VIRUS RESTRICTION GENE VARIANTS AND THEIR POSSIBLE ROLE IN NEUROCOGNITIVE FUNCTION IN CHILDREN BORN TO HIV-INFECTED MOTHERS

Kudakwashe Mhandire

Student number: MHNKUD004

Supervised by Dr Collet Dandara

Submitted for MSc (Med) in Human Genetics

Division of Human Genetics

Department of Clinical Laboratory Sciences

Faculty of Health Sciences

University of Cape Town

August 2012

PLAGIARISM DECLARATION

1. I know that plagiarism is wrong. Plagiarism is using another's work and to pretend that it is one's own.

2. I have used the Harvard British-Standard 2010 as the convention for citation and referencing. Each significant contribution to this dissertation from the work of other people has been attributed and has cited and referenced.

3. This project is my own work.

4. I have not allowed, and will not allow, anyone to copy my work with the intention of passing it off as his or her own work.

SIGNATURE: _Signed by candidate_____

DATE: _____

ABSTRACT

Introduction: Host genetic variation is an important determinant of HIV infection, disease progression and HIV-associated neurocognitive deficits. However, there is no sufficient knowledge on the role of genetic variants especially among African populations. This study is focused on investigating variation in HIV/AIDS restriction genes; *CCR2, CX3CR1, SDF1, RANTES, APOBEC3G* and *MBL2* and their possible role in HIV infection and neurocognitive function among children born to HIV infected mothers, recruited in Harare, Zimbabwe.

Methods: A total of 116 children comprising of 73 perinatally exposed to HIV (34 who were born infected and 39 who were uninfected) and 43 unexposed controls were recruited in 2011(at ages 7-9 years) from a cohort of mother-baby pairs that has been followed up since 2002. The demographic characteristics of the recruited children were captured from their medical records. A McCarthy Scale of Children's Abilities (MSCA) was administered to determine each child's neurocognitive status. Genotyping for allelic variants was done using PCR-RFLP, SNaPshot® and Sanger DNA sequencing. Statistical analysis was carried out to determine association between genotypes, HIV status and neurocognitive function.

Results: Baseline allele frequencies were established for 23 genetic variants in this Zimbabwean population and these included the discovery of a novel *MBL2* -595G>A SNP. The distribution of the following genotypes and genotype combinations were significantly different when compared between the HIV-infected and HIV-uninfected groups, both perinatally exposed to HIV; *CCR2* 190G/A (P=0.02), *CCR2* 190G/A-*CX3CR1* 745G/G (P=0.0002), *CCR2* 190G/A-*APOBEC3G* 557A/A (P=0.015), *MBL2* -221G/G-+4C/T (P=0.03) occurring in 15% v 39%, 0% v 33%, 0% v 17% and 44% v 22%, respectively. In addition, haplotypes were generated for *APOBEC3G*. The haplotype C-G-G-C-C with regard to -571G>C, -90C>G, 557G>A, 197T>C and 199C>G SNPs was present in 6% of HIV infected group but was not observed in the uninfected group (P=0.007). We also observed that the *MBL2* genotypes +4C/C (P=0.03) and -221C/G (P=0.02) were significantly associated with poor cognitive ability regardless of HIV status.

Conclusion: The observation of different genetic variants or combinations of genotypes between the HIV-exposed and infected group and that of the HIV-exposed but uninfected group may be a pointer to critical pathways in differential HIV susceptibility. Exposure and infection with HIV is controlled by a multitude of genes/processes, thus, SNPs are unlikely to show statistically significant effects individually and may be more useful in a multifactorial model, as observed from comparisons of genotype combinations and haplotypes. The role of host genetic variation on neurocognitive function remains disputed but our observations suggest innate immune factors such as *MBL2* may have a pronounced effect. Therefore, it may be possible to genotype for a suite of genes and use them as markers of either HIV susceptibility or neuro-developmental patterns.

I dedicate this work to my late parents and brother

IRENE, NOAH & JOHN MHANDIRE

ACKNOWLEDGEMENTS

I wish to express my heartfelt gratitude to the following:

- My Supervisor, Dr Collet Dandara for introducing me to the deep-end of Genetic and taking his time to bring the most out of me.
- My co-supervisors, Mrs K. Duri and Dr Nyasha Chin'ombe for unwavering support.
- Co-investigators, Dr G. Kandawasvika and P. Chandiwana for the clinical aspect of the study.
- My colleagues Gavin, Marelize, Russel and Alltalents for assisting with some of the work.
- Pharmacogenetics and Cancer Research Group members Collet, Nyasha, Michelle, Elizabeth, Luke, Naseeha, Marelize and Gavin, for advice, interaction and those highly informative Thursday afternoon sessions.
- Prof Raj and the entire Division of Human Genetics for allowing me to be a part of the great team.
- The Letten Foundation Research Group, Harare especially Mr W. Soko, Dr Kurewa, Dr N. Midzi, Auxillia and Phineas for their support in different but valid ways.
- University of Zimbabwe, DMLS, will always be my family
- Ministry of Health NIH lab member Dr Mtambu, Joseph, Munashe and Chris.
- My Cape Town friends Leah, Elizabeth, KK, Tavonga, Wisdom, Lebo, Shelly, Martha and 6G Obz Square for providing the little home away from home.
- My family Connie, Tate and Yewu (and their families), Praise, Mudhara Sam and other Mhandires for their relentless support over the years.
- To my love, friend, advisor and pillar of strength, Doreen for enduring the long months apart. I love you.
- Most of all to Prof Babil Stray-Pedersen and Letten Foundation, Norway for the funding and vision that has led to 10 successful years of research in Zimbabwe.

Table of Contents

List of Appendices

List of figures

List of tables

List of Abbreviations

˚C	Degrees celcius
A	Adenine
AIDS	Acquired immune-deficiency syndrome
AML	Acute Myeloid Leukaemia
AP	Alkaline phosphatase
APOBEC	Apolipoprotein B mRNA-editing, enzyme-catalytic
ARGs	AIDS Restriction Genes
ART	Anti-retroviral therapy
ARV	Anti-retroviral
ASN	African ancestry in Southwest USA
C	Cytosine
CCR2	C-C chemokine receptor type 2
CCR5	C-C chemokine receptor 5
CD4+	Cluster differentiation 4 marker
CD8+	Cluster differentiation 8 marker
cDNA	complementary Deoxy ribonucleic acid
CI	Confidence interval
CIS	Cognitive index score
CNS	Central nervous system
CRD	Carbohydrate recognition domain
CRF	Circulatory recombinant forms
CX3CR1	C-X3-C chemokine receptor 1
CXCR4	C-X-C chemokine receptor 4
D	Aspartic acid
dbSNP	Single Nucleotide Polymorphism database
ddNTPs	dideoxy-nucleotide triphosphates
DMSO	Dimethyl sulphoxide
DNA	Deoxy ribonucleic acid

dNTPs	deoxy-nucleotide triphosphates
E	Glutamic acid
EDTA	Ethylenediaminetetraacetic acid
EI	HIV-exposed infected
ELISA	Enzyme linked immuno-sorbernt assay
EtBr	Ethidium bromide
EU	HIV-exposed uninfected
G	Guanine/glycine
gp120	glycoprotein 120
-H	Hydrogen
HAD	HIV-associated Dementia
HAND	HIV-associated neurocognitive disorder
Hb	Haemoglobin
HCB	Han Chinese in Beijing
HIV	Human Immune-deficiency virus
I	Isoleucine
LD	Linkage disequilibrium
LWK	Luhya in Webuye, Kenya
MAC	Membrane attack complex
MASPs	MBL-associated serine proteases
MBL	Mannose Binding Lectin
MKK	Maasai in Kenyawa, Kenya
mM	millimolar
mRNA	messenger Ribonucleic acid
MSCA	McCarthy Scale of Children's Abilities
MTCT	Mother-to-child transmission
NCBI	National Centre for Biotechnology Information
NF-AT	Nuclear factor of activated T-cells
OD	Optical density
-OH	hydroxyl group

OR	Odds ratio
PCR	Polymerase chain reaction
RANTES	Regulated on Activation, Normal T-cell Expressed and Secreted
RFLP	Restriction Fragment Length Polymorphism
RNA	Ribonucleic acid
RT	Reverse Transcriptase
SDF1	Stromal cell-derived factor-1
sdNVP	single dose Nevirapine
SIV	Simian Immunodeficiency virus
SNPs	Single nucleotide polymorphism
SSA	Sub-Saharan Africa
T	Thymine/threonine
TBE	Tris-Borate EDTA
UEUI	Unexposed uninfected
UNAIDS	United Nations Programme on HIV/AIDS
UNGASS	United Nations General Assembly
URF	Unique recombinant forms
UV	Ultra-violet
V	Valine
WHO	World Health Organisation
YRI	Yomba of Ibadan Nigeria
ZDHS	Zimbabwe Demographic and Health Survey
µ	microlitre

1 CHAPTER ONE: BACKGROUND

1.1 Introduction

Almost three decades after its discovery, human immune-deficiency virus (HIV) remains a major health burden in the world. At the end of year 2011, the World Health Organisation (WHO) reported an estimated 34 million people to be living with HIV worldwide . The wide use of antiretroviral therapy (ART) has significantly lowered the number of deaths and morbidities related to HIV/AIDS. Despite this, the pandemic currently contributes almost two million deaths annually (WHO 2011). According to the Joint United Nations Programme on HIV/AIDS (UNAIDS 2010), sub-Saharan Africa (SSA) is the worst affected region with an estimated 70% of the world's HIV cases living in the region yet it carries only 10% of the world population . Figure 1.1 is a map showing the distribution of HIV cases across the world.

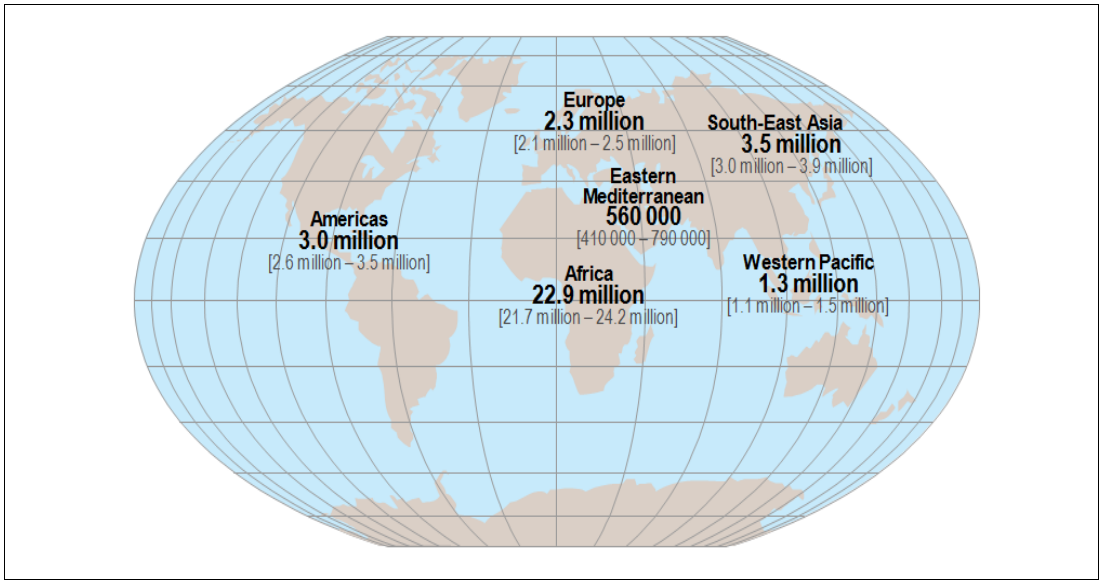


Figure 1.1: World map showing the distribution of HIV cases in WHO regions. Sub-Saharan Africa carries the highest HIV burden in the world (adapted from WHO. 2011).

An estimated 60% of the 22 million HIV cases in SSA are women (UNAIDS 2010). Most of the HIV-infected women fall within the reproductive age group of 15-49 years, thus, the number of children perinatally exposed to HIV is very high (UNAIDS 2010). This increases chances of mother-to-child transmission (MTCT) of HIV. Generally, the rate of HIV MTCT ranges from about 10-40% in the absence of intervention (De Cock et al. 2000). The use of ART, avoidance of breastfeeding and alternative delivery methods may reduce MTCT to almost 2% as witnessed in the developed world (Mofenson et al. 2002). However, these practices are largely impracticable in the developing world therefore MTCT remains high in SSA accounting for up to 90% of paediatric HIV infections. In 2010 alone, there were 400 000 infants who were born HIV-infected worldwide (UNAIDS 2010).

The countries worst affected by HIV in SSA are Swaziland, Botswana, Lesotho, South Africa, Namibia, Zambia and Zimbabwe. At least 10% of the population in each of these countries is HIV-infected with Swaziland leading the pack with a prevalence of more than 25% (UNAIDS 2010). Zimbabwe noted its first HIV case in 1985 and by the end of the 1980s, the prevalence had risen to about 10% (UNAIDS. 2005). This was the beginning of an escalation of the deadly disease in the country with figures reaching beyond 25% in 1997 as reported by the United Nations General Assembly Special Session (UNGASS 2010). The country has since witnessed a gradual drop in the HIV prevalence to 14% in 2010 according to the Zimbabwe Demographic and Healthy Survey 2010. More than 10% of pregnant women in Zimbabwe are HIV-positive and an estimated 15 000 children are consequently infected by HIV each year (United Nations Population Fund. 2010). It was recently shown that that even with single dose nevirapine prophylaxis, the rate of HIV MTCT is as high as 22% in Zimbabwe (Gumbo et al. 2010).

The introduction of ART has meant that HIV/AIDS patients now survive longer compared to the pre-ART time, thus the long term effects of living with the HIV on target organs are only starting to be understood now but there are many gaps. HIV has been shown to target the central nervous system (CNS) in its early stages of infection (Davis et al. 1992). This is thought to lead to neurological complications such as motor impairments and cognitive deficits at a later stage (Zhou et al. 2009). An estimated 30% of HIV infected individuals and 50% of those with AIDS have been reported to exhibit signs of neurologic disease (Heaton et al. 2004). A study among Americans reported that even in the presence of ART, HIV-associated neurocognitive impairments exist in up to 40% of HIV patients with increase as the disease progresses (Heaton et al. 2011). Neurological impairments interfere with cognitively demanding activities such as work, driving, operation of machinery and management of medication and also increase the risk of early mortality in HIV patients (Heaton et al. 2004, Hinkin et al. 2004).

The occurrence and extent of damage caused by neurological impairments and other HIV/AIDS related morbidities differs amongst individuals and populations. Since the early years of HIV/AIDS, it has been observed that disease progression varies among infected individuals (De Rossi et al. 1996). Some individuals may survive asymptomatically for up to 15 years from time of infection even without ART (long term non-progressors) whilst some deteriorate rapidly reaching the AIDS stage within months to two years (rapid progressors) (Pantaleo et al. 1996). However, about 70-80% of HIV infected people are intermediate progressors whose immune system declines at a steady rate reaching the symptomatic stage in about five to ten years (Pantaleo et al. 1996, Langford et al. 2007).

In addition, a fraction of the world's population seems to be resistant to HIV infection as observed in people who remained uninfected despite repeated exposure to HIV (Fowke et al. 1996, Tomescu et al. 2011, Kroner et al. 1994). Another study on patients with haemophilia A who were transfused with HIV contaminated factor VIII concentrates before the discovery of HIV reported that about 5% remained uninfected despite constant exposure to HIV (Kroner et al. 1994). These findings point to inter-individual differences that influence the host response to HIV exposure. Factors affecting the risk of HIV infection and/or disease progression include viral, immunological and host genetic variation (Langford et al. 2007). This study sets out to determine if inter-individual genetic variation plays a role in one's risk of HIV infection and neurocognitive function in children born to HIV-infected mothers. In order to understand the role of host genetics in fighting HIV, one needs to understand the structure and pathophysiology of HIV.

1.2 HIV structure

HIV is a member of the retroviridae family which are RNA viruses that require a reverse transcriptase enzyme to replicate (Barré-Sinoussi et al. 1983). HIV is made up of two classes, HIV-1 which is common worldwide and HIV-2 which is largely restricted to West Africa. HIV-1 has subtypes A to K, up to 49 circulatory recombinant forms (CRF) and several unique recombinant forms (URF). The HIV-1 subtypes exhibit distinctive geographical distribution across the world with subtype A predominant found in East Africa, subtype B is in Europe, America and Australia and subtype C which makes up >80% of HIV-1 infections worldwide being the main strain in Southern Africa and India (Buonaguro et al. 2007).

The distribution of HIV subtypes appears to follow patterns of HIV prevalence in different parts of the world suggesting that host genetic variation may have shaped the distribution of HIV subtypes by exerting selective pressure (Taylor et al. 2008).

The HIV genome is made up of two identical copies of positive strand RNA held together by hydrogen bonds (Huet et al. 1990). The HIV genome carries three structural genes named, group antigen (*gag*), polymerase (*pol*) and envelope (*env*). The three genes code for five major structural proteins (gp120, gp41, p24, p17 and p7) and three non-structural proteins namely reverse transcriptase (RT), integrase and protease which facilitate the viral life cycle (Vanhée-Brossollet et al. 1995). HIV also has six regulatory genes which code for proteins involved at various stages of its cycle in a host cell. These are; viral infectivity factor (*vif*), viral protein R (*vpr*), viral protein U (*vpu*), trans-activator of transcription (*tat*), negative factor (*nef*) and regulator of expression of viral proteins (*rev*) (Vanhée-Brossollet et al. 1995). Despite having a 9.1 kilo base (kb) genome, HIV is highly virulent and infectious.

HIV is transmitted from one individual to the next when infected body fluids such as semen, vaginal fluid and blood find their way into an uninfected person's system. The main modes of HIV transmission are sexual intercourse and mother-to-child transmission. Sharing of sharp objects and blood transfusion are minor contributors. This study focuses on children who were born to HIV-infected mothers, thus, MTCT was the route of infection.

1.3 Mother-to-child transmission of HIV

MTCT can occur during pregnancy (in-utero), at delivery (intrapartum) and after birth (postpartum). Studies have shown that in-utero HIV transmission accounts for 15-20% of the cases, breastfeeding 30-40% whilst approximately 50% of the events take place around delivery (Bryson et al. 1992, De Cock et al. 2000). In-utero HIV transmission occurs mainly through trans-placental passage of HIV (Backe et al. 1992). Even though the placenta forms a barrier that protects the developing foetus from HIV, the virus is able to penetrate through breaches in the barrier and/or transcytosis of intracellular HIV (Backe et al. 1992). HIV transmission during delivery takes place when the infant's skin and mucous membranes are exposed to infected maternal blood, amniotic fluid and vaginal secretions during passage through the birth canal (Perry *et al.*, 2010). Breastfeeding forms the main mode of vertical HIV transmission postpartum. The risk of transmission is highest just after delivery due to the high viral load in colostrum (Rousseau *et al.*, 2003).

The risk of HIV MTCT at any of these stages depends on both maternal and foetal factors. Factors determining risk of HIV MTCT include clinical and immunological status of the mother (HIV disease stage, co-infections), viral characteristics, tissue integrity in both mother and child and host genetic make-up of both mother and child (Renjifo et al. 2004, Singh et al. 2009). Co-infections such as sexually transmitted micro-organisms, malaria and tuberculosis in the mother increase the risk of transmitting HIV to the child (Gumbo et al. 2010, Gupta et al. 2011). This study focuses on the possible effect of child genetic variation on their risk of being infected from an HIV-infected mother. Control of HIV MTCT is dependent on the correct and timely diagnosis of HIV which guides subsequent treatment during pregnancy, delivery and the postpartum period.

1.4 HIV Diagnosis

Diagnosis of HIV is based on detection of the virus (virological) or antibodies produced against the virus (serological) in body fluids such as blood and blood products. Antibody detection based methods include enzyme linked immune-sorbent assay (ELISA), immunochromatography, and western blotting (Burke et al. 1987). Viral detection methods include p24 antigen test and nucleic acid amplification based tests (Sherman et al. 2005). HIV diagnosis using serological methods in children less than 18 months of age is difficult because sero-conversion is not certain at that stage due to their immature immune system therefore WHO recommends virological tests in infant HIV diagnosis (WHO 2010).

Once diagnosed of HIV, patients need to be monitored on a regular basis for effective treatment purposes. CD4+ T-cell counts and HIV quantification are used in patient monitoring and initiation of ART in HIV patients. In infants, WHO recommends immediate initiation of ART once the child has been diagnosed (WHO 2011) as studies have shown better survival in children who start therapy before they develop clinical symptoms (Chiappini et al. 2006, Violari et al. 2008). Despite the wide use of ART in HIV treatment, patients still progress to AIDS at different rates indicating the role of inter-individual variation at genetic and immunological level in HIV pathogenesis.

1.5 HIV pathogenesis and host genetics

1.5.1 HIV life cycle

HIV can only replicate inside a living host cell because in its complex life cycle, viral components have to interact with host proteins to execute their roles. Since the HIV genome only codes for about 15 proteins, its life-cycle like all retroviruses, is highly dependent on the host cell machinery (Frankel et al. 1998). The main stages in HIV life cycle are attachment and entry, reverse transcription, integration, transcription, assembly and budding. Figure 1.2 is an illustration of the HIV life cycle in a host cell.

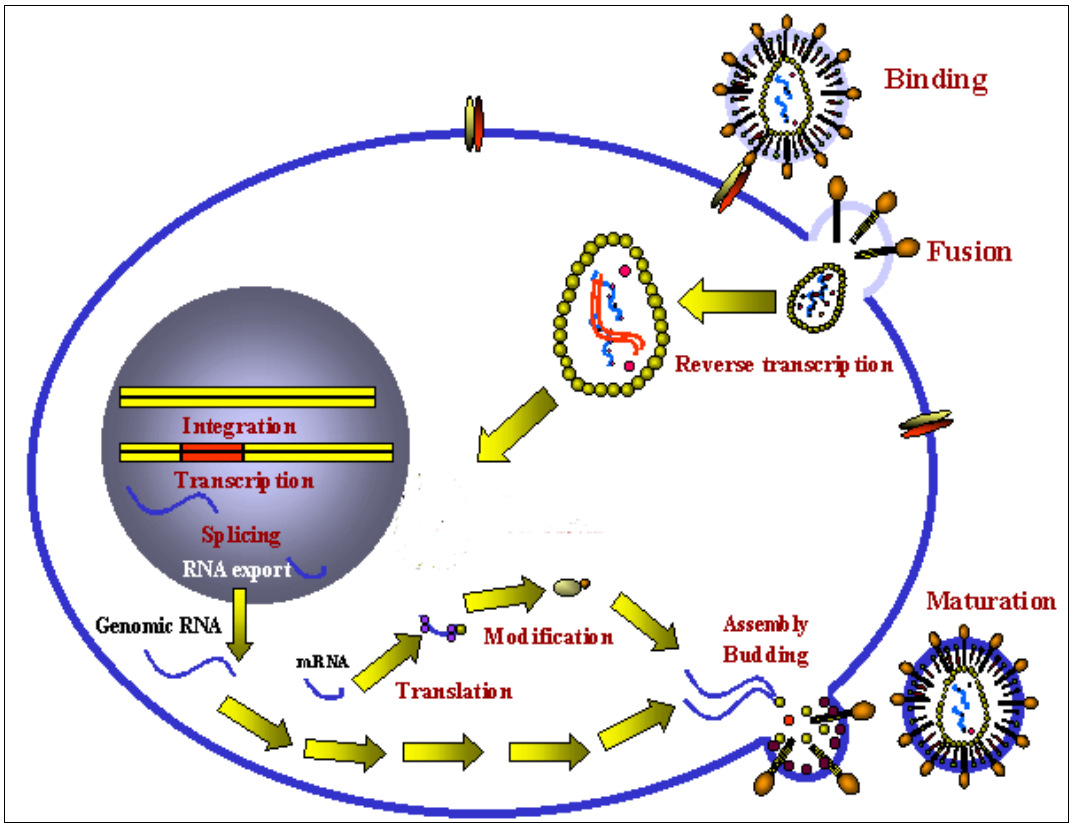


Figure 1.2: HIV life cycle in the human host cell (adapted from Hunt. 2009)

HIV infects the human cell by attaching its envelope glycoprotein 120 (gp120) to CD4 receptors on host cells such as T-lymphocytes, macrophages, langerhans cells and dendritic cells (Orloff et al. 1991, Rosenberg et al. 1991). This attachment causes gp120 to undergo a structural change that results in engagement of one of the major co-receptors, CCR5 or CXCR4. Gp41 an HIV transmembrane glycoprotein, then facilitates the transfer of the viral capsid into the host cell (McClure et al. 1988). Post-entry, the viral capsid is released into the cytoplasm where the outer viral lipid envelope is removed releasing the viral genome for reverse transcription to complementary DNA (cDNA) by viral reverse transcriptase enzyme (Chan et al. 1998). Once viral cDNA is made, it is combined with host proteins to form a pre-integration complex (PIC). This complex is necessary for translocation of the nucleic acid material to the nucleus where it is integrated into the human DNA by enzyme DNA integrase (Shun et al. 2007). The integrated DNA may lie dormant or is replicated simultaneously when human DNA multiplies before undergoing transcription to produce viral mRNA (Pace et al. 2011).

Tat is the only transcription factor encoded by HIV genome therefore human transcription factors are recruited to assist in the process (Calman et al. 1988). HIV protein Rev attaches to newly formed mRNA and shuttles it from the nucleus to the endoplasmic reticulum for further processing and translation (Pollard et al. 1998). Structural proteins encoded by viral genome such as Gag and Env bind to the viral RNA (Pollard et al. 1998). At the cell membrane, gp41 and gp120 are anchored into the membrane whilst the viral genome and proteins are assembled before a virion buds off, matures and goes on to infect other cells where they will undergo the same cycle in a repeated process (Gabuzda et al. 1992). Since HIV life cycle is heavily dependent on its interaction with host proteins, polymorphism in human genes that encode the proteins may impede or enhance the pathway.

1.5.2 HIV/AIDS restriction genes

Host genes whose variation influences course of HIV infection and disease are known as AIDS restriction genes (ARGs) (O'Brien et al. 2004). Some ARGs directly interfere with the HIV life cycle in a host cell whilst others may affect the host immune system. ARGs can be categorised into gene encoding the following; (1) chemokine receptors (e.g. *CCR5*, *CCR2* and *CX3CR1),* (2) chemokines (e.g. stromal cell derived factor, *SDF1*; regulated on activation normal T-cell expressed and secreted, *RANTES* and monocyte chemoattractant protein, *MCP-1*) and (3) innate immune factors (e.g. apolipoprotein B mRNA-editing enzyme-catalytic polypeptide-like 3G, *APOBEC3G*; tripartite motif 5α, *TRIM5α* and mannose binding lectin, *MBL2*) (Singh et al. 2009, Hutcheson et al. 2008).

An example of chemokine receptor polymorphism is a 32 base pair deletion in the human CCR5 receptor gene (*CCR5* Δ32) which leads to production of an aberrant protein that reduces HIV binding affinity to the receptor, thus, reducing chances of HIV infecting the host cell (Carrington et al. 1997). Chemokines on the other end, compete with HIV for co-receptor binding, thus, host polymorphisms that increase chemokine expression, for example, that of the *RANTES* and *SDF1* genes may confer a protective advantage against HIV whilst the opposite is true (An et al. 2002, Amara et al. 2010). In addition to SNPs, gene copy number variation (CNV) in a chemokine gene *CCL3L1* has also been reported to influence susceptibility to HIV/AIDS through the competitive block of the chemokine's receptor CCR5 (Gonzalez et al. 2005). CCL3L1 CNV however, does not form part of our current study.

Variation in genes that encode immunological factors may affect HIV infection by differential regulation of proteins involved in mounting a response against HIV such as chemotaxis and complement activation (Holmskov et al. 1994).

For example, low expression and poor oligomerisation of the mannose binding lectin (MBL) due to polymorphism in the gene encoding the protein (*MBL2*) have been linked with increased susceptibility to HIV infection and rapid disease progression in HIV patients (Boniotto et al. 2000, Mangano et al. 2008). The focus of this study is narrowed to two chemokine receptor genes (*CCR2* and *CX3CR1*), two chemokine genes (*RANTES* and *SDF1*) and two innate immune factor genes (*APOBEC3G* and *MBL2*). The study therefore was aimed at describing the distribution of their allelic variants in children born to HIV-infected mothers and their possible role in HIV infection and neurocognitive function. Figure 1.3 shows the HIV restriction factors investigated and the pathways affected.

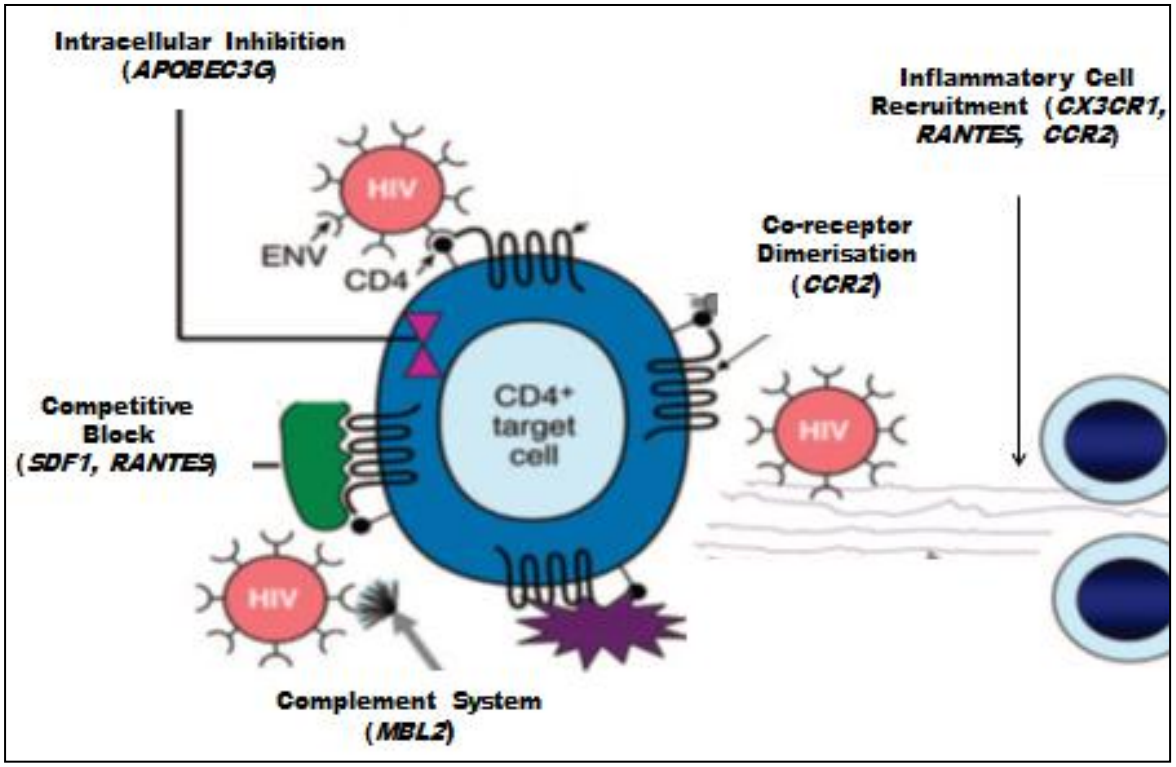


Figure 1.3: Schematic showing the viral restriction factors investigated in this study and their putative roles (adapted from Singh et al. 2009)

1.5.3 Distribution of HIV restriction genetic variants

There are several genetic variants that have been linked with differential HIV-related outcomes and their distribution varies amongst populations. This has made it difficult to come to a general consensus of their role in HIV pathogenesis (Martinson et al. 2000). The distribution of some genetic variants show trends across geographical boundaries suggesting that environmental factors could have influenced the polymorphism before HIV was discovered. For example, the *CCR5* Δ32 is prevalent in up to 15% of Caucasians but has not been found in African populations (Lucotte et al. 1998). In Europe where *CCR5* Δ32 allele is most common, it exhibits a decreasing trend in frequency from the North to South and West to East of the continent (Faure et al. 2008). Some authors have suggested that epidemics such as the bubonic plague in 1346-52 could have caused the deletion (Stephens et al. 1998) but others have argued that the distribution of *CCR5* Δ32 allele does not follow that of the plague (Cohn Jr et al. 2006).

Other variants such as *CCR2* 190A and *SDF1*-3′A do not show such defined trends among world populations and their frequencies tend to differ even among closely related populations (Ma et al. 2005, Qian et al. 2008). For example, the frequencies of both *CCR2* 190A and *SDF1* 801A alleles range from an estimated 1% to 30% in East Asian populations (Su et al. 1999) whist *CCR2* 190A allele frequencies range from 10-30% in seven Cameroonian tribes (Ma et al. 2005). Given the varied distribution of ARGs, it is important to investigate their frequency in all populations so that we may be able to approximate how they may be contributing to the HIV burden. The information may be useful in predicting HIV risk and disease progression in populations or individuals.

1.6 HIV disease progression

Despite the influence of host genetic variation in HIV life cycle and host immune mechanisms, HIV generally follows defined stages in its progression from infection to AIDS and subsequently death (Figure 1.4). Two to four weeks after infection, an infected person goes through the acute infection stage when HIV replication is very high. HIV attacks and kills CD4+ T-cells at a fast rate and the infected person may experience several illnesses including "flu like" symptoms (Cohen et al. 2011). The innate and CD8+ T-cell mediated immune factors mount a response against HIV that reduces viraemia to set point and allows CD4+ T-cells to recuperate (Deeks et al. 2004). This is followed by an asymptomatic HIV latency stage when the body is able to control HIV even though the virus still multiplies at a slow rate and CD4+ cell counts fall gradually. This period may last for several years and the length of time differs from one individual to the next which is suggestive of genetic variation influence (Siliciano et al. 2011).

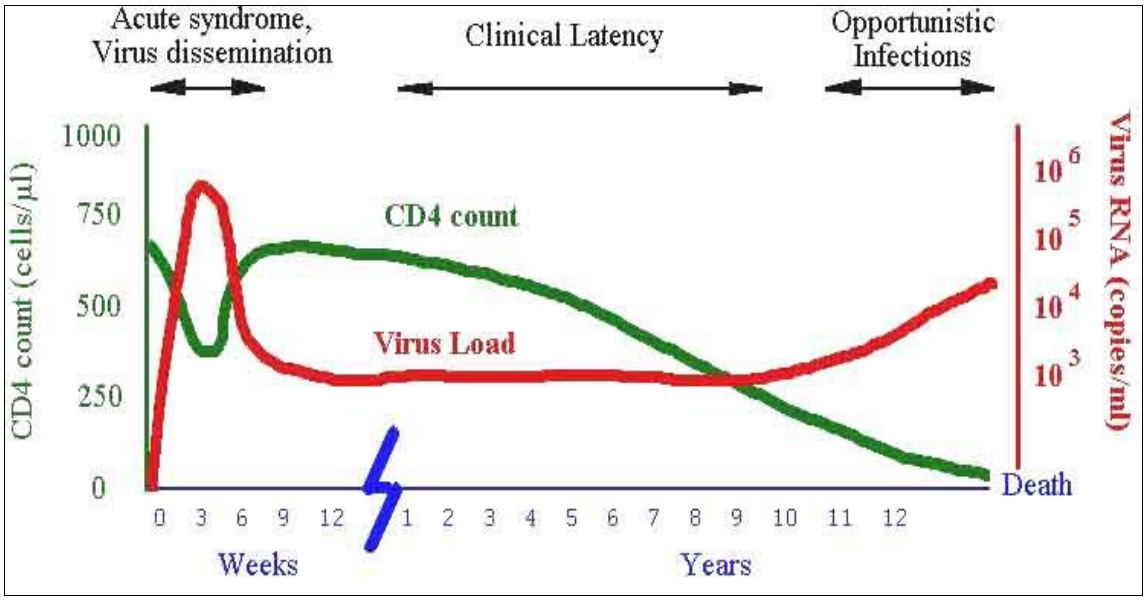


Figure 1.4: Stages in HIV disease progression. Host genetic variation in HIV/AIDS restriction genes may alter this course of disease progression (adapted from Dewhurst et al. 2000).

After several years of latency, one may progress to an AIDS stage. AIDS is marked by severe reduction of CD4+ T-cell count, high viraemia and opportunistic infections. The course of HIV disease progression in children is different from that in adults. HIV tends to generally progress faster with more serious effects in infants because of their immature immune system (Kuhn et al. 1999). Accelerated HIV disease could also be because children born to HIV infected mothers tend to acquire a virus with history of resistance (Tiemessen et al. 2006). HIV strains resistant to the maternal immune system are more efficiently transmitted from mother to child compared to well-controlled ones (Dickover et al. 2001).

HIV disease progression is very rapid during the first year of life and about 10-15% of the infected children are likely to die before by their first birthday (Spira et al. 1999, Pliner et al. 1998). The children exhibit a high viral loads which decrease at a slower rate compared to adults (Frederick et al. 1994). Contrary to adult immune systems where cytotoxic T lymphocyte responses reduce viraemia, infant immune responses produce virus specific CD8+ cells but they do not seem to reduce the amount of circulating virus (Lohman et al. 2005). Because of the poor ability to control HIV, HIV-infected children become a target for a wide spectrum of opportunistic infections. They experience recurrent bacterial infections, pneumonia, meningitis, viral respiratory infections to mention a few (Owor et al. 2004). They may show stunted growth, walking difficulties, poor social development and a range of neurological complications (McAthur et al. 2005). We therefore aim to determine if genetic variation has an effect on markers of disease progression such as CD4+ T-cell count and neurocognitive impairment.

1.7 Pathogenesis of HIV-associated neurocognitive disorders

HIV-associated neurocognitive disorders have been observed since the early stages of HIV (Grant et al. 1987, Mayeux et al. 1993). During the middle 1990's, an estimated 60% of HIV-infected patients had neuro-pathological conditions detected during post-mortem (Levy 1997). Ten years later, 15% of patients still developed HIV-associated dementia (HAD) whilst from 30-60% had the less severe forms (Gendelman et al. 2005; McArthur et al. 2004). These neurological complications are more aggravated in children in whom 50% of the untreated develop HAD and their intellectual ability is heavily compromised (González-Scarano et al. 2005).

In affected children problems include inefficient problem solving, poor planning and lack of general organisation. More complex manifestations involving visual-spatial processing, impaired fine motor skills as shown by poor arithmetic skills, difficulties carrying out daily activities such as dressing and poor handwriting develop with progression in the condition (Bisiacchi et al. 2000). With time, psychiatric complications can be observed in the children characterised by depression, delirium, anxiety, disorders, attention deficit-hyperactivity disorder (Brouwers et al. 1995, Krebs et al. 2000). The use of ART has improved the clinical outcome and reduced the prevalence of the serious forms of HIV-associated neurocognitive deficits but mild forms are still rampant with persistent impairment to the patients (Cysique et al. 2009). The first step in HIV neuropathogenesis involves HIV crossing the blood brain barrier (BBB) into the CNS.

The BBB is fashioned to provide biologic, physiological and immunological separation between the CNS and the rest of the circulation (Strazza et al. 2011). HIV exploits host machinery to cross this BBB using the "Trojan Horse" strategy (Williams et al. 1990, Garcia-

Garcia et al. 2005). The process begins with HIV infecting circulating monocytes where it remains latent. Whilst inside the monocytes, the virus enters the CNS during immune-surveillance and replacement of perivascular macrophages (Williams et al. 1990, Garcia-Garcia et al. 2005). This process is regulated by the secretion of chemokines such as macrophage inflammatory protein 1, monocyte chemoattractant protein 1 and RANTES. Other mechanisms of entry include macropinocytosis (Liu et al. 2002) and adoptive endocytosis mechanism mediated by gp120 (Banks et al. 2001).

