The effect sizes (odds ratios) of G1 and G2 alleles shown in Figure 3.5 clearly indicated the effect size of G2 was weaker than G1. However, we should note that the effect G1/G2 compound heterozygotes is equivalent in effect size to G1/G1.

**Table 3.9:** Association between *APOL1* risk alleles and various glomerular diseases

| | **HIV positive cases and controls** | | | | | **HIV negative cases and controls** | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | **Number of *APOL1* risk alleles** | | | **OR (95% CI)** | **P $^{FET}$** | **Number of *APOL1* risk alleles** | | | **OR (95% CI)** | **P $^{FET}$** |
| ***APOL1* risk allele distribution** | **0** | **1** | **2** | **2 vs 1 or 0 risk alleles** | **P $^{FET}$** | **0** | **1** | **2** | **2 vs 1 or 0 risk alleles** | **P $^{FET}$** |
| **Controls (N=108)** | 34 (63%) | 18 (33%) | 2 (4%) | - | - | 36 (67%) | 17 (32%) | 1 (2%) | - | - |
| **HIVAN (N=38)** | 2 (5%) | 6 (16%) | 30 (79%) | **89 (17.7, 912)** | **1.2 X $10^{-14}$** | - | - | - | - | - |
| **Other CKD (N=78)** | 22 (57%) | 11 (28%) | 6 (15%) | 3.8 (0.6, 42) | 0.13 | 25 (64%) | 13 (33%) | 1(3%) | 1.4 (0.02, 11) | 1.0 |
| **FSGS (N=22)** | 9 (69%) | 3 (23%) | 1 (8%) | 2.1 (0.03, 44) | 0.48 | 5 (56%) | 3 (33%) | 1 (11%) | 6 .3 (0.08, 527) | 0.26 |
| **HIVICK (N=12)** | 4 (33%) | 5 (42%) | 3 (25%) | 5.6 (0.4, 86) | 0.13 | - | - | - | - | - |
| **Other glomerulonephritis (N=27)** | 7 (70%) | 3 (30%) | 0 (0%) | 0.0 (0, 30) | 1.0 | 10 (59%) | 7 (41%) | 0 (0%) | 0 (0, 124) | 1.0 |
| **Other kidney diseases (N=17)** | 2 (50%) | 0 (0%) | 2 (50%) | 21 (0.2, 2029) | 0.11 | 10 (77%) | 3 (23%) | 0 (0%) | 0 (0, 210) | 1.0 |

*APOL1* genotype frequencies and associations, tested with the Fisher exact test (FET), for various glomerular diseases among HIV+ and HIV- subjects and general population controls. The only glomerular disease that showed a significant association with *APOL1* risk alleles was HIVAN. Of note, HIV+ FSGS, HIV- FSGS, and HIV immune complex disease (HIVICK) were not significantly associated with *APOL1* risk alleles, although group sizes were small. Analyses were adjusted for age, sex, and ancestry. $P^{FET}$ denotes probability by Fisher exact test. **Other Glomerulonephritis** (HIV positive; Membranoproliferative glomerulonephritis N=1, Membranoproliferative glomerulonephritis consistent with C3 glomerulopathy N=1, Membranous glomerulonephritis N=8. HIV negative; IgA nephropathy N=1, Lupus nephritis N=3, Membranoproliferative glomerulonephritis N=2, Membranous glomerulonephritis N=11). **Other kidney diseases** (HIV positive; Benign nephrosclerosis N=1, Global glomerulosclerosis N=1, Minimal Change disease N=1, Thrombotic microangiopathy N=1. HIV negative; Benign nephrosclerosis N=1, Oxalosis N=1, Global glomerulosclerosis N=2, Minimal Change disease N=5, Nodular Glomerulosclerosis N=1, Severe arterial nephrosclerosis (ANS) with secondary FSGS N=1, Thrombotic microangiopathy N=2). Permission obtained from the American Society of Nephrology publishing group (Journal of the American Society of Nephrology. (Permission obtained from the American Society of Nephrology publishing group). Appendix 15 F

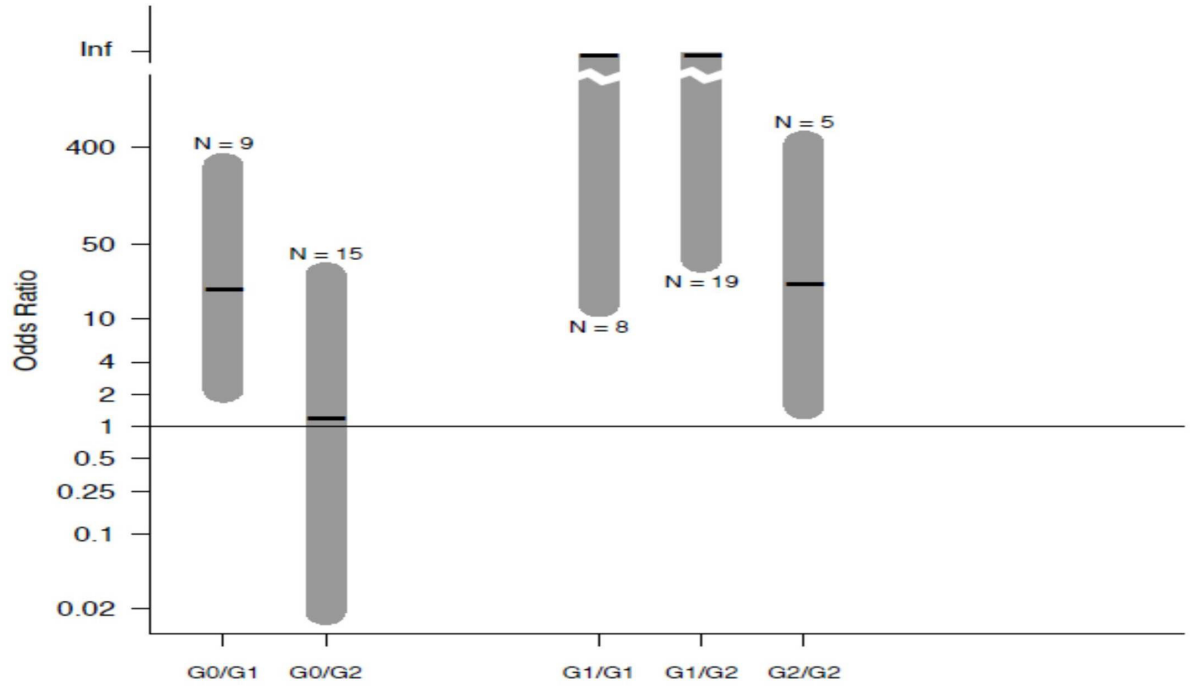



**Figure 3.5:** Odds ratios for the effect sizes for HIVAN.
Odds ratios with confidence intervals for association of different strata of G1-G2 genotypes with HIVAN cases compared with subjects carrying no G1 or G2 risk alleles (+/+). For G1/G1 and G1/G2, the odds ratio is infinity, since these genotypes were only observed in cases. Permission obtained from the American Society of Nephrology publishing group (Journal of the American Society of Nephrology. (Permission obtained from the American Society of Nephrology publishing group). Appendix 15 F

## 3.5 Combined haplotype analysis of the *MYH9* and *APOL1* SNPs

Haplotype analysis is the type of analysis that tests for the closely linked SNPs in LD within the genes or chromosome region. In order to validate the effect of the *MYH9* E1 haplotype SNPs (rs4821480, rs4821481 and rs3752462) since they showed no individual significant association, we included them together with SNPs that showed individual significant association with HIVAN in a combined haplotype analysis of the *MYH9* and *APOL1*. In total, there were 15 SNPs from both *MYH9* and *APOL1* gene regions that were involved in combined haplotype analysis (Tables 3.4 and 3.7). There were a total of twelve haplotypes each with a frequency of at least 0.10 (Figure 3.6). Five SNPs (*APOL3*-rs11089781, *APOL1*G2-rs71785313, *MYH9*-rs3752462,

rs16996668 and rs5756152) were not included in any of the blocks. Therefore, there were three blocks, one block consisting of four *APOL1* SNPs and two blocks consisting of three *MYH9* SNPs each. The results of the covariate-adjusted haplotype associations are presented in Table 3.10 showing haplotype frequencies which range from 0.012 to 0.676 in the study sample. The association results indicated that T-T-G-G haplotype consisting of *APOL1* SNPs rs7284919-rs136147-rs73885319-rs60910145, showed the highest significant risk factor for HIVAN (OR= 52.20 $P=2.75x10^{-06}$) after correcting for multiple comparisons. This haplotype consists of two G1 risk alleles and this could explain its high association with HIVAN since has been shown in Figure 3.4 where the G1 risk allele had the highest association with HIVAN and also using logistic regression models in both dominant and additive association models (Table 3.7). The T-T-G-T haplotype of the same *APOL1* SNPs rs7284919-rs136147-rs73885319-rs60910145, indicated significant protective association with HIVAN (OR=3.30 *P=0.999*). This shows that, as much as there is no significant high risk for HIVAN for individuals carrying the zero or one risk allele G1: T-T-G-T, there might be a potential of developing HIVAN considering the odds ratio of 3.30. After accounting for multiple testing, we observed a significant association of haplotype block C-G-C association consisting of *MYH9* SNPs rs16996648-rs4821480-rs4821481, two of which are E1 haplotype SNPs, showed a significant risk to HIVAN (OR=3.45 *P=0.003*) as shown in Table 3.10. We also observed haplotype C-C-G of *MYH9* SNPs rs16996672-16996674-16996677, indicated significant protective association with HIVAN (OR=0.25 *P=0.0007*). However, haplotype T-T-A of the same SNPs, indicated a high association for HIVAN (OR=5.87 $P=6.81x10^{-05}$). The implication is that, there is a significantly high risk for HIVAN for those individual carrying risk alleles, T-T-A (Table 3.10). These *MYH9*

SNPs (rs16996672-16996674-16996677) do not pose a significant threat since, they do not form part of the risk haplotype, hence not predisposing to kidney disease.

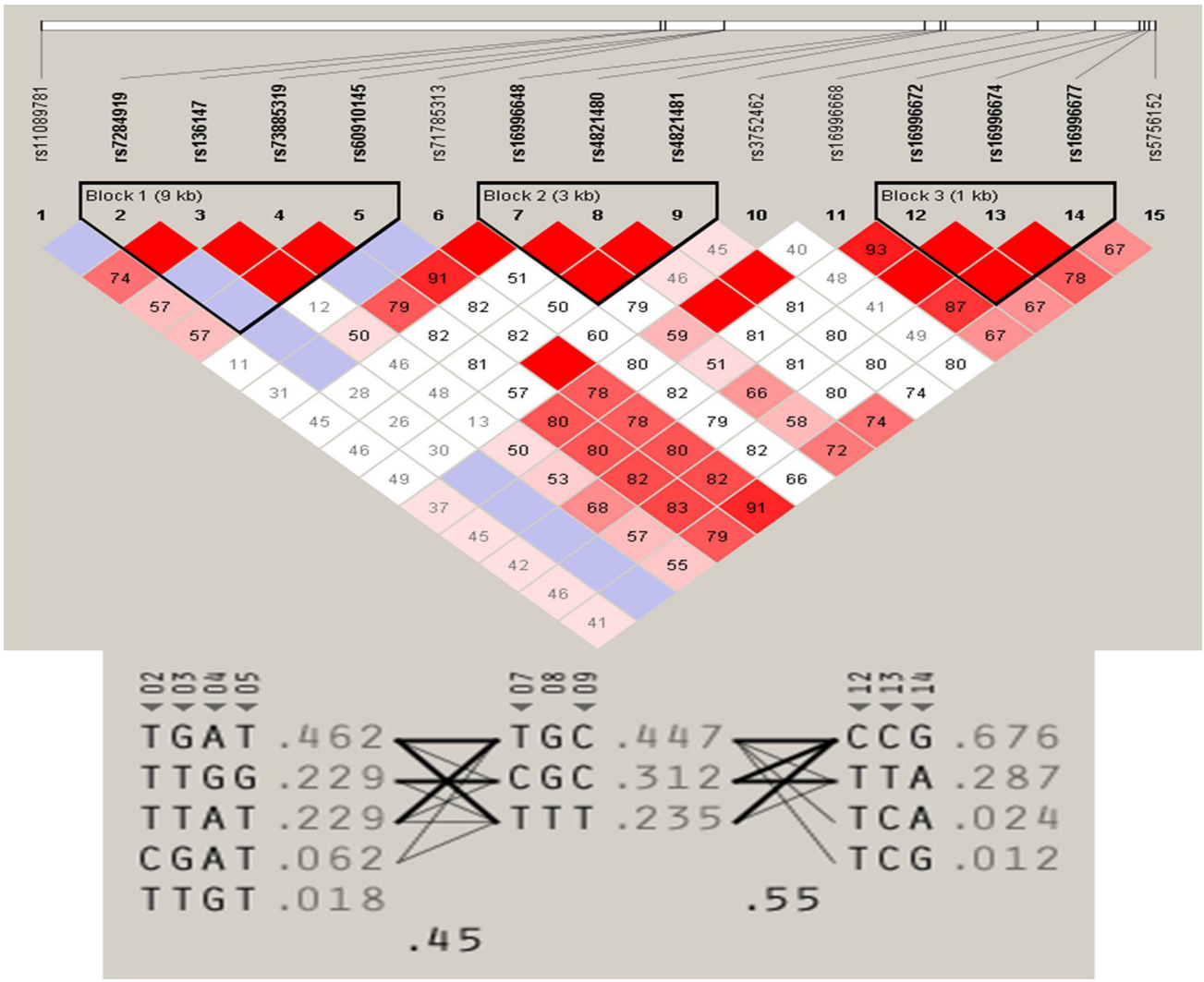



**Figure 3.6:** Gene structure and haploview linkage disequilibrium plot**.**
Generated by 15 SNPs associated with HIVAN from the *APOL3*, *APOL1* and *MYH9* gene regions (at the top) and their haplotypes (below). This is an indication of the position of the SNPs along with the diagram of the block structure of the region. The LD blocks were identified with a strong LD. Red and Pink colour indicates the computed pairwise D' value; deeper pink indicates a higher D' value. The blue blocks indicate a higher LD but a lower LOD values. Inter SNP D' values are displayed on the plot.

**Table 3.10:** Combined haplotype association with 38 HIVAN and 54 HIV (+) Controls.

| **Haplotype Blocks** | | | | **Haplotypes** | **Haplotype Frequencies** | | | | ***P values*** | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| ***APOL1* SNPs** | | | | | **Cases** | **Controls** | **All** | **OR** | **Unadjusted** | **Adjusted** | **Corrected**** |
| **rs7284919** | **rs136147** | **rs73885319** | **rs60910145** | **T-G-A-T** | **0.302** | **0.579** | **0.462** | **0.265** | **0.005** | **0.007** | **0.005** |
| **rs7284919** | **rs136147** | **rs73885319** | **rs60910145** | **T-T-G-G** | **0.486** | **0.041** | **0.229** | **52.20** | **$1.76x10^{-07}$** | **$2.75x10^{-06}$** | **0.001** |
| rs7284919 | rs136147 | rs73885319 | rs60910145 | T-T-A-T | 0.153 | 0.286 | 0.229 | 0.385 | 0.039 | 0.031 | 0.133 |
| rs7284919 | rs136147 | rs73885319 | rs60910145 | C-G-A-T | 0.017 | 0.094 | 0.062 | 0.14 | 0.055 | 0.060 | 0.245 |
| rs7284919 | rs136147 | rs73885319 | rs60910145 | T-T-G-T | 0.042 | 0.000 | 0.018 | 3.30 | 0.999 | 0.999 | 1.000 |
| ***MYH9* SNPs** | | | | | | | | | | | |
| rs16996648 | rs4821480 | rs4821481 | | T-G-C | 0.377 | 0.498 | 0.447 | 0.53 | 0.079 | 0.018 | 0.222 |
| **rs16996648** | **rs4821480** | **rs4821481** | | **C-G-C** | **0.442** | **0.217** | **0.312** | **3.45** | **0.002** | **0.003** | **0.008** |
| rs16996648 | rs4821480 | rs4821481 | | T-T-T | 0.167 | 0.258 | 0.235 | 0.47 | 0.059 | 0.079 | 0.219 |
| **rs16996672** | **rs16996674** | **rs16996677** | | **C-C-G** | **0.528** | **0.786** | **0.676** | **0.25** | **0.0006** | **0.0007** | **0.001** |
| **rs16996672** | **rs16996674** | **rs16996677** | | **T-T-A** | **0.472** | **0.151** | **0.287** | **5.87** | **$4.73x10^{-05}$** | **$6.81x10^{-05}$** | **0.001** |
| rs16996672 | rs16996674 | rs16996677 | | T-C-A | 0.000 | 0.042 | 0.024 | 0.46 | 0.994 | 0.994 | 1.000 |
| rs16996672 | rs16996674 | rs16996677 | | T-C-G | 0.000 | 0.021 | 0.012 | 7.40 | 0.999 | 0.999 | 1.000 |

Adjusted for age and gender, SNP: Single nucleotide polymorphism, OR: Odds ratio, CI: Confidence Interval, ***P values* corrected for multiple comparisons, rs4821480 and rs4821481 SNPs (*MYH9* E1 risk haplotype)

## 3.6 Population stratification

We performed population stratification analyses using the HIVAN cases and HIV (+) controls in order to make sure that the high association of *APOL1* risk alleles was as a result of genetic variation in the gene and not as a result of population substructure. Population stratification analyses using HIVAN and HIV (+) controls showed that the distribution of HIVAN cases and HIV (+) controls group was not significantly different along the top two Eigen axes ($P > 0.05$) as shown in **Figure 3.7,** hence confirming that the high association of *APOL1* risk alleles with HIVAN was as a result of genetic variation in the gene and not population substructure.

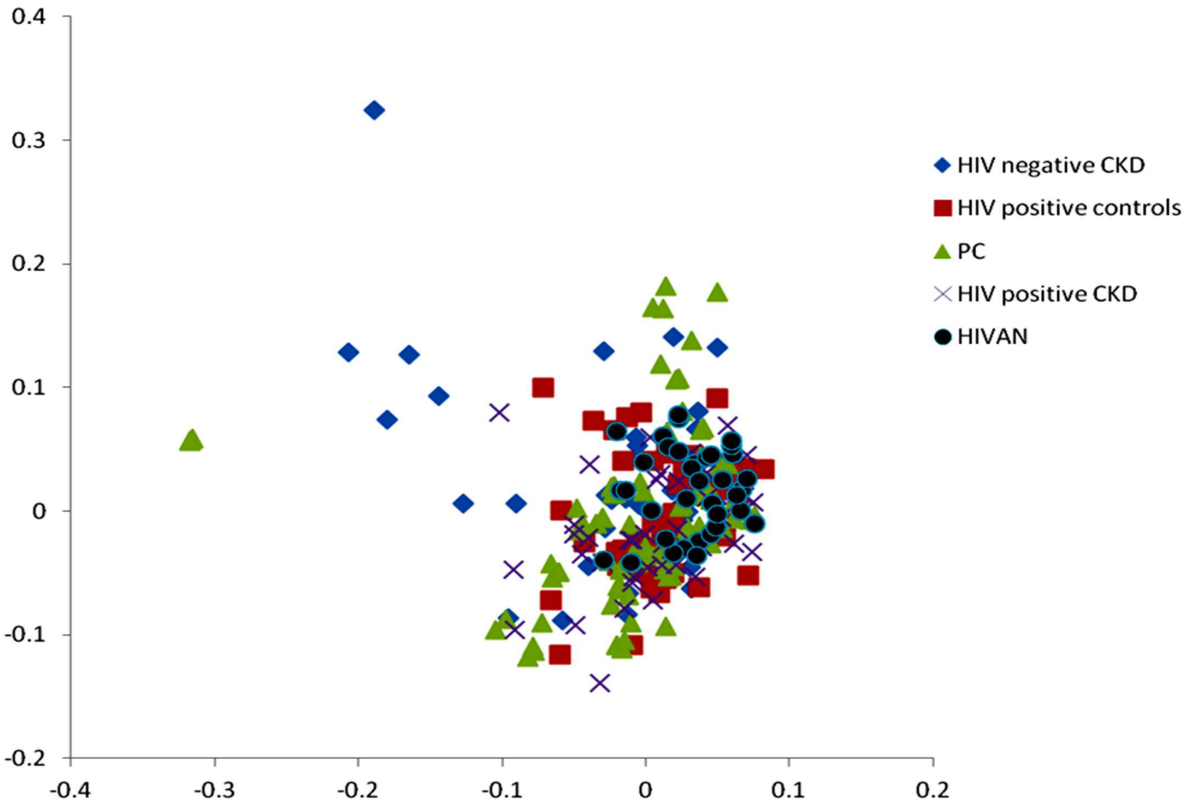
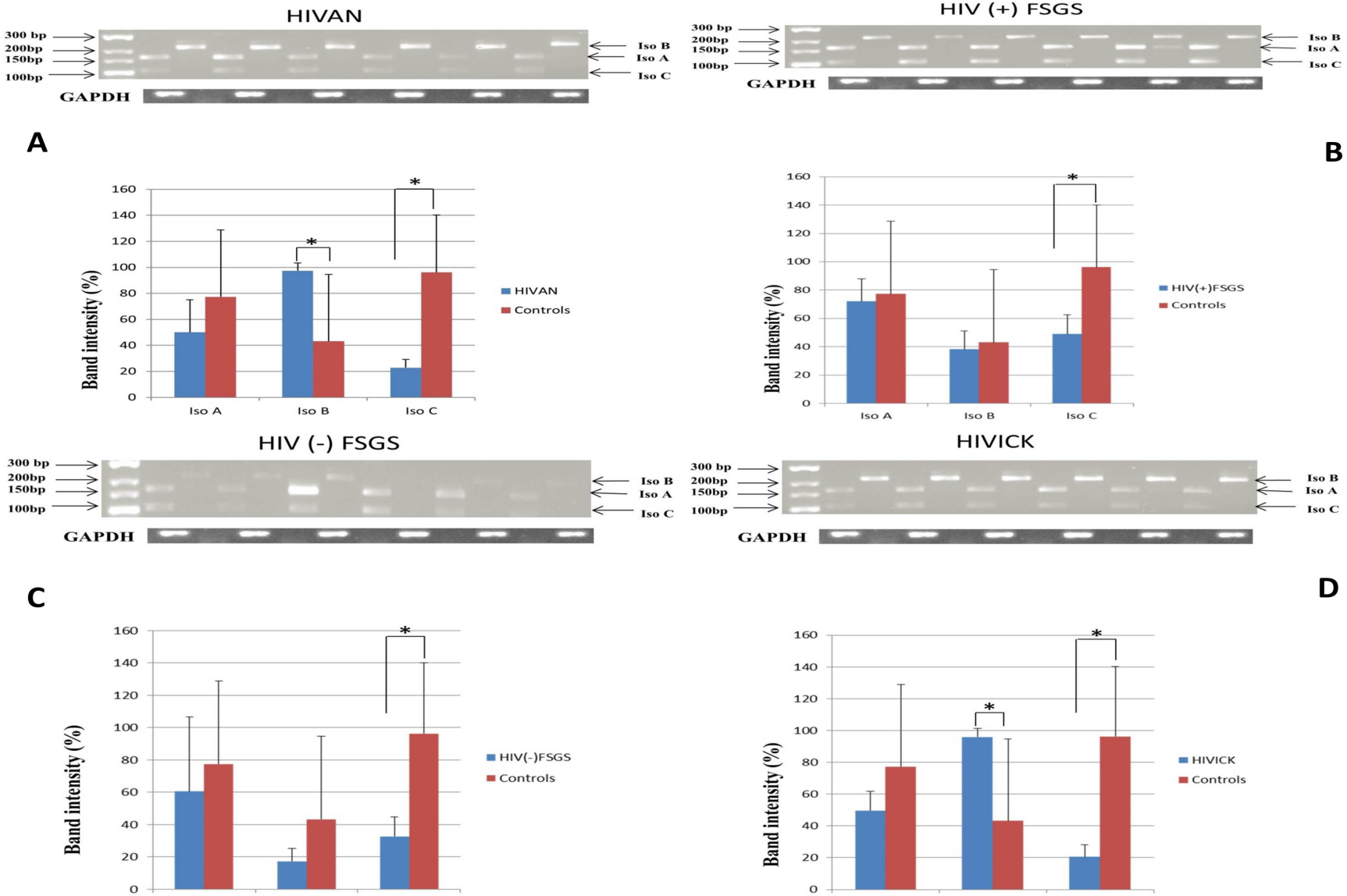



**Figure 3.7:** Principal component analysis of all study participants along the top two eigen values.
The distribution of HIVAN cases and HIV (+) control groups was not significantly different along the top two eigen axes ($P > 0.05$) suggesting that the high association observed is not due to an undetected population substructure. PC (Population Controls). HIVAN (HIV-associated Nephropathy). CKD (Chronic kidney Disease). (Permission obtained from the American Society of Nephrology publishing group). Appendix 15 F

## 3.7 mRNA expression profiles of N-terminal-apolipoprotein L1 isoforms

### 3.7.1 N-terminal *APOL1* mRNA isoform expression patterns

The N-terminal-apolipoprotein L1 mRNA isoform expression was determined using the RT-PCR method. Isoform A (Iso A) amplified as a 177bp product, isoform B (Iso B) a 225bp product and isoform C (Iso C), 123bps. The N-terminal-apolipoprotein L1 expression differed among kidney disease phenotypes and controls. All isoforms were expressed in all the cases and controls although the pattern of expression appeared to be different. There was no significant difference in expression of Iso A between cases and controls (Figure 3.8). However, there was a notable highly significant expression Iso B in HIVAN (*P=0.012*) and HIVICK (*P=0.010*) cases compared to controls (Figure 3.8A and D). Iso C N-terminal mRNA levels were elevated in normal kidney compared to cases. There was no correlation in terms of the mRNA N-terminal isoform expression and the number of risk alleles, since individuals with zero risk alleles also expressed all the three mRNA isoforms.


HIVAN
300 bp
200bp
150bp
100bp
Iso B
Iso A
Iso C
GAPDH
A
Band intensity (%)
HIVAN
Controls
Iso A
Iso B
Iso C
HIV (+) FSGS
B
HIV(+)FSGS
HIV (-) FSGS
C
HIV(-)FSGS
HIVICK
D

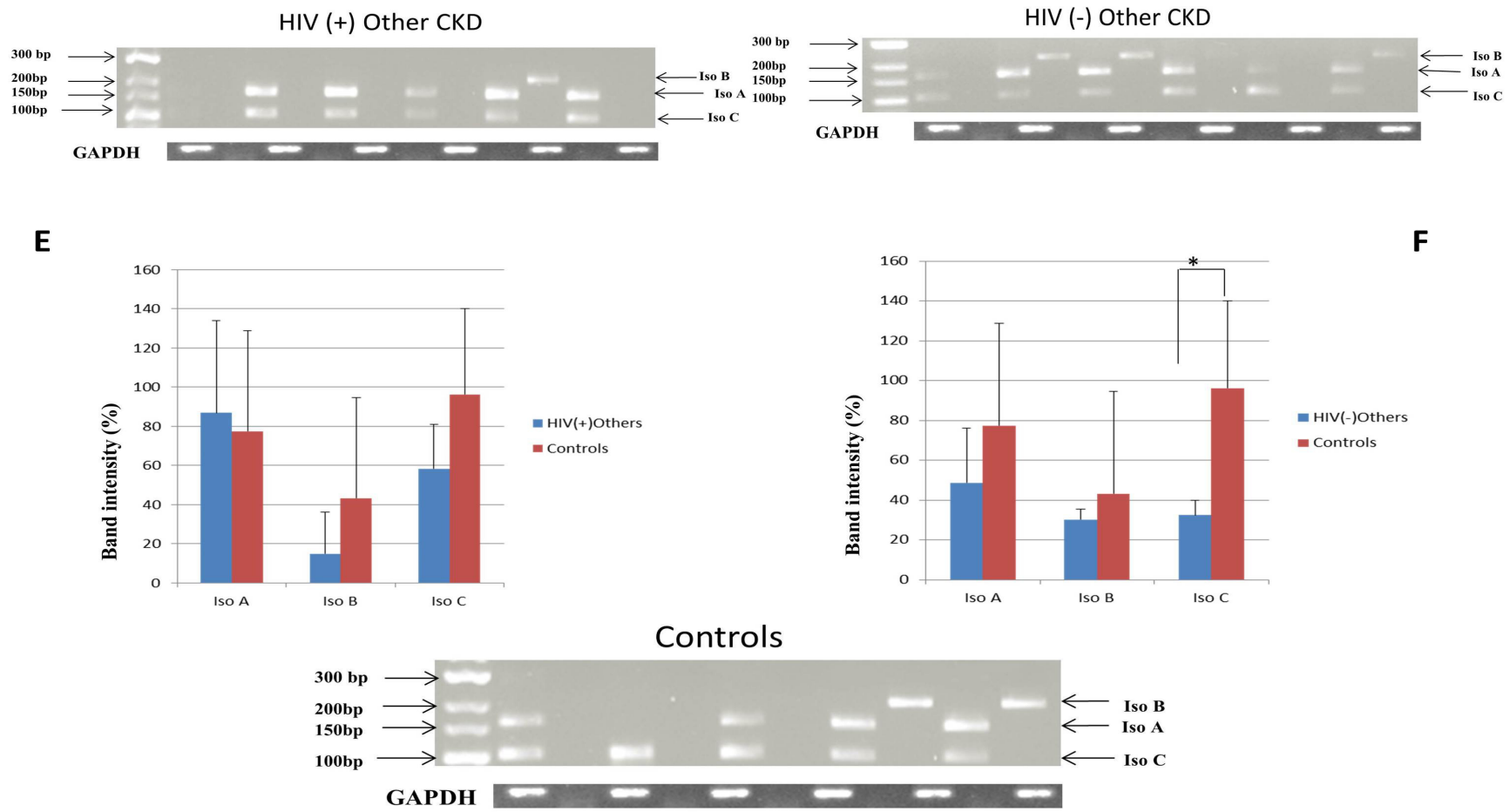



**Figure 3.8:** mRNA expression of the N-terminal-apolipoprotein L1 in kidney disease phenotypes.
**(A)** The band intensity values for Iso B mRNA at 177bp significantly increased in HIVAN compared to controls (n=6, *P=0.012*). There was no significant elevation of mRNA expression in **(B)** HIV (+) FSGS and in **(C)** HIV (-) FSGS. **(D)** The band intensity values for Iso B mRNA was significantly increased in HIVICK compared to controls (n=6, *P=0.010)*. There was no significant elevation of mRNA expression in **(E)** HIV (+) other CKD and in **(F)** HIV (-) other CKD. Iso A; Isoform A. Iso B; Isoform B. Iso C; Isoform C. * indicates that there is significant difference in mRNA expression. HIV (+) other CKD and HIV (-) other CKD; see Table 2.5 for detailed diagnosis and genotype of each individual. GAPDH was used as a reference gene.

## 3.8 Expression of the N-terminal- apolipoprotein L1 isoforms in kidney cells

### 3.8.1 Screening of clones expressing *N-terminal-ApoL1* isoforms in *pJET 1.2/blunt* vector

We examined the expression of N-terminal-*ApoL1* isoforms in HEK 293 cells by creating a series of isoform-DsRed2-N1 fusion constructs using a stepwise PCR-based cloning approach. To achieve this all the *N-terminal-ApoL1* isoforms PCR products were initially successfully cloned into the *pJET1.2/blunt* vector. *N-terminal-ApoL1* isoforms A, B and C were successfully cloned as shown in Figure 3.9. Colony PCR using *pJET* specific primers was used to identify clones containing the correct size of inserted isoform PCR product: Isoform A (294bp), isoform B (342bp) and isoform C (240bp) from the colony PCR (Figure 3.10).

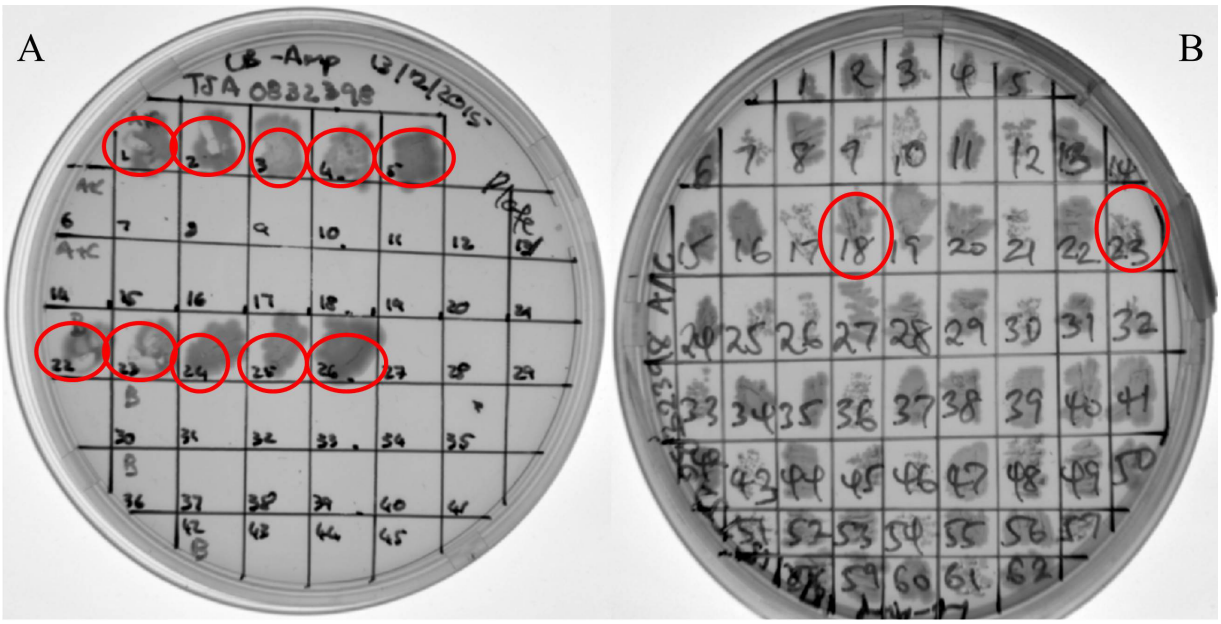



**Figure 3.9:** Colonies grown on LB-Ampincilin agar plate. Positive colonies are circled in red. N-terminal-*ApoL1* Iso A and B colonies are shown on plate **(A)** while N-terminal-*ApoL1* Iso C colonies are shown on plate **(B).**

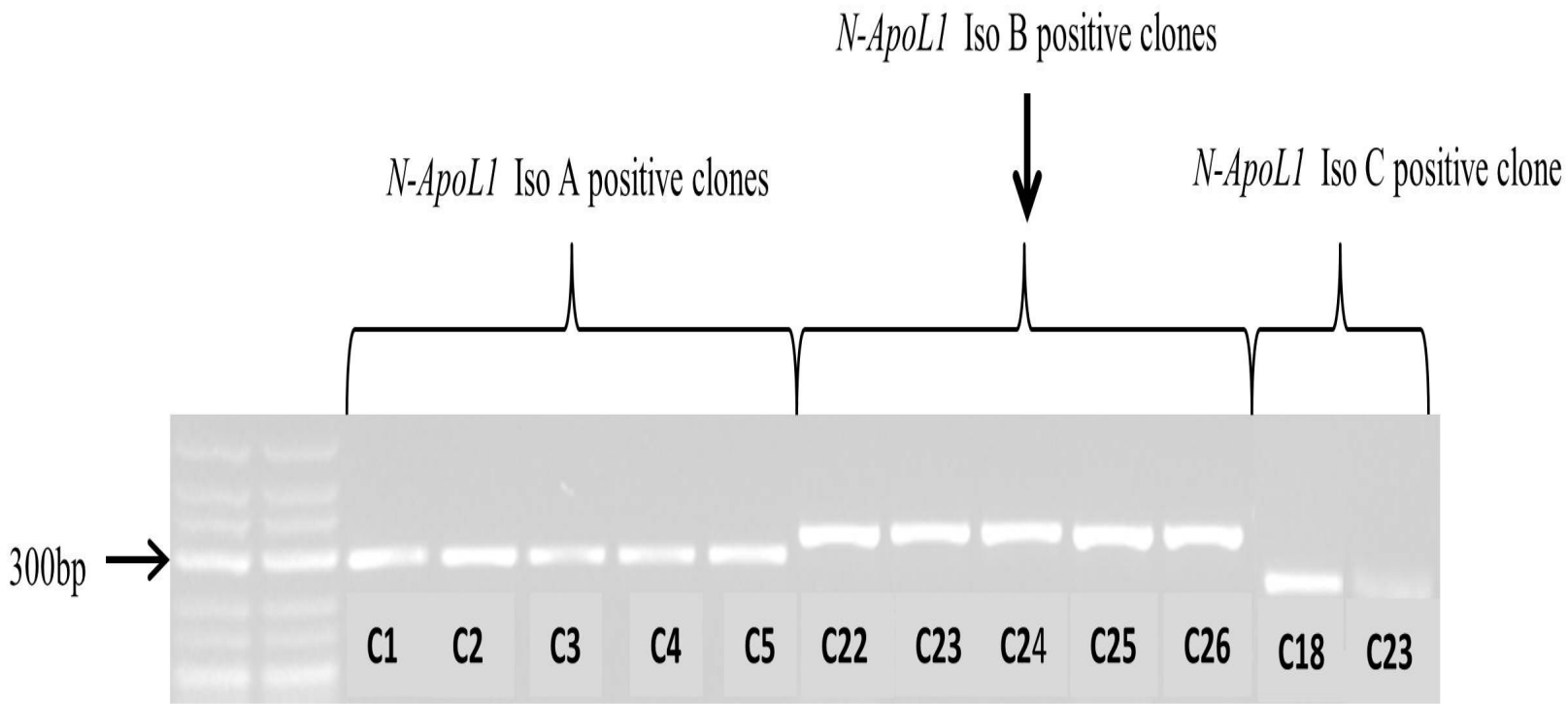



**Figure 3.10:** Positive clones from pJet1.2/blunt vector of the N-terminal-apolipoprotein L1 isoforms A, B and C.
C: Colonies

One positive colony of each isoform was grown overnight on LB agar (Appendix 10 B) with Amp100 at 37$^{o}$C for plasmid propagation thereafter grown in LB media (Appendix 10 C) overnight at 37$^{o}$C in preparation for plasmid extraction.

Purified plasmids were subjected to Sanger sequencing (Figure 3.11) to verify the successful cloning of Isoform A, B and C in the *pJET 1.2/blunt* vector. The recombinant plasmids were then digested with BglII restriction enzyme in order to restrict the isoform inserts from the vector in preparation for the ligation into the linearized, phosphorylated *pDsred2-N1* vector.

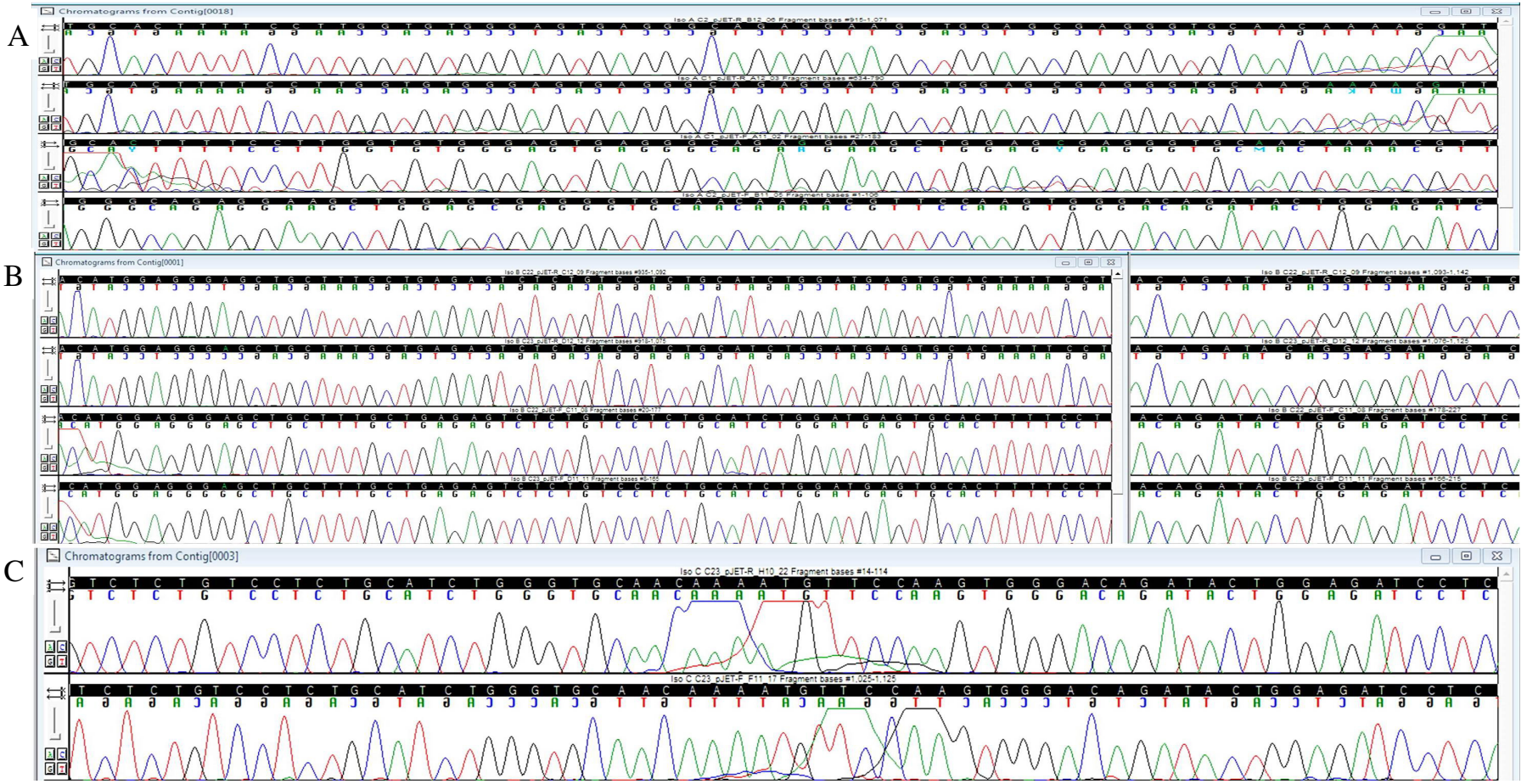



**Figure 3.11:** Sanger sequencing to verify successful cloning of isoforms A, B and C.
**(A)** Nucleotide sequence for *N-terminal-ApoL1*Isoform A inserts. **(B)** The nucleotide sequence for *N-terminal-ApoL1*Isoform B inserts. **(C)** The nucleotide sequence for the *N-terminal-ApoL1*Isoform C insert.