Presence of HIV in the CNS introduces viral proteins such as gp120, Tat and Nef which increase secretion of chemokines, chemokine receptors, cytokines, nitric oxide and prostaglandins (Annunziata 2003, Strazza et al. 2011). Under normal physiological conditions, selected chemokines are expressed in the brain (Levine et al. 2009). These chemokines aid in leukocyte trafficking, cytokine activation, cell adhesion and other immunological roles. In the presence of HIV, this controlled environment is disturbed and non-specific chemokine expression is increased. This results in general inflammation of the organ and disruption of tight junctions of the BBB allowing more rapid entry of the HIV infected macrophages into the brain (Strazza et al. 2011).

Although disruption of the BBB and entry of the virus occurs early in infection, at that point it doesn't cause damage as HIV does not infect neurons and its penetration of endothelial cells occurs without productive infection (Liu et al. 2002). The fact that neurocognitive impairments mainly occur in the later stage of infection suggests that neurocognitive impairment may be a factor of disease progression (Liu et al. 2002). HIV-infected individuals exhibit inter-individual differences in the severity of neurocognitive deficits strongly pointing to the possible effect of host genetic variation on neurocognitive status.

In a different dynamic, Bagenda et al., (2006) have reported neurocognitive impairment among HIV-uninfected children suggesting that other etiological factors besides HIV/AIDS are important in neurocognitive deficits. The possible effects of human genetic variation on neurocognitive function are not well described. A study among American children reported an accelerated tendency to neurocognitive impairment among HIV-infected individuals carrying the chemokine gene variant *SDF1*-3′ A/A compared to the *SDF1*-3′G/G and G/A genotypes but did not observe any relationship with polymorphisms in genes encoding chemokine receptors CCR5 and CCR2 (Singh et al. 2003). However, another study reported better neurodevelopmental outcomes among children with the mutant *CCR5* genotypes compared to the wild-types and any of the *CCR2* genotypes (Llorente et al. 2006).

In addition to variations in chemokines and chemokine receptors, polymorphism in a human gene encoding the MBL protein called *MBL2* have been implicated in HIV-associated neurocognitive impairments (Singh et al. 2008, Spector et al. 2010). Here, the MBL2 variants associated with poor oligomerisation of MBL polypeptides to form functional units tend to present a higher risk of one developing neurocognitive impairments. This observation highlights the importance of the innate immune system in neuropathogenesis especially in HIV-infected individuals who are constantly faced with opportunistic infections that target the central nervous system. Several other studies have reported a number of SNPs involved in several pathways to influence the outcomes of HAD/HAND (Gonzalez et al. 2002, Pemberton et al. 2008, Bol et al. 2012). However, the first genome wide associated study (GWAS) on neurocognitive impairment hosting 1287 participants has dismissed most of the associations (Levine et al. 2012). Not much work has been reported among African populations therefore this study attempts to elucidate the correlation between host genetic variation and neurocognitive status.

1.8 Chemokines and their receptors

Chemokines are a type of cytokines made up of structurally related, small and highly basic proteins (Zlotnik et al. 2000). Chemokines and their receptors direct cells towards locations where they carry out specific functions in a process called leukocyte trafficking (Mahalingam et al. 2001). There are about 53 chemokines that interact with over 23 receptors in a promiscuous system where a chemokine may bind to more than one receptor and vice versa (Bonavia et al. 2003). In summary, chemokines can be grouped into four classes CC, XC, CX3C and CXC according to the spacing of cysteine residues on conserved sequences of the N-terminal region that influence formation of secondary structures (Murphy et al. 2000). Table 1.1 shows the classification of chemokines and their receptors with a few examples cited.

Table 1.1: Classification of chemokines and chemokine receptors

Class	Chemokine receptor	Chemokine
C	XCR1	XCL1 (Lymphotactin)
C-C	**CCR2**	CCL2 (MCP-1), CCL8 (MCP-2)
	CCR5	**CCL5 (RANTES)**, MCP-2, CCL3 (MIP-1α)
C-X-C	CXCR4	**Stromal cell-derived factor (SDF-1)**
	CXCR1	Interleukin 8, CXCL6 (GCP-2)
C-X3-C	**CX3CR1**	CX3CL1 (Fractalkine)

Key: MCP- monocyte chemoattractant protein, RANTES- regulated on activation, normal T-cell excreted and secreted, MIP- macrophage inflammatory protein, GCP- granulocyte chemotactic protein.
NB: Chemokines and chemokine receptors investigated in this study are shown in bold.

Chemokines carry out a wide range of functions by binding to their G protein-coupled receptors (GPCR) on cells. When a chemokine binds to its receptor, a protein attached to the receptor is activated and released to initiate a cascade of reactions that involves phopholipases and change in cytosolic calcium (Bokoch 1995). The eventual result of the chain is activation of several transcription factors such as CREB and STAT which increase gene expression and proliferation (Bokoch 1995, Kuang et al. 1996). Chemokines are actively involved in the immune system where they modulate both the innate and adaptive immune processes such as primary leukocyte trafficking, immune surveillance and inflammatory cell recruitment (Bajetto et al. 2001). Chemokines are also involved in embryonic development, wound healing, angiogenesis, leukocyte infiltration, metastasis of tumours and apoptosis (Belperio et al. 2000, Ishida et al. 2008, Lazennec et al. 2010).

In addition to the above mentioned functions, chemokine receptors also act as co-receptors to HIV e.g. CCR5, CXCR4 and CCR2 (Deng et al. 1996, Feng et al. 1996). This led to investigation into the role of these co-receptors in HIV disease mechanisms. Sequencing of genes encoding CCR5, CCR2, CX3CR1 and other minor co-receptors has shown several polymorphisms that seem to affect HIV infection and disease progression. Scientists have invested a lot of effort in describing the distribution of the polymorphisms in different populations across the world but Africans remain understudied. Zimbabwe is no exception and to our knowledge, this study will be the first to describe most of the variations in chemokine/chemokine receptor genes among Zimbabweans. The focus of this study is on polymorphism in genes encoding two chemokine receptors (*CX3CR1* and *CCR2),* two chemokines *(RANTES* and *SDF1)* and two innate immune system players (*APOBEC3G* and *MBL2).*

1.8.1 C-X3-C chemokine receptor type 1 (CX3CR1)

Chemokine receptor CX3CR1 was first reported in 1997 (Imai et al. 1997) as a high-affinity ligand for the chemokine of CX3C motif called fractalkine (CX3CL1) (Bazan et al. 1997). The receptor-ligand combination, CX3CR1-fractalkine facilitates leukocyte trafficking as cells carrying the CX3CR1 receptor migrate towards regions of high fractalkine expression. CX3CR1 is highly expressed on monocytes, natural killer cells and a selection of T-lymphocytes and therefore plays a part in their recruitment during immune responses (Imai et al. 1997, Nishimura et al. 2002). CX3CR1 is also expressed in astrocytes and glial cells of the CNS where it is involved in neuronal activity (Nishiyori et al. 1998, Meucci et al. 2000). In addition to leukocyte trafficking, CX3CR1 also acts as an HIV co-receptor (Combadiere et al. 1998). Expression of its ligand fractalkine has been reported to differ between groups of HIV-infected and HIV-uninfected individuals suggesting that CX3CR1 and fractalkine have an active role in immunity against HIV (Foussat et al. 2001). Fractalkine is also highly expressed in the brain of HIV infected patients and may be a culprit in HAD (Cotter et al. 2002).

CX3CR1 is encoded by an 18kb long gene located with 4 other chemokine receptor genes on chromosome 3 (Maho et al. 1999). Two non-synonymous SNPs, *CX3CR1* 745G>A (V249I) and 839T>C (T280M) have been reported in the sixth and seven transmembrane domains of CX3CR1 protein (Faure et al. 2000). The resulting amino acid substitutions have been reported to reduce the binding of CX3CR1 to its cognate ligand fractalkine (Faure et al. 2000). Alleles 745A and 839T were reported to be in complete linkage disequilibrium (LD) among French Caucasians (Faure et al. 2000).

The resulting 745-839 (A-T) haplotype is associated with decreased monocyte binding and poor fractalkine-dependant cell to cell adhesion (McDermott et al. 2003). Individuals carrying the homozygous 839T/T genotype therefore exhibit reduced fractalkine binding and fractalkine induced leukocyte chemotaxis (McDermott et al. 2003). The reduced potency of CX3CR1 in leukocyte trafficking negatively affects its ability to activate the immune system. This offers a possible explanation why the A-T haplotype was reported to be associated with accelerated HIV/AIDS disease progression in one of the earliest studies on *CX3CR1* variation and HIV (Faure et al. 2000).

Subsequent studies have reported conflicting findings on the effect of *CX3CR1* 745A and 839T alleles on HIV/AIDS disease progression. Some reported the 745A and 839T alleles to be associated with faster HIV disease progression compared to 745G and 839C alleles, respectively (Faure et al. 2003, Brumme et al. 2003, Singh et al. 2005). Others have not found any significant association (Hendel et al. 2001, Kwa et al. 2003, Suresh et al. 2006) whilst one study demonstrated association between the alleles (745A and 839T) and long term non-progression in HIV-infected patients (Vidal et al. 2005). However, most of these studies were carried out in adult populations and not much has been done in paediatric populations especially in sub-Saharan Africa.

One of few studies done in children was a multicentre study in the USA which showed that children with 745A/A genotype experienced more rapid disease progression and a trend towards neurocognitive impairment compared to those with the 745G/G genotype (Singh et al. 2005). Another study involving Malawian, South African and Ugandan pregnant women, some treatment naive and others on zidovudine prophylaxis, did not find any association between the genetic variants and risk of HIV MTCT (Singh et al. 2008).

There is no general consensus on the effect of *CX3CR1* polymorphism on risk of HIV infection or disease progression. In addition to the *CX3CR1* gene, polymorphism in another chemokine receptor gene, *CCR2* was also investigated.

1.8.2 C-C Chemokine receptor type 2 (CCR2)

CCR2 is a seven transmembrane chemokine receptor found on several cells including monocytes, myeloid precursor cells and activated T-cells (Myers et al. 1995). CCR2 is involved in several functions including migration of monocytes from the bone marrow where they are made to peripheral circulation during inflammation, bone homeostasis and maintenance of physiologically correct chemokines levels through scavenging of excess chemokines (Palframan et al. 2001, Cardona et al. 2008, Binder et al. 2009).

In addition to its physiological roles, CCR2 is also a minor HIV co-receptor and has been implicated in HIV-associated dementia and other neurological disorders such as Alzheimer's disease and multiple sclerosis (Vos et al. 2000, Gonzalez et al. 2002). The *CCR2* is a 7kb gene on chromosome 3p21 (Daugherty et al. 1997). A 190G>A polymorphism in the protein coding region of *CCR2* gene results in a valine to isoleucine change at amino acid 64 of the receptor (V64I) (Charo et al. 1994, Smith et al. 1997). The *CCR2* 190>A has been reported in most populations and is one of the most widely studied HIV/AIDS restriction genetic variants (Martinson et al. 2000). The *CCR2* 190A allele enhances gene expression and prolonged half-life of its CCR2A isoform (Nakayama et al. 2004). The 64I containing variant of CCR2 receptor dimerises with CXCR4 receptor after binding of its ligand whereas the 64Vvariant has reduced affinity. This reduces the amount of CXCR4 available to bind the X4 HIV strain on the cell surface among *CCR2* 190A carriers, thus conferring a protective advantage against HIV infection (Mellado et al. 1999).

A study among Cameroonians reported a protective effect against HIV infection in males carrying the *CCR2* 190A allele compared to the 190G allele. However this association was not observed in females where other factors such as multiple pregnancies and other maternal stressors may have masked the genetic effect (Ma et al. 2005). The protective effect of the *CCR2* 190A variant is further supported by reports that *CCR2-CCR5* haplotypes containing the 190A allele were more frequent in HIV-exposed but uninfected adults compared to the HIV-infected among Thais and Zambians (Malhotra et al. 2011, Wichukchinda et al. 2008). However, some studies have failed to observe any association between *CCR2* 190G>A variants and HIV transmission (Smith et al. 1997, Alagarasu et al. 2009). The bulk of studies have been done on adult populations but the effect of *CCR2* 190G>A variants on HIV transmission in children is not well described.

Some of the few studies involving children reported a protective effect against HIV MTCT in children carrying the 190A allele compared to those with the 190G allele (Mangano et al. 2000, Mabuka et al. 2009). In contrast, a multicentre study involving Malawians, South Africans and Ugandans reported an increased risk of HIV MTCT in children with the homozygous 190A/A genotype compared to 190G containing genotypes (Singh et al. 2008). Our study aims to elucidate the effect of *CCR2* 190G>A variation on the risk of HIV infection among who are perinatally exposed to HIV.

In individuals already infected by HIV, several studies have suggested a better prognosis in HIV/AIDS patients carrying the *CCR2* 190A variant (Smith et al. 1997, Mangano et al. 2000, Kostrikis et al. 1998, Kostrikis et al. 1998, Vieira et al. 2011, Vieira et al. 2011, Mulherin et al. 2003) whilst others oppose these findings (Schinkel et al. 1999, Michael et al. 1997).

Moreover HIV/AIDS, the 190G>A variants have been shown to influence the outcome of several other clinical conditions such as cancer (Chatterjee et al. 2010, Chen et al. 2011), preeclampsia (Agachan et al. 2010), myocardial infarction (Karaali et al. 2010) and atherosclerosis (Nyquist et al. 2009). This can be attributed to the receptor's wide spectrum of roles in the immune system through activation of immune cells. In addition to *CX3CR1* and *CCR2* receptor genes, the study also investigated genetic variation in two chemokine genes *SDF1* and *RANTES*. SDF1 and RANTES are the natural ligands of the two main HIV co-receptors CXCR4 and CCR5 respectively (Deng et al. 1996).

1.8.3 Stromal cell derived factor 1 (SDF1)

Stromal cell derived factor (SDF) is a CXC chemokine originally described as a secreted product of bone marrow stromal cell line (Tashiro et al. 1993). SDF is encoded by the *CXCL12 (SDF1)* gene. The *SDF1* gene is located on chromosome 10 spanning a 17kb region. SDF1 is chemoattractant involved in leukocyte trafficking of several cell lines including T-cells, monocytes, pre-B cells, dendritic cells, and haematopoietic progenitor cells (Sozzani et al. 1997, Möhle et al. 1998, Bleul et al. 1996). The SDF1-CXCR4 combination is vital for developmental processes such as cardiogenesis, brain development and guidance of primordial germ cell (Tachibana et al. 1998, Ara et al. 2003). Furthermore, SDF1 is also active in CNS physiology with phylogenetic studies showing that the CXC family has ancestral roles in the CNS opposed to the immune system (Huising et al. 2003, Guyon et al. 2007). Both CXCR4 and SDF-1 are constitutively expressed by glial and neuronal cells in the CNS (Bonavia et al. 2003, Bajetto et al. 2001) with co-distribution of the pair in a number of brain regions indicating a functional system (Guyon et al. 2007).

Despite the multiple role of SDF1, this study focused on the how its genetic variation might affect HIV infection and neurocognitive development. The chemokine SDF1 inhibits infection of the host cell by HIV through competitive binding of the HIV co-receptor CXCR4 (Oberlin et al. 1996, Yuan et al. 2000). SDF1 has also been implicated in HIV-HAND due to its active role in CNS physiology. Investigations on variation in the *SDF1* gene led to the discovery of a G>A change at position 801 in the 3′ untranslated (3′-UTR) region of the gene (*SDF*1 801G>A; *SDF1*-3′G>A) (Oberlin et al. 1996). The *SDF*1 801G>A SNP has been implicated in differential HIV transmission and disease progression. This is because the 801A allele up-regulates the synthesis of the SDF1 protein by stabilising the mRNA allowing for more competitive binding against HIV (Winkler et al. 1998, Shirozu et al. 1995).

Reports on the role of *SDF1* variants in HIV infection and disease progression are conflicting. One of the earliest studies reported slower disease progression in individuals carrying the *SDF1* 801A allele compared to those with the 801G variant (Winkler et al. 1998). This observation was disputed by several subsequent studies that found the 801A to be associated with accelerated progression from infection to AIDS (Amara et al. 2010, Mummidi et al. 1998, van Rij et al. 1998) whilst others have not found any association (Watanabe et al. 2003, Reiche et al. 2006).The relationship between 801G>A variants and either HIV infection or disease progression in African and paediatric populations is not well described. Among the Xhosa speaking people of South Africa, the *SDF1* 801A was reported to increase risk of HIV infection in adults (Petersen et al. 2005). In one of the few African studies involving children Singh et al. (2003) reported rapid disease progression in individuals with homozygous *SDF1* 801A/A compared to 801G/G and 801G/A carrying individuals (Singh et al. 2003).

Despite SDF1's active involvement in CNS pathophysiology, the effect of variation in its encoding gene on HIV associated neurocognitive impairments is not well described, thus, its inclusion in this study. Whilst SDF1 binds to HIV co-receptor CXCR4, another chemokine RANTES binds to the other main HIV co-receptor, CCR5.

1.8.4 Regulated on Activation Normal T-cell Expressed and Secreted

Regulated on Activation, Normal T-cell Expressed and Secreted (RANTES); CCL3L1 and Monocyte Inflammatory Proteins, MIP-1α (CCL3) and MIP-1β (CCL4) encoded by host genes *CCL5* (*RANTES*), *CCL3L1*, *CCL3* and *CCL4*, respectively, are ligands to the chemokine receptor and HIV R5 co-receptor CCR5 (Table 1.1) (Deng et al. 1996, Cocchi et al. 1995, Torre et al. 2000, Gonzalez et al., 2005). RANTES, CCL3, CCL4 and CCL3L1 exhibit anti-HIV activities through competitive block of CCR5. Of the genes encoding the four chemokines, *RANTES* is the most studied and SNPs that influence HIV outcomes show a worldwide distribution, thus was selected to form part of this study.

RANTES also inhibits HIV infection by reducing cell surface CCR5 levels (Mack et al. 1998). The discovery of the interaction between RANTES and CCR5 presented a potential therapeutic strategy against HIV. The strategy is based on co-receptor blockage and has led to the first CCR5 inhibitor drug called maraviroc (Fätkenheuer *et al*., 2005). CCR5 antagonism has also been exploited in the "cure" of the Berlin patient with HIV and acute myeloid leukaemia in whom CCR5 negative stem cells were successfully used to clear both leukemic cells and HIV (Hütter *et al.*, 2009).

Primarily RANTES is a chemoattractant for a variety of cell types including monocytes, natural killer cells, memory T cells, eosinophils and dendritic cells (Loetscher et al. 1994, Schall et al. 1990, Rot et al. 1992). RANTES plays an important role in immune regulation against respiratory viral infections such as the common influenza and respiratory synctia virus (Hao et al. 2008, Culley et al. 2006). It was also shown that in RANTES deficient mice, virus-specific CD8+ T cells had poor cytokine production with reduced cytotoxic ability (Crawford et al. 2011). In the CNS, RANTES is produced in astrocytes and up-regulated in HIV infection. This elevation of RANTES in the CNS has been reported to be associated with SIV encephalitis (Barnes et al. 1996). Execution of RANTES function depends on how much of the chemokine is available in the circulation for certain pathways.

RANTES levels in circulation show inter-individuals differences and in the context of HIV/AIDS, different levels of RANTES could result in differences in inhibition of HIV-CCR5 binding and therefore risk of infection or disease progression. The differences in RANTES levels can be attributed to variation in the *RANTES* gene, thus, *RANTES* polymorphisms have been linked with HIV transmission and/or disease progression. Three *RANTES* SNPs namely -403G>A and -28C>G in the promoter region and In1.1T>C in intron 1 have been reported to affect differential expression and protein levels of RANTES (An et al. 2002, Bleul et al. 1996). The -403A and -28G alleles have been reported to up-regulate the expression of *RANTES* (Liu et al. 1999, Nickel et al. 2000). Increased expression would mean that more chemokines are available for blocking of HIV binding to CCR5 co-receptor, thus, protection against HIV infection and disease progression. On the contrary, reduction of RANTES expression would result in more CCR5 being left available for HIV binding and therefore increased chances of infection by HIV.

Reports on association studies between *RANTES* gene variants and HIV disease progression are conflicting. Several studies have reported slow disease progression in individuals carrying the -403A and/or -28G alleles (Liu et al. 1999, Koizumi et al. 2007, McDermott et al. 2000) whilst others have not found any association (Vidal et al. 2006). Studies on the effect of *RANTES* variants on risk of HIV infection have also reported conflicting findings (McDermott et al. 2000, Rathore et al. 2008, Katz et al. 2010, Ahlenstiel et al. 2005, Liu et al. 2004). *RANTES* -403G>A and -28C>G variants have also been implicated in susceptibility to other conditions such as hepatitis B and C, asthma (Muro et al. 2008) and myocardial infarction (Irina P et al. 2011). A study among Tunisians reported that the -28G/G genotype was significantly associated with increased risk of pulmonary TB infection whilst -403G allele increased chances of extra pulmonary TB (Ben-Selma et al. 2011).

Contrary to the increased expression associated with the two promoter SNPs, *RANTES* In1.1C allele affects the binding of nuclear protein to its regulatory sequence resulting in down-regulation of *RANTES* expression (An et al. 2002). The In1.1C allele has been linked with accelerated disease progression in HIV/AIDS patients in several studies (An et al. 2002, Wichukchinda et al. 2006, Gonzalez et al. 2001). A study among Indians reported the In1.1T allele to be a risk factor for HIV transmission whilst the In1.1C allele was associated with accelerated HIV disease progression. Other studies have failed to observe any association between the In1.1T>C variants and neither disease progression nor HIV transmission (Koizumi et al. 2007, Katz et al. 2010).

There is limited information on *RANTES* SNPs in African population. A study among Ugandans reported longer survival among individuals carrying the homozygous In1.1C/C genotype compared to those who had the In1.1T/T genotype (Cooke et al. 2006).

This study aimed to investigate the distribution of *RANTES* -403G>A and In1.1T>C variants in Zimbabwean population and how they might be associated with HIV infection and neurocognitive status. Even though chemokines and chemokine receptors are involved in HIV attachment to the host cell during infection, they are basically part of the antiviral innate immune system. The innate immune system is a vital player in HIV control and pathogenesis.

1.9 Innate immune system

The immune system forms the body's line of defence against infections and disease (Lynn et al. 2010). The immune system may be broadly classified into anatomical and physiological barriers, innate and adaptive immune systems. The barriers physically keep pathogens out of the body while the adaptive immune system is an antigen recognition-dependent, specific response mounted against a specific pathogen by activating monoclonal production of specific antibodies by B-cells or effector T cells (Iwasaki et al. 2010, Turvey et al. 2010).

The innate arm of the immune system is a non-specific defence line that uses a spectrum of pattern recognition receptors to detect pathogen-associated molecular patterns on micro-organism (Carrington et al. 2012). The pattern recognition results in activation of inflammatory pathways involving cells such as macrophages, dendritic cells, mast cells, neutrophils and natural killer cells (Turvey et al. 2010). Non cellular pathways involved in the innate immune system include; lipopolysaccharide binding, C reactive protein (CRP) and collectins (Turvey et al. 2010). Mannose binding lectin (MBL) is part of the collectins (Holmskov et al. 1994). Other innate immune factors act intracellularly by targeting particular components of the cell.

An example is the APOBEC family of cytidine deaminases which acts through recognition of double stranded RNA of invading pathogens (Bishop et al. 2004).

In HIV infection, the body only manages to mount a specific immune response some weeks after infection leaving the innate immune system to face the virus during the early stages of infection. Thus, the innate immune system is the major determinant of the viral set-point. The viral set point in turn determines the rate of disease progression (Pantaleo et al. 1996). The current study investigated variation in genes encoding two innate immune factors, mannose binding lectin (*MBL2*) and apolipoprotein B m-RNA enzyme-editing catalytic polypeptide-like 3G (*APOBEC3G*) and their possible role in HIV infection and neurocognitive function.

1.9.1 Apolipoprotein B mRNA-editing catalytic polypeptide like-3G

Apolipoprotein B mRNA-editing catalytic polypeptide like-3G (APOBEC3G/APOBEC3G) belongs to the APOBEC family of antiviral cytidine deaminases (Bishop et al. 2004). The antiviral properties of APOBEC are affected by presence of HIV regulatory protein Vif which is potent to viral infectivity (Sheehy et al. 2002). Early studies reported Vif to be present in selected cell lines called “permissive” cells where HIV is able to proliferate but absent in “non-permissive” where the virus can’t replicate (Gabuzda et al. 1992, von Schwedler et al. 1993). The discovery of these different cell types led to the hunt for host genes that were differentially expressed in these two sets of cells. It was then demonstrated that the *APOBEC3G* gene expresses antiviral function which are inhibited by Vif (Sheehy et al. 2002). It is now agreed that in the absence of the Vif, APOBEC3G protein is encapsulated into new virions that bud off an infected cell.

When the virion infects another cell, the APOBEC3G deaminate cytidine to uracil (C>U) on the viral minus DNA strand during reverse transcription resulting in G>A mutation in the complementary coding strand (Zhang et al. 2003, Lecossier et al. 2003).This is commonly known as G>A hypermutation. The G>A hypermutation cripples HIV infectivity by introducing multiple stop codons in the HIV genome (Zhang et al. 2003, Lecossier et al. 2003). It has been reported that the 5′-CCCA-3′ sequence on the viral minus strand is a 'hot spot for cytidine deamination (Bishop et al. 2004, Yu et al. 2004). Mutation of this sequence leads to the complementary 5′-TGGG-3′ changing to 5′-TAGG-3′ which translates from a tryptophan codon to a stop codon (TAG). Other mechanisms of APOBEC3G antiviral action include inhibition of viral DNA synthesis by hindering the translocation of reverse transcriptase along the RNA template (Malim 2009).

The APOBEC3G protein is encoded by a 12.7kb gene on chromosome 22. A study by An et al. (2004) reported seven SNPs in the *APOBEC3G* gene, three in the promoter region, two in exons and two in introns. *APOBEC3G* 557A>G (H186R) in exon 4 is the most studied SNP on the gene (An et al. 2004). The *APOBEC3G* 557G/G genotype has been linked to accelerated HIV disease progression in adults of African origin but not in Caucasians (An et al. 2004). This observation was supported by later studies that also reported accelerated HIV disease in 557G allele carrying individuals compared to those who had the 557A allele (Do et al. 2005, Reddy et al. 2010). Most studies have failed to observe association between any of the *APOBEC3G* variants and HIV transmission (An et al. 2004, Do et al. 2005, Reddy et al. 2010, Valcke et al. 2006). Cho et al. (2006) attempted to elucidate the relationship between *APOBEC3G* expression and disease progression using mRNA but failed to observe any association.

Contrary to this, a study measuring mRNA levels in human cells from different groups of HIV-positive people reported that mRNA abundance followed the order; long term non-progressors>HIV uninfected> progressors. *APOBEC3G* mRNA abundance correlated with CD4+ T-cell counts and inversely with viral loads (Jin et al. 2007). These observations suggest that increased APOBEC3G expression may slow the progression of disease in HIV infected individuals.

This study investigated polymorphism at five loci (SNPs) in the *APOBEC3G* gene; two in the promoter (-571G>C and -90C>G), one in exon 4 (557A>G) and the two in introns (197T>C and 199C>G) to determine their baseline frequencies and possible association with HIV infection and neurocognitive development. Whilst APOBEC3G acts through recognition of double stranded viral RNA, other innate immune antiviral factors act on virus through recognition of surface patterns such polysaccharide chains for example, mannose binding lectin which also formed part of this study.

1.9.2 Mannose Binding Lectin (MBL)

Collectins are a family of soluble mammalian pattern recognition receptors that contain a C-type lectin attached to a collagen-like domain (Holmskov et al. 1994). Collectins are key players in the innate immune system where they kill invading organisms by recognising pathogen-associated molecular patterns before triggering a wide spectrum of microbial clearance mechanisms such as aggregation, complement activation and phagocytosis (Holmskov et al. 1994, Turner 1996). Mannose binding lectin (MBL) is a member of the collectins family that is encoded by the *MBL2* gene. MBL has affinity for a range of carbohydrates such as D-mannose, N-acetyl-D-glucosamine on the surface of pathogens.

By binding polysaccharide chains on pathogen surfaces, MBL activates the lectin pathway of the complement system to kill invading micro-organism (Holmskov et al. 1994). The complement system has three activation pathways (Figure 1.5) which are (1) classical which is initiated by antibodies binding target antigens, (2) alternative pathway which is capable of auto-activation through a spontaneous 'tick over' of the C3 protein when it binds to a microbe and (3) the lectin pathway which involves MBL and ficolins (Endo et al. 2006). MBL binds to carbohydrates on the surface of a micro-organism in a process that activates MBL-associated serine proteases (MASPs) to cleave the C4 protein of the complement system (Matsushita et al. 2000). Cleavage of C4 results in formation of the C3 convertase, C4bC2a (for the classical and lectin pathways). The C3 convertase stage is where the activating pathways converge and the subsequent steps are similar in all three. The eventual result is the formation of a membrane attack complex C5b-9 which kills the pathogen by bursting its membrane or wall (Biesecker et al. 1979, Podack et al. 1979).

Each MBL polypeptide is made up of 4 regions, C-terminal domain, collagen-like region, α-helical coiled-coil neck and the carbohydrate-recognition domain (CRD). Figure 1.6 shows the structure of the MBL protein and the encoding *MBL2* gene. The CRD part of the MBL as its name suggests is involved in the selective binding to specific carbohydrates on the pathogen surface (Weis et al. 1992, Håkansson et al. 1999). The α-helical neck region bridges the CRD and the collagen-like domain. The neck region is involved in the trimerisation of the polypeptide units of the MBL and the subsequent oligomerisation of the trimers into multimers (Kishore et al. 1997). The collagen-like domain is mainly responsible for interaction with MASPs in the cascade that leads to complement activation (Thiel et al. 1997).

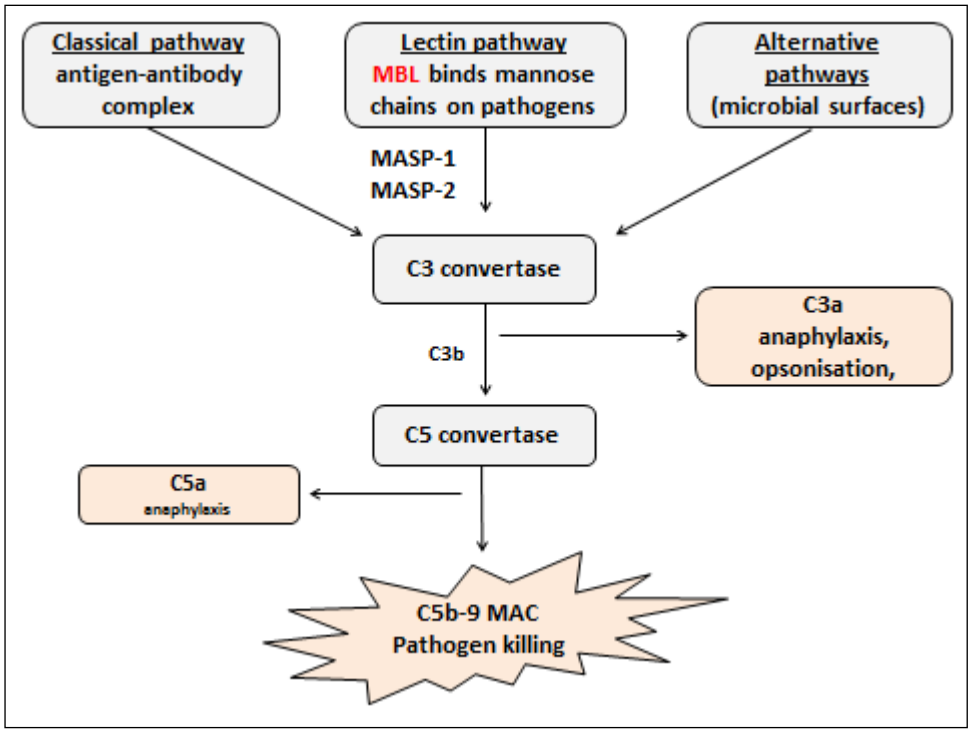


Figure 1.5: Complement system activation pathways. There are three main activation pathways that converge at the C3 convertase stage. The final product of all pathways is the formation of the membrane attack complex (MAC) which kills the invading pathogen whilst other by-products facilitate microbial opsonisation and other anaphylactic processes.

Stabilisation of multimers is a function of the cysteine rich N-terminal domain of the MBL which forms strong disulphide bridges between monomers in the folded structure (Wallis et al. 2008). Through multimerisation, MBL forms structures with six trimers giving an 18 polypeptide complex that resembles a bouquet of flowers (Garred et al. 2006). Multimerisation is crucial for the overall affinity of the lectin towards carbohydrates on pathogen surfaces and failure to assemble into these structures can result in poor innate immunity (Clark et al. 2000).

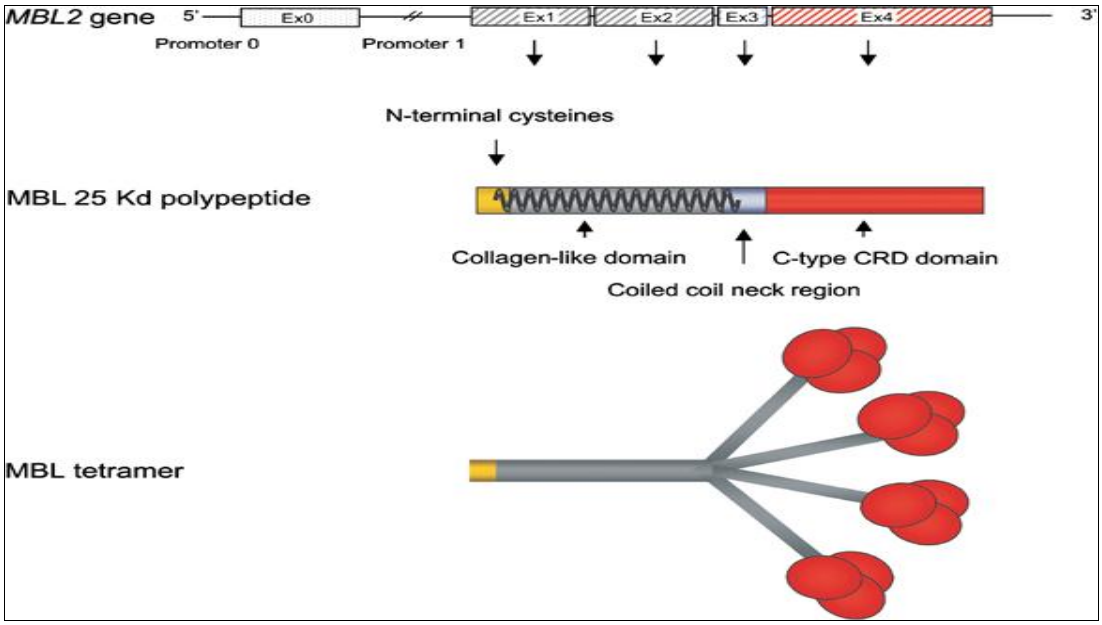


Figure 1.6: Schematic representation of *MBL2* gene, MBL polypeptide and tetramer structures. The gene shows the exons and their corresponding polypeptide regions whilst the bottom panel shows polypeptides combine into trimers which in turn combine to form the tetramers **(**adapted from Garred et al. 2006).

The *MBL2* is a 6.32 kb gene located on chromosome 10q11.2-21 (NCBI). Polymorphisms in its promoter region and exon1 have been reported to interfere with the overall function of the lectin. Three independent missense SNPs in exon 1 (154C>T, 161G>A, 170G>A) result in the following amino acid changes; R52C, G54D, G57E alternatively named '*D*', '*B*' and '*C*' respectively whilst the wild-type allele is "*A"*. The *B, C* and *D* alleles collectively make up the "*O*" genotype (Sumiya et al. 1991, Lipscombe et al. 1992, Madsen et al. 1998). These SNPs alter the amino acid sequence in the first part of the collagenous region of the MBL near the cysteine rich domain which interferes with oligomerisation of the MBL trimers (Madsen et al. 1998).

The distribution of the *MBL2* exon 1 alleles differs in different populations with African populations predominantly carrying the *C* allele whilst Caucasians have the B allele mostly (Lipscombe et al. 1992, Madsen et al. 1998). In the *MBL2* promoter region SNPs, -550G>C (-550*H/L*), -221C>G (-221*X/Y*) and +4C>T (+4*P/Q*) in 5` UTR have been reported to affect *MBL2* expression (Madsen et al. 1995). Due to these polymorphisms, MBL levels in healthy individuals may vary up to 1000 fold (Crosdale et al. 2000). The *-550G*, *-221C* and *+4T* variants have been reported to result in low and poorly oligomerised MBL which is associated with poor prognosis in meningococcal disease and viral hepatitis (Koutsounaki et al. 2008, Brouwer et al. 2009).

Antiviral mechanisms of MBL against HIV involve the lectin binding to high mannose oligosaccharides on HIV gp120 making the gp120 unavailable for HIV attachment to host cell, thus, blocking the virus from entering the host cell (Saifuddin et al. 2000). MBL also neutralises HIV by blocking the virus' interaction with dendritic cell-specific intercellular adhesion molecule-3-grabbing non-integrin (DC-SIGN) (Spear et al. 2003). DC-SIGN is a C-type lectin receptor expressed on macrophages and dendritic cells that binds to mannose chains on HIV in a process that aids the virus in infecting host cells (Geijtenbeck et al. 2000). DC-SIGN is expressed on macrophage subsets that infest the placenta and therefore has been implicated in HIV MTCT (Soilleux et al. 2001). Studies among Zimbabweans have reported increased risk of MTCT in pregnant women carrying certain genetic variants in the promoters and exons of *DC-SIGN* and *DC-SIGNR* genes (Boily-Larouche et al. 2009, 2012). The current study focuses on the possible role of *MBL2* polymorphism on HIV transmission and neurocognitive function among Zimbabwean children.

A study among Argentinians demonstrated that *MBL2* haplotype *XA/XA* associated with risk of acquiring HIV and accelerated progression to AIDS in children (Mangano et al. 2008). Another study among children born to HIV-infected mothers, reported that possession of an *MBL2 O* allele was associated with accelerated HIV disease and a tendency to develop neurocognitive impairment in children younger than 2 years of age (Singh et al. 2008). *MBL2* genetic variations have also been linked with opportunistic organisms that tend to cause neurological complications such as cytomegalovirus (Hu et al. 2010), pneumococcal meningitis (Ou et al. 2011) and tuberculosis (TB) (Thye et al. 2011). This suggests that MBL may be involved in protecting the CNS from HIV-associated neurological infections.

1.10 Aims and Objectives

1.10.1 Aims

Given the qualitative and quantitative differences in the distribution of HIV/AIDS restriction gene variants among populations and the lack of concordance on their effects on HIV infection, disease progression and neurocognitive function, the overall aims of the study were to describe variation in HIV/AIDS restriction genes in children born to HIV-infected mothers and to determine the possible role of the genetic variants in HIV infection and neurocognitive function.

1.10.2 Objectives

In order to realise our aims, the following objectives were set;

- To genotype for genetic variants in *CCR2, CX3CR1, RANTES, SDF1, APOBEC3G* and *MBL2* genes among HIV-exposed infected (EI), HIV-exposed uninfected (EU) and healthy unexposed (UEUI) children
- To determine association between genetic variants and HIV infection by comparing genotype and allele frequencies between HIV EI and EU children
- To investigate association between genetic variation and neurocognitive function
- To correlate genotypes with phenotypes such as CD4+ T-cell count, haemoglobin and body mass index

1.10.3 Rationale

There is a large body of information linking genetic polymorphism to HIV outcomes but most studies have been done on Caucasian populations with limited information available on Africans. The distribution of antiviral gene variants significantly differs between populations and conclusions cannot be made based on Western studies alone. In addition, the distribution of HIV subtypes differs among populations with Europe and America having subtype B predominantly whilst subtype C is predominant in Africa (Buonaguro et al. 2007). The virulence of HIV subtypes is not the same as evidenced by the difference in HIV prevalence across the world. Given this it is important to investigate the distribution and effects of HIV/AIDS restriction genes in the African populations.

HIV is transmitted through sexual contact mainly, thus, research has been prioritised in adults compared to children. However genetic variation seems to play a significant role in paediatric HIV infection. Up to 60% of children perinatally exposed to HIV are born uninfected without any intervention suggesting that inter-individual differences play a part in evading infection (De Cock et al. 2000). We therefore hypothesise that if certain allelic variants are protective against HIV infection then their frequencies would be higher among the HIV exposed but uninfected children compared to the exposed and infected.

Identifying genetic variants that influence HIV transmission and disease progression can help predict disease course and guide therapy. The information will also provide new therapeutic targets for vaccine and drug development. With current ARVs threatened by drug resistance and adverse side effects, new antiviral therapeutics are a priority and some promising targets include the use of co-receptor gene mutations such as evidenced by the CCR5 antagonist drugs and the prospects of CCR5 delta 32 zinc finger gene therapy (Fätkenheuer et al. 2005).

In this ART era, children born infected with HIV have a better chance of surviving into adolescence and adulthood where they enrol into different academic and intellectual curricula. Knowledge of the relationship between genetic factors and neurocognitive development means a forecast of who is vulnerable can help policy makers and health providers to provide appropriate care and guide the children into possible careers. A lot of work has been done on the effect of genetic variants on HIV disease progression but little is known about how the genes affect neurocognitive development yet cognitive deficits are the common among HIV infected people.

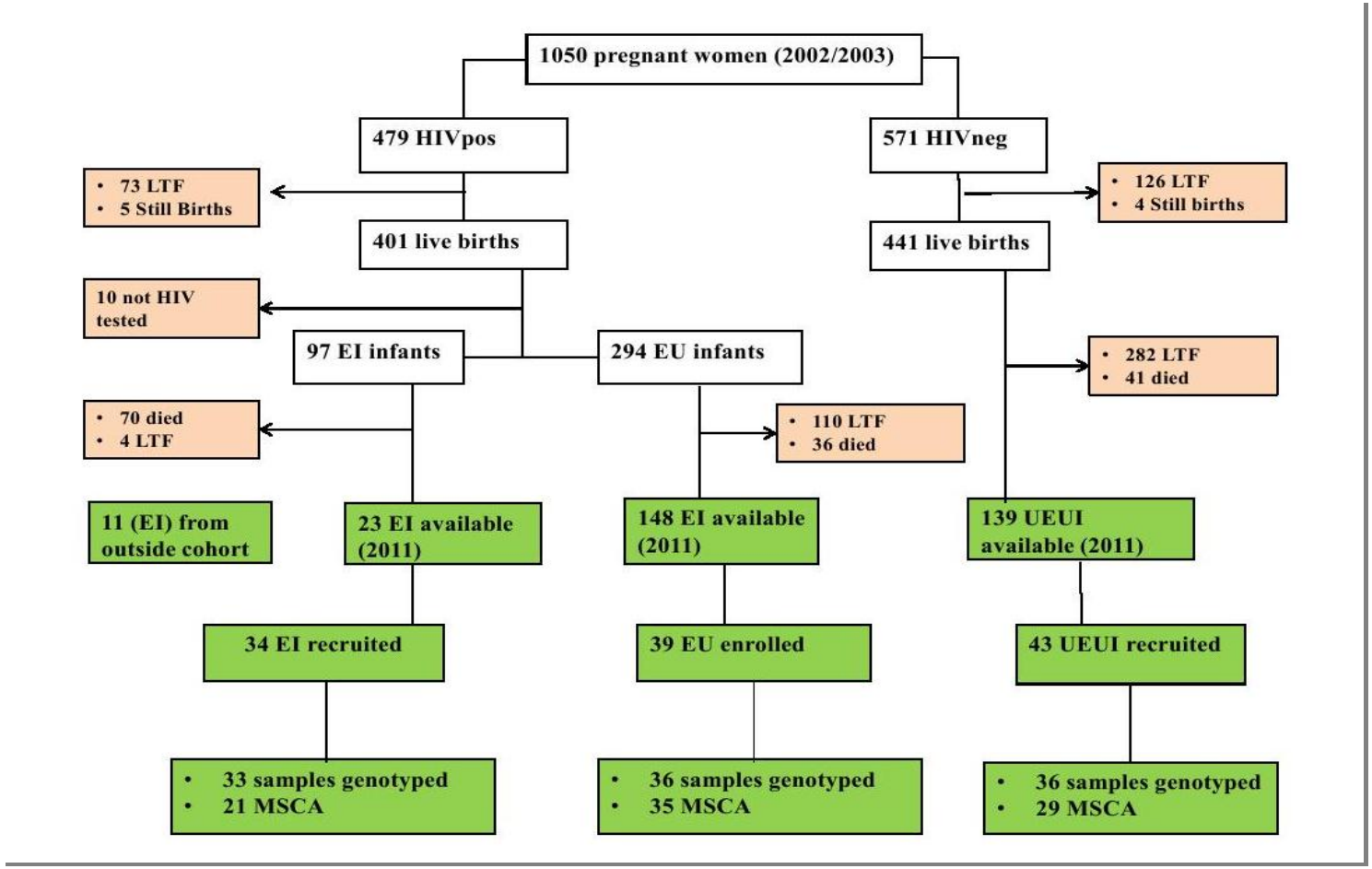
2 CHAPTER TWO: MATERIALS AND METHODS

2.1 Study participants and sample collection

2.1.1 Historical perspective of cohort

Participants were recruited from the Better Health for The African Mother and Child (BHAMC) cohort (Kurewa et al. 2011). This is a longitudinal study of mother-baby pairs who have been followed up since 2002 at three Harare peri-urban clinics in Epworth, St Mary′s Chitungwiza and Seke North. The history of the cohort and recruitment of participants for the current study is shown in Figure 2.1. The cohort started with 1050 pregnant women presenting for antenatal care who were recruited into a Prevention of Mother-to-Child Transmission (PMTCT) program at 36 weeks of gestation. 479 (46%) tested positive for HIV whilst 571 (54%) were HIV seronegative on enrolment into study. The HIV-infected mothers and their infants were given 200mg and 2mg/kg single dose nevirapine (sdNVP) respectively within 72 hours of delivery as prophylaxis to prevent infection of the child.

Of the 571 HIV-negative mothers, 441 (77%) gave birth to live babies, 126 (22%) were lost to follow up and four (1%) had still births. On the other hand, the 479 HIV seropositive mothers gave birth to 401(84%) live infants, 73 (15%) were lost to follow up and five had still births (1%). After five years of follow-up, 391(97%) of the HIV-exposed children had been screened for HIV at least once with 97 (25%) of them testing positive (Kurewa et al. 2011). A total of 23 (23%) of the HIV-infected children died within the first year of life and at five years of follow-up, more than 60% of them had died.



Key: LTF- loss to follow-up, EI-exposed infected, EU- exposed uninfected, UEUI-unexposed uninfected, MSCA-McCarty Scale of Children's Abilities

Figure 2.1: Schematic diagram of the BHMAC cohort from 2002 to 2011. The green boxes show how participants were recruited for the current study.

2.1.2 Current study participants

A total of 116 children who were 7-9 years of age were recruited for this sub-study. These included 34 children perinatally-infected with HIV (EI) who were on ART for at least six months prior to sample collection and officially enrolled at a primary school and 82 healthy HIV-uninfected children (EU+UEUI) of the same age who were recruited as controls. The HIV-uninfected children fell into two groups, one comprising of 39 born to HIV-infected mothers (exposed uninfected, EU) and the other of 43 born to uninfected mothers (UEUI). All participants were of indigenous black African origin. The study received ethics clearance from the Medical Research Council of Zimbabwe and the Ethics Research Committee of Faculty of Health Sciences, UCT. Written consent was obtained from the parents/guardian of each child before a sample could be collected. A 2-5ml EDTA blood sample was collected from each child and transported to the laboratory for CD4+ T-cell count within six hours of collection. The remaining sample was stored at -80˚C for genotyping purposes.

A questionnaire was administered to the parent/guardian to determine each child's demographic information and medical history for the three months prior to the data collection. Occurrence of any of the common morbidities such as fever, diarrhoea, cough, ear discharge and vomiting at least once in the three months preceding sample collection was recorded. Neurocognitive function was tested using the McCarthy Scale of Children's Abilities (MSCA) as part of a bigger study titled; Neurodevelopment and Growth among Peri-urban Zimbabwean School Age Children from a High Prevalence HIV Community. The MSCA tool was translated into Shona language with help of a linguist and then administered to each child by a trained nurse in the presence of a paediatrician.

The MSCA tested three attributes of cognitive function (verbal, perceptive-performance and quantitative) and a general cognitive ability which reflected the overall cognitive function of the child. The tool assesses performance of each child on assigned tasks and gives a score from which a cognitive index scores can be calculated. The verbal aspect assesses cognitive tasks related to verbal information processing whilst perceptive-performance assesses information processing during manual tasks. The general cognitive index score takes into account all aspects of cognitive function including memory, gross and fine motor skills (Guxens et al. 2009). For analysis, cut-off points were set at 30 for verbal, perceptive-performance and quantitative cognitive index scores. The cut-off for the general cognitive index score was 68. The children were categorised into cognitively unimpaired and impaired if they fell above and below cut-off, respectively for each attribute. Genotype frequencies were compared between the groups of neuro-cognitively impaired and the normal to determine association between genetic variation and neurocognitive function in HIV infected or uninfected children.

2.2 Preparation of biological samples for genotyping

2.2.1 Extraction of genomic DNA from whole blood

Genomic DNA was extracted from blood using the Nucleospin® Blood L kit (Macherey-Nagel, Germany) according to the manufacturer's instructions. In summary, 2ml of blood was mixed 150µl of enzyme proteinase K in a Tris-hydrochloride buffered reaction containing chaotropic ions. The mixture was vigorously vortexed to lyse blood cells before incubation at 56˚C for 15 minutes (min) to allow proteinase K to digest the proteins in the lysate. 2ml of absolute ethanol was then added to the lysate to facilitate binding of DNA released from white blood cells to a silica membrane (Nucleospin® Blood columns).

The chaotropic ions together with ethanol destabilise the remaining protein contaminants making them water insoluble. This allows the DNA to bind to the membrane whilst ethanol washes away all the destabilised macromolecules. The binding stage was followed by two ethanol washes to remove any extra contaminants. The silica membrane was transferred to a clean tube into which the silica membrane bound DNA was eluted using a 200µl of 5mM alkaline Tris-Ethyleneiaminetetraacetic acid (Tris-EDTA) buffer. The EDTA chelates calcium ions to prevent any residual DNase enzyme from degrading the purified DNA. Before the DNA was stored or used for genotyping, it was checked for quality and quantified as described in the next section.

2.2.2 Determination of DNA integrity

To check integrity of the extracted genomic DNA, 5µl of the eluted DNA solution was electrophoresed on 1% agarose gel stained with ethidium bromide (EtBr) for 1 hour at 100V in Tris-Borate EDTA (TBE) buffer. Agarose gel electrophoresis is a method of separating DNA fragments according to their size using direct electric current. Agarose is a polysaccharide whose chains bind non-covalently creating a gel with sieving pores of different sizes depending on the concentration of the agarose. Since DNA has a negative phosphate backbone and uniform mass/charge ratio it migrates towards the positive charge at rates governed by the size of fragments (Helling et al. 1974). For visualisation of the DNA fragments, the gel is stained with a nucleic dye that intercalates between the bases of the DNA such as ethidium bromide (EtBr) or SYBR Green. These dyes emit fluorescent light which is visible under ultra-violet (UV) light. To approximate the sizes of the DNA fragments, a molecular weight marker which is a mixture of DNA fragments of known sizes is run simultaneously with the DNA to be sized.

Good quality DNA shows clear and clean band with no smearing after electrophoresis. After confirming the quality of DNA, it is important to measure the amount of DNA in each sample so that a known amount can used for subsequent experiments.

2.2.3 Quantification of isolated genomic DNA

Genomic DNA was quantified using spectrophotometry (NanoDrop® ND-500) by measuring absorbance of the DNA at 230nm, 260nm and 280nm wavelengths. All nucleic acids, RNA, DNA, single stranded or double stranded absorb light at 260nm but at different rates. Calculation of DNA concentration is based on the equation that one optical density at 260nm (OD_{260}) equals 50ng/µl of double stranded DNA (Gallagher et al. 1989). The 260/280 and 260/230 ratios are used to check for presence of contaminants in the DNA solution. Values of 1.8 and 2.0-2.22 for OD_{260}/OD_{280} and OD_{260}/OD_{230} ratios respectively indicate high DNA purity. Protein and phenol absorb light at 280nm wavelength whilst carbohydrates, phenol and residual guanidine from DNA extraction absorb at 230nm (Gallagher et al. 1989). Presence of contaminants lowers the expected ratios. The concentration of each DNA sample was adjusted to a working concentration of 50ng/µl using sterile DNA free water. The working solutions were then used in PCR based genotyping methods.

2.3 Genotyping for single nucleotide polymorphisms

2.3.1 Primer designing

The success of PCR based genotyping is dependent on use of primers that achieve high specificity and yield during amplification. To achieve this careful designing of primers is a pre-requisite and several bioinformatics tools have been designed for this purpose. PCR primers were designed using the Primer 3 online package http://frodo.wi.mit.edu/.

Primer 3 allows one to input DNA sequence spanning a region of interest, input the desired specifications for the primer and the program selects pairs of potentially useful primers. Preferred primers were of length 18 to 21 bases, with guanine-cytidine (GC) content between 40 to 60% and melting temperatures (T_m) ranging from 57˚C to 63˚C. Primer pairs obtained from Primer 3 were analysed using additional online bioinformatics softwares, Primer Blast (NCBI) and OligoAnalyzer SciTool (Integrated DNA Technologies®; http://eu.idtdna.com/analyzer/Applications/OligoAnalyzer/). The Primer Blast package was used to confirm if the gene region amplified by a selected pair of primers actually flanked the region intended.

Furthermore, Primer Blast shows the length of the fragment to be amplified, melting temperatures of the primers and if the primer pair will also amplify any undesired non-specific genomic regions. OligoAnalyzer was primarily used to determine possible formation of secondary structures such as hairpins, self-dimers and hetero-dimers within and between the primer pairs. If secondary structures detected formed bonds requiring total enthalpy change (ΔG) of more than 5kcal/mole to be broken then the primers were excluded. Hairpin structures containing three or more C-G bonds or requiring temperatures above 40 ˚C to melt them were also excluded. PCR primer designing however differed slightly from that of SNaPshot primers.

SNaPshot primers have to anneal immediately adjacent to the polymorphic position to be genotyped, therefore, only the sense or antisense sequences immediate of that nucleotide position can be used. If the adjacent DNA sequences made primers with undesirable secondary structures then bases involved in the secondary structure formation were substituted with a different one to eliminate the bonds.

Sense primers were preferred but where they remained unfavourable after base adjustment, the antisense sequences were opted for. The primers designed were used in PCR to amplify DNA fragments for downstream genotyping techniques such as restriction fragment length polymorphism and DNA sequencing.

2.3.2 Polymerase Chain Reaction

PCR was used to amplify target DNA regions on the genomic DNA template. PCRs were optimised to achieve high specificity and yield by adjusting annealing temperatures and magnesium ions (Mg^{2+}) concentration. A range of annealing temperatures between 50-65˚C depending on melting temperature of primers, were tested in a temperature gradient to determine the most optimal for each reaction. The annealing temperature achieving clean and sharp DNA bands on agarose gel was used in the PCR. If the most optimal annealing temperature showed non-specificity or poor yield, the concentration of Mg^{2+} was titrated to further optimise the reaction. The primer extension time and number of thermal cycles were also considered as part of optimisation. Dimethyl sulfoxide (DMSO) was used to enhance PCR where reaction products remained non-specific and/or with low yield after thermal cycling conditions adjustment. Once optimal PCR conditions were achieved, DNA fragments were amplified in preparation for PCR based genotyping methods such as restriction fragment polymorphism and DNA sequencing. The optimal cycling conditions for all PCRs done in this study are listed in Table 2.2 and all PCR reactions were confirmed by electrophoresing 5µl of the product on 2% agarose gel stained with ethidium bromide at 100V for an hour.

2.3.3 Restriction Fragment Length Polymorphism (RFLP)

Restriction fragment polymorphism is a genotyping method that involves cleavage of DNA fragments at target sites using restriction enzymes that recognise specific sequences called restriction sites. Fragments generated by the enzyme after incubation with the DNA are separated according to size using gel electrophoresis. DNA has a net negative charge due to phosphate groups. DNA therefore migrates towards the positive charge under electrical current in electrophoresis. The digested DNA forms banding patterns that are consistent with the positions that are targeted for restriction or digestion by the restriction enzyme. Polymorphisms either create or abolish restriction enzyme sites and this is observed from the pattern on the gel. A molecular weight maker is included during the electrophoresis so that the DNA fragment sizes may be approximated. Restriction sites on DNA fragments analysed in this study were confirmed using an online package called Nebcutter V2.0 (New England Bio systems®). In this study, the following SNPs were genotyped using PCR-RFLP; *CCR2* 190G>A, *CX3CR1* 745G>A, *CX3CR1* 839C>T, *SDF1* 801G>A, *RANTES* In1.1T>C, *APOBEC3G* 557>G, *APOBEC3G* -90C<G and *MBL2* 170G>A. In addition to RFLP, genotyping was also done using DNA sequencing (Sanger) and primer extension mini-sequencing (SNaPshot®).

2.3.4 Genotyping by Sanger chain termination sequencing

DNA sequencing is a genotyping method that involves determination of linear arrangement of nucleotides on a DNA fragment. In the Sanger chain termination sequencing, template DNA to be sequenced is mixed in a reaction with deoxy-nucleotide triphosphates (dNTPs), labelled dideoxy-nucleotide triphosphates (ddNTPs), DNA polymerase enzyme and a primer that anneals on the 5′ end of the strand to be sequenced.

The primer anneals to the template and polymerase enzyme adds dNTPs to the primer's 3′ end in a process that extends the primer. The ddNTPs in the reaction are also randomly added to extending strands but ddNTPs have a hydrogen (–H) instead of a hydroxyl group (–OH) on their 3′ end and therefore cannot form phosphodiester bonds with an incoming dNTP. The addition of a ddNTP to the strand therefore terminates the extension. Because dNTPs and ddNTP are added randomly to extending DNA strands, the process generates a pool of terminated fragments of varying lengths (Sanger 1977, Sanger 1980). Each of the four bases (G, C, T and A) on the ddNTPs is labelled with a different colour dye. The terminated DNA strands are then separated according to size using capillary electrophoresis for analysis giving a chain of coloured peaks representative of nucleotide bases at each position. Sanger's chain termination DNA sequencing technique was used to genotype variants in the promoter and exon 1 regions of the *MBL2* gene. The principle of Sanger chain termination sequencing has been adapted in another genotyping method called minisequencing.

2.3.5 Genotyping by primer extension minisequencing (SNaPshot®)

Primer extension mini-sequencing is a SNP detection method based on the annealing of a primer adjacent of the 5′ of the nucleotide to be identified on a DNA template. A single labelled dideoxy nucleoside triphosphate (ddNTP) is then added to the primer in a process catalysed by DNA polymerase enzyme (Syvänen et al. 1990). The ddNTP lacks the hydroxyl group needed for formation of the phosphodiester linkage, thus, its incorporation terminates the extension hence the term "minisequencing". SNaPshot® primer extension minisequencing is patented by Applied Biosystems. In SNaPshot®, several SNPs can be genotyped simultaneously using different primer sizes and ddNTPs labelled with different fluorescent dyes (Table 2.1).

The extended primers are separated according to size using automated capillary electrophoresis (ABI PRISM® 3130 Genetic Analysers) and the SNPs can be identified by size on the electrophoregram whilst the fluorescent dye on the incorporated ddNTP identifies the nucleotide base. To determine the genotypes for all SNaPshot reactions in this study, 5µl of SNaPshot products were mixed with 0.3µl GeneScan™-120 LIZ™ size standards and 4.7µl Hi-Di™ formamide and analysed using capillary electrophoresis on the ABI PRISM® 3130 Genetic Analyser. The output of the electrophoresis was analysed using the GeneMapper® v4 software. The program generates an electrophoregram where alleles appear as peaks of different colours and size depending on the base of the nucleotide added and length of primer used in the SNaPshot. SNPs, *RANTES* -403G>A, *APOBEC3G* (-571G>C, 197T>C and 199C>G) and *MBL2* (-595G>A, -550C>G, -221C>G and +4C>T) were investigated using the SNaPshot genotyping technique in the current study.

Table 2.1: List of fluorescent dyes used in SNaPshot and their corresponding colours

ddNTP	Dye label	Colour of analysed data
A	dR6G	Green
C	dTAMRA™	Black
G	dR110	Blue
T	dROX™	Red

2.4 Genotyping of chemokine receptor genes *CCR2* and *CX3CR1*

Genotyping of CCR2 variants was done using PCR based RFLP. Table 2.2 is a summary of the primers, cyclic conditions and enzyme restriction methods used in all RFLP assays in this study. The CCR2 190G>A (rs1799864) SNP was genotyped according to a method described elsewhere (Martinson et al., 2000) but with minor modifications. To determine *CCR2* 190G>A alleles, a 337 base pair (bp) fragment flanking the SNP was amplified in a 25µL volume containing 10 picomoles of each of the primers (Table2.2), 200µM dNTP mix, 1X buffer (pH 8.3), 1.5mM Mg^{2+}, 1U Taq polymerase and 50ng genomic DNA. The cycling conditions were as shown in Table 2.2. A 10µl volume of the PCR product was digested 16 hoursusing 2U of *BseG*I (*Fok*I) restriction enzyme and electrophosed on 2% agarose gel stained with EtBr at 120V for 1h. The *CCR2* 190A allele is susceptible to *BseG*I restriction activity, thus the PCR product is cleaved to give two fragments of 228bp and 109bp whilst *CCR2* 190G allele is resistant to *BseG*I activity leaving the 337bp PCR product undigested. The structure of the CCR2 gene and position of the 190G>A SNP are shown in Figure 2.2.

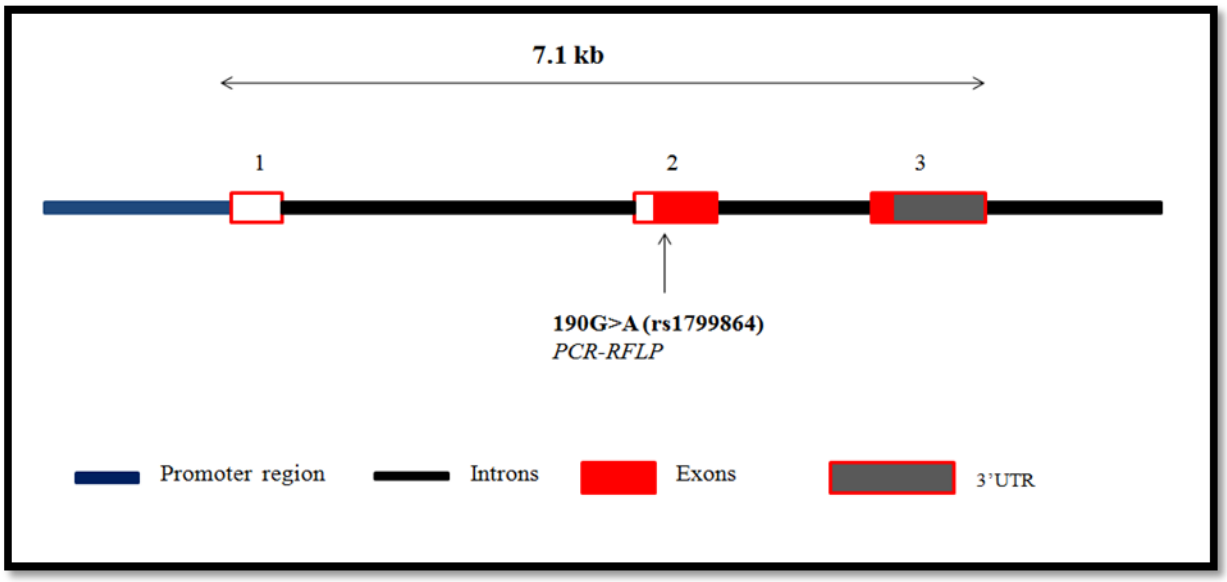


Figure 2.2: Schematic representation of the *CCR2* gene

In addition to *CCR2*, another chemokine receptor gene *CX3CR1* also formed part of this study. Two SNPs in *CX3CR1* 745G>A (rs3732379) and *CX3CR1* 839C>T (rs3732378) were genotyped using PCR-RFLP (Table 2.2). The positions of the two SNPs on the gene are shown on the schematic diagram of *CX3CR1* shown in Figure 2.3. The 745G>A genotyping was carried out according to the method of Nasar et al. (2008) with minor modifications (Nassar et al. 2008). A 311bp gene region flanking the 745G>A position was amplified in a reaction mixture consisting of 10 picomoles of forward and reverse primers (Table 2.2), 200µM dNTP mixture, 1X buffer (pH 8.3), 1.5mM Mg^{2+}, 1U of Taq polymerase enzyme and 50ng genomic DNA in a 25µl volume. The thermal-cycling conditions are shown in Table 2.2. To determine the genotypes, 10µl PCR fragment were digested using 5U of *Psp*1406I *(AclI)* restriction enzyme. For successful digestion, 3U of enzyme were added at beginning of reaction and 2U more supplemented three hours into the 16h digestion to ensure complete digestion. In the presence of the 745G allele, the enzyme cleaves the PCR product into two bands of 204bp and 107bp. Presence of the A allele disrupts the cleavage site leaving the 311bp PCR fragment undigested.

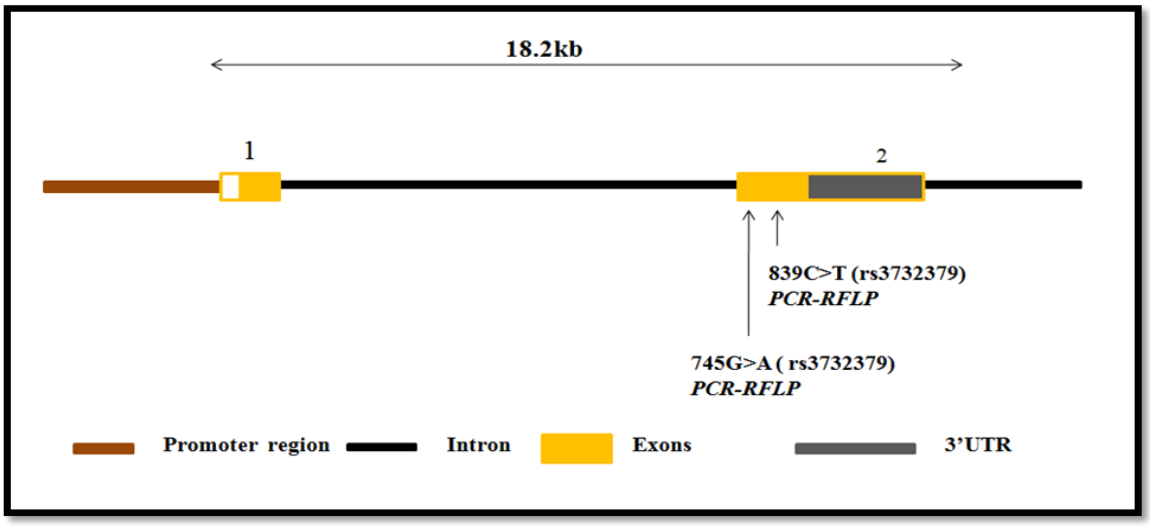


Figure 2.3: Schematic representation of the *CX3CR1* gene showing the position of the two SNPs investigated.

Table 2.2: PCR-RFLP conditions for all SNPs investigated using the genotyping method

SNP	Primer (5′-3′)	*Cycling conditions	RE	RE site	Product size	Definitive fragments
***CCR2* 190G>A rs17141036**	F-GGATTGAACAAGGACGCATT R-GAGCCCACAATGGGAGAGTA	94˚C-30s, 60˚C-30s, 72˚C-30s, 72˚C-5min	*Bse*GI	5′...GGATGNN↓N...3′ 3′...CCTAC↑NNN...5′	337bp	G-337 A-228, 109
***CX3CR1* 745G>A rs3732379**	F-AGAATCATCCAGACGCTGTTTTCC R-CACAGGACAGCCAGGCATTTCC	94˚C-30s, 60˚C-30s, 72˚C-30s, 72˚C-5min	*Psp*1406I	5′...AA↓CGTT...3′ 3′...TTGC↑AA...5′	311bp	G-204, 107 A-311
***CX3CR1* 839C>T rs3732378**	F-CATCCAGACGCTGTTTTCCT R-TGCTCAGAACACTTCCATGC	94˚C-30s, 55˚C-30s, 72˚C-30s, 72˚C-5min	*Bsm*BI	5′...GAATGCN↓N...3′ 3′...CTTAC↑GNN...5′	355bp	C-189, 113, 75 T-113, 264
***SDF1* 801G>A rs1801157**	F-CAGTCAACCTGGGCAAAGCC R-AGCTTTGGTCCTGAGAGTCC	94˚C-30s, 62˚C-30s, 72˚C-30s, 72˚C-5min	*Hpa*II	5′...C↓CGG...3′ 3′...GGC↑C...5′	302bp	G-202, 100 A-302
***RANTES* In1.1T>C rs2280789**	F-CCTGGTCCTGACCACCACA R-GCTGACAGGCATGAGTCAGA	94˚C-30s, 62˚C-30s, 72˚C-30s, 72˚C-5min	*Mbo*II	5′...GAAGANNNNNNNN↓N...3′ 3′...CTTCTNNNNNNN↑NN...5′	343bp	T-343 C-225, 118
***A3G* 557A>G rs8177832**	F-ACCTGTGGGTCTGCTCTGAT R-CAGGAGGGAAGGCAGGAG	94˚C-30s, 60˚C-30s, 72˚C-30s, 72˚C-5min	*Hha*I	5′...GCG↓C...3′ 3′...C↑GCG...5′	409bp	G-194, 145, 70 A-339, 70
***A3G* -90C>G rs5750743**	F-TGTCTGTCTTGATGGTGGAGAG R-CAAGCTTCTGGTTTTAGGACA	94˚C-30s, 55˚C-30s, 72˚C-30s, 72˚C-5min	*Ava*I	5′...C↓YCGRG...3′ 3′...GRGCY↑C...5′	355bp	C-259, 96 G-355
***MBL2* 170G>A rs1800451**	F-TGGCAGCGTCTTACTCAGAA R-TCCTCATATCCCCAGGCAGT	94˚C-30s, 57˚C-30s, 72˚C-30s, 72˚C-5min	*Mbo*II	5′...GAAGANNNNNNN↓N...3′ 3′...CTTCTNNNNNNN↑NN...5′	213bp	G-213 A-138, 75

Key: RE- Restriction enzyme, *A3G- APOBEC3G*

*Cycling conditions in order, denaturation, primer annealing, primer extension, final extension.

*All cycling conditions started with an initial denaturation step at 94˚C for 5min which is not shown.

*All PCRs involved 35 cycles except *APOBEC3G* 557A>G which had 40.

To genotype *CX3CR1* 839C>T variants, 10 picomoles of each of the primers sense and antisense (Table 2.2) were mixed with 200µM dNTP mixture, 1X PCR buffer (pH 8.3), 1.5mM $MgCl^{2+}$, 1U of Taq polymerase enzyme and 50ng genomic DNA in a 25µl volume PCR mix to amplify a 377bp fragment. A 10µl volume of the PCR product was digested using 6U *of BsmB*I restriction enzyme for 16 hours at 37˚C. The digestion products were electrophoresed on 2% agarose gel for 1 hour at 100V in the presence of EtBr as the DNA staining chemical. In the presence of the 839C allele, the enzyme has two restriction sites resulting in three fragments of 189bp, 113bp and 75bp whilst the 839T allele disrupts one of the restriction sites giving two fragments 113 and 264bp long.

2.5 Genotyping of chemokine genes *SDF1* and *RANTES*

The genotyping of *SDF1* 801G>A (rs1801157) was done as described elsewhere (Balotta et al. 1999) but with minor modifications. 10 picomoles of each of forward and reverse primers (Table 2.2) were added to a PCR reaction mixture consisting of, 200µM dNTP mixture, 1 X PCR buffer, 1.0mM MgCl2, 5% DMSO, 1U of Taq polymerase enzyme and 50ng genomic DNA in a 25µl reaction volume. The cycling conditions were as shown in Table 2.2. A 302bp fragment was the PCR product of interest confirmed by electrophoresis of 5µl of the PCR product on 2% agarose gel with EtBr as DNA staining chemical for an hour at 100V. 10µl of PCR product were then digested 16 hours using 2U *Msp*I (*Hpa*II) restriction enzyme to determine the genotypes. The products of the digestion were run on 2% agarose gel containing EtBr stain for 1h at 100V to separate them according to size. The 801G allele is digested into two fragments of 202 and 100bp sizes. The A variant disrupts the *Msp*I restriction site leaving the 302bp fragment undigested. Figure 2.4 shows a schematic of the SDF1 gene structure.

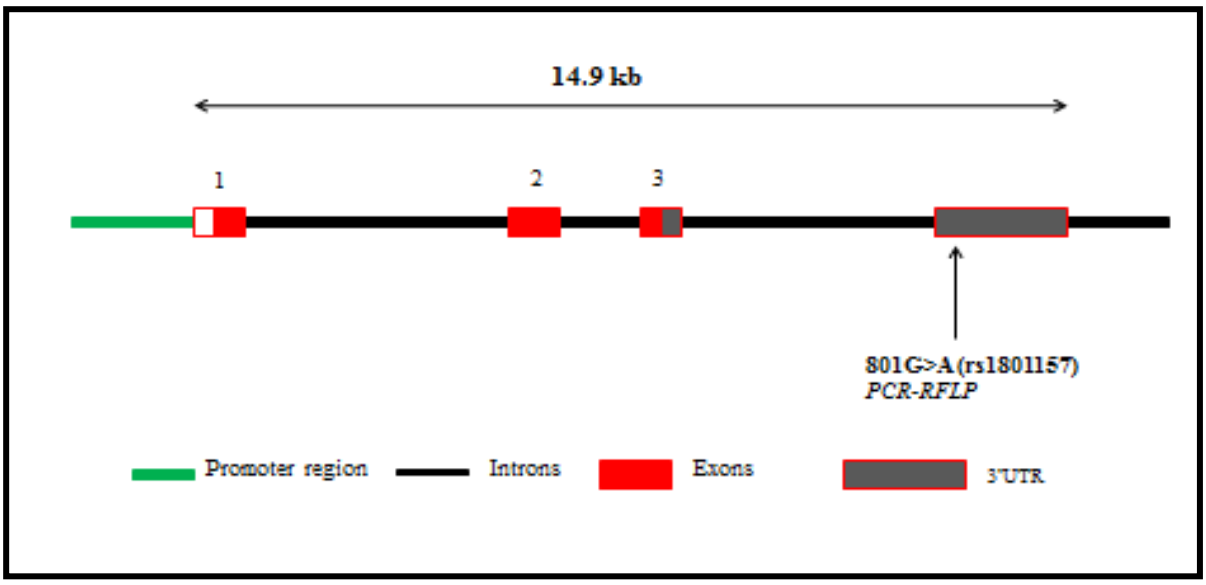


Figure 2.4: Schematic representation of the *SDF1* gene structure showing the position of the 801G>A polymorphism

Two SNPs, *RANTES* In1.1T>C and *RANTES* -403G>A (Figure 2.5) were studied using PCR-RFLP and SNaPshot® respectively. In1.1T>C genotyping was done according to a method described by elsewhere but with a few modifications (Qian et al. 2008). A 343bp DNA region incorporating the In1.1T>C polymorphism was amplified using a PCR reaction mixture comprising of 10 picomoles of the each primer, 200µl of dNTP mixture, 1X PCR buffer, 1.0 mM Mg2+, 1U of Taq polymerase enzyme and 50ng genomic DNA in a 25µl volume. A 10µl volume of the PCR products was then digested using 3U of *Mbo*II restriction enzyme by incubating at 37˚C overnight. The digestion products were electrophoresed on 2% agarose gel stained with EtBr for 1 hour at 120V in TBE buffer. The 343bp PCR product remains undigested in the presence of In1.1T allele but the In1.1C allele introduces an *Mbo*II restriction site yielding two fragments, 225bp and 118bp.

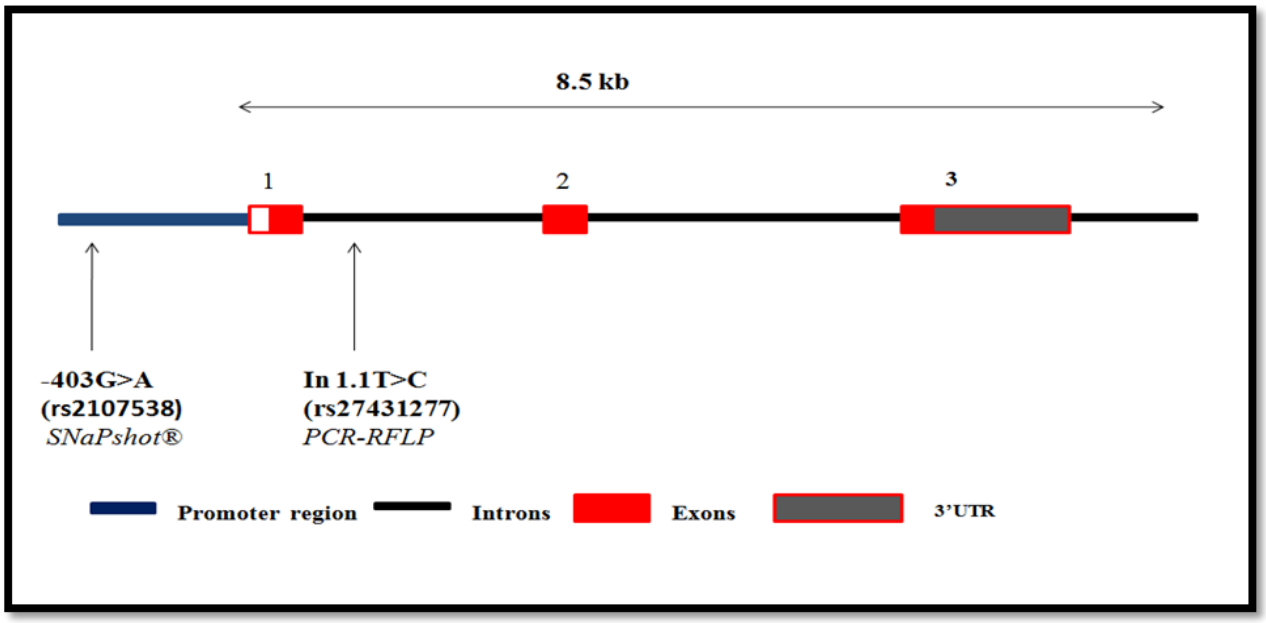


Figure 2.5: Schematic representation of the *RANTES* gene structure showing the position of the two SNPs investigated

PCR based SNaPshot was the method of choice for *RANTES* -403G>A genotyping. The primers, PCR and SNaPshot cycling conditions for all SNaPshot reactions done in this study are shown in Table 2.3. The PCR reaction used to amplify a 527bp fragment flanking the *RANTES* -403G>A SNP consisted of 10 picomoles of each of the forward and reverse primers, 200µM dNTP mixture, 1X PCR buffer, 1.5mM $MgCl^{2+}$, 1U of Taq polymerase enzyme and 50ng genomic DNA in a 25µl volume. In preparation for SNaPshot, the PCR products were purified using exonuclease (Fermentas®) and Thermo sensitive Alkaline Phosphatase enzymes (FastAP™). The exonuclease (exo) digests excess primer whilst the FastAP™ neutralises extra dNTPs by removing the phosphate group. For each purification reaction, 5µl purified PCR product were mixed with 1U FastAP™ and 2U exo enzymes in a 10µl volume and incubated for 1 hour at 37˚C before deactivating the enzyme by heating at 75˚C for 15 min. The enzymes were then heat-inactivated at 75˚C for 15 min.