### 3.8.2 Screening of clones expressing N-terminal-*ApoL1* isoforms in *pDsRed2-N1* vector

Colony PCR using N-terminal-*APOL1* gene-specific primers amplified bands of the predicted size for each isoforms [Isoform A (177bp), isoform B (225bp) and isoform C (123bp)]. The following figures show the successful PCR cloning of the *N-terminal-ApoL1* isoforms into *pDsred2-N1* vector:

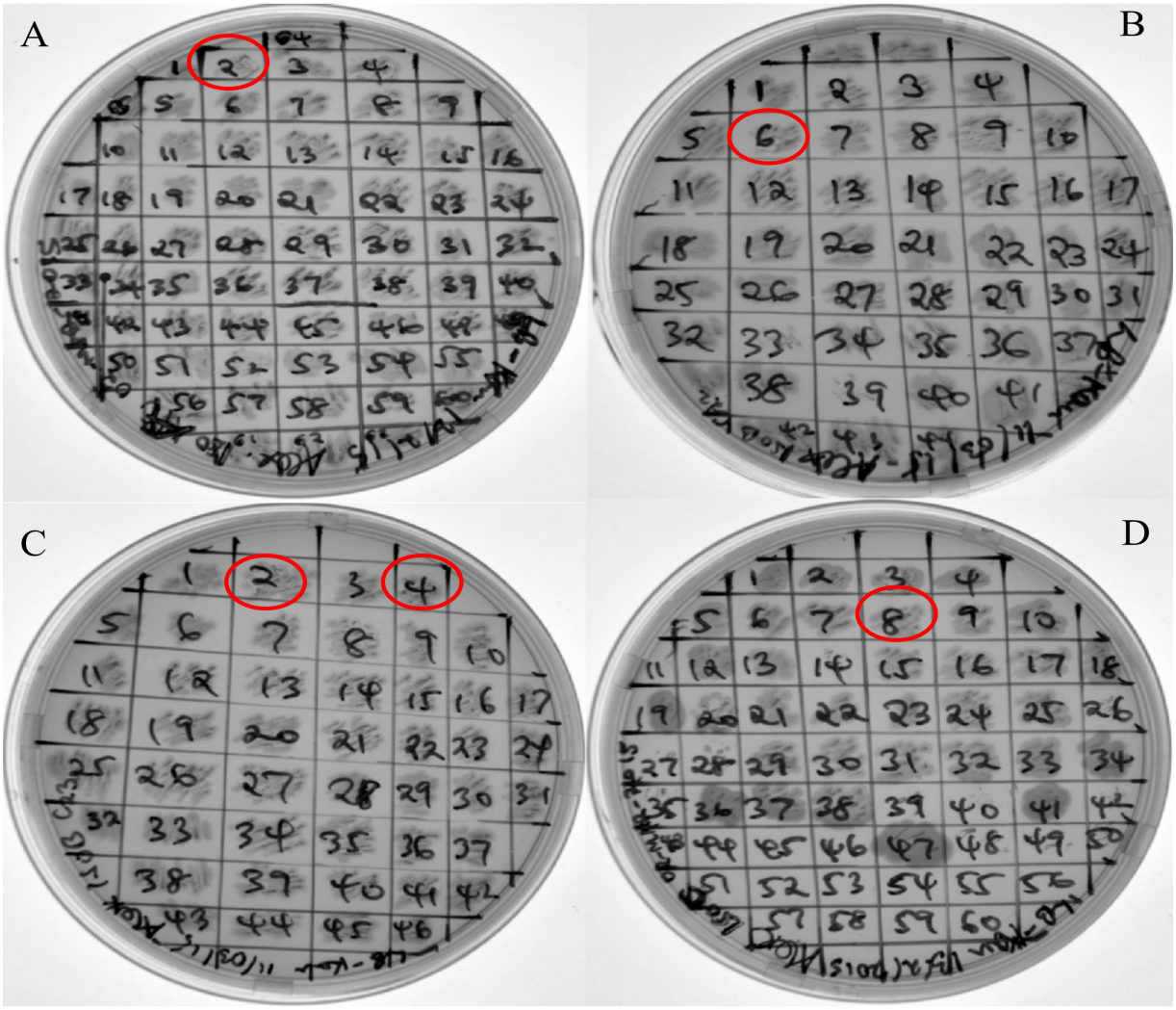



**Figure 3.12:** Colonies grown on LB-Kanamycin agar plate. Positive colonies are circled in red. *N-terminal-ApoL1* Iso A colonies are shown on plate **(A)**, *N-terminal-ApoL1* Iso B colonies are shown on plates **(B)** and **(C)** while *N-terminal-ApoL1* Iso C colonies are shown on plate **(D)**.

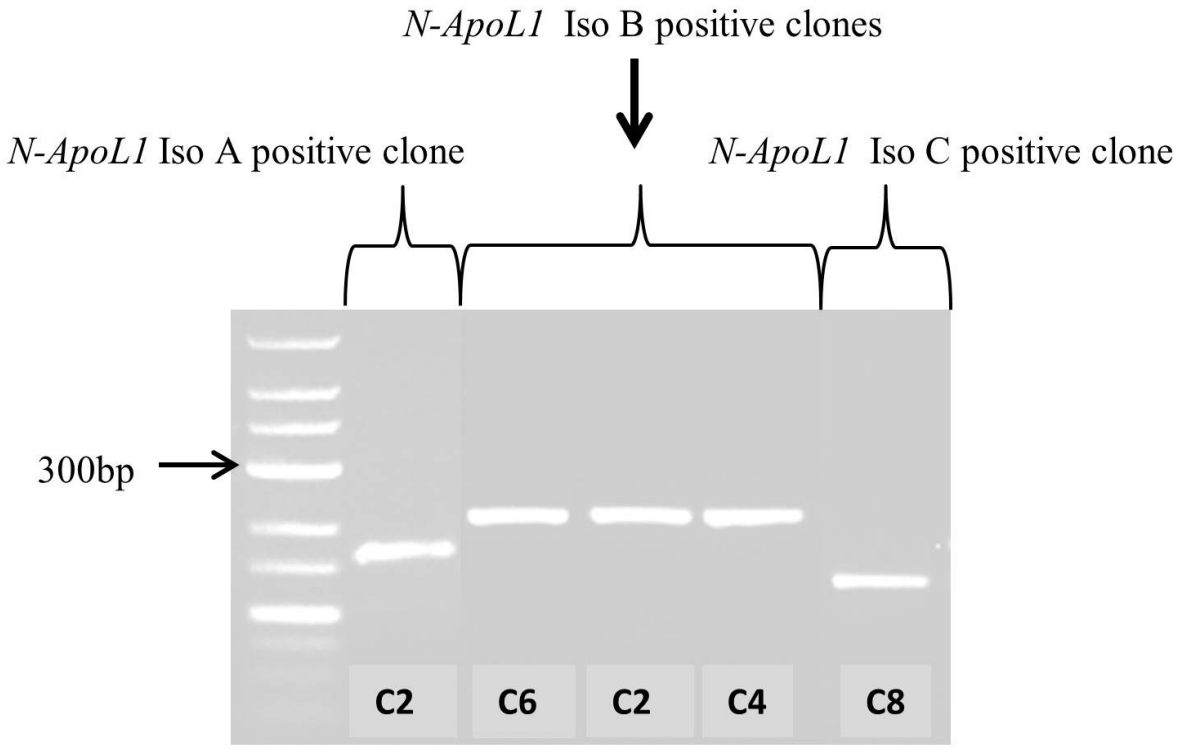



**Figure 3.13:** PCR amplification of positive clones in *pDsred2-N1*containing the N-terminal-apolipoprotein L1 isoforms A, B and C.
C: Colonies

Only one plasmid expressing each *N-terminal-ApoL1* isoform was purified according to "Plasmid DNA extraction protocol" in Appendix 11 and used for transfection of 80% confluent HEK 293 cells.

### 3.8.3 Transfected HEK 293 cells

The seeded HEK 293 cells (125,000cells/ml) were transfected with each plasmid expressing the *N-terminal-ApoL1* isoforms and after 72 hour incubation at 37$^{o}$C, we viewed the transgene expression of the different *N-terminal-ApoL1* isoforms in cells using confocal microscope (Carl Zeiss cell observer CD). There were more cells expressing *N-terminal-ApoL1* isoforms B and C compared to *N-terminal-ApoL1* isoform A (Figure 3.14) after 72 hours post transfection.

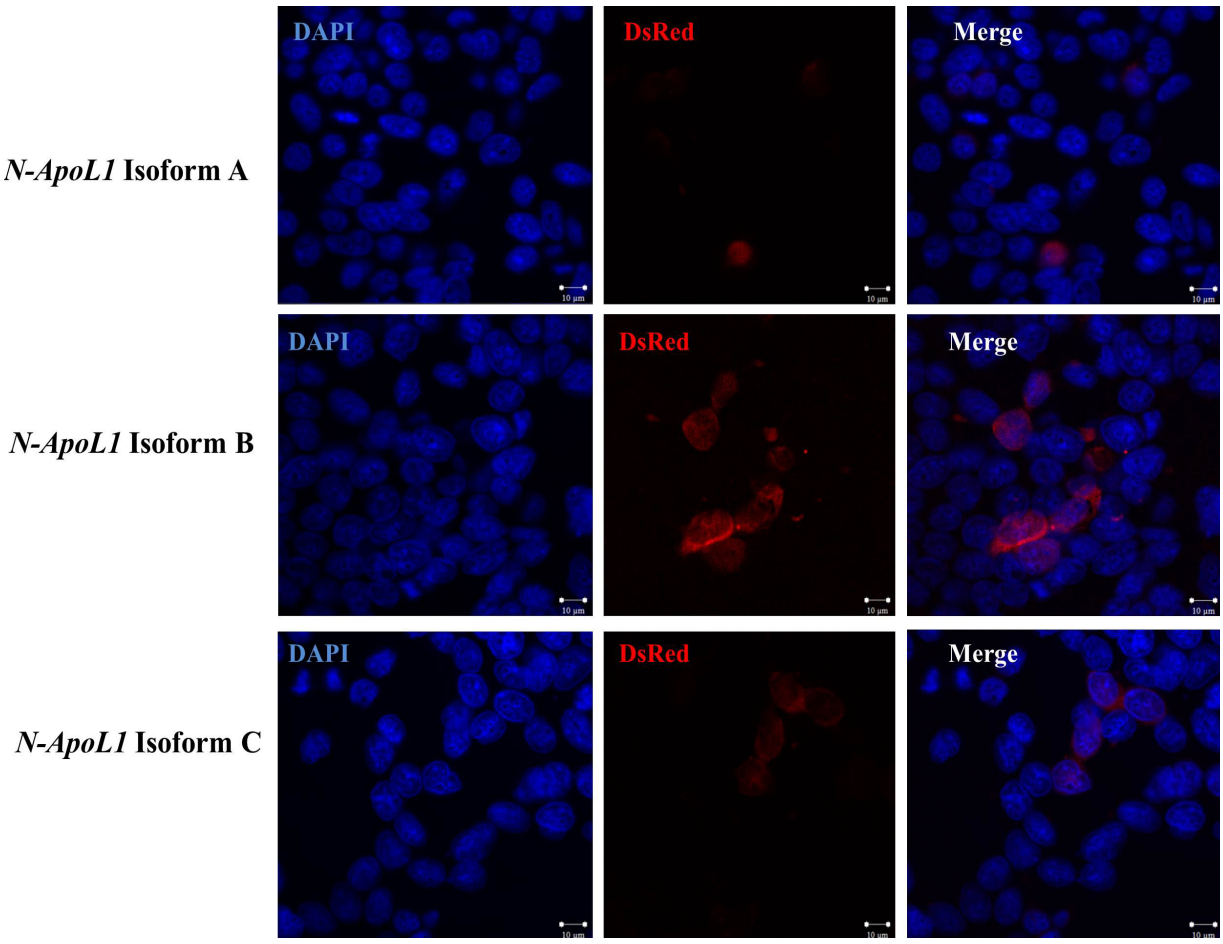



**Figure 3.14:** N-terminal apolipoprotein L1 (*ApoL1*) expression in HEK 293 cells.
There were more cells that expressed *N-terminal ApoL1* isoform B and C compared to isoforms A. The cells were stained with DAPI (4', 6-diamidino-2-phenylindole).

# Chapter 4

## 4) DISCUSSION

High risk *APOL1* variants have been shown to predispose to HIV-associated nephropathy, probably as a result of *APOL1* gene and *APOL1*-gene environment interactions. This has been driven by positive selection of *APOL1* high-risk variants of with a combination of a high frequency and strong effect size (Friedman & Pollak, 2011). In this study, we first show findings from a case-control association analysis with variants in the *APOL1* and *MYH9* genes in a South African black population. This is the first report to statistically test the association of biopsy-proven forms of kidney diseases in a case-control study in Africa using DNA from formalin-fixed paraffin embedded kidney tissues. *APOL1* and *MYH9* variants are known to be associated with non-diabetic CKD among African Americans and Hispanic Americans (Genovese *et al.,* 2010; Kopp *et al.*, 2011; Kopp *et al.*, 2008; Kopp & Winkler, 2003; Tzur *et al.*, 2010; Winkler *et al.*, 2010). However, none of these variants had been shown previously to be associated with kidney disease in Southern Africa, a region with high prevalence of HIV-1 subtype C.

Our study was unique in that we explored SNPs from *APOL1* and *MYH9* genes covering 1000 basepairs upstream and downstream through a very stringent and thorough bioinformatic process described in chapter 2. We also included SNPs that had been previously used in other studies (Freedman *et al.*, 2009a; Freedman *et al.*, 2012; Kopp *et al.,* 2008; Oleksyk *et al.*, 2010; Winkler *et al.,* 2010), as well as ancestry informative markers to test for the possibility of bias in the strong association with HIVAN in this population.

## 4.1 *MYH9* association

The *MYH9* gene was found to be the leading gene responsible for mapping by admixture linkage disequilibrium signal highly associated with a large and important group of ESRD risk phenotypes regarded as *MYH9*-associated nephropathies (Bostrom & Freedman, 2010) in African Americans (Figure 4.1). These included HIVAN and FSGS. In our population, individual *MYH9* SNPs including SNPs constituting the *MYH9* E1 haplotype, showed no association with either of the kidney disease forms. This is in contrast to other studies (Cooke *et al.*, 2012; Freedman *et al.,* 2009a; Kopp *et al.*, 2008; Kopp *et al.,* 2010; O'Seaghdha *et al.*, 2011). However, in the haplotype analysis, the block (C-G-C) containing *MYH9* E1 haplotype SNPs (rs16996648-rs4821480-rs4821481) showed a weaker association. This showed that *MYH9* could be having some potential effect on kidney disease in our population. Further studies that included extending the *MYH9* association analyses to earlier stages and related kidney disease phenotypes and intensive re-sequencing of the *MYH9* gene, with the expectation of finding a functional African ancestry causative variant, were not successful (Freedman *et al.,* 2010; Kopp *et al.*, 2008; Tzur *et al.*, 2010). This led to a conclusion that there must be a region close to the *MYH9* that could be carrying the higher effect of predisposition to kidney disease. This region was later discovered to be the *APOL1* gene (Tzur *et al.,* 2010). Risk variants in this gene are said to be in high linkage disequilibrium with those of *MYH9* and this could probably be the reason why *MYH9* risk variants had been earlier thought to be responsible in the predisposition to kidney disease. Figure 4.1 shows the history of the timeline of the discovery of the genetic association with CKD in individuals of African ancestry.

## 4.2 *APOL1* association

This study demonstrated the highest association of HIVAN against other forms of kidney diseases, including FSGS and HIVICK, with *APOL1* risk variants in a black South African population by correlating genotype with clinical phenotype. This is due to the fact that the *APOL1* region has been found to harbor novel missense mutations with predicted functional effects specifically found in the African genome and that are more significantly associated with ESRD than all previously reported SNPs. This significant association with HIVAN is the strongest so far reported for common variants with complex diseases (Hindorff *et al.*, 2009). We established the highest association of HIVAN with two *APOL1* risk variants by determining the effect size of the combined *APOL1* G1 and G2 risk alleles for HIVAN with an OR of 89 (95% CI 17.68-911.72; $P=1.24X10^{-14}$) in a recessive model. This association was higher than that found in the African American population (OR 29.2, 95% CI 13.1-68.5; $P=6.0X10^{-22}$) (Kopp *et al.*, 2011) despite the fact that the confidence intervals overlap, there is no statistical difference between effect sizes in South Africans with HIVAN compared to African Americans with HIVAN. From our study, both the frequencies of *APOL1* risk alleles (79%) and OR (89) were elevated compared to that of African American population in the United States (72% and 29 respectively). This suggests that *APOL1* may have a strong gene-environment interaction in this HIV-1 subtype C setting. This HIV-1 subtype is the dominant strain in Southern and Eastern Africa and South Asia. HIV-1 subtype A is common in West Africa and subtype B is dominant in Europe, Australia, and the Americas (Hemelaar *et al.,* 2006). HIV-1 subtype C is highly virulent and accounts for up to 98% of HIV infections in South Africa, with corresponding higher viral loads (Novitsky *et al.*, 2011) and lower CD4 cells with the development of HIVAN. Late initiation of ART in resource-limited settings also has a part to play in predisposing at risk

individuals to HIVAN since previous studies have shown that effective roll-out of ART could reduce the occurrence of HIVAN (Hays T, 2012; Kalayjian, 2010; Wearne *et al.*, 2012). There could also be the possibilities of other environmental exposures, nutrition, genetic factors that are unique to or more common in this population coupled with other infections that may modify the effects of *APOL1* variants on the kidney.

In addition, our study also showed a weak effect with one *APOL1* risk allele (G1 or G2) in HIVAN (OR 5.49, 95% CI 0.87-61.14; *P=0.05*), which is consistent with previous studies (Genovese *et al.,* 2010; Kanji *et al.,* 2011; Kopp *et al.,* 2011). This suggests that there is a gain of injury and/or toxicity of these risk variants on kidney cells which is driven by an environmental factor such as HIV infection. Notably, there was no significant *APOL1* association with other forms of CKD (histology phenotypes), HIV (+) CKD and HIV (-) CKD including FSGS although previous studies on African Americans have shown significant associations of two *APOL1* risk alleles with primary FSGS with odds ratios of 10.5 (95% CI 6.0-18.1) and 17 (95% CI 11-26.5; $P=1.3X10^{-48}$) (Genovese *et al.,* 2010; Kopp *et al.,* 2011). Other studies have also shown *APOL1* associations with FSGS (Fine *et al.,* 2012; Papeta *et al.,* 2011). One of the reasons that could explain lack of *APOL1* risk alleles association with FSGS in our study is the high prevalence of HIV-1 in this setting such that, susceptible HIV-infected ART naïve individuals may develop HIVAN and not FSGS. Others have suggested that FSGS may be part of the spectrum of histology in response to HIV infection (Meehan *et al.,* 2012) or that HIV could be an incidental finding to primary FSGS in our population since FSGS is the commonest form of histology reported in adults and children with kidney disease (Marone *et al.,* 2001; Zou *et al.,* 2003). We also found that *APOL1* risk alleles were not associated with HIVICK. This was in

concordance with a study on African Americans that observed two *APOL1* risk alleles in only one of the thirty one cases of HIVICK (Fine *et al.,* 2012 & Foy *et al.*, 2013).

In the general population, individuals carrying two *APOL1* risk alleles occurred at a prevalence of 2-4% which was lower than the prevalence of 12-14% in the African American general population (Kopp *et al.,* 2011). In addition, it has been previously estimated that 50% of ART naïve HIV infected individuals with 2 *APOL1* risk alleles will develop HIVAN in their lifetime, supporting the magnitude of the interaction (Kopp *et al.,* 2011). However, it is worth noting that not all ART naïve HIV infected individuals with two risk alleles develop HIVAN and that approximately 20% of patients with HIVAN carry one or no *APOL1* risk allele, suggesting that there are additional factors to be considered.

There are distinct distributions of *APOL1* G1 and G2 risk alleles among various African and African-derived populations. We established the G1 allele (rs73885319) frequency to be approximately 7.3%, which is much lower than that reported in West Africa, in whom the frequencies are ~50% for Yoruba and 23.3% for Igbo or in African Americans where the G1 frequency is ~20% (Kopp *et al.,* 2011; Ulasi *et al.,* 2013). This could be as a result of selection pressure caused by trypanosomiasis or any other pathogens that could have been present in regions inhabited by ancestors of this South African population, or could be as a result of the *APOL1* variant alleles introduced into southern African populations as a result of gene flow from the Bantu expansion (de Filippo *et al.*, 2011; Thomson *et al.,* 2014). Due to host-pathogen interactions, these *APOL1* mutations provided a selective advantage to carriers of the risk alleles while at the same time increasing the risk for kidney disease. Genovese *et al.* (2010) (Genovese *et al.,* 2010), demonstrated the trypanolytic effect of serum from carriers of *APOL1* G1 and G2 risk alleles compared to serum from individuals lacking the risk alleles. They reported that

apolipoprotein L1 with both G1 and G2 risk variants restored the lytic activity of the serum, hence providing a selective advantage to carriers against Trypanosomiasis as a result of infection with *Trypanosoma brucei rhodesiense*. In addition, evidence of recent evolution of *APOL1* risk variants occurring in the past 10,000 years suggest that these variants were selected within Africa since they conferred protection against lethal Trypanosomiasis at a cost of increased susceptibility to kidney disease (Freedman *et al.*, 2010; Genovese *et al.*, 2010). This could also explain the high frequencies of these nephropathy risk variants in our HIVAN patients. G1 and/or G2 risk variants have been reported to be present in more than 30% of the African American chromosomes, approximately half of the 60% frequency for the *MYH9* E1 haplotype previously reported in this group. However, the effect of *APOL1* risk variants is approximately three-times greater (Freedman *et al.*, 2010; Genovese *et al.*, 2010).

Kopp *et al.* (2011) (Kopp *et al.,* 2011) showed that the effect size of the $G1^{GM}$ and $G1^{G+}$ haplotypes are equivalent and the renal risk of $G1^{+M}$ cannot be determined and therefore the individual with the rare haplotype $G1^{+M}$ (A-G-I) detected in the heterozygous state ($G1^{GM}/G1^{+M}$) among the HIV (+) CKD group was excluded from the analyses. This rare haplotype had also not been reported in the 1000 Genomes Project or in other previous studies of African populations, summarized in Thomson et al. (2014) and Limou et al. (2014) (Limou *et al.,* 2014; Thomson *et al.,* 2014).

The effect of *APOL1* risk variants is also shown in the combined haplotype analysis. *APOL1* and *MYH9* SNPs (T-T-G-G haplotype consisting of *APOL1* SNPs rs7284919-rs136147-rs73885319-rs60910145), showed the highest significant risk factor for HIVAN (OR 52.20; $P=2.75x10^{-06}$). This haplotype consists of two G1 risk alleles and this could explain its high association with HIVAN. In addition, the T-T-G-T haplotype of the same *APOL1* SNPs rs7284919-rs136147-

rs73885319-rs60910145, indicated significant protective association with HIVAN (OR 3.30 *P=0.999*). This shows that, as much as there is no significant high risk for HIVAN for individuals carrying zero or one risk allele G1: T-T-G-T, there might be a potential of developing HIVAN considering the odds ratio of 3.30. The highly associated G1 (rs73885319) risk variant with HIVAN, makes the haplotype consisting of the missense risk allele at rs73885319 ($G1^{G+}$) able to complement the $G1^{GM}$ haplotype containing both the G1 risk alleles in the pathogenesis of HIVAN. This proves that HIV infection interacts strongly with *APOL1* variants, although the molecular mechanism is unknown. This interaction is similar in South Africa and the USA. In addition to the *APOL1* association with HIVAN, we also observed a significant association of rs11089781 SNP in the *APOL3* gene with HIVAN under dominant and additive modes of association but not recessive mode. This is a very important association and it could be taken into consideration in future studies where an interaction effect from APOL3 could be examined in addition to its single risk association parameters.

The molecular mechanism by which *APOL1* variants cause kidney disease remains unclear. Apolipoprotein L1, a protein that is part of high density lipoprotein (HDL) class 3 family, and absent in other classes, circulates in the body and is responsible for a number of functions, among them protection of low density lipoprotein from oxidation. However, whether *APOL1* has a definitive role in the anti-oxidation is still unclear (Davidson *et al.,* 2009). This protein is also expressed in the kidney cells of HIVAN and FSGS biopsies but it remains to be determined whether circulating apolipoprotein L1, kidney-expressed apolipoprotein L1, or both, are involved in kidney injury (Madhavan *et al.,* 2011). G1 and G2 are located in the C-terminal domain (residues 342 to 398) of apolipoprotein, the binding region of the trypanosomal serum resistant associated (SRA) protein (Vanhamme *et al.*, 2003) and therefore it suggests that this domain

with *APOL1* risk variants act by gain-of-injury or toxicity mechanism on podocytes or vascular cells even though its main function is still unknown.

During apolipoprotein L1 expression, there is alternative splicing that takes place at the transcription level where three isoforms (A, B and C) are expressed (Nichols *et al.,* 2014). Isoform A is the predominant/principal isoform with a signal secretory peptide in the N-terminal region. Isoform C lacks this signal secretory peptide while in Isoform B the signal secretory peptide is too far from the N terminus to promote cleavage. N-terminal signal sequences dictate the secretory and membrane proteins to the endoplasmic reticulum (ER) in a signal recognition particle (SRP)-dependent manner. The signal sequence of genes consists of a hydrophobic core region (h-region), and the signal peptidase (SPase) cleavage site where cleavage occurs, resulting in cleaved signal sequences termed as signal peptides. Signal sequences are variable in their length and in their amino acid composition. This variability affects the characteristics of the protein including ER targeting, translocation, signal peptidase and post-targeting function cleavage, among others. Some signal peptides are processed by an intramembranous cleaving protease known as signal peptide peptidase (SPP), resulting in the N-terminal signal peptide fragments that are released into the cytosol. Alternatively, signal peptides remain membrane-inserted or are released from the ER membrane. These signal peptides are known to have wide-spread functions, either together with or independent of their corresponding mature proteins (Hegde & Bernstein, 2006; von Heijne, 1983).

In our study, not only do we show that the three isoforms can be expressed in HEK cells but we also show that there was high expression of isoforms B and C as opposed to isoform A in the HEK cells. The high expression seen in cells transfected with isoforms B and C suggest that these are the isoforms that are retained in the cell offering an environment that best suits the

development of kidney disease in the presence of the C-terminal risk alleles. However, it was beyond the scope of this thesis to show the trafficking and the localization of this protein within the HEK 293 cells which is an area that needs to be explored in future studies. Assuming apolipoprotein L1 uses the same mechanism on kidney cells as its trypanolytic effect, isoform A could be exported outside the cell while Isoforms B and C are retained within the cell and exported to the lysosome where they could utilize the pore forming domain (PFD) to create anion-specific pores that enable the movement of chloride ions ($Cl^-$) from the cytoplasm into the lysosome creating an anion imbalance. This leads to the movement of water into the lysosome, resulting in uncontrolled osmotic swelling of the lysosome which increases the intracellular pressure resulting in damage to the kidney cells, hence lysis and damage of plasma membrane, triggering the autophagic response manifestation of kidney disease. This could effectively occur in the presence of risk alleles in the last exon of the gene. Following this explanation, we demonstrate that the splice variant isoforms B and C that lack an effective signal peptide , have elevated expression in the kidney at the mRNA level in the presence of HIV infection as seen in the HIVAN case group and also following transfection of HEK 293 cells. This could possibly be explained in the cases of HIVAN and HIVICK, where there was a significantly high mRNA expression of isoform B, which coupled with the high frequency of *APOL1* risk alleles in the last exon, could have triggered the development of these diseases although there could be other factors involved that put HIVAN ahead of HIVICK in the severity of the diseases.

Expression of mRNA isoforms B and C in normal kidney suggest that, these individuals are at risk of developing kidney disease. However, lack of risk alleles at the last exon, and environmental factors such as HIV and other pathogens, limits the development of kidney disease suggesting a two-hit model for *APOL1*-associated kidney disease.

Although a gene's mRNA expression does not usually predict its protein level, it is equally important for its quantification. It is important to note that the cell can control the levels of protein at both the transcription and translation levels and the variability of mRNA expression is an indication that the cell controls mRNA expression at different points of the cell cycle. There are reasons why the two cannot be correlated; there could be a number of complex steps that are controlled by the cells during post-transcription that affect the ultimate protein levels. Greenbaum et al. (2003) found that the open reading frames (ORFs) that showed a large degree of variation in their expression, resulted in a high correlation (r=0.89) between the mRNA and the respective ORFs protein levels while there was low variation (r=0.20) between the mRNA and respective ORFs protein levels among the ORFs that showed minimal variation in their expression (Greenbaum *et al.,* 2003). There could also be substantial difference in the half-lives of different proteins *in vivo* hence affecting their ultimate levels of expression in cells.

A recent study showed that there is evidence suggesting that *APOL1* risk variants act as toxic gain-of-function mutations and that HIV triggers the immune system by activating exogenous and endogenous interferons such as interferon-gamma (IFN-γ) and tumor necrosis factor- alpha (TNF-α), which in turn increase the expression of apolipoprotein L1 Isoforms B and C in cells, leading to the development of kidney disease (Nichols *et al.*, 2014). This therefore shows that HIV is a powerful driver of *APOL1*-related nephropathy especially in individuals with HIVAN who carry two risk alleles. In addition, having carried out this study in a setting of high HIV prevalence with HIV-1 subtype C, explains the environmental factor inducing the expression of Isoforms B and C, hence leading to *APOL1*-mediated kidney disease. This supports a two-hit model for *APOL1*-associated disease expression (Freedman & Skorecki, 2014).

## 4.3 Suggestion for future study

The molecular mechanism of how apolipoprotein L1 causes kidney disease, especially in the context of HIV, is an area of research that still needs to be explored. Our results give a positive direction towards the link between *APOL1* high risk variants, apolipoprotein L1 isoforms and specific clinical entities such as HIVAN, HIVICK and possibly, other kidney diseases. Most importantly, our study highlights the areas of focus in determining the mechanism of HIV-associated apolipoprotein L1-mediated kidney disease. However, it is beyond the scope of this thesis to look at the effects of distinct pathways by specific pathogens in different cell compartments and/or cell types in specific *APOL1* phenotypes and also to look at the localization and trafficking of these isoforms which is an area that needs to be explored. In addition, for effective explanation of full-range apolipoprotein L1-mediated kidney disease, an exploration in many systems including clinical studies, human cells and animal models would be required. It is also important to note that there may be more than one means by which the mutant apolipoprotein L1 leads to pathology despite the presence of high risk genotypes. It is also important to look broadly at the roles of viral recognition pathways in the upregulation of the retained apolipoprotein L1 that contains risk variants, hence understanding pathways governing its expression that could help determine therapeutic interventions.

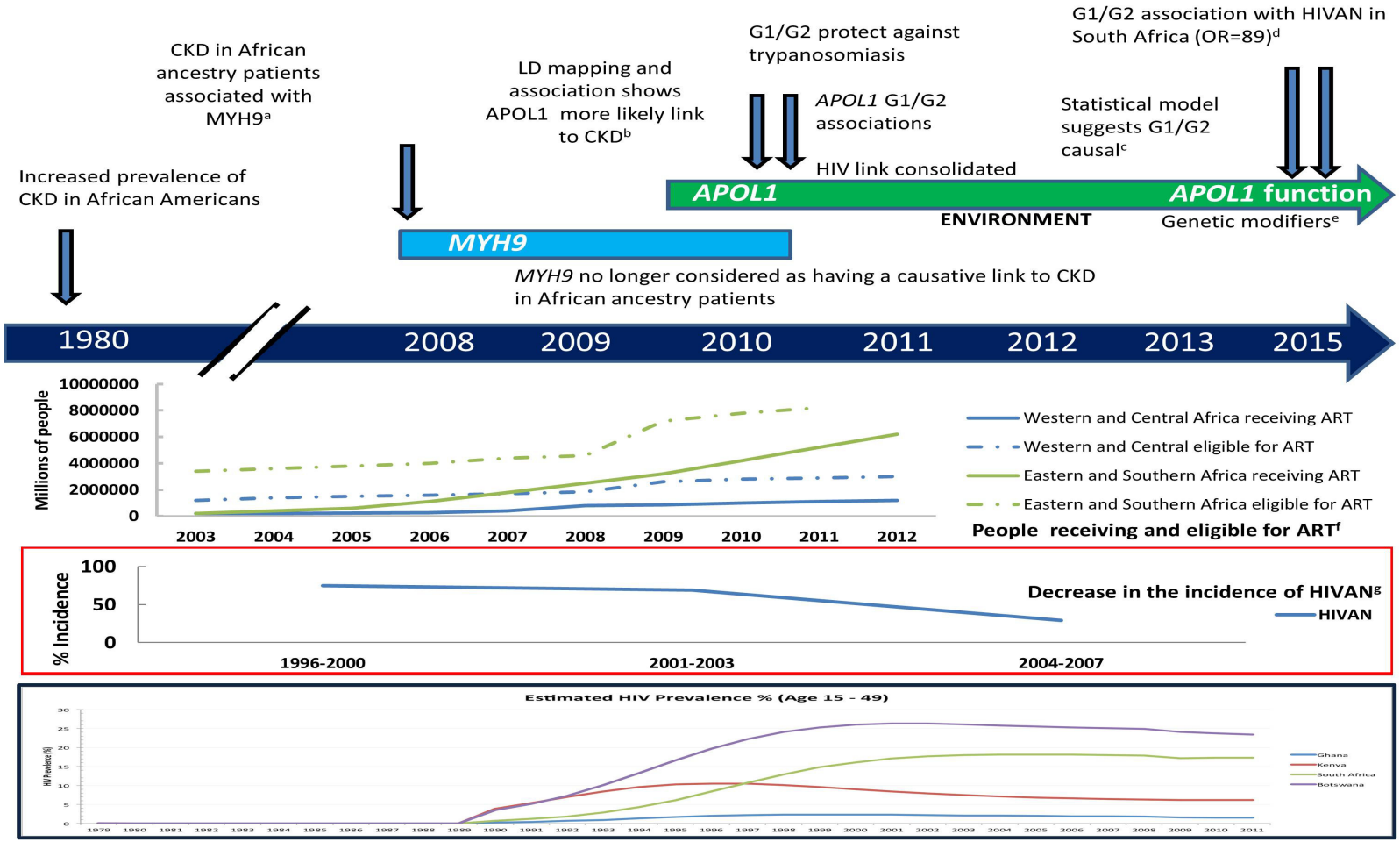



**Figure 4.1:** Timeline showing the discovery of genetic association to CKD in African ancestry populations and how an effective ART rollout has reduced HIVAN incidence and the plateaued increase of HIV infection prevalence.
[a]Kopp et al. (2008) (Kopp *et al.,* 2008); Kao et al. (2008) (Kao *et al.,* 2008). [b]Freedman et al. (2010) (Freedman *et al.*, 2010); Genovese *et al.* (2010) (Genovese *et al.,* 2010);Tzur et al. (2010) (Tzur *et al.*, 2010). [c]Genovese *et al.* (2013) (Genovese et al., 2013). [d]Kasembeli et al. (2015) (Kasembeli et al., 2015). [e] Freedman et al. 2014(Freedman & Skorecki, 2014). [f]UNAIDS report on global AIDS epidemic(UNAIDS). [g]Vermeulen A. MMed, University of the Witwatersrand, 2014. JHB; Johannesburg. *APOL1*; Apolipoprotein L1 gene. *MYH9*; non-muscle myosin heavy chain 9 gene. ART; Antiretroviral therapy.

## 4.4 Proposed pathogenesis of *APOL1*-mediated nephropathy

Several studies have affirmed that *APOL1* gene play a vital role in the development and faster progression of kidney disease. Having *N-terminal* apolipoprotein L1 isoforms B/C coupled with two *APOL1* risk alleles in the *C-terminal* of the same gene, being HIV-infected and of African descent, confers a greater risk of developing CKD as illustrated in Figure 4.2. Presence of HIV infection with other environmental factors increases the chance of developing HIVAN by 89-fold in our population. Thirty-three years after AIDS was first reported, HIVAN remains an important kidney disease in HIV-infected individuals (Gerntholtz *et al.*, 2006; Han *et al.*, 2006; Rao *et al.*, 1984). The pathogenesis could involve direct HIV infection of the kidney and infection of the lymphocytes and macrophages that enter the kidney resulting in the release of inflammatory lymphokines or cytokines such as IFN-γ which promote injury and fibrosis (Eitner *et al.*, 1998; Eitner *et al.*, 2000). This warrants close attention because South Africa has the highest number of people living with HIV. Without ART, genetically high risk HIV infected individuals would develop HIVAN which progresses rapidly to ESRD and death. Several studies have shown that effective roll-out of anti-retroviral (ARV) drugs can reduce the occurrence of HIVAN (Kalayjian, 2010; Wearne *et al.,* 2012; Wyatt & Klotman, 2007; Wyatt *et al.*, 2012) (Figure 4.1). Reasons that could explain the high effect size in HIVAN in South Africa compared to African Americans could possibly be due to co-infections with other pathogens and an ART treatment effect since the roll-out in the United States was earlier than the South African one.

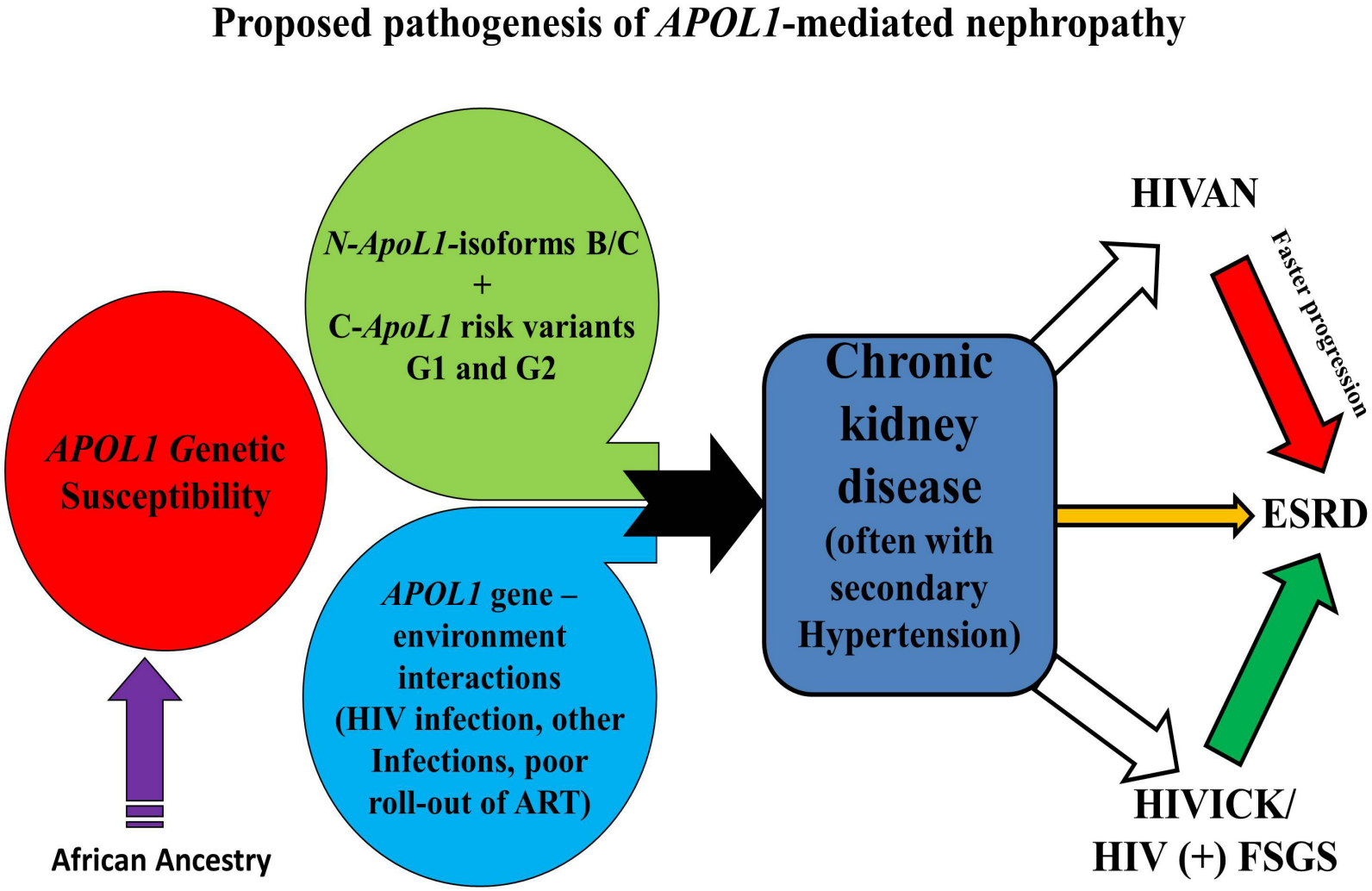



**Figure 4.2:** Proposed pathogenesis of *APOL1*-mediated nephropathy.
*APOL1*: Apolipoprotein L1 gene. *N-ApoL1*: N-terminal apolipoprotein L1 protein. *C-ApoL1:* C-terminal apolipoprotein L1 protein. HIVAN: HIV-associated nephropathy. HIVICK: HIV immune complex kidney disease. FSGS: Focal segmental glomerulosclerosis. ESRD: End stage renal disease. Adapted from Bostrom and Freedman et al. 2010 (Bostrom & Freedman, 2010).

# LIMITATIONS OF THE STUDY

Since there is diversity in the African population, the SNPs selected from the *International HapMap* and *1000 genomes Project*, of the Yoruba from Ibadan Nigeria, Luhya from Webuye, Kenya and Maasai from Kinyawa, Kenya, could not be the perfect proxies for the black South African population. We did not have sufficient numbers of primary or HIV (+) FSGS cases in order to detect the effects of the G1 and G2 risk alleles. We required a sample size of 95 FSGS cases and 95 controls in order to have 80% power to detect an OR 5, alpha 0.05. Our healthy population controls in the genetic study lacked clinical data hence assumed to have normal kidney function. It is in accordance to use random controls in genetic studies and it is unlikely that healthy individuals in this group had undetected glomerulopathy especially given the low frequency of individuals carrying two *APOL1* renal risk alleles in this group. It is also theoretically possible that undetected renal disease could reduce power.

# CONCLUSIONS

Our data suggest that there is a strikingly higher prevalence of *APOL1* risk variants among HIVAN as compared to other forms of kidney disease, including FSGS, and we demonstrated that HIV-infected South African blacks with two *APOL1* risk alleles, not receiving ART, have the greatest risk of developing HIVAN. Our study also highlights the importance of HIV as an environmental factor that triggers the expression of apolipoprotein L1 isoforms B and C that are retained in the cells hence acting as the initiating factors that lead to development of kidney disease. There is a need for HIV screening, implementation of World Health Organization (WHO) recommendations in order to initiate ART earlier to reduce the burden of other forms of HIV-related kidney disease in Africa especially HIVAN. Currently, the treatment of HIV infection guidelines for South Africa suggest that ART should be initiated in individuals with clinical stage 3 and 4 of HIV disease and HIV-specific manifestations (like HIV-related malignancies, kidney or other organ involvement, and hepatitis B and C infection) and in all HIV-positive individuals with CD4 counts <500 cells/μL (Society, 2012). We also demonstrated high isoform B mRNA levels in HIVAN and HIVICK individuals indicating high levels of apolipoprotein L1 isoform B expression in these patients. The high expression of the N-terminal region of *APOL1* gene for isoforms B and C in kidney cells compared to that of isoform A indicating that due to the secretory domain sequence cleavage site being too far to promote cleavage in isoform B, and lack of the secretory domain in isoform C, enables these two isoforms to be retained in the cell resulting in the process of cell damage manifested by the presence of risk variants at the C-terminal region. Furthermore, having the fluorescently labelled isoform constructs we would be able to study the dynamics of apolipoprotein L1 trafficking in the cell.