The SNaPshot reaction mixture was made up of 20 picomoles of primer (Table 2.3), 1µl purified PCR product, 1µl of the SNaPshot multiplex ready reaction mix and made up to 10µl volume with DNA-free water. The reaction mixture was incubated for 25 cycles of; denaturation at 96˚C for 10s, primer annealing at 50˚C for 5s and primer extension at 60˚C for 30s. Excess fluorescently labelled ddNTPs were neutralised by digesting the reaction products with 1U FastAP™ at 37˚C for 1 hour before heat-inactivating the enzyme at 75˚C for 15 min. The SNaPshot products were then analysed on the ABI PRISM® 3130 Genetic Analyser as described in section 2.3.5 Blue peak of size 34.2 on the electrophoregram was interpreted as -403G allele while a green peak at the same position represented the -403A allele. In addition to chemokine receptors and chemokines, this study also investigated polymorphism in two genes involved in the innate immune system through pathogen pattern recognition, *APOBEC3G* and *MBL2*.

Table 2.3: Primers and cycling conditions for PCR based SNaPshot genotyping

SNP	Primer (5′-3′)	PCR Cycling conditions (35 cycles)	PCR Product	SNaPshot cycling conditions (25 cycles)
***RANTES* -403G>A rs2107538**	F-CACCTCCTTTGGGGACTGTA R-CGTGCTGTCTTGATCCTCTG **SS**- CCATAGATGAGGGAAAGGAG	94˚C-5m, 94 ˚C-30s, 60 ˚C-30s, 72 ˚C-30s, 72 ˚C-30s	527bp	96˚C-10s, 50˚C-5s, 60˚C- 30s.
***A3G* -571G>C rs5757463**	F-GCAAATGCATCCCTTGTGTA F-CCTCCTCTCCACCATCAAGA **SS**-AATTTGTAGGTCACCACGCCATAGGAACACA CTACCA	94˚C-5m, 94 ˚C-30s, 55 ˚C-30s, 72 ˚C-30s, 72 ˚C-30s	467bp	96˚C-10s, 50˚C-5s, 60˚C- 30s.
***A3G* 197T>C rs3736685**	F-TACCCTGACCATCTTTGTTGC R-ACACGAACTTGCTCCAACAGT **SS**-TGGCACTGACTGTAACTAGTATC	94˚C-5m, 94 ˚C-30s, 60 ˚C-30s, 72 ˚C-30s, 72 ˚C-30s	398bp	96˚C-10s, 50˚C-5s, 60˚C- 30s.
***A3G* 199C>G rs2294367**	F-GTGAGGCCAGGGAAGAAGA R-TGAAAGTGAATGTGGGTGGA **SS**-ACATCCCTTAGAATCTGTCAGAAGAGGTTC CCACTTACTTGCTA	94˚C-5m, 94 ˚C-30s, 60 ˚C-30s, 72 ˚C-30s, 72 ˚C-30s	352bp	96˚C-10s, 50˚C-5s, 60˚C- 30s.
MBL2	F-AGGCTGCTGAGGTTTCTTAGG R-ATGCCAGAGAATGAGAGCTGA	94˚C-5m, 94 ˚C-30s, 64 ˚C-30s, 72 ˚C-30s, 72 ˚C-30s	1187bp	N/A
****MBL2* -595 novel**	**SS**-GGGTCAGCATTTTCTCTGGATATTTC TTACTA	N/A		96˚C-10s, 50˚C-5s, 60˚C- 30s.
****MBL2* -221 rs7096206**	**SS**-GCTGGAAGACTATACACTTGCTTTC	N/A		96˚C-10s, 50˚C-5s, 60˚C- 30s.
****MBL2* +4C>T rs7095891**	**SS**-CAGGGAAGGTTAATCTCAGTTAATGAA CAC ATATTT ACC	N/A		96˚C-10s, 50˚C-5s, 60˚C- 30s.

Key: *A3G-APOBEC3G*, SS- SNaPshot primer, N/A-not applicable

*SNaPshot primers only are shown because all SNPs are on the same fragment amplified by the primer pair named *MBL2*.

2.6 Genotyping of innate immune factor genes

2.6.1 Genotyping of *APOBEC3G* gene variants

Five SNPs were investigated in the *APOBEC3G* gene; two in the promoter region (-571G>C; rs5757463 and -90C>G; rs5750743), one in exon 4 (557A>G; rs8177832) and two in introns (197T>C/ rs3736685 and 199C>G/ rs2294367). The five SNPs are shown in the schematic diagram of the gene in Figure 2.6. Variation at 557A>G and -90C>G were genotyped using PCR-RFLP whilst SNaPshot was used to genotype the remaining three.

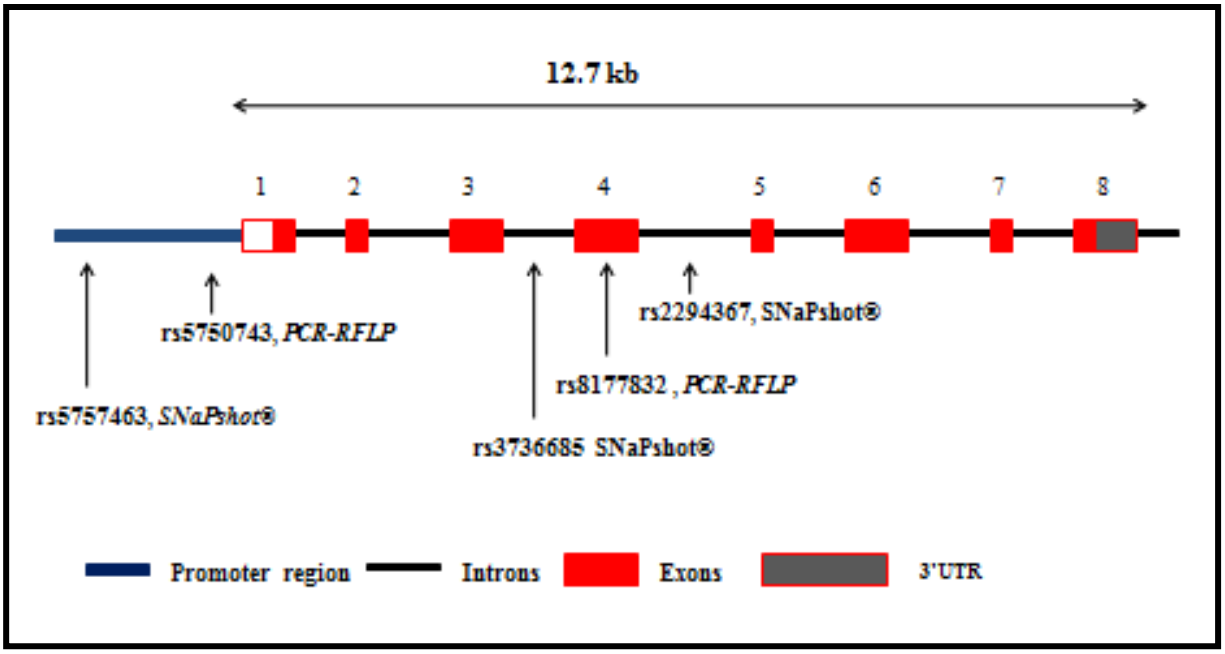


Figure 2.6: Schematic representation of the *APOBEC3G* gene showing all the SNPs studied

The *APOBEC3G* 557A>G polymorphism in exon 4 of *APOBEC3G* gene was determined using a method previously described elsewhere but with minor modifications (An et al. 2004). A 409bp DNA fragment flanking the 557A>G site was generated in a 25µl PCR reaction mixture containing 10 picomoles of each primer, 200µM of dNTP mixture, 1X PCR buffer, 0.5mM Mg^{2+}, 5% DMSO, 1U Taq polymerase and 50ng genomic DNA.

The PCR cycling conditions are shown in Table 2.2. In the final step of the genotyping procedure, a 10µl volume of the PCR product was subsequently digested using 2U of *Hha*I restriction enzyme by incubating 16 hoursat 37˚C. The G allele has two restriction sites resulting in the generation of three fragments; 194bp, 145bp and 70bp. The A allele results in the loss of one restriction site, thus, after digestion with *Hha*I only two bands, 339bp and 70bp are observed.

Genotyping for *APOBEC3G*-90C>G variants involved the amplification of a 355bp fragment flanking the SNP in a PCR reaction mixture consisting of 10 picomoles of each of the primers (Table 2.2), 200µM dNTPs, 1X buffer (pH 8.3), 1.5mM Mg^{2+}, 1U Taq polymerase and 50ng genomic DNA in a total volume of 25µl. Cycling conditions were as shown in Table 2.2. PCR specificity and quantity was checked by electrophoresing 5µl of the PCR product on 2% agarose gel with EtBr stain at 100V for 1 hour before visualisation under UV light. For allele determination, 10µl of PCR products were incubated with 5U of *Ava*I restriction enzyme at 37˚C 16 hoursin a 20µl volume. The digested DNA was then analysed after electrophoresis on 2% agarose gel stained with EtBr. The enzyme cleaved the C variant into two fragments, 259bp and 96bp whilst the G variant led to the loss of the restriction site, thus, resulting in an undigested PCR product of 355bp fragment. In addition to the two *APOBEC3G* SNPs genotyped by PCR-RFLP, three more were genotyped using SNaPshot.

The PCR based primer extension method was used to genotype three SNPs in the *APOBEC3G* gene; -571G>C, 197T>C and 199C>G. Three sets of sense and antisense primer pairs were designed to amplify three DNA fragments in *APOBEC3G* promoter, intron 3 and intron 4 flanking SNPs, rs5757463G>C, rs3736685T>C and rs2294367C>G respectively (Table 2.3). PCR conditions for amplification of all three fragments were the similar.

Each PCR consisted of 10 picomoles of each of the respective forward and reverse primers, 50ng of genomic DNA, 1X PCR buffer (pH 8.3), 1.5mM Mg^{2+}, 200µM dNTPs mixture, 1U Taq polymerase enzyme and 50ng genomic DNA in a 25µL volume. The thermal cycling conditions are shown in Table 2.3. The PCR products were confirmed by electrophoresing 5µl on 2% agarose gel stained with EtBr at 100V for 1hour before they could be used in SNaPshot reaction. To prepare the PCR products for SNaPshot reaction, excess primers and dNTPs from the PCR were digested by incubating 5µl of each PCR product with 1U FastAP™ and 2U exonuclease enzymes in a 10µl volume at 37˚C for 1 hour.

Genotyping of *APOBEC3G* -571G>C and 199C>G variants was done simultaneously in a multiplex SNaPshot reaction. The reaction comprised of 13.3 picomoles of the -571G>C primer , 6.7 picomoles of 199C>G primer (Table 2.3), 5µl of *APOBEC3G* promoter and exon 4 purified PCR products mixed in a 1:1 ratio and 2µl of SNaPshot® Multiplex Ready Reaction Mixture in a 10µl reaction. Thermal cycling conditions consisted of 25 cycles of; denaturation at 96˚C for 10s, primer annealing at 50˚C for 5s and primer extension at 60˚C for 30s. Unattached ddNTPs from the SNaPshot reaction were inactivated by digesting the products with 1U FastAP™ for 1h before heat inactivating the enzyme at 75 ˚C for 15 min. To determine the -571G>C and 199 alleles, the SNaPshot products were analysed through capillary electrophoresis as described in section 2.3.5. Blue and black peaks at position 40.5 on the electrophoregram represented -571G and -571C alleles respectively. Because an antisense primer was used to genotype 199C>G variants, the blue and black peaks at position 49 represented the 199C and G alleles respectively. Unlike -571G>C and 199C>G SNPs that were genotyped in a multiplex SNaPshot, 197T>C variants were determined in a single-plex SNaPshot.

The *APOBEC3G* 197T>C SNaPshot reaction mixture was made up of 20 picomoles of the primer (Table 2.3), 1µl purified PCR product, 1µl of the SNaPshot® Multiplex Ready Reaction Mixture made up to 10µl volume with DNA-free water. The cycling conditions consisted of 25 cycles of denaturation at 96˚C for 10s, annealing at 50˚C for 5s and primer extension at 60˚C for 30s. Excess ddNTPs from the SNaPshot were digested using 1U of fast AP incubated at 37˚C for 60 and 75˚C for 15min to heat-inactivate the enzyme. To determine the 197T>C alleles, 5µl of the SNaPshot products were mixed with 0.3µl GeneScan™-120 LIZ™ size standard and 4.7µl Hi-Di™ formamide and separated according to size using capillary electrophoresis on the ABI PRISM® 3130 Genetic Analyser 3130. Presence of a red peak on position 35.5 on the electrophoregram represented a 197T allele while a black peak was interpreted as 197C allele.

2.6.2 Genotyping of *MBL2* gene

The expression of the mannose binding lectin (MBL) protein and assembly of its polypeptides into functional units seems to be affected by polymorphism in the promoter and exon 1 regions of the gene encoding the protein *(MBL2).*To investigate variation of previously reported and possible novel SNPs, a 1187bp region of the *MBL2* stretching from promoter to intron 1 was amplified and subsequently sequenced using PCR based Sanger chain termination sequencing. Figure 2.7 is a schematic depiction of *MBL2* showing the region sequenced and positions of SNPs of investigated further.

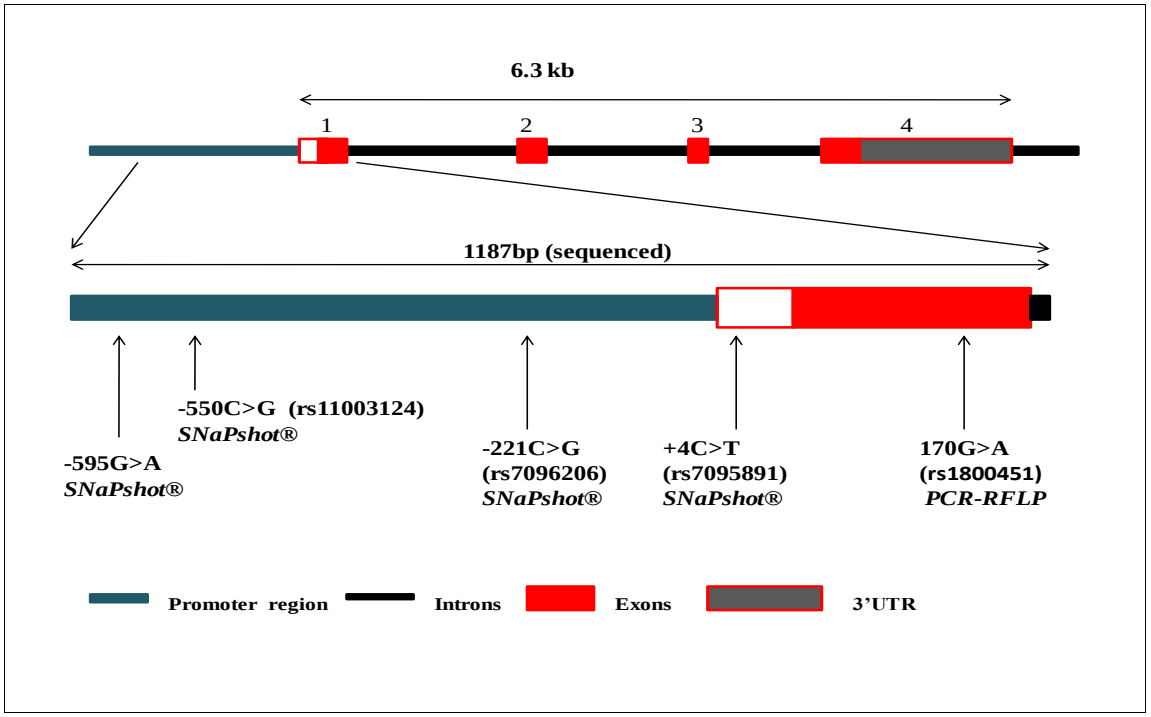


Figure 2.7: Schematic representation of the *MBL2* gene. The top panel shows the full gene structure whilst the bottom one zooms in on the region that was investigated in this study. The arrows indicate the positions of the SNPs that were genotyped and the genotyping methods employed.

Ten picomoles of each of the sense and antisense primers were mixed with; 200µM dNTPs, 1X TE buffer , 1.5mM Mg2+, 1U Taq polymerase and 50ng genomic DNA in a 25µl PCR reaction mixture. The cycling conditions were as follows; initial denaturation at 94˚C for 5min, followed by 35 cycles denaturation , primer annealing and primer extension at 94˚C, 64˚Cand 72˚C respectively for 30s at each step. The reaction was finalised by an extension step of 5 min at 72˚C. The PCR was confirmed by electrophoresis of 5µl of the PCR product on 2% agarose gel containing GR green nucleic acid stain at 100V for 1 hour in TBE buffer. In preparation for the sequencing, 5µl of the PCR products were incubated with1U FastAP™ and 2U exonuclease enzymes in a 20µl volume at 37˚C for 1 hour before heat-inactivating the enzyme at 75˚C for 15 min. To allow the appropriate amount of PCR amplified DNA to be added to subsequent sequencing reaction, the purified PCR product was quantified by measuring absorbance at 260nm wavelength using a spectrophotometer.

The sequencing reaction mixture for the sense strand genotyping consisted of; 10 picomoles of the sense primer, 300ng of the purified PCR product, 2µl of BigDye® Terminator Ready Reaction Mix (containing fluorescently labelled ddNTPs, dNTPs and Taq polymerase) and 4µl BigDye® sequencing buffer (ABI® PRISM) in a 20µl volume. For antisense strand sequencing, a similar reaction mixture was prepared but with the antisense primer in place of the sense primer. Thermal cycling was done as follows; initial denaturation at 98˚C for 5 min followed by 30 cycles of denaturation at 96˚C, primer annealing at 50˚C and extension at 60 ˚C for 30s, 15s and 4 min respectively. The sequenced DNA was then precipitated out of solution using absolute alcohol and 0.3M sodium acetate (pH 5.5) at -20˚C for 2 hours.

The precipitate was pelleted by spinning down at 5000rpm for 10 min and the supernatant decanted. The ethanol in the precipitation disrupts the balance of charge in water resulting in negatively charged DNA binding to positively charged sodium ions provided by the low sodium acetate forcing the DNA to precipitate out of solution. The DNA pellet was washed using 50µl of 70% alcohol to remove excess salts by spinning down in a centrifuge at 5000rpm for 10 min before decanting the supernatant. The DNA pellet was then air dried and resuspended in 10µl DNA free water. A volume of 5µl of the resuspended DNA was mixed with 8µl of HIDI™ (BigDye® Terminator sequencing standard) for capillary electrophoresis of the DNA on the ABI® PRISM 3130 Genetic Analyser.

The sequences were analysed using the Lasergene® 10 software (DNA star). Lasergene® generates an electrophoregram where nucleotide bases are read as C for blue peaks, T for red, G for black and A for the green peaks. The sequences were manually inspected and edited in cases where the software incorrectly called the bases due to blobs, low signal or other artefacts. Alignment with a reference sequence from NCBI databases was used to detect polymorphism on the sequenced fragment.

Sequencing of *MBL2* promoter and exon 1 region served as a screening process for the presence of previously reported SNPs whose allelic variants have been linked with HIV infection or disease progression. Sequencing was also aimed at detection of possible novel variants in the *MBL2* promoter. Using the sequencing results as a guideline for further investigation, three previously reported SNPs and a novel SNPs were genotyped using other methods in samples that had not been sequenced. *MBL2* -595G>A (novel), -221C>G (rs7096206) and +4C>T (rs7095891), where genotyped using SNaPshot while PCR-RFLP was the method of choice for 170G>A (rs1800451) genotyping.

MBL2 -595G>A, -221C>G and +4C>T are located within 600bp of each other, thus, the same 1187bp PCR fragment used for sequencing was used for the SNaPshot. In preparation for SNaPshot, the PCR products were purified using exonuclease and FastAP™ enzymes. The purification reaction mixture contained; 5µl purified PCR product, 1U FastAP™ and 2U exo in a 10µl volume. This was incubated for 1 hour at 37˚C before inactivating the enzymes by heating at 75˚C for 15 min. Each of the three SNPs (-595G>A, -221C>G and +4C>T) was genotyped in a separate single-plex SNaPshot reaction. However, reaction conditions were the same for all. Each SNaPshot reaction mix was made up of 20 picomoles of the respective primer, 1µl purified PCR product and 1µl of the SNaPshot Multiplex Ready Reaction Mix made up to a 10µl volume with DNA-free water. The reaction mixture was incubated for 25 cycles of; denaturation at 96˚C for 10s, primer annealing at 50˚C for 5s and primer extension at 60˚C for 30s. Excess fluorescently labelled ddNTPs from the SNaPshot reaction were neutralised by digesting the reaction products with 1U FastAP at 37˚C for 1 hour before heat-inactivating the enzyme at 75˚C for 15 min.

To determine the nucleotide added to each primer, 5µl of the SNaPshot products were mixed with 0.3µl GeneScan™-120 LIZ™ size standard and 4.7µl Hi-Di™ formamide and separated according to size using capillary electrophoresis on the ABI PRISM® 3130 Genetic Analyser 3130. Output of the electrophoresis is on an electrophoregram generated using the Gene Mapper software. Because all three SNaPshot reactions were carried out using antisense primers the colour peaks seen on the electrophoregram were representative of the reverse complement of the alleles. The -595G and -595A alleles were therefore seen as black and red peaks respectively at position 35.5 on the electrophoregram. Blue and black peaks at 32 represented -221C and -221G alleles respectively whilst blue and green at 41 were read as +4C and +4T alleles.

The *MBL2* 170G>A SNP was genotyped using PCR-RFLP. A 213bp DNA region incorporating the 170G>A polymorphism was amplified in a PCR reaction mixture composed of; 10 picomoles of the each primer, 200µl of dNTP mixture, 1X PCR buffer, 1.0 mM Mg^{2+}, 1U of Taq polymerase enzyme and 50ng genomic DNA in a 25µl volume. The cycling program involved a genomic DNA denaturation step at 94˚C for 5min followed by 35 cycles of; denaturation at 94˚C for 30s, annealing at 62˚C for 30s, extension at 72˚C for 30s and final extension for 5min at 72˚C. A 10µl volume of the PCR products was digested using 3U of *Mbo*II restriction enzyme 16 hoursat 37˚C. The digestion products were electrophoresed on 2% agarose gel stained containing EtBr stain for 1 hour at 100V in TBE. The 213bp PCR product remains undigested in the presence of 170G allele while the 170A allele introduces a *Mbo*II restriction site yielding two fragments, 218bp and 75bp.

2.7 Statistical Analysis

Genotype and allele frequencies in the HIV exposed infected (EI), exposed uninfected (EU) and unexposed children cases and controls were calculated using Stata 11.2 (StataCorp LP, Texas USA) and/or SHEsis online version (Shi & He. 2005). Genotype frequency calculation involves counting the number of individuals carrying a genotype in a group and dividing by the total number of individuals in the group. To determine minor allele frequency the number of homozygotes for the allele is multiplied by two, added to the number of heterozygotes and divided by twice the number of individuals in the group. Fit to Hardy Weinberg Equilibrium (HWE) was tested using chi-squared (X^2)/Fisher′s exact with one degree of freedom. Chi-squared (X^2) and Fisher′s exact test where appropriate, were used to test homogeneity in alleles and genotypes among the participant groups. The same tests were used to determine whether genotypes influenced other categorical phenotypes such as neurocognitive status and haemoglobin. A $P<0.05$ was considered statistically significant. The correlation between CD4+ T-cell counts with genotypes and risk to neurocognitive impairment were analysed using Kruskall-Wallis and Wilcoxon ranking tests respectively. Differences in demographic and clinical characteristics between the infected and uninfected were calculated using the X^2 test for categorical data and analysis of variance (ANOVA) for continuous variables. Pairwise Linkage disequilibrium (LD) analyses between loci located on the same gene was carried out using SHEsis (Li et al. 2009). Lewontin's D′value and r^2 were used to quantify the level of LD ranging from zero for independence to 1 for complete co-inheritance. Haplotypes were inferred using the expected maximisation algorithm. The bulk of the statistical analyses were done using Stata 11.2 (StataCorp LP, Texas USA) except where SHEsis is indicated as the package of choice.

3 CHAPTER THREE: RESULTS

3.1 Demographic features

A total of 116 children met the inclusion criteria and were enrolled in the study. Of these, 34 (29%) were HIV-exposed and HIV-infected (exposed infected; EI) while 82 (71%) were HIV-negative controls comprising of 39 (34%) HIV-exposed (exposed but uninfected; EU) and 43 (37%) not exposed to HIV (unexposed and uninfected; UEUI). Of the 116 blood samples, 106 (91%) had DNA successfully extracted and formed part of the subsequent genotyping study (Figure 2.1). Table 3.1 is a summary of the demographic features and common morbidities recorded among the participants. There were no significant differences observed between the EU and UEUI groups (appendix D) so the two were combined to form the HIV-uninfected and compared to the HIV-infected (EI) group as shown in Table 3.1. HIV-uninfected children were significantly taller and weighed more than the HIV-infected. However, body mass index (BMI) did not differ significantly with HIV status despite the HIV-infected group having more underweight individuals. There were a higher proportion of children with low Hb in the HIV-infected group compared to the HIV-uninfected group (20% v 11%) although the difference was not significant. Occurrence of morbidities was generally higher among HIV-infected children compared to the HIV-uninfected with differences of 10% or more in all five clinical conditions.

Table 3.1: Demographic and clinical characteristics of study participants

Characteristics	HIV-infected (EI)	HIV-uninfected (EU+UEUI)	Odds Ratio	P-value
Mean age in years ±Stdev (range)	8.22 ±0.57 (7.25-9.08	8.43 ±0.55 (7.5-9.08)	N/A	0.28
Mean height in cm ±Stdev (range)	117.11 ±6.22 (109-131)	120.37 ±6.00 (108-134)	N/A	**0.03**
Mean weight in kg ±Stdev (range)	20.31 ±1.99 (17-24)	22.35 ±3.01 (15-31)	N/A	**0.005**
Mean head circum in cm ±Stdev (range)	51.14 ±1.57 (49-54)	51.37 ±1.2 (48-54)	N/A-	0.48
Body mass index	n=22	n=62		
Normal (5-85%)	18 (0.82)	55 (0.89)	1.00	
Underweight (<5%)	4 (0.18)	7 (0.11)	1.75 (0.33-7.80)	0.41
Sex	n=34	n=72		
Female	20 (0.59)	36 (0.50)	1.00	0.48
Male	15 (0.41)	36 (0.50)	0.75 (0.31-1.82)	
Haemoglobin	n=20	n=65		
Normal (11-16g/dL)	15 (0.75)	58 (0.89)	1.00	
Low (<11g/dL)	4 (0.20)	7 (0.11)	2.07 (0.39-9.36)	0.28
High (>16g/dL)	1 (0.05)	0 (0.00)		
Morbidities	n=24	n=72		
Diarrhoea				
No	16 (0.67)	59 (0.82)	1.00	
Yes	8 (0.33)	13 (0.18)	2.27 (0.68-7.13)	0.12
Ear Discharge				
No	20 (0.83)	67 (0.93)	1.00	
Yes	4 (0.17)	5 (0.07)	2.68 (0.47-13.6)	0.16
Vomiting				
No	18 (0.75)	61 (0.85)	1.00	
Yes	6 (0.25)	11 (0.15)	1.85 (0.48-6.37)	0.28
Cough				
No	12 (0.50)	44 (0.61)	1.00	
Yes	12 (0.50)	28 (0.39)	1.57 (0.55-4.41)	0.34
Fever				
No	10 (0.42)	43 (0.60)	1.00	
Yes	14 (0.58)	29 (0.40)	3.48 (1.26-9.75)	0.12

3.2 Genomic DNA quantification and integrity

After DNA was extracted from blood in preparation for genotyping as described in the method section 2.2, DNA yield was measured using spectrophotometry and 106 (91%) samples had detectable DNA with concentrations averaging 79.04 ng/µl (range; 3.32 to 290.68 ng/µl). Five percent of the samples had less than 20ng/µl DNA. DNA integrity check using gel electrophoresis showed that for all the 106 samples, intact DNA was successfully extracted. The genomic DNA was used to determine variation in selected HIV/AIDS restriction genes.

3.3 Genotyping for single nucleotide polymorphisms

As part of genotyping assay designing, primers were designed using bioinformatics tools as described in the methods section 2.3.1. Nine pairs of PCR primers were designed and successfully amplified DNA fragment flanking *CX3CR1* 839C>T, *RANTES* -403G>A, *APOBEC3G* -571G>C, *APOBEC3G* -90C>G, *APOBEC3G* 197193T>C (197T>C), *APOBEC3G* 199376C>G (199C>G), *MBL2* 170G>A SNPs and *MBL2* promoter and exon 1 regions. In addition, seven SNaPshot® primers were designed for genotyping of the following SNPs; *RANTES* -403G>A, *APOBEC3G* -571G>C, *APOBEC3G* 197T>C, *APOBEC3G* 199C>T, *MBL2* -595G>A, *MBL2* -221C>G and *MBL2* +4C>T. Despite careful designing of primers, PCR reactions had to be optimised for successful amplification of DNA fragments. Of the 14 PCRs that were successfully optimised with annealing temperatures ranging from 55˚C to 63˚C, additional Mg^{2+} was adjusted only for two assays, *APOBEC3G* 557A>G and *SDF1* 801G>A which required 0.5mM and 1.0mM, respectively. Five percent dimethyl sulfoxide (DMSO) was also added to the two reactions, *APOBEC3G* 557A>G and *SDF1* 801G>A to enhance amplification. Post PCR genotyping techniques used included restriction fragment length polymorphism (RFLP), SNaPshot® and Sanger chain termination sequencing.

Figures 3.1, 3.2, 3.3 and 3.4 show examples of PCR, RFLP, SNaPshot® and Sanger chain termination sequencing genotyping techniques, respectively.

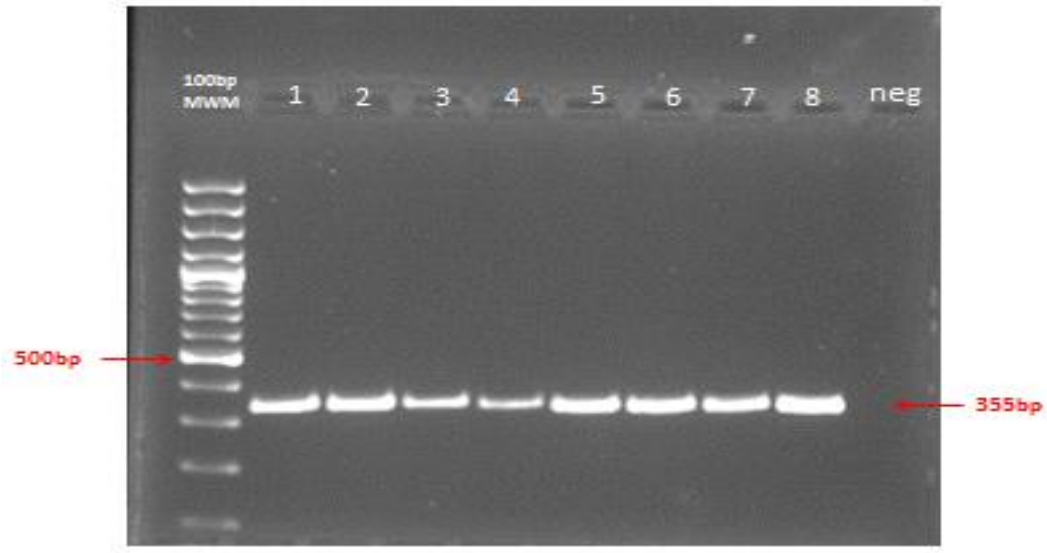


Figure 3.1: A 2% agarose gel showing a representative PCR. The picture shows a 355bp fragment of the *CX3CR1* gene flanking the 839C>T (rs3732378) SNP amplified using PCR. The products were electrophoresed at on 2% agarose gel stained with EtBr at100V for 1hour in TBE buffer.

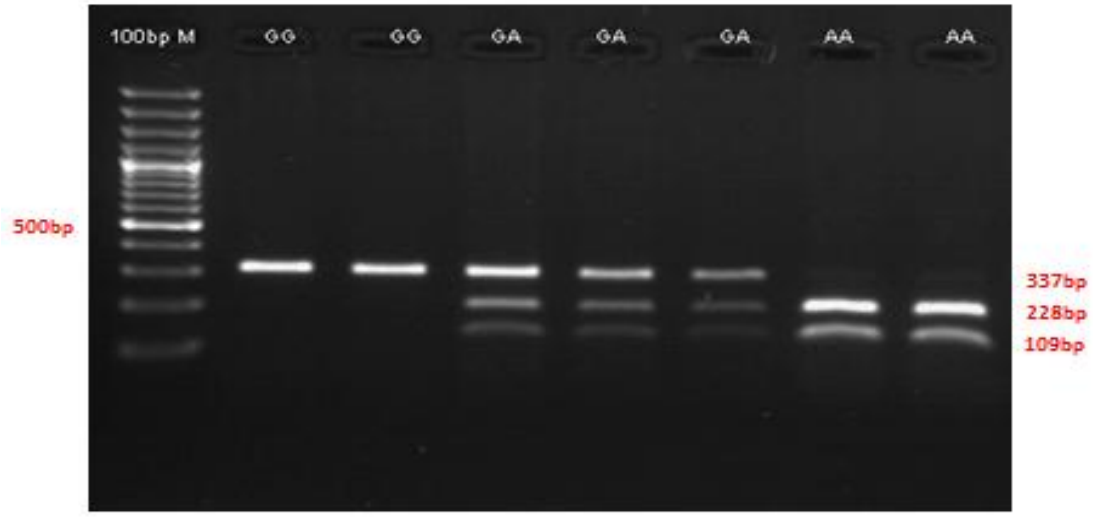


Figure 3.2: 2% agarose gel electrophoresis showing a representative RFLP. The *CCR2* 190G>A (rs17141036) SNP was genotyped using *BseG*I enzyme restriction. The 190A allele is digested into two fragments whilst the 190G remains undigested (Method described by Martinson et al., 2000).

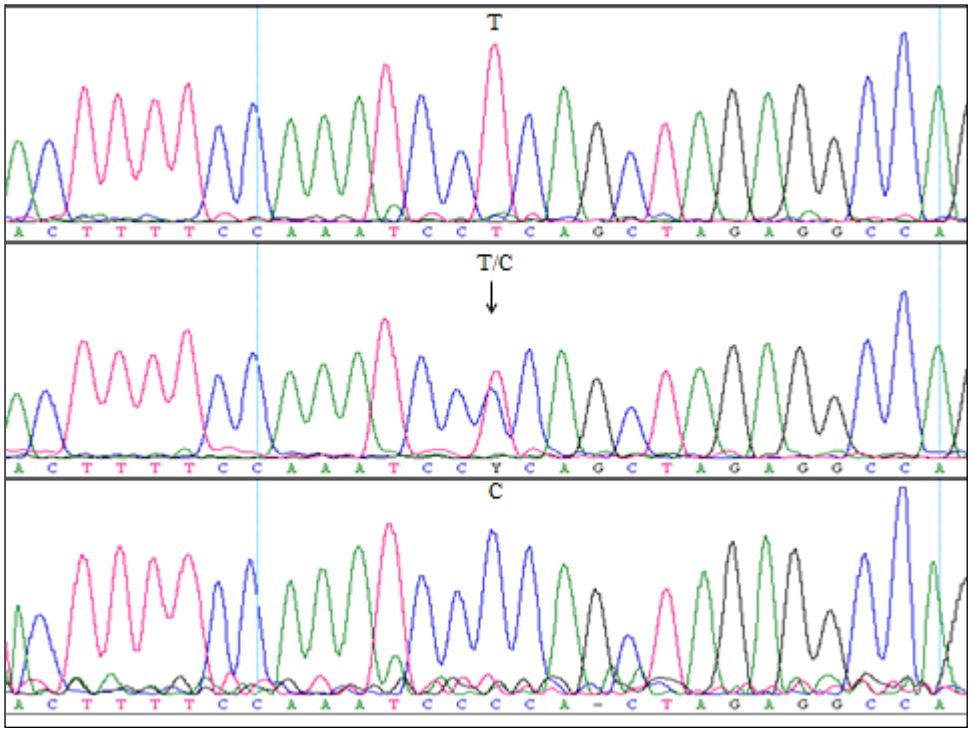


Figure 3.3: Electrophoregram illustrating genotyping by DNA sequencing. Each peak represents a nucleotide in the DNA sequence and the black text indicates the allelic variants at a polymorphic site.

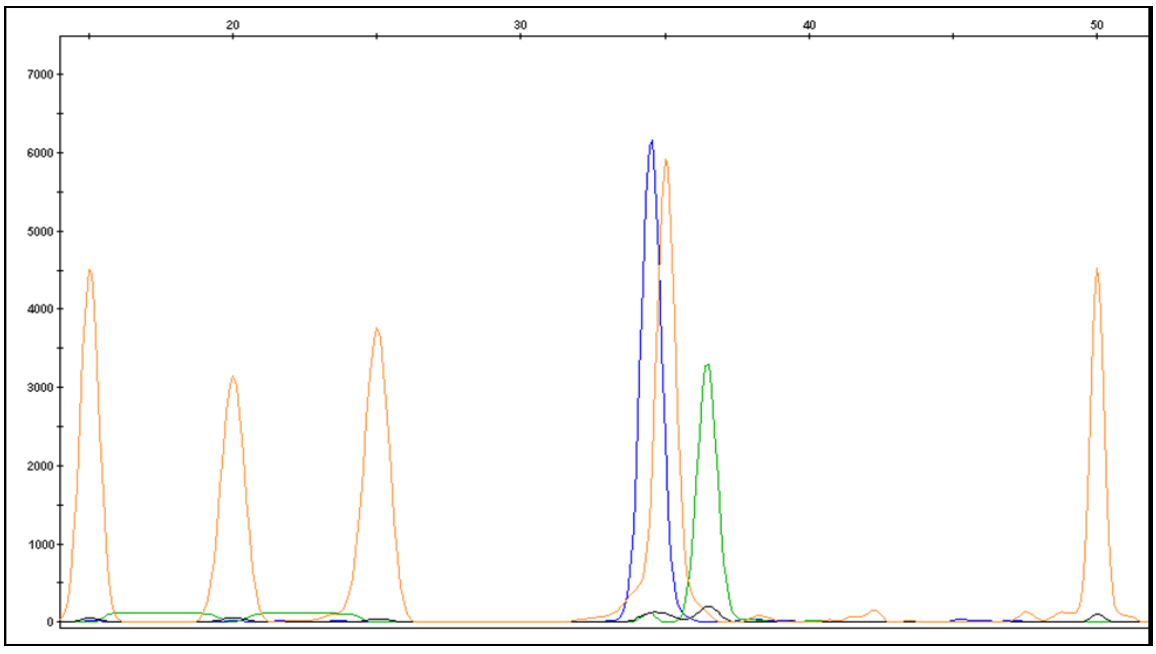


Figure 3.4: Electrophoregram illustrating genotyping by SNaPshot®. The blue and green peaks are representative of G (dR110) and A (dR6G) alleles respectively whilst the orange peaks are size standards.

3.4 Baseline allele frequencies and their comparison to other populations

After genotyping for variants in six HIV/AIDS restriction genes *CCR2, CX3CR1,SDF1 (CXCL12), RANTES (CCL5), APOBEC3G (APOBEC3G)* and *MBL2*, baseline frequencies in the Zimbabweans were calculated using the combined HIV-uninfected group (EU+UEUI). We report frequencies for the following minor alleles in the Zimbabwean population; *CCR2* 190A (16%), *CX3CR1* 745A (9%), *CX3CR1* 839T (1%), *SDF1* 801A (2%), *RANTES* In1.1C (20%), *RANTES* -403A (44%), *APOBEC3G* 557G (40%), *APOBEC3G* -90C (32%), *APOBEC3G* -571C (12%), *APOBEC3G* 197C (38%), *APOBEC3G* 199G (6%), *MBL2* -595A (0.02), *MBL2* 170A (23%), *MBL2* -221C (12%), and *MBL2* +4C (46%) (Table 3.2). To our knowledge all or most of these variants are being reported for the first time in the population. All SNPs conformed to the Hardy-Weinberg Equilibrium in the HIV uninfected group used in calculation of baseline allele frequencies (EUI + UEUI).

The frequencies of the alleles in our study cohort were compared to six other populations published on HapMap and NCBI dbSNP databases. The populations include the Yoruba of Ibadan Nigeria (YRI), African ancestry in Southwest USA (ASW), Luhya in Webuye, Kenya (LWK), Maasai in Kinyawa, Kenya (MKK), Han Chinese in Beijing (HCB) and the Utah residents with Northern and Western European ancestry (CEU). Table 3.2 shows the minor allele frequencies in the Zimbabwean population and how they compare to the other populations. The prevalence of minor alleles in our study population closely resembled that of the YRI, ASW and LWK populations because they are all of African origin. However, there were exceptions for example the *MBL2* -221G allele frequency had borderline significance for difference between the YRI (19%) and this Zimbabwean group (11%) (P=0.05). Another *MBL2* variant, 170A was significantly lower in the LWK (15%) compared to Zimbabweans (23%) (P=0.05).

Table 3.2: Comparison of allele frequencies between the HIV-uninfected (EUI+UEUI) group and other populations published on HapMap and NCBI dbSNP databases

Minor Allele	Current study	YRI	LWK	MKK	ASW	HCB	CEU
***CCR2* 190A (rs17141036A)**	0.16	0.18	NA	NA	N/A	0.26	0.11
***CX3CR1* 745A (rs3732379A)**	0.09	0.12	0.09	0.11	0.11	0.00 **P=0.003**	0.261 **P=0.001**
***CX3CR1* 839T (rs3732378T)**	0.01	0.00	NA	0.03	0.03	0.00	0.15 **P=0.001**
***SDF1* 801A (rs1801157A)**	0.02	0.02	0.05	0.08 **P=0.01**	0.08	0.24 **P<0.001**	0.208 P=0.001
***RANTES* In1.1C (rs2280789C)**	0.204	0.17	0.23	0.14	0.22	0.34 **P=0.02**	0.10.2 **P=0.006**
***RANTES*-403A (rs2107538A)**	0.44	0.43	0.48	0.42	0.39	0.35	0.155 P=0.001
***A3G* 557G (rs8177832G)**	0.40	0.48	0.34	0.18 **P<0.001**	0.31	0.05 **P<0.001**	0.031 **P<0.001**
***A3G* -90C (rs5757463C)**	0.32	0.44	NA	NA	NA	N/A	0.32
***A3G* 197C (rs3736685C)**	0.42	0.48	0.34	0.18 **P<0.001**	0.31 **P=0.02**	0.05 **P<0.001**	0.031 **P<0.001**
***A3G* -571C (rs5757463C)**	0.12	0.04	NA	NA	NA	N/A	0.10
***A3G* 199G (rs2294367G)**	0.06	0.04	NA	NA	NA	0.66 **P<0.001**	0.523 **P<0.001**
***MBL2* 170A (rs1800451A)**	0.23	0.25	0.15 **P=0.05**	0.15 **P=0.043**	0.20	0.01 **P<0.001**	0.018 **P<0.001**
***MBL2* -221G (rs7096206G)**	0.11	0.19 **P=0.05**	0.16	0.14	0.20 **P=0.043**	0.15	0.22
***MBL2* +4C (rs7095891C)**	0.44	0.503	0.51	0.43 **P=0.01**	0.45	0.11 **P<0.001**	0.21 **P<0.001**

Key: *A3G- APOBEC3G*, Zim-Zimbabweans; YRI-Yoruba of Ibadan, Nigeria, LWK-Luhya of Webuye,Kenya, MKK-Masaai of Kinyawa, Kenya, ASW-African ancestry in Southwest, USA, HCB-Han Chinese in Beijing, CEU- Utah residents with Northern and Western European ancestry.
P-values are only shown where the difference from our study population was significant

In addition, two variants, *MBL2* -221G and APOBEC3G 197C showed statistically significant differences between the ASW (20% & 31%, respectively) and our study population (11% & 42%, respectively). We also observed that the MKK genetic profile seemed to drift much more from the Zimbabwean compared to the other populations of African origin (YRI, LWK and ASW). Five out of the ten (50%) allele frequencies that were compared between the MKK and our study population showed statistically significant differences. As expected, the Asian and Caucasian populations (HCB and CEU respectively) were largely different from the Zimbabwean population. Eight out of 13 (62%) and 10 out of 14 (71%) variants in HCB and CEU respectively were significantly different from our study population. With baseline allele frequencies now established, distribution of genotypes was compared between HIV EI and EU groups to determine possible association between genetic variation and risk of HIV infection.

3.5 Genetic polymorphism and their association with HIV status

3.5.1 Genetic polymorphism of chemokines receptor genes *CCR2* and *CX3CR1*

Three SNPs (*CCR2* 190G>A, rs17141036; *CX3CR1* 745G>A, rs3732379; *CX3CR1* 839T>C; rs3732378) on two genes encoding for chemokine receptors CCR2 and CX3CR1 were genotyped using PCR-RFLP. Their genotype frequencies in the HIV EI and EU groups and their measures of association with HIV infection are shown in Table 3.3. The heterozygous *CCR2* 190G/A genotype when compared to the 190G/G genotype was significantly associated with reduced risk of infection when compared between the HIV EI (15%) and EU (39%) with an odds ratio of 0.27 (95% CI: 0.07-0.97; p=0.02). However, the 190A/A genotype appeared more in the HIV exposed infected (EI) children (11.8%) compared to the HIV exposed uninfected (EU) (3%) but this difference was not statistically significant.

There were no significant differences observed between the EI and EU groups with a dominant model for the 190A allele. Genotyping of the two non-synonymous SNPs *CX3CR1* 745G>A and *CX3CR1* 839C>T *(*T280M) showed the homozygous 745A/A genotype occurred more frequently in the EI (10%) compared to the EU (3%) but the difference was not statistically significant. Due to the low frequency of the 839T, the 839T/T genotype was not observed in the Zimbabwean population.

3.5.2 Genetic polymorphism of chemokine genes *RANTES* and *SDF 1*

Frequencies of *SDF1* 801G>A, *RANTES* -403G>A and *RANTES* In1.1T>C genotypes were compared between HIV EI and EU children to determine the relationship between genetic variants and HIV infection (Table 3.3). A higher frequency of the homozygous -403A/A genotype was observed in the HIV exposed infected (17%) compared to the HIV exposed uninfected (12%) but the difference was not significant. No association was observed between HIV infection and variants of the SNPs, *SDF1* 801G>A, *RANTES* -403G>A and *RANTES* In1.1T>C. Since the genotypes resulting from the In1.1T>C and -403G>A SNPs did not show significant differences in distribution among the HIV groups individually their co-inheritance patterns were considered.

Table 3.3: Frequency and distribution of chemokine and chemokine receptor genotypes and their association with HIV status

GENOTYPES	EI	EU	UEUI	OR (95% CI)	EI v EU p-value
***CCR2* 190G>A (rs17141036)**	n=34	n=36	N=36		
190G/G	25 (0.74)	21 (0.58)	27 (0.75)	1.00	
190G/A	5 (0.15)	14 (0.39)	8 (0.22)	**0.27 (0.07-0.97)**	**0.02**
190A/A	4 (0.11)	1 (0.03)	1 (0.03)	4.67 (0.42-236)	0.14
***CX3CR1* 745G>A (rs3732379)**	n=31	n=36	n=36		
745G/G	24 (0.77)	29 (0.81)	31 (0.86)	1.00	
745G/A	4 (0.13)	6 (0.17)	5 (0.14)	0.74 (0.14-3.53)	0.67
745A/A	3 (0.10)	1 (0.03)	0 (0.00)	3.75 (0.28-202)	0.23
CX3CR1 839C>T (rs3732378)	n=31	n=36	n=36		
839C/C	31 (1.00)	36 (1.00)	35 (0.97)	N/A	1.00
839C/T	0	0	1 (0.03)	N/A	N/A
***SDF1* 801G>A (rs1801157)**	n=33	n=36	n=36		
801G/G	31 (0.94)	35 (0.97)	35 (0.97)	1.00	
801G/A	2 (0.06)	1 (0.03)	0 (0.00)	2.26 (0.11-136)	0.50
801A/A	0	0 (0.00)	1 (0.03)	N/A	N/A
***RANTES* In1.1T>C (rs2280789)**	n=34	n=36	n=35		
In1.1T/T	19 (0.56)	20 (0.56)	23 (0.66)	1.00	
In1.1T/C	14 (0.41)	16 (0.44)	11 (0.31)	0.88 (0.30-2.51)	0.78
In1.1C/C	1 (0.03)	0 (0.00)	1 (0.03)	N/A	0.30
RANTES -403G>A (rs2107538)	n=33	n=36	n=36		
-403G/G	7 (0.21)	8 (0.22)	11 (0.31)	1.00	
-403G/A	22 (0.67)	22 (0.61)	21 (0.58)	1.27 (0.43-3.84)	0.63
-403A/A	4 (0.12)	6 (0.17)	4 (0.11)	0.74 (0.14-3.5)	0.67

Key: EI-HIV exposed infected, EU-HIV exposed uninfected, UEUI- HIV unexposed, N/A- non-applicable, OR (95%CI) - odds ratio (95% confidence interval).

Haplotype formation and linkage disequilibrium (LD) between the *RANTES In1.1T>C* and *-403G>A* were calculated using SHEsis online package (Li et al. 2009). For correct approximation of LD between In1.1T>C and -403G>A variants in the population, the HIV-uninfected group comprising of HIV-exposed and unexposed children (EU+UEUI) was used to calculate the LD. The SHEsis program uses the expectation maximisation algorithm to infer haplotypes from provided frequencies. Haplotype frequencies were compared between the EI children and the combined EU+UEUI groups. Table 3.4 is a summary of the haplotypes constructed between *RANTES* In1.1T>C and -403G>A allelic variants, their distribution between the two groups and measures of association. All four expected haplotypes were observed in EI and EI+UEUI but the In1.1-403 (C-G) haplotype which is the combination of two RANTES down-regulating alleles made up only 2.4% and 1% of the EI and EU+UEUI, respectively. No statistically significant differences in haplotype frequencies between the HIV-infected and HIV-uninfected were observed. LD analysis (Figure 3.5) showed that the In1.1C allele is in weak disequilibrium with the -403A (r^2 =0.277, p<0.0001).

Table 3.4: Frequency and distribution of *RANTES* haplotypes between HIV infected and uninfected groups

Haplotype		**EI (n=66)**	**EU+UEUI (n=136)**	**OR (95% CI)**	**P-value**
In 1.1T>C	-403G>A				
T	G	34 (0.53)	78 (0.56)	0.90 (0.50-1.64)	0.74
T	A	17 (0.25)	34 (0.24)	1.07 (0.55-2.11)	0.84
C	A	13(0.20)	28 (0.19)	1.08 (0.52-2.24)	0.85
C	G	2 (0.02)	1 (0.01)	4.22 (0.21-250)	0.21

Key: EI- HIV exposed infected, EU- exposed uninfected, EU+UEUI-HIV negative

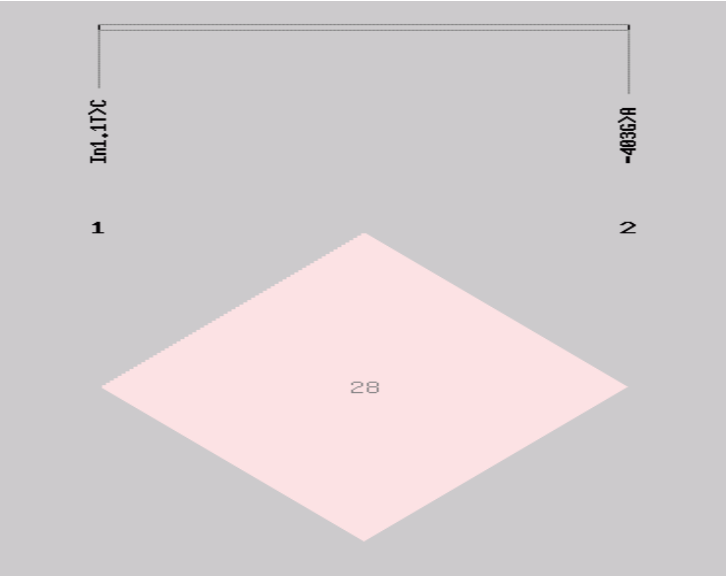


Figure 3.5: LD plot of *RANTES* SNPs, -40G>A and In1.1T>C. The number in the middle of the box shows the r^2 value expressed as a percentage.

3.5.3 Polymorphism in innate immune system genes

3.5.3.1 *APOBEC3G* polymorphism

We genotyped five *APOBEC3G* (*APOBEC3G*) SNPs including 557A>G (rs8177832), -571G>C (rs5757463), -90C>G (rs5757463), 197T>C (rs3736685) and 199C>G (rs2294367) and compared their genotype frequencies between HIV EI and HI EU children to determine possible association between genetic variants and HIV infection. Table 3.5 shows the distribution of *APOBEC3G* variants in EI and EU children and the measures of association with HIV status. Although all the SNPs in the *APOBEC3G* were polymorphic the -571C/C and 199G/G genotypes were not observed in the whole study group and among EI children, respectively. The *APOBEC3G* 557G/G (H186R) genotype occurred more frequently in EI children (59%) compared EU children (47%) but the difference was not significantly significant. None of the *APOBEC3G* variants investigated in this study was individually associated with risk of HIV infection. Therefore the possible effect of co-inheritance of *APOBEC3G* variants on risk of HIV infection was investigated through haplotype analysis.

Table 3.5: Frequency and distribution of *APOBEC3G* and *MBL2* genotypes and their association with HIV status

GENOTYPES	HIV EI	HIV EU	UEUI	OR (95% CI)	P-value EI v EU
***A3G* 557A>G (rs8177832)**	n=32	n=36	n=36		
557A/A	9 (0.28)	12 (0.33)	15 (0.42)	1.00	
557A/G	19 (0.59)	17 (0.47)	16 (0.44)	1.63 (0.56-4.7)	0.32
557G/G	4 (0.13)	7 (0.19)	5 (0.14)	0.59 (0.11-2.66)	0.44
***A3G* -90C>G (rs5750743)**	n=32	n=36	n=36		
-90G/G	16 (0.50)	17 (0.47)	16 (0.44)	1.00	
-90C/G	14 (0.42)	17 (0.47)	15 (0.42)	0.87 (0.30-2.52)	0.77
-90C/C	2 (0.06)	2 (0.06)	5 (0.14)	1.13 (0.08-16.50)	0.90
***A3G* 197T>C (rs3736685)**	n=33	n=36	n=36		
197T/T	9 (0.27)	13 (0.36)	16 (0.44)	1.00	
197T/C	20 (0.61)	16 (0.44)	15 (0.42)	1.92 (0.67-5.61)	0.18
197C/C	4 (0.12)	7 (0.19)	5 (0.14)	0.57 (0.11-2.56)	0.41
***A3G* 199C>G (rs2294367)**	n=31	n=32	n=34		
199C/C	29 (0.94)	30 (0.94)	29 (0.85)	1.00	
199C/G	2 (0.06)	1 (0.03)	5 (0.15)	2.07 (0.10-126)	0.55
199G/G	0 (0.00)	1 (0.03)	0 (0.00)	N/A	0.32
***A3G* -571G>C (rs5757463)**	n=30	n=31	n=36		
-571G/G	23 (0.77)	28 (0.90)	23 (0.64)	1.00	
-571G/C	7 (0.23)	3 (0.10)	13 (0.36)	2.84 (0.56-18.63)	0.15
***MBL2* -595G>A (novel)**	n=32	n=36	n=36		
-595G/G	30 (0.94)	34 (0.94)	35 (0.97)	1.00	
-595G/A	2 (0.07)	2 (0.06)	1 (0.03)	1.13 (0.08-17.32)	0.90
***MBL2* -221C>G (rs7096206)**	n=32	n=36	n=36		
-221G/G	23 (0.73)	27 (0.75)	28 (0.78)	1.00	
-221C/G	9 (0.27)	9 (0.25)	8 (0.22)	1.17 (0.34-3.97)	0.77
***MBL2* 170G>A (rs1800451G)**	n=32	n=36	n=36		
170G/G	19 (0.59)	21 (0.58)	20 (0.56)	1.00	
170G/A	9 (0.28)	13 (0.36)	16 (0.44)	0.69(0.22-2.17)	0.48
170A/A	4 (0.13)	2 (0.06)	0	2.43(0.32-28.34)	0.31
***MBL2* +4C>T (rs7095891)**	n=32	n=36	n=36		
+4T/T	11 (0.34)	11 (0.31)	13 (0.36)	1.00	
+4C/T	12 (0.38)	21 (0.58)	12 (0.33)	0.42 (0.14-1.26)	0.09
+4C/C	9 (0.28)	4 (0.11)	11 (0.36)	3.0 (0.72-14.59)	0.08

Key: *A3G- APOBEC3G*, EI- exposed infected, EU- exposed uninfected, UEUI- unexposed uninfected.

Haplotypes in the *APOBEC3G* gene with respect to -571G>C, -90C>G, 557A>G, 197T>C and 199C>G SNPs were inferred using SHEsis online software (Li et al. 2009). Haplotype frequencies were compared between the HIV-infected children and the HIV-uninfected. Table 3.6 shows the haplotypes observed in the HIV-infected and uninfected groups and their frequencies. Haplotype C-G-G-C-C (number 7 in Table 3.6) was present in 6% of the HIV-infected children but absent among the HIV-uninfected (P=0.007). In contrast, haplotypes C-C-A-T-C (number 5) and G-G-A-T-G (number 6) were observed in 4% of the HIV-uninfected group each but none in the HIV-infected but the differences were not statistically significant. To determine the inheritance patterns of *APOBEC3G* SNPs in the population, pair-wise linkage disequilibrium was calculated in the HIV uninfected group using the correlation coefficient, r^2 using SHEsis online package. Figure 3.6 is an LD plot showing the results of the pair-wise LD analysis. The strongest LD was observed between 557G and 197C (r^2=0.83) while the rest were in weak LD with r^2 <0.2.

Table 3.6: Haplotype formation with respect to SNPs on the *APOBEC3G* gene

	-571G>C	-90C>G	557A>G	197T>C	199C>G	HIV-pos	HIV-neg	P-value
Hap 1	G	G	G	C	C	18 (0.31)	44 (0.34)	0.68
Hap 2	G	G	A	T	C	18 (0.30)	26 (0.20)	0.13
Hap 3	G	C	A	T	C	10 (0.17)	28 (0.22)	0.40
Hap 4	C	G	A	T	C	3 (0.05)	7 (0.06)	0.72
Hap 5	C	C	A	T	C	0	5 (0.04)	0.11
Hap 6	G	G	A	T	G	0	5 (0.04)	0.13
Hap 7	C	G	G	C	C	4 (0.06)	0	**0.007**
Hap 8	G	C	G	C	C	2 (0.04)	3 (0.02)	0.56

Key: Hap- haplotype

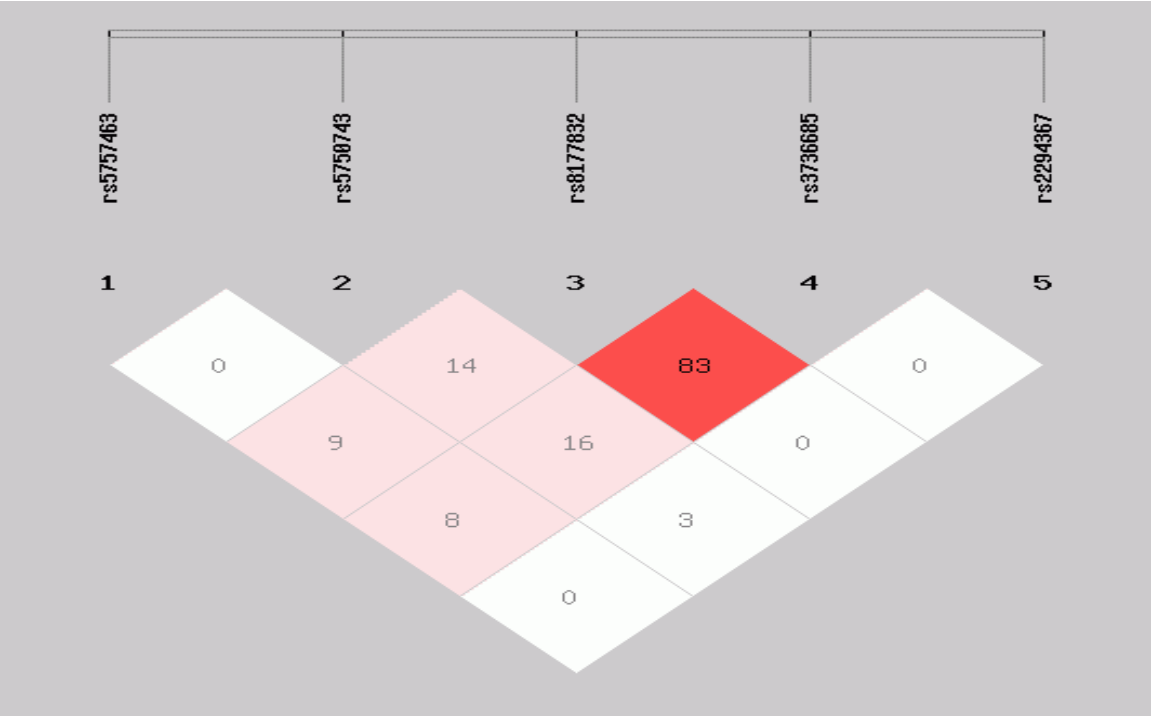


Figure 3.6: Linkage disequilibrium plot of *APOBEC3G* variants. The boxes show the r^2 for LD between two SNPs expressed as a percentage. The pink and red boxes indicate where minor alleles are in LD and the darker the box, the stronger the LD.

3.5.3.2 Polymorphism in the *MBL2* gene

After sequencing the promoter and exon 1 of *MBL2* gene using Sanger chain termination sequencing as described in the methods section, 12 polymorphic sites were detected comprising of one novel SNP and 11 previously reported SNPs which included a multiple base deletion. Table 3.7 shows the polymorphisms detected in the sequenced region and their genotype and minor allele frequencies. We report a novel *G>A* variation at position -595 upstream of the exon 1 start site (*MBL2* -595G>A) that has not been previously reported (Figure 3.7). Two out of the 26 samples sequenced, two carried the heterozygous -595G/A genotype but no homozygous -595A/A genotype was observed.

To confirm the frequency of the novel SNP, a SNaPshot® assay was designed to genotype all samples in our study cohort (n=106) and it was observed that the -595A variant occurred with a frequency of 3% and still no -595A/A homozygotes were observed. The -595G>A genotype frequencies were not significantly different between HIV-exposed infected (EI) and exposed uninfected (EU) (Table 3.5). Functional analysis of -595G>A variants done using a predictive bioinformatics software called TFSearch (Heinemeyer et al. 1998) showed that the *MBL2* -595G>A position does not carry any transcription factor binding site regardless of the presence of G or A allele.

Table 3.7: SNPs detected by DNA sequencing of the *MBL2* gene promoter and exon 1 regions.

SNP	N.t change	n	F	wt/wt*	wt/mt*	mt/mt*	MA (Freq)
Novel	-595G>A	104[a]		99 (0.93)	5 (0.07)	0	A (0.02)
rs11003125	-550C>G	28		27 (0.96)	1 (0.04)	0	G (0.02)
rs7100749	-435G>A	26	2	22 (0.85)	4 (0.15)	0	A (0.08)
rs11003124	-428A>C	28		6 (0.21)	10 (0.36)	12 (0.43)	C (0.39)
rs7084554	-349A>G	23	5	6 (0.26)	6 (0.26)	11 (0.48)	A (0.39)
rs45560739	-328AGAGAA/del	23	5	8 (0.33)	5 (0.21)	11 (0.46)	AGAGAA (0.44)
rs35236971	-245G>A	21	7	19 (0.90)	2 (0.10)	0	A (0.05)
rs7096206	-221C>G	104[a]		0	26 (0.25)	78 (0.75)	C (0.12)
rs67990116	-111A>T	21	7	19 (0.90)	2 (0.10)	0	T (0.10)
rs11003123	-70C>T	23	5	7 (0.31)	7 (0.30)	9 (0.39)	C (0.46)
rs7095891	+4C>T	104[a]		24 (0.23)	45 (0.43)	35 (0.34)	C (0.45)
rs5030737	154C>T	26	2	26 (1.00)	0	0	T (0)
rs1800450	161G>A	26	2	26 (1.00)	0	0	A (0)
rs1800451	170G>A	104[a]		60 (0.58)	38 (0.36)	6 (0.06)	A (0.24)

Key: N- total number of samples genotyped, F- failed samples, N.t- nucleotide, wt-wild type allele, mt- mutant allele, MA (Freq) - minor allele frequency.

[a] the SNPs were further genotyped in the samples remaining after sequencing.

*wt= refers to the starting allele as indicated in the nucleotide base substitution column and mt= to the second allele. This designation has nothing to do with functional studies.

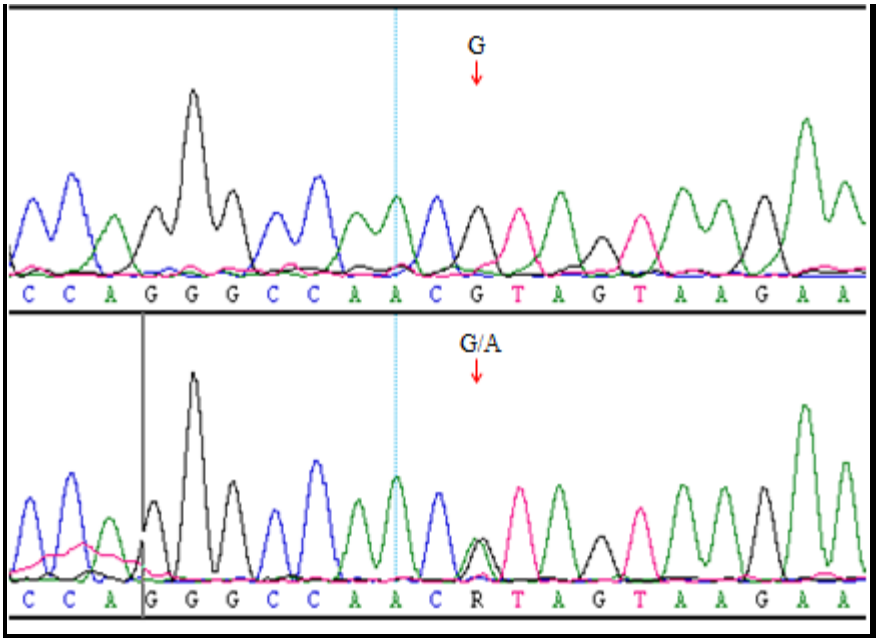


Figure 3.7: Electrophoregram showing the novel *MBL2* -595G>A SNP. The top panel shows the G allele whilst the bottom panel shows the heterozygous A/G genotype at the position highlighted

Of the eleven previously reported SNPs detected from the sequencing, three are part of a battery of *MBL2* polymorphisms that have been linked with variation in MBL protein levels and various diseases including HIV. These SNPs include -221C>G (-221*X/Y*) in the promoter region, +4C>T (+4*P/Q*) in the non-coding 5′-UTR and 170G>A *(*G57E*)* in the coding region of exon 1. Two other exon 1 SNPs that have been linked with HIV restriction, 154C>T (R52C) and 161G>A (G54E) were monomorphic during sequencing and therefore not pursued any further. The SNPs -221C>G, +4C>T and 170G>A were genotyped in the samples that had not been sequenced using other genotyping techniques as described in the methods chapter. The genotype and allele frequencies distribution for the SNPs are shown in Tables 3.2 and 3.5 respectively.

In the *MBL2* 5-UTR, the homozygous +4C/C genotype occurred more in HIV EI children (28%) compared to the EU (11%) with near borderline significance under a recessive model (OR=2.9 (95% CI, 0.70-14.59), P=0.09). In the promoter region, SNP -221C>G showed a higher frequency of heterozygous -221C/G genotype in HIV EI children (27%) compared to the EU (25%) although the differences were not statistically significant. Despite the presence of up to 20% of heterozygous -221C/G genotypes, the homozygous -221G/G genotype was not observed in the whole study population. In exon 1, the distribution of 170G>A genotypes between EI and EU groups was not significantly different.

MBL2 genotypes were combined and their frequencies compared between HIV-exposed infected (EI) and HIV-exposed uninfected (EU) children to determine if co-occurrence of variants had a possible effect on HIV infection. The *MBL2* -221G/G-+4C/T (*Y/P-Y/Q*) genotype combination was associated with reduced risk of HIV infection compared to other -221C>G-+4C>T genotypes combinations because its frequency was significantly lower in the EI group (22%) compared to the EU (44%) group (OR=0.32, 95%C.I; 0.09-1.04, P=0.03). In addition, the +4C/C-170G/A (*P/A-P/O*) genotype combination was observed in 9% of EI children but not the HIV exposed uninfected (P=0.06)

Pair-wise LD was calculated for all SNPs detected in the *MBL2* promoter and exon 1 regions. Figures 3.8 is an LD plot showing the level of LD between each pair of SNPs. LD is strongest between the *MBL2* -328del and -349G alleles (r^2 =1) which are located 21 bases apart in the genome. Inspection of the sequences showed that the homozygous -328del/del almost always occurs with the homozygous *MBL2*-349G/G whilst the heterozygotes are also inherited together in the Zimbabwean population.

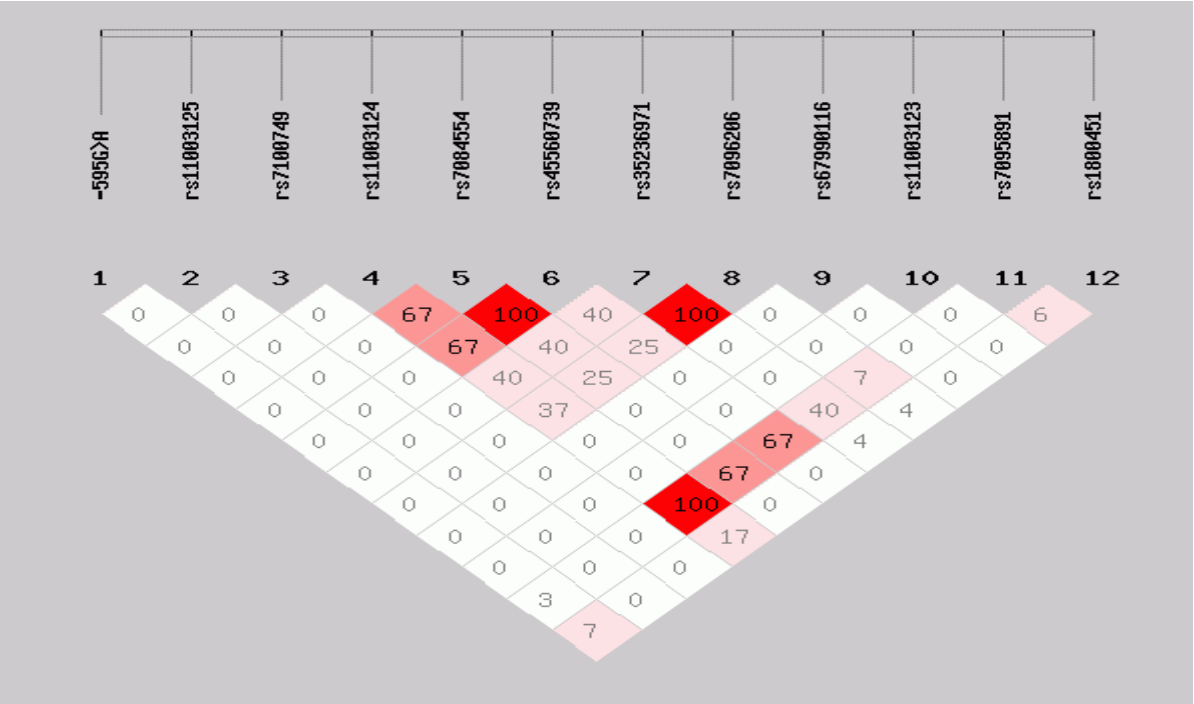


Figure 3.8: Linkage disequilibrium plot of *MBL2* SNPs. The plot shows pair-wise LD of *MBL2* variants in the promoter and exon1. The numbers in the boxes are percentage r^2 values indicating strength of LD.

Haplotype formation with respect to the -221C>G, +4C>T and 170G>A SNPs was also investigated. Seven haplotypes were observed, five in both HIV-infected and HIV-uninfected groups and the other two appearing in one of the HIV groups but not in the other. Table 3.8 shows the frequencies and distribution of the haplotypes between the HIV-infected and HIV-uninfected children. We observed a haplotype that has not been previously reported, C-C- A (*X-P-O* or haplotype 6 in Table 3.8) in the HIV-infected group. This however, was found in two chromatids only (minimum of one person and maximum of two possible).

Table 3.8: Haplotype formation with respect to SNPs in the *MBL2* gene and their comparison between HIV infected and uninfected children

***MBL2* haplotypes**					**HIV pos**	**HIV neg**	**OR (95% CI)**	**P-value**
	-221C>G	**4C>T**	**170G>A**	**AN**				
Hap 1	G	T	G	YQA	21 (0.33)	52 (0.36)	0.85 (0.46-1.5)	0.62
Hap 2	G	C	G	YPA	19 (0.30)	42 (0.30)	1.04 (0.55-1.9)	0.90
Hap 3	G	T	A	YQA	13 (0.20)	26 (0.18)	1.17 (0.56-2.4)	0.68
Hap 4	C	C	G	YQC	7 (0.10)	13 (0.09)	1.12 (0.42-2.9)	0.81
Hap 5	G	C	A	XPA	2 (0.03)	6 (0.06)	0.56 (0.10-2.6)	0.51
Hap 6	C	C	A	XPO	2 (0.04)	0	8.06 (0.91-7.0)	0.08
Hap 7	C	T	G	XQA	0	3 (0.02)	N/A	0.46

Key: Hap- haplotype, AN- alternative haplotype nomenclature

3.6 Genotype combinations and HIV status

The effect of single SNPs on HIV transmission or disease progression might be difficult to ascertain given the wide range of host proteins involved in the host parasite interaction pathways. However, when several pathways are disrupted due to multiple polymorphisms in different genes, a notable effect may be exerted. Given this, genotypes of selected SNPs were compared between HIV exposed infected (EI) and exposed uninfected (EU) children to determine their association with risk of HIV infection (Table 3.9).

The *CCR2* 190G/A-*CX3CR1* 745G/G combination was significantly associated with reduced risk of HIV infection compared to other *CCR2* 190G>A-*CX3CR*1 745G>A combinations as it occurred in 33% of HIV-exposed uninfected children but was not observed in than the EI group (P<0.001). The combination of two homozygous genotypes *CCR2* 190A/A-*CX3CR1* 745A/A was not observed in our study population. Combination between *CCR2* 190G>A and *APOBEC3G* 557A>G genotypes were also compared among the HIV groups.

Table 3.9: Genotype combinations and their association with HIV status

Genotype combination	HIV EI	HIV EU	OR (95% CI)	P-value
***CCR2* 190G>A v *CX3CR1* 745G>A**	N=31	N=36		
190G/G + 745G/G	22 (0.71)	17 (0.47)	1.00	**0.037**
190G/G + 745G/A	1 (0.03)	3 (0.08)	0.38 (0.0-5.06)	0.394
190G/G + 745A/A	1 (0.03)	1 (0.03)	1.2 (0.01-96.77	0.90
190G/A + 745G/G	0	13 (0.33)	N/A	**0.0002**
190G/A + 745G/A	3 (0.10)	2 (0.0.6)	1.86 (0.20-23.68)	0.501
190G/A + 745A/A	2 (0.06)	0	N/A	0.17
190A/A + 745G/G	2 (0.06)	0	N/A	0.17
***CCR2* 190G>A v *A3G* 557A>G**	N=32	N=36		
190G/G + 557A/A	7 (0.22)	5 (0.14)	1.00	0.39
190G/G + 557A/G	15 (0.47)	11 (0.31)	2.00 (0.67-6.10)	0.1
190G/G + 557G/G	2 (0.06)	5 (0.14)	0.41 (0.04-2.8)	0.30
190G/A + 557A/A	0	6 (0.17)	N/A	**0.015**
190G/A + 557A/G	4 (0.13)	6 (0.17)	0.71 (0.13-3.40)	0.63
190G/A + 557G/G	1 (0.03)	2 (0.06)	0.55 (0.09-11.11)	0.63
190A/A + 557A/A	2 (0.06)	1 (0.03)	2.33 (0.11-141)	0.49
190A/A + 557G/G	1 (0.03)	0	N/A	0.29
***CX3CR1* 745G>A v *RANTES* In1.1T>C**	N=31	N=36		
745G/G + In1.1T/T	13 (0.42)	17 (0.47)	1.00	0.66
745G/G + In1.1T/C	10 (0.32)	12 (0.33)	0.95 (0.30-2.98)	0.93
745GG + In1.1C/C	1 (0.03)	0	N/A	0.28
745G/A + In1.1T/T	3 (0.10)	3 (0.08)	1.18 (0.15-9.49)	0.85
745G/A + In1.1T/C	0	3 (0.08)	N/A	0.10
745A/A + In1.1T/T	3 (0.10)	0	N/A	**0.056**
745A/A + In1.1T/C	1 (0.03)	1 (0.03)	1.1 (0.01-94.16)	0.91

Key: EU- exposed uninfected, EI-exposed infected, N/A- not applicable, *A3G –APOBEC3G*

Analyses of *CCR2* 190G>A-*APOBEC3G* 557A>G genotype combinations revealed that the *CCR2* 190G/A-*APOBEC3G* 557A/A genotype combination was associated with protection against HIV infection as it was observed in 17% of the EU and but none in the EI group (P=0.015). The *CCR2* 190 A/A-*APOBEC3G* 557A/G combination was not found in the whole study population. An attempt to analyse genotype combinations amongst the three SNPs *CCR2 160G>A, APOBEC3G 557 A>G and CX3CR1 745G>A* was not possible because of small sample size. However, *CX3CR1* 745G>A-*APOBEC3G* 557A>G genotype combination analysis showed that the occurrence of *CX3CR1* 745G/G-*APOBEC3G* 557G/G was notably lower in EI children compared to EU (19.44% v 6.45 %) even though this difference was not statistically significant.

3.7 Genetic polymorphism and laboratory markers of disease progression

Possible association between genetic polymorphism and the following indicators of HIV disease progression; CD4+ T-cell counts, haemoglobin (Hb) and BMI was investigated. CD4+ T-cell counts in HIV-infected children were categorised into "low" for figures below 350cells/mm^3 and "normal" for counts above 350cells/mm^3. This grouping was done with the intention to determine association between genetic variation and CD4+ T-cell count status. However, this proved difficult as there were only three children with a CD4+ T-cell count below 350cells/mm^3. When the Kruskall-Wallis equality of proportion rank test was alternatively used to calculate differences in CD4+ T-cell counts among HIV-infected children with different genotypes, no significant differences in CD4+ T-cell count median values in children with different genotypes were observed for all SNPs investigated.

Among the indicators of disease progression investigated was haemoglobin (Hb). To determine if genetic polymorphism has a role to play in preservation of haemoglobin levels, we compared Hb level in individuals carrying different genotypes (Appendix C). Hb was grouped into low and normal using 11g/dl-16g/dl as the normal range in children. The *APOBEC3G* -571G/C genotype was associated with a low Hb (Hb<11g/dl) compared to the homozygous *APOBEC3G* -571G/G regardless of HIV status (P=0.01). The frequency of *RANTES* -403G/A genotype was higher in individuals with low Hb (80%) compared to those with normal Hb (57%). No association was observed between Hb levels and the rest of the SNPs investigated in this study. BMI also did not show any association with genetic variants investigated in this study.