Genetic susceptibility to kidney disease in the context of HIV infection is influenced by variations in the *APOL1* gene in black South Africans and HIV could be a trigger in the expression of apolipoprotein L1 isoforms B and C that play a very crucial role in the development of kidney disease. Further studies would be required to assess the clinical utility of *APOL1* genetic testing and if at all in future, a patient's genetic risk profile could be established, to determine whether increased surveillance and early detection could alter the clinical course of the disease and prompt clinicians to initiate ART early in patients who otherwise would not meet the criteria for treatment. By doing this, complications of HIV infection could be controlled resulting in reduction of the expression of apolipoprotein L1 isoform B and C. These studies should be focused in regions with the highest HIV prevalence and where high frequencies of these risk alleles pose a greater need for the earlier initiation of ART. This potentially makes the *APOL1* gene variants a good candidate for personalized medicine in sub Saharan Africa.

# REFERENCES


Barrett, J. C., Fry, B., Maller, J., & Daly, M. J. (2005). Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics, 21(2), 263-265. doi:10.1093/bioinformatics/bth457bth457 [pii]

Behar, D.M., Kedem, E., Rosset, S., Haileselassie, Y., Tzur, S., Kra-Oz, Z., et al. (2011). Absence of APOL1 risk variants protects against HIV-associated nephropathy in the Ethiopian population. Am J Nephrol, 34(5), 452-459. doi: 10.1159/000332378

Behar, D. M., Rosset, S., Tzur, S., Selig, S., Yudkovsky, G., Bercovici, S., et al. (2010). African ancestry allelic variation at the MYH9 gene contributes to increased susceptibility to non-diabetic end-stage kidney disease in Hispanic Americans. Hum Mol Genet, 19(9), 1816-1827. doi: ddq040 [pii]10.1093/hmg/ddq040

Bishop, J. R., Shimamura, M., & Hajduk, S. L. (2001). Insight into the mechanism of trypanosome lytic factor-1 killing of Trypanosoma brucei brucei. [Research Support, Non-U.S. Gov't Research Support, U.S. Gov't, P.H.S.]. Mol Biochem Parasitol, 118(1), 33-40.

Bostrom, M. A., & Freedman, B. I. (2010). The spectrum of MYH9-associated nephropathy. [Research Support, N.I.H., Extramural Review]. Clin J Am Soc Nephrol, 5(6), 1107-1113. doi: 10.2215/CJN.08721209

Capewell, P., Cooper, A., Clucas, C., Weir, W., & Macleod, A. (2015). A co-evolutionary arms race: trypanosomes shaping the human genome, humans shaping the trypanosome genome. Parasitology, 142 (1), S108-S119. doi: 10.1017/S003118201400062

Chan, D. C., Fass, D., Berger, J. M., & Kim, P. S. (1997). Core structure of gp41 from the HIV envelope glycoprotein. Cell, 89(2), 263-273. doi: S0092-8674(00)80205-6 [pii]

Conrad, D. F., Keebler, J. E., DePristo, M. A., Lindsay, S. J., Zhang, Y., Casals, F., et al. (2011). Variation in genome-wide mutation rates within and between human families. Nat Genet, 43(7), 712-714. doi: ng.862 [pii]10.1038/ng.862

Cooke, J. N., Bostrom, M. A., Hicks, P. J., Ng, M. C., Hellwege, J. N., Comeau, M. E., et al. (2012). Polymorphisms in MYH9 are associated with diabetic nephropathy in European Americans. [Comparative Study Research Support, N.I.H., Extramural]. Nephrol Dial Transplant, 27(4), 1505-1511. doi: 10.1093/ndt/gfr522

Cunningham, A. L., Donaghy, H., Harman, A. N., Kim, M., & Turville, S. G. (2010). Manipulation of dendritic cell function by viruses. Curr Opin Microbiol, 13(4), 524-529. doi: S1369-5274(10)00095-0 [pii]10.1016/j.mib.2010.06.002

D'Agati, V., & Appel, G. B. (1997). HIV infection and the kidney. [Review]. J Am Soc Nephrol, 8(1), 138-152.

Davidson, W. S., Silva, R. A., Chantepie, S., Lagor, W. R., Chapman, M. J., & Kontush, A. (2009). Proteomic analysis of defined HDL subpopulations reveals particle-specific protein clusters: relevance to antioxidative function. Arterioscler Thromb Vasc Biol, 29(6), 870-876. doi: ATVBAHA.109.186031 [pii]10.1161/ATVBAHA.109.186031

de Filippo, C., Barbieri, C., Whitten, M., Mpoloka, S. W., Gunnarsdottir, E. D., Bostoen, K., et al. (2011). Y-chromosomal variation in sub-Saharan Africa: insights into the history of Niger-Congo groups. [Historical Article Research Support, Non-U.S. Gov't]. Mol Biol Evol, 28(3), 1255-1269. doi: 10.1093/molbev/msq312

De Rocco, D., Zieger, B., Platokouki, H., Heller, P. G., Pastore, A., Bottega, R., et al. (2013). MYH9-related disease: five novel mutations expanding the spectrum of causative mutations and confirming genotype/phenotype correlations. [Case Reports Research Support, Non-U.S. Gov't]. Eur J Med Genet, 56(1), 7-12. doi: 10.1016/j.ejmg.2012.10.009

de Silva, T. I., Post, F. A., Griffin, M. D., & Dockrell, D. H. (2007). HIV-1 infection and the kidney: an evolving challenge in HIV medicine. [Review]. Mayo Clin Proc, 82(9), 1103-1116. doi: 10.4065/82.9.1103

Dean, M., Carrington, M., Winkler, C., Huttley, G. A., Smith, M. W., Allikmets, R., et al. (1996). Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the CKR5 structural gene. Hemophilia Growth and Development Study, Multicenter AIDS Cohort Study, Multicenter Hemophilia Cohort Study, San Francisco City Cohort, ALIVE Study. [Multicenter Study Research Support, U.S. Gov't, Non-P.H.S. Research Support, U.S. Gov't, P.H.S.]. Science, 273(5283), 1856-1862.

Department of Health, S. A. (2010). COUNTRY PROGRESS REPORT ON THE DECLARATION OF COMMITMENT ON HIV/AIDS.

Douek, D. C., Roederer, M., & Koup, R. A. (2009). Emerging concepts in the immunopathogenesis of AIDS. Annu Rev Med, 60, 471-484. doi: 10.1146/annurev.med.60.041807.123549

Drawz, P. E., & Sedor, J. R. (2011). The genetics of common kidney disease: a pathway toward clinical relevance. [Review]. Nat Rev Nephrol, 7(8), 458-468. doi: 10.1038/nrneph.2011.85

Dunham, I., Shimizu, N., Roe, B. A., Chissoe, S., Hunt, A. R., Collins, J. E., et al. (1999). The DNA sequence of human chromosome 22. Nature, 402(6761), 489-495. doi: 10.1038/990031

Eitner, F., Cui, Y., Hudkins, K. L., & Alpers, C. E. (1998). Chemokine receptor (CXCR4) mRNA-expressing leukocytes are increased in human renal allograft rejection. Transplantation, 66(11), 1551-1557.

Eitner, F., Cui, Y., Hudkins, K. L., Stokes, M. B., Segerer, S., Mack, M., et al. (2000). Chemokine receptor CCR5 and CXCR4 expression in HIV-associated kidney disease. [Research Support, U.S. Gov't, P.H.S.]. J Am Soc Nephrol, 11(5), 856-867.

Enanga, B., Burchmore, R. J., Stewart, M. L., & Barrett, M. P. (2002). Sleeping sickness and the brain. Cell Mol Life Sci, 59(5), 845-858.

ESRD Patients in 2013: A Global Perspective. cited 2014-07-07. Available from: URL: http://www.vision-fmc.com/files/ESRD_Patients_in_2013.pdf

Estrella, M. M., Fine, D. M., & Atta, M. G. (2010). Recent developments in HIV-related kidney disease. HIV Ther, 4(5), 589-603. doi: 10.2217/hiv.10.42

Fabian, J., Katz, I., Gerntholtz, T., Goetsch, S., & Naicker, S. (2007). Chronic kidney disease in human immunodeficiency virus infection. [Review]. Panminerva Med, 49(2), 51-66.

Fan, J. B., Chee, M. S., & Gunderson, K. L. (2006). Highly parallel genomic assays. Nat Rev Genet, 7(8), 632-644. doi: nrg1901 [pii]10.1038/nrg1901

Faul, C., Asanuma, K., Yanagida-Asanuma, E., Kim, K., & Mundel, P. (2007). Actin up: regulation of podocyte structure and function by components of the actin cytoskeleton. Trends Cell Biol, 17(9), 428-437. doi: S0962-8924(07)00167-5 [pii]10.1016/j.tcb.2007.06.006

Fine, D. M., Wasser, W. G., Estrella, M. M., Atta, M. G., Kuperman, M., Shemer, R., et al. (2012). APOL1 risk variants predict histopathology and progression to ESRD in HIV-related kidney disease. [Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't]. J Am Soc Nephrol, 23(2), 343-350. doi: 10.1681/ASN.2011060562

Foundation, N. K. (2002). K/DOQI clinical practice guidelines for chronic kidney disease: evaluation, classification, and stratification. Am J Kidney Dis, 39(2 Suppl 1), S1-266. doi: S0272638602093563 [pii]

Foy, M. C., Estrella, M. M., Lucas, G. M., Tahir, F., Fine, D. M., Moore, R. D., et al. (2013). Comparison of risk factors and outcomes in HIV immune complex kidney disease and HIV-associated nephropathy. [Research Support, N.I.H., Extramural]. Clin J Am Soc Nephrol, 8(9), 1524-1532. doi: 10.2215/CJN.10991012

Freedman, B. I., Hicks, P. J., Bostrom, M. A., Comeau, M. E., Divers, J., Bleyer, A. J., et al. (2009a). Non-muscle myosin heavy chain 9 gene MYH9 associations in African Americans with

clinically diagnosed type 2 diabetes mellitus-associated ESRD. Nephrol Dial Transplant, 24(11), 3366-3371. doi: gfp316 [pii]10.1093/ndt/gfp316

Freedman, B. I., Hicks, P. J., Bostrom, M. A., Cunningham, M. E., Liu, Y., Divers, J., et al. (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 Int, 75(7), 736-745. doi: ki2008701 [pii]10.1038/ki.2008.701

Freedman, B. I., Kopp, J. B., Langefeld, C. D., Genovese, G., Friedman, D. J., Nelson, G. W., et al. (2010). The apolipoprotein L1 (APOL1) gene and nondiabetic nephropathy in African Americans. J Am Soc Nephrol, 21(9), 1422-1426. doi: ASN.2010070730 [pii]10.1681/ASN.2010070730

Freedman, B. I., Langefeld, C. D., Turner, J., Nunez, M., High, K. P., Spainhour, M., et al. (2012). Association of APOL1 variants with mild kidney disease in the first-degree relatives of African American patients with non-diabetic end-stage renal disease. [Comparative Study Research Support, N.I.H., Extramural]. Kidney Int, 82(7), 805-811. doi: 10.1038/ki.2012.217

Freedman, B. I., & Skorecki, K. (2014). Gene-Gene and Gene-Environment Interactions in Apolipoprotein L1 Gene-Associated Nephropathy. [Review]. Clin J Am Soc Nephrol. 9(11):2006-13. doi: 10.2215/CJN.01330214

Friedman, D. J., & Pollak, M. R. (2011). Genetics of kidney failure and the evolving story of APOL1. [Review]. J Clin Invest, 121(9), 3367-3374. doi: 10.1172/JCI46263

Gallo, R. C., & Montagnier, L. (2003). The discovery of HIV as the cause of AIDS. N Engl J Med, 349(24), 2283-2285. doi: 10.1056/NEJMp038194349/24/2283 [pii]

Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. (2007). [Multicenter Study Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't Research Support, U.S. Gov't, Non-P.H.S.]. Nature, 447(7145), 661-678. doi:10.1038/nature05911

Genovese, G., Friedman, D. J., Ross, M. D., Lecordier, L., Uzureau, P., Freedman, B. I., et al. (2010). Association of trypanolytic ApoL1 variants with kidney disease in African Americans. Science, 329(5993), 841-845. doi: science.1193032 [pii]10.1126/science.1193032

Genovese, G., Handsaker, R. E., Li, H., Kenny, E. E., & McCarroll, S. A. (2013). Mapping the human reference genome's missing sequence by three-way admixture in Latino genomes. [Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't]. Am J Hum Genet, 93(3), 411-421. doi: 10.1016/j.ajhg.2013.07.002

Gerntholtz, T. E., Goetsch, S. J., & Katz, I. (2006). HIV-related nephropathy: a South African perspective. [Comparative Study]. Kidney Int, 69(10), 1885-1891. doi: 10.1038/sj.ki.5000351

Gillett, M. P., & Owen, J. S. (1991). Trypanosoma brucei brucei: differences in the trypanocidal activity of human plasma and its relationship to the level of high density lipoproteins. [Research Support, Non-U.S. Gov't]. Trans R Soc Trop Med Hyg, 85(5), 612-616.

Greenbaum, D., Colangelo, C., Williams, K., & Gerstein, M. (2003). Comparing protein abundance and mRNA expression levels on a genomic scale. [Comparative Study Research Support, U.S. Gov't, P.H.S. Review]. Genome Biol, 4(9), 117. doi: 10.1186/gb-2003-4-9-117

Hajduk, S. L., Moore, D. R., Vasudevacharya, J., Siqueira, H., Torri, A. F., Tytler, E. M., et al. (1989). Lysis of Trypanosoma brucei by a toxic subspecies of human high density lipoprotein. [Research Support, Non-U.S. Gov't Research Support, U.S. Gov't, P.H.S.]. J Biol Chem, 264(9), 5210-5217.

Han, T. M., Naicker, S., Ramdial, P. K., & Assounga, A. G. (2006). A cross-sectional study of HIV-seropositive patients with varying degrees of proteinuria in South Africa. [Research Support, Non-U.S. Gov't]. Kidney Int, 69(12), 2243-2250. doi: 10.1038/sj.ki.5000339

Happo, L., Strasser, A., & Cory, S. (2012). BH3-only proteins in apoptosis at a glance. [Research Support, Non-U.S. Gov't Review]. J Cell Sci, 125(Pt 5), 1081-1087. doi: 10.1242/jcs.090514

Hays, T., & Wyatt, C. M. (2012). APOL1 variants in HIV-associated nephropathy: just one piece of the puzzle. Kidney Int, 82(3), 259-260. doi: ki2012129 [pii]10.1038/ki.2012.129

Hegde, R. S., & Bernstein, H. D. (2006). The surprising complexity of signal sequences. [Research Support, N.I.H., Intramural Review]. Trends Biochem Sci, 31(10), 563-571. doi: 10.1016/j.tibs.2006.08.004

Hemelaar, J., Gouws, E., Ghys, P. D., & Osmanov, S. (2006). Global and regional distribution of HIV-1 genetic subtypes and recombinants in 2004. AIDS, 20(16), W13-23. doi: 10.1097/01.aids.0000247564.73009.bc

Hindorff, L. A., Sethupathy, P., Junkins, H. A., Ramos, E. M., Mehta, J. P., Collins, F. S., et al. (2009). Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. [Research Support, N.I.H., Intramural]. Proc Natl Acad Sci U S A, 106(23), 9362-9367. doi: 10.1073/pnas.0903103106

Huang, Y., Paxton, W. A., Wolinsky, S. M., Neumann, A. U., Zhang, L., He, T., et al. (1996). The role of a mutant CCR5 allele in HIV-1 transmission and disease progression. Nat Med, 2(11), 1240-1243.

Idem. (2005). Epidemiology of ESRD: a world-wide perspective. In: El Nahas M, ed. Kidney diseases in the developing world and ethnic minorities. London, UK: Taylor & Francis. 1-13.

Kalayjian, R. C. (2010). The treatment of HIV-associated nephropathy. Adv Chronic Kidney Dis, 17(1), 59-71. doi: S1548-5595(09)00171-2 [pii]10.1053/j.ackd.2009.08.013

Kamau, E., Alemayehu, S., Feghali, K. C., Tolbert, L. S., Ogutu, B., & Ockenhouse, C. F. (2012). Development of a TaqMan Allelic Discrimination assay for detection of single nucleotides polymorphisms associated with anti-malarial drug resistance. Malar J, 11, 23. doi: 1475-2875-11-23 [pii]10.1186/1475-2875-11-23

Kanji, Z., Powe, C. E., Wenger, J. B., Huang, C., Ankers, E., Sullivan, D. A., et al. (2011). Genetic variation in APOL1 associates with younger age at hemodialysis initiation. [Research Support, N.I.H., Extramural]. J Am Soc Nephrol, 22(11), 2091-2097. doi: 10.1681/ASN.2010121234

Kao, W. H., Klag, M. J., Meoni, L. A., Reich, D., Berthier-Schaad, Y., Li, M., et al. (2008). MYH9 is associated with nondiabetic end-stage renal disease in African Americans. Nat Genet, 40(10), 1185-1192. doi: ng.232 [pii]10.1038/ng.232

Kasembeli, A. N., Duarte, R., Ramsay, M., Mosiane, P., Dickens, C., Dix-Peek, T., et al. (2015). APOL1 Risk Variants Are Strongly Associated with HIV-Associated Nephropathy in Black South Africans. J Am Soc Nephrol. doi: 10.1681/ASN.2014050469

Kher, V. (2002). End-stage renal disease in developing countries. [Case Reports Clinical Conference Research Support, Non-U.S. Gov't]. Kidney Int, 62(1), 350-362. doi: 10.1046/j.1523-1755.2002.00426.x

Kiryluk, K., Martino, J., & Gharavi, A. G. (2007). Genetic susceptibility, HIV infection, and the kidney. [Research Support, N.I.H., Extramural Review]. Clin J Am Soc Nephrol, 2 Suppl 1, S25-35. doi: 10.2215/CJN.00320107

Klag, M. J., Whelton, P. K., Randall, B. L., Neaton, J. D., Brancati, F. L., & Stamler, J. (1997). End-stage renal disease in African-American and white men. 16-year MRFIT findings. [Comparative Study Research Support, U.S. Gov't, P.H.S.]. JAMA, 277(16), 1293-1298.

Ko, D. C., & Urban, T. J. (2013). Understanding Human Variation in Infectious Disease Susceptibility through Clinical and Cellular GWAS. Plos Pathogens, 9(8). doi: ARTN e1003424DOI 10.1371/journal.ppat.1003424

Kopp, J. B., Nelson, G. W., Sampath, K., Johnson, R. C., Genovese, G., An, P., et al. (2011). APOL1 Genetic Variants in Focal Segmental Glomerulosclerosis and HIV-Associated Nephropathy. J Am Soc Nephrol, 22(11), 2129-2137. doi: ASN.2011040388 [pii]10.1681/ASN.2011040388

Kopp, J. B., Smith, M. W., Nelson, G. W., Johnson, R. C., Freedman, B. I., Bowden, D. W., et al. (2008). MYH9 is a major-effect risk gene for focal segmental glomerulosclerosis. Nat Genet, 40(10), 1175-1184. doi: ng.226 [pii]10.1038/ng.226

Kopp, J. B., & Winkler, C. (2003). HIV-associated nephropathy in African Americans. Kidney Int Suppl(83), S43-49. doi: kid8309 [pii]

Kopp, J. B., Winkler, C. A., & Nelson, G. W. (2010). MYH9 genetic variants associated with glomerular disease: what is the role for genetic testing? Semin Nephrol, 30(4), 409-417. doi: S0270-9295(10)00099-9 [pii]10.1016/j.semnephrol.2010.06.007

Kottgen, A. (2010). Genome-wide association studies in nephrology research. [Research Support, Non-U.S. Gov't Review]. Am J Kidney Dis, 56(4), 743-758. doi: 10.1053/j.ajkd.2010.05.018

Kronenberg, F. (2011). APOL1 variants and kidney disease. There is no such thing as a free lunch. Nephrol Dial Transplant, 26(3), 775-778. doi: gfq753 [pii]10.1093/ndt/gfq753

Lao, O., van Duijn, K., Kersbergen, P., de Knijff, P., & Kayser, M. (2006). Proportioning whole-genome single-nucleotide-polymorphism diversity for the identification of geographic population structure and genetic ancestry. Am J Hum Genet. 78(4):680-90.

Landi, D., Lockhart, E., Miller, S. E., Datto, M., Rehder, C., Kanaly, A., et al. (2012). Report of a young girl with MYH9 mutation and review of the literature. [Case Reports Review]. J Pediatr Hematol Oncol, 34(7), 538-540. doi: 10.1097/MPH.0b013e3182678fc9

Levey, A. S., Bosch, J. P., Lewis, J. B., Greene, T., Rogers, N., & Roth, D. (1999). A more accurate method to estimate glomerular filtration rate from serum creatinine: a new prediction equation. Modification of Diet in Renal Disease Study Group. Ann Intern Med, 130(6), 461-470. doi: 199903160-00002 [pii]

Limou, S., Nelson, G. W., Kopp, J. B., & Winkler, C. A. (2014). APOL1 kidney risk alleles: population genetics and disease associations. [Research Support, N.I.H., Extramural Research Support, N.I.H., Intramural Review]. Adv Chronic Kidney Dis, 21(5), 426-433. doi: 10.1053/j.ackd.2014.06.005

Lipkowitz, M. S., Freedman, B. I., Langefeld, C. D., Comeau, M. E., Bowden, D. W., Kao, W. H., et al. (2013). Apolipoprotein L1 gene variants associate with hypertension-attributed nephropathy and the rate of kidney function decline in African Americans. Kidney Int, 83(1), 114-120. doi: ki2012263 [pii]10.1038/ki.2012.263

Livak, K. J. (1999). Allelic discrimination using fluorogenic probes and the 5' nuclease assay. Genet Anal, 14(5-6), 143-149.

Lombard, Z., Crowther, N. J., van der Merwe, L., Pitamber, P., Norris, S. A., & Ramsay, M. (2012). Appetite regulation genes are associated with body mass index in black South African adolescents: a genetic association study. BMJ Open, 2(3). pii: e000873. doi: 10.1136/bmjopen-2012-000873.Lugli, E. B., Pouliot, M., Portela Mdel, P., Loomis, M. R., & Raper, J. (2004). Characterization of primate trypanosome lytic factors. [Research Support, Non-U.S. Gov't Research Support, U.S. Gov't, P.H.S.]. Mol Biochem Parasitol, 138(1), 9-20. doi: 10.1016/j.molbiopara.2004.07.004

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. [Research Support, N.I.H., Extramural]. J Am Soc Nephrol, 22(11), 2119-2128. doi: 10.1681/ASN.2011010069

Marone, M., Mozzetti, S., De Ritis, D., Pierelli, L., & Scambia, G. (2001). Semiquantitative RT-PCR analysis to assess the expression levels of multiple transcripts from the same sample. Biol Proced Online, 3, 19-25. doi: 10.1251/bpo20

May, A., Hazelhurst, S., Li, Y., Norris, S. A., Govind, N., Tikly, M., et al. (2013). Genetic diversity in black South Africans from Soweto. BMC Genomics, 14(1), 644. doi: 1471-2164-14-644 [pii]10.1186/1471-2164-14-644

Meehan, S. M., Kim, L., & Chang, A. (2012). A spectrum of morphologic lesions of focal segmental glomerulosclerosis by Columbia criteria in human immunodeficiency virus infection. Virchows Arch, 460(4), 429-435. doi: 10.1007/s00428-012-1213-3

Michelle M Estrella, M., MHS and Mohamed G Atta, MD, MPH. (2009). Kidney Disease in HIV Infection in the Late Era of Highly Active Antiretroviral Therapy. US Nephrology, 4(5): 589–603. doi: 10.2217/hiv.10.42

Miller, S. A., Dykes, D. D., & Polesky, H. F. (1988). A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res ., 16(3): 1215.

Nasiri, H., Forouzandeh, M., Rasaee, M. J., & Rahbarizadeh, F. (2005). Modified salting-out method: high-yield, high-quality genomic DNA extraction from whole blood using laundry detergent. [Evaluation Studies]. J Clin Lab Anal, 19(6), 229-232. doi: 10.1002/jcla.20083

Ng, P. C., & Henikoff, S. (2001). Predicting deleterious amino acid substitutions. Genome Res, 11(5), 863-874. doi: 10.1101/gr.176601

Ng, P. C., & Henikoff, S. (2002). Accounting for human polymorphisms predicted to affect protein function. Genome Res, 12(3), 436-446. doi: 10.1101/gr.212802

Ng, P. C., & Henikoff, S. (2003). SIFT: Predicting amino acid changes that affect protein function. Nucleic Acids Res, 31(13), 3812-3814.

Nichols, B., Jog, P., Lee, J. H., Blackler, D., Wilmot, M., D'Agati, V., et al. (2014). Innate immunity pathways regulate the nephropathy gene Apolipoprotein L1. Kidney Int., Feb; 87(2):332-42.doi: 10.1038/ki.2014.270

NIDDK. (2012). United States Renal Data System.

Nochy, D., Glotz, D., Dosquet, P., Pruna, A., Lemoine, R., Guettier, C., et al. (1993). Renal lesions associated with human immunodeficiency virus infection: North American vs. European experience. Adv Nephrol Necker Hosp, 22, 269-286.

Novitsky, V., Ndung'u, T., Wang, R., Bussmann, H., Chonco, F., Makhema, J., et al. (2011). Extended high viremics: a substantial fraction of individuals maintain high plasma viral RNA levels after acute HIV-1 subtype C infection. [Multicenter Study Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't]. AIDS, 25(12), 1515-1522. doi: 10.1097/QAD.0b013e3283471eb2

O'Seaghdha, C. M., Parekh, R. S., Hwang, S. J., Li, M., Kottgen, A., Coresh, J., et al. (2011). The MYH9/APOL1 region and chronic kidney disease in European-Americans. [Meta-Analysis Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't]. Hum Mol Genet, 20(12), 2450-2456. doi: 10.1093/hmg/ddr118

Oleksyk, T. K., Nelson, G. W., An, P., Kopp, J. B., & Winkler, C. A. (2010). Worldwide distribution of the MYH9 kidney disease susceptibility alleles and haplotypes: evidence of historical selection in Africa. PLoS One, 5(7), e11474. doi: 10.1371/journal.pone.0011474

Organization, W. H. (2011). Progress report 2011: Global HIV/AIDS response.

Owen, J. S., Gillett, M. P., & Hughes, T. E. (1992). Transgenic mice expressing human apolipoprotein A-I have sera with modest trypanolytic activity in vitro but remain susceptible to infection by Trypanosoma brucei brucei. [In Vitro Research Support, Non-U.S. Gov't]. J Lipid Res, 33(11), 1639-1646.

Page, N. M., Butlin, D. J., Lomthaisong, K., & Lowry, P. J. (2001). The human apolipoprotein L gene cluster: identification, classification, and sites of distribution. [Research Support, Non-U.S. Gov't]. Genomics, 74(1), 71-78. doi: 10.1006/geno.2001.6534

Papeta, N., Kiryluk, K., Patel, A., Sterken, R., Kacak, N., Snyder, H. J., et al. (2011). APOL1 variants increase risk for FSGS and HIVAN but not IgA nephropathy. [Research Support, N.I.H.,

Extramural Research Support, Non-U.S. Gov't]. J Am Soc Nephrol, 22(11), 1991-1996. doi: 10.1681/ASN.2011040434

Pays, E., Vanhollebeke, B., Vanhamme, L., Paturiaux-Hanocq, F., Nolan, D. P., & Perez-Morga, D. (2006). The trypanolytic factor of human serum. [Research Support, Non-U.S. Gov't Review]. Nat Rev Microbiol, 4(6), 477-486. doi: 10.1038/nrmicro1428

Perez-Morga, D., Vanhollebeke, B., Paturiaux-Hanocq, F., Nolan, D. P., Lins, L., Homble, F., et al. (2005). Apolipoprotein L-I promotes trypanosome lysis by forming pores in lysosomal membranes. [Research Support, Non-U.S. Gov't]. Science, 309(5733), 469-472. doi: 10.1126/science.1114566

Population Division of the Department of Economic and Social Aff airs of the United Nations Secretariat. World population prospects: the 2007 revision and world urbanization prospects. cited 2014-06-13. Available from: URL: http://www.un.org/esa/population/unpop.htm

Post, F. A., & Holt, S. G. (2009). Recent developments in HIV and the kidney. [Research Support, Non-U.S. Gov't Review]. Curr Opin Infect Dis, 22(1), 43-48. doi: 10.1097/QCO.0b013e328320ffec

Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., & Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. [Research Support, Non-U.S. Gov't]. Nat Genet, 38(8), 904-909. doi: 10.1038/ng1847

Ramsay, M. (2012). Africa: continent of genome contrasts with implications for biomedical research and health. FEBS Lett, 586(18), 2813-2819. doi: S0014-5793(12)00629-1 [pii]10.1016/j.febslet.2012.07.061

Rao, T. K., Filippone, E. J., Nicastri, A. D., Landesman, S. H., Frank, E., Chen, C. K., et al. (1984). Associated focal and segmental glomerulosclerosis in the acquired immunodeficiency syndrome. [Case Reports]. N Engl J Med, 310(11), 669-673. doi: 10.1056/NEJM198403153101101

Ray, P. E. (2012a). HIV-associated nephropathy: a diagnosis in evolution. [Comment Editorial Research Support, N.I.H., Extramural]. Nephrol Dial Transplant, 27(11), 3969-3972. doi: 10.1093/ndt/gfs114

Ray, P. E. (2012b). HIV-associated nephropathy: a diagnosis in evolution. Nephrol Dial Transplant. doi: gfs114 [pii]10.1093/ndt/gfs114

Reich, D. E., Cargill, M., Bolk, S., Ireland, J., Sabeti, P. C., Richter, D. J., et al. (2001). Linkage disequilibrium in the human genome. Nature, 411(6834), 199-204. doi: 10.1038/3507559035075590 [pii]

Rodriguez, J., & Lazebnik, Y. (1999). Caspase-9 and APAF-1 form an active holoenzyme. [Research Support, Non-U.S. Gov't Research Support, U.S. Gov't, P.H.S.]. Genes Dev, 13(24), 3179-3184.

Rosset, S., Tzur, S., Behar, D. M., Wasser, W. G., & Skorecki, K. (2011). The population genetics of chronic kidney disease: insights from the MYH9-APOL1 locus. Nat Rev Nephrol, 7(6), 313-326. doi: nrneph.2011.52 [pii]10.1038/nrneph.2011.52

Rostand, S. G., Kirk, K. A., Rutsky, E. A., & Pate, B. A. (1982). Racial differences in the incidence of treatment for end-stage renal disease. [Research Support, Non-U.S. Gov't Research Support, U.S. Gov't, Non-P.H.S.]. N Engl J Med, 306(21), 1276-1279. doi: 10.1056/NEJM198205273062106

Satko, S. G., Freedman, B. I., & Moossavi, S. (2005). Genetic factors in end-stage renal disease. [Research Support, U.S. Gov't, P.H.S. Review]. Kidney Int Suppl(94), S46-49. doi: 10.1111/j.1523-1755.2005.09411.x

Savoia, A., Germeshausen, M., De Rocco, D., Henschel, B., Kratz, C. P., Kuhlen, M., et al. (2010). MYH9-related disease: Report on five German families and description of a novel mutation. Ann Hematol, 89(10), 1057-1059. doi: 10.1007/s00277-010-0928-y

Schlebusch, C. M., & Soodyall, H. (2012). Extensive population structure in San, Khoe, and mixed ancestry populations from southern Africa revealed by 44 short 5-SNP haplotypes. Hum Biol, 84(6):695-724. doi: http://dx.doi.org/10.3378/027.084.0603

Shlush, L. I., Bercovici, S., Wasser, W. G., Yudkovsky, G., Templeton, A., Geiger, D., et al. (2010). Admixture mapping of end stage kidney disease genetic susceptibility using estimated mutual information ancestry informative markers. [Research Support, Non-U.S. Gov't]. BMC Med Genomics, 3, 47. doi: 10.1186/1755-8794-3-47

Sim, N. L., Kumar, P., Hu, J., Henikoff, S., Schneider, G., & Ng, P. C. (2012). SIFT web server: predicting effects of amino acid substitutions on proteins. Nucleic Acids Res, 40(Web Server issue), W452-457. doi: gks539 [pii]10.1093/nar/gks539

Smith, A. B., Esko, J. D., & Hajduk, S. L. (1995). Killing of trypanosomes by the human haptoglobin-related protein. [Research Support, Non-U.S. Gov't Research Support, U.S. Gov't, P.H.S.]. Science, 268(5208), 284-286.

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. [Research Support, Non-U.S. Gov't]. Genome Res, 19(5), 850-858. doi: 10.1101/gr.085647.108

Smith, M. W., & O'Brien, S. J. (2005). Mapping by admixture linkage disequilibrium: advances, limitations and guidelines. Nat Rev Genet, 6(8), 623-632. doi: nrg1657 [pii]10.1038/nrg1657

Society, S. A. H. C. (2012). Guidelines for antiretroviral therapy in adults. SAJHIVMED, Vol. 13, No. 3, 114-133.

Strasser, A. (2005). The role of BH3-only proteins in the immune system. [Research Support, Non-U.S. Gov't Research Support, U.S. Gov't, P.H.S. Review]. Nat Rev Immunol, 5(3), 189-200. doi: 10.1038/nri1568

Tayo, B. O., Kramer, H., Salako, B. L., Gottesman, O., McKenzie, C. A., Ogunniyi, A., et al. (2013). Genetic variation in APOL1 and MYH9 genes is associated with chronic kidney disease among Nigerians. Int Urol Nephrol, 45(2), 485-494. doi: 10.1007/s11255-012-0263-4

Thomas, D. B., Franceschini, N., Hogan, S. L., Ten Holder, S., Jennette, C. E., Falk, R. J., et al. (2006). Clinical and pathologic characteristics of focal segmental glomerulosclerosis pathologic variants. Kidney Int, 69(5), 920-926. doi: 5000160 [pii]10.1038/sj.ki.5000160

Thome, M., & Tschopp, J. (2001). Regulation of lymphocyte proliferation and death by FLIP. [Research Support, Non-U.S. Gov't Review]. Nat Rev Immunol, 1(1), 50-58. doi: 10.1038/35095508

Thomson, R., & Finkelstein, A. (2015). Human trypanolytic factor APOL1 forms pH gated cation-selective channels in planar lipid bilayers: relevance to trypanosome lysis. Proc Natl Acad Sci U S A, 112(9):2894-2899. doi: 10.1073/pnas.1421953112

Thomson, R., Genovese, G., Canon, C., Kovacsics, D., Higgins, M. K., Carrington, M., et al. (2014). Evolution of the primate trypanolytic factor APOL1. Proc Natl Acad Sci U S A, 111(20), E2130-2139. doi: 10.1073/pnas.1400699111

Tomlinson, S., Muranjan, M., Nussenzweig, V., & Raper, J. (1997). Haptoglobin-related protein and apolipoprotein AI are components of the two trypanolytic factors in human serum. [Research Support, Non-U.S. Gov't]. Mol Biochem Parasitol, 86(1), 117-120.

Trenaman, A., Hartley, C., Prorocic, M., Passos-Silva, D. G., Hoek, M., Nechyporuk-Zloy, V., et al. (2013). Trypanosoma brucei BRCA2 acts in a life cycle-specific genome stability process and dictates BRC repeat number-dependent RAD51 subnuclear dynamics. Nucleic Acids Res, 41(2), 943-960. doi: gks1192 [pii]10.1093/nar/gks1192

Tzur, S., Rosset, S., Shemer, R., Yudkovsky, G., Selig, S., Tarekegn, A., et al. (2010). Missense mutations in the APOL1 gene are highly associated with end stage kidney disease risk previously attributed to the MYH9 gene. Hum Genet, 128(3), 345-350. doi: 10.1007/s00439-010-0861-0

Ulasi, II, Tzur, S., Wasser, W. G., Shemer, R., Kruzel, E., Feigin, E., et al. (2013). High Population Frequencies of APOL1 Risk Variants Are Associated with Increased Prevalence of Non-Diabetic Chronic Kidney Disease in the Igbo People from South-Eastern Nigeria. Nephron Clin Pract, 123(1-2), 123-128. doi: 10.1159/000353223

UNAIDS. report on the global AIDS epidemic. cited 2014-05-20. Available from : URL: http://www.unaids.org/en/media/unaids/contentassets/documents/epidemiology/2013/gr2013/UNAIDS_Global_Report_2013_en.pdf.

UNAIDS. (2011). UNAIDS World AIDS Day Report 2011.

Vanhamme, L., Paturiaux-Hanocq, F., Poelvoorde, P., Nolan, D. P., Lins, L., Van Den Abbeele, J., et al. (2003). Apolipoprotein L-I is the trypanosome lytic factor of human serum. [Research Support, Non-U.S. Gov't]. Nature, 422(6927), 83-87. doi: 10.1038/nature01461

Vanhollebeke, B., & Pays, E. (2010). The trypanolytic factor of human serum: many ways to enter the parasite, a single way to kill. [Research Support, Non-U.S. Gov't Review]. Mol Microbiol, 76(4), 806-814. doi: 10.1111/j.1365-2958.2010.07156.x

von Heijne, G. (1983). Patterns of amino acids near signal-sequence cleavage sites. [Research Support, Non-U.S. Gov't]. Eur J Biochem, 133(1), 17-21.

Wan, G., Zhaorigetu, S., Liu, Z., Kaini, R., Jiang, Z., & Hu, C. A. (2008). Apolipoprotein L1, a novel Bcl-2 homology domain 3-only lipid-binding protein, induces autophagic cell death. [Research Support, N.I.H., Extramural]. J Biol Chem, 283(31), 21540-21549. doi: 10.1074/jbc.M800214200

Wearne, N., Swanepoel, C. R., Boulle, A., Duffield, M. S., & Rayner, B. L. (2012). The spectrum of renal histologies seen in HIV with outcomes, prognostic indicators and clinical correlations. Nephrol Dial Transplant, 27(11), 4109-4118. doi: 10.1093/ndt/gfr702

Wheeler, R. J. (2010). The trypanolytic factor-mechanism, impacts and applications. [Research Support, Non-U.S. Gov't Review]. Trends Parasitol, 26(9), 457-464. doi: 10.1016/j.pt.2010.05.005

Williams, S. E., Carmichael, T. R., Allingham, R.R., Hauser, M., & Ramsay, M. (2015). The genetics of POAG in black South Africans: a candidate gene association study. Sci. Rep, 5:8378. doi: 10.1038/srep08378

Winkler, C. A., Nelson, G., Oleksyk, T. K., Nava, M. B., & Kopp, J. B. (2010). Genetics of focal segmental glomerulosclerosis and human immunodeficiency virus-associated collapsing glomerulopathy: the role of MYH9 genetic variation. Semin Nephrol, 30(2), 111-125. doi: S0270-9295(10)00011-2 [pii]10.1016/j.semnephrol.2010.01.003

Wyatt, C. M., & Klotman, P. E. (2007). HIV-1 and HIV-Associated Nephropathy 25 Years Later. Clin J Am Soc Nephrol, 2 Suppl 1, S20-24. doi: 2/Supplement_1/S20 [pii]10.2215/CJN.03561006

Wyatt, C. M., Meliambro, K., & Klotman, P. E. (2012). Recent progress in HIV-associated nephropathy. Annu Rev Med, 63, 147-159. doi: 10.1146/annurev-med-041610-134224

Zou, N., Ditty, S., Li, B., & Lo, S. C. (2003). Random priming PCR strategy to amplify and clone trace amounts of DNA. [Evaluation Studies Technical Report Validation Studies]. Biotechniques, 35(4), 758-760, 762-755.

# APPENDICES

## APPENDIX 1

CLINICAL RESEARCH www.jasn.org

# *APOL1* Risk Variants Are Strongly Associated with HIV-Associated Nephropathy in Black South Africans

Alex N. Kasembeli,* Raquel Duarte,* Michèle Ramsay,† Pulane Mosiane,‡ Caroline Dickens,* Thérèse Dix-Peek,* Sophie Limou,§ Efe Sezgin,§‖ George W. Nelson,§ Agnes B. Fogo,¶ Stewart Goetsch,** Jeffrey B. Kopp,†† Cheryl A. Winkler,§ and Saraladevi Naicker*‡‡

*Department of Internal Medicine, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa; †Division of Human Genetics, National Health Laboratory Service, School of Pathology, Faculty of Health Sciences, Sydney Brenner Institute for Molecular Bioscience, University of the Witwatersrand, Johannesburg, South Africa; ‡Department of Anatomical Pathology, National Health Laboratory Service, University of the Witwatersrand, Johannesburg, South Africa; §Basic Research Laboratory, Center for Cancer Research, National Cancer Institute, Frederick National Laboratory, Leidos Biomedical Inc., Frederick, Maryland; ‖Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland; ¶Division of Renal Pathology/Electron Microscopy, Department of Pathology, Microbiology, and Immunology, Vanderbilt University Medical Center, Nashville, Tennessee; **Lancet Pathology Laboratory, Johannesburg, South Africa; ††Kidney Disease Section, National Institute of Diabetes and Digestive and Kidney Disease, National Institutes of Health, Bethesda, Maryland; and ‡‡Division of Nephrology, University of the Witwatersrand, Johannesburg, South Africa

**ABSTRACT**

*APOL1* variants are associated with HIV-associated nephropathy and FSGS in African Americans. The prevalence of these variants in African populations with CKD in HIV-1 infection has not been investigated. We determined the role of *APOL1* variants in 120 patients with HIV-associated nephropathy and CKD and 108 controls from a South-African black population. Patients with CKD were selected on the basis of histology. Genotypes were successfully determined for *APOL1* G1 and G2 variants and 42 single nucleotide polymorphisms, including 18 ancestry informative markers, for 116 patients with CKD (96.7%; 38 patients with HIV-associated nephropathy, 39 patients with HIV-positive CKD, and 39 patients with HIV-negative CKD), and 108 controls (100%). Overall, 79% of patients with HIV-associated nephropathy and 2% of population controls carried two risk alleles. In a recessive model, individuals carrying any combination of two *APOL1* risk alleles had 89-fold higher odds (95% confidence interval, 18 to 912; $P<0.001$) of developing HIV-associated nephropathy compared with HIV-positive controls. Population allele frequencies were 7.3% for G1 and 11.1% for G2. *APOL1* risk alleles were not significantly associated with other forms of CKD. These results indicate HIV-positive, antiretroviral therapy–naïve South-African blacks with two *APOL1* risk alleles are at very high risk for developing HIV-associated nephropathy. Further studies are required to determine the effect of *APOL1* risk variants on kidney diseases in other regions of sub-Saharan Africa.