3.8 Genetic polymorphism and neurocognitive function

Neurocognitive deficits are common among HIV-infected individuals with more aggravating effects on children compared to adults. In this study, possible association between genetic variation and neurocognitive function as determined by the McCarthy Scale of Children's Abilities (MSCA) among HIV-infected and HIV-uninfected children aged seven to nine years was investigated. The MSCA neurocognitive index scores were available for 21 HIV-exposed infected (EI) and 64 HIV-uninfected (EU+UEUI) children only at the time of writing as they were administered as part of a bigger study. Children were divided into unimpaired and impaired using cut-off values for each neurodevelopmental attribute as described in the methods. The results are shown in Table 3.10 below. The number of children falling below cut-off was generally higher in the HIV-infected compared to the HIV-uninfected for all four categories assessed using the MSCA which are; verbal, perceptive performance, quantitative and general index scores (Table 3.10).

The highest number of children falling below cut-off was in the quantitative index score category followed by perceptive and the verbal category had the least. CD4+ T-cell counts were compared between HIV-positive children falling below and above the cut-off values for all neurocognitive attributes but no significant differences were observed.

Table 3.10: Neurocognitive index scores in HIV infected and uninfected children

Attribute	HIV-infected	HIV-uninfected	OR (95% C.I)	P-value
General CIS				
>68	14 (0.67)	53 (0.83)	1.00	
<68	7 (0.33)	11 (0.17)	2.41 (0.66-8.30)	0.12
Verbal CIS				
>30	17 (0.81)	58 (0.91)	1.00	
<30	4 (0.20)	6 (0.09)	2.27 (0.42-10.82)	0.23
Perceptive CIS				
>30	14 (0.67)	54 (0.84)	1.00	
<30	7 (0.33)	10 (0.16)	2.7 (0.72-9.50)	0.08
Quantitative CIS				
>30	11 (0.52)	40 (0.63)	1.00	
<30	10 (0.48)	24 (0.37)	1.52 (0.49-4.59	0.4

Key: CIS- cognitive index score

To determine possible association between genetic variation and neurocognitive function, genotype frequencies were compared between cognitively-impaired and unimpaired groups of children for each of the four categories (Table not shown). Because HIV has been linked with neurocognitive deterioration, analysis was done on HIV-infected and HIV-uninfected groups separately. In the HIV-infected group, the *APOBEC3G* -90C/G genotype was associated with poor cognitive ability compared to the -90C/C and -90G/G genotypes as it occurred more in individuals who had a general cognitive index score (CIS) of 68 and below (71%) compared to those above 68 (25%) (P=0.048).

The *APOBEC3G* -90C/G was also observed more frequently in the verbally (73%) and perceptively impaired (71%) compared to the unimpaired (33% and 25%, respectively) but the differences were not significant. In addition, the *MBL2* 170G/A (*A/O*) genotype was significantly associated with high perceptive performance compared to the 170G/G and 170A/A as its occurrence was significantly higher in the unimpaired group (42%) compared to the impaired (0%) with regard to the perspective index score (P=0.046).

Other genetic variants showed weak association with neurocognitive impairment. For example, the *RANTES* In1.1T/C genotype was observed in 57% of the generally impaired HIV-positive children compared to 23% in the unimpaired but difference was not statistically significant. Another heterozygous *RANTES* variant, -403G/A was more frequent in generally impaired HIV positive children (86%) compared to the unimpaired HIV positive group (54%) but the difference was also not statistically significant. A fraction of HIV uninfected children develop neurocognitive deficits suggesting that etiological agents other than HIV are at play and genetic variation maybe one of the factors. Thus, genotype frequencies were compared between cognitively impaired and unimpaired children who are HIV-uninfected.

MBL2 polymorphisms appeared to have a significant contribution to neurocognitive function in children in the absence of HIV (Table 3.11). The frequency of the *MBL2* +4C/C (*P/P*) genotype was significantly higher in the verbally impaired HIV-uninfected group (50%) than the unimpaired (14%) (P=0.03). It was also observed that all verbally impaired children had at least one +4C allele. A dominant model of was therefore considered for the +4C allele (+4T/C + +4C/C vs +4T/T) and the frequency of +4C carrying genotypes remained significantly higher in the impaired group compared to the unimpaired (P=0.05).

Another *MBL2* promoter region variant -221C/G (*X/Y*) also occurred more frequently in the verbally impaired HIV-negative children (50%) compared to the unimpaired with borderline significance (17%) (P=0.057). The frequency of *MBL2* -221C/G genotype was also statistically higher in individuals with a general NCI score below cut-off (45%) compared to those above (15%) (P=0.02).

Table 3.11: *MBL2* genotypes and their association with verbal and general neurocognitive impairment in HIV-uninfected children

GENOTYPES	**Verbal cognition**		**P-value**	**General Cognition**		**P-value**
	Unimpaired	**Impaired**		**Unimpaired**	**Impaired**	
***MBL2* -595G>A (novel)**						
-595G/G	56 (0.97)	6 (1.0)	1.00	51 (0.96)	11 (1.00)	1.00
-595G/A	2 (0.03)	0 (0)		2 (0.04)	0	
***MBL2* -221C>G (rs7096206)**						
-221G/G	48 (0.83)	3 (0.50)	0.092	45 (0.85)	6 (0.54)	**0.023**
-221C/G	10 (0.17)	3 (0.50)		8 (0.15)	5 (0.45)	
***MBL2* 170G>A (rs1800451G)**						
170G/G	34 (0.59)	2 (0.33)	0.234	31 (0.58)	5 (0.45)	0.4277
170G/A	22 (0.38)	4 (0.67)	0.172	20 (0.38)	6 (0.55)	0.3016
170A/A	2 (0.03)	0	0.644	2 (0.04)	0	0.5127
***MBL2* +4C>T (rs7095891)**						
+4T/T	23 (0.40)	0	**0.05**	20 (0.38)	3 (0.27)	0.5104
+4C/T	27 (0.46)	3 (0.50)	0.872	25 (0.47)	5 (0.46)	0.9174
+4C/C	8 (0.14)	3 (0.50)	**0.0252**	8 (0.15)	3 (0.27)	0.3299

4 CHAPTER FOUR: DISCUSSION AND CONCLUSION

Since the discovery of HIV as cause of AIDS in the middle 1980s, it has been noted that the virus is heavily-dependent on the host machinery for the completion of its life cycle (Frankel et al. 1998). This led to investigations on what role the host might play in facilitating its own destruction or mounting a defence against HIV. The observation that some individuals, despite constant exposure to HIV either remained uninfected or if infected took much longer to develop HIV-associated symptoms even in the absence of antiretroviral therapy pointed to a possible role of host genetic variation in HIV infection and disease progression (Fowke et al. 1996, Kroner et al. 1994). In this study, we investigated the frequency and distribution of genetic variants in six genes; *CCR2, CX3CR1, SDF1, RANTES, APOBEC3G* and *MBL2* in HIV-exposed (infected and uninfected) children and HIV unexposed controls and their possible association with HIV infection and neurocognitive function.

The demographic features of the children enrolled showed that HIV-uninfected children were significantly taller and weighed more than the HIV-infected indicating possible slowed growth due to HIV-related challenges. HIV-infected children are constantly faced with a wide range of opportunistic infections and organ damage which may slow their growth compared to the uninfected (Owor et al. 2004, Bailey et al. 1999). Body mass index calculations also showed HIV-infected children to be underweight compared to the uninfected even though the difference was not statistically significant. Occurrence of morbidities was generally higher in HIV-infected children compared to the HIV-uninfected with differences of 10% or more in all five clinical conditions. This also may be a result of opportunistic infections and the immune-compromising effects of HIV.

However, the focus of the study was to describe human genetic variation and its correlation with HIV-infection and neurocognitive function. Baseline allele frequencies for the SNPs genotyped were established using the HIV-uninfected group (EU+UEUI). HIV-infected children were excluded from baseline frequency calculations because their genetic profile may not be representative of the population. For example, if host genetic variation has a role in HIV transmission then HIV-infected individuals are likely to carry more of the susceptibility alleles compared to the uninfected. The distribution of alleles in the HIV negative group of our study population conformed to the HWE for all SNPs studied indicating that the group can be correctly used to approximate allele frequencies for the population. Allele frequency data forms a baseline for further studies on the effect of genetic polymorphism on disease susceptibility, treatment response and disease progression in different populations. Distribution of alleles in populations may also be useful in estimating disease burden in the different populations.

After establishing baseline frequencies in the Zimbabwean children, the distribution of genotypes between HIV exposed infected (EI) and HIV exposed uninfected (EU) children was compared to determine possible association between host genetic variants and risk of HIV infection. Our study involved three groups of viral restriction genes which are chemokine receptors (*CCR2* and *CX3CR1*), chemokines (*RANTES* and *SDF1*) and innate immune factors (*APOBEC3G* and *MBL2*)

4.1 Chemokine receptor polymorphism and their association with HIV infection

CCR2 is a chemokine receptor and minor HIV co-receptor that is involved in leukocyte trafficking during inflammation and several other physiological processes (Palframan et al. 2001, Cardona et al. 2008). We genotyped the *CCR2* 190G>A (V64I) SNP and observed a frequency of 16% for the 190A allele in the study population which does not differ much from what has been observed in other African populations. Studies on Africa populations show frequencies ranging from about 13% among black South Africans (Williamson et al. 2000) to 23% among Kenyans (Anzala et al.1998). The occurrence of the *CCR2* 190A allele has been reported to vary considerably among ethnic groups in the same population as shown in Cameroonians where it ranged from 10 to 30% among seven ethnic groups (Ma et al. 2005). Frequencies of the *CCR2* 190A allele in Zimbabweans and among other African populations are slightly higher than what has been observed in Caucasians (7-10%) (Williamson et al. 2000, Struyf et al. 2000). Asian populations show a wide range of *CCR2* 190A allele frequencies (1-30%) (Qian et al. 2008, Su et al. 1999) which overlaps with both African and Caucasian frequencies. The distribution of *CCR2* 190A allele among the Caucasians, Africans and Asians does not seem to exhibit trends suggestive of the variant's role in HIV restriction at population level.

Frequencies of *CCR2* 190G>A genotypes were compared between HIV-exposed infected and HIV-exposed uninfected groups to determine association between variants and risk of HIV infection among individuals. We found the *CCR2* 190G/A genotype frequency significantly higher in HIV-exposed uninfected (EU) children compared to the HIV-exposed infected (EI) ($p=0.02$) suggesting that the genotype might be protective against HIV infection.

A possible mechanism of protection is that the 190A (64I) allele encodes a protein variant of the CCR2 receptor that is able to dimerise with HIV main co-receptor CXCR4 whereas the 190G (64V) has reduced affinity (Mellado et al. 1999). The dimerization reduces the amount of CXCR4 available for HIV binding therefore reducing the chances of HIV entering the cell. Surprisingly, neither the homozygous 190A/A genotype nor the 190A allele frequencies were significantly different between the HIV EI and EU. However, it has been suggested that the 190A allele is dominant over the 190G allele (Winkler et al. 1998), thus, the heterozygous 190G/A may be capable of exerting a full effect even with a single 190A allele. We however, did not find statistically significant differences between the EI and EU under the dominant model. These observations led to the conclusion that *CCR2* 190G>A SNPs might not have a significant individual role in risk of HIV infection.

Our observation support reports from two studies among Kenyans and French children where no significant differences were observed between HIV-infected and uninfected children infants born to HIV-infected mothers in the distribution of the 190A variant (Teglas et al. 1999, Brouwer et al. 2005). Contrary to our findings, a multicentre study involving Ugandans (n=358), Malawians (n=322) and South Africans (n=300) reported an increased risk of HIV infection in children homozygous for the 190A/A genotype when compared to 190G/G and 190G/A genotypes (Singh et al. 2008). The findings of Singh et al., (2008) however may have been distorted by the heterogeneous sample drawn from the three countries. In a separate study, our research group has demonstrated differences in distribution of genetic variants among African populations therefore pooling samples from different countries in gene association studies may introduce bias (Swart et al. 2012).

The chances of infection per child in the multicentre study could not have been the same given that Uganda has HIV subtypes A and D predominantly whilst South Africa and Malawi have subtype C (Buonaguro et al. 2007).

HIV subtype C has been shown to be more transmissible compared to both subtype A and D (Renjifo et al. 2004). In contrast to Singh et al. (2008), a study among Argentinians reported that the *CCR2* 190A/A genotype was significantly higher in HIV-exposed uninfected compared to exposed infected children again supporting a protective role against infection (Mangano et al. 2000). Given that CCR2 is a minor HIV co-receptor and its chemotactic roles are also executed by alternative pathways, *CCR2* 190G>A variation may be contributing little to the antiviral activity, thus, its role in HIV infection remains a subject of debate. Therefore in addition to *CCR2*, another chemokine receptor gene *CX3CR1* was also investigated.

CX3CR1 is a receptor for fractalkine which also acts as a minor HIV co-receptor (Imai et al. 1997, Combadiere et al. 1998). The CX3CR1-fractalkine pair is important for leukocyte trafficking and chemotaxis during inflammation (Nishimura et al. 2002). Two non-synonymous SNPs; *CX3CR1* 745G>A (V249I) and *CX3CR1* 839T>C (T280M) in the coding region of the gene were investigated in this study. We report frequencies of 9% and 1% for the 745A and 839T alleles, respectively in Zimbabweans. The occurrence of the 745A allele in our study population is lower than what has been reported among Caucasians (20.2%) (McDermott et al. 2000) but higher than that in Asian populations (2-5%) (Li et al. 2005); (Qian et al. 2008). The 839T allele was also less frequent in our study population compared to Caucasians (15%) but comparable to observations among Asians (2%) (Li et al. 2005). The high frequency of *CX3CR1* variants among Caucasians compared to that among Africans and Asians suggests its role in disease outcome but it is highly unlikely to be HIV.

This is because the resultant amino acid alterations (249I and 280M for 745A and 839T, respectively) are associated with less efficient binding of CX3CR1 receptors on peripheral blood mononuclear cells to ligand fractalkine. This reduces fractalkine-induced chemotaxis and therefore weakens the immune response mounted against HIV and other pathogens through the pathway (Faure et al. 2000) yet HIV prevalence is much lower in Caucasians compared to both Africans and Asians. We therefore speculate that *CX3CR1* variants may be risk factors for some inflammatory diseases that are more common among Caucasians as opposed to HIV/AIDS. For example, *CX3CR1* polymorphisms have been implicated in conditions like Crohn's disease, multiple sclerosis and bronchiolitis in Caucasian populations (Brand et al. 2006, Amanatidou et al. 2006, Stojković et al. 2012).

There were no significant differences in *CX3CR1* 745G>A genotype and allele frequencies between HIV EI and EU children. This implies that 745G>A variation alone may not have a significant role in HIV transmission. Our observations are supported by a study involving three sub-Saharan countries, South Africa, Uganda and Malawi where they also did not find any association between 745G>A variation and HIV infection in children born to ART naïve mothers (Singh et al. 2008). The same study however, observed that the *CX3CR1* 745A allele was associated with greater risk of HIV infection compared to the *CX3CR1* 745G in children born to infected mothers on ART prophylaxis (Singh et al. 2008). This observation suggests that the effect of *CX3CR1* 745G>A variation on HIV infection might require a minimum threshold of other protective mechanisms to be noticed. Not many studies have investigated the effect of 745G>A variation on HIV infection in children but a few have reported on adult population where no association has been observed (Faure et al. 2000, Suresh et al. 2006). Co-inheritance of alleles at the two loci may have more significant effect.

CX3CR1 745A and 839T have been shown to be in complete LD in Caucasian and Chinese populations (Qian et al. 2008, Faure et al. 2000) which is in contrast to our observation where 745A and 839T seem to be in complete equilibrium (r^2=0.001). Because of the complete LD in the populations, 745A and 839T alleles have been studied as haplotype *CX3CR1* 745-839 (A-T) and found to be associated with faster HIV disease progression in both adults and children (Faure et al. 2000, Brumme et al. 2003, Singh et al. 2005) as explained by the resulting reduced chemotactic activity of CX3CR1 described above. In addition to chemokine receptors, chemokines have also been implicated in HIV infection therefore two chemokines RANTES and SDF1 also formed part this study.

4.2 Chemokine receptor variants and their association with HIV infection

RANTES and SDF1 are natural ligands of the major HIV co-receptors CCR5 and CXCR4, respectively, thus, are able to inhibit HIV infection by competitively binding to these receptors (Bajetto et al. 2001, Choe et al. 1996). The frequency of *SDF1* 801G>A variants varies widely in different populations. We report a frequency of 2% for the *SDF1* 801A allele in the Zimbabwean population. This is comparable to what has been reported among black South Africans (1%) and the Yoruba of Ibadan Nigeria (2%) (Williamson et al. 2000); dbSNP). In contrast, the *SDF1* 801A allele occurs at frequencies of between 15% and 30% among Caucasians (Williamson et al. 2000, Apostolakis et al. 2005, Wasik et al. 2005). Asian populations exhibit a wider range of *SDF1* 801A allele frequencies, ranging between 10% and 40% (Amara et al. 2010, Qian et al. 2008, Khabour et al. 2012).

The effects of *SDF1* 801G>A genotypes on HIV infection in our study population could not be determined because of the low frequency of the 801A variant. The 801A allele makes the *SDF1* mRNA more stable post transcription, thus, up-regulating the synthesis of the SDF1 protein. This makes more SDF1 protein available for competitive binding against HIV (Winkler et al. 1998). It can therefore be speculated that the differential distribution of the *SDF1* 801A allele between Africans and Caucasians maybe contributing to the higher HIV prevalence among Africans compared to Caucasians. However, high frequency of *SDF1* 801A in Asians where HIV prevalence is also high casts doubt on its protective effect against HIV. HIV is a multifactorial condition hence there might be other underlying factors that work in concert with the 801A variant to give its differential role in different settings.

In contrast to *SDF1* 801G>A, the two *RANTES* SNPs (In1.1T>C and -403G>A) genotyped in this study were highly polymorphic in the population with minor allele frequencies of 20% (In1.1C) and 44% (-403A). The frequency of the *RANTES* In1.1C allele in our study population was lower than what has been observed among Asians (28%) (Qian et al. 2008) but higher than that among Caucasians (10%) (Laplana et al. 2012). The occurrence of the *RANTES* -403A allele in Zimbabweans (44%) was lower than that among Caucasians (16%) (Ahlenstiel et al. 2005) but comparable to that observed among Asians (40%) (Zhao et al. 2004). The In1.1T>C polymorphism is located in an intronic regulatory sequence and the In1.1C variant is associated with down-regulation of RANTES expression (An et al. 2002) resulting in less chemokine to inhibit HIV binding or internalise CCR5 receptors on T-cells, thus, favouring HIV infection. The high frequency of the In1.1C allele in Africans and Asians may argue for similar susceptibilities to HIV infection in the two populations when compared to Caucasians (An et al. 2002).

In contrast, *RANTES* -403A increases expression of *RANTES* (Liu et al. 1999, Nickel et al. 2000) thus its presence may reverse the down-regulating effect of *RANTES* In1.1C.

When *RANTES* In1.1T>C and -403G>A genotype frequencies were compared between HIV EI and EU groups, there were no significant differences observed. Our results support reports among Kenyans where they did not observe any significant differences between HIV exposed infected and uninfected in *RANTES* -403G>A genotype frequencies (Katz et al. 2010). Not much has been reported on the distribution of *RANTES* variants between EI and EU children thus our study attempts to understand the effect of *RANTES* variants in perinatal HIV infection of exposed children. However there are several reports on the association between *RANTES* polymorphism and risk of HIV infection in adults.

A study among Indians reported that the In1.1T allele was a risk factor for HIV infection but did not observe any association between -403G>A variants and HIV transmission (Rathore et al. 2008). In contrast to this and our findings, other studies have reported association between the *RANTES* -403A allele and increased risk of HIV infection under different settings (McDermott et al. 2000, Ahlenstiel et al. 2005). However, both studies used HIV negative controls that had no known history of HIV exposure. There is no general consensus on the effect of *RANTES* polymorphism on risk of HIV infection and our study adds to the growing body of conflicting findings.

The frequency of *RANTES* haplotypes with respect to In1.1T>C and -403G>A loci did not show any significant differences between the HIV EI and EU children. This observation suggests the effect of the co-inheritance of *RANTES* In1.1T>C and -403G>A alleles on the risk of HIV infection may be negligible in the population.

In contrast to our findings, An et al. (2002) reported an association between haplotypes containing In1.1C and -403A alleles and increased susceptibility to HIV infection among African American adults (An et al. 2002). Another study supported the observation that *RANTES* haplotypes may influence risk of HIV infection by reporting association between In1.1T-403A containing haplotypes and increased risk of HIV infection among Indian adults (Rathore et al. 2008).

The *RANTES* In1.1-403 (C-G) haplotype which carries two RANTES down regulating alleles was observed at very low frequency in the Zimbabwean population (1%) which is comparable to observations in Caucasians, Indians and African Americans where the C-G haplotype was not reported (An et al. 2002). The C-G haplotype is rare possibly because it carries both *RANTES* down-regulating variants which results in excessively low levels of RANTES protein that might not be able to sustain its multiple roles in the immunity and organ development, thus, may be aborted during development. This study also investigated genetic polymorphism in two innate antiviral factor genes *APOBEC3G* and *MBL2*.

4.3 Polymorphism in genes encoding innate immune system factors

4.3.1 *APOBEC3G* polymorphism

APOBEC3G is a cytidine deaminase that converts cytidine bases to uracil on the new HIV minus strand during reverse transcription (Bishop et al. 2004). This results in multiple G>A changes on the HIV DNA sense strand, a process called hypermutation. Hypermutation introduces stop codons on the HIV genome which disrupt production of new virions (Zhang et al. 2003). This study investigated genetic variation at five *APOBEC3G* loci and their association with HIV infection and neurocognitive function.

APOBEC3G 557A>G is the most studied of the *APOBEC3G* SNPs. We report a frequency of 40% for the 557G allele in the Zimbabwean population which is slightly higher to that observed among black South Africans (30%) and African Americans (37%) (An et al. 2004, Reddy et al. 2010) but much higher than among Caucasians (<5%) (An et al. 2004) and Asians (0%) (Rathore et al. 2008b). The *APOBEC3G* 557A>G polymorphism results in an H186R amino acid change on the APOBEC3G protein that alters its catalytic activity. The proposed mechanism is that the polymorphic site is involved in protein-protein interaction thus, the 557G (186R) variant leads to poor protein assembly which reduces APOBEC3G's catalytic functions (Jarmuz et al. 2002, MacGinnitie et al. 1995). In addition to the possible effect of the H186R amino acid alteration, APOBEC3G activity is determined by amount of the enzyme available in the circulation (Jin et al. 2007, Mariani et al. 2003). Inter-individual variation in APOBEC3G level in circulation may be a result of polymorphism in the gene's promoter region and other regulatory sequences such as those close to splice variants in introns. Thus two promoter SNPs *APOBEC3G* -90C>G and *APOBEC3G* -571G>C and two intron SNPs *APOBEC3G* 197193T>C (197T>C) and *APOBEC3G* 199376C>G (199) were investigated in this study.

We report frequencies of 32%, 12%, 42% and 6% for -90C, -571C, 197C and 199G alleles respectively in the Zimbabweans. There is limited information on the frequency and distribution of these alleles in different populations but this is how our observations compared to what has been published on dbSNP, NCBI. The occurrence of the -90C allele in our study is lower than that observed among the Yoruba of Nigeria (YRI) (32% v 44%) but higher than among Caucasians (4%). The frequency of the -571C was higher among the Zimbabweans in our study (12%) compared to the West Africans YRI (4%) and Caucasians (10%).

The distribution of *APOBEC3G* 197C allele frequency was comparable between Zimbabweans (42%) and YRI (48%) but higher than that observed among Caucasians (4%) and Asians (5%). In contrast to -90C, -571C and 197C alleles whose frequencies were generally higher in our study population and other Africans compared to Caucasians, the 199G allele was more frequent among Caucasians (58%) and Asians (66%) compared to what we observed (6%) and that among the YRI (4%). The distribution of *APOBEC3G* alleles, 557G, -90C, 197C and 199G in the world almost mirrors that of HIV cases where the highest prevalence is in Sub-Saharan Africa where 67% of the HIV cases live compared to Europe (7%) and East Asia (10%) (Figure 4.1). This suggests that *APOBEC3G* genetic profile may have contributed to the differences in susceptibility or resistance to HIV infection observed in different populations.

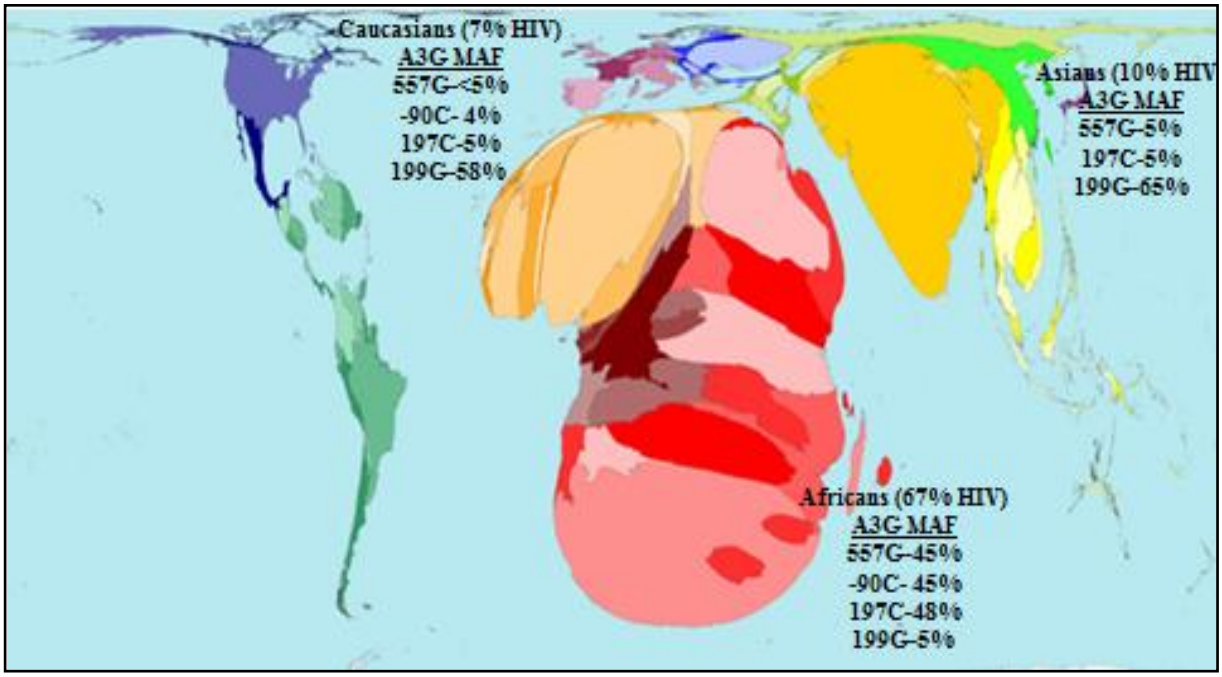


Figure 4.1: The co-distribution of *APOBEC3G* variants and HIV in different populations. This is a world map showing the how *APOBEC3G* variants seem to follow the trend in HIV distribution. The big body in the middle is the African continent whilst the green and blue strip to the left shows the Americas (adapted from www.worldmapper.org).

Despite the relationship between *APOBEC3G* variants and HIV prevalence in different populations, we did not observe any association between any of the *APOBEC3G* SNPs investigated in this study and risk of HIV infection. Our findings support observations by An et al. (2004), who in a multicentre study involving highly exposed seronegative and HIV seropositive individuals among African Americans and European Americans, also failed to observe any association between *APOBEC3G* variants and HIV status. Other studies on the 557G>A SNP have reported no association between its variants and risk of HIV infection (Reddy et al. 2010, Valcke et al. 2006, De Maio et al. 2011). However, the 40683T allele of a C40693T SNP (rs17496018) in exon 4 of the *APOBEC3G* gene which did not form part of this study was associated with increased risk of HIV transmission (Valcke et al. 2006). These observations on *APOBEC3G* gene variation suggest that individual allelic variants may fail to show any effect on their own but their collective contribution to gene expression or overall protein structure may be important.

In an effort to understand the combinational effect of *APOBEC3G* polymorphism on HIV transmission haplotype frequencies were compared between HIV infected and uninfected groups of children. LD analysis showed that only two variants, 557G and 197C which are positioned 177bp apart were in strong LD (r^2=0.83) in the population. This is in contrast to studies in Argentinians, Caucasians and African Americans where all five SNPs investigated as part of this study have been reported to be in strong LD (An et al. 2004, Bizinoto et al. 2011). This is not surprising because African populations have been shown to exhibit very low LD blocks compared to other world populations because they are more genetically diverse (Tishkoff et al. 2009).

We compared haplotype frequencies and observed that the haplotype for C-G-G-C-C with regard to -571G>C, -90C>G, 557A>G, 197T>C and 199C>G SNPs was associated with increased risk of HIV infection as it occurred in 6% of the HIV-infected children but was not seen in the HIV-uninfected (P=0.007). A possible explanation for this observation is that the haplotype carries two promoter variants -571C and -90G which according to a predictive bioinformatics tool TFSearch (Heinemeyer et al. 1998) result in disruption of ADRI transcription binding sites which we speculate might interrupt with *APOBEC3G* expression. In addition, the haplotype contains the 557G allele which codes for the 186R variant in the APOBEC3G protein which results in reduced enzyme activity due to poor protein-protein structure. Therefore the C-G-G-C-C haplotype is likely to result in reduced antiviral activity of the APOBEC3G enzyme.

Contrary to this haplotypes C-C-A-T-C and G-G-A-T-G were only observed in the HIV uninfected group suggesting they could be protective against HIV. In support of the explanation above, C-C-A-T-C and G-G-A-T-G carry only one transcription factor binding site altering variants each (either -571C or -90G and not both) and a 557A (H186) allele which means they have the correct amino acid at position 186 for protein-protein interaction and their expression is unlikely to be significantly altered. We therefore suggest that *APOBEC3G* haplotypes may have a significant role in HIV infection among the Zimbabweans.

However, the effect of variation in the *APOBEC3G* gene on HIV outcomes depends on underlying factors such as the activity of other APOBEC proteins that also have antiviral properties (Henriet et al. 2009, OhAinle et al. 2006). For example, genetic polymorphisms and alternative splicing of *APOBEC3F* and *APOBEC3H* genes have been linked with differential antiretroviral activity of their protein products (Harari et al. 2009, Li et al. 2011).

The activity of APOBEC3G is also dependent on the functional activity of HIV Vif which antagonises the APOBEC3G antiviral properties by degrading the enzyme. APOBEC3G only exerts its antiviral effect in Vif deficient cells (Yu et al. 2004), thus, *APOBEC3G* variation does not matter in the presence of Vif. A study by Gourrand et al. (2012) showed that haplotype variation in *APOBEC3F* made the enzyme differentially susceptible to suppression of its antiviral activity by Vif (Gourraud et al. 2011). HIV-Vif makes use of the host Cullin 5-Elongin C E3 ubiquitin ligase complex to degrade APOBEC3G. Polymorphisms on the *CUL5* gene that encodes the ligase complex have also been shown to affect HIV disease progression (An et al. 2007). Given this complex network of interaction of proteins in the cytidine deamination antiviral pathway, determination of association between *APOBEC3G* genetic variants and HIV infection or disease progression would require large sample sizes and adjusting for other proteins involved, a motivation for genome-wide association studies. Another innate immune factor MBL was also studied in addition to the APOBEC3G.

4.3.2 *MBL2* polymorphism

Mannose binding lectin (MBL) is an acute phase protein involved in the non-specific response to invading micro-organisms through recognition of carbohydrate structures on their surface (Holmskov et al. 1994). MBL kills pathogens by activating the lectin pathway of the complement system, opsonisation, chemotaxis and/or direct killing (Vang Petersen et al. 2001, Kuhlman et al. 1989). MBL recognises and binds to high mannose residues on HIV gp120 resulting in activation of the complement system to kill HIV (Ezekowitz et al. 1989). By binding the gp120, MBL also blocks the attachment of HIV to host cells bearing gp120 receptors, thus, inhibiting infection of the host cell (Saifuddin et al. 2000). Deficiency in MBL protein has been linked with increased susceptibility to HIV.

MBL deficiency is a result of low *MBL2* gene expression and poor oligomerisation of MBL polypeptide units caused by polymorphism in the gene's promoter and exon 1regions (Boniotto et al. 2000, Garred et al. 2006). Therefore, this study investigated polymorphism in a 1187bp region of the *MBL2* gene spanning the promoter and exon 1.

Using the Sanger chain termination DNA sequencing, twelve polymorphic sites were detected in the region, one novel SNP, 10 previously reported and a multiple base deletion. We reported a new *MBL2* -595G>A SNP with a minor allele (-595A) frequency of 2% in the Zimbabwean population. According to a predictive bioinformatics tool TFSearch, the polymorphism does not interact with any transcription factor binding site and may not play a significant role in *MBL2* expression. We further investigated three of the four previously reported SNPs that have been linked with differential expression and oligomerisation of the MBL to determine their possible role in HIV transmission and neurocognitive development.

A frequency of 24% for the *MBL2* 170A allele in the Zimbabwean population was observed which is comparable to that observed among other Africans (20-30%) (Lipscombe et al. 1992, Thye et al. 2011). The mutant alleles of *MBL2* exon1 non-synonymous SNPs 170G>A (G57E), 154C>T (R52C) and 161G>A (G54D) are alternatively termed "*C*", "*D*" and "*B*" respectively whilst the promoter and 5′-UTR variants -221C/G and +4C>T are termed -221X/Y and +4P/Q respectively. Presence of any of the *C, D* or *B* allele is designated "*O*" allele while the wild-type is the "*A*" allele (Sumiya et al. 1991, Lipscombe et al. 1992, Madsen et al. 1994). The *C, B* and *D* alleles are unevenly distributed in populations of different ethnic origin (Garred et al. 2006). Thus, despite high frequency of the 170A allele in Zimbabweans and Africans in general, the allele is very rare in the both Asians (Ou et al. 2011) and Caucasians (1%) (Lipscombe et al. 1996).

Out of the 26 samples sequenced in our study, neither 154T nor 161A alleles were observed yet the 154T occurs in 15% and 25% of Caucasians and Asians respectively whilst the 161A is rare in most populations (<5%) (Lipscombe et al. 1996, Chen et al. 2009).

The uneven distribution of the 170A and 154T alleles in Caucasians and Africans does not seem to explain the differences in their HIV prevalences because the effect of the *B* and *C* alleles on MBL oligomerisation is almost the similar (Israels 2012). The resulting amino acid changes disrupt the α-helical structure of the MBL polypeptide chain, thus, interfering with the formation of functional oligomers (Eisen et al. 2003). This reduces the protein's antiviral activity and may increase susceptibility to a wide range of diseases.

In the 5′-UTR, the *MBL2* +4C (*P*) allele was observed in 45% of our study population which was comparable to what has been reported among other Africans (38-50%) but higher than that among Caucasians (17%) and Asians (12%) (Ou et al. 2011, Thye et al. 2011, Mangano et al. 2008). The distribution of the +4C allele in different populations suggests a possible contribution to the higher prevalence of HIV among Africans compared to Caucasians. However, this is contradicted by the lower frequency of the +4C among Asians (12%) compared to Caucasians (17%) because HIV prevalence is higher among Asians than Caucasians (Singh et al. 2008, Ou et al. 2011). In addition, the individual effect of the +4C allele on *MBL2* gene expression has been reported to be mild and often overshadowed by -221C>G variants (Israels et al. 2012).

A frequency of the 12% for the -221C allele was observed which was in the same range with what has been reported among other Africans (12-17%), Caucasians (17%) and Asians (15%) (Singh et al. 2008, Ou et al. 2011, Thye et al. 2011).

We therefore speculate that -221C>G variants do not have a significant individual role in the distribution of HIV in different populations. Therefore, the distribution of individual *MBL2* variants does not comprehensively explain differences in HIV distribution in world populations.

The distribution of *MBL2* haplotypes, like that of the individual SNPs is uneven among world populations but it also does not show a trend that mirrors HIV distribution. *MBL2* polymorphism is likely to have been influenced by other micro-organisms more ancient than HIV. For example, the *MBL2* 170A (*C*) allele which results in MBL deficiency may have been selected for in African populations because it protects the host against MBL-mediated *Mycobacterium africanum* infection (Thye et al. 2011, Søborg et al. 2007). MBL-mediated opsonisation is used by some intracellular organisms such as *M. africanum* to enter host cells.

Several other organisms have also been linked to *MBL2* polymorphism (Eisen et al. 2003). In opposition of this linked to ancient organisms, reports on the evolutionary analysis of *MBL2* support neutral evolution involving migration and genetic drift to have shaped *MBL2* allele distribution as opposed to natural selection (Verdu et al.). It is also possible that because in addition to the lectin (MBL) pathway, complement system can also be activated by two other pathways (Endo et al. 2006) reduced MBL in circulation may not be detrimental to life, thus, unlikely to exert selective pressure (Mangano et al. 2008). To determine the possible effect of *MBL2* variation on HIV infection in our study population, we compared genotype frequencies of 170G>A, -221C>G and +4C>T SNPs between HIV-exposed infected (EI) and exposed-uninfected (EU) groups.

The homozygous +4C/C genotype was associated with increased risk of HIV infection compared to +4T/T and +4C/T genotypes but fell short of statistical significance (p=0.08). This can be explained by the down-regulatory effect of the +4C allele on *MBL2* expression. The individual effect of +4C allele on HIV transmission has been down-played due the stronger effect of -221C>G and *MBL2* A>O loci on MBL function (Mangano et al. 2008). However, our study provides evidence for an independent association of the +4C variant with HIV risk. When *MBL2* genotypes at -221C>G, +4C>T and 170G>A loci where combined in a pair-wise fashion and their association with HIV status analysed the *MBL2* -221G/G-4C/T combination was found to have a protective role against HIV infection as its frequency was significantly higher in the HIV EU group (44%) compared to the EI (22%) children (P=0.03). This is because the -221G and +4T alleles up-regulate *MBL2* gene expression, thus, the -221G/G-4C/T genotype is likely to encode sufficient MBL for antiviral activity.

On the contrary, the +4C/C-170G/A genotype combination was associated with increased risk of HIV infection with borderline significance (p=0.06). The +4C/C-170G/A genotype carries alleles that both down regulate *MBL2* expression (+4C) and interrupt MBL multimetrisation making an MBL deficiency genotype. This leads to impaired MBL-mediated complement activation, viral opsonisation and killing, thus leaving the body exposed to HIV infection (Garred et al. 2006, Saifuddin et al. 2000). Our observations are supported by reports that MBL deficiency variants such as *MBL2 O/O* and *MBL2 XA/XA* may be strong risk factors for HIV infection compared to MBL sufficiency variants (Kuhn et al. 2006, Boniotto et al. 2003). Others have reported accelerated disease progression in individuals carrying MBL deficiency variants (Singh et al. 2008, Mangano et al. 2008). A study among Americans reported that MBL deficiency due to genetic polymorphism had a more detrimental effect in children below two years of age (Singh et al. 2008).

4.4 Genotype Combinations and HIV infection

HIV/AIDS is a multifactorial condition affected by several pathways in the innate and adaptive immune systems. Several antiviral factors work simultaneously in an attempt to mount immunity against HIV. Therefore, co-occurrence of HIV/AIDS restriction gene variants in an individual may give an amplified effect compared to the individual SNPs. Given this, genotypes of some SNPs were combined in a pair-wise fashion to determine their possible role in HIV infection.

The frequency of *CCR2* 190G/A-*CX3CR*1 745G/G genotype combination was significantly (P=0.002) higher in HIV EU children (33%) than EI children (0%) suggesting that this genotype combination may be having a protective role against HIV infection (Table 3.9). This observation and reasoning is supported by the functional roles of alleles in the individual genotypes. The *CCR2* 190A has been reported to confer a protective role by facilitating the dimerization of the CCR2 64I protein to CXCR4 making CXCR4 less available for HIV attachment. The allele has been reported to act in a dominant manner where its presence even in a heterozygote is fully expressed. The protein product of *CX3CR1* 745G (249V) has stronger capacity to bind to fractalkine compared to the 249I, resulting in a stronger chemotactic effect conferring a more viable immune response. The combined effect of these two protective mechanisms may account for the HIV resistance exhibited by individuals carrying *CCR2* 190G/A-*CX3CR1* 745G/G.

The *CCR2* 190G/A-*APOBEC3G* 557A/A combination was associated with low risk of HIV infection as it occurred in 17% of the HIV EU individuals whilst it was absent in the HIV EI group (P=0.015) (Table 3.9). *APOBEC3G* 557G variant has been associated with accelerated disease progression (An et al. 2002, Reddy et al. 2010), thus, the *APOBEC3G* 557A variant may be protective against HIV as it gives the correct amino acid configuration for protein-protein interaction of APOBEC3G enzyme in a reversal of the mechanism discussed earlier. The *CCR2* 190G/A as mentioned earlier carries the protective 190A allele. Therefore, our study points out that given a *CCR2* 190G/A background, the protective effect of 557A/A may reduce risk HIV infection.

In contrast, *CX3CR1* 745A/A-*RANTES* In1.1T/T genotype was significantly higher in HIV EI children (10%) compared to the EU (0%) suggesting that it might be a risk factor for HIV infection (p=0.05) (Table 3.9). Paradoxically *RANTES* In1.1T allele has been described to up-regulate the expression of *RANTES* therefore making available more of the chemokine for competitive inhibition of HIV binding. Despite this, Rathore et al. (2008) reported the In1.1T to be a risk factor for HIV infection (Rathore et al. 2008). Our study supports Rathore's observation but on a *CX3CR1* 745A/A genotype background. The *CX3CR1* 745A/A genotype background has a villain role in HIV infection where its resultant protein has weaker binding to its ligand and therefore poor immune activation. Genotype combinations are therefore more likely to have an effect on HIV-infection compared to individual SNPs because of the multifactorial nature of anti-HIV immunity.

4.5 Genetic polymorphism and neurocognitive function

Neurocognitive impairments ranging from mild cognitive deficits to HAD are a common feature in HIV/AIDS patients (Heaton et al. 2011). Although HIV enters the CNS in the early stages of infection neurocognitive complications only surface later when the clinical state of the patient deteriorates (Gendelman et al. 1997, Hazleton et al. 2010). This suggests that neurocognitive impairments in HIV maybe a result of disease progression. There seems to be inter-individual variation in the susceptibility and prognosis of neurocognitive disorders even in patients who are on ART. A fraction of HIV-negative children also experience neurocognitive impairments suggesting that etiological factors other than HIV are at play (Bagenda et al. 2006). We therefore investigated the relationship between genetic variation and neurocognitive function in HIV-infected and HIV-uninfected children.

We observed a general tendency to poor neurocognitive performance in HIV-infected children compared to the uninfected (Table 3.10). This is not surprising as presence of HIV protein such as gp120 and Tat in the central nervous system causes microglia and astrocytes to secrete neurotoxins that result in neuronal damage (Strazza et al. 2011, Fritz-French et al. 2012). HIV patients are also at risk of developing encephalitis due to cytokine overexpression and opportunistic infections such as cytomegalovirus, Epstein-Barr virus and others (Strazza et al. 2011, Polilli et al. 2010). These are likely to interfere with neurological function resulting in HIV-associated neurocognitive disorders. Our observations are supported by other reports on higher risk of neurocognitive impairments among HIV-infected children compared to the HIV-uninfected (Boivin et al. 1995, Kandawasvika et al. 2011).

Other factors such as nutrition, family and general environment have been reported to influence cognitive ability in children (Boivin et al. 1995, Eskenazi et al. 1999, Mendola et al. 2002). However, the focus of our study was the association between genetic variation and neurocognitive impairment.

We found the frequency of *APOBEC3G* -90C/G and *MBL2* 170G/A (A/C) significantly higher in children exhibiting poor cognitive ability compared to those with normal cognitive ability (P=0.048 and 0.046 respectively) among HIV infected children. These observations suggest that *APOBEC3G* -90C/G and *MBL2* 170 G/A genotypes are possible risk factors for neurocognitive impairment in HIV-infected children compared to the *APOBEC3G* -90G/G and *MBL2* G/A respectively. APOBEC3G and MBL are both innate immune factors involved in antiviral activities against HIV. To our knowledge there is no evidence in literature linking APOBEC3G -90C>G variants to HIV-associated neurocognitive impairment. However, the APOBEC3G -90C/G may be involved in the differential expression of the APOBEC3G protein and studies have linked low APOBEC3G protein levels to faster HIV disease progression (Jin et al. 2007). Another possibility is that since deficiency of apolipoprotein B containing lipoproteins in serum leads to peripheral neuropathy (Rampoldi et al. 2002) and APOBEC3G is an apolipoprotein B mRNA enzyme there could be a link to neurocognitive ability.

The association between *MBL2* 170G/A genotype and neurocognitive impairment in HIV-infected children can be explained by the MBL deficiency resulting from carrying the 170A (*C*) allele (Garred et al. 2006). This observation is supported by reports of a tendency towards neurocognitive impairment among American HIV-infected children with the *MBL2 O/O* genotype compared to the *A/A* genotype (Singh et al. 2008).

Another study also reported an association between *O/O* genotype and increased risk of neurocognitive decline among Chinese HIV-infected adults compared to the *A/A* genotype (Spector et al. 2010), However, a recent genome wide association study (GWAS) on an adult population did not find any SNP to be associated with neurocognitive impairment and/or HAD (Levine et al. 2012)(). In addition to the possible role of *MBL2* polymorphism in neurocognitive function in HIV-infected children, we also observed its effect among the HIV-negative children.

We report a significantly higher frequency of the *MBL2* +4C/C (*P/P*) and -221C/G (*X/Y*) genotypes in HIV-uninfected children who are verbally impaired compared to the HIV negative verbally unimpaired (P=0.03 and 0.05 respectively) (Table 3.11). *MBL2* -221C/G was also observed more frequently in HIV-uninfected children with a general NCI score below 68 when compared to those above 68. These observations provide evidence that +4C/C and -221C/G genotypes are possible risk factors for neurocognitive impairment regardless of HIV status. The -221C allele is associated with MBL deficiency which may allow infection by micro-organisms that target the CNS (Garred et al. 2006). A study among Chinese HIV-negative adults reported an association between the MBL deficiency *O/O* genotype and cryptococcal meningitis (Ou et al. 2011) further supporting our observation that *MBL2* polymorphism may affect neurocognitive outcomes regardless of HIV status. We therefore report a possible role of *MBL2* polymorphism in both HIV-associated and HIV-free neurocognitive status. Our study makes a great contribution towards understanding host genetic variation with regards to HIV infection and neurocognitive development among Zimbabweans and Africans in general. However, like any other study we had a few limitations.

4.6 Limitations of the study and potential redress

The major challenge in our study was the limited sample size. We could not do anything to increase the number of HIV-exposed infected children because the study drew samples from a decade long cohort and numbers were restricted to participants who were accessible and meeting the inclusion criteria. However, to boost the sample size, a 1:2 ratio of HIV positive to HIV negative was used. This allowed for allele frequencies to be approximated more accurately as evidenced by the fact that allele distribution in our control population conformed to Hardy Weinberg Equilibrium for all SNPs studied meaning there was no apparent sampling bias.

In addition to sample size, another limitation of the study was the possibility of survivor bias in our study population. Since the study population was limited to individuals alive after close to a decade of follow-up, information on the genetic make-up of those children who died during follow-up period was not available. This means that comparison of genotype and allele frequencies between HIV-exposed infected and HIV-exposed uninfected might not be a correct approximation of the relationship between genetic variation and risk of HIV infection. A large fraction of HIV-infected children succumb to effects of the virus before the age of five. It can be speculated that if any genetic variants are associated with faster disease progression and possible death then these could have been reduced more in the available HIV-infected population compared to the uninfected after nine years of follow-up.

It has also been argued that maternal virological, immunological and genetic factors may influence the chances of a mother transmitting HIV to their child. Mothers exhibiting advanced clinical signs and symptoms are more likely to transmit HIV to their children (Tiemessen et al. 2006).

Differences in maternal viral loads mean that children's exposure to HIV during pregnancy, birth and breastfeeding is different and therefore may determine the risk of infection despite child's genetic make-up. HIV co-infections such as STIs, malaria and TB may also increase chances of a mother transmitting HIV to their child (Gumbo et al. 2010, Gupta et al. 2011, Kliks et al. 1994). Depleted immunological status of the mother as reflected by low CD4+ T-cell count also indicates increased exposure to HIV (Singh et al. 2009). These factors, if considered in a study like ours could have given a more holistic picture.

The role of host genetic variation in HIV transmission could have been elucidated better if maternal genetics had also been included in the study. It has been shown that KIR-HLA heterogeneity between mother and child influenced the risk of vertical transmission of HIV (Paximadis et al. 2011). Reports among Kenyans show that infants were more likely to be infected perinatally by HIV if the mother had the *SDF*1 801G/A genotype (Mabuka et al. 2009). It would be important to consider maternal genetic variation in future work.

4.7 Conclusion and future prospects

Despite the mentioned limitations our study makes a huge contribution in understanding the genetic correlation of HIV infection and neurodevelopmental deficits. Our findings pave way for larger and more focused studies taking a genomics rather than a candidate gene approach. It is learnt from our findings that host gene variants are more likely to work together as opposed to individual SNPs in influencing outcomes of multifactorial conditions such as HIV/AIDS.

RANTES, *CX3CR1*and *APOBEC3G* variants did not show association with HIV infection when analysed individually but combination of genotypes that enhance their antiviral activities where more frequent in the HIV-exposed uninfected children than the infected strongly arguing for a synergistic antiviral effect. This combinational effect points to critical pathways in HIV infection that may be exploited for therapeutic benefits.

HIV is a multifactorial condition where different immune pathways work simultaneously with a possibility of masking or enhancing one another. We therefore recommend that future work takes advantage of advances in DNA technology such as microarrays and next generation sequencing to carry out GWAS on HIV and HIV-related outcomes. Protein profiling to confirm effect of polymorphism on disease outcomes is also recommended to enhance the interpretation of raw DNA sequences. Proteomics and transcriptomes have shown new pathways involving the dysregulation of gene expression due to presence of HIV in the host (Kartvelishvili et al. 2004) so there is need to integrate genetic variants and their resultant protein profiles.

We observed an association between *MBL2* gene variants and neurocognitive function regardless of HIV suggesting that gene variants that regulate the immune system are a potentially useful tool in predicting the outcome of neurodevelopmental status and diseases. Due to the wide use of ART, children who are perinatally infected by HIV now have a good chance to survive into adulthood therefore *MBL2*, chemokine and chemokine receptor variants can be assembled into a predictive tool for HIV-related and neurocognitive outcomes. This way, the vulnerable or high risk individuals can be identified and monitored closely. An accurate correlation of clinical, genetic and immunological profiling in such cases could prove useful in prioritising care. We therefore recommend more research goes into the role of genetic profiles in neurocognitive status for potential use in risk prediction.

A novel MBL2 -595G>A SNP was observed after sequencing 1187bp in the *MBL2* gene. It can be deduced from this observation that there may be more undetected gene variants in African populations as most have not been studied at genome level. These SNPs may be able to answer some disease related questions in the African population. This further argues for GWAS among Africans.

5 REFERENCES

AGACHAN, B., et al, 2010. Association of Monocyte Chemotactic Protein-1 and CC Chemokine Receptor 2 Gene Variants with Preeclampsia. *Journal of Interferon & Cytokine Research*, vol. 30, no. 9, pp. 673-676.

AHLENSTIEL, G., et al, 2005. Distribution and Effects of Polymorphic RANTES Gene Alleles in HIV/HCV Coinfection-a Prospective Cross-Sectional Study. *World Journal of Gastroenterology*, vol. 11, no. 48, pp. 7631.

ALAGARASU, K., et al, 2009. CCR2, MCP-1, SDF-1α & DC-SIGN Gene Polymorphisms in HIV-1 Infected Patients with & without Tuberculosis. *Indian Journal of Medical Research*, vol. 130, no. 4, pp. 444.

AMANATIDOU, V., et al, 2006. T280M Variation of the CX3C Receptor Gene is Associated with Increased Risk for Severe Respiratory Syncytial Virus Bronchiolitis. *The Pediatric Infectious Disease Journal*, vol. 25, no. 5, pp. 410.

AMARA, S., DOMENECH, J. and JENHANI, F., 2010. Stromal Cell-Derived Factor 1 Polymorphism in Patients Infected with HIV and Implications for AIDS Progression in Tunisia. *HIV/AIDS (Auckland, NZ)*, vol. 2, pp. 203.

AN, P., et al, 2004. APOBEC3G Genetic Variants and their Influence on the Progression to AIDS. *Journal of Virology*, vol. 78, no. 20, pp. 11070-11076.

AN, P., et al, 2007. Polymorphisms of CUL5 are Associated with CD4 T Cell Loss in HIV-1 Infected Individuals. *PLoS Genetics*, vol. 3, no. 1, pp. e19.

AN, P., et al, 2002. Modulating Influence on HIV/AIDS by Interacting RANTES Gene Variants. *Proceedings of the National Academy of Sciences*, vol. 99, no. 15, pp. 10002.

ANNUNZIATA, P., 2003. Blood-Brain Barrier Changes during Invasion of the Central Nervous System by HIV-1. *Journal of Neurology*, vol. 250, no. 8, pp. 901-906.

ANZALA, A., et al, 1998. CCR2-64I Allele and Genotype Association with Delayed AIDS Progression in African Women. *Lancet*, vol. 351, pp. 1632-1633.

APOSTOLAKIS, S., BARITAKI, S., KRAMBOVITIS, E. and SPANDIDOS, D., 2005. Distribution of HIV/AIDS Protective SDF1, CCR5 and CCR2 Gene Variants within Cretan Population. *Journal of Clinical Virology*, vol. 34, no. 4, pp. 310-314.

ARA, T., et al, 2003. Impaired Colonization of the Gonads by Primordial Germ Cells in Mice Lacking a Chemokine, Stromal Cell-Derived Factor-1 (SDF-1). *Proceedings of the National Academy of Sciences*, vol. 100, no. 9, pp. 5319.

BACKE, E., et al, 1992. Demonstration of HIV-1 Infected Cells in Human Placenta by in Situ Hybridisation and Immunostaining. *Journal of Clinical Pathology*, vol. 45, no. 10, pp. 871-874.

BAGENDA, D., et al, 2006. Health, Neurologic, and Cognitive Status of HIV-Infected, Long-Surviving, and Antiretroviral-Naive Ugandan Children. *Pediatrics*, vol. 117, no. 3, pp. 729-740.

BAILEY, R.C., et al, 1999. Growth of Children According to Maternal and Child HIV, Immunological and Disease Characteristics: A Prospective Cohort Study in Kinshasa, Democratic Republic of Congo. *International Journal of Epidemiology*, vol. 28, no. 3, pp. 532-540.

BAJETTO, A., et al, 2001. Chemokines and their Receptors in the Central Nervous System. *Frontiers in Neuroendocrinology*, vol. 22, no. 3, pp. 147-184.

BALOTTA, C., et al, 1999. Identification of Two Distinct Subsets of Long-Term Nonprogressors with Divergent Viral Activity by Stromal-Derived Factor 1 Chemokine Gene Polymorphism Analysis. *Journal of Infectious Diseases*, vol. 180, no. 2, pp. 285-289.

BANKS, W.A., et al, 2001. Transport of Human Immunodeficiency Virus Type 1 Pseudoviruses Across the Blood-Brain Barrier: Role of Envelope Proteins and Adsorptive Endocytosis. *Journal of Virology*, vol. 75, no. 10, pp. 4681-4691.

BARNES, D.A., et al, 1996. Induction of RANTES Expression by Astrocytes and Astrocytoma Cell Lines. *Journal of Neuroimmunology*, vol. 71, no. 1, pp. 207-214.

BARRÉ-SINOUSSI, F., et al, 1983. Isolation of a T-Lymphotropic Retrovirus from a Patient at Risk for Acquired Immune Deficiency Syndrome (AIDS). *Science*, vol. 220, no. 4599, pp. 868-871.

BAZAN, J.F., et al, 1997. A New Class of Membrane-Bound Chemokine with a CX3C Motif.

BELPERIO, J.A., et al, 2000. CXC Chemokines in Angiogenesis. *Journal of Leukocyte Biology*, vol. 68, no. 1, pp. 1-8.

BEN-SELMA, W., et al, 2011. Polymorphisms in the RANTES Gene Increase Susceptibility to Active Tuberculosis in Tunisia. *DNA and Cell Biology*, vol. 30, no. 10, pp. 789-800.

BIESECKER, G., PODACK, E., HALVERSON, C. and MÜLLER-EBERHARD, H., 1979. C5b-9 Dimer: Isolation from Complement Lysed Cells and Ultrastructural Identification with Complement-Dependent Membrane Lesions. *The Journal of Experimental Medicine*, vol. 149, no. 2, pp. 448-458.

BINDER, N.B., et al, 2009. Estrogen-Dependent and CC Chemokine Receptor-2–dependent Pathways Determine Osteoclast Behavior in Osteoporosis. *Nature Medicine*, vol. 15, no. 4, pp. 417-424.

BISHOP, K.N., et al, 2004. Cytidine Deamination of Retroviral DNA by Diverse APOBEC Proteins. *Current Biology*, vol. 14, no. 15, pp. 1392-1396.

BISIACCHI, P.S., SUPPIEJ, A. and LAVERDA, A., 2000. Neuropsychological Evaluation of Neurologically Asymptomatic HIV-Infected Children. *Brain and Cognition*, Jun-Aug, vol. 43, no. 1-3, pp. 49-52 ISSN 0278-2626; 0278-2626.

BIZINOTO, M.C., LEAL, É., DIAZ, R.S. and JANINI, L.M., 2011. Loci Polymorphisms of the APOBEC3G Gene in HIV Type 1-Infected Brazilians. *AIDS Research and Human Retroviruses*, vol. 27, no. 2, pp. 137-141.

BLEUL, C.C., et al, 1996. The Lymphocyte Chemoattractant SDF-1 is a Ligand for LESTR/fusin and Blocks HIV-1 Entry.

BLEUL, C.C., et al, 1996b. A Highly Efficacious Lymphocyte Chemoattractant, Stromal Cell-Derived Factor 1 (SDF-1). *The Journal of Experimental Medicine*, vol. 184, no. 3, pp. 1101-1109.

BOILY-LAROUCHE, G., et al, 2009. Functional genetic variants in DC-SIGNR are associated with mother-tochild transmission of HIV-1. *PLoS One*, vol 4, e7211.

BOILY-LAROUCHE, G., et al, 2012. Naturally-Occurring Genetic Variants in Human DC-SIGN Increase HIV-1 Capture, Cell-Transfer and Risk of MotherTo-Child Transmission. *PLoS ONE*, vol. 7, no. 7, e40706.

BOIVIN, M.J., et al, 1995. A Preliminary Evaluation of the Cognitive and Motor Effects on Pediatric HIV Infection in Zairian Children. *Health Psychology*, vol. 14, no. 1, pp. 13.

BOKOCH, G.M., 1995. Chemoattractant Signaling and Leukocyte Activation. *Blood*, vol. 86, no. 5, pp. 1649-1660.

BOL, SM., et al, 2012. Single Nucleotide Polymorphism in Gene Encoding Transcription Factor Prep1 is Associated with HIV-1-Assocaited Dementia. PLoS ONE, vol. 7, no. 2. e30990.

BONAVIA, R., et al, 2003. Chemokines and their Receptors in the CNS: Expression of CXCL12/SDF-1 and CXCR4 and their Role in Astrocyte Proliferation. *Toxicology Letters*, vol. 139, no. 2-3, pp. 181-189.

BONIOTTO, M., et al, 2003. MBL2 Polymorphisms are Involved in HIV-1 Infection in Brazilian Perinatally Infected Children. *Aids*, vol. 17, no. 5, pp. 779.

BONIOTTO, M., et al, 2000. Polymorphisms in the MBL2 Promoter Correlated with Risk of HIV-1 Vertical Transmission and AIDS Progression. *Genes and Immunity*, vol. 1, no. 5, pp. 346.

BRAND, S., et al, 2006. Increased Expression of the Chemokine Fractalkine in Crohn's Disease and Association of the Fractalkine Receptor T280M Polymorphism with a Fibrostenosing Disease Phenotype. *The American Journal of Gastroenterology*, vol. 101, no. 1, pp. 99-106.

BROUWER, K.C., et al, 2005. Effect of CCR2 Chemokine Receptor Polymorphism on HIV Type 1 Mother-to-Child Transmission and Child Survival in Western Kenya. *AIDS Research & Human Retroviruses*, vol. 21, no. 5, pp. 358-362.

BROUWER, M.C., et al, 2009. Host Genetic Susceptibility to Pneumococcal and Meningococcal Disease: A Systematic Review and Meta-Analysis. *The Lancet Infectious Diseases*, vol. 9, no. 1, pp. 31-44.

BROUWERS, P., TUDOR-WILLIAMS, G., DECARLI, C. and MOSS, H.A., 1995. Relation between Stage of Disease and Neurobehavioral Measures in Children with Symptomatic HIV Disease. *AIDS*.

BRUMME, Z.L., et al, 2003. Influence of Polymorphisms within the CX3CR1 and MDR-1 Genes on Initial Antiretroviral Therapy Response. *Aids*, vol. 17, no. 2, pp. 201.

BRYSON, Y.J., LUZURIAGA, K., SULLIVAN, J.L. and WARA, D.W., 1992. Proposed Definitions for in Utero Versus Intrapartum Transmission of HIV-1. *New England Journal of Medicine*, vol. 327, no. 17, pp. 1246-1247.

BUONAGURO, L., TORNESELLO, M. and BUONAGURO, F., 2007. Human Immunodeficiency Virus Type 1 Subtype Distribution in the Worldwide Epidemic: Pathogenetic and Therapeutic Implications. *Journal of Virology*, vol. 81, no. 19, pp. 10209-10219.

BURKE, D.S., REDFIELD, R.R., PUTMAN, P. and ALEXANDER, S.S., 1987. Variations in Western Blot Banding Patterns of Human T-Cell Lymphotropic Virus Type III/lymphadenopathy-Associated Virus. *Journal of Clinical Microbiology*, vol. 25, no. 1, pp. 81-84.

CALMAN, A., et al, 1988. Transcription and Replication of Human Immunodeficiency Virus-1 in B Lymphocytes in Vitro. *AIDS (London, England)*, vol. 2, no. 3, pp. 185.

CARDONA, A.E., et al, 2008. Scavenging Roles of Chemokine Receptors: Chemokine Receptor Deficiency is Associated with Increased Levels of Ligand in Circulation and Tissues. *Blood*, vol. 112, no. 2, pp. 256.

CARRINGTON, M. and ALTER, G., 2012. Innate Immune Control of HIV. *Cold Spring Harbor Perspectives in Medicine*, vol. 2, no. 7.

CARRINGTON, M., et al, 1997. Novel Alleles of the Chemokine-Receptor Gene< i> CCR5</i>. *The American Journal of Human Genetics*, vol. 61, no. 6, pp. 1261-1267.

CHAN, D.C. and KIM, P.S., 1998. HIV Entry and its Inhibition. *Cell*, vol. 93, no. 5, pp. 681.

CHARO, I.F., et al, 1994. Molecular Cloning and Functional Expression of Two Monocyte Chemoattractant Protein 1 Receptors Reveals Alternative Splicing of the Carboxyl-Terminal Tails. *Proceedings of the National Academy of Sciences*, vol. 91, no. 7, pp. 2752.

CHATTERJEE, K., DANDARA, C., HOFFMAN, M. and WILLIAMSON, A.L., 2010. CCR2-V64I Polymorphism is Associated with Increased Risk of Cervical Cancer but Not with HPV Infection Or Pre-Cancerous Lesions in African Women. *BMC Cancer*, vol. 10, no. 1, pp. 278.

CHEN, J., et al, 2009. Mannose-Binding Lectin Polymorphisms and Recurrent Respiratory Tract Infection in Chinese Children. *European Journal of Pediatrics*, vol. 168, no. 11, pp. 1305-1313.

CHEN, M.K., et al, 2011. CCR2-64I Gene Polymorphism Increase Susceptibility to Oral Cancer. *Oral Oncology*.

CHIAPPINI, E., et al, 2006. Virologic, Immunologic, and Clinical Benefits from Early Combined Antiretroviral Therapy in Infants with Perinatal HIV-1 Infection. *Aids*, vol. 20, no. 2, pp. 207.

CHOE, H., et al, 1996. The β-Chemokine Receptors CCR3 and CCR5 Facilitate Infection by Primary HIV-1 Isolates. *Cell*, vol. 85, no. 7, pp. 1135-1148.

CLARK, H.W., REID, K. and SIM, R.B., 2000. Collectins and Innate Immunity in the Lung. *Microbes and Infection*, vol. 2, no. 3, pp. 273-278.

COHEN, M.S., SHAW, G.M., MCMICHAEL, A.J. and HAYNES, B.F., 2011. Acute HIV-1 Infection. *New England Journal of Medicine*, vol. 364, no. 20, pp. 1943-1954.

COHN JR, S. and WEAVER, L., 2006. The Black Death and AIDS: CCR5-Δ32 in Genetics and History. *QJM*, vol. 99, no. 8, pp. 497-503.

COMBADIERE, C., et al, 1998. Identification of CX 3CR1. *Journal of Biological Chemistry*, vol. 273, no. 37, pp. 23799-23804.

COOKE, G.S., et al, 2006. A Polymorphism that Reduces RANTES Expression is Associated with Protection from Death in HIV-Seropositive Ugandans with Advanced Disease. *Journal of Infectious Diseases*, vol. 194, no. 5, pp. 666-669.

COTTER, R., et al, 2002. Fractalkine (CX3CL1) and Brain Inflammation: Implications for HIV-1-Associated Dementia. *Journal of Neurovirology*, vol. 8, no. 6, pp. 585-598.

CRAWFORD, A., et al, 2011. A Role for the Chemokine RANTES in Regulating CD8 T Cell Responses during Chronic Viral Infection. *PLoS Pathogens*, vol. 7, no. 7, pp. e1002098.

CROSDALE, D., et al, 2000. Mannose Binding Lectin (MBL) Genotype Distributions with Relation to Serum Levels in UK Caucasoids. *European Journal of Immunogenetics*, vol. 27, no. 3, pp. 111-117.

CULLEY, F.J., et al, 2006. Role of CCL5 (RANTES) in Viral Lung Disease. *Journal of Virology*, vol. 80, no. 16, pp. 8151-8157.

CYSIQUE, L., et al, 2009. Dynamics of Cognitive Change in Impaired HIV-Positive Patients Initiating Antiretroviral Therapy. *Neurology*, vol. 73, no. 5, pp. 342-348.

DAUGHERTY, B.L. and SPRINGER, M.S., 1997. The Beta-Chemokine Receptor Genes CCR1 (CMKBR1), CCR2 (CMKBR2), and CCR3 (CMKBR3) Cluster within 285 Kb on Human Chromosome 3p21. *Genomics*, Apr 15, vol. 41, no. 2, pp. 294-295 ISSN 0888-7543; 0888-7543. DOI 10.1006/geno.1997.4626.

DAVIS, L., et al, 1992. Early Viral Brain Invasion in Iatrogenic Human Immunodeficiency Virus Infection. *Neurology*, vol. 42, no. 9, pp. 1736-1736.

DE COCK, K.M., et al, 2000. Prevention of Mother-to-Child HIV Transmission in Resource-Poor Countries. *JAMA: The Journal of the American Medical Association*, vol. 283, no. 9, pp. 1175-1182.

DE MAIO, F.A., et al, 2011. Effect of HIV-1 Vif Variability on Progression to Pediatric AIDS and its Association with APOBEC3G and CUL5 Polymorphisms. *Infection, Genetics and Evolution*.

DE ROSSI, A., et al, 1996. Dynamics of Viral Replication in Infants with Vertically Acquired Human Immunodeficiency Virus Type 1 Infection. *Journal of Clinical Investigation*, vol. 97, no. 2, pp. 323.

DEEKS, S.G., et al, 2004. Immune Activation Set Point during Early HIV Infection Predicts Subsequent CD4 T-Cell Changes Independent of Viral Load. *Blood*, vol. 104, no. 4, pp. 942-947.

DENG, H.K., et al, 1996. Of a Major Co-Receptor for Primary Isolates of HIV-1. *Nature*, vol. 381, pp. 20.

DICKOVER, R.E., GARRATTY, E.M., PLAEGER, S. and BRYSON, Y.J., 2001. Perinatal Transmission of Major, Minor, and Multiple Maternal Human Immunodeficiency Virus Type 1 Variants in Utero and Intrapartum. *Journal of Virology*, vol. 75, no. 5, pp. 2194-2203.

DO, H., et al, 2005. Exhaustive Genotyping of the CEM15 (APOBEC3G) Gene and Absence of Association with AIDS Progression in a French Cohort. *Journal of Infectious Diseases*, vol. 191, no. 2, pp. 159.

EISEN, D.P. and MINCHINTON, R.M., 2003. Impact of Mannose-Binding Lectin on Susceptibility to Infectious Diseases. *Clinical Infectious Diseases*, vol. 37, no. 11, pp. 1496-1505.

ENDO, Y., TAKAHASHI, M. and FUJITA, T., 2006. Lectin Complement System and Pattern Recognition. *Immunobiology*, vol. 211, no. 4, pp. 283-293.

ESKENAZI, B. and CASTORINA, R., 1999. Association of Prenatal Maternal Or Postnatal Child Environmental Tobacco Smoke Exposure and Neurodevelopmental and Behavioral Problems in Children. *Environmental Health Perspectives*, vol. 107, no. 12, pp. 991.

EZEKOWITZ, R., KUHLMAN, M., GROOPMAN, J.E. and BYRN, R.A., 1989. A Human Serum Mannose-Binding Protein Inhibits in Vitro Infection by the Human Immunodeficiency Virus. *The Journal of Experimental Medicine*, vol. 169, no. 1, pp. 185-196.

FÄTKENHEUER, G., et al, 2005a. Efficacy of Short-Term Monotherapy with Maraviroc, a New CCR5 Antagonist, in Patients Infected with HIV-1. *Nature Medicine*, vol. 11, no. 11, pp. 1170-1172.

FAURE, E. and ROYER-CARENZI, M., 2008. Is the European Spatial Distribution of the HIV-1-Resistant CCR5-[Delta] 32 Allele Formed by a Breakdown of the Pathocenosis due to the Historical Roman Expansion?. *Infection, Genetics and Evolution*, vol. 8, no. 6, pp. 864-874.

FAURE, S., et al, 2000. Rapid Progression to AIDS in HIV Individuals with a Structural Variant of the Chemokine Receptor CX3CR1. *Science*, vol. 287, no. 5461, pp. 2274-2277.

FAURE, S., et al, 2003. Deleterious Genetic Influence of CX3CR1 Genotypes on HIV-1 Disease Progression. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, vol. 32, no. 3, pp. 335.

FENG, Y., BRODER, C.C., KENNEDY, P.E. and BERGER, E.A., 1996. HIV-1 Entry Cofactor: Functional cDNA Cloning of a Seven-Transmembrane, G Protein-Coupled Receptor. *Science*, vol. 272, no. 5263, pp. 872-877.

FOUSSAT, A., et al, 2001. Deregulation of the Expression of the fractalkine/fractalkine Receptor Complex in HIV-1–infected Patients. *Blood*, vol. 98, no. 6, pp. 1678-1686.

FOWKE, K.R., et al, 1996. Resistance to HIV-1 Infection among Persistently Seronegative Prostitutes in Nairobi, Kenya. *The Lancet*, vol. 348, no. 9038, pp. 1347-1351.

FRANKEL, A.D. and YOUNG, J.A.T., 1998. HIV-1: Fifteen Proteins and an RNA. *Annual Review of Biochemistry*, vol. 67, no. 1, pp. 1-25.

FREDERICK, T., et al, 1994. Progression of Human Immunodeficiency Virus Disease among Infants and Children Infected Perinatally with Human Immunodeficiency Virus Or through Neonatal Blood Transfusion. Los Angeles County Pediatric AIDS Consortium and the Los Angeles County-University of Southern California Medical Center and the University of Southern California School of Medicine. *The Pediatric Infectious Disease Journal*, vol. 13, no. 12, pp. 1091.

FRITZ-FRENCH, C. and TYOR, W., 2012. Interferon-α (IFNα) Neurotoxicity. *Cytokine & Growth Factor Reviews*.

GABUZDA, D.H., et al, 1992. Role of Vif in Replication of Human Immunodeficiency Virus Type 1 in CD4 T Lymphocytes. *Journal of Virology*, vol. 66, no. 11, pp. 6489-6495.

GALLAGHER, S.R. and DESJARDINS, P.R., 1989. Quantitation of DNA and RNA with Absorption and Fluorescence Spectroscopy. *Current Protocol in Molecular Biology*.

GARCIA-GARCIA, E., ANDRIEUX, K., GIL, S. and COUVREUR, P., 2005. Colloidal Carriers and Blood-Brain Barrier (BBB) Translocation: A Way to Deliver Drugs to the Brain?. *International Journal of Pharmaceutics*, vol. 298, no. 2, pp. 274-292.

GARRED, P., et al, 2006. Mannose-Binding Lectin and its Genetic Variants. *Genes and Immunity*, vol. 7, no. 2, pp. 85-94.

GEIJTENBEEK, TB., et al, 2000. DC-SIGN, a dendritic cell-specific HIV-1-binding protein that enhances trans-infection of T-cells. *Cell*, vol. 100, pp. 587

GENDELMAN, H.E., et al, 1997. The Neuropathogenesis of the AIDS Dementia Complex. *AIDS (London, England)*, vol. 11 Suppl A, pp. S35-45 ISSN 0269-9370; 0269-9370.

GONZALEZ, E., et al, 2001. Global Survey of Genetic Variation in CCR5, RANTES, and MIP-1α: Impact on the Epidemiology of the HIV-1 Pandemic. *Proceedings of the National Academy of Sciences*, vol. 98, no. 9, pp. 5199.

GONZALEZ, E., et al, 2002. HIV-1 Infection and AIDS Dementia are Influenced by a Mutant MCP-1 Allele Linked to Increased Monocyte Infiltration of Tissues and MCP-1 Levels. *Proceedings of the National Academy of Sciences*, vol. 99, no. 21, pp. 13795.

GOURRAUD, P., et al, 2011. APOBEC3H Haplotypes and HIV-1 Pro-Viral< i> vif</i> DNA Sequence Diversity in Early Untreated Human Immunodeficiency virus–1 Infection. *Human Immunology*, vol. 72, no. 3, pp. 207-212.

GRANT, I., et al, 1987. Evidence for Early Central Nervous System Involvement in the Acquired Immunodeficiency Syndrome (AIDS) and Other Human Immunodeficiency Virus (HIV) Infections. *Annals of Internal Medicine*, vol. 107, no. 6, pp. 828-836.

GUMBO, F.Z., et al, 2010. Rising Mother-to-Child HIV Transmission in a Resource-Limited Breastfeeding Population. *Tropical Doctor*, vol. 40, no. 2, pp. 70-73.

GUPTA, A., et al, 2011. Maternal Tuberculosis: A Risk Factor for Mother-to-Child Transmission of Human Immunodeficiency Virus. *Journal of Infectious Diseases*, vol. 203, no. 3, pp. 358.

GUXENS, M., et al, 2009. Cognitive Function and Overweight in Preschool Children. *American Journal of Epidemiology*, vol. 170, no. 4, pp. 438-446.

GUYON, A. and NAHON, J.L., 2007. Multiple Actions of the Chemokine Stromal Cell-Derived Factor-1α on Neuronal Activity. *Journal of Molecular Endocrinology*, vol. 38, no. 3, pp. 365-376.

HÅKANSSON, K., LIM, N.K., HOPPE, H.J. and REID, K., 1999. Crystal Structure of the Trimeric [Alpha]-Helical Coiled-Coil and the Three Lectin Domains of Human Lung Surfactant Protein D. *Structure*, vol. 7, no. 3, pp. 255-264.

HAO, X., KIM, T.S. and BRACIALE, T.J., 2008. Differential Response of Respiratory Dendritic Cell Subsets to Influenza Virus Infection. *Journal of Virology*, vol. 82, no. 10, pp. 4908-4919.

HARARI, A., OOMS, M., MULDER, L.C.F. and SIMON, V., 2009. Polymorphisms and Splice Variants Influence the Antiretroviral Activity of Human APOBEC3H. *Journal of Virology*, vol. 83, no. 1, pp. 295-303.

HAZLETON, J.E., BERMAN, J.W. and EUGENIN, E.A., 2010. Novel Mechanisms of Central Nervous System Damage in HIV Infection. *HIV/AIDS (Auckland, NZ)*, vol. 2, pp. 39.

HEATON, R.K., et al, 2011. HIV-Associated Neurocognitive Disorders before and during the Era of Combination Antiretroviral Therapy: Differences in Rates, Nature, and Predictors. *Journal of Neurovirology*, vol. 17, no. 1, pp. 3-16.

HEATON, R.K., et al, 2004. The Impact of HIV-Associated Neuropsychological Impairment on Everyday Functioning. *Psychology Faculty Publications*, pp. 15.

HEATON, R., et al, 2010. HIV-Associated Neurocognitive Disorders Persist in the Era of Potent Antiretroviral Therapy. *Neurology*, vol. 75, no. 23, pp. 2087-2096.

HEINEMEYER, T., et al, 1998. Databases on Transcriptional Regulation: TRANSFAC, TRRD and COMPEL. *Nucleic Acids Research*, vol. 26, no. 1, pp. 362-367.

HELLING, R.B., GOODMAN, H.M. and BOYER, H.W., 1974. Analysis of Endonuclease R· EcoRI Fragments of DNA from Lambdoid Bacteriophages and Other Viruses by Agarose-Gel Electrophoresis. *Journal of Virology*, vol. 14, no. 5, pp. 1235-1244.

HENDEL, H., et al, 2001. Validation of Genetic Case-Control Studies in AIDS and Application to the CX3CR1 Polymorphism. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, vol. 26, no. 5, pp. 507.

HENRIET, S., et al, 2009. Tumultuous Relationship between the Human Immunodeficiency Virus Type 1 Viral Infectivity Factor (Vif) and the Human APOBEC-3G and APOBEC-3F Restriction Factors. *Microbiology and Molecular Biology Reviews*, vol. 73, no. 2, pp. 211-232.

HINKIN, C.H., et al, 2004. Medication Adherence in HIV-Infected Adults: Effect of Patient Age, Cognitive Status, and Substance Abuse. *AIDS (London, England)*, vol. 18, no. Suppl 1, pp. S19.

HOLMSKOV, U., MALHOTRA, R., SIM, R.B. and JENSENIUS, J.C., 1994. Collectins: Collagenous C-Type Lectins of the Innate Immune Defense System. *Immunology Today*, vol. 15, no. 2, pp. 67-74.

HU, Y., WU, D., TAO, R. and SHANG, S., 2010. Association between Mannose-Binding Lectin Gene Polymorphism and Pediatric Cytomegalovirus Infection. *Viral Immunology*, vol. 23, no. 4, pp. 443-447.

HUET, T., et al, 1990. Genetic Organization of a Chimpanzee Lentivirus Related to HIV-1.

HUISING, M.O., et al, 2003. Molecular Evolution of CXC Chemokines: Extant CXC Chemokines Originate from the CNS. *Trends in Immunology*, vol. 24, no. 6, pp. 306-312.

HUTCHESON, H., et al, 2008. Detecting AIDS Restriction Genes: From Candidate Genes to Genome-Wide Association Discovery. *Vaccine*, vol. 26, no. 24, pp. 2951-2965.

HÜTTER, G., et al, 2009. Long-Term Control of HIV by CCR5 Delta32/Delta32 Stem-Cell Transplantation. *New England Journal of Medicine*, vol. 360, no. 7, pp. 692-698.

IMAI, T., et al, 1997. Identification and Molecular Characterization of Fractalkine Receptor CX< Sub> 3</sub> CR1, which Mediates both Leukocyte Migration and Adhesion. *Cell*, vol. 91, no. 4, pp. 521-530.