*J Am Soc Nephrol* 26: ●●●–●●●, 2015. doi: 10.1681/ASN.2014050469

CKD is an important public health problem worldwide. African Americans experience high rates of CKD arising from diabetic nephropathy, hypertension-attributed CKD, FSGS, and HIV-associated nephropathy (HIVAN).[1–3] HIV-positive individuals of African descent have an 18-fold to 50-fold increased risk of developing CKD[4] and an 18-fold increased risk of developing HIVAN, suggesting that genetic factors play an important role in susceptibility for HIVAN.[2]

HIVAN is considered to be a morphologic variant of FSGS characterized by segmental and global

Received May 13, 2014. Accepted January 15, 2015.

S.N and C.A.W. contributed equally to this work.

Published online ahead of print. Publication date available at www.jasn.org.

**Correspondence:** Dr. Saraladevi Naicker, School of Clinical Medicine, Faculty of Health Sciences, University of the Witwatersrand, 7 York Road, Room 4B41, Parktown, Johannesburg 2193, South Africa, or Dr. Cheryl A. Winkler, Center for Cancer Research, National Cancer Institute, Frederick National Laboratory, Building 560, Room 21-104, Frederick, MD 20878. Email: Saraladevi.Naicker@wits.ac.za or winklerc@mail.nih.gov

Copyright © 2015 by the American Society of Nephrology

glomerular collapse, hypertrophy and hyperplasia of visceral glomerular epithelial cells, and severe tubulointerstitial disease.[5] In the absence of effective antiretroviral therapy (ART), approximately 50% of patients with HIVAN progress to ESRD within 2 years.[6,7] The introduction of ART has led to a substantial decline in the incidence of HIVAN in the United States.[6,7] In sub-Saharan Africa, HIVAN is the most common cause of kidney disease morbidity and mortality in HIV-positive ART-naïve patients.[8,9]

Mapping by admixture linkage disequilibrium studies for FSGS, HIVAN, and nondiabetic ESRD among African Americans led to the identification of a genomic region on chromosome 22q12 with very strong association with HIVAN, FSGS, and other forms of nondiabetic and hypertension-attributed ESRD.[10,11] Further studies identified three codon-changing variants in the *APOL1* gene, encoding ApoL1, with extremely strong associations with HIVAN (odds ratio [OR], 29; 95% CI, 13.1 to 68.5), FSGS (OR, 17; 95% CI, 11 to 26.5), and ESRD (OR, 7; 95% CI, 6 to 10) for homozygotes or compound heterozygotes carrying two risk alleles in Americans with African ancestry.[12–14] *APOL1* renal risk variants have been shown to increase the rate of progression to ESRD in persons with CKD, including FSGS, diabetic CKD, and hypertension-attributed CKD.[13,15]

The *APOL1* risk alleles comprise two tightly linked missense variants G1 (rs73885319; p.S342G and rs60910145; p.I384M) and G2, a 6-bp in-frame deletion removing two amino acids (rs71785313; p.N388del:Y389del) in the last exon of the *APOL1* gene.[12] The G1 and G2 alleles range in frequency from 0% to approximately 45% for G1 and to approximately 20% for G2 in different geographical regions of sub-Saharan Africa with the lowest frequencies in East Africa and the highest frequencies observed in West Africa.[13,16–19] The prevailing hypothesis is that the G1, and to a lesser extent G2, alleles have been under recent selection in West Africa by trypanosomes or other extracellular pathogens.[12] This study aimed to determine the prevalence of *APOL1* risk variants and the effect of these variants on HIVAN and CKD in black South Africans in a setting of high HIV-1 prevalence.

## RESULTS

This study included 228 adult individuals of black African ancestry from Johannesburg, South Africa. Characteristics of the CKD patient groups, all of whom were verified by histology, and HIV-positive controls are presented in Table 1. HIVAN was distinguished from HIV-positive FSGS by the presence of glomerular capillary collapse and glomerular visceral epithelial cell proliferation affecting at least one glomerulus, together with microcystic tubular dilation and interstitial inflammation. Median eGFR was significantly lower in patients with HIVAN (11.9 ml/min per 1.73 $m^2$) compared with all other patients with CKD (33.5 ml/min per 1.73 $m^2$ and 56.6 ml/min per 1.73 $m^2$ for HIV-positive and HIV-negative patients, respectively) and HIV-positive controls ($P$=0.001). Measures of kidney function were not available for the population control group; they were all apparently healthy individuals who gave consent to be enrolled for population genetic studies as control participants. The CKD patient groups and HIV-positive control group did not differ significantly by age (34.6, 36.1, 36.4, and 38.8 years, respectively) and sex at biopsy or study enrollment ($P$>0.05) (Table 1). All HIV-positive patients with CKD were ART naïve; however, 75% of the HIV-positive control group had been initiated on ART before recruitment into the study. Median viral load (log10 copies/ml) levels were highest in patients with HIVAN (5.1 log copies/ml) compared with HIV-positive patients with CKD (4.7 log copies/ml) and undetectable median levels in HIV-positive controls. Median CD4 counts were lowest in patients with HIVAN (92 cells/$mm^3$) compared with HIV-positive patients with CKD (244.5 cells/$mm^3$) and HIV-positive controls (371.0 cells/$mm^3$) (Table 1). *APOL1* genotypes were determined for 96.7% of patients with CKD (HIVAN, $n$=38; HIV-positive CKD, $n$=39; and HIV-negative CKD, $n$=39) and 100% of controls (HIV-positive controls, $n$=54; and population controls, $n$=54) (Table 2). Genotype distributions did not differ significantly from Hardy–Weinberg equilibrium expectations in HIV-positive and population control groups ($P$>0.05). Single nucleotide polymorphism (SNP) and haplotype frequencies are shown in Figure 1. Two infrequent haplotypes were observed for G1; these were confirmed by Sanger sequencing and by a repeat of the TaqMan assay. Because of the close proximity of the two G1 SNPs, recombination between them is a rare event. The haplotype G-G-I ($G1^{GM}$), comprising the derived allele at both rs73885319 (p.S342G) and rs60910145 (p.I384M) and the ancestral insertion (I) allele at rs71785313, was the most frequent G1 configuration (6.4% in the population controls and 52.6% in the HIVAN group). A novel haplotype, $G1^{+M}$ (A-G-I), confirmed by Sanger sequencing of the PCR product, was observed as a compound heterozygote ($G1^{+M}$/$G1^{GM}$) in a single HIV-positive patient with FSGS. Although Kopp *et al.*[13] showed that the effect size of the $G1^{GM}$ and $G1^{G+}$ haplotypes are equivalent, the renal risk of $G1^{+M}$ cannot be determined; therefore, we excluded this individual from further analysis in this study.

In the population control group, the allele frequencies were 7.3% for the G1 risk allele for rs73885319 and 11.1% for the G2 deletion allele (Figure 1). The combined frequency of the G1 and G2 risk alleles was highly enriched in patients with HIVAN (90.6%) compared with HIV-positive controls (20.8%) ($P$=8.0×$10^{-20}$) and the population control group (18.4%) ($P$=8.0×$10^{-22}$) (Figure 1). Similar distortions were observed for carriage of two *APOL1* risk alleles; 78.9% of the patients with HIVAN were homozygous (G1/G1 or G2/G2) or compound heterozygotes (G1/G2) compared with 3.7% of the HIV-positive controls ($P$=1.2×$10^{-14}$) and 1.9% of the population controls ($P$=8.9×$10^{-16}$) (Table 2). The distribution of G0, G1, and G2 genotypes and risk alleles among patients with CKD and controls are shown in Table 2. The distribution of genotypes was not statistically different when comparing the HIV-positive patients with CKD with HIV-positive controls or

**Table 1.** Characteristics of study participants and controls

| Characteristic | HIVAN ($n$=39)[d] | HIV-Positive CKD ($n$=40)[e] | HIV-Positive Controls ($n$=54)[f] | $P$ Value[g] | HIV-Negative CKD ($n$=41)[e] | Population Controls ($n$=54)[h] |
|---|---|---|---|---|---|---|
| Sex, $n$ (%) | | | | | | |
| Men | 16.0 (41.0) | 18.0 (42.9) | 15.0 (28.0) | 0.07 | 24.0 (58.1) | 24.0 (44.4) |
| Women | 23.0 (59.0) | 22.0 (57.1) | 39.0 (72.0) | 0.73 | 17.0 (41.9) | 30.0 (55.6) |
| Mean age (SD) | 34.6 (8.3) | 36.1 (9.3) | 38.8 (7.1) | 0.62 | 36.4 (11.0) | 38.5 (8.6) |
| Viral load (log copies/ml)[a] | 5.1 (4.5–5.3) | 4.7 (3.8–4.9) | Undetectable | 0.001 | — | — |
| CD4 (cells/mm$^3$)[b] | 92.0 (76–195) | 244.5 (71.5–351.5) | 371.0 (215.0–545.0) | 0.001 | — | — |
| eGFR (ml/min per 1.73 m$^2$)[c] | 11.9 (7.5–22.6) | 33.5 (15.9–68.8) | 121.6 (103.3–146.9) | 0.001 | 56.6 (8.3–101.7) | Not available |
| Serum creatinine ($\mu$mol/L) | 534.0 (262.0–799.0) | 233.0 (106.0–388.0) | 66.0 (55.0–77.0) | 0.001 | 155.0 (91.0–778.0) | Not available |

Data are given as the median (interquartile range) unless otherwise indicated.
[a]Patients with HIVAN ($n$=30), HIV-positive patients with CKD ($n$=27), and HIV-positive controls ($n$=45).
[b]Patients with HIVAN ($n$=26), HIV-positive patients with CKD ($n$=32), and HIV-positive controls ($n$=41).
[c]eGFR was calculated according to the modified Modification of Diet in Renal Disease formula.
[d]HIV-associated nephropathy (collapsing FSGS).
[e]See Table 3 for histologic diagnosis for HIV-positive and HIV-negative CKD.
[f]HIV-positive controls with no kidney disease.
[g]Kruskal–Wallis test comparing HIV-positive patients with HIV-positive controls.
[h]Population controls were from the Division of Human Genetics, National Health Laboratory Service, School of Pathology, Faculty of Health Sciences, University of the Witwatersrand.

**Table 2.** *APOL1* genotype distribution among patients with HIVAN, HIV-positive patients with CKD, HIV-negative patients with CKD, and controls

| Genotype | HIV-Positive Patients and Controls | | | HIV-Negative Patients and Controls | |
|---|---|---|---|---|---|
| | HIVAN[a] | CKD[b] | HIV-Positive Controls | CKD[c] | Population Controls |
| 0 risk alleles | | | | | |
| G0/G0 | 2 (5.3) | 22 (56.4) | 34 (63.0) | 25 (64.1) | 36 (66.7) |
| 1 risk allele | 6 (15.8) | 11 (28.2) | 18 (33.3) | 13 (33.3) | 17 (31.5) |
| G0/G1 | 5 (13.2) | 4 (10.3) | 4 (7.4) | 6 (15.4) | 7 (13.0) |
| G0/G2 | 1 (2.6) | 7 (17.9) | 14 (25.9) | 7 (17.9) | 10 (18.5) |
| 2 risk alleles | 30 (78.9) | 6 (15.4) | 2 (3.7) | 1 (2.6) | 1 (1.9) |
| G1/G1 | 8 (21.0) | 2 (5.1) | 0 (0.0) | 0 (0.0) | 0 (0.0) |
| G1/G2 | 19 (50.0) | 2 (5.1) | 0 (0.0) | 1 (2.6) | 0 (0.0) |
| G2/G2 | 3 (7.9) | 2 (5.1) | 2 (3.7) | 0 (0.0) | 1 (1.9) |
| Total | 38 | 39 | 54 | 39 | 54 |

Data are given as $n$ (%). The distribution of *APOL1* genotypes and fractions with one or two risk alleles for the study groups is shown.
[a]One patient did not pass genotyping.
[b]The single HIV-positive patient with CKD carrying the G1$^{+M}$ (A-G-I) haplotype is excluded from this table and one patient did not pass genotyping.
[c]Two patients did not pass genotyping.

comparing HIV-negative patients with CKD with population controls (Tables 2).

We explored the distribution of *APOL1* G1 and G2 alleles among the histologically diagnosed HIV-positive and HIV-negative patients with CKD (Table 3, Supplemental Table 1). For the HIV-positive group, there was no significant association between *APOL1* genotypes and HIV-positive patients with FSGS (OR, 2.13; 95% CI, 0.03 to 44.30; $P$=0.48) and HIV-associated immune complex kidney disease (HIVICK) (OR=5.60; 95% CI, 0.4 to 86; $P$=0.13) compared with HIV-positive controls. There was also no significant association for primary FSGS in the HIV-negative group (OR, 6.30; 95% CI, 0.04 to 248.70; $P$=0.26) compared with population controls (Tables 3 and 4). To increase statistical power, we combined HIV-positive patients with FSGS ($n$=13) and primary FSGS ($n$=9) and compared this with all controls ($n$=108) and still did not observe any significant association with *APOL1* G1 and G2 alleles in a recessive model ($P$=0.20) (Table 4). Because the point ORs for *APOL1* association with primary FSGS and HIVAN differed in this study from those reported by Kopp *et al.*,[13] we used a Woolf test for homogeneity to determine whether the ORs were significantly different. Although the ORs in this study were higher for HIVAN and lower for FSGS compared with the ratios observed in African Americans, the differences were not statistically significant ($P$=0.56 and $P$=0.21 for FSGS and HIVAN, respectively). We also compared ORs for HIV-positive FSGS seen in this study with data from Fine *et al.*[20] for African Americans with HIV-positive FSGS compared with a population control group from the study by Kopp *et al.*[13] (OR, 24.5; 95% confidence interval [95% CI], 9.7 to 65.1); $P=4.3\times10^{-14}$); the Woolf test gave a $P$ value of 0.07.

We assessed the effect size and statistical significance for carriage of one or two *APOL1* risk variants by comparing the distribution of two versus one or zero copies of *APOL1* G1 and G2 risk alleles in patients and controls in a recessive model (Table 4). Consistent with previous studies, the association of

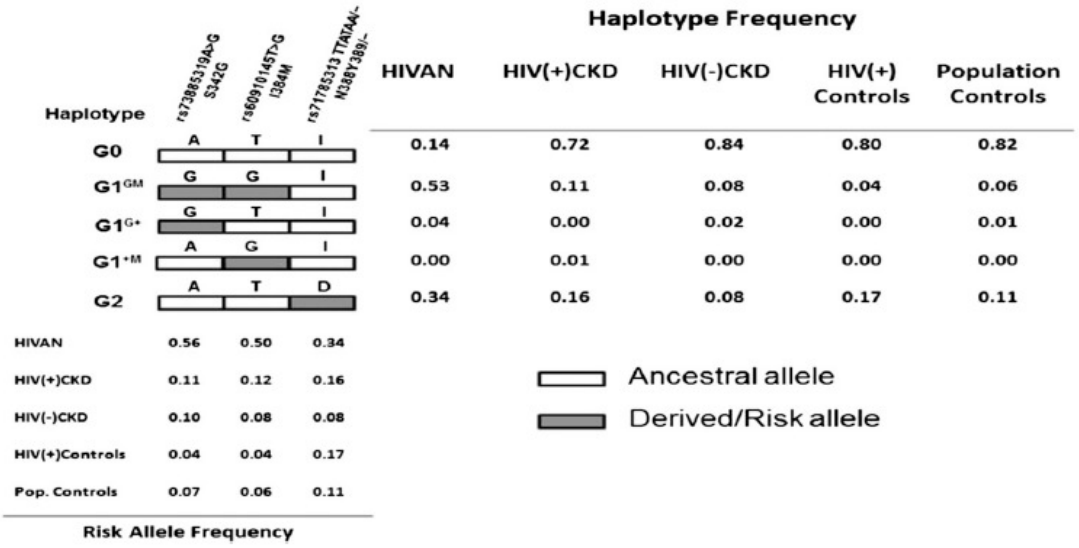



**Figure 1.** Distribution of *APOL1* haplotypes and risk alleles. Five *APOL1* haplotypes are observed. The ancestral haplotype, denoted as G0, has three ancestral alleles. The $G1^{GM}$ haplotype has two missense alleles. The $G1^{G+}$ haplotype has one missense risk allele at rs73885319. The $G1^{+M}$ has one missense variant at rs60910145. The G2 haplotype has the 6-bp deletion risk allele at rs71785313. The risk allele and haplotype frequencies are shown for patients with HIVAN, HIV-positive patients with CKD and HIV-negative patients with CKD, and HIV-positive controls and population controls (PCs).

*APOL1* genotypes with HIVAN best fits a recessive genetic model. Carriage of two copies of *APOL1* renal risk alleles was a strong predictor of HIVAN (OR, 89.1; 95% CI, 17.68 to 911.72; $P=1.24\times10^{-14}$). The predictor was so strong that carriage of even one renal risk allele increased the risk of HIVAN with marginal statistical support (OR, 5.49; 95% CI, 0.87 to 61.14; $P=0.05$) (Table 4). The results for HIVAN remained highly significant after correcting for age, sex, and genetic ancestry (Table 4).

Examination of the independent effect sizes (ORs) of the G1 and G2 alleles (Figure 2) indicated that the effect size of

**Table 3.** Association between *APOL1* risk alleles and various glomerular diseases

| Glomerular Disease | HIV-Positive Patients and Controls | | | | | HIV-Negative Patients and Controls | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | No. of *APOL1* Risk Alleles | | | OR (95% CI) 2 versus 1 or 0 risk alleles | *P* Value | No. of *APOL1* Risk Alleles | | | OR (95% CI) 2 versus 1 or 0 risk alleles | *P* Value |
| | 0 | 1 | 2 | | | 0 | 1 | 2 | | |
| Controls (*n*=108) | 34 (63) | 18 (33) | 2 (4) | — | — | 36 (67) | 17 (32) | 1 (2) | — | — |
| HIVAN (*n*=38) | 2 (5) | 6 (16) | 30 (79) | 89 (17.7 to 912) | $1.2\times10^{-14}$ | — | — | — | — | — |
| Other CKD (*n*=78) | 22 (57) | 11 (28) | 6 (15) | 3.8 (0.6 to 42) | 0.13 | 25 (64) | 13 (33) | 1(3) | 1.4 (0.02 to 11) | >0.99 |
| FSGS (*n*=22) | 9 (69) | 3 (23) | 1 (8) | 2.1 (0.03 to 44) | 0.48 | 5 (56) | 3 (33) | 1 (11) | 6 0.3 (0.08 to 527) | 0.26 |
| HIVICK (*n*=12) | 4 (33) | 5 (42) | 3 (25) | 5.6 (0.4 to 86) | 0.13 | — | — | — | — | — |
| Other GN (*n*=27)[a] | 7 (70) | 3 (30) | 0 (0) | 0.0 (0 to 30) | >0.99 | 10 (59) | 7 (41) | 0 (0) | 0 (0 to 124) | >0.99 |
| Other kidney diseases (*n*=17)[b] | 2 (50) | 0 (0) | 2 (50) | 21 (0.2 to 2029) | 0.11 | 10 (77) | 3 (23) | 0 (0) | 0 (0 to 210) | >0.99 |

Data are given as n (%) unless otherwise indicated. The *APOL1* genotype frequencies and associations, tested with the Fisher exact test, are shown for various glomerular diseases among HIV-positive and HIV-negative patients and general population controls. The only glomerular disease that showed a significant association with *APOL1* risk alleles was HIVAN. Of note, HIV-positive FSGS, HIV-negative FSGS, and HIVICK were not significantly associated with *APOL1* risk alleles, although group sizes were small. Analyses were adjusted for age, sex, and ancestry. *P* values indicate probability by the Fisher exact test.

[a]Other GN is as follows: HIV-positive: membranoproliferative GN (*n*=1), membranoproliferative GN consistent with C3 glomerulopathy (*n*=1), and membranous GN (*n*=8); and HIV-negative: IgA nephropathy (*n*=1), lupus nephritis (*n*=3), membranoproliferative GN (*n*=2), and membranous GN (*n*=11).

[b]Other kidney diseases are as follows: HIV-positive: benign nephrosclerosis (*n*=1), global glomerulosclerosis (*n*=1), minimal change disease (*n*=1), and thrombotic microangiopathy (*n*=1); and HIV-negative: benign nephrosclerosis (*n*=1), oxalosis (*n*=1), global glomerulosclerosis (*n*=2), minimal change disease (*n*=5), nodular glomerulosclerosis (*n*=1), severe arterial nephrosclerosis with secondary FSGS (*n*=1), and thrombotic microangiopathy (*n*=2).

**Table 4.** *APOL1* risk allele associations with HIVAN, HIV-associated FSGS, and HIV-negative FSGS (primary FSGS)

| No. of Risk Alleles | HIV-Positive CKD Patient Groups (*n*=51) versus HIV-Positive Controls (*n*=54) | | | | | | Primary FSGS (*n*=9) versus Population Controls (*n*=54) | | All FSGS (*n*=22) versus All Controls (*n*=108) | |
|---|---|---|---|---|---|---|---|---|---|---|
| | HIVAN (*n*=38) | | | HIV-Positive FSGS (*n*=13) | | | Primary FSGS (*n*=9) | | HIV-Positive FSGS (*n*=13) and Primary FSGS (*n*=9) | |
| | OR (95% CI) | *P* Value | Adjusted *P* Value[a] | OR (95% CI) | *P* Value | | OR (95% CI) | *P* Value | OR (95% CI) | *P* Value |
| 1 versus 0 | 5.49 (0.87 to 61.14) | 0.05 | 0.13 | 0.63 (0.10 to 2.98) | 0.74 | | 1.3 (0.18 to 7.42) | >0.99 | 0.86 (0.25 to 2.64) | >0.99 |
| 2 versus 0 | 200.87 (27.62 to 3119.17) | $1.5\times10^{-14}$ | $1.9\times10^{-3}$ | 1.85 (0.03 to 39.57) | 0.53 | | 6.70 (0.08 to 577.00) | 0.26 | 3.27 (0.25 to 31.46) | 0.22 |
| 2 versus 1 | 40.50 (7.09 to 446.94) | $8.7\times10^{-08}$ | $1.8\times10^{-4}$ | 2.82 (0.04 to 73.20) | 0.44 | | 5.0 (0.05 to 467.90) | 0.34 | 3.74 (0.26 to 40.67) | 0.20 |
| 2 versus 1 or 0 recessive model | 89.1 (17.68 to 911.72) | $1.2\times10^{-14}$ | $1.5\times10^{-5}$ | 2.13 (0.03 to 44.30) | 0.48 | | 6.30 (0.04 to 248.70) | 0.26 | 3.45 (0.27 to 32.23) | 0.20 |

ORs (95% CIs) are shown for the association between *APOL1* risk alleles and particular glomerular diseases. *P* values are from Fisher's exact test unless otherwise noted.
[a]Analyses were adjusted for age, sex, and ancestry.

G2 was weaker than G1; however, it is worth noting that G1/G2 compound heterozygotes are equivalent in effect size to G1/G1.

Population stratification analyses using a principal components analysis (PCA) showed that the distribution of patients with HIVAN and HIV-positive control groups was not significantly different along the top-two eigen axes ($P>0.05$), confirming that the high association observed is not due to an undetected population substructure that might cause false positive associations (Supplemental Figure 1).

## DISCUSSION

This is the first study that correlates genotypic information with phenotypes in HIV-associated CKD in Africa, a region with the highest prevalence of HIV-1 infection.[21] The study shows an extremely strong association between *APOL1* variants, found only on African chromosomes, and biopsy-confirmed HIVAN, an aggressive form of kidney disease that rapidly progresses to ESRD if untreated. Among the patients with HIVAN, 79% carried two copies of *APOL1* risk alleles compared with 2% in the general population. Notably, the effect size of *APOL1* risk alleles for HIVAN in black South Africans (OR, 89.1; 95% CI, 17.8 to 911.72; $P=1.24\times10^{-14}$) was numerically higher than in African Americans (OR, 29.2; 95% CI, 13.1 to 68.5; $P=6.0\times10^{-22}$), although the CIs overlapped.[13] To our knowledge, this is the strongest effect size ever reported for common variants with complex diseases.[22]

Although there is no statistical difference between effect sizes in South Africans with HIVAN compared with African Americans with HIVAN,[13] both the frequency of *APOL1* risk alleles (79% versus 72%) and ORs (89 versus 29) are elevated in South Africa. Although these differences may reflect statistical fluctuation, *APOL1* may also have a stronger gene–environment interaction in this setting, which warrants further investigation. First, this strong effect might be related to HIV-1 subtype C. HIV-1 subtype C is the dominant strain in Southern and Eastern Africa and South Asia, whereas HIV-1 subtype A is common in West Africa and subtype B is dominant in Europe, Australia, and the Americas.[23] HIV-1 subtype C is highly virulent and accounts for approximately 50% of all HIV infections worldwide and 98% of HIV infections in South Africa, with corresponding higher viral loads.[23,24] High HIV RNA levels are correlated with a decline in CD4 T cells and with the development of HIVAN in individuals of African ancestry.[25] Second, in resource-limited settings, late initiation of ART may predispose at-risk individuals to HIVAN; several studies have shown that effective rollout of ART can reduce the occurrence of HIVAN.[6,7,9] Furthermore, all of the HIV-positive patients with CKD in this study were ART naïve. Third, there are regional differences in nutrition, environmental exposures, and other viral infections that may modify the effects of *APOL1* variants on the kidney. Fourth, there might be other genetic variants unique to or more common in this population that interacts with *APOL1*-driven susceptibility to HIVAN.

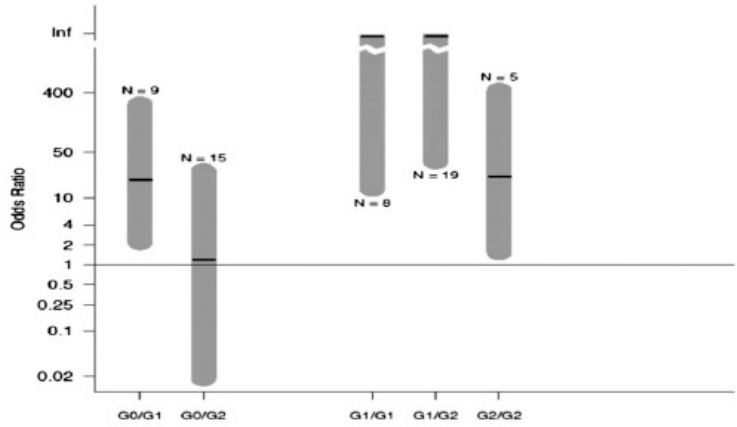



**Figure 2.** ORs for the effect sizes for HIVAN with different combination of *APOL1* G1 and G2 alleles. ORs with CIs for association of different strata of G1–G2 genotypes with HIVAN patients compared with participants carrying no G1 or G2 risk alleles (+/+). For G1/G1 and G1/G2, the OR is infinity, because these genotypes were only observed in patients.

This study also showed a weak effect with one *APOL1* risk allele in HIVAN (OR, 5.49; 95% CI, 0.87 to 61.14; $P$=0.05), consistent with previous studies.[12,13,26] An effect from one risk allele would be consistent with a gain of injury and/or toxicity of these variants in kidney cells that manifests with a strong environmental insult, such as HIV infection, as opposed to a purely recessive effect that might suggest a loss of function.

There was little or no significant *APOL1* association with multiple histologic phenotypes of HIV-positive patients with CKD and HIV-negative patients with CKD. Previous studies in African Americans have reported ORs for primary FSGS with two *APOL1* risk alleles as 10.5 (95% CI, 6.0 to 18.1) and 17 (95% CI, 11 to 26.5; $P=1.3\times10^{-48}$).[12,13] In addition, other studies also demonstrated higher association of *APOL1* risk alleles with FSGS.[20,27] It is possible that HIV infection is an incidental finding to primary FSGS in black South Africans, with FSGS being the commonest histology reported in children and adults worldwide.[28,29] However, others have suggested that FSGS may be part of the spectrum of histology in response to HIV infection.[30] It is also possible that due to the high HIV prevalence in South Africa, susceptible HIV-infected individuals who are ART naïve may develop HIVAN and not FSGS. In addition, we found that *APOL1* risk allele status was not associated with HIVICK, a result in concordance with that of a study in African Americans that observed two *APOL1* risk alleles in only 1 of 31 patients with HIVICK.[31]

HIV is a powerful driver of *APOL1* nephropathy, as seen by the large proportion of individuals with HIVAN who carry two risk allele genotypes and the strong effect sizes observed. Furthermore, the frequencies were similar for two risk allele genotypes in HIVAN (79% in this study and 72% in Kopp *et al.*[13] study), although the two risk allele genotype frequencies are markedly lower in the South-African (2%–4%) compared with the African-American general population (12%–14%).[13] It was previously estimated that 50% of HIV-positive patients with two *APOL1* risk alleles who do not receive ART will develop HIVAN in their lifetime, suggesting the magnitude of the interaction.[13] Nevertheless, it is worth noting that HIV-positive patients with two risk alleles do not necessarily develop HIVAN and that approximately 20% of patients with HIVAN carry one or no *APOL1* risk allele, suggesting that there are more factors to be considered. Clearly, HIV infection interacts strongly with *APOL1* variants, although the molecular mechanism is largely unknown, and this interaction is similar in South Africa and the United States.

The *APOL1* G1 and G2 alleles show distinct distributions among various African and African-derived populations. With a robust population sample comprising 216 chromosomes (including HIV-positive controls), we were able to establish that the G1 allele (rs73885319) frequency is approximately 7.3%, much lower than reported in West Africa (Yoruba and Igbo from Nigeria), in whom the frequencies are approximately 45% and 30%, respectively, or in African Americans where the G1 frequency is approximately 20%.[13,17] This may be as a result of relaxation of selection pressures exerted by trypanosomiasis or other pathogens in regions inhabited by ancestors of this South-African population or may represent introgression of the variant alleles into southern African populations as a result of gene flow from the Bantu expansion.[32,33] In this study, G1 and G2 alleles have a combined allele frequency of 18.4%, indicating that in the general black population, 32% of individuals carry one risk allele and 2% carry two risk alleles. The risk of developing HIVAN is substantial for individuals living with HIV infection and possessing two *APOL1* risk alleles.[13,27] We also detected a rare, recombinant haplotype, $G1^{+M}$ (A-G-I), in the heterozygous state ($G1^{GM}/G1^{+M}$), in a single individual with HIV-positive CKD, which has not been previously reported in the 1000 Genomes Project or in previous studies of African populations, summarized by Thomson *et al.*[33] and Limou *et al.*[34]

The molecular mechanisms by which *APOL1* variants cause kidney disease remain unclear. This protein is expressed in podocytes and proximal tubular epithelial cells in normal kidneys and arteriolar endothelial cells and in vascular smooth muscle cells in HIVAN and FSGS biopsies.[35] The G1 and G2 variant proteins, respectively, have altered and deleted amino acids in the *APOL1*-binding region of the trypanosomal serum resistance–associated (SRA) protein.[12] By so doing, the trypanosomal SRA is not able to bind *APOL1*, and hence is exposed to the trypanolytic activity of ApoL1. This suggests that the risk variants

in the SRA domain of *APOL1* act by a gain of injury or a toxicity mechanism on podocytes or vascular cells.

A study by Nichols *et al.*[36] showed that despite the recessive mode of inheritance observed in *APOL1* kidney disease, there is evidence suggesting that *APOL1* risk variants act as toxic gain-of-function mutations and that HIV triggers the immune system by activating exogenous IFNs that in turn increase ApoL1 expression in cells, leading to the development of kidney disease.[36] Having carried out our study in a high HIV prevalence setting, with predominantly HIV-1 subtype C, puts HIV as an important environmental factor inducing *APOL1*-mediated kidney disease, supporting a two-hit model for *APOL1*-associated disease expression.[37] This study highlights the need to consider gene–environment interaction in the interpretation of genome-wide association studies for other common diseases.

Our study also highlights the need for HIV screening, surveillance, and implementation of World Health Organization (WHO) recommendations for earlier ART initiation to reduce the burden of HIVAN and other forms of HIV-related kidney disease in Africa. At present, guidelines for the treatment of HIV infection in South Africa suggest that ART be instituted for individuals with WHO clinical stage 3 and 4 disease and HIV-specific manifestations (*e.g.,* HIV-related malignancies, kidney or other organ involvement, and hepatitis B and C infection) and in all HIV-positive individuals with CD4 counts $<350$ cells/$\mu$l.[38]

There are several limitations to our study. We lacked sufficient numbers of primary or HIV-positive patients with FSGS to detect the effects of the G1 and G2 risk alleles (a sample size of 95 patients with FSGS and 95 controls would be required for 80% power to detect an OR of 5, $\alpha$ 0.05). Our control group comprised healthy population controls who were assumed to have normal kidney function; however, kidney function was not assessed. Although it is theoretically possible that undetected renal disease could reduce power, random controls are commonly used in genetic studies and it is unlikely that healthy individuals in this group had undetected glomerulopathy particularly given the low frequency of individuals carrying two *APOL1* renal risk alleles in our control group.

In conclusion, this study shows an extremely strong association between *APOL1* variants and biopsy-confirmed HIVAN driven by interaction between *APOL1* and untreated HIV infection, the first such study in Southern Africa.

## CONCISE METHODS

### Study Participants

Archived kidney biopsies obtained from patients aged >18 years were retrieved from the Division of Anatomic Pathology at the University of the Witwatersrand, according to their histologic diagnosis and HIV status (Table 1). Renal histologic diagnoses were made on the basis of standard diagnostic criteria. Histology was confirmed by an independent pathologist (A.B.F.). HIVAN was defined by the presence of glomerular capillary collapse and glomerular visceral epithelial cell proliferation affecting at least one glomerulus, together with microcystic tubular dilation and interstitial inflammation. HIVICK was defined by the presence of glomerular endocapillary cell proliferation together with an increased deposition of at least one Ig, thus including HIV-positive IgA nephropathy. Other histopathologic groups included the following: HIV-positive and HIV-negative FSGS, other GN (which included membranoproliferative GN and membranoproliferative GN consistent with C3 glomerulopathy, membranous GN, HIV-negative IgA nephropathy, and lupus nephritis,) and other kidney disease (which included thrombotic microangiopathy, nodular glomerulosclerosis, minimal change disease, arterionephrosclerosis and benign nephrosclerosis, oxalosis, and global glomerulosclerosis).

Blood samples were collected prospectively from 54 volunteers of black ethnicity and of similar age and sex to the patients with CKD for each of the two control groups: a control group (HIV-positive controls) comprising HIV-positive individuals with no clinical evidence of CKD (normal renal function; no proteinuria) from Charlotte Maxeke Johannesburg Academic Hospital and a population control group from the National Health Laboratory Service who were all apparently healthy individuals who gave consent to be enrolled for population genetic studies as control participants. All HIV-positive individuals with CKD were ART naïve and 75% of the HIV-positive control group was receiving ART. Exclusion criteria included patients aged <18 years, patients with a history of diabetes, and persons who were nonblack South Africans. This study was approved by the University of the Witwatersrand Human Research Ethics Committee (medical; approval numbers M111185 and M10745). All prospective participants (controls) provided signed informed consent.

### Genomic DNA Extraction from Formalin-Fixed Paraffin-Embedded Kidney Tissue and Peripheral Blood Samples

The All Prep QIAamp DNA formalin-fixed paraffin-embedded tissue kit (Qiagen, Chatsworth, CA) was used to purify genomic DNA from formalin-fixed paraffin-embedded kidney samples, and a modified in-house salting-out procedure[39] was used to extract genomic DNA from peripheral blood samples.

### SNP Genotyping

The *APOL1* SNPs G1 (rs73885319 and rs60910145) and the G2 indel (rs71785313) were genotyped using the TaqMan SNP genotyping system (Applied Biosystems, Foster City, CA), whereas 42 SNPs (18 ancestry informative markers and 24 other SNPs from chromosome 22 not in linkage disequilibrium with the risk variants) were genotyped using the Illumina BeadXpress SNP system (Illumina Inc., San Diego, CA). Sanger sequencing was used to confirm unusual haplotype configurations. Each SNP was tested for deviation from Hardy–Weinberg equilibrium using a chi-squared goodness-of-fit test in the population control group.

### Population Stratification

Population stratification was assessed based on 42 SNPs distributed throughout the genome including chromosome 22 using the PCA module of Eigensoft software.[40] For these analyses, chromosome 22

SNPs were limited to $r^2<0.05$ with each other. SNPs defining the *APOL1* risk genotypes were excluded from the population stratification analyses.

### Terminology

For *APOL1* haplotypes, we used the terms $G1^{G+}$ (carrying the derived risk allele at S342G only), $G1^{GM}$ (containing the derived risk alleles at S342G and I384M), and $G1^{+M}$ (containing only the derived allele at I384M). It has not been determined whether I384M confers risk in the absence of S342G.[13] For *APOL1* alleles, we use the terms G0 for the ancestral, nonrisk (+/+) allele, G1 for the risk allele of S342G that occurs on two G1 haplotypes ($G1^{G+}$ and $G1^{GM}$), and G2 for the deletion allele (Figure 1).

### Statistical Analyses

The effect of *APOL1* risk genotypes on kidney disease was tested by using the Fisher's exact test, and by logistic regression adjusting for age, sex, and the first two eigen values obtained from a PCA of genetic ancestry. Informed by previous association results, we primarily tested recessive models (comparing individuals with two risk alleles, *i.e.*, G1/G1, G1/G2, or G2/G2, against all others), but also tested other models (dominant; and one versus zero, two versus zero, two versus one, and two versus one or zero risk alleles). Comparisons were made between HIV-positive patients and HIV-positive controls and between HIV-negative patients and population controls, comparing the distribution of two risk alleles (explanatory exposure) to zero or two risk alleles (no exposure). The Woolf test was used to test for homogeneity of the ORs in South-African and African-American FSGS populations.

## ACKNOWLEDGMENTS

The authors thank Dr. Nadia Carstens and Angela Hobbs Steyn (University of the Witwatersrand) and Victor David and Elizabeth Binns-Roemer (Frederick National Laboratory) for excellent technical support, as well as Dr. Francois Venter for access to the HIV clinic at Charlotte Maxeke Johannesburg Academic Hospital.

This project was funded by the National Institutes of Health (NIH) Fogarty International Center (Grant 1D43TW008330-01A1 [July 1, 2010– June 30, 2015] Millennium Promise Award), the National Institute of Diabetes and Digestive and Kidney Diseases Intramural Research Programs, the Non-communicable Chronic Diseases Leadership Training Program, the Medical Research Council of South Africa, the National Research Foundation of South Africa, the National Health Laboratory Service Research Trust, the Columbia University–South Africa Training Program for Research on AIDS-Related Malignancies, the NIH National Cancer Institute (Grant 1D43CA153715-03), the Division of Nephrology of University of the Witwatersrand and the Faculty Research Committee individual grant, University of the Witwatersrand. This project was also funded in part with federal funds from the NIH National Cancer Institute (Contract HHSN26120080001E), and was supported in part by the Intramural Research Program of the NIH National Cancer Institute Center for Cancer Research.

The content of this publication does not necessarily reflect the views or policies of the US Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the US Government.

## DISCLOSURES

None.

## REFERENCES

1. Kitiyakara C, Kopp JB, Eggers P: Trends in the epidemiology of focal segmental glomerulosclerosis. *Semin Nephrol* 23: 172–182, 2003
2. Kopp JB, Winkler C: HIV-associated nephropathy in African Americans. *Kidney Int Suppl* 83: S43–S49, 2003
3. Mohanram A, Zhang Z, Shahinfar S, Keane WF, Brenner BM, Toto RD: Anemia and end-stage renal disease in patients with type 2 diabetes and nephropathy. *Kidney Int* 66: 1131–1138, 2004
4. Collins AJ, Foley RN, Chavers B, Gilbertson D, Herzog C, Johansen K, Kasiske B, Kutner N, Liu J, St Peter W, Guo H, Gustafson S, Heubner B, Lamb K, Li S, Li S, Peng Y, Qiu Y, Roberts T, Skeans M, Snyder J, Solid C, Thompson B, Wang C, Weinhandl E, Zaun D, Arko C, Chen SC, Daniels F, Ebben J, Frazier E, Hanzlik C, Johnson R, Sheets D, Wang X, Forrest B, Constantini E, Everson S, Eggers P, Agodoa L: United States Renal Data System 2011 Annual Data Report: Atlas of chronic kidney disease and end-stage renal disease in the United States. *Am J Kidney Dis* 59 [Suppl 1]: A7, e1–e420, 2012
5. Laurinavicius A, Hurwitz S, Rennke HG: Collapsing glomerulopathy in HIV and non-HIV patients: A clinicopathological and follow-up study. *Kidney Int* 56: 2203–2213, 1999
6. Kalayjian RC: The treatment of HIV-associated nephropathy. *Adv Chronic Kidney Dis* 17: 59–71, 2010
7. Hays T, Wyatt CM: APOL1 variants in HIV-associated nephropathy: Just one piece of the puzzle. *Kidney Int* 82: 259–260, 2012
8. Han TM, Naicker S, Ramdial PK, Assounga AG: A cross-sectional study of HIV-seropositive patients with varying degrees of proteinuria in South Africa. *Kidney Int* 69: 2243–2250, 2006
9. Wearne N, Swanepoel CR, Boulle A, Duffield MS, Rayner BL: The spectrum of renal histologies seen in HIV with outcomes, prognostic indicators and clinical correlations. *Nephrol Dial Transplant* 27: 4109–4118, 2012
10. Kopp JB, Smith MW, Nelson GW, Johnson RC, Freedman BI, Bowden DW, Oleksyk T, McKenzie LM, Kajiyama H, Ahuja TS, Berns JS, Briggs W, Cho ME, Dart RA, Kimmel PL, Korbet SM, Michel DM, Mokrzycki MH, Schelling JR, Simon E, Trachtman H, Vlahov D, Winkler CA: MYH9 is a major-effect risk gene for focal segmental glomerulosclerosis. *Nat Genet* 40: 1175–1184, 2008
11. Kao WH, Klag MJ, Meoni LA, Reich D, Berthier-Schaad Y, Li M, Coresh J, Patterson N, Tandon A, Powe NR, Fink NE, Sadler JH, Weir MR, Abboud HE, Adler SG, Divers J, Iyengar SK, Freedman BI, Kimmel PL, Knowler WC, Kohn OF, Kramp K, Leehey DJ, Nicholas SB, Pahl MV, Schelling JR, Sedor JR, Thomley-Brown D, Winkler CA, Smith MW, Parekh RS; Family Investigation of Nephropathy and Diabetes Research Group: MYH9 is associated with nondiabetic end-stage renal disease in African Americans. *Nat Genet* 40: 1185–1192, 2008
12. Genovese G, Friedman DJ, Ross MD, Lecordier L, Uzureau P, Freedman BI, Bowden DW, Langefeld CD, Oleksyk TK, Uscinski Knob AL, Bernhardy AJ, Hicks PJ, Nelson GW, Vanhollebeke B, Winkler CA, Kopp JB, Pays E, Pollak MR: Association of trypanolytic ApoL1 variants with kidney disease in African Americans. *Science* 329: 841–845, 2010
13. Kopp JB, Nelson GW, Sampath K, Johnson RC, Genovese G, An P, Friedman D, Briggs W, Dart R, Korbet S, Mokrzycki MH, Kimmel PL,

Limou S, Ahuja TS, Berns JS, Fryc J, Simon EE, Smith MC, Trachtman H, Michel DM, Schelling JR, Vlahov D, Pollak M, Winkler CA: APOL1 genetic variants in focal segmental glomerulosclerosis and HIV-associated nephropathy. *J Am Soc Nephrol* 22: 2129–2137, 2011

14. Tzur S, Rosset S, Shemer R, Yudkovsky G, Selig S, Tarekegn A, Bekele E, Bradman N, Wasser WG, Behar DM, Skorecki K: Missense mutations in the APOL1 gene are highly associated with end stage kidney disease risk previously attributed to the MYH9 gene. *Hum Genet* 128: 345–350, 2010
15. Parsa A, Kao WH, Xie D, Astor BC, Li M, Hsu CY, Feldman HI, Parekh RS, Kusek JW, Greene TH, Fink JC, Anderson AH, Choi MJ, Wright JT Jr, Lash JP, Freedman BI, Ojo A, Winkler CA, Raj DS, Kopp JB, He J, Jensvold NG, Tao K, Lipkowitz MS, Appel LJ; AASK Study InvestigatorsCRIC Study Investigators: APOL1 risk variants, race, and progression of chronic kidney disease. *N Engl J Med* 369: 2183–2196, 2013
16. Ko WY, Rajan P, Gomez F, Scheinfeldt L, An P, Winkler CA, Froment A, Nyambo TB, Omar SA, Wambebe C, Ranciaro A, Hirbo JB, Tishkoff SA: Identifying Darwinian selection acting on different human APOL1 variants among diverse African populations. *Am J Hum Genet* 93: 54–66, 2013
17. Ulasi II, Tzur S, Wasser WG, Shemer R, Kruzel E, Feigin E, Ijoma CK, Onodugo OD, Okoye JU, Arodiwe EB, Ifebunandu NA, Chukwuka CJ, Onyedum CC, Ijoma UN, Nna E, Onuigbo M, Rosset S, Skorecki K: High population frequencies of APOL1 risk variants are associated with increased prevalence of non-diabetic chronic kidney disease in the Igbo people from south-eastern Nigeria. *Nephron Clin Pract* 123: 123–128, 2013
18. Tzur S, Rosset S, Skorecki K, Wasser WG: APOL1 allelic variants are associated with lower age of dialysis initiation and thereby increased dialysis vintage in African and Hispanic Americans with non-diabetic end-stage kidney disease. *Nephrol Dial Transplant* 27: 1498–1505, 2012
19. Behar DM, Kedem E, Rosset S, Haileselassie Y, Tzur S, Kra-Oz Z, Wasser WG, Shenhar Y, Shahar E, Hassoun G, Maor C, Wolday D, Pollack S, Skorecki K: Absence of APOL1 risk variants protects against HIV-associated nephropathy in the Ethiopian population. *Am J Nephrol* 34: 452–459, 2011
20. Fine DM, Wasser WG, Estrella MM, Atta MG, Kuperman M, Shemer R, Rajasekaran A, Tzur S, Racusen LC, Skorecki K: APOL1 risk variants predict histopathology and progression to ESRD in HIV-related kidney disease. *J Am Soc Nephrol* 23: 343–350, 2012
21. World Health Organization, Joint United Nations Programme on HIV/AIDS, United Nations Children's Fund: Global HIV/AIDS Response: Epidemic Update and Health Sector Progress Towards Universal Access, Geneva, Switzerland, World Health Organization, 2011
22. Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, Manolio TA: Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci U S A* 106: 9362–9367, 2009
23. Hemelaar J, Gouws E, Ghys PD, Osmanov S: Global and regional distribution of HIV-1 genetic subtypes and recombinants in 2004. *AIDS* 20: W13–W23, 2006
24. Novitsky V, Ndung'u T, Wang R, Bussmann H, Chonco F, Makhema J, De Gruttola V, Walker BD, Essex M: Extended high viremics: A substantial fraction of individuals maintain high plasma viral RNA levels after acute HIV-1 subtype C infection. *AIDS* 25: 1515–1522, 2011
25. Izzedine H, Wirden M, Launay-Vacher V: Viral load and HIV-associated nephropathy. *N Engl J Med* 353: 1072–1074, 2005
26. Kanji Z, Powe CE, Wenger JB, Huang C, Ankers E, Sullivan DA, Collerone G, Powe NR, Tonelli M, Bhan I, Bernhardy AJ, Dibartolo S, Friedman D, Genovese G, Pollak MR, Thadhani R: Genetic variation in APOL1 associates with younger age at hemodialysis initiation. *J Am Soc Nephrol* 22: 2091–2097, 2011
27. Papeta N, Kiryluk K, Patel A, Sterken R, Kacak N, Snyder HJ, Imus PH, Mhatre AN, Lawani AK, Julian BA, Wyatt RJ, Novak J, Wyatt CM, Ross MJ, Winston JA, Klotman ME, Cohen DJ, Appel GB, D'Agati VD, Klotman PE, Gharavi AG: APOL1 variants increase risk for FSGS and HIVAN but not IgA nephropathy. *J Am Soc Nephrol* 22: 1991–1996, 2011
28. Adhikari M, Bhimma R, Coovadia HM: Focal segmental glomerulosclerosis in children from KwaZulu/Natal, South Africa. *Clin Nephrol* 55: 16–24, 2001
29. Chou YH, Lien YC, Hu FC, Lin WC, Kao CC, Lai CF, Chiang WC, Lin SL, Tsai TJ, Wu KD, Chen YM: Clinical outcomes and predictors for ESRD and mortality in primary GN. *Clin J Am Soc Nephrol* 7: 1401–1408, 2012
30. Meehan SM, Kim L, Chang A: A spectrum of morphologic lesions of focal segmental glomerulosclerosis by Columbia criteria in human immunodeficiency virus infection. *Virchows Arch* 460: 429–435, 2012
31. Foy MC, Estrella MM, Lucas GM, Tahir F, Fine DM, Moore RD, Atta MG: Comparison of risk factors and outcomes in HIV immune complex kidney disease and HIV-associated nephropathy. *Clin J Am Soc Nephrol* 8: 1524–1532, 2013
32. de Filippo C, Barbieri C, Whitten M, Mpoloka SW, Gunnarsdóttir ED, Bostoen K, Nyambe T, Beyer K, Schreiber H, de Knijff P, Luiselli D, Stoneking M, Pakendorf B: Y-chromosomal variation in sub-Saharan Africa: Insights into the history of Niger-Congo groups. *Mol Biol Evol* 28: 1255–1269, 2011
33. Thomson R, Genovese G, Canon C, Kovacsics D, Higgins MK, Carrington M, Winkler CA, Kopp J, Rotimi C, Adeyemo A, Doumatey A, Ayodo G, Alper SL, Pollak MR, Friedman DJ, Raper J: Evolution of the primate trypanolytic factor APOL1. *Proc Natl Acad Sci U S A* 111: E2130–E2139, 2014
34. Limou S, Nelson GW, Kopp JB, Winkler CA: APOL1 kidney risk alleles: Population genetics and disease associations. *Adv Chronic Kidney Dis* 21: 426–433, 2014
35. Madhavan SM, O'Toole JF, Konieczkowski M, Ganesan S, Bruggeman LA, Sedor JR: APOL1 localization in normal kidney and nondiabetic kidney disease. *J Am Soc Nephrol* 22: 2119–2128, 2011
36. Nichols B, Jog P, Lee JH, Blackler D, Wilmot M, D'Agati V, Markowitz G, Kopp JB, Alper SL, Pollak MR, Friedman DJ: Innate immunity pathways regulate the nephropathy gene Apolipoprotein L1 [published online ahead of print August 6, 2014]. *Kidney Int* doi:10.1038/ki.2014.270
37. Freedman BI, Skorecki K: Gene-gene and gene-environment interactions in apolipoprotein L1 gene-associated nephropathy. *Clin J Am Soc Nephrol* 9: 2006–2013, 2014
38. Meintjes G, Maartens G; Southern African HIV Clinicians Society: Guidelines for antiretroviral therapy in adults. *S Afr J HIV Med* 13: 114–133, 2012
39. Nasiri H, Forouzandeh M, Rasaee MJ, Rahbarizadeh F: Modified salting-out method: High-yield, high-quality genomic DNA extraction from whole blood using laundry detergent. *J Clin Lab Anal* 19: 229–232, 2005
40. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D: Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 38: 904–909, 2006

This article contains supplemental material online at http://jasn.asnjournals.org/lookup/suppl/doi:10.1681/ASN.2014050469/-/DCSupplemental.

## APPENDIX 2

**Name of journal:** *World Journal of Nephrology*
**ESPS Manuscript NO: 12457**
**Columns: MINIREVIEWS**

**African origins and chronic kidney disease susceptibility in the HIV era**

Kasembeli AN *et al*. CKD genetics in the HIV era

Alex N Kasembeli, Raquel Duarte, Michèle Ramsay, Saraladevi Naicker

**Alex N Kasembeli, Raquel Duarte, Saraladevi Naicker,** Department of Internal Medicine, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg 2193, South Africa

**Michèle Ramsay,** Division of Human Genetics, National Health Laboratory Service and School of Pathology, Faculty of Health Sciences and The Sydney Brenner Institute for Molecular Bioscience, University of the Witwatersrand, Johannesburg 2193, South Africa

**Saraladevi Naicker,** Division of Nephrology, University of the Witwatersrand, Johannesburg 2193, South Africa

**Author contributions:** Kasembeli AN, Duarte R, Ramsay M and Naicker S planned the review, wrote specific sections and reviewed the paper; Kasembeli AN was responsible for producing the first complete draft.

**Supported by** NIH Fogarty International Center Grant, No. 1D43TW008330-01A1, Millennium Promise Award, Non-communicable Chronic Diseases Leadership Training Program, NHLS Research Trust, Division of Nephrology Research Fund, University of the Witwatersrand, Johannesburg and FRC Individual grant, University of the Witwatersrand, Johannesburg.

**Conflict-of-interest**: None.

**Open-Access**: This article is an open-access article which was selected by an in-house editor and fully peer-reviewed by external reviewers. It is distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/

**Correspondence to: Saraladevi Naicker, Professor,** Department of Internal Medicine, Faculty of Health Sciences, University of the Witwatersrand, 7 York Road, Parktown, Johannesburg 2193, South Africa. saraladevi.naicker@wits.ac.za

**Telephone**: +27-11-717 2129

**Fax**: +27-11-643 8777

**Received**: July 9, 2014

**Peer-review started**: July 10, 2014

**First decision**: October 27, 2014

**Revised**: January 6, 2015

**Accepted**: Janurary 18, 2015

**Article in press**:

**Published online**:

**Abstract**

Chronic kidney disease (CKD) is a major public health problem worldwide with the estimated incidence growing by approximately 6% annually. There are striking ethnic differences in the prevalence of CKD such that, in the United States, African Americans have the highest prevalence of CKD, four times the incidence of end stage renal disease when compared to Americans of European ancestry suggestive of genetic predisposition. Diabetes mellitus, hypertension and human immunodeficiency virus (HIV) infection are the major causes of CKD. HIV-

associated nephropathy (HIVAN) is an irreversible form of CKD with considerable morbidity and mortality and is present predominantly in people of African ancestry. The *APOL1* G1 and G2 alleles were more strongly associated with the risk for CKD than the previously examined *MYH9* E1 risk haplotype in individuals of African ancestry. A strong association was reported in HIVAN, suggesting that 50% of African Americans with 2 *APOL1* risk alleles, if untreated, would develop HIVAN. However these two variants are not enough to cause disease. The prevailing belief is that modifying factors or second hits (including genetic hits) underlie the pathogenesis of kidney disease. This work reviews the history of genetic susceptibility of CKD and outlines current theories regarding the role for *APOL1* in CKD in the HIV era.

**Key words:** Chronic kidney disease; Human immunodeficiency virus; Human immunodeficiency virus-associated nephropathy; African ancestry; Genetics; *APOL1*; *MYH9*

© **The Author(s) 2015.** Published by Baishideng Publishing Group Inc. All rights reserved.

**Core tip:** There are striking ethnic differences in the prevalence of chronic kidney disease CKD, including human immunodeficiency virus (HIV)-associated nephropathy (HIVAN), in people of African ancestry suggestive of genetic predisposition. The *APOL1* G1 and G2 alleles are more strongly associated with the risk for HIVAN than the previously reported *MYH9* E1 risk haplotype in individuals of African ancestry in United States and African studies. The high prevalence of HIVAN among individuals of African ancestry could be a result of high frequencies of *APOL1* risk variants as well as the prevalence of HIV-1 subtypes and modifying factors or second hits underlying the pathogenesis of kidney disease.

Kasembeli AN, Duarte R, Ramsay M, Naicker S. African origins and chronic kidney disease susceptibility in the HIV era. *World J Nephrol* 2015; In press

## INTRODUCTION

Chronic kidney disease (CKD) is a major public health problem worldwide[1]. Mortality due to CKD nearly doubled worldwide between 1990 and 2010, and is now positioned at 18$^{th}$ as a cause of death in the Global Burden of Disease Study[2] and at 5$^{th}$ position in South Africa[3]. An estimated 3.2 million people were on renal replacement therapy by the end of 2013, approximately 2522000 people undergoing dialysis treatment (haemodialysis or peritoneal dialysis) and 678000 people living with renal transplants[4] and it is also estimated that CKD incidence grows by approximately 6% annually[4]. There are striking ethnic differences in the prevalence of CKD such that, in the United States, African Americans have the highest prevalence of CKD[5]. Diabetes and hypertension, which have been considered the two leading causes of CKD, together with differences in clinical, social-demographic or lifestyle factors, are insufficient to account satisfactorily for the excess risk of end stage renal disease (ESRD) in African Americans[6,7]. Africa, the second largest and the world's second most populous continent, is approximately 30.2 million square km$^2$ and composed of 54 countries[8] and more than 1.1 billion people as of 2013, accounting for 15% of the world's population[9]. It has been postulated that, by 2030, ~70% of the patients with ESRD will be living in low income countries such as those in sub Saharan Africa where majority of people leave on a less than one dollar-a-day[10,11]. The increased burden of CKD in Africa could be as a result of various communicable diseases such as leishmaniasis, schistosomiasis, infectious glomerulonephritis and importantly, human immunodeficiency virus (HIV) infection superimposed on non-communicable diseases such as hypertension and diabetes mellitus. These have resulted in the increase in CKD; several studies have shown that there is a four-fold increase in CKD in HIV uninfected individuals, compared to 18-50 fold increase in CKD in HIV positive individuals of African descent[12,13].

## HIV AND CHRONIC KIDNEY DISEASE

Acquired immune deficiency syndrome (AIDS)-associated nephropathy was originally reported in AIDS patients in the United States in 1984. Subsequently, asymptomatic HIV-infected individuals showed similar clinical and histological features, and the name was later changed to HIV-associated nephropathy (HIVAN)[14]. HIVAN is an irreversible form of CKD which is a pathologically distinct complication of HIV infection with considerable morbidity and mortality[15,16].The odds of developing HIVAN have increased in the recent years to fifty according to the United States Renal Data System[17]. There is a huge regional variation in the prevalence of HIV infection; globally, an estimated 35.3 million people were living with HIV as at 2012; in North Africa, approximately 260000 people are living with HIV; while in sub-Saharan Africa, which comprises two thirds of all people living with HIV, an estimated 25 million people are living with HIV[18]. CKD occurs in approximately 6.0%-48.5% of HIV positive patients in Africa[19]. About 24-83% of these cases are classic HIVAN in South Africa[20-22]. HIVAN is a clinicopathological condition characterized by the presence of focal glomerulosclerosis with collapsing glomerulopathy and glomerular epithelial cell proliferation, together with microcystic tubular dilatation and interstitial inflammation[23]. Risk factors for HIVAN are older age, lower CD4 counts, high viral load, co-morbidity (such as diabetes mellitus, hypertension and hepatitis C co-infection)[24]. HIVAN is present predominantly in people of African ancestry, indicating a possible genetic predisposition[25,26]. Renal histology in HIV infected patients in South Africa is shown in Table 1. A 30 year review of 1848 renal biopsies by Vermeulen at Chris Hani Baragwanath Hospital in Johannesburg, South Africa found that focal segmental glomerulosclerosis (FSGS) comprised 29.6% of primary glomerulonephritis (GN), 24.4% of membranous GN, 23.8% of membranoproliferative GN, 10.3% of minimal change disease, 4.1% of mesangial proliferative GN and 2.7% of IgA nephritis; 19.7 % of the biopsies were in HIV positive individuals (Vermeulen A, MMed, University of the Witwatersrand, 2014).

The mechanism by which HIV induces glomerular injury leading to the pathologic syndrome of HIVAN is not well understood, hence a number of theories have been postulated as to how HIV causes renal injury. First, a direct viral infection of podocytes, renal parenchymal cells, especially the visceral epithelial cells of the glomerulus, and the tubular epithelial cells and as a result, this elicits cytopathic effects including proliferation and apoptosis[27]. Secondly, HIV infects the lymphocytes and macrophages that enter the kidney, resulting in the release of inflammatory lymphokines or cytokines which promote injury and fibrosis [27]. In addition, there are studies that have demonstrated that *CCR5* and *CXCR4*, the two main HIV co-receptors, that mediate entry of HIV strains into susceptible cells, are not expressed by intrinsic renal cells, but are expressed in circulating and infiltrating leukocytes at sites of tubulo-interstitial inflammation[28].

## GENETIC PREDISPOSITION TO CHRONIC KIDNEY DISEASE IN PATIENTS OF AFRICAN ANCESTRY

Genetic variation plays an important role in susceptibility to common forms of disease such as diabetes, hypertension and kidney disease, with marked differences in the prevalence and sometimes the presentation, according to ethnicity and ancestry. African Americans have four times the incidence of ESRD when compared to Americans of European ancestry, supporting a causal role for genetics in the aetiology of kidney disease[12,13,29]. These observations led to the use of ancestry informative population variation data to help explain this disparity. In 2008, two groups published papers back to back in Nature Genetics, heralding the discovery of genetic association of markers in the non-muscle myosin heavy chain 9 (*MYH9*) gene on chromosome 22 with non-diabetic ESRD[30] and FSGS[31] in African Americans (Figure 1 and Table 2). Both groups used genome wide admixture mapping approaches in their analysis, showing that increased African ancestry was correlated with increased susceptibility.

The transatlantic slave trade in the 16$^{th}$ to 19$^{th}$ centuries brought in an estimated 12 million individuals from Africa (mainly West Africa) to enslavement in

America[32,33] and this was the driver for the introduction of African genetic variation to America. As a consequence of this population relocation, admixture occurred with Native Americans (Amerindians) and Europeans leading to mixed genomic profiles among the group now referred to as African Americans. African Americans have, on average, about 80% African ancestry, although there are regional differences across the country[34]. Differences in allele frequencies of common and rare variants have occurred as a result of random genetic drift, selection and other forces over thousands of years of separation of the ancestral populations, with Europeans having separated roughly 40000 years ago from African populations[35]. Computational approaches take advantage of ancestry informative markers (AIMs), which are Single nucleotide polymorphisms (SNPs) that show marked allele frequency differences among the ancestral populations to infer the global ancestry of individuals. Studies utilizing AIMs have shown that American populations with African, Hispanic and Caribbean origins are admixed with varying substantial components of African continental ancestry[36]. This effect of admixture helped in identifying genetic regions that affect one ancestral population and not others, which drive phenotypic associations[37] and can be measured using mapping by admixture linkage disequilibrium (MALD), which quantifies the degree of ancestry of each locus[37-39]. MALD studies were used to identify genomic regions where admixed African American patients with CKD had an excess of African genomic markers compared to unaffected individual controls[30,31]. These studies identified CKD susceptibility loci in African Americans localised to a specific genomic region on chromosome 22q12 that contains more than 21 genes, and proceeded to pinpoint the association with non-diabetic and non-hypertensive CKD, but not type 2 diabetic ESRD, to markers in non-muscle myosin heavy chain 9 (*MYH9*) gene.

The *MYH9* gene was an excellent biologically plausible candidate as it has a direct link to the structure of podocytes since it codes for a 1960 amino acid protein (Myosin IIA) expressed in the podocytes and widely-distributed cellular motor protein that is essential for cytoskeleton rearrangement, cell motility, division, and

cell-cell adhesion[40]. As a result of these observations, researchers concluded that *MYH9* variants that are associated with susceptibility to CKD likely play a causal role in disease pathology[41]. The high frequency of the *MYH9* associated haplotypes in African populations led to speculations of selection in Africa (Figure 2)[42]. The ethnic specificity of the association was explored with different phenotypes and different populations (Table 2) and subsequent studies provided evidence for a contribution of *MYH9* variants in early stages of CKD as well as diabetic and hypertensive-related CKD in both African Americans and Europeans, but not in Native Americans[43-45]. However, biopsy-proven forms of CKD were lacking in some of these studies and therefore, as in the case of diabetic and hypertensive nephropathies, researchers suggested that the association with *MYH9* could also reflect the presence of non-diabetic and non-hypertensive CKD. This hypothesis was supported by association with *MYH9* SNPs that were strongly associated with these types of CKD[46].

The *MYH9* SNPs with the strongest associations were categorised into three haplotype groups termed as E, S and F[42,46] and E1 was defined as the risk haplotype (rs4821480, rs2032487, rs4821481 and rs3752462), being highly associated with CKD in individuals of African descent. E1 was then further found to be highly distributed in Africa as compared to other regions of the world (Figure 2) [42]. The *MYH9* E1 haplotype explained nearly the entire excess burden of major forms of CKD in African Americans with attributable risks of 100% and 70% for HIVAN and FSGS, respectively, and a significant percentage for hypertensive nephrosclerosis risk[31,40,47]. However, despite major scientific efforts, including *MYH9* re-sequencing experiments and detailed-intense genotyping, no mutations with a clear predicted functional effect could be identified that would impact kidney function and the field began to shift toward exploring neighbouring genes on chromosome 22.

Two research groups re-analyzed the chromosome 22q12 genomic region using data from International HapMap and 1000 Genomes Projects[29,48]. The data from these studies played a vital role in the discovery of candidate SNPs in the neighbouring apolipoprotein L1 (*APOL1*) gene, approximately 20kb downstream

from the 3′ end of *MYH9*, that were statistically powered to explain an increased risk of CKD in individuals of African ancestry[13,29]. The studies yielded 7479 SNPs, four of which were non-synonymous mutations in the coding region of the genes in high linkage disequilibrium with the *MYH9* E1 risk haplotype. Two of these (rs73885319 and rs60910145) were missense mutations in the last exon (exon 7) of the *APOL1* gene, which result in amino acid substitutions: Ser342Gly and Ile384Met. These two missense mutations are referred to as the G1 alleles since they were in almost complete linkage disequilibrium ($r^2$ = 1.0) with each other, and are both highly associated with CKD susceptibility. Another SNP, rs71785313, was also found in exon 7 of *APOL1* and represents a six base pair deletion resulting in loss of two amino acids (Asn388-Tyr389del), and is referred to as the G2 allele. These three codon-changing variants in the *APOL1* gene, encoding apolipoprotein L1, were found to be in strong association with HIVAN (odds ratio (OR = 29), FSGS (OR = 17) and ESRD (OR = 7) in African Americans, for homozygotes or compound heterozygotes carrying two risk alleles[12,13,29]. They were also absent in individuals of European ancestry but common in African populations. The G1 and G2 alleles were strongly associated with the risk for CKD than the previously examined *MYH9* E1 risk haplotype in a sample of individuals of African ancestry. They are in perfect negative linkage disequilibrium, never occurring on the same parental chromosome, suggesting that these variants arose independently and due to their proximity and high linkage disequilibrium, they have remained mutually exclusive in almost all haplotypes observed[12,13]. The historical timeline reflecting the discovery of genetic association to CKD in populations with African ancestry is shown in Figure 1. The link with increased susceptibility to kidney dysfunction in the presence of HIV infection[15,16] is now well established and the association of *APOL1* risk alleles with HIVAN is of particular concern in sub-Saharan Africa, where the risk allele frequency is high (Figure 2). *APOL1* association with CKD and the postulated mechanism of action and the functional role of *APOL1* in kidney disease are explored further in the next sections.

## *APOL1* ASSOCIATION WITH CKD IN INDIVIDUALS OF AFRICAN ANCESTRY

The coding variants (G1 and G2) are suggested to be causally related to CKD and provide an explanation for selection of *APOL1*-associated CKD risk polymorphisms as a protective measure against Trypanosomiasis, an infectious disease that was common in Africa. A study by Genovese et al., (2010), comparing 205 African Americans with biopsy-proven FSGS with 180 African Americans without kidney disease as controls was performed, using SNPs from the 1000 Genomes Project belonging to Yoruba as proxies to the African American population[13]. These SNPs revealed evidence of a strong association within a 10kb region in the exon 7 of the *APOL1* gene. The strong signal was found to be at non-synonymous coding variants, rs73885319 (S342G) and rs60910145 (I384M) which were in perfect linkage disequilibrium ($r^2$ = 1.0). The frequency of these variants was 52% in patients and 18% in controls. They further controlled for the effects of these two variants and found a second strong *APOL1* signal, 12 base pairs from I384M. This signal is a 6-base pair deletion represented by rs71785313 which removes two amino acid residues (Asparagine-N and Tyrosine-Y). The frequency of this variant was 23% in patients and 15% in controls. The odds ratio (OR) of association for carrying at least one risk (G1-G2) allele was 10.5 (95% confidence interval [CI] 6.0 to 18.4). Controlling for both G1 and G2 did not result in significant association with *MYH9*. However, controlling for *MYH9* variants maintained significant *APOL1* signal at G1 and G2. Kopp et al., (2011), in a larger FSGS and HIVAN cohort confirmed the association but this time, a greater association was observed in HIVAN (OR=29.2 CI 13.1 to 68.5 $P=6 \times 10^{-22}$) compared to FSGS (OR=16.9 CI 11.0 to 26.5 $P=1.3 \times 10^{-48}$)[12]. The authors reported that 50% of African Americans with 2 *APOL1* risk alleles would develop HIVAN if not on antiretroviral therapy. A study in an indigenous South African black cohort showed an independent high association with HIVAN susceptibility for these G1 and G2 variants of 89-fold CI 17.68 to 911.72 $P=1.2x10^{-14}$ (Kasembeli et al, unpublished observations). In all these studies, there was strong evidence for a contribution of *APOL1* variants to CKD (Figure 1 and Table 2).

The observed mode of inheritance of the *APOL1* risk variants was fully recessive in both FSGS and HIVAN cohorts. However, there have been instances where mild dominant effects (OR of 1.26 for one risk allele and OR of 7.3 for 2 risk alleles) have been observed in larger cohorts of hypertensive-associated CKD and geographically matched control subjects[13]. We therefore cannot fully exclude this mild dominant inheritance because this could be explained by yet "undiscovered variants" or by sporadic mutations that might occur in patients with recessive model. Furthermore, there could be a possibility of additional rare variants in *APOL1*, *MYH9* or other neighbouring genes that could be involved in CKD susceptibility since extended linkage disequilibrium exists in this region as a result of selective pressure[49].

## POSITIVE SELECTION FOR *APOL1*-ASSOCIATED CKD RISK VARIANTS AS A RESULT OF TRYPANOSOMIASIS EPIDEMIC IN AFRICA

It has been shown that harbouring of *APOL1* risk variants protects against Trypanosomiasis disease [Human African Trypanosomiasis (H.A.T)], otherwise known as sleeping sickness, that was epidemic in Africa many years ago and still affects millions of Africans today. This effect explains the high frequencies of these variants in the general African American and indigenous African population (Figure 2)[12,42]. The *APOL1* G1 and G2 alleles show distinct distributions among various African and African-derived populations and evidence shows these mutations to be maintained in these populations. In Yoruba from Nigeria in West Africa, the frequency is greater than 45 % for G1 (Figure 2) while in African Americans the G1 frequency is ~20%. The battle between host and pathogen (*Trypanosoma* species), resulted in development of *APOL1* mutations (G1 and G2) that provided positive selective advantage to carriers at the expense of increased risk for CKD (Figure 3).

There are three main *Trypanosoma* species; *Trypanosoma brucei brucei*, *Trypanosoma brucei rhodesiense* and *Trypanosoma brucei gambiense*. *Trypanosoma brucei brucei* is unable to infect humans because of the complex, trypanolytic factor (TLF) comprising of apolipoprotein L1, high density lipoprotein (HDL) particles, haptoglobin-related protein and apoliprotein A1 that is present in the human serum.

This confers innate protection against *Trypanosoma brucei brucei*. However, both *Trypanosoma brucei rhodesiense* and *Trypanosoma brucei gambiense* have evolved a mechanism to evade lysis by the TLF leading to infection, hence sleeping sickness (Figure 3)[50,51]. Apolipoprotein L1, a protein product of *APOL1* gene is usually part of TLF circulating in the blood. The *APOL1* gene is a member of the *APOL* gene family which is composed of six genes in humans (*APOL1*, *APOL2*, *APOL3*, *APOL4*, *APOL5* and *APOL6*), grouped within 619kb on chromosome 22[52]. This protein has five functional and structural domains: a secretory domain (SD), pore forming domain (PFD), B-cell lymphoma 2 homology domain 3 (BH3), membrane addressing domain (MAD), leucine zipper domain (LZ) and serum resistant-associated interacting domain (SRA), listed from the N-terminal to C-terminal respectively.

The trypanolytic function of apolipoprotein L1 is the most widely studied function of apolipoprotein L1[13,53]. The secretory domain allows it to be expressed as a circulating protein, which makes it the only circulating *APOL* protein[53-56]. *Trypanosoma brucei rhodesiense* and *Trypanosoma brucei gambiense* evade the TLF lysis by expressing SRA protein that binds to the C-terminus domain of the apolipoprotein L1, in the process neutralizing its lytic activity[57,58]. However, *APOL1* variants G1 and G2, powerfully associated with CKD arose to modify the C-terminus SRA binding site of the *APOL1* gene resulting in a mutated apolipoprotein L1 that evades neutralization by the *Trypanosoma* SRA protein[58]. By so doing, apolipoprotein L1 exercised its trypanolytic activity, conferring an adaptive advantage in the endemic regions of Africa (Figure 3). This explains the distribution of G1 and G2 risk variants in Africa.

Genovese *et al*[13] reported that both G1 and G2 variants restored the lytic activity of human serum and this provides the selective advantage to carriers of two *APOL1* risk variants against sleeping sickness. These findings corroborated the evidence of the recent evolution of *APOL1* which occurred in the last 10,000 years and also suggesting that these variants were selected for within Africa because they conferred protection against lethal trypanosomiasis while at the same time increasing

susceptibility to CKD (Figure 3)[49]. A more recent study in a South African black population (Kasembeli et al., unpublished data, 2014) has found the odds to have almost doubled. The prevalence of HIVAN in Africa is variable, 24%-83% in South Africa, while in the United States, it is highest in the African American population (15.5%). This is eight-fold greater than that of HIV-infected European Americans[59]. This high prevalence of HIVAN among individuals of African ancestry could be, not only as a result of high frequencies of *APOL1* risk variants, but also the prevalence of HIV-1 subtypes circulating in Africa. For instance, HIV-1 subtype C is highly virulent, accounting for ~50% of all HIV infections worldwide and 98% of HIV infections in South, West and East Africa, with corresponding higher viral loads[60,61]. Another more important reason could be because sub-Saharan African countries are resource limited, and therefore roll-out of antiretroviral therapy (ART) may have been delayed, giving more room for the development of HIVAN among individuals carrying two *APOL1* risk variants. An effective roll-out of ART has been shown to reduce the occurrence of HIVAN[22,62,63]; Figure 1. Therefore, there is need for HIV screening, surveillance, and strict implementation of World Health Organization (WHO) recommendations for ART initiation to reduce the burden of HIVAN and other forms of HIV-related CKD in Africa. At present, WHO ART guidelines 2013 for the treatment of HIV infection in Africa suggest that ART be instituted for individuals with WHO clinical stage 3 and 4 disease and in all HIV positive individuals with CD4 counts < 350 cells/μL and viral load of > 10000 copies/mL.

## GENE-ENVIRONMENTAL MODIFIERS OF *APOL1* SUSCEPTIBILITY

*APOL1*-environmental interactions play a vital role in CKD susceptibility in individuals of African ancestry (Figure 3). These could be social demographic status, lifestyle or presence of other communicable diseases and most importantly, HIV infection coupled with poor ART roll-out. Environmental exposure to HIV was initially thought to trigger HIVAN. However, observations in African American family studies showed that relatives of HIVAN patients, in the absence of HIV infection, suffered ESRD due to other aetiologies[64]. This illustrates that there could

be other environmental factors that drive the process. As described, patients with HIVAN harbouring two *APOL1* risk alleles that are untreated or undertreated for HIV infection will suffer rapid progression to ESRD[12]. In striking contrast, HIV patients without the *APOL1* risk genotype are protected from HIVAN even when HIV infection is not properly controlled. This effect is well illustrated in the Ethiopian population who appear to be protected from HIVAN since their genomes lack the *APOL1* risk variants. A study by Behar *et al*[65], in HIV infected individuals of Ethiopian origin reported complete absence of HIVAN. This led to the emphasis of skewed ethnic distribution, inter-individual variability and/or familial aggregation of HIVAN suggesting that host genetic susceptibility plays a major contributing factor. This study genotyped 676 African individuals from 12 populations, including 304 Ethiopians, for mutations in the *MYH9* and *APOL1* risk clusters. The frequency of the G1 and G2 *APOL1* risk variants was zero [65]. However, there was an increasing trend in frequency of the risk variants proceeding towards the west and south in Africa (Figure2). This led researchers to conclude that the risk of developing HIVAN is not a wide African problem but rather restricted to Western, Central and Southern Africa, and absent in regions of the North and North-East parts of Africa including Ethiopia. Since sleeping sickness was not an epidemic in Southern Africa, a possible explanation for the increase in prevalence of *APOL1* risk variants could be the result of migration of the bantu-speaking populations from West Africa and East Africa[66,67].

In the United States there has been a steady decline in the incidence of HIVAN with the introduction of HAART, in spite of stable frequencies of the risk variants[68]. Risk factors for progression to ESRD in HIVAN are severity of renal dysfunction, percentage of sclerotic glomeruli[25,69], lack of viral suppression[26,70], 2 *APOL1* risk alleles[63,71], while use of renin angiotensin system blockers were reported to be protective[25]. HIV-infected individuals with non-HIVAN pathology and two *APOL1* risk alleles had an almost 3-fold risk of ESRD, in spite of effective ART-suppression of viral load and use of renin-angiotensin aldosterone blockers; baseline kidney function was the strongest predictor of progression to ESRD in this study[71].

Investigators reviewing the African American study on Kidney Disease and Hypertension (AASK) and The Chronic Renal insufficiency Cohort (CRIC) found that *APOL1* risk variants in black patients were associated with higher rates of ESRD and progression of CKD[72].

Thus HIV is considered a risk factor for HIVAN when presented with the appropriate genetic susceptibility. Either genetic risk or viral infections alone do not cause the kidney disease. Instead it is the gene-environment interaction that is fundamental for the pathogenesis. The rapidly changing natural history of *APOL1*-associated HIVAN provides further support that HIV is an environmental risk factor[68]. Additional viral environmental modifiers have been proposed. The John Cunningham (JC) polyoma virus has been shown to maintain a reservoir in the uroepithelium of the kidney after infection and has been proposed to interact with the genetic risk posed by *APOL1* variants[73]. Divers *et al*[73] studied the relationship between the JC virus and genetic risk for kidney disease hypothesising that the presence of the *APOL1* risk variants may predispose individuals to JC infection and that this second hit may act as an additional environmental factor increasing kidney disease risk. However, paradoxically the reverse scenario was observed where the presence of the high risk *APOL1* variants in the presence of JC virus resulted in less kidney disease. The JC virus in the kidney was postulated to either protect against other nephropathic viruses or alter cellular function as protection against other sources of glomerular injury.

## GENE-GENE INTERACTIONS AS MODIFIERS OF *APOL1*-ASSOCIATED NEPHROPATHY

Whilst nephrology research in African ancestry populations has been hampered by the lack of large genome wide- association studies, a number of gene-gene interaction studies have been conducted on pooled GWAS data in non-diabetic ESRD in African Americans and non-nephropathy controls. Results of a gene-gene interaction analysis identified several SNPs that interacted with *APOL1* risk variants (Figure 3)[74]. *MYH9* has been shown to be one of the genes linked to *APOL1* to cause

CKD susceptibility. Other studies have also shown a possibility of other genes interaction with *APOL1* gene. In a replication study, eleven SNPs were validated and three genes, podocin (NPHS2; rs16854341); serologically defined colon cancer antigen 8 (SDCCAG8; rs2802723) and SNP near the bone morphogenetic protein 4 (BMP4; rs8014363) were significant. These interactions were quantified and all show effects on *APOL1* association[75,76]. These three genes show expression in podocytes and are linked to renal disease characterised by FSGS. It has thus been postulated that they play a role in inducing podocyturia and glomerular damage.

## *APOL1*-ASSOCIATED CKD: THE FUTURE

Since there is evidence of *APOL1* association with non-diabetic forms of CKD and the role of selection in the increase in frequencies of the risk variants, it is necessary to move beyond the statistical tests of association to molecular cellular characterization to evaluate the effects of these risk variants in CKD. It has been postulated that cellular and physiologic activities of apolipoprotein L1 include involvement in autophagic and apoptosis pathways[52,77-79]. There is a general agreement that *APOL1* expression occurs in podocytes[80]. But whether there is apolipoprotein L1 expression in the tubular cells, glomerular endothelial cells, and the tunica intima and media of the renal blood vessels is uncertain. Currently, there is no definitive mechanism by which *APOL1* variants cause kidney injury but several possibilities have been proposed. Firstly, apolipoprotein L1 isoform expressed in the kidney cells may be retained in the cells and cause cell destruction via the apoptotic pathway since they share structural and functional similarities with proteins from the BcL2 family[77,78]. Secondly, *APOL1* as part of TLF, is directed to lysosomes to induce programmed cell death via the autophagic response[13,56,81]. Thirdly, circulating apolipoprotein L1 may also be important in the pathogenesis of CKD since the presence of G1 and G2 could lead to dysfunctional HDL particles leading to inflammation of vascular endothelial cells, with arteriolar nephrosclerosis[49]. There is a proven race specific relationship between *APOL1* genotype and HDL cholesterol concentration and kidney function[82]. In Han Chinese and European American

populations, there was a higher HDL level in association with higher eGFR. The inverse association was observed in West Africans and African Americans. However, a significant effect was observed only in African Americans with *APOL1* risk variants (but not in West Africans). These observations have led to the view that the mechanism underlying *APOL1* nephropathy most likely involves HDL cholesterol. In addition, circulating *APOL1*, may be available for uptake by podocytes after passage across the glomerular filtration barrier and exercising their effect on the podocytes. Future studies should define the role of *APOL1* in the pathogenesis of kidney disease.

## LESSONS LEARNT IN POPULATION GENETICS OF CKD

The remarkable advances in molecular genetics have enabled researchers to unravel the underlying genetic susceptibility to kidney disease in African Ancestry populations. Identification of region of 22q12.3 using MALD studies and identification of *APOL1* risk variants have raised the possibility of a personalised approach to treat several forms of kidney disease that are prevalent in African populations. As regards HIVAN, the priority for the African continent should be targeting of the modifying trigger of the disease, HIV infection, through effective treatment and prevention campaigns. Greater advances in understanding the mechanisms underlying *APOL1* pathogenesis, the identification of modifiable environmental factors and interacting genes offer the promise of novel preventive, prognostic and therapeutic measures to treat *APOL1* associated forms of kidney disease in the genetically susceptible and therefore vulnerable African descent individual.

## REFERENCES

1 **United Nations.** General Assembly political declaration of the high-level meeting of the general assembly on the prevention and control of non-communicable diseases. Cited 2014-06-13. Available from: URL: http: //www.who.int/nmh/events/un_ncd_summit2011/political_declaration_en.pdf

2 **Lozano R,** Naghavi M, Foreman K, Lim S, Shibuya K, Aboyans V, Abraham J, Adair T, Aggarwal R, Ahn SY, Alvarado M, Anderson HR, Anderson LM, Andrews KG, Atkinson C, Baddour LM, Barker-Collo S, Bartels DH, Bell ML, Benjamin EJ, Bennett D, Bhalla K, Bikbov B, Bin Abdulhak A, Birbeck G, Blyth F, Bolliger I, Boufous S, Bucello C, Burch M, Burney P, Carapetis J, Chen H, Chou D, Chugh SS, Coffeng LE, Colan SD, Colquhoun S, Colson KE, Condon J, Connor MD, Cooper LT, Corriere M, Cortinovis M, de Vaccaro KC, Couser W, Cowie BC, Criqui MH, Cross M, Dabhadkar KC, Dahodwala N, De Leo D, Degenhardt L, Delossantos A, Denenberg J, Des Jarlais DC, Dharmaratne SD, Dorsey ER, Driscoll T, Duber H, Ebel B, Erwin PJ, Espindola P, Ezzati M, Feigin V, Flaxman AD, Forouzanfar MH, Fowkes FG, Franklin R, Fransen M, Freeman MK, Gabriel SE, Gakidou E, Gaspari F, Gillum RF, Gonzalez-Medina D, Halasa YA, Haring D, Harrison JE, Havmoeller R, Hay RJ, Hoen B, Hotez PJ, Hoy D, Jacobsen KH, James SL, Jasrasaria R, Jayaraman S, Johns N, Karthikeyan G, Kassebaum N, Keren A, Khoo JP, Knowlton LM, Kobusingye O, Koranteng A, Krishnamurthi R, Lipnick M, Lipshultz SE, Ohno SL, Mabweijano J, MacIntyre MF, Mallinger L, March L, Marks GB, Marks R, Matsumori A, Matzopoulos R, Mayosi BM, McAnulty JH, McDermott MM, McGrath J, Mensah GA, Merriman TR, Michaud C, Miller M, Miller TR, Mock C, Mocumbi AO, Mokdad AA, Moran A, Mulholland K, Nair MN, Naldi L, Narayan KM, Nasseri K, Norman P, O'Donnell M, Omer SB, Ortblad K, Osborne R, Ozgediz D, Pahari B, Pandian JD, Rivero AP, Padilla RP, Perez-Ruiz F, Perico N, Phillips D, Pierce K, Pope CA 3rd, Porrini E, Pourmalek F, Raju M, Ranganathan D, Rehm JT, Rein DB, Remuzzi G, Rivara FP, Roberts T, De León FR, Rosenfeld LC, Rushton L, Sacco RL, Salomon JA, Sampson U, Sanman E, Schwebel DC, Segui-Gomez M, Shepard DS, Singh D, Singleton J, Sliwa K, Smith E, Steer A, Taylor JA, Thomas B, Tleyjeh IM, Towbin JA,

Truelsen T, Undurraga EA, Venketasubramanian N, Vijayakumar L, Vos T, Wagner GR, Wang M, Wang W, Watt K, Weinstock MA, Weintraub R, Wilkinson JD, Woolf AD, Wulf S, Yeh PH, Yip P, Zabetian A, Zheng ZJ, Lopez AD, Murray CJ, AlMazroa MA, Memish ZA. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. *Lancet* 2012; **380**: 2095-2128 [PMID: 23245604 DOI: 10.1016/S0140-6736(12)61728-0]

3 **Mayosi BM**, Flisher AJ, Lalloo UG, Sitas F, Tollman SM, Bradshaw D. The burden of non-communicable diseases in South Africa. *Lancet* 2009; **374**: 934-947 [PMID: 19709736 DOI: 10.1016/S0140-6736(09)61087-4]

4 ESRD Patients in 2013: A Global Perspective. Cited 2014-07-07. Available from: URL: http: //www.vision-fmc.com/files/ESRD_Patients_in_2013.pdf

5 **Rostand SG**, Kirk KA, Rutsky EA, Pate BA. Racial differences in the incidence of treatment for end-stage renal disease. *N Engl J Med* 1982; **306**: 1276-1279 [PMID: 7040967 DOI: 10.1056/NEJM198205273062106]

6 **Shlush LI**, Bercovici S, Wasser WG, Yudkovsky G, Templeton A, Geiger D, Skorecki K. Admixture mapping of end stage kidney disease genetic susceptibility using estimated mutual information ancestry informative markers. *BMC Med Genomics* 2010; **3**: 47 [PMID: 20955568 DOI: 10.1186/1755-8794-3-47]

7 **Klag MJ**, Whelton PK, Randall BL, Neaton JD, Brancati FL, Stamler J. End-stage renal disease in African-American and white men. 16-year MRFIT findings. *JAMA* 1997; **277**: 1293-1298 [PMID: 9109467 DOI: 10.1001/jama.1997.03540400043029]

8 **Sayre AP.** Africa: The Seven Continent. Cited 2014-05-20. Available from: URL: http: //www.amazon.com/Africa-Seven-Continents-April-Pulley/dp/0761313672

9 World Population Data Sheet. Cited 2014-06-13. Available from: URL: http: //www.prb.org/Publications/Datasheets/2013/2013-world-population-data-sheet/world-map.aspx#map/world/population/2013

10 **Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat.** World population prospects: the 2007 revision and

world urbanization prospects. Cited 2014-06-13. Available from: URL: http://www.un.org/en/development/desa/index.html

11 Idem. Epidemiology of ESRD: a world-wide perspective. El Nahas M, editor. Kidney diseases in the developing world and ethnic minorities. London, UK: Taylor & Francis, 2005: 1-13 [DOI: 10.1201/b14128]

12 **Kopp JB**, Nelson GW, Sampath K, Johnson RC, Genovese G, An P, Friedman D, Briggs W, Dart R, Korbet S, Mokrzycki MH, Kimmel PL, Limou S, Ahuja TS, Berns JS, Fryc J, Simon EE, Smith MC, Trachtman H, Michel DM, Schelling JR, Vlahov D, Pollak M, Winkler CA. APOL1 genetic variants in focal segmental glomerulosclerosis and HIV-associated nephropathy. *J Am Soc Nephrol* 2011; **22**: 2129-2137 [PMID: 21997394 DOI: 10.1681/ASN.2011040388]

13 **Genovese G**, Friedman DJ, Ross MD, Lecordier L, Uzureau P, Freedman BI, Bowden DW, Langefeld CD, Oleksyk TK, Uscinski Knob AL, Bernhardy AJ, Hicks PJ, Nelson GW, Vanhollebeke B, Winkler CA, Kopp JB, Pays E, Pollak MR. Association of trypanolytic ApoL1 variants with kidney disease in African Americans. *Science* 2010; **329**: 841-845 [PMID: 20647424 DOI: 10.1126/science.1193032]

14 **Rao TK**, Filippone EJ, Nicastri AD, Landesman SH, Frank E, Chen CK, Friedman EA. Associated focal and segmental glomerulosclerosis in the acquired immunodeficiency syndrome. *N Engl J Med* 1984; **310**: 669-673 [PMID: 6700641 DOI: 10.1056/NEJM198403153101101]

15 **Estrella MM,** Fine DM, Atta MG. Recent developments in HIV-related kidney disease. *HIV Ther* 2010; **4**: 589-603 [PMID: 21331321 DOI: 10.2217/hiv.10.42]

16 **Estrella MM**, Parekh RS, Abraham A, Astor BC, Szczech LA, Anastos K, Dehovitz JA, Merenstein DJ, Pearce CL, Tien PC, Cohen MH, Gange SJ. The impact of kidney function at highly active antiretroviral therapy initiation on mortality in HIV-infected women. *J Acquir Immune Defic Syndr* 2010; **55**: 217-220 [PMID: 20581688 DOI: 10.1097/QAI.0b013e3181e674f4]

17 **Nation Kindey Foundations.** 2013 USRDS Annual Data Report: Atlas of Chronic Kidney Disease and End-Stage Renal Disease in the United States. American Journel of Kidney diesease 2014; 63 Suppl: e1-e478

18 UNAIDS report on the global AIDS epidemic 2013. Cited 2014-05-20. Available from: URL: http: //www.unaids.org/en/media/unaids/contentassets/documents/epidemiology/2013/gr2013/UNAIDS_Global_Report_2013_en.pdf

19 **Fabian J**, Naicker S. HIV and kidney disease in sub-Saharan Africa. *Nat Rev Nephrol* 2009; **5**: 591-598 [PMID: 19776781 DOI: 10.1038/nrneph.2009.141]

20 **Gerntholtz TE**, Goetsch SJ, Katz I. HIV-related nephropathy: a South African perspective. *Kidney Int* 2006; **69**: 1885-1891 [PMID: 16625149 DOI: 10.1038/sj.ki.5000351]

21 **Han TM**, Naicker S, Ramdial PK, Assounga AG. A cross-sectional study of HIV-seropositive patients with varying degrees of proteinuria in South Africa. *Kidney Int* 2006; **69**: 2243-2250 [PMID: 16672914 DOI: 10.1038/sj.ki.5000339]

22 **Wearne N**, Swanepoel CR, Boulle A, Duffield MS, Rayner BL. The spectrum of renal histologies seen in HIV with outcomes, prognostic indicators and clinical correlations. *Nephrol Dial Transplant* 2012; **27**: 4109-4118 [PMID: 22200584 DOI: 10.1093/ndt/gfr702]

23 **D'Agati V**, Appel GB. HIV infection and the kidney. *J Am Soc Nephrol* 1997; **8**: 138-152 [PMID: 9013459]

24 **Ray PE**. HIV-associated nephropathy: a diagnosis in evolution. *Nephrol Dial Transplant* 2012; **27**: 3969-3972 [PMID: 22584790 DOI: 10.1093/ndt/gfs114]

25 **Bigé N**, Lanternier F, Viard JP, Kamgang P, Daugas E, Elie C, Jidar K, Walker-Combrouze F, Peraldi MN, Isnard-Bagnis C, Servais A, Lortholary O, Noël LH, Bollée G. Presentation of HIV-associated nephropathy and outcome in HAART-treated patients. *Nephrol Dial Transplant* 2012; **27**: 1114-1121 [PMID: 21745806 DOI: 10.1093/ndt/gfr376]

26 **Szczech LA**, Gupta SK, Habash R, Guasch A, Kalayjian R, Appel R, Fields TA, Svetkey LP, Flanagan KH, Klotman PE, Winston JA. The clinical epidemiology and

course of the spectrum of renal diseases associated with HIV infection. *Kidney Int* 2004; **66**: 1145-1152 [PMID: 15327410 DOI: 10.1111/j.1523-1755.2004.00865.x]

27 **Ray PE**, Liu XH, Henry D, Dye L, Xu L, Orenstein JM, Schuztbank TE. Infection of human primary renal epithelial cells with HIV-1 from children with HIV-associated nephropathy. *Kidney Int* 1998; **53**: 1217-1229 [PMID: 9573536 DOI: 10.1046/j.1523-1755.1998.00900.x]

28 **Eitner F**, Cui Y, Hudkins KL, Stokes MB, Segerer S, Mack M, Lewis PL, Abraham AA, Schlöndorff D, Gallo G, Kimmel PL, Alpers CE. Chemokine receptor CCR5 and CXCR4 expression in HIV-associated kidney disease. *J Am Soc Nephrol* 2000; **11**: 856-867 [PMID: 10770963]

29 **Tzur S**, Rosset S, Shemer R, Yudkovsky G, Selig S, Tarekegn A, Bekele E, Bradman N, Wasser WG, Behar DM, Skorecki K. Missense mutations in the APOL1 gene are highly associated with end stage kidney disease risk previously attributed to the MYH9 gene. *Hum Genet* 2010; **128**: 345-350 [PMID: 20635188 DOI: 10.1007/s00439-010-0861-0]

30 **Kao WH**, Klag MJ, Meoni LA, Reich D, Berthier-Schaad Y, Li M, Coresh J, Patterson N, Tandon A, Powe NR, Fink NE, Sadler JH, Weir MR, Abboud HE, Adler SG, Divers J, Iyengar SK, Freedman BI, Kimmel PL, Knowler WC, Kohn OF, Kramp K, Leehey DJ, Nicholas SB, Pahl MV, Schelling JR, Sedor JR, Thornley-Brown D, Winkler CA, Smith MW, Parekh RS. MYH9 is associated with nondiabetic end-stage renal disease in African Americans. *Nat Genet* 2008; **40**: 1185-1192 [PMID: 18794854 DOI: 10.1038/ng.232]

31 **Kopp JB**, Smith MW, Nelson GW, Johnson RC, Freedman BI, Bowden DW, Oleksyk T, McKenzie LM, Kajiyama H, Ahuja TS, Berns JS, Briggs W, Cho ME, Dart RA, Kimmel PL, Korbet SM, Michel DM, Mokrzycki MH, Schelling JR, Simon E, Trachtman H, Vlahov D, Winkler CA. MYH9 is a major-effect risk gene for focal segmental glomerulosclerosis. *Nat Genet* 2008; **40**: 1175-1184 [PMID: 18794856 DOI: 10.1038/ng.226]

32 **Salas A**, Carracedo A, Richards M, Macaulay V. Charting the ancestry of African Americans. *Am J Hum Genet* 2005; **77**: 676-680 [PMID: 16175514 DOI: 10.1086/491675]

33 **Tishkoff SA**, Reed FA, Friedlaender FR, Ehret C, Ranciaro A, Froment A, Hirbo JB, Awomoyi AA, Bodo JM, Doumbo O, Ibrahim M, Juma AT, Kotze MJ, Lema G, Moore JH, Mortensen H, Nyambo TB, Omar SA, Powell K, Pretorius GS, Smith MW, Thera MA, Wambebe C, Weber JL, Williams SM. The genetic structure and history of Africans and African Americans. *Science* 2009; **324**: 1035-1044 [PMID: 19407144 DOI: 10.1126/science.1172257]

34 **Zakharia F**, Basu A, Absher D, Assimes TL, Go AS, Hlatky MA, Iribarren C, Knowles JW, Li J, Narasimhan B, Sidney S, Southwick A, Myers RM, Quertermous T, Risch N, Tang H. Characterizing the admixed African ancestry of African Americans. *Genome Biol* 2009; **10**: R141 [PMID: 20025784 DOI: 10.1186/gb-2009-10-12-r141]

35 **Macaulay V**, Hill C, Achilli A, Rengo C, Clarke D, Meehan W, Blackburn J, Semino O, Scozzari R, Cruciani F, Taha A, Shaari NK, Raja JM, Ismail P, Zainuddin Z, Goodwin W, Bulbeck D, Bandelt HJ, Oppenheimer S, Torroni A, Richards M. Single, rapid coastal settlement of Asia revealed by analysis of complete mitochondrial genomes. *Science* 2005; **308**: 1034-1036 [PMID: 15890885 DOI: 10.1126/science.1109792]

36 **Tang H**, Peng J, Wang P, Risch NJ. Estimation of individual admixture: analytical and study design considerations. *Genet Epidemiol* 2005; **28**: 289-301 [PMID: 15712363 DOI: 10.1002/gepi.20064]

37 **Tian C**, Gregersen PK, Seldin MF. Accounting for ancestry: population substructure and genome-wide association studies. *Hum Mol Genet* 2008; **17**: R143-R150 [PMID: 18852203 DOI: 10.1093/hmg/ddn268]

38 **Hoggart CJ**, Shriver MD, Kittles RA, Clayton DG, McKeigue PM. Design and analysis of admixture mapping studies. *Am J Hum Genet* 2004; **74**: 965-978 [PMID: 15088268 DOI: 10.1086/420855]

39 **Patterson N**, Hattangadi N, Lane B, Lohmueller KE, Hafler DA, Oksenberg JR, Hauser SL, Smith MW, O'Brien SJ, Altshuler D, Daly MJ, Reich D. Methods for high-density admixture mapping of disease genes. *Am J Hum Genet* 2004; **74**: 979-1000 [PMID: 15088269 DOI: 10.1086/420871]

40 **Winkler CA**, Nelson G, Oleksyk TK, Nava MB, Kopp JB. Genetics of focal segmental glomerulosclerosis and human immunodeficiency virus-associated collapsing glomerulopathy: the role of MYH9 genetic variation. *Semin Nephrol* 2010; **30**: 111-125 [PMID: 20347641 DOI: 10.1016/j.semnephrol.2010.01.003]

41 **Kopp JB**. Glomerular pathology in autosomal dominant MYH9 spectrum disorders: what are the clues telling us about disease mechanism? *Kidney Int* 2010; **78**: 130-133 [PMID: 20588287 DOI: 10.1038/ki.2010.82]

42 **Oleksyk TK**, Nelson GW, An P, Kopp JB, Winkler CA. Worldwide distribution of the MYH9 kidney disease susceptibility alleles and haplotypes: evidence of historical selection in Africa. *PLoS One* 2010; **5**: e11474 [PMID: 20634883 DOI: 10.1371/journal.pone.0011474]

43 **Franceschini N**, Voruganti VS, Haack K, Almasy L, Laston S, Goring HH, Umans JG, Lee ET, Best LG, Fabsitz RR, MacCluer JW, Howard BV, North KE, Cole SA. The association of the MYH9 gene and kidney outcomes in American Indians: the Strong Heart Family Study. *Hum Genet* 2010; **127**: 295-301 [PMID: 19921264 DOI: 10.1007/s00439-009-0769-8]

44 **Freedman BI**, Hicks PJ, Bostrom MA, Comeau ME, Divers J, Bleyer AJ, Kopp JB, Winkler CA, Nelson GW, Langefeld CD, Bowden DW. Non-muscle myosin heavy chain 9 gene MYH9 associations in African Americans with clinically diagnosed type 2 diabetes mellitus-associated ESRD. *Nephrol Dial Transplant* 2009; **24**: 3366-3371 [PMID: 19567477 DOI: 10.1093/ndt/gfp316]

45 **Freedman BI**, Kopp JB, Winkler CA, Nelson GW, Rao DC, Eckfeldt JH, Leppert MF, Hicks PJ, Divers J, Langefeld CD, Hunt SC. Polymorphisms in the nonmuscle myosin heavy chain 9 gene (MYH9) are associated with albuminuria in hypertensive African Americans: the HyperGEN study. *Am J Nephrol* 2009; **29**: 626-632 [PMID: 19153477 DOI: 10.1159/000194791]

46 **Behar DM**, Rosset S, Tzur S, Selig S, Yudkovsky G, Bercovici S, Kopp JB, Winkler CA, Nelson GW, Wasser WG, Skorecki K. African ancestry allelic variation at the MYH9 gene contributes to increased susceptibility to non-diabetic end-stage kidney

disease in Hispanic Americans. *Hum Mol Genet* 2010; **19**: 1816-1827 [PMID: 20144966 DOI: 10.1093/hmg/ddq040]

47 **Freedman BI**, Hicks PJ, Bostrom MA, Cunningham ME, Liu Y, Divers J, Kopp JB, Winkler CA, Nelson GW, Langefeld CD, Bowden DW. Polymorphisms in the non-muscle myosin heavy chain 9 gene (MYH9) are strongly associated with end-stage renal disease historically attributed to hypertension in African Americans. *Kidney Int* 2009; **75**: 736-745 [PMID: 19177153 DOI: 10.1038/ki.2008.701]

48 **Abecasis GR**, Altshuler D, Auton A, Brooks LD, Durbin RM, Gibbs RA, Hurles ME, McVean GA. A map of human genome variation from population-scale sequencing. *Nature* 2010; **467**: 1061-1073 [PMID: 20981092 DOI: 10.1038/nature09534]

49 **Freedman BI**, Kopp JB, Langefeld CD, Genovese G, Friedman DJ, Nelson GW, Winkler CA, Bowden DW, Pollak MR. The apolipoprotein L1 (APOL1) gene and nondiabetic nephropathy in African Americans. *J Am Soc Nephrol* 2010; **21**: 1422-1426 [PMID: 20688934 DOI: 10.1681/ASN.2010070730]

50 **Brun R**, Blum J. Human African trypanosomiasis. *Infect Dis Clin North Am* 2012; **26**: 261-273 [PMID: 22632638 DOI: 10.1016/j.idc.2012.03.003]

51 **Brun R**, Blum J, Chappuis F, Burri C. Human African trypanosomiasis. *Lancet* 2010; **375**: 148-159 [PMID: 19833383 DOI: 10.1016/S0140-6736(09)60829-1]

52 **Smith EE**, Malik HS. The apolipoprotein L family of programmed cell death and immunity genes rapidly evolved in primates at discrete sites of host-pathogen interactions. *Genome Res* 2009; **19**: 850-858 [PMID: 19299565 DOI: 10.1101/gr.085647.108]

53 **Vanhamme L**, Paturiaux-Hanocq F, Poelvoorde P, Nolan DP, Lins L, Van Den Abbeele J, Pays A, Tebabi P, Van Xong H, Jacquet A, Moguilevsky N, Dieu M, Kane JP, De Baetselier P, Brasseur R, Pays E. Apolipoprotein L-I is the trypanosome lytic factor of human serum. *Nature* 2003; **422**: 83-87 [PMID: 12621437 DOI: 10.1038/nature01461]

54 **Madhavan SM**, O'Toole JF. The biology of APOL1 with insights into the association between APOL1 variants and chronic kidney disease. *Clin Exp Nephrol* 2014; **18**: 238-242 [PMID: 24233469 DOI: 10.1007/s10157-013-0907-4]

55 **Vanhamme L**, Pays E. The trypanosome lytic factor of human serum and the molecular basis of sleeping sickness. *Int J Parasitol* 2004; **34**: 887-898 [PMID: 15217727 DOI: 10.1016/j.ijpara.2004.04.008]

56 **Vanhollebeke B**, Pays E. The trypanolytic factor of human serum: many ways to enter the parasite, a single way to kill. *Mol Microbiol* 2010; **76**: 806-814 [PMID: 20398209 DOI: 10.1111/j.1365-2958.2010.07156.x]

57 **De Greef C**, Chimfwembe E, Kihang'a Wabacha J, Bajyana Songa E, Hamers R. Only the serum-resistant bloodstream forms of Trypanosoma brucei rhodesiense express the serum resistance associated (SRA) protein. *Ann Soc Belg Med Trop* 1992; **72** Suppl 1: 13-21 [PMID: 1417165]

58 **Lecordier L**, Vanhollebeke B, Poelvoorde P, Tebabi P, Paturiaux-Hanocq F, Andris F, Lins L, Pays E. C-terminal mutants of apolipoprotein L-I efficiently kill both Trypanosoma brucei brucei and Trypanosoma brucei rhodesiense. *PLoS Pathog* 2009; **5**: e1000685 [PMID: 19997494 DOI: 10.1371/journal.ppat.1000685]

59 **Ikpeme EE**, Ekrikpo UE, Akpan MU, Ekaidem SI. Determining the prevalence of human immunodeficiency virus-associated nephropathy (HIVAN) using proteinuria and ultrasound findings in a Nigerian paediatric HIV population. *Pan Afr Med J* 2012; **11**: 13 [PMID: 22368756]

60 **Hemelaar J**, Gouws E, Ghys PD, Osmanov S. Global and regional distribution of HIV-1 genetic subtypes and recombinants in 2004. *AIDS* 2006; **20**: W13-W23 [PMID: 17053344 DOI: 10.1097/01.aids.0000247564.73009.bc]

61 **Novitsky V**, Ndung'u T, Wang R, Bussmann H, Chonco F, Makhema J, De Gruttola V, Walker BD, Essex M. Extended high viremics: a substantial fraction of individuals maintain high plasma viral RNA levels after acute HIV-1 subtype C infection. *AIDS* 2011; **25**: 1515-1522 [PMID: 21505307 DOI: 10.1097/QAD.0b013e3283471eb2]

62 **Hays T**, Wyatt CM. APOL1 variants in HIV-associated nephropathy: just one piece of the puzzle. *Kidney Int* 2012; **82**: 259-260 [PMID: 22791322 DOI: 10.1038/ki.2012.129]

63 **Kalayjian RC**. The treatment of HIV-associated nephropathy. *Adv Chronic Kidney Dis* 2010; **17**: 59-71 [PMID: 20005490 DOI: 10.1053/j.ackd.2009.08.013]

64 **Freedman BI**, Skorecki K. Gene-gene and gene-environment interactions in apolipoprotein L1 gene-associated nephropathy. *Clin J Am Soc Nephrol* 2014; **9**: 2006-2013 [PMID: 24903390 DOI: 10.2215/CJN.01330214]

65 **Behar DM**, Shlush LI, Maor C, Lorber M, Skorecki K. Absence of HIV-associated nephropathy in Ethiopians. *Am J Kidney Dis* 2006; **47**: 88-94 [PMID: 16377389 DOI: 10.1053/j.ajkd.2005.09.023]

66 **de Filippo C**, Barbieri C, Whitten M, Mpoloka SW, Gunnarsdóttir ED, Bostoen K, Nyambe T, Beyer K, Schreiber H, de Knijff P, Luiselli D, Stoneking M, Pakendorf B. Y-chromosomal variation in sub-Saharan Africa: insights into the history of Niger-Congo groups. *Mol Biol Evol* 2011; **28**: 1255-1269 [PMID: 21109585 DOI: 10.1093/molbev/msq312]

67 **Tishkoff SA**, Williams SM. Genetic analysis of African populations: human evolution and complex disease. *Nat Rev Genet* 2002; **3**: 611-621 [PMID: 12154384 DOI: 10.1038/nrg865]

68 USRSD 2012 Annual Data Report. Cited 2014-05-10. Available from: URL: http://www.usrds.org/atlas12.aspx

69 **Post FA**, Campbell LJ, Hamzah L, Collins L, Jones R, Siwani R, Johnson L, Fisher M, Holt SG, Bhagani S, Frankel AH, Wilkins E, Ainsworth JG, Larbalestier N, Macallan DC, Banerjee D, Baily G, Thuraisingham RC, Donohoe P, Hendry BM, Hilton RM, Edwards SG, Hangartner R, Howie AJ, Connolly JO, Easterbrook PJ. Predictors of renal outcome in HIV-associated nephropathy. *Clin Infect Dis* 2008; **46**: 1282-1289 [PMID: 18444868 DOI: 10.1086/529385]

70 **Lucas GM**, Eustace JA, Sozio S, Mentari EK, Appiah KA, Moore RD. Highly active antiretroviral therapy and the incidence of HIV-1-associated nephropathy: a 12-year cohort study. *AIDS* 2004; **18**: 541-546 [PMID: 15090808 DOI: 10.1097/00002030-200402200-00022]

71 **Fine DM**, Wasser WG, Estrella MM, Atta MG, Kuperman M, Shemer R, Rajasekaran A, Tzur S, Racusen LC, Skorecki K. APOL1 risk variants predict

histopathology and progression to ESRD in HIV-related kidney disease. *J Am Soc Nephrol* 2012; **23**: 343-350 [PMID: 22135313 DOI: 10.1681/ASN.2011060562]

72 **Parsa A**, Kao WH, Xie D, Astor BC, Li M, Hsu CY, Feldman HI, Parekh RS, Kusek JW, Greene TH, Fink JC, Anderson AH, Choi MJ, Wright JT, Lash JP, Freedman BI, Ojo A, Winkler CA, Raj DS, Kopp JB, He J, Jensvold NG, Tao K, Lipkowitz MS, Appel LJ. APOL1 risk variants, race, and progression of chronic kidney disease. *N Engl J Med* 2013; **369**: 2183-2196 [PMID: 24206458 DOI: 10.1056/NEJMoa1310345]

73 **Divers J**, Núñez M, High KP, Murea M, Rocco MV, Ma L, Bowden DW, Hicks PJ, Spainhour M, Ornelles DA, Kleiboeker SB, Duncan K, Langefeld CD, Turner J, Freedman BI. JC polyoma virus interacts with APOL1 in African Americans with nondiabetic nephropathy. *Kidney Int* 2013; **84**: 1207-1213 [PMID: 23677244 DOI: 10.1038/ki.2013.173]

74 **Bostrom MA**, Freedman BI. The spectrum of MYH9-associated nephropathy. *Clin J Am Soc Nephrol* 2010; **5**: 1107-1113 [PMID: 20299374 DOI: 10.2215/CJN.08721209]

75 **Divers J**, Freedman BI. Genetics in kidney disease in 2013: Susceptibility genes for renal and urological disorders. *Nat Rev Nephrol* 2014; **10**: 69-70 [PMID: 24296626 DOI: 10.1038/nrneph.2013.259]

76 **Divers J**, Palmer ND, Lu L, Langefeld CD, Rocco MV, Hicks PJ, Murea M, Ma L, Bowden DW, Freedman BI. Gene-gene interactions in APOL1-associated nephropathy. *Nephrol Dial Transplant* 2014; **29**: 587-594 [PMID: 24157943 DOI: 10.1093/ndt/gft423]

77 **Zhaorigetu S**, Wan G, Kaini R, Jiang Z, Hu CA. ApoL1, a BH3-only lipid-binding protein, induces autophagic cell death. *Autophagy* 2008; **4**: 1079-1082 [PMID: 18927493 DOI: 10.4161/auto.7066]

78 **Wan G**, Zhaorigetu S, Liu Z, Kaini R, Jiang Z, Hu CA. Apolipoprotein L1, a novel Bcl-2 homology domain 3-only lipid-binding protein, induces autophagic cell death. *J Biol Chem* 2008; **283**: 21540-21549 [PMID: 18505729 DOI: 10.1074/jbc.M800214200]

79 **Pérez-Morga D**, Vanhollebeke B, Paturiaux-Hanocq F, Nolan DP, Lins L, Homblé F, Vanhamme L, Tebabi P, Pays A, Poelvoorde P, Jacquet A, Brasseur R, Pays E. Apolipoprotein L-I promotes trypanosome lysis by forming pores in lysosomal

membranes. *Science* 2005; **309**: 469-472 [PMID: 16020735 DOI: 10.1126/science.1114566]

80 **Madhavan SM**, O'Toole JF, Konieczkowski M, Ganesan S, Bruggeman LA, Sedor JR. APOL1 localization in normal kidney and nondiabetic kidney disease. *J Am Soc Nephrol* 2011; **22**: 2119-2128 [PMID: 21997392 DOI: 10.1681/ASN.2011010069]

81 **Pays E**, Vanhollebeke B, Vanhamme L, Paturiaux-Hanocq F, Nolan DP, Pérez-Morga D. The trypanolytic factor of human serum. *Nat Rev Microbiol* 2006; **4**: 477-486 [PMID: 16710327 DOI: 10.1038/nrmicro1428]

82 **Bentley AR**, Doumatey AP, Chen G, Huang H, Zhou J, Shriner D, Jiang C, Zhang Z, Liu G, Fasanmade O, Johnson T, Oli J, Okafor G, Eghan BA, Agyenim-Boateng K, Adeleye J, Balogun W, Adebamowo C, Amoah A, Acheampong J, Adeyemo A, Rotimi CN. Variation in APOL1 Contributes to Ancestry-Level Differences in HDLc-Kidney Function Association. *Int J Nephrol* 2012; **2012**: 748984 [PMID: 22973513 DOI: 10.1155/2012/748984]

83 **Vermeulen A.** A Review of Patterns of Renal Disease at Chris Hani Baragwanath Academic Hospital from 1982 to 2011. University of the Witwatersrand: MMED Research report. Cited 2014-06-30. Available from: URL: http://wiredspace.wits.ac.za/jspui/bitstream/10539/15463/2/Vermeulen_ResReport.pdf

84 **Genovese G**, Handsaker RE, Li H, Kenny EE, McCarroll SA. Mapping the human reference genome's missing sequence by three-way admixture in Latino genomes. *Am J Hum Genet* 2013; **93**: 411-421 [PMID: 23932108 DOI: 10.1016/j.ajhg.2013.07.002]

**P-Reviewer:** Futrakul P, Watanabe T **S-Editor:** Tian YL

**L-Editor:** **E-Editor:**

**Table 1 Spectrum of renal histology in HIV in South Africa**

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

[1]Adapted from Vermeulen Alda, MMed Research report, University of the Witwatersrand, 2014[83]. JHB: Johannesburg; HIVAN: HIV-associated nephropathy; TTP: Thrombotic thrombocytopaenic purpura; HUS: Haemolytic syndrome.

**Table 2 Summary of the studies of *MYH9* and *APOL1* variants**

| Year | Population (Ancestry) | Disease | Variant | Freq. | OR (95%CI) | Refs. |
|---|---|---|---|---|---|---|
| **2008** | African Americans | Hypertensive ESRD | *MYH9* E1 | 0.67 | 1.9 (1.25-2.87) | Kopp *et al*[31] |
| | | HIVAN | *MYH9* E1 | 0.67 | 5.3 (2.40-12.90) | |
| | | FSGS | *MYH9* E1 | 0.67 | 4.5 (2.92,7.19) | |
| | European Americans | T2DM ESRD | *MYH9* E1 | 0.04 | NS | |
| | | FSGS | *MYH9* E1 | 0.04 | 9.7 (1.07,463) | |
| **2008** | African Americans | Hypertensive ESRD | *MYH9* E1 | 0.30 | 2.1(1.56,2.74) | Kao *et al*[30] |
| | | Non-diabetic ESRD | *MYH9* E1 | 0.30 | 2.2 (1.73,2.73) | |
| | | FSGS | *MYH9* E1 | 0.30 | 3.7(2.11,6.34) | |
| **2009** | African Americans | Hypertensive ESRD | *MYH9* E1 | 0.75 | 2.4 (NS) | Freedman *et al*[47] |
| | | Non-diabetic ESRD | *MYH9* E1 | 0.76 | 2.5 (NS) | |
| **2009** | African Americans | T2DM ESRD | *MYH9* E1 | 0.67 | 1.4 (NS) | Freedman *et al*[44] |
| **2010** | African Americans | Non-diabetic ESRD | *MYH9* E1 | NS | 2.0 (1.37,2.92) | Behar *et al*[46] |
| | Hispanic Americans | Non-diabetic ESRD | *MYH9* E1 | NS | 3.7 (1.67,8.20) | |



| | | | | | | |
|---|---|---|---|---|---|---|
| **2010** | American Indians | Kidney dysfunction | *MYH9* SNPs | 0.43 | 1.04 (0.79,1.36) | Franceschini *et al*[43] Strong Heart Family Study |
| **2010** | African Americans | Non-diabetic ESRD | *APOL1* G1 | 0.46 | 4.86 (2.35,10.06) | Tzur *et al*[29] |
| | Hispanic Americans | Non-diabetic ESRD | *APOL1* G1 | 0.00 | 15.48 (4.00,60.00) | |
| **2010** | African Americans | Hypertensive ESRD | *APOL1* G1/G2 | 0.41/0.21 | 7.3 (5.60,9.50) | Genovese *et al*[13] |
| | | FSGS | *APOL1* G1/G2 | 0.47/0.25 | 10.5 (6.0,18.4) | |
| **2011** | African Americans | HIVAN | *APOL1* G1/G2 | 0.54/0.28 | 29.2 (13.10,68.50) | Kopp *et al*[12] |
| | | FSGS | *APOL1* G1/G2 | 0.55/0.25 | 16.9 (11.00,26.50) | |
| **2014** | South African blacks | HIVAN | *MYH9* E1 *APOL1* G1/G2 | 0.83 0.56/0.34 | 2.10 (0.07-60.99) 89.10 (17.68,911.72) | Kasembeli *et al* (Unpublished observations) |

NS: Not stated; SNPs: Single nucleotide polymorphisms; OR: Odds ratio; Freq: Frequencies; HIVAN: HIV-associated nephropathy; FSGS: Focal segmental glomerulosclerosis; T2DM: Type 2 diabetes mellitus; ESRD: End stage renal disease; *APOL1*: Apolipoprotein L1; *MYH9*: Non-muscle myosin heavy chain 9.

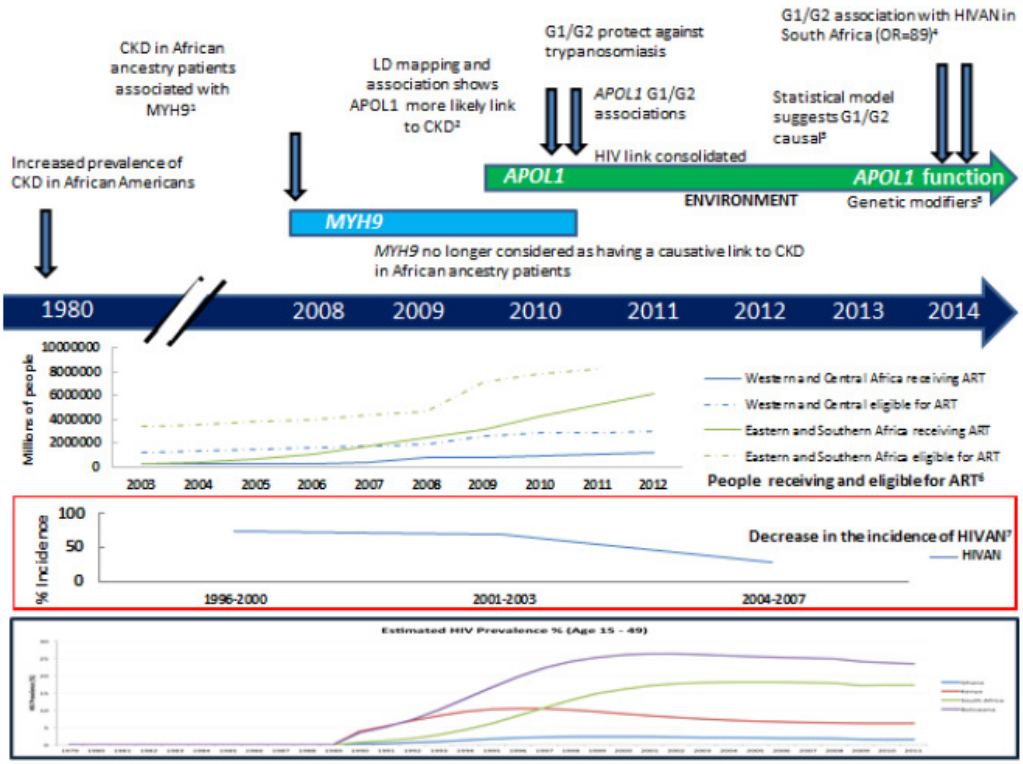



**Figure 1 Historical timeline reflecting the discovery of genetic association to CKD in populations with African ancestry.** [1]Adapted from Kopp *et al*[31] and Kao *et al*[30]; [2]Adapted from Freedman *et al*[49], Genovese *et al*[13], Tzur *et al*[29]; [3]Adapted from Genovese *et al*[84]; [4]Adapted from Kasembeli *et al* (2014 unpublished observations); [5]Adapted from Freedman *et al*[64]; [6]Adapted from UNAIDS report on global AIDS epidemic[18]; [7]Adapted from Vermeulen A, MMed Research report, University of the Witwatersrand, 2014. JHB: Johannesburg; *APOL1*: Apolipoprotein L1; *MYH9*: Non-muscle myosin heavy chain 9; ART: Antiretroviral therapy.

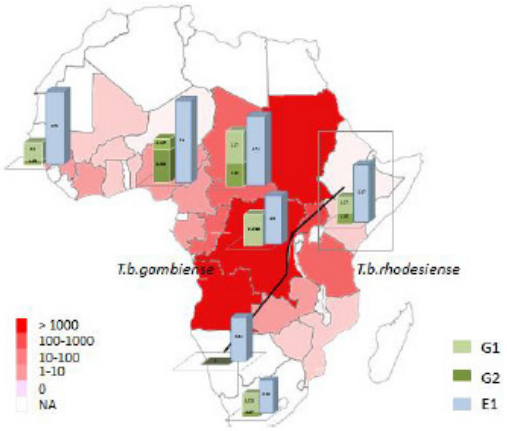



**Figure 2 Distribution of Human African Trypanosomiasis (T.b *gambiense* and T.b *rhodesiense*), *MYH9* E1 and *APOL1* G1 and G2 risk alleles in Africa[12,42].** The frequency of distribution of *APOL1* risk variants in Africa are represented by bar charts and overlap the areas distribution of Human African Trypanosomiasis. The numbers reflect the reported cases of Trypanosomiasis from the WHO, 2010. T.b: *Trypanosoma brucei*.

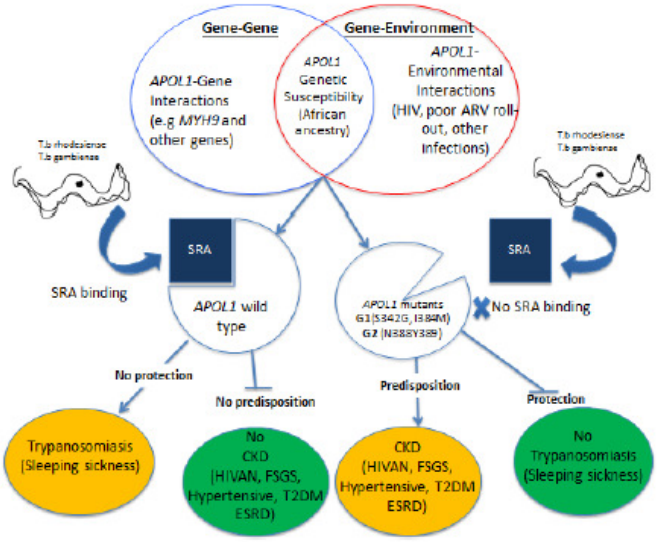



**Figure 3 Gene-Gene, Gene-Environment steering contribution to *APOL1* associated CKD and the positive selection of *APOL1* associated CKD variants as a result of Trypanosomiasis.** SRA: Serum resistant associated protein; HIV: Human immunodeficiency virus; T.b: *Trypanosoma brucei*; *APOL1*: Apolipoprotein L1; *MYH9*: Non-muscle myosin heavy chain 9; HIVAN: HIV-associated nephropathy; FSGS:

Focal segmental glomerulosclerosis; T2DM: Type 2 diabetes mellitus; ESRD: End stage renal disease; CKD: Chronic kidney disease.

# APPENDIX 3

# ETHICS CLEARANCE CERTIFICATE M111185

**UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG**
Division of the Deputy Registrar (Research)

**HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)**
R14/49 Mr Alex Kasembeli

| | |
|---|---|
| **CLEARANCE CERTIFICATE** | M111185 |
| **PROJECT** | Genetic Predisposition to Kidney Disease in HIV Positive South African Blacks: Genetic Dissection of the MYH9 and APOL1 Gene Regulation |
| **INVESTIGATORS** | Mr Alex Kasembeli. |
| **DEPARTMENT** | Department of Internal Medicine/Nephrology |
| **DATE CONSIDERED** | 25/11/2011 |
| **M1111850DECISION OF THE COMMITTEE*** | Approved unconditionally |

**Unless otherwise specified this ethical clearance is valid for 5 years and may be renewed upon application.**

**DATE** 20/01/2012 **CHAIRPERSON** .....
(Professor PE Cleaton-Jones)

*Guidelines for written 'informed consent' attached where applicable
cc: Supervisor :

---

---

**DECLARATION OF INVESTIGATOR(S)**
To be completed in duplicate and **ONE COPY** returned to the Secretary at Room 10004, 10th Floor, Senate House, University.
I/We fully understand the conditions under which I am/we are authorized to carry out the abovementioned research and I/we guarantee to ensure compliance with these conditions. Should any departure to be contemplated from the research procedure as approved I/we undertake to resubmit the protocol to the Committee. **I agree to a completion of a yearly progress report.**

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES...

**UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG**
Division of the Deputy Registrar (Research)

**HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)**
R14/49 Professor M Ramsay et al

| | |
|---|---|
| **CLEARANCE CERTIFICATE** | **M10745** |
| **PROJECT** | Genetic Diseases Related Studies in Southern African Populations: Control Samples (Previously M050706) |
| **INVESTIGATORS** | Professor M Ramsay et al. |
| **DEPARTMENT** | Division of Human Genetics |
| **DATE CONSIDERED** | 06/08/2010 |
| **DECISION OF THE COMMITTEE*** | Renewal approved |

**Unless otherwise specified this ethical clearance is valid for 5 years and may be renewed upon application.**

**DATE** 06/08/2010 **CHAIRPERSON** .....
(Professor PE Cleaton-Jones)

*Guidelines for written 'informed consent' attached where applicable
cc: Supervisor : Prof A Krause

---

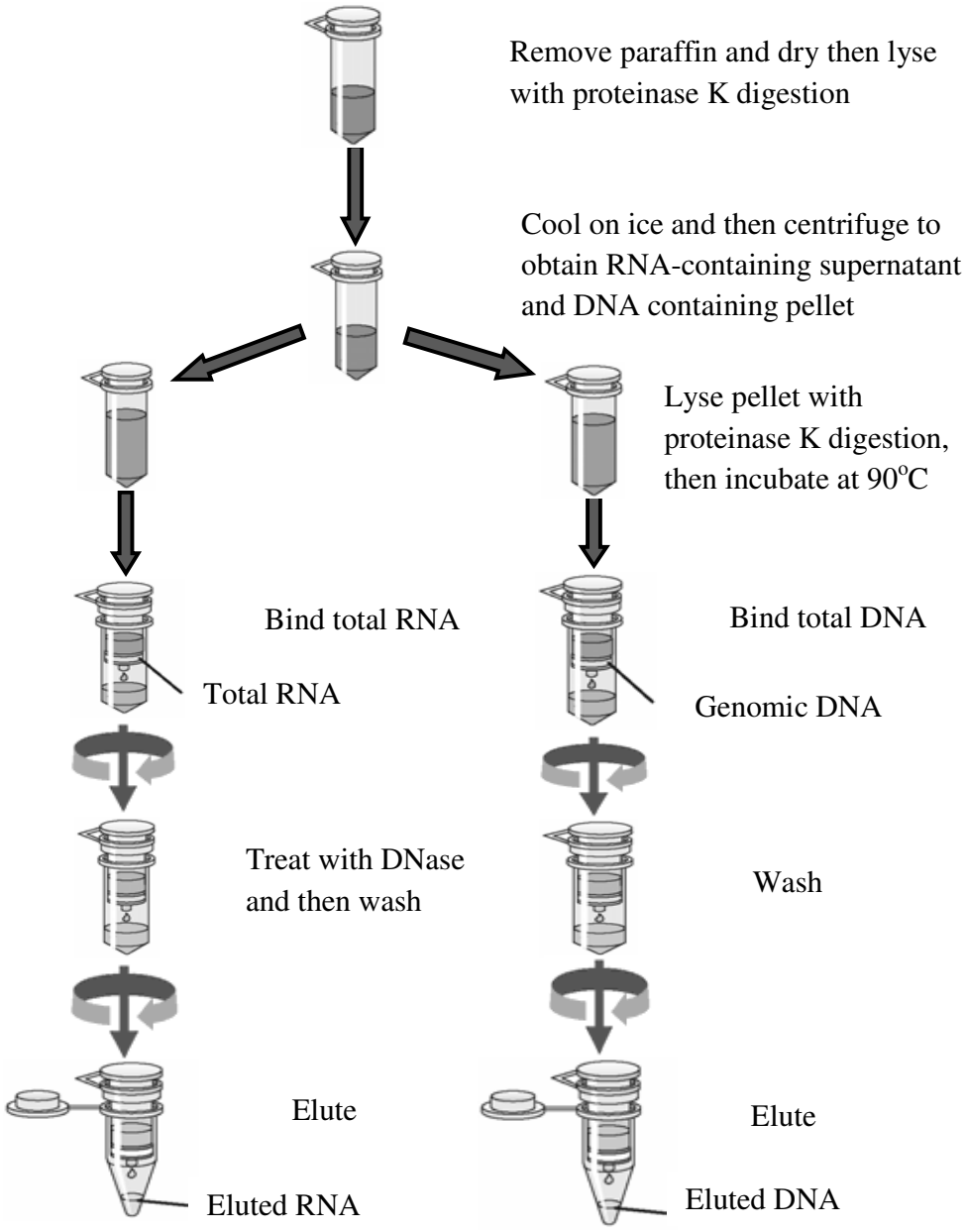
**DECLARATION OF INVESTIGATOR(S)**

To be completed in duplicate and **ONE COPY** returned to the Secretary at Room 10004, 10th Floor, Senate House, University.
I/We fully understand the conditions under which I am/we are authorized to carry out the abovementioned research and I/we guarantee to ensure compliance with these conditions. Should any departure to be contemplated from the research procedure as approved I/we undertake to resubmit the protocol to the Committee. **I agree to a completion of a yearly progress report.**
PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES...

M. Ramsay
16/8/2010

## APPENDIX 4

### All Prep DNA/RNA FFPE schematic laboratory procedure



## APPENDIX 5

### DNA EXTRACTION FROM HUMAN NUCLEATED CELLS IN WHOLE BLOOD USING SALTING OUT PROCEDURE

**Solutions Needed:**

1. **Sucrose-Triton-X Lysis Buffer**
   In a one liter bottle, 10ml of 1M Tris-HCL pH 8.0 was added followed by 5ml 1M $MgCl_2$ and 10ml 1X Triton X-100 solution.
   This was made-up to 1 litre with distilled water, autoclaved and stored at $4^oC$.
   109.5g of sucrose was added just before use. Once the sucrose has been added, the solution does not keep, so only sufficient amounts of each batch was made. Approximately 70ml/ 10ml sample is needed.

2. **20mM Tris-HCL and 5mM EDTA (T 20 E 5) Solution**
   In a 1 litre bottle, 20ml of 1M Tris-HCL was added (to make 20 mM of Tris-HCL) followed by 10mL of 0.5M EDTA (to make 5mM of EDTA). Distilled water was added to 800ml mark, pH 8.0 and made up to 1 litre mark. Autoclaved and stored at room temperature.

3. **Saturated Sodium Chloride (NaCl) solution.**
   40g of NaCl was slowly added to 100ml sterile distilled water until completely saturated and then store at room temperature.

4. **Proteinase K Master Mix.**
   10% Sodium dodecyl sulphate (SDS) made in sterile distilled water.
   0.5M EDTA made in sterile distilled water
   20mg/ml Proteinase K

| | Stock Solution | 2 mL | 4 mL | 8 mL |
|---|---|---|---|---|
| 1% SDS | 10% | 200µl | 400 µl | 800 µl |
| 2mM EDTA | 0.5M | 8 µl | 16 µl | 32 µl |
| 2mg/mL Proteinase K | 20mg/mL | 200 µl | 400 µl | 800 µl |
| Water ($H_2O$) | | 1600 µl | 3200 µl | 6400 µl |

Prepare as per the table above depending on the number of samples. Make fresh each time because you need 500 µl /10ml sample.

### PROCEDURE

**Day One**

1. 5 to 10ml of whole blood was placed in 50ml polypropylene tubes (Falcon tubes). The extraction procedure was done in a batch of 4 samples
2. Each Falcon tube containing whole blood was filled to the 45ml mark with cold lysis buffer containing sucrose in 1M Tris-HcL pH 8.0, 1M $MgCl_2$ and 1X Triton X-100 solutions.
3. The Falcon tubes were mixed well by inverting several times and then centrifuged at 2300rpm for 10 minutes at 4$^{o}$C.
4. The supernatant containing lysed red blood cells were poured off as "blood waste". The pellet (reddish-white) containing white blood cells were left in the tube.
5. There were then re-suspended in another 20ml of cold lysis buffer and placed on ice for 5 minutes.
6. The tubes were then centrifuged for a further 10 minutes at 2300 rpm at 4$^{o}$C. This wash step was done to ensure that most of the red blood cell debris was removed in the supernatant.
7. 3 ml of T20E5 solution, 0.2ml 10% SDS and 0.5ml Proteinase K mix were added to each tube containing sample. This process lyses the white blood cells and degrades the proteins. SDS facilitates the breaking up of the lipid bilayer of the cell membrane.
8. The tube contents were then mixed thoroughly by vortexing and incubated in water bath at 42$^{o}$C overnight.

**Day Two**

1. 1ml of saturated NaCl solution was added to each tube, mixed well by vortexing and placed on ice for 5 minutes.
2. The tubes with salt were then centrifuged at 2300 rpm for 30 min. The pellet at this step consists of proteins precipitated by the salt and the DNA is present in the supernatant. Cloudy supernatant were placed into clean tubes and centrifuged for a further 20 minutes.
3. The DNA-containing supernatant was transferred into clean 50ml falcon tubes and 20ml absolute ethanol (96-100%) added at room temperature. In the event of high salt concentration, DNA is not soluble in ethanol hence DNA precipitated out of solution with addition of the absolute ethanol.
4. After gentle agitation, the white aggregate visible DNA was spooled out/ fished out using a clean P200 pipette tip and washed gently in ice cold 70% ethanol to remove excess salt that could interfere with enzymatic reactions. The DNA was then placed in clean labeled 1.5 ml eppendorf tube and ethanol allowed to evaporate off (air-dried) at room temperature. The DNA was NOT allowed to over dry and dissolved in an appropriate amount of 1X Tris-EDTA (TE) buffer DNA (100-200µl).
5. In case the DNA was not visible, the DNA was precipitated at minus 70$^{o}$C for 30 minutes and centrifuged at 2300rpm for 10 minutes to pellet DNA. The supernatant was discarded and the DNA pellet was washed with ice cold 70% ethanol and centrifuged again for 10 minutes. The supernatant was discarded and the falcon tube was inverted to a paper towel to dry the pellet before resuspending the DNA pellet in an appropriate amount of 1X TE buffer DNA (20-100µl). This was then transferred to labeled clean 1.5ml eppendorf tube

## APPENDIX 6

## ILLUMINA GOLDENGATE BEADXPRESS GENOTYPING ASSAY

### 1. Materials

- Three 96-plex VeraCode bead plates where each well containing the beads providing identifiers for each SNP
- Query oligonucleotides targeting the specific 96 SNPs selected
- Resuspension and precipitation solutions
- Oligonucleotide hybridization and cDNA and genomic DNA binding buffers
- Universal buffers for washing magnetic beads and particles
- Uracyl DNA Glycosylase to help prevent contamination of PCR products
- Reagents that:
    - activate DNA
    - extend and ligate
    - elute extended and ligated products
    - bind double-stranded PCR products to magnetic particles
    - are used to make hybridization plate
    - wash away excess and non-specifically hybridized oligonucleotides from the genomic DNA
    - are used to wash the beads

## 2. Procedure

*DAY ONE*

## 2.1 Pre-PCR

### 2.1.1 Making DNA plate

There was an addition of 5µl appropriate DNA activating reagent containing Biotin on each well of the 96-well plate. This was followed by addition of sufficient genomic DNA of each individual normalized to 50ng/µl. This process is to activate the DNA. The plate was then pulse centrifuged at 250xg to collect contents at the bottom of the wells, vortexed at 2300rpm for 20 seconds and incubated at 95$^{o}$C for 30 minutes. In this process, the DNA is denatured to biotinylated single-stranded DNA in the plate.

### 2.1.2 Precipitation of the single-stranded biotinylated DNA

Precipitation of the single-stranded biotinylated DNA was performed by addition of 15µl Isopropanol to each sample on the well since DNA is insoluble in Isopropanol. The plate was then centrifuged at 3000xg for 20 minutes. The supernatant was decanted over an absorbent pad by smacking the inverted plate down hard onto the absorbent pad to blot off the excess supernatant. The precipitated single-stranded DNA was then allowed to air-dry for 15 minutes.

### 2.1.3 Resuspension of the single-stranded DNA in the 96-well plate

To each well containing the single-stranded DNA, 10µl of resuspension reagent solution resuspended was added, the plate was pulse centrifuged at 250xg and vortexed at 2300rpm for 1 minute or until the blue pellet is completely resuspended. The plate was then pulse centrifuged at 250xg. The DNA here is fully resuspended.

### 2.1.4 Making allele-specific extension plate

To a new 96-well plate, 10µl of the custom Illumina® GoldenGate BeadXpress assay oligo pool containing two allele-specific oligonucleotides and one locus-specific oligonucleotide for each of the 96 selected SNPs was added. This was followed by 30µl of the oligo hybridization and DNA binding buffer containing streptavidin-conjugated magnetic particles was added.

Transfer of 10µl of the biotinylated DNA samples from the DNA plate to the complimentary wells of the allele-specific extension plate. This plate was heat-sealed for 3 seconds and pulse centrifuged at 250xg, vortexed at 1600rpm for 1 minute or until beads are completely resuspended. The plate was then placed on a pre-heated heating block at 70$^{o}$C and immediately the heating block allowed to drop in temperature to 30$^{o}$C. This takes about 2 hours and this allows the oligonucleotides for each sequence target of interest to anneal to the biotinylated genomic DNA samples and at the same time simultaneously captured by magnetic particles through the binding of biotin to streptavidin.

### 2.1.5 Master Mix addition for extension and ligation

Once the query oligos were hybridized on the genomic DNA, mishybridized and excess oligos were washed away in repeated steps using a raised bar magnetic capture plate to allow the magnetic beads to remain in the wells while all liquid discarded and replaced. A washing reagent was used to resuspended the beads in each well twice before repeating the process twice again with a universal buffer. To each washed genomic DNA, 37µl of extension and ligation master mix was added, the plate sealed

with a clear adhesive film, vortexed at 1650rpm for 1 minute to resuspend the beads and incubated on a pre-heated heating block at 45$^{o}$C. This is to allow the extension and ligation reaction to take place.

### 2.1.6 Making of the polymerase chain reaction (PCR) plate

There was an addition of 64µl of Titanium Taq DNA polymerase and 50µl of Uracyl DNA Glycosylase to the reagent for master mix (containing three universal primers: two are labelled with fluorescent dyes and the third is biotinylated which allows capture of the PCR products and elution of the strands containing fluorescent signals. 30µl of this mixture was transferred into a new 96 well plate "PCR plate", the plate sealed with a clear adhesive film, pulse centrifuged at 250xg to bring reagent mixtures to the bottom of the wells.

### 2.1.7 Inoculation of the PCR plate

The allele-specific extension plate was placed on the raised bar magnetic plate for approximately 2 minutes or until the beads were completely captured. The clear adhesive film from the plate was removed and the supernatant removed and discarded. The pipette tips were visually inspected after removing the supernatant from each column of the plate to ensure that no beads were removed. In case the beads were visible in the pipette tips, the supernatant were replaced into the same wells, the magnet allowed to re-collect beads and the pipette tips changed. The beads were then washed with universal buffer as described above.

35µl of the reagent for inoculating PCR plate was added to each column of the plate, plate covered with clear adhesive film, sealed, vortexed at 1850rpm for 1 minute or until the beads were resuspended and placed on the pre-heated 95$^{o}$C heating block for

1 minute. The temperature here denatures the DNA and releases the templates formed in the extension and ligation process from genomic DNA.

The plate was placed on the raised magnetic plate for 2 minutes or until the beads were completely resuspended and approximately 30µl of the supernatant containing the complementary templates were transferred to corresponding wells on the PCR plate while the plate containing the beads discarded. The PCR plate was then sealed with a clear adhesive film and immediately transferred to the thermal cycler.

## 2.2 Post-PCR

### 2.2.1 Thermal cycling the PCR plate

The sealed PCR plate was placed in the thermal cycler and the thermal cycler run at 37$^{o}$C for 10 minutes, 95$^{o}$C for 3 minutes, followed by 34 cycles (of 95$^{o}$C for 35 seconds, 56$^{o}$C for 35 seconds and 72$^{o}$C for 2 minutes) then 72oC for 10 minutes and finally 4$^{o}$C for 5 minutes and after the end of the PCR, the PCR plate was removed from the thermal cycler placed at -20$^{o}$C overnight. This process thermal cycled the PCR plate to fluorescently label and amplifies the templates generated in the pre-PCR process.

***DAY TWO***

### 2.2.2 Binding PCR products

The PCR plate was removed from -20$^{o}$C, and thawed on ice. It was then pulse centrifuged at 250xg. Using new pipette tips, 20µl of magnetic particle reagent that is used to bind the biotinylated double-stranded PCR products was added to each wells of the PCR plate. Using new pipette tips, the mixture was pipetted up and down to mix the beads with the PCR products, and then transferred the mixed solution to the

filter plate, changing tips between column dispenses. The empty PCR plate was discarded. The filter plate was covered and stored in the dark (protected from the light) at room temperature for 60 minutes in order for the PCR products to bind the magnetic particles.

### 2.2.3 Making intermediate plate for VeraCode Bead plate

The filter plate containing the bound PCR products is placed onto a filter plate adapter and centrifuged at 1000xg for 5 minutes at 25$^{o}$C in order to discard the waste fluid into the waste plate as bound PCR products were captured by the filter. This process was repeated after addition of universal buffer.

The waste plate was replaced by new empty 96-well plate with hybridization buffer (intermediate plate) added to each well of the intermediate plate. 0.1N Sodium hydroxide was then added to each well of the filter plate, covered and centrifuged at 1000xg for 5 minutes at 25$^{o}$C. Sodium hydroxide lowers the pH and breaks the bonds between the biotinylated PCR products and the streptavidin-conjugated magnetic particles. In this step, the solution containing the PCR products was dispensed into the intermediate plate containing hybridization buffer. The filter plate containing the magnetic particles was discarded.

### 2.2.4 Hybridizing the VeraCode Bead plate

A second hybridization buffer was neutralized by 0.1N Sodium hydroxide. Using clean pipette tips, 50µl of the neutralized hybridization buffer was transferred to each wells of the intermediate plate above. Using pipette tips, 100µl of each assay product was added into the corresponding wells of the VeraCode bead plate containing the inscribed microbeads.

The VeraCode bead plate with samples was then placed in the pre-heated VeraCode vortex incubator (incubating microplate shaker) at 45$^{o}$C at a speed of 850rpm for 3 hours. In this step, the PCR products were hybridized to the beads on the VeraCode bead plate.

### 2.2.5 Washing VeraCode Bead plate

The VeraCode vortex incubator was stopped after 3 hours and when the speed indicator reached zero, the VeraCode bead plate was removed from the VeraCode vortex incubator, pulse centrifuged at 250xg to collect any condensation and cover removed. Using clean pipette tips, 200µl of wash buffer to each well making sure that bead pellets are agitated. The plate was gently swirled in a circular motion on the bench-top, and then the plate was left for 2 minutes for the beads to collect in the bottom of the well. The supernatant was then aspirated with the vacuum manifold with a pressure of 50 mb. This process was repeated one more time.

### 2.2.6 Scanning VeraCode Bead plate

The scan settings file containing the information of the samples was prepared on the set up of the BeadXpress Reader. The VeraCode bead plate was then placed in the BeadXpress Reader to be scanned.

## APPENDIX 7

### A) 5% DNA Acrylamide Gel (1.5mm wells)

**Materials:**

Distilled water

10X Tris Borate EDTA (TBE)

1X Tris Borate EDTA (TBE)

30% Acrylamide

10% Ammonium persulfate (APS)

N, N, N', N'-tetramethylethylenediamine (TEMED)

Glass plates, spacers and 1.5mm combs for pouring the gels

Gel electrophoresis apparatus

DC power supply

**Set up:**

Gel electrophoresis apparatus, glass plates, spacers, combs and casting systems (*BIO-RAD*) were assembled according to the manufacturer's instructions. These apparatus have been designed to avoid any leakages.

The following reagents were mixed in a 15ml tubes according to the template and in the order below:

| REAGENTS | VOLUME (ML) |
|---|---|
| Distilled water | 6.6 |
| 10X Tris Borate EDTA | 1.0 |
| 30% Acrylamide | 2.4 |
| 10% APS | 0.1 |
| TEMED | 0.01 |

10% APS and TEMED were added last because polymerization starts immediately thereafter and therefore, immediately after the addition of 10% APS and TEMED, the reagent mixture was

mixed thoroughly and poured between the gel plates and 1.5mm comb inserted The combs were then clamped in place at the top of the gels to avoid separation of the gel from the plates as the acrylamide polymerizes. The gels were allowed to polymerize for 30 minutes. After the polymerization was complete, the combs and the bottom spacers were removed from the gel and the gel plates wiped and made sure that the spacers are clean and properly seated. The gel was placed into the lower tank at an angle to avoid air bubbles forming between the plates and the gel bottom. The gel plates were then clamped to the top of the electrophoresis tank and the upper reservoir filled with 1X TBE so that to cover the wells. The lower reservoir of the electrophoresis tank was also filled with 1X TBE. In case of any air bubbles in the wells, a syringe was used to remove the air bubbles trapped because this might disrupt the current flow. 10µl of the samples were mixed with a 6X loading dye (ThermoScientific$^{TM}$, Bishop Meadow Road, Loughborough, Leicestershire, LE11 5RG, United Kingdom) in a ratio of 1:1 and loaded onto the appropriate wells. A DC power supply was connected to the casting system and the samples electrophoresed at 60V/cm for 45 minutes the PCR products had migrated one-half to three-fourths the way through the gel.

**B) 2.5% Agarose gel**

**Materials:**

Agarose powder

Gel box

1X Tris Borate EDTA (TBE)

Combs

DC power supply

**Set up:**

2.5 grams of agarose was weighed into a conical flask followed by 100ml of 1X TBE solution. The agarose/buffer mixture was heated in a microwave in order to dissolve the agarose. At 30 seconds intervals, the flask was removed and swirled to mix. This was repeated until the agarose has completely dissolved.

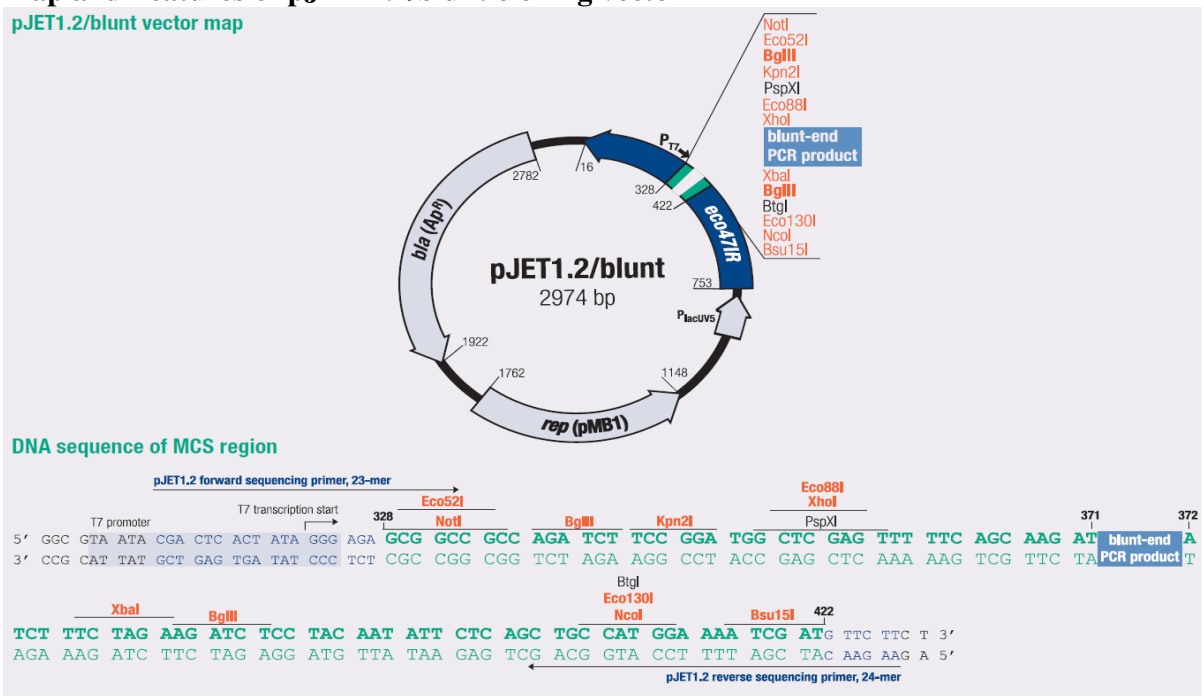
The agarose mixture was allowed to cool to about 60$^{o}$C and 3µl of ethidium bromide added for staining and visualization of PCR products.

The gel tray was placed into the casting apparatus, taped the open edges of a gel tray to create a mold and an appropriate comb placed into the gel mold to create the wells.

The molten agarose was then poured into the gel mold and the agarose allowed to set at room temperature for approximately 1 hour. Once set, the comb was removed and the gel placed in the gel box. Running buffer (1X TBE) was then added onto the molten gel and into the gel box in order to cover the surface of the molten gel. 10µl of the samples were mixed with a 6X loading dye (ThermoScientific$^{TM}$, Bishop Meadow Road, Loughborough, Leicestershire, LE11 5RG, United Kingdom) in a ratio of 1:1 and loaded onto the appropriate wells. Loading dye helps to track how far your DNA sample has travelled, and also allows the sample to sink into the gel. A DC power supply was connected to the casting system in such a way that the cathode (black leads) was closer to the wells than the anode (red leads). The DC power was turned on and the gel with samples run until the dye had migrated one-half to three-fourths the way through the gel.

# APPENDIX 8

## Map and Features of pJET 1.2/blunt cloning vector




## Genetic Elements of the pJET 1.2/blunt cloning vector

| Element | Function | Position, bp |
|---|---|---|
| rep (pMB1) | Replicon (rep) from the pMBI plasmid is responsible for the replication of pJET1.2 | 1762-1148 |
| Replication start | Initiation of the replication | 1162 ±1 |
| *bla* (Ap$^R$) | ß-lactamase gene conferring resistance to ampicillin. Used for selection and maintenance of recombinant *E.coli* cells | 2782-1922 |
| *eco47IR* | Lethal gene *eco47IR* enables positive selection of the recombinants | 753-16 |
| $P_{lacUV5}$ | Modified $P_{lac}$ promoter for expression of the *eco47IR* gene at a level sufficient to provide positive selection | 892-769 |
| T7 promoter | T7 RNA polymerase promoter for *in vitro* transcription of the cloned insert | 305-324 |
| Multiple cloning site (MCS) | Mapping, screening and excision of the cloned insert | 422-328 |
| Insertion site | Blunt DNA ends for ligation with insert | 371-372 |
| Primer binding sites: | | |
| pJET1.2 forward sequencing | Sequencing of insert, colony PCR | 310-332 |
| pJET1.2 reverse sequencing | Sequencing of insert, colony PCR | 428-405 |

**Primer Sequences**

| Primer | Sequence |
|---|---|
| pJET1.2 Forward Sequencing Primer, 23-mer | 5'-CGACTCACTATAGGGAGAGCGGC-3' |
| pJET1.2 Reverse Sequencing Primer, 24-mer | 5'-AAGAACATCGATTTTCCATGGCAG-3' |

## APPENDIX 9

**Preparing E. coli DH5 α competent cells**

**Recipes for Media**

| **LB (Psi) broth** (per liter) | |
|---|---|
| **Reagents** | **Amount** |
| Bacto yeast extract | 5 g |
| Bacto Tryptone | 20 g |
| magnesium sulfate | 5 g |
| *pH 7.6 with potassium hydroxide and a*utoclaved for 40 min. | |

| **TfbI** (per 200 ml) | | |
|---|---|---|
| **Reagents** | **Amount** | **Final molarity/conc.** |
| Potassium acetate | .588 g | 30 mM |
| Rubidium chloride | 2.42 g | 100 mM |
| Calcium chloride | 0.294 g | 10 mM |
| Manganese chloride | 2.0 g | 50 mM |
| Glycerol | 30 ml | 15% v/v |
| *pH 5.8 with dilute acetic acid* and Filter-sterilized | | |

| **TfbII** (per 100 ml) | | |
|---|---|---|
| **Reagents** | **Amount** | **Final molarity/conc.** |
| MOPS | 0.21 g | 10 mM |
| Calcium chloride | 1.1 g | 75 mM |
| Rubidium chloride | 0.121 g | 10 mM |
| Glycerol | 15 ml | 15% v/v |
| *pH 6.5 with dilute NaOH* and Filter-sterilized | | |

1. Cells were streaked out from a glycerol stock onto a LB agar plate using an inoculation loop and incubated at 37°C overnight.
2. A colony was picked and grown overnight at 37 °C in 5 ml psi (ψ) broth.
3. 1 ml of the overnight culture was placed in 25 ml pre-warmed ψ broth and incubated at 37 °C overnight on a shaker at ~100rpm. The $OD_{600}$ was checked to make sure it was between 0.52 and 0.58.
4. All the apparatus to be used, including pipettes for re-suspending, were placed in the cold room. **SUBSEQUENT STEPS WERE PERFORMED IN THE COLD ROOM.**
5. The cells were then chilled on ice for 5 minutes and centrifuged at 4000 rpm (~4000 x g) for 10 minutes at 4 °C or at 3000 rpm (~2000 x g) for 20 minutes at 4 °C.
6. The supernatant was decanted in the cold room and the pellet re-suspended in 5 ml TfbI buffer. Make up to 70 ml with TfbI buffer and left on ice for 5 minutes.
7. The cells were harvested again at 4000 rpm for 10 minutes in the cold room once again, the supernatant decanted and the cells resuspended in 7 ml TfbII buffer and left on ice for 15 minutes.
8. Cells were then aliquoted (~100µl) into pre-cooled tubes in the cold room.
9. The cells were then snap frozen in liquid nitrogen and stored at -80 °C.
10. We tested for competency efficiency by transforming 10pg and 100pg of supercoiled plasmid/vector into 50 µl cells and compared efficiency to that of old cells that are competent.

## APPENDIX 10

### A) Super optimum culture (SOC) Medium

**Recipes**

In order to make 250mM potassium chloride (KCl), 1.86g of potassium chloride was dissolved in 100ml of distilled water.

In order to make 20% glucose, 100 g of glucose was dissolved in 400 ml distilled water completely and the volume adjusted to 500 ml with distilled water and filter sterilized using a 0.22-micron filter.

In order to make 2M $MgCl_2$: 19 g of $MgCl_2$ was dissolved in 90 ml of distilled water and the volume adjusted to 100 ml with distilled water and sterilized by autoclaving for 20 minutes. This solution was left to cool before use.

| **Per 200 ml solution** | |
|---|---|
| **Reagents** | **Amount** |
| Tryptone | 4g |
| Yeast Extract | 1g |
| Sodium chloride | 0.1g |

The above contents were measured and placed in a 250ml bottle and 150ml distilled water added. The solutes were shaken until dissolved.

2 ml of 250mM KCl was then added and the pH adjusted to 7.0 with 1N NaOH. The overall volume was then adjusted to 200ml with distilled water and sterilized by autoclaving for 20 minutes. The medium was allowed to cool to 60$^{o}$C or less and then 3.6ml of filter-sterilized 20% Glucose solution was added followed by 1 ml of sterile 2 M $MgCl_2$ solution.

The final SOC medium was then filter-sterilized and stored at 4$^{o}$C.

### B) Luria Burtani (LB) Agar

**Materials:**

Tryptone

Yeast

Sodium Chloride

Bacto Agar

5 grams of Tryptone, 2.5 grams of yeast, 5 grams of sodium chloride and 7.5 grams of bacto agar were mixed in a 500ml bottle followed by a top-up to the 500ml mark with distilled water. This solution was then autoclaved for 30 minutes and let to cool at room temperature to approximately 60$^{o}$C. 500µl of appropriate antibiotic was added and poured onto bacterial plates as the bottle is flamed using a Bunsen burner to avoid contamination of the LB agar.

### C) Luria Burtani (LB) broth

**Materials:**

Tryptone

Yeast

Sodium Chloride

5 grams of Tryptone, 2.5 grams of yeast and 5 grams of sodium chloride were mixed in a 500ml bottle followed by a top-up to the 500ml mark with distilled water. This solution was then autoclaved for 30 minutes and let to cool at room temperature to approximately 60$^{o}$C. 500µl of appropriate antibiotic was added and aliquoted in 50ml tubes when growing in cultures as the bottle is flamed using a Bunsen burner to avoid contamination of the LB broth.

## APPENDIX 11

**Plasmid DNA extraction protocol**

3ml of the overnight grown bacterial culture was centrifuged for 30 seconds at 14,000rpm and the supernatant discarded. 600µl of filter sterilized distilled water was added to the bacterial cell pellet and re-suspended completely. The re-suspended bacterial cell pellet were then transferred to 1.5ml microcentrifuge tubes and 100µl of 7X Lysis Buffer added to each tube, mixed well by inverting the tubes up-to 6 times. Addition of 7X Lysis Buffer changes the colour of the solution from opaque to clear blue indicating complete lysis. This step should take exactly 2 minutes. 350µl of cold Neutralizing Buffer was then added in all the tubes and mixed thoroughly by inverting the tubes up-to 6 times. The samples turned yellow when neutralization was complete and yellowish precipitates formed. The samples were then inverted an addition 3 times to ensure complete neutralization followed by centrifugation at 14,000 rpm for 2 minutes and the supernatant transferred to Zymo-Spin™ IIN columns. The cell debris pellets were not disturbed at this step. The columns were then placed into collection tubes and centrifuged at 14,000 rpm for 15 seconds. The flow-through was discarded and the columns placed back into the same collection tubes. 200µl of Endo-Wash Buffer was added to each tube and centrifuged at 14,000 rpm for 30 seconds followed by 400µl of Zyppy™ Wash Buffer and centrifuged for 1 minute at 14,000 rpm. The columns were transferred into each properly labelled 1.5ml microcentrifuge tube and 30µl of Zyppy™ Elution Buffer added directly to each column matrix and let to stand for one minute at room temperature and centrifuged for 30 seconds at 14,000 rpm to elute the plasmid DNA. The purified plasmid DNA concentrations were measured using NanoDrop ND 2000 spectrophotometer (ThermoScientific™, USA). All positive recombinant plasmids were

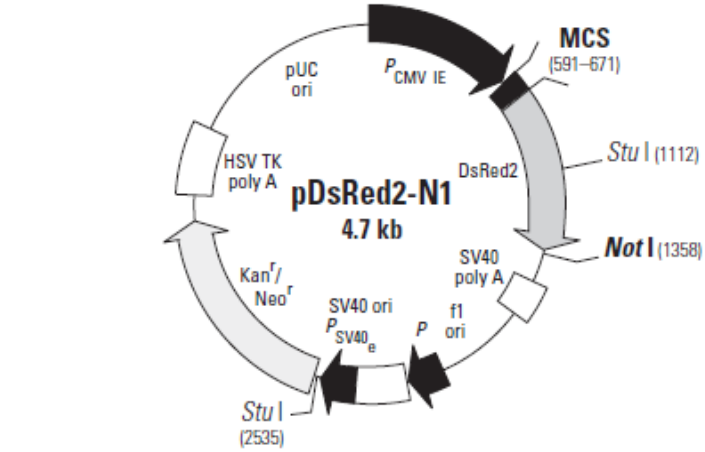
verified using Sanger Sequencing (Inqaba biotic, 525 Justice Mahomed street (formerly Walker Street) Muckleneuk, Pretoria. 0002).

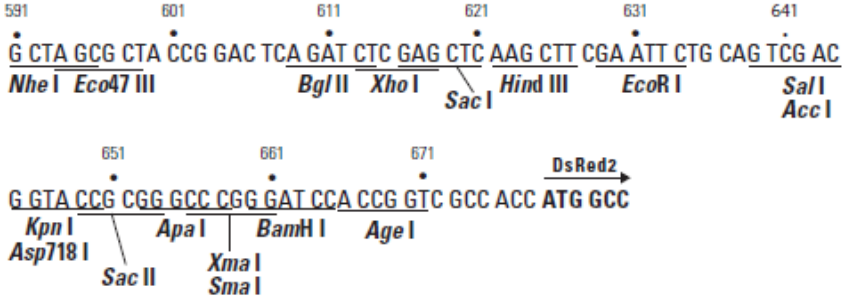
## APPENDIX 12

### Map and Features of pDsRed2-N1 cloning vector

**pDsRed2-N1 Vector Information** PT3604-5

Cat. No. 632406




**Restriction Map and Multiple Cloning Site (MCS) of pDsRed2-N1 Vector.** Unique restriction sites are in bold. The *Not* I site follows the DsRed2 stop codon.

# APPENDIX 13

## FINAL SNP LIST

**Table A13.1: Final SNP list**

| | Gene | Chromosome | SNP | Position | Genome build |
|---|---|---|---|---|---|
| **1** | *MYH9* | 22 | rs1005570 | 36715274 | 37 |
| **2** | *MYH9* | 22 | rs1009150 | 36702300 | 37 |
| **3** | *MYH9* | 22 | rs12107 | 36677982 | 37 |
| **4** | *MYH9* | 22 | rs16996648 | 36692752 | 37 |
| **5** | *MYH9* | 22 | rs16996668 | 36719185 | 37 |
| **6** | *MYH9* | 22 | rs16996672 | 36725970 | 37 |
| **7** | *MYH9* | 22 | rs16996674 | 36726652 | 37 |
| **8** | *MYH9* | 22 | rs16996677 | 36727283 | 37 |
| **9** | *MYH9* | 22 | rs17806513 | 36700386 | 37 |
| **10** | *MYH9* | 22 | rs2071730 | 36690026 | 37 |
| **11** | *MYH9* | 22 | rs2071731 | 36718858 | 37 |
| **12** | *MYH9* | 22 | rs2157256 | 36707661 | 37 |
| **13** | *MYH9* | 22 | rs2157257 | 36708338 | 37 |
| **14** | *MYH9* | 22 | rs2239784 | 36714635 | 37 |
| **15** | *MYH9* | 22 | rs2481 | 36677400 | 37 |
| **16** | *MYH9* | 22 | rs3752462 | 36710183 | 37 |
| **17** | *MYH9* | 22 | rs4821480 | 36695247 | 37 |
| **18** | *MYH9* | 22 | rs4821481 | 36695942 | 37 |
| **19** | *MYH9* | 22 | rs5750248 | 36702892 | 37 |
| **20** | *MYH9* | 22 | rs5750250 | 36708483 | 37 |
| **21** | *MYH9* | 22 | rs5750252 | 36728256 | 37 |
| **22** | *MYH9* | 22 | rs5756129 | 36684092 | 37 |
| **23** | *MYH9* | 22 | rs5756133 | 36693980 | 37 |
| **24** | *MYH9* | 22 | rs5756152 | 36712472 | 37 |
| **25** | *MYH9* | 22 | rs5756153 | 36719012 | 37 |
| **26** | *MYH9* | 22 | rs5756154 | 36720424 | 37 |
| **27** | *MYH9* | 22 | rs5756156 | 36720779 | 37 |
| **28** | *MYH9* | 22 | rs6000226 | 36684567 | 37 |
| **29** | *MYH9* | 22 | rs6000229 | 36686159 | 37 |
| **30** | *MYH9* | 22 | rs6000233 | 36687962 | 37 |
| **31** | *MYH9* | 22 | rs6000239 | 36716029 | 37 |
| **32** | *MYH9* | 22 | rs6000240 | 36728034 | 37 |
| **33** | *MYH9* | 22 | rs6000243 | 36741565 | 37 |
| **34** | *MYH9* | 22 | rs7078 | 36677914 | 37 |
| **35** | *MYH9* | 22 | rs735853 | 36679215 | 37 |
| **36** | *MYH9* | 22 | rs735854 | 36679058 | 37 |
| **37** | *MYH9* | 22 | rs739096 | 36741740 | 37 |
| **38** | *MYH9* | 22 | rs762964 | 36679652 | 37 |
| **39** | *MYH9* | 22 | rs8137674 | 36702102 | 37 |
| **40** | *MYH9* | 22 | rs875725 | 36691691 | 37 |
| **41** | *MYH9* | 22 | rs9610486 | 36693442 | 37 |
| **42** | *APOL3* | 22 | rs11089781 | 36556768 | 37 |
| **43** | *APOL1* | 22 | rs129423 | 36657329 | 37 |
| **44** | *APOL1* | 22 | rs136147 | 36652889 | 37 |
| **45** | *APOL1* | 22 | rs136159 | 36657023 | 37 |
| **46** | *APOL1* | 22 | rs136160 | 36657119 | 37 |
| **47** | *APOL1* | 22 | rs136161 | 36657432 | 37 |
| **48** | *APOL1* | 22 | rs2012928 | 36664002 | 37 |
| **49** | *APOL1* | 22 | rs4820224 | 36653275 | 37 |

| | | | | | |
|---|---|---|---|---|---|
| **50** | *APOL1* | 22 | rs6000220 | 36649966 | 37 |
| **51** | *APOL1* | 22 | rs60910145 | 36662034 | 37 |
| **52** | *APOL1* | 22 | rs7284919 | 36652164 | 37 |
| **53** | *APOL1* | 22 | rs73885319 | 36661906 | 37 |
| **54** | *APOL1* | 22 | rs9610467 | 36649574 | 37 |
| **55** | *APOL1* | 22 | rs71785313 | 36662046 | 37 |
| **56** | *APOL1* | 22 | rs713929 | 36657596 | 37 |
| **57** | *APOL1* | 22 | rs2413395 | 36654716 | 37 |
| **58** | *APOL1* | 22 | rs136150 | 36654145 | 37 |
| **59** | *APOL1* | 22 | rs3886200 | 36663627 | 37 |
| **60** | *AIM* | 3 | rs1344870 | 21307401 | 37 |
| **61** | *AIM* | 5 | rs1363448 | 140783596 | 37 |
| **62** | *AIM* | 15 | rs1823718 | 74147244 | 37 |
| **63** | *AIM* | 17 | rs1858465 | 51142920 | 37 |
| **64** | *AIM* | 2 | rs1876482 | 17362568 | 37 |
| **65** | *AIM* | 9 | rs2077559 | 36014850 | 37 |
| **66** | *AIM* | 15 | rs2089740 | 36310531 | 37 |
| **67** | *AIM* | 19 | rs2112527 | 9603751 | 37 |
| **68** | *AIM* | 6 | rs217538 | 108483470 | 37 |
| **69** | *AIM* | 7 | rs65264 | 28545611 | 37 |
| **70** | *AIM* | 9 | rs679047 | 12883664 | 37 |
| **71** | *AIM* | 11 | rs714857 | 15974389 | 37 |
| **72** | *AIM* | 4 | rs720096 | 179551071 | 37 |
| **73** | *AIM* | 14 | rs722869 | 97277005 | 37 |
| **74** | *AIM* | 1 | rs723854 | 192511012 | 37 |
| **75** | *AIM* | 15 | rs735612 | 34076642 | 37 |
| **76** | *AIM* | 2 | rs952718 | 215888624 | 37 |
| **77** | *AIM* | 13 | rs953386 | 110943692 | 37 |
| ***78** | *MYH9* | 22 | rs1154997 | 123839334 | 37 |
| ***79** | *MYH9* | 22 | rs12160045 | 36719360 | 37 |
| ***80** | *APOL1* | 22 | rs132736 | 36598058 | 37 |
| ***81** | *APOL1* | 22 | rs136175 | 36661566 | 37 |
| ***82** | *APOL1* | 22 | rs136176 | 36661646 | 37 |
| ***83** | *MYH9* | 22 | rs1476009 | 36686056 | 37 |
| ***84** | *MYH9* | 22 | rs2032487 | 36695428 | 37 |
| ***85** | *MYH9* | 22 | rs2187776 | 36695173 | 37 |
| ***86** | *APOL1* | 22 | rs2239785 | 36661330 | 37 |
| ***87** | *MYH9* | 22 | rs2269529 | 36684354 | 37 |
| ***88** | *MYH9* | 22 | rs2413396 | 36708084 | 37 |
| ***89** | *APOL1* | 22 | rs4419330 | 36658855 | 37 |
| ***90** | *APOL1* | 22 | rs6000218 | 36648352 | 37 |
| ***91** | *MYH9* | 22 | rs6000244 | 36741886 | 37 |
| ***92** | *MYH9* | 22 | rs713659 | 36728692 | 37 |
| ***93** | *MYH9* | 22 | rs739101 | 36694333 | 37 |
| ***94** | *MYH9* | 22 | rs8143119 | 36727551 | 37 |
| ***95** | *MYH9* | 22 | rs11913874 | 36737703 | 37 |
| ***96** | *MYH9* | 22 | rs1557539 | 36714089 | 37 |
| ***97** | *MYH9* | 22 | rs3830104 | 36708624 | 37 |

AIM: Ancestry informative marker. *SNPs that did not pass quality control procedures.

# APPENDIX 14

## Allele frequencies and association of the *MYH9*, *APOL1* and AIMs SNPs

**Table A 14.1: SNP Association with 38 HIVAN and 54 HIV positive without CKD**

| | | | Risk Allele Frequency | | | | *P-values* | |
|---|---|---|---|---|---|---|---|---|
| **Gene** | **SNP** | **Chromosome** | **Risk Allele** | **Cases** | **Controls** | **Odds Ratio** | **Unadjusted** | **Adjusted** |
| *APOL3* | rs11089781 | 22 | A | 0.65625 | 0.234043 | 6.0815502 | 0.00037 | 0.002259905 |
| *APOL1* | rs9610467 | 22 | G | 0.966667 | 0.8125 | 6.5695444 | 0.0789818 | 0.05381787 |
| *APOL1* | rs6000220 | 22 | C | 0.62069 | 0.469388 | 2.4637728 | 0.645677 | 0.238214174 |
| *APOL1* | rs7284919 | 22 | C | 0 | 0 | 0.142432 | 0.0446304 | 0.033554943 |
| *APOL1* | rs136147 | 22 | T | 0.969697 | 0.511628 | 29.391042 | $6.39x10^{-06}$ | 0.00063313 |
| *APOL1* | rs4820224 | 22 | G | 1 | 0.816327 | NE | NE | NE |
| *APOL1* | rs136150 | 22 | A | 1 | 0.976744 | Inf | 1 | 0.991603218 |
| *APOL1* | rs2413395 | 22 | G | 0.903226 | 0.836735 | NE | NE | NE |
| *APOL1* | rs136159 | 22 | C | 0.965517 | 0.897959 | NE | NE | NE |
| *APOL1* | rs136160 | 22 | C | 0.969697 | 0.836735 | 6.1370333 | 0.0773909 | 0.075015021 |
| *APOL1* | rs129423 | 22 | C | 0.933333 | 0.895833 | NE | NE | NE |
| *APOL1* | rs136161 | 22 | G | 0.896552 | 0.44898 | Inf | 0.1510788 | 0.991676037 |
| *APOL1* | rs713929 | 22 | G | 0.875 | 0.866667 | 1.0752076 | 1 | 0.364325208 |
| *APOL1* | rs73885319 | 22 | G | 0.5556 | 0.03704 | 98.56 | $1.614x10^{-14}$ | $7.203x10^{-07}$ |
| *APOL1* | rs60910145 | 22 | G | 0.4872 | 0.03704 | 46.08 | $4.313x10^{-13}$ | $2.832x10^{-07}$ |
| *APOL1* | rs71785313 | 22 | D | 0.3421 | 0.1667 | 2.076 | 0.006014 | 0.00672 |
| *APOL1* | rs3886200 | 22 | C | NE | 0.44 | 0 | 1 | 1 |
| *APOL1* | rs2012928 | 22 | A | 0.655172 | 0.612245 | 1.2004938 | 0.8101471 | 0.701915321 |
| *MYH9* | rs2481 | 22 | C | 0.787879 | 0.708333 | 0.6841826 | 1 | 0.661499921 |
| *MYH9* | rs7078 | 22 | C | 0.090909 | 0 | 0.5047766 | 0.2794321 | 0.096620573 |
| *MYH9* | rs12107 | 22 | C | 0.851852 | 0.708333 | 0.557856 | 1 | 0.581700941 |
| *MYH9* | rs735854 | 22 | T | 1 | 0.979167 | Inf | 1 | 0.992104142 |
| *MYH9* | rs735853 | 22 | C | 0.965517 | 0.897959 | Inf | 1 | NE |
| *MYH9* | rs762964 | 22 | A | 0.928571 | 0.829787 | Inf | 0.5257658 | 0.992724757 |
| *MYH9* | rs5756129 | 22 | C | 0.030303 | 0.020408 | 0.4310823 | 0.1025239 | 0.06476053 |
| *MYH9* | rs6000226 | 22 | C | 0.482759 | 0.44898 | 5.3727725 | 0.1422447 | 0.115269085 |
| *MYH9* | rs6000229 | 22 | C | 0.76 | 0.520833 | 0.5156936 | 1 | 0.81236158 |
| *MYH9* | rs6000233 | 22 | T | 0.90625 | 0.77551 | 2.7661804 | 0.1476288 | 0.201075813 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *MYH9* | rs2071730 | 22 | G | 0.733333 | 0.55102 | 0.6082172 | 1 | 0.869411768 |
| *MYH9* | rs16996648 | 22 | C | 0.225806 | 0 | 3.8114543 | 0.0099959 | 0.005596493 |
| *MYH9* | rs9610486 | 22 | G | 0.833333 | 0.791667 | NE | NE | NE |
| *MYH9* | rs5756133 | 22 | T | 1 | 1 | NE | NE | NE |
| *MYH9* | rs4821480 | 22 | G | 0.666667 | 0.458333 | 0.6841826 | 1 | 0.944956866 |
| *MYH9* | rs4821481 | 22 | C | 0.636364 | 0.44898 | 0.6700763 | 1 | 0.954693465 |
| *MYH9* | rs17806513 | 22 | G | 0.862069 | 0.612245 | NE | NE | NE |
| *MYH9* | rs8137674 | 22 | A | 1 | 0.979592 | Inf | 1 | 0.991774077 |
| *MYH9* | rs1009150 | 22 | C | 0.8125 | 0.708333 | NE | NE | NE |
| *MYH9* | rs5750248 | 22 | T | 0.848485 | 0.77551 | 1.6118157 | 0.5716498 | 0.409184888 |
| *MYH9* | rs2157256 | 22 | A | 0.909091 | 0.895833 | 1.1606395 | 1 | 0.548387639 |
| *MYH9* | rs2157257 | 22 | A | 1 | 1 | NE | NE | NE |
| *MYH9* | rs5750250 | 22 | G | 0.533333 | 0.229167 | 1.8430219 | 0.4088952 | 0.274705633 |
| *MYH9* | rs3752462 | 22 | T | 0.967742 | 0.938776 | 1.9416199 | 1 | 0.432859764 |
| *MYH9* | rs5756152 | 22 | A | 0.75 | 0.354167 | 5.3389163 | 0.001739 | 0.002709614 |
| *MYH9* | rs2239784 | 22 | T | 0.96875 | 0.938776 | 2.0058091 | 1 | 0.272021858 |
| *MYH9* | rs1005570 | 22 | A | 0.875 | 0.693878 | 3.0483653 | 0.0674624 | 0.065611096 |
| *MYH9* | rs6000239 | 22 | C | 0.516129 | 0.326531 | 2.9894431 | 0.1426529 | 0.103506983 |
| *MYH9* | rs2071731 | 22 | G | 0.5625 | 0.489362 | 2.0955011 | 0.6431443 | 0.549690327 |
| *MYH9* | rs5756153 | 22 | G | 0.709677 | 0.673469 | NE | NE | NE |
| *MYH9* | rs16996668 | 22 | C | 0.21875 | 0.020408 | 5.1523472 | 0.0006835 | 0.001051309 |
| *MYH9* | rs5756154 | 22 | C | 0.677419 | 0.469388 | Inf | 0.2789557 | 0.98973063 |
| *MYH9* | rs5756156 | 22 | C | 0.6875 | 0.658537 | 0.5190915 | 0.370336 | 0.367900459 |
| *MYH9* | rs16996672 | 22 | T | 0.727273 | 0.416667 | 3.6702252 | 0.0069864 | 0.010119483 |
| *MYH9* | rs16996674 | 22 | T | 0.727273 | 0.276596 | 6.7800858 | 9.40E-05 | 0.000284294 |
| *MYH9* | rs16996677 | 22 | A | 0.733333 | 0.387755 | 4.2571797 | 0.0049768 | 0.008398059 |
| *MYH9* | rs6000240 | 22 | T | 1 | 0.959184 | Inf | 0.5130985 | 0.992953397 |
| *MYH9* | rs5750252 | 22 | G | 0.46875 | 0.408163 | 2.8925731 | 0.3010376 | 0.213139756 |
| *MYH9* | rs713659 | 22 | G | 0 | 0 | 0.6651641 | 0.4721014 | 0.275199801 |
| *MYH9* | rs6000243 | 22 | T | 1 | 1 | NE | NE | NE |
| *MYH9* | rs739096 | 22 | G | 0.966667 | 0.9375 | NE | NE | NE |
| *AIM* | rs2112527 | 19 | A | 0.9354839 | 0.8979591 | 1.63804300 | 0.700199 | 0.419756567 |
| *AIM* | rs679047 | 9 | T | 0.53125 | 0.4897959 | 1.32874478 | 1 | 0.690300168 |
| *AIM* | rs714857 | 11 | T | 1 | 0.9591836 | Inf | 0.526807 | 0.992351574 |
| *AIM* | rs1876482 | 2 | T | 0 | 0 | NE | NE | NE |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *AIM* | rs1344870 | 3 | NE | NE | NE | NE | NE | NE |
| *AIM* | rs65264 | 7 | C | 0.0606061 | 0.0408163 | 1.07467148 | 1 | 0.965648373 |
| *AIM* | rs735612 | 15 | T | 0.7575758 | 0.7083333 | 1.28280185 | 0.799926 | 0.736774276 |
| *AIM* | rs2077559 | 9 | C | 0.1515152 | 0.0816326 | 0.41629809 | 0.07165 | 0.094880304 |
| *AIM* | rs2089740 | 15 | T | 0.6333333 | 0.5714285 | 1.29122629 | 0.642215 | 0.513362392 |
| *AIM* | rs1858465 | 17 | A | 0.84375 | 0.75 | 1.34297093 | 1 | 0.637235783 |
| *AIM* | rs1823718 | 15 | C | 0.0967742 | 0 | 1.41200771 | 0.489499 | 0.35312344 |
| *AIM* | rs217538 | 6 | G | 0.0666667 | 0.0204081 | 3.37343353 | 0.553864 | 0.127939402 |
| *AIM* | rs953386 | 13 | A | 0.71875 | 0.6734693 | 1.23580550 | 0.806632 | 0.669459101 |
| *AIM* | rs1154997 | 2 | C | 0.5333333 | 0.3541666 | 3.63111954 | 0.115486 | 0.19444374 |
| *AIM* | rs1363448 | 5 | C | 0 | 0 | 0.69943618 | 0.771709 | 0.488680133 |
| *AIM* | rs720096 | 4 | G | 0.96875 | 0.9375 | 2.03311459 | 1 | 0.619290544 |
| *AIM* | rs723854 | 1 | C | 0.6774194 | 0.6326530 | 0.61264889 | 0.672139 | 0.52094187 |
| *AIM* | rs952718 | 2 | A | 0.8787879 | 0.8571428 | 1.20558763 | 1 | 0.604778962 |

Allele frequencies and association obtained from linear regression in the dominant model. Adjusted for age and gender. NE: Not Estimated
SNPs: Single nucleotide polymorphisms. AIM: Ancestry informative marker.

**Table A 14.2: SNP Association with 39 HIV (+) CKD and 54 HIV positive without CKD**

| | | | | Risk Allele Frequency | | | *P-values* | |
|---|---|---|---|---|---|---|---|---|
| **Gene** | **SNP** | **Chromosome** | **Risk Allele** | **Cases** | **Controls** | **Odds Ratio** | **Unadjusted** | **Adjusted** |
| *APOL3* | rs11089781 | 22 | A | 0.135135 | 0.234043 | 0.515347036 | 0.279309515 | 0.043519 |
| *APOL1* | rs9610467 | 22 | G | 0.8 | 0.8125 | 0.923971622 | 1 | 0.563245 |
| *APOL1* | rs6000220 | 22 | C | 0.444444 | 0.469388 | 0.555095137 | 0.484690323 | 0.348586 |
| *APOL1* | rs7284919 | 22 | C | 0.027778 | 0 | 1.656292125 | 0.4305617 | 0.627986 |
| *APOL1* | rs136147 | 22 | T | 0.441176 | 0.511628 | 0.756378713 | 0.647259246 | 0.670825 |
| *APOL1* | rs4820224 | 22 | G | 0.885714 | 0.816327 | NE | NE | NE |
| *APOL1* | rs136150 | 22 | A | 1 | 0.976744 | Inf | 1 | 0.991569 |
| *APOL1* | rs2413395 | 22 | G | 0.891892 | 0.836735 | NE | NE | NE |
| *APOL1* | rs136159 | 22 | C | 1 | 0.897959 | NE | NE | NE |
| *APOL1* | rs136160 | 22 | C | 0.702703 | 0.836735 | 0.46549816 | 0.190008735 | 0.193475 |
| *APOL1* | rs129423 | 22 | C | 1 | 0.895833 | NE | NE | NE |
| *APOL1* | rs136161 | 22 | G | 0.558824 | 0.44898 | 1.805902908 | 0.694631302 | 0.425706 |
| *APOL1* | rs713929 | 22 | G | 1 | 0.866667 | Inf | 0.134615385 | 0.994515 |
| *APOL1* | rs73885319 | 22 | G | 0.01098 | 0.03704 | 2.774 | 0.0341 | 0.1252 |
| *APOL1* | rs60910145 | 22 | G | 0.1163 | 0.3704 | 2.617 | 0.04925 | 0.1346 |
| *APOL1* | rs71785313 | 22 | D | 0.1744 | 0.1667 | 0.9659 | 0.8865 | 0.9282 |
| *APOL1* | rs3886200 | 22 | C | 0.583333 | 0.44 | 1.484647797 | 1 | 0.894736 |
| *APOL1* | rs2012928 | 22 | A | 0.628571 | 0.612245 | 1.07088609 | 1 | 0.597221 |
| *MYH9* | rs2481 | 22 | C | 0.621622 | 0.708333 | 0.768397947 | 1 | 0.987484 |
| *MYH9* | rs7078 | 22 | C | 0 | 0 | 0.609978972 | 0.427583279 | 0.31745 |
| *MYH9* | rs12107 | 22 | C | 0.611111 | 0.708333 | 0.747343939 | 1 | 0.96699 |
| *MYH9* | rs735854 | 22 | T | 1 | 0.979167 | Inf | 1 | 0.992055 |
| *MYH9* | rs735853 | 22 | C | 0.909091 | 0.897959 | Inf | 1 | NE |
| *MYH9* | rs762964 | 22 | A | 0.805556 | 0.829787 | 0.758150261 | 1 | 0.471028 |
| *MYH9* | rs5756129 | 22 | C | 0.054054 | 0.020408 | 0.643337143 | 0.375261187 | 0.296386 |
| *MYH9* | rs6000226 | 22 | C | 0.333333 | 0.44898 | 0.9758842 | 1 | 0.748379 |
| *MYH9* | rs6000229 | 22 | C | 0.666667 | 0.520833 | 0.237983633 | 0.308994398 | 0.069931 |
| *MYH9* | rs6000233 | 22 | T | 0.638889 | 0.77551 | 0.516311281 | 0.223567211 | 0.171956 |
| *MYH9* | rs2071730 | 22 | G | 0.675676 | 0.55102 | 0.239962767 | 0.310164324 | 0.065378 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *MYH9* | rs16996648 | 22 | C | 0.029412 | 0 | 1.064945985 | 1 | 0.976029 |
| *MYH9* | rs9610486 | 22 | G | 0.777778 | 0.791667 | 0 | 0.428571429 | 0.990452 |
| *MYH9* | rs5756133 | 22 | T | 1 | 1 | NE | NE | NE |
| *MYH9* | rs4821480 | 22 | G | 0.675676 | 0.458333 | 0.245041061 | 0.312914705 | 0.094839 |
| *MYH9* | rs4821481 | 22 | C | 0.675676 | 0.44898 | 0.239962767 | 0.310164324 | 0.090678 |
| *MYH9* | rs17806513 | 22 | G | 0.594595 | 0.612245 | 0 | 0.430232558 | 0.991522 |
| *MYH9* | rs8137674 | 22 | A | 1 | 0.979592 | Inf | 1 | 0.991722 |
| *MYH9* | rs1009150 | 22 | C | 0.756757 | 0.708333 | 0 | 0.435294118 | 0.990462 |
| *MYH9* | rs5750248 | 22 | T | 0.864865 | 0.77551 | 1.839794813 | 0.403438351 | 0.548518 |
| *MYH9* | rs2157256 | 22 | A | 0.918919 | 0.895833 | 1.313630873 | 1 | 0.896609 |
| *MYH9* | rs2157257 | 22 | G | 0.027027 | 0 | 1.27575129 | 0.800364917 | 0.935333 |
| *MYH9* | rs5750250 | 22 | G | 0.388889 | 0.229167 | 1.295977537 | 0.799542245 | 0.671538 |
| *MYH9* | rs3752462 | 22 | T | 0.972222 | 0.938776 | 2.262546892 | 0.634094484 | 0.618366 |
| *MYH9* | rs5756152 | 22 | A | 0.352941 | 0.354167 | 0.994717187 | 1 | 0.811836 |
| *MYH9* | rs2239784 | 22 | T | 0.891892 | 0.938776 | 0.542021318 | 0.457258098 | 0.650449 |
| *MYH9* | rs1005570 | 22 | A | 0.833333 | 0.693878 | 2.186037761 | 0.203494827 | 0.194045 |
| *MYH9* | rs6000239 | 22 | C | 0.277778 | 0.326531 | 0.973275287 | 1 | 0.978193 |
| *MYH9* | rs2071731 | 22 | G | 0.432432 | 0.489362 | 1.190702621 | 1 | 0.735224 |
| *MYH9* | rs5756153 | 22 | G | 0.675676 | 0.673469 | NE | NE | NE |
| *MYH9* | rs16996668 | 22 | C | 0.054054 | 0.020408 | 1.400597356 | 0.501094198 | 0.906194 |
| *MYH9* | rs5756154 | 22 | C | 0.527778 | 0.469388 | 2.262546892 | 0.634094484 | 0.415503 |
| *MYH9* | rs5756156 | 22 | C | 0.666667 | 0.658537 | 0.713394727 | 0.761681622 | 0.554936 |
| *MYH9* | rs16996672 | 22 | T | 0.486486 | 0.416667 | 1.321873954 | 0.660274142 | 0.948455 |
| *MYH9* | rs16996674 | 22 | T | 0.405405 | 0.276596 | 1.770760243 | 0.248872667 | 0.622479 |
| *MYH9* | rs16996677 | 22 | A | 0.441176 | 0.387755 | 1.243201536 | 0.655901584 | 0.858469 |
| *MYH9* | rs6000240 | 22 | T | 1 | 0.959184 | Inf | 0.503967168 | 0.992902 |
| *MYH9* | rs5750252 | 22 | G | 0.378378 | 0.408163 | 1.008035568 | 1 | 0.84776 |
| *MYH9* | rs713659 | 22 | G | 0 | 0 | 0.848732189 | 0.818610295 | 0.935214 |
| *MYH9* | rs6000243 | 22 | T | 1 | 1 | NE | NE | NE |
| *MYH9* | rs739096 | 22 | G | 0.861111 | 0.9375 | NE | NE | NE |
| *AIM* | rs2112527 | 19 | A | 0.875 | 0.89795914 | 0.797754 | 0.73414 | 0.79762069 |
| *AIM* | rs679047 | 9 | T | 0.52777777 | 0.48979591 | 0.978029 | 1 | 0.98940372 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *AIM* | rs714857 | 11 | T | 0.93548387 | 0.95918367 | 0.620916 | 0.638878 | 0.89175711 |
| *AIM* | rs1876482 | 2 | T | 0 | 0 | NE | NE | NE |
| *AIM* | rs1344870 | 3 | NE | NE | NE | NE | NE | NE |
| *AIM* | rs65264 | 7 | C | 0 | 0.04081632 | 1.143959 | 0.822624 | 0.75086515 |
| *AIM* | rs735612 | 15 | T | 0.81081081 | 0.70833333 | 1.753132 | 0.319335 | 0.27147919 |
| *AIM* | rs2077559 | 9 | C | 0.13513513 | 0.08163265 | 0.603612 | 0.278259 | 0.22247409 |
| *AIM* | rs1823718 | 15 | C | 0.02857142 | 0 | 0.791571 | 0.648878 | 0.86657293 |
| *AIM* | rs217538 | 6 | G | 0.02702702 | 0.02040816 | 1.328807 | 1 | 0.80132369 |
| *AIM* | rs953386 | 13 | A | 0.88888888 | 0.67346938 | 3.821057 | 0.022582 | 0.03267542 |
| *AIM* | rs1154997 | 2 | C | 0.58333333 | 0.35416666 | 2.087529 | 0.375535 | 0.39813565 |
| *AIM* | rs1363448 | 5 | C | 0 | 0 | 0.476673 | 0.376789 | 0.26112323 |
| *AIM* | rs720096 | 4 | G | 1 | 0.9375 | Inf | 0.390244 | 0.99423081 |
| *AIM* | rs723854 | 1 | C | 0.8 | 0.63265306 | 1.075171 | 1 | 0.93973825 |
| *AIM* | rs952718 | 2 | A | 0.91891891 | 0.85714285 | 1.875616 | 0.504429 | 0.36076108 |

Allele frequencies and association obtained from linear regression in the dominant model. Adjusted for age and gender. NE: Not Estimated
SNPs: Single nucleotide polymorphisms. AIM: Ancestry informative marker.

**Table A 14.3: SNP Association with 39 HIV (-) CKD and 54 Population Controls.**

| | | | | Risk Allele Frequency | | | *P-values* | |
|---|---|---|---|---|---|---|---|---|
| **Gene** | **SNP** | **Chromosome** | **Risk Allele** | **Cases** | **Controls** | **Odds Ratio** | **Unadjusted** | **Adjusted** |
| *APOL3* | rs11089781 | 22 | A | 0.137931 | 0.304347826 | 0.370349 | 0.164144 | 1 |
| *APOL1* | rs9610467 | 22 | G | 0.8518519 | 0.739130435 | 2.01088 | 0.381231 | 1 |
| *APOL1* | rs6000220 | 22 | C | 0.5862069 | 0.456521739 | 0.305017 | 0.555498 | 1 |
| *APOL1* | rs7284919 | 22 | C | 0 | 0 | 2.026196 | 0.263 | 1 |
| *APOL1* | rs136147 | 22 | T | 0.625 | 0.583333333 | 1.187609 | 0.802531 | 1 |
| *APOL1* | rs4820224 | 22 | G | 0.6785714 | 0.782608696 | 2.543105 | 0.643794 | 1 |
| *APOL1* | rs136150 | 22 | A | 0.9090909 | 0.857142857 | 1.65429 | 0.704026 | 1 |
| *APOL1* | rs2413395 | 22 | G | 0.7241379 | 0.790697674 | Inf | 0.512128 | NE |
| *APOL1* | rs136159 | 22 | C | 0.6896552 | 0.957446809 | 0 | 0.381579 | 1 |
| *APOL1* | rs136160 | 22 | C | 0.6896552 | 0.70212766 | 0.943495 | 1 | 1 |
| *APOL1* | rs129423 | 22 | C | 0.7142857 | 0.957446809 | 0 | 0.373333 | 1 |
| *APOL1* | rs136161 | 22 | G | 0.3214286 | 0.375 | 0.660947 | 0.520923 | 1 |
| *APOL1* | rs713929 | 22 | G | NE | 1 | NE | NE | NE |
| *APOL1* | rs73885319 | 22 | G | 0.09091 | 0.06481 | 1.41 | 0.4943 | 0.5289 |
| *APOL1* | rs60910145 | 22 | G | 0.06522 | 0.05556 | 1.153 | 0.7743 | 0.8204 |
| *APOL1* | rs71785313 | 22 | D | 0.0889 | 0.1111 | 0.7481 | 0.6054 | 0.5584 |
| *APOL1* | rs3886200 | 22 | C | 0.25 | 0.238095238 | 0 | 0.487805 | 1 |
| *APOL1* | rs2012928 | 22 | A | 0.4666667 | 0.583333333 | 0.628841 | 0.356599 | 1 |
| *MYH9* | rs2481 | 22 | C | 0.7307692 | 0.636363636 | 0 | 0.371429 | 1 |
| *MYH9* | rs7078 | 22 | C | 0 | 0.026315789 | 0.626995 | 0.542164 | 1 |
| *MYH9* | rs12107 | 22 | C | 0.7407407 | 0.625 | 0 | 0.36 | 1 |
| *MYH9* | rs735854 | 22 | T | 0.9032258 | 0.869565217 | 1.394069 | 0.732643 | 1 |
| *MYH9* | rs735853 | 22 | C | 0.7419355 | 0.869565217 | Inf | 0.512645 | NE |
| *MYH9* | rs762964 | 22 | A | 0.8076923 | 0.638297872 | Inf | 0.082591 | NE |
| *MYH9* | rs5756129 | 22 | C | 0.0645161 | 0.0625 | 0.653558 | 0.483986 | 1 |
| *MYH9* | rs6000226 | 22 | C | 0.3870968 | 0.5 | 0.599322 | 0.523846 | 1 |
| *MYH9* | rs6000229 | 22 | C | 0.6206897 | 0.489361702 | 0.575857 | 0.492069 | 1 |
| *MYH9* | rs6000233 | 22 | T | 0.6206897 | 0.770833333 | 0.491278 | 0.196413 | 1 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *MYH9* | rs2071730 | 22 | G | 0.6129032 | 0.489361702 | 0.623008 | 0.506837 | 1 |
| *MYH9* | rs16996648 | 22 | C | 0.0384615 | 0.111111111 | 0.46552 | 0.205073 | 1 |
| *MYH9* | rs9610486 | 22 | G | 0.7419355 | 0.680851064 | Inf | 0.514819 | NE |
| *MYH9* | rs5756133 | 22 | T | 1 | 0.958333333 | Inf | 0.51704 | NE |
| *MYH9* | rs4821480 | 22 | G | 0.6206897 | 0.520833333 | 0.689213 | 0.738275 | 1 |
| *MYH9* | rs4821481 | 22 | C | 0.6 | 0.510638298 | 0.589628 | 0.52162 | 1 |
| *MYH9* | rs17806513 | 22 | G | 0.8333333 | 0.5625 | 1.918489 | 1 | 1 |
| *MYH9* | rs8137674 | 22 | G | 0 | 0 | 1.641978 | 0.63496 | 1 |
| *MYH9* | rs1009150 | 22 | C | 0.6333333 | 0.723404255 | Inf | 0.518113 | NE |
| *MYH9* | rs5750248 | 22 | T | 0.8 | 0.833333333 | 0.80234 | 0.766542 | 1 |
| *MYH9* | rs2157256 | 22 | A | 0.962963 | 0.958333333 | 1.12864 | 1 | 1 |
| *MYH9* | rs2157257 | 22 | A | 1 | 0.957446809 | Inf | 0.518113 | NE |
| *MYH9* | rs5750250 | 22 | G | 0.3548387 | 0.282608696 | 0.519241 | 0.26929 | 1 |
| *MYH9* | rs3752462 | 22 | T | 0.9677419 | 0.957446809 | 1.328582 | 1 | 1 |
| *MYH9* | rs5756152 | 22 | A | 0.4285714 | 0.382978723 | 1.205261 | 0.808482 | 1 |
| *MYH9* | rs2239784 | 22 | T | 0.7586207 | 0.8 | 0.788329 | 0.774696 | 1 |
| *MYH9* | rs1005570 | 22 | A | 0.7419355 | 0.75 | 0.95888 | 1 | 1 |
| *MYH9* | rs6000239 | 22 | C | 0.4482759 | 0.234042553 | 0.648568 | 0.555666 | 1 |
| *MYH9* | rs2071731 | 22 | G | 0.6 | 0.5 | 0 | 0.062631 | 1 |
| *MYH9* | rs5756153 | 22 | G | 0.6774194 | 0.711111111 | 1.035241 | 1 | 1 |
| *MYH9* | rs16996668 | 22 | C | 0.0967742 | 0.020833333 | 0.952953 | 1 | 1 |
| *MYH9* | rs5756154 | 22 | C | 0.5 | 0.5625 | 0.4594 | 0.293851 | 1 |
| *MYH9* | rs5756156 | 22 | C | 0.64 | 0.7 | 4.990404 | 0.138971 | 1 |
| *MYH9* | rs16996672 | 22 | T | 0.4 | 0.395833333 | 1.017322 | 1 | 1 |
| *MYH9* | rs16996674 | 22 | T | 0.3333333 | 0.326086957 | 1.032891 | 1 | 1 |
| *MYH9* | rs16996677 | 22 | A | 0.4333333 | 0.404255319 | 1.12518 | 0.816933 | 1 |
| *MYH9* | rs6000240 | 22 | T | 1 | 0.979166667 | Inf | 1 | NE |
| *MYH9* | rs5750252 | 22 | G | 0.516129 | 0.291666667 | 0.454773 | 0.423594 | 1 |
| *MYH9* | rs713659 | 22 | G | 0.0333333 | 0.023809524 | 1.113686 | 1 | 1 |
| *MYH9* | rs6000243 | 22 | T | 0.9333333 | 0.914893617 | 1.297989 | 1 | 1 |
| *MYH9* | rs739096 | 22 | G | 0.8387097 | 0.847826087 | Inf | 0.512645 | NE |
| *AIM* | rs2112527 | 19 | A | 0.884615 | 0.717391 | 2.978623 | 0.142404 | 1 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| *AIM* | rs679047 | 9 | T | 0.6 | 0.520833 | 1.618323 | 0.700975 | 1 |
| *AIM* | rs714857 | 11 | T | 0.866667 | 1 | 0 | 0.021361 | 1 |
| *AIM* | rs1876482 | 2 | T | 0.064516 | 0 | Inf | 0.150925 | 1 |
| *AIM* | rs1344870 | 3 | NE | NE | NE | NE | NE | NE |
| *AIM* | rs65264 | 7 | C | 0.032258 | 0.191489 | 1.064463 | 1 | 1 |
| *AIM* | rs735612 | 15 | T | 0.724138 | 0.531915 | 2.284784 | 0.146762 | 1 |
| *AIM* | rs2077559 | 9 | C | 0.1 | 0.083333 | 1.123297 | 0.819198 | 1 |
| *AIM* | rs2089740 | 15 | T | 0.733333 | 0.595745 | 1.851219 | 0.234089 | 1 |
| *AIM* | rs1858465 | 17 | A | 0.5 | 0.76087 | 0.354008 | 0.344153 | 1 |
| *AIM* | rs1823718 | 15 | C | 0.066667 | 0.041667 | 0.441581 | 0.10409 | 1 |
| *AIM* | rs217538 | 6 | G | 0.129032 | 0.06383 | 2.150478 | 0.426537 | 1 |
| *AIM* | rs953386 | 13 | A | 0.612903 | 0.702128 | 0.675237 | 0.466939 | 1 |
| *AIM* | rs1154997 | 2 | C | 0.37931 | 0.659574 | 0.282751 | 0.19426 | 1 |
| *AIM* | rs1363448 | 5 | C | 0.064516 | 0.041667 | 3.757031 | 0.052881 | 1 |
| *AIM* | rs720096 | 4 | C | NE | 0 | 0 | 1 | NE |
| *AIM* | rs723854 | 1 | C | NE | NE | NE | NE | NE |
| *AIM* | rs952718 | 2 | A | NE | NE | NE | NE | NE |

Allele frequencies and association obtained from linear regression in the dominant model. Adjusted for age and gender. NE: Not Estimated. SNPs: Single nucleotide polymorphisms. AIM: Ancestry informative marker.

# APPENDIX 15

## A) Permission to use for Figure 1.5

**NATURE PUBLISHING GROUP LICENSE TERMS AND CONDITIONS**

Mar 31, 2015

---

This is a License Agreement between Alex N Kasembeli ("You") and Nature Publishing Group ("Nature Publishing Group") provided by Copyright Clearance Center ("CCC"). Tl license consists of your order details, the terms and conditions provided by Nature Publishing Group, and the payment terms and conditions.

**All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.**

License Number

3597160727936

License date

Mar 27, 2015

Licensed content publisher

Nature Publishing Group

Licensed content publication

Nature Reviews Nephrology

Licensed content title

The population genetics of chronic kidney disease: insights from the MYH9-APOL1 locus

Licensed content author

Saharon Rosset, Shay Tzur, Doron M. Behar, Walter G. Wasser and Karl Skorecki

Licensed content date

May 3, 2011

Volume number

7

Issue number

6

Type of Use

reuse in a dissertation / thesis

Requestor type

academic/educational

Format

print and electronic

Portion

figures/tables/illustrations

Number of figures/tables/illustrations

1

High-res required

no

Figures

Figure 1 | The effect of recombination on linkage disequilibrium and haplotype blocks.

Author of this NPG article

no

Your reference number

None

Title of your thesis / dissertation

Genetic and molecular characterization of HIV-associated kidney disease in black South Africans

Expected completion date

Mar 2015

Estimated size (number of pages)

210

Total

0.00 USD

Terms and Conditions

Terms and Conditions for Permissions

B) **Permission to use for Figure 1.7**

**NATURE PUBLISHING GROUP LICENSE TERMS AND CONDITIONS**

Mar 31, 2015

---

This is a License Agreement between Alex N Kasembeli ("You") and Nature Publishing Group ("Nature Publishing Group") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Nature Publishing Group, and the payment terms and conditions.

**All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.**

| | |
|---|---|
| License Number | 3597020054117 |
| License date | Mar 27, 2015 |
| Licensed content publisher | Nature Publishing Group |
| Licensed content publication | Nature Reviews Nephrology |
| Licensed content title | The population genetics of chronic kidney disease: insights from the MYH9-APOL1 locus |
| Licensed content author | Saharon Rosset, Shay Tzur, Doron M. Behar, Walter G. Wasser and Karl Skorecki |
| Licensed content date | May 3, 2011 |
| Volume number | 7 |
| Issue number | 6 |
| Type of Use | reuse in a dissertation / thesis |
| Requestor type | academic/educational |
| Format | print and electronic |
| Portion | figures/tables/illustrations |
| Number of figures/tables/illustrations | 1 |
| High-res required | no |
| Figures | Figure 3 \| Chromosomal region associated with diabetes-unrelated ESRD phenotypes in African Americans. b \| These studies highlighted a region of particularly strong associations with ESRD containing more than 20 genes, including MYH9 and APOL1. |
| Author of this NPG article | no |
| Your reference number | None |
| Title of your thesis / dissertation | Genetic and molecular characterization of HIV-associated kidney disease in black South Africans |
| Expected completion date | Mar 2015 |
| Estimated size (number of pages) | 210 |
| Total | 0.00 USD |
| Terms and Conditions | |

C) **Permission to use for Figure 1.15**

**NATURE PUBLISHING GROUP LICENSE TERMS AND CONDITIONS**

Mar 31, 2015

---

This is a License Agreement between Alex N Kasembeli ("You") and Nature Publishing Group ("Nature Publishing Group") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Nature Publishing Group, and the payment terms and conditions.

**All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.**

License Number

3597021203243

License date

Mar 27, 2015

Licensed content publisher

Nature Publishing Group

Licensed content publication

Nature Reviews Microbiology

Licensed content title

The trypanolytic factor of human serum

Licensed content author

Etienne Pays, Benoit Vanhollebeke, Luc Vanhamme, Françoise Paturiaux-Hanocq, Derek P. Nolan et al.

Licensed content date

Jun 1, 2006

Volume number

4

Issue number

6

Type of Use

reuse in a dissertation / thesis

Requestor type

academic/educational

Format

print and electronic

Portion

figures/tables/illustrations

Number of figures/tables/illustrations

1

High-res required

no

Figures

Figure 5 | Model of the mechanism of trypanolysis by apolipoprotein L1 (APOL1).

Author of this NPG article

no

Your reference number

None

Title of your thesis / dissertation

Genetic and molecular characterization of HIV-associated kidney disease in black South Africans

Expected completion date

Mar 2015

Estimated size (number of pages)

210

Total

0.00 USD

Terms and Conditions

Terms and Conditions for Permissions

D) **Permission to use for Figure 1.17**

**Customer Information**

**Customer:** Alex Kasembeli
**Account Number:** 3000903207
**Organization:** Alex Kasembeli
**Email:** alekasesh@yahoo.com
**Phone:** +27 791674847

Search order details by:

**This is not an invoice**

**Order Details**

›urnal of the American Society of Nephrology

Billing Status:
**N/A**

**Order detail ID:**66590871
**ISSN:**1533-3450
**Publication Type:**e-Journal
**Volume:**
**Issue:**
**Start page:**
**Publisher:**AMERICAN SOCIETY OF NEPHROLOGY
**Author/Editor:**American Society of Nephrology
**Permission Status:** **Granted**

**Permission type:**Republish or display content
**Type of use:**Thesis/Dissertation

**Order License Id:** 3598701179344

Hide details

| | |
|---|---|
| **Requestor type** | Author of requested content |
| **Format** | Print, Electronic |
| **Portion** | chart/graph/table/figure |
| **Number of charts/graphs/tables/figures** | 1 |
| **Title or numeric reference of the portion(s)** | Figure 5. Worldwide frequency distribution of APOL1 variants G1 and G2. Genotypes of G1 and G2 are determined for 51 populations in the Centre dâ□□Etude du Polymorphisme Humain (CEPH) HGDP, for the HapMap Luhya population from Kenya (International HapMap Project [Phase II]), and for African Americans (AA) and European American (EA) controls in the NIH FSGS study cohort. The allele frequencies of G1 (red), |

| | |
|---|---|
| | G2 (orange), and wild-type alleles (light blue) in each population are depicted in pie charts overlaid upon a world map. |
| **Title of the article or chapter the portion is from** | APOL1 Genetic Variants in Focal Segmental Glomerulosclerosis and HIV-Associated Nephropathy |
| **Editor of portion(s)** | N/A |
| **Author of portion(s)** | N/A |
| **Volume of serial or monograph** | N/A |
| **Page range of portion** | 2135 |
| **Publication date of portion** | 2011 |
| **Rights for** | Main product |
| **Duration of use** | Life of current edition |
| **Creation of copies for the disabled** | no |
| **With minor editing privileges** | no |
| **For distribution to** | Worldwide |
| **In the following language(s)** | Original language of publication |
| **With incidental promotional use** | no |
| **Lifetime unit quantity of new product** | Up to 499 |
| **Made available in the following markets** | education |
| **The requesting person/organization** | Alex Nganga Kasembeli |
| **Order reference number** | |
| **Author/Editor** | Alex Nganga Kasembeli |
| **The standard identifier** | PhD Thesis |
| **Title** | Genetic and molecular characterization of HIV-associated kidney disease in black South Africans |
| **Publisher** | University of the Witwatersrand |
| **Expected publication date** | Mar 2015 |
| **Estimated size (pages)** | 215 |

**Note:** This item was invoiced separately through our **RightsLink service.** More info **$ 0.00**

**Total order items: 1**

**Order Total: $0.00**

E) **Permission to use for Figure 1.18**

**NATURE PUBLISHING GROUP LICENSE TERMS AND CONDITIONS**

Mar 31, 2015

---

This is a License Agreement between Alex N Kasembeli ("You") and Nature Publishing Group ("Nature Publishing Group") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Nature Publishing Group, and the payment terms and conditions.

**All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.**

| | |
|---|---|
| License Number | 3597190274425 |
| License date | Mar 27, 2015 |
| Licensed content publisher | Nature Publishing Group |
| Licensed content publication | Nature Reviews Nephrology |
| Licensed content title | The genetics of common kidney disease: a pathway toward clinical relevance |
| Licensed content author | Paul E. Drawz and John R. Sedor |
| Licensed content date | Aug 1, 2011 |
| Volume number | 7 |
| Issue number | 8 |
| Type of Use | reuse in a dissertation / thesis |
| Requestor type | non-commercial (non-profit) |
| Format | print and electronic |
| Portion | figures/tables/illustrations |
| Number of figures/tables/illustrations | 1 |
| High-res required | no |
| Figures | Figure 2 \| The relationship between APOL1 and MYH9 genetic variants associated with common forms of kidney disease. |
| Author of this NPG article | no |
| Your reference number | None |
| Title of your thesis / dissertation | Genetic and molecular characterization of HIV-associated kidney disease in black South Africans |
| Expected completion date | Mar 2015 |
| Estimated size (number of pages) | 210 |
| Total | 0.00 USD |
| Terms and Conditions | |

Terms and Conditions for Permissions

F) **Permission to use for Tables 3.1, 3.6, 3.8 and 3.9 and Figures 3.3, 3.5 and 3.7**

**Confirmation Number: 11328265**
**Order Date: 03/30/2015**

Print this page
Print terms & conditions
Print citation information(What's this?)

-

**Customer Information**

**Customer:** Alex Kasembeli
**Account Number:** 3000903207
**Organization:** Alex Kasembeli
**Email:** alekasesh@yahoo.com
**Phone:** +27 791674847

Search order details by:

**This is not an invoice**

**Order Details**

urnal of the American Society of Nephrology : JASN

Billing Status:
**N/A**

**Order detail ID:**66590891
**ISSN:**1046-6673
**Publication Type:**Journal
**Volume:**
**Issue:**
**Start page:**
**Publisher:**AMERICAN SOCIETY OF NEPHROLOGY
**Author/Editor:**AMERICAN SOCIETY OF NEPHROLOGY
**Permission Status:** **Granted**

**Permission type:**Republish or display content
**Type of use:**Thesis/Dissertation

**Order License Id:** 3598720072592
Hide details

| | |
|---|---|
| **Requestor type** | Author of requested content |
| **Format** | Print, Electronic |
| **Portion** | chapter/article |
| **Title or numeric reference of the portion(s)** | Table 1. Characteristics of study participants and controls, Table 2. APOL1 genotype distribution among patients with HIVAN, HIVpositive patients with CKD, HIV-negative patients with CKD, and controls, Table 3. Association between APOL1 risk alleles and various glomerular diseases, Table 4. APOL1 risk allele associations with HIVAN, HIV-associated FSGS, and HIV-negative |

| | |
|---|---|
| | FSGS (primary FSGS), Figure 1. Distribution of APOL1 haplotypes and risk alleles. Five APOL1 haplotypes are observed. The ancestral haplotype, denoted as G0, has three ancestral alleles. The G1GM haplotype has two missense alleles. The G1G+ haplotype has one missense risk allele at rs73885319. The G1+M has one missense variant at rs60910145. The G2 haplotype has the 6-bp deletion risk allele at rs71785313. The risk allele and haplotype frequencies are shown for patients with HIVAN, HIV-positive patients with CKD and HIV-negative patients with CKD, and HIV-positive controls and population controls (PCs) and Figure 2. ORs for the effect sizes for HIVAN with different combination of APOL1 G1 and G2 alleles. ORs with CIs for association of different strata of G1-G2 genotypes with HIVAN patients compared with participants carrying no G1 or G2 risk alleles (+/+). For G1/G1 and G1/G2, the OR is infinity, because these genotypes were only observed in patients. |
| **Title of the article or chapter the portion is from** | APOL1 Risk Variants Are Strongly Associated with HIV-Associated Nephropathy in Black South Africans |
| **Editor of portion(s)** | N/A |
| **Author of portion(s)** | Alex Nganga Kasembeli |
| **Volume of serial or monograph** | N/A |
| **Page range of portion** | |
| **Publication date of portion** | 2015 |
| **Rights for** | Main product |
| **Duration of use** | Life of current edition |
| **Creation of copies for the disabled** | no |
| **With minor editing privileges** | no |
| **For distribution to** | Worldwide |
| **In the following language(s)** | Original language of publication |
| **With incidental promotional use** | no |
| **Lifetime unit quantity of new product** | Up to 499 |
| **Made available in the following markets** | education |
| **The requesting person/organization** | Alex Nganga Kasembeli |
| **Order reference number** | |

| | |
|---|---|
| **Author/Editor** | Alex Nganga Kasembeli |
| **The standard identifier** | PhD Thesis |
| **Title** | Genetic and molecular characterization of HIV-associated kidney disease in black South Africans |
| **Publisher** | University of the Witwatersrand |
| **Expected publication date** | Mar 2015 |
| **Estimated size (pages)** | 215 |

**Note:** This item was invoiced separately through our **RightsLink service.** More info **$ 0.00**

**Total order items: 1**

**Order Total: $0.00**