IRINA P, T., et al, 2011. CCL5/RANTES Gene Polymorphisms in Slavonic Patients with Myocardial Infarction. *Mediators of Inflammation*, vol. 2011.

ISHIDA, Y., GAO, J.L. and MURPHY, P.M., 2008. Chemokine Receptor CX3CR1 Mediates Skin Wound Healing by Promoting Macrophage and Fibroblast Accumulation and Function. *The Journal of Immunology*, vol. 180, no. 1, pp. 569.

IWASAKI, A. and MEDZHITOV, R., 2010. Regulation of Adaptive Immunity by the Innate Immune System. *Science's STKE*, vol. 327, no. 5963, pp. 291.

JARMUZ, A., et al, 2002. An Anthropoid-Specific Locus of Orphan C to U RNA-Editing Enzymes on Chromosome 22. *Genomics*, vol. 79, no. 3, pp. 285-296.

JIN, X., WU, H. and SMITH, H., 2007. APOBEC3G Levels Predict Rates of Progression to AIDS. *Retrovirology*, vol. 4, no. 1, pp. 20.

KANDAWASVIKA, G.Q., et al, 2011. Neurodevelopmental Impairment among Infants Born to Mothers Infected with Human Immunodeficiency Virus and Uninfected Mothers from Three peri-urban Primary Care Clinics in Harare, Zimbabwe. *Developmental Medicine & Child Neurology*, vol. 53, no. 11, pp. 1046-1052.

KARAALI, Z.E., et al, 2010. Effect of Genetic Variants of Chemokine Receptors on the Development of Myocardial Infarction in Turkish Population. *Molecular Biology Reports*, vol. 37, no. 7, pp. 3615-3619.

KARTVELISHVILI, A., LESNER, A., SZPONAR, M. and SIMM, M., 2004. Microarray Analysis of Differentially Expressed Genes in Cells Resistant to HIV-1. *Immunology Letters*, vol. 93, no. 1, pp. 79-86.

KATZ, D., et al, 2010. CCR5, RANTES and SDF-1 Polymorphisms and mother-to-child HIV-1 Transmission. *International Journal of Immunogenetics*, vol. 37, no. 4, pp. 301-305.

KHABOUR, O.F., ABU-HAWELEH, L.J. and ALZOUBI, K.H., 2012. Distribution of CCR-5 Δ 32, CCR2-64I, and SDF-1-3′ A Alleles among Jordanians. *AIDS Research and Human Retroviruses*.

KISHORE, U., EGGLETON, P. and REID, K., 1997. Modular Organization of Carbohydrate Recognition Domains in Animal Lectins. *Matrix Biology*, vol. 15, no. 8-9, pp. 583-592.

KLIKS, S.C., WARA, D.W., LANDERS, D.V. and LEVY, J.A., 1994. Features of HIV-1 that could Influence Maternal-Child Transmission. *JAMA: The Journal of the American Medical Association*, vol. 272, no. 6, pp. 467-474.

KOIZUMI, Y., et al, 2007. RANTES-28G Delays and DC-SIGN-139C Enhances AIDS Progression in HIV Type 1-Infected Japanese Hemophiliacs. *AIDS Research and Human Retroviruses*, vol. 23, no. 5, pp. 713-719.

KOSTRIKIS, L.G., et al, 1998. A Chemokine Receptor CCR2 Allele Delays HIV-1 Disease Progression and is Associated with a CCR5 Promoter Mutation. *Nature Medicine*, vol. 4, no. 3, pp. 350-353.

KOUTSOUNAKI, E., et al, 2008. Mannose-Binding Lectin MBL2 Gene Polymorphisms and Outcome of Hepatitis C Virus-Infected Patients. *Journal of Clinical Immunology*, vol. 28, no. 5, pp. 495-500.

KREBS, F.C., ROSS, H., MCALLISTER, J. and WIGDAHL, B., 2000. HIV-I-Associated Central Nervous System Dysfunction. *Advances in Pharmacology*, vol. 49, pp. 315-385.

KRONER, B.L., et al, 1994. HIV-1 Infection Incidence among Persons with Hemophilia in the United States and Western Europe, 1978-1990. Multicenter Hemophilia Cohort Study. *Journal of Acquired Immune Deficiency Syndromes*, vol. 7, no. 3, pp. 279.

KUANG, Y., WU, Y., JIANG, H. and WU, D., 1996. Selective G Protein Coupling by CC Chemokine Receptors. *Journal of Biological Chemistry*, vol. 271, no. 8, pp. 3975.

KUHLMAN, M., JOINER, K. and EZEKOWITZ, R., 1989. The Human Mannose-Binding Protein Functions as an Opsonin. *The Journal of Experimental Medicine*, vol. 169, no. 5, pp. 1733-1745.

KUHN, L., et al, 2006. Synergy between Mannose-Binding Lectin Gene Polymorphisms and Supplementation with Vitamin A Influences Susceptibility to HIV Infection in Infants Born to HIV-Positive Mothers. *The American Journal of Clinical Nutrition*, vol. 84, no. 3, pp. 610-615.

KUHN, L., et al, 1999. Distinct Risk Factors for Intrauterine and Intrapartum Human Immunodeficiency Virus Transmission and Consequences for Disease Progression in Infected Children. *Journal of Infectious Diseases*, vol. 179, no. 1, pp. 52-58..

KUREWA, EN., et al, 2011. Realities and Challenges of a Five Year Follow Up of Mother and Child Pairs on a PMTCT Program in Zimbabwe. *The Open AIDS Journal*, vol. 5, pp. 51-58.

KWA, D., BOESER-NUNNINK, B. and SCHUITEMAKER, H., 2003. Lack of Evidence for an Association between a Polymorphism in CX3CR1 and the Clinical Course of HIV Infection Or Virus Phenotype Evolution. *AIDS*, vol. 17, no. 5, pp. 759.

LANGFORD, S.E., ANANWORANICH, J. and COOPER, D.A., 2007. Predictors of Disease Progression in HIV Infection: A Review. *AIDS Research and Therapy*, vol. 4, no. 1, pp. 11.

LAPLANA, M. and FIBLA, J., 2012. Distribution of Functional Polymorphic Variants of Inflammation-Related Genes< i> RANTES</i> and< i> CCR5</i> in Long-Lived Individuals. *Cytokine*.

LAZENNEC, G. and RICHMOND, A., 2010. Chemokines and Chemokine Receptors: New Insights into Cancer-Related Inflammation. *Trends in Molecular Medicine*, vol. 16, no. 3, pp. 133-144.

LECOSSIER, D., BOUCHONNET, F., CLAVEL, F. and HANCE, A.J., 2003. Hypermutation of HIV-1 DNA in the Absence of the Vif Protein. *Science*, vol. 300, no. 5622, pp. 1112-1112.

LEVINE, A.J., et al, 2012. Genome-wide Association Study of Neurocognitive Impairment and Dementia in HIV-infected Adults. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*.

LEVINE, A.J., SINGER, E.J. and SHAPSHAK, P., 2009. The Role of Host Genetics in the Susceptibility for HIV-Associated Neurocognitive Disorders. *AIDS and Behavior*, vol. 13, no. 1, pp. 118-132.

LI, C., et al, 2005. Distribution of Human Chemokine (C-X3-C) Receptor 1 (CX3CR1) Gene Polymorphisms and Haplotypes of the CC Chemokine Receptor 5 (CCR5) Promoter in Chinese People, and the Effects of CCR5 Haplotypes on CCR5 Expression. *International Journal of Immunogenetics*, vol. 32, no. 2, pp. 99-106.

LI, M.M.H. and EMERMAN, M., 2011. Polymorphism in Human APOBEC3H Affects a Phenotype Dominant for Subcellular Localization and Antiviral Activity. *Journal of Virology*, vol. 85, no. 16, pp. 8197-8207.

LI, Z., et al, 2009. A Partition-Ligation-Combination-Subdivision EM Algorithm for Haplotype Inference with Multiallelic Markers: Update of the SHEsis (Http://analysis. Bio-x. Cn). *Cell Research*, vol. 19, no. 4, pp. 519-523.

LIPSCOMBE, R., et al, 1996. Mutations in the Human Mannose-Binding Protein Gene: Frequencies in several Population Groups. *European Journal of Human Genetics: EJHG*, vol. 4, no. 1, pp. 13.

LIPSCOMBE, R., et al, 1992. High Frequencies in African and Non-African Populations of Independent Mutations in the Mannose Binding Protein Gene. *Human Molecular Genetics*, vol. 1, no. 9, pp. 709-715.

LIU, H., et al, 1999. Polymorphism in RANTES Chemokine Promoter Affects HIV-1 Disease Progression. *Proceedings of the National Academy of Sciences*, vol. 96, no. 8, pp. 4581.

LIU, H., et al, 2004. Analysis of Genetic Polymorphisms in CCR5, CCR2, Stromal Cell-Derived Factor-1, RANTES, and Dendritic Cell-Specific Intercellular Adhesion Molecule-3-Grabbing Nonintegrin in Seronegative Individuals Repeatedly Exposed to HIV-1. *Journal of Infectious Diseases*, vol. 190, no. 6, pp. 1055-1058.

LIU, N.Q., et al, 2002. Human Immunodeficiency Virus Type 1 Enters Brain Microvascular Endothelia by Macropinocytosis Dependent on Lipid Rafts and the Mitogen-Activated Protein Kinase Signaling Pathway. *Journal of Virology*, vol. 76, no. 13, pp. 6689-6700.

LLORENTE, A., et al, 2006. Effects of Polymorphisms of Chemokine Receptors on Neurodevelopment and the Onset of Encephalopathy in Children with Perinatal HIV-1 Infection. *Appl Neuropsycho,*. vol. 13, no. 3, pp. 180-9.

LOCHET, P., et al, 2003. Long-term Assessment of Neuropsychiatric Adverse Reactions Associated with Efavirenz. *HIV Medicine*, vol. 4, no. 1, pp. 62-66.

LOETSCHER, P., et al, 1994. Monocyte Chemotactic Proteins MCP-1, MCP-2, and MCP-3 are Major Attractants for Human CD4 and CD8 T Lymphocytes. *The FASEB Journal*, vol. 8, no. 13, pp. 1055-1060.

LOHMAN, B.L., et al, 2005. Longitudinal Assessment of Human Immunodeficiency Virus Type 1 (HIV-1)-Specific Gamma Interferon Responses during the First Year of Life in HIV-1-Infected Infants. *Journal of Virology*, vol. 79, no. 13, pp. 8121-8130.

LUCOTTE, G. and MERCIER, G., 1998. Distribution of the CCR5 Gene 32-Bp Deletion in Europe. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, vol. 19, no. 2, pp. 174.

LYNN, D., et al, 2010. Curating the Innate Immunity Interactome. *BMC Systems Biology*, vol. 4, no. 1, pp. 117.

MA, L., et al, 2005. Distribution of CCR2-64I and SDF1-3'A Alleles and HIV Status in 7 Ethnic Populations of Cameroon. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, vol. 40, no. 1, pp. 89.

MABUKA, J.M., et al, 2009. CCR2-64I Polymorphism is Associated with Lower Maternal HIV-1 Viral Load and Reduced Vertical HIV-1 Transmission. *Journal of Acquired Immune Deficiency Syndromes (1999)*, vol. 51, no. 2, pp. 235.

MACGINNITIE, A.J., ANANT, S. and DAVIDSON, N.O., 1995. Mutagenesis of Apobec-1, the Catalytic Subunit of the Mammalian Apolipoprotein B mRNA Editing Enzyme, Reveals Distinct Domains that Mediate Cytosine Nucleoside Deaminase, RNA Binding, and RNA Editing Activity. *Journal of Biological Chemistry*, vol. 270, no. 24, pp. 14768-14775.

MACK, M., et al, 1998. Aminooxypentane-RANTES Induces CCR5 Internalization but Inhibits Recycling: A Novel Inhibitory Mechanism of HIV Infectivity. *The Journal of Experimental Medicine*, vol. 187, no. 8, pp. 1215.

MADSEN, H.O., et al, 1994. A New Frequent Allele is the Missing Link in the Structural Polymorphism of the Human Mannan-Binding Protein. *Immunogenetics*, vol. 40, no. 1, pp. 37-44.

MADSEN, H.O., et al, 1995. Interplay between Promoter and Structural Gene Variants Control Basal Serum Level of Mannan-Binding Protein. *The Journal of Immunology*, vol. 155, no. 6, pp. 3013.

MADSEN, H.O., et al, 1998. Different Molecular Events Result in Low Protein Levels of Mannan-Binding Lectin in Populations from Southeast Africa and South America. *The Journal of Immunology*, vol. 161, no. 6, pp. 3169-3175.

MAHALINGAM, S., CLARK, K., MATTHAEI, K.I. and FOSTER, P.S., 2001. Antiviral Potential of Chemokines. *Bioessays*, vol. 23, no. 5, pp. 428-435.

MALHOTRA, R., et al, 2011. Association of Chemokine Receptor Gene (CCR2-CCR5) Haplotypes with Acquisition and Control of HIV-1 Infection in Zambians. *Retrovirology*, vol. 8, no. 1, pp. 22.

MALIM, M.H., 2009. APOBEC Proteins and Intrinsic Resistance to HIV-1 Infection. *Philosophical Transactions of the Royal Society B: Biological Sciences*, vol. 364, no. 1517, pp. 675-687.

MANGANO, A., et al, 2000. Protective Effect of CCR2-64I and Not of CCR5-[DELTA] 32 and SDF1-3 [Spacing Acute] A in Pediatric HIV-1 Infection. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, vol. 23, no. 1, pp. 52.

MANGANO, A., et al, 2008. Detrimental Effects of Mannose-Binding Lectin (MBL2) Promoter Genotype XA/XA on HIV-1 Vertical Transmission and AIDS Progression. *Journal of Infectious Diseases*, vol. 198, no. 5, pp. 694-700.

MARIANI, R., et al, 2003. Species-Specific Exclusion of APOBEC3G from HIV-1 Virions by Vif. *Cell*, vol. 114, no. 1, pp. 21-31.

MARTINSON, J.J., et al, 2000. Global Distribution of the CCR2-64I/CCR5-59653T HIV-1 Disease-Protective Haplotype. *Aids*, vol. 14, no. 5, pp. 483.

MATSUSHITA, M., ENDO, Y. and FUJITA, T., 2000. Cutting Edge: Complement-Activating Complex of Ficolin and Mannose-Binding Lectin-Associated Serine Protease. *The Journal of Immunology*, vol. 164, no. 5, pp. 2281-2284.

MAYEUX, R., et al, 1993. Mortality Risks in Gay Men with Human Immunodeficiency Virus Infection and Cognitive Impairment. *Neurology*, vol. 43, no. 1 Part 1, pp. 176-176.

MCCLURE, M.O., MARSH, M. and WEISS, R.A., 1988. Human Immunodeficiency Virus Infection of CD4-Bearing Cells Occurs by a pH-Independent Mechanism. *The EMBO Journal*, vol. 7, no. 2, pp. 513.

MCDERMOTT, D.H., et al, 2000. Chemokine RANTES Promoter Polymorphism Affects Risk of both HIV Infection and Disease Progression in the Multicenter AIDS Cohort Study. *Aids*, vol. 14, no. 17, pp. 2671.

MCDERMOTT, D.H., et al, 2003. Chemokine Receptor Mutant CX3CR1-M280 has Impaired Adhesive Function and Correlates with Protection from Cardiovascular Disease in Humans. *Journal of Clinical Investigation*, vol. 111, no. 8, pp. 1241-1250.

MELLADO, M., et al, 1999. Chemokine Control of HIV-1 Infection. *Nature*, vol. 400, no. 6746, pp. 723-723.

MENDOLA, P., SELEVAN, S.G., GUTTER, S. and RICE, D., 2002. Environmental Factors Associated with a Spectrum of Neurodevelopmental Deficits. *Mental Retardation and Developmental Disabilities Research Reviews*, vol. 8, no. 3, pp. 188-197.

MEUCCI, O., FATATIS, A., SIMEN, A.A. and MILLER, R.J., 2000. Expression of CX3CR1 Chemokine Receptors on Neurons and their Role in Neuronal Survival. *Proceedings of the National Academy of Sciences*, vol. 97, no. 14, pp. 8075.

MICHAEL, N.L., et al, 1997. The Role of CCR5 and CCR2 Polymorphisms in HIV-1 Transmission and Disease Progression. *Nature Medicine*, vol. 3, no. 10, pp. 1160-1162.

MOFENSON, L.M. and MUNDERI, P., 2002. Safety of Antiretroviral Prophylaxis of Perinatal Transmission for HIV-Infected Pregnant Women and their Infants. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, vol. 30, no. 2, pp. 200.

MÖHLE, R., et al, 1998. The Chemokine Receptor CXCR-4 is Expressed on CD34 Hematopoietic Progenitors and Leukemic Cells and Mediates Transendothelial Migration Induced by Stromal Cell-Derived Factor-1. *Blood*, vol. 91, no. 12, pp. 4523-4530.

MULHERIN, S.A., et al, 2003. Effects of CCR5-[DELTA] 32 and CCR2-64I Alleles on HIV-1 Disease Progression: The Protection Varies with Duration of Infection. *Aids*, vol. 17, no. 3, pp. 377.

MUMMIDI, S., et al, 1998. Genealogy of the CCR5 Locus and Chemokine System Gene Variants Associated with Altered Rates of HIV-1 Disease Progression. *Nature Medicine*, vol. 4, no. 7, pp. 786-793.

MURO, M., et al, 2008. CCL5/RANTES Chemokine Gene Promoter Polymorphisms are Not Associated with Atopic and Nonatopic Asthma in a Spanish Population. *International Journal of Immunogenetics*, vol. 35, no. 1, pp. 19-23.

MURPHY, P.M., et al, 2000. International Union of Pharmacology. XXII. Nomenclature for Chemokine Receptors. *Pharmacological Reviews*, vol. 52, no. 1, pp. 145-176.

MYERS, S.J., WONG, L.M. and CHARO, I.F., 1995. Signal Transduction and Ligand Specificity of the Human Monocyte Chemoattractant Protein-1 Receptor in Transfected Embryonic Kidney Cells. *Journal of Biological Chemistry*, vol. 270, no. 11, pp. 5786-5792.

NAKAYAMA, E.E., et al, 2004. A CCR2-V64I Polymorphism Affects Stability of CCR2A Isoform. *Aids*, vol. 18, no. 5, pp. 729.

NASSAR, B., et al, 2008. Role of the Fractalkine Receptor CX3CR1 Polymorphisms V249I and T280M as Risk Factors for Early-Onset Coronary Artery Disease in Patients with no Classic Risk Factors. *Scandinavian Journal of Clinical & Laboratory Investigation*, vol. 68, no. 4, pp. 286-291.

NICKEL, R.G., et al, 2000. Atopic Dermatitis is Associated with a Functional Mutation in the Promoter of the CC Chemokine RANTES. *The Journal of Immunology*, vol. 164, no. 3, pp. 1612.

NISHIMURA, M., et al, 2002. Dual Functions of fractalkine/CX3C Ligand 1 in Trafficking of Perforin /granzyme B Cytotoxic Effector Lymphocytes that are Defined by CX3CR1 Expression. *The Journal of Immunology*, vol. 168, no. 12, pp. 6173-6180.

NISHIYORI, A., et al, 1998. Localization of Fractalkine and CX3CR1 mRNAs in Rat Brain: Does Fractalkine Play a Role in Signaling from Neuron to Microglia?. *FEBS Letters*, vol. 429, no. 2, pp. 167-172.

NYQUIST, P.A., et al, 2009. Single Nucleotide Polymorphisms in Monocyte Chemoattractant Protein-1 and its Receptor Act Synergistically to Increase the Risk of Carotid Atherosclerosis. *Cerebrovascular Diseases*, vol. 28, no. 2, pp. 124-130.

OBERLIN, E., et al, 1996. The CXC Chemokine, Stromal Cell Derived Factor 1 (SDF-1), is the Ligand for LESTR/fusin and Prevents Infection by Lymphocyte-Tropic HIV-1 Syncytium-Inducing Strains. *Nature*, vol. 382, pp. 833-835.

O'BRIEN, S.J. and NELSON, G.W., 2004. Human Genes that Limit AIDS. *Nature Genetics*, vol. 36, no. 6, pp. 565-574.

OHAINLE, M., KERNS, J.A., MALIK, H.S. and EMERMAN, M., 2006. Adaptive Evolution and Antiviral Activity of the Conserved Mammalian Cytidine Deaminase APOBEC3H. *Journal of Virology*, vol. 80, no. 8, pp. 3853-3862.

ORLOFF, G.M., et al, 1991. Penetration of CD4 T Cells by HIV-1. the CD4 Receptor does Not Internalize with HIV, and CD4-Related Signal Transduction Events are Not Required for Entry. *The Journal of Immunology*, vol. 146, no. 8, pp. 2578-2587.

OU, X.T., et al, 2011. Genotypes Coding for Mannose-Binding Lectin Deficiency Correlated with Cryptococcal Meningitis in HIV-Uninfected Chinese Patients. *Journal of Infectious Diseases*, vol. 203, no. 11, pp. 1686.

OWOR, M. and FOWLER, M.G., 2004. Treatment and Prevention of Opportunistic Infections in Children Living with HIV. *From the Ground Up*.

PACE, M.J., AGOSTO, L., GRAF, E.H. and O'DOHERTY, U., 2011. HIV Reservoirs and Latency Models. *Virology*.

PALFRAMAN, R.T., et al, 2001. Inflammatory Chemokine Transport and Presentation in HEV. *The Journal of Experimental Medicine*, vol. 194, no. 9, pp. 1361.

PEMBERTON, LA., et al, 2008. The Relationship between ApoE, TNFA, IL1a, IL1b and IL12b Genes and HIV-1-associated Dementia. *HIV Medicine* ,vol. 9, pp. 677–680

PANTALEO, G. and FAUCI, A., 1996. Immunopathogenesis of HIV Infection 1. *Annual Reviews in Microbiology*, vol. 50, no. 1, pp. 825-854.

PAXIMADIS, M., et al, 2011. KIR-HLA and Maternal-Infant HIV-1 Transmission in Sub-Saharan Africa. *PLoS One*, vol. 6, no. 2, pp. e16541.

PERRY, S.E., et al, 2010. The Newborn at Risk: Acquired and Congenital Problems. *Maternal Child Nursing Care*, 4th ed, St Louis, MO: Mosby-Elsevier, pp. 886-931.

PETERSEN, D.C., et al, 2005. Risk for HIV-1 Infection Associated with a Common CXCL12 (SDF1) Polymorphism and CXCR4 Variation in an African Population. *Journal of Acquired Immune Deficiency Syndromes (1999)*, vol. 40, no. 5, pp. 521.

PLINER, V., et al, 1998. Incubation Period of HIV-1 in Perinatally Infected Children. *Aids*, vol. 12, no. 7, pp. 759.

PODACK, E.R., BIESECKER, G. and MÜLLER-EBERHARD, H.J., 1979. Membrane Attack Complex of Complement: Generation of High-Affinity Phospholipid Binding Sites by Fusion of Five Hydrophilic Plasma Proteins. *Proceedings of the National Academy of Sciences*, vol. 76, no. 2, pp. 897.

POLILLI, E., et al, 2010. Rapidly Progressive and Fatal EBV-Related Encephalitis in a Patient with Advanced HIV-1 Infection at Presentation: A Case Report and Review of the Literature. *The New Microbiologica*, vol. 33, no. 3, pp. 275.

POLLARD, V.W. and MALIM, M.H., 1998. The HIV-1 Rev Protein. *Annual Reviews in Microbiology*, vol. 52, no. 1, pp. 491-532.

QIAN, Y., et al, 2008. Distribution of CCR5-A32, CCR2-641, SDF1-3'A, CX3CR1-249I, and CX3CR1-280M in Chinese Populations. *AIDS Research and Human Retroviruses*, vol. 24, no. 11, pp. 1391-1397.

RAMPOLDI, L., DANEK, A. and MONACO, A.P., 2002. Clinical Features and Molecular Bases of Neuroacanthocytosis. *Journal of Molecular Medicine*, vol. 80, no. 8, pp. 475-491.

RATHORE, A., et al, 2008. Association of RANTES− 403 G/A,− 28 C/G and In1. 1 T/C Polymorphism with HIV-1 Transmission and Progression among North Indians. *Journal of Medical Virology*, vol. 80, no. 7, pp. 1133-1141.

RATHORE, A., CHATTERJEE, A., YAMAMOTO, N. and DHOLE, T.N., 2008b. Absence of H186R Polymorphism in Exon 4 of the APOBEC3G Gene among North Indian Individuals. *Genetic Testing*, vol. 12, no. 3, pp. 453-456.

REDDY, K., et al, 2010. APOBEC3G Expression is Dysregulated in Primary HIV-1 Infection and Polymorphic Variants Influence CD4 T-Cell Counts and Plasma Viral Load. *AIDS*, vol. 24, no. 2, pp. 195.

REICHE, E., et al, 2006. Stromal cell-derived Factor 1 (SDF1) Genetic Polymorphism in a Sample of Healthy Individuals, Seronegative Individuals Exposed to Human Immunodeficiency Virus Type 1 (HIV-1) and Patients Infected with HIV-1 from the Brazilian Population. *International Journal of Immunogenetics*, vol. 33, no. 2, pp. 127-133.

RENJIFO, B., et al, 2004. Preferential in-Utero Transmission of HIV-1 Subtype C as Compared to HIV-1 Subtype A Or D. *Aids*, vol. 18, no. 12, pp. 1629.

ROSENBERG, Z.F. and FAUCI, A.S., 1991. Immunopathogenesis of HIV Infection. *The FASEB Journal*, vol. 5, no. 10, pp. 2382-2390.

ROT, A., et al, 1992. RANTES and Macrophage Inflammatory Protein 1 Alpha Induce the Migration and Activation of Normal Human Eosinophil Granulocytes. *The Journal of Experimental Medicine*, vol. 176, no. 6, pp. 1489-1495.

ROUSSEAU, C.M., et al, 2003. Longitudinal Analysis of Human Immunodeficiency Virus Type 1 RNA in Breast Milk and of its Relationship to Infant Infection and Maternal Disease. *Journal of Infectious Diseases*, vol. 187, no. 5, pp. 741-747.

SAIFUDDIN, M., et al, 2000. Interaction of Mannose-Binding Lectin with Primary Isolates of Human Immunodeficiency Virus Type 1. *Journal of General Virology*, vol. 81, no. 4, pp. 949-955.

SCHALL, T.J., BACON, K., TOY, K.J. and GOEDDEL, D.V., 1990. Selective Attraction of Monocytes and T Lymphocytes of the Memory Phenotype by Cytokine RANTES. *Nature*, vol. 347, no. 6294, pp. 669-671.

SCHINKEL, J., et al, 1999. No Evidence for an Effect of the CCR5 Δ32/ and CCR2b 64I/ Mutations on Human Immunodeficiency Virus (HIV)-1 Disease Progression among HIV-1-Infected Injecting Drug Users. *Journal of Infectious Diseases*, vol. 179, no. 4, pp. 825-831.

SHEEHY, A.M., GADDIS, N.C., CHOI, J.D. and MALIM, M.H., 2002. Isolation of a Human Gene that Inhibits HIV-1 Infection and is Suppressed by the Viral Vif Protein. *Nature*, vol. 418, no. 6898, pp. 646-650.

SHERMAN, G.G., et al, 2005. Dried Blood Spots Improve Access to HIV Diagnosis and Care for Infants in Low-Resource Settings. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, vol. 38, no. 5, pp. 615.

SHIROZU, M., et al, 1995. Structure and Chromosomal Localization of the Human Stromal Cell-Derived Factor 1 (SDF1) Gene. *Genomics*, vol. 28, no. 3, pp. 495-500.

SHUN, M.C., et al, 2007. LEDGF/p75 Functions Downstream from Preintegration Complex Formation to Effect Gene-Specific HIV-1 Integration. *Genes & Development*, vol. 21, no. 14, pp. 1767-1778.

SILICIANO, R.F. and GREENE, W.C., 2011. HIV Latency. *Cold Spring Harbor Perspectives in Medicine*, vol. 1, no. 1.

SINGH, K.K., et al, 2003. Genetic Influence of CCR5, CCR2 and SDF1 Variants on Human Immunodeficiency Virus 1 (HIV-1)–Related Disease Progression and Neurological Impairment, in Children with Symptomatic HIV-1 Infection. *Journal of Infectious Diseases*, vol. 188, no. 10, pp. 1461-1472.

SINGH, K.K., et al, 2008. Associations of Chemokine Receptor Polymorphisms with HIV-1 Mother-to-Child Transmission in Sub-Saharan Africa: Possible Modulation of Genetic Effects by Antiretrovirals. *Journal of Acquired Immune Deficiency Syndromes (1999)*, vol. 49, no. 3, pp. 259.

SINGH, K.K., HUGHES, M.D., CHEN, J. and SPECTOR, S.A., 2005. Genetic Polymorphisms in CX 3 CR1 Predict HIV-1 Disease Progression in Children Independently of CD4 Lymphocyte Count and HIV-1 Rna Load. *Journal of Infectious Diseases*, vol. 191, no. 11, pp. 1971.

SINGH, K.K., et al, 2008. An Age-Dependent Association of Mannose-Binding Lectin-2 Genetic Variants on HIV-1–related Disease in Children. *Journal of Allergy and Clinical Immunology*, vol. 122, no. 1, pp. 173-180. e2.

SINGH, K.K. and SPECTOR, S.A., 2009. Host Genetic Determinants of Human Immunodeficiency Virus Infection and Disease Progression in Children. *Pediatric Research*, vol. 65, pp. 55R-63R.

SMITH, M.W., et al, 1997. Contrasting Genetic Influence of CCR2 and CCR5 Variants on HIV-1 Infection and Disease Progression. *Science*, vol. 277, no. 5328, pp. 959.

SØBORG, C., et al, 2007. Influence of Candidate Susceptibility Genes on Tuberculosis in a High Endemic Region. *Molecular Immunology*, vol. 44, no. 9, pp. 2213-2220.

SOILLEUX, EJ., et al, 2001. Placental expression of DC-SIGN may mediate intrauterine vertical transmission of HIV. *Journal of Pathology*, vol. 195, pp.586–592.

SOZZANI, S., et al, 1997. Receptor Expression and Responsiveness of Human Dendritic Cells to a Defined Set of CC and CXC Chemokines. *The Journal of Immunology*, vol. 159, no. 4, pp. 1993.

SPEAR, GT., et al, 2003. Inhibition of DC-SIGN-mediated Trans-infection of T-cells by Mannose-binding Lectin. *Immunology*, vol. 110, p.80-85.

SPECTOR, S.A., et al, 2010. 195APOE ε4 and MBL-2 O/O Genotypes are Associated with Neurocognitive Impairment in HIV-Infected Plasma Donors from Anhui Province, China. *AIDS (London, England)*, vol. 24, no. 10, pp. 1471.

SPIRA, R., et al, 1999. Natural History of Human Immunodefiency Virus Type 1 Infection in Children: A Five-Year Prospective Study in Rwanda. *Pediatrics*, vol. 104, no. 5, pp. e56-e56.

STEPHENS, J.C., et al, 1998. Dating the Origin of the< i> CCR5</i>-Δ< i> 32</i> AIDS-Resistance Allele by the Coalescence of Haplotypes. *The American Journal of Human Genetics*, vol. 62, no. 6, pp. 1507-1515.

STOJKOVIĆ, L., et al, 2012. The Association of V249I and T280M Fractalkine Receptor Haplotypes with Disease Course of Multiple Sclerosis. *Journal of Neuroimmunology*.

STRAZZA, M., PIRRONE, V., WIGDAHL, B. and NONNEMACHER, M.R., 2011. Breaking Down the Barrier: The Effects of HIV-1 on the Blood-Brain Barrier. *Brain Research*.

STRUYF, F., et al, 2000. Prevalence of CCR5 and CCR2 HIV-Coreceptor Gene Polymorphisms in Belgium. *Human Heredity*, vol. 50, no. 5, pp. 304-307.

SU, B., et al, 1999. Distribution of Two HIV-1–Resistant Polymorphisms (SDF1-3′ A and CCR2-64I) in East Asian and World Populations and its Implication in AIDS Epidemiology. *The American Journal of Human Genetics*, vol. 65, no. 4, pp. 1047-1053.

SUMIYA, M., et al, 1991. Molecular Basis of Opsonic Defect in Immunodeficient Children. *The Lancet*, vol. 337, no. 8757, pp. 1569-1570.

SURESH, P., WANCHU, A., SACHDEVA, R.K. and BHATNAGAR, A., 2006a. Gene Polymorphisms in CCR5, CCR2, CX3CR1, SDF-1 and RANTES in Exposed but Uninfected Partners of HIV-1 Infected Individuals in North India. *Journal of Clinical Immunology*, vol. 26, no. 5, pp. 476-484.

SWART, M., et al, 2012. CYP1A2, CYP2A6, CYP2B6, CYP3A4 and CYP3A5 Polymorphisms in Two Bantu-Speaking Populations from Cameroon and South Africa: Implications for Global Pharmacogenetics. *Current Pharmacogenomics and Personalized Medicine (Formerly Current Pharmacog*, vol. 10, no. 1, pp. 43-53.

SYVÄNEN, A.C., et al, 1990. A Primer-Guided Nucleotide Incorporation Assay in the Genotyping of Apolipoprotein E. *Genomics*, vol. 8, no. 4, pp. 684-692.

TACHIBANA, K., et al, 1998. The Chemokine Receptor CXCR4 is Essential for Vascularization of the Gastrointestinal Tract. *Nature*, vol. 393, no. 6685, pp. 591-594.

TAYLOR, B.S., SOBIESZCZYK, M.E., MCCUTCHAN, F.E. and HAMMER, S.M., 2008. The Challenge of HIV-1 Subtype Diversity. *New England Journal of Medicine*, vol. 358, no. 15, pp. 1590-1602.

TEGLAS, J.P., et al, 1999. CCR2B-64I Chemokine Receptor Allele and Mother-to-Child HIV-1 Transmission Or Disease Progression in Children. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, vol. 22, no. 3, pp. 267.

THIEL, S., et al, 1997. A Second Serine Protease Associated with Mannan-Binding Lectin that Activates Complement. *Nature*, vol. 386, no. 6624, pp. 506-510.

THYE, T., et al, 2011. Variant G57E of Mannose Binding Lectin Associated with Protection Against Tuberculosis Caused by Mycobacterium Africanum but Not by M. Tuberculosis. *PloS One*, vol. 6, no. 6, pp. e20908.

TIEMESSEN, C.T. and KUHN, L., 2006. Immune Pathogenesis of Pediatric HIV-1 Infection. *Current HIV/AIDS Reports*, vol. 3, no. 1, pp. 13-19.

TISHKOFF, S.A., et al, 2009. The Genetic Structure and History of Africans and African Americans. *Science*, vol. 324, no. 5930, pp. 1035-1044.

TOMESCU, C., ABDULHAQQ, S. and MONTANER, L., 2011. Evidence for the Innate Immune Response as a Correlate of Protection in Human Immunodeficiency Virus (HIV)-1 Highly Exposed Seronegative Subjects (HESN). *Clinical & Experimental Immunology*.

TORRE, V.S., et al, 2000. Variable Sensitivity of CCR5-Tropic Human Immunodeficiency Virus Type 1 Isolates to Inhibition by RANTES Analogs. *Journal of Virology*, vol. 74, no. 10, pp. 4868-4876.

TURNER, M.W., 1996. Mannose-Binding Lectin: The Pluripotent Molecule of the Innate Immune System. *Immunology Today*, vol. 17, no. 11, pp. 532-540.

TURVEY, S.E. and BROIDE, D.H., 2010. Innate Immunity. *Journal of Allergy and Clinical Immunology*, vol. 125, no. 2, pp. S24-S32.

UNAIDS., 2005. Evidence for HIV Decline in Zimbabwe: A Comprehensive Review of the Epidemiological Data.

UNAIDS., 2010. UNAIDS Report on the Global AIDS Epidemic 2010.

UNFPA Zimbabwe., 2011. Promoting evidence based HIV Prevention Services.

UNGASS., 2010. Report on HIV and AIDS: A follow-up to the Declaration of Commitment on HIV and AIDS. Zimbabwe Country Report.

VALCKE, H.S., et al, 2006. APOBEC3G Genetic Variants and their Association with Risk of HIV Infection in Highly Exposed Caucasians. *Aids*, vol. 20, no. 15, pp. 1984.

VAN RIJ, R.P., et al, 1998. The Role of a Stromal Cell-Derived Factor-1 Chemokine Gene Variant in the Clinical Course of HIV-1 Infection. *Aids*, vol. 12, no. 9, pp. F85.

VANG PETERSEN, S., THIEL, S. and JENSENIUS, J.C., 2001. The Mannan-Binding Lectin Pathway of Complement Activation: Biology and Disease Association. *Molecular Immunology*, vol. 38, no. 2-3, pp. 133-149.

VANHÉE-BROSSOLLET, C., et al, 1995. A Natural Antisense RNA Derived from the HIV-1 Env Gene Encodes a Protein which is Recognized by Circulating Antibodies of HIV Individuals. *Virology*, vol. 206, no. 1, pp. 196-202.

VERDU, P., et al, 2006. Evolutionary Insights into the High Worldwide Prevalence of *MBL2* Deficiency Alleles. *Human Molecular Genetics*, vol. 15, no. 17, pp. 2650–2658.

VIDAL, F., et al, 2006. Polymorphism of RANTES Chemokine Gene Promoter is Not Associated with Long-Term Nonprogressive HIV-1 Infection of More than 16 Years. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, vol. 41, no. 1, pp. 17.

VIDAL, F., et al, 2005. Spanish HIV-1-Infected Long-Term Nonprogressors of More than 15 Years have an Increased Frequency of the CX3CR1 249I Variant Allele. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, vol. 40, no. 5, pp. 527.

VIEIRA, V.C., et al, 2011. The Effect of Combined Polymorphisms in Chemokines and Chemokine Receptors on the Clinical Course of HIV-1 Infection in a Brazilian Population. *Memórias do Instituto Oswaldo Cruz*, vol. 106, no. 4, pp. 408-414.

VIOLARI, A., et al, 2008. Early Antiretroviral Therapy and Mortality among HIV-Infected Infants. *New England Journal of Medicine*, vol. 359, no. 21, pp. 2233-2244.

VON SCHWEDLER, U., SONG, J., AIKEN, C. and TRONO, D., 1993. Vif is Crucial for Human Immunodeficiency Virus Type 1 Proviral DNA Synthesis in Infected Cells. *Journal of Virology*, vol. 67, no. 8, pp. 4945-4955.

VOS, C.M.P., et al, 2000. Matrix Metalloprotease-9 Release from Monocytes Increases as a Function of Differentiation: Implications for Neuroinflammation and Neurodegeneration. *Journal of Neuroimmunology*, vol. 109, no. 2, pp. 221-227.

WALLIS, R., KRARUP, A. and GIRIJA, U.V., 2008. The Structure and Function of Ficolins, MBLs and MASPs. *Molecular Aspects of Innate and Adaptive Immunity.RSC Publishing, Cambridge UK*, pp. 147-166.

WASIK, T.J., et al, 2005. Effects of CCR5-delta32, CCR2-64I and SDF-1-3'A Polymorphic Alleles on Human Immunodeficiency Virus 1 (HIV-1) Infection in the Polish Population. *Wiadomosci Lekarskie (Warsaw, Poland : 1960)*, vol. 58, no. 9-10, pp. 500-507 ISSN 0043-5147; 0043-5147.

WATANABE, M.A.E., et al, 2003. SDF-1 Gene Polymorphisms and Syncytia Induction in Brazilian HIV-1 Infected Individuals. *Microbial Pathogenesis*, vol. 35, no. 1, pp. 31-34.

WEIS, W.I., DRICKAMER, K. and HENDRICKSON, W.A., 1992. Structure of a C-Type Mannose-Binding Protein Complexed with an Oligosaccharide.

WHO., 2011. Global HIV/AIDS Response: Epidemic Update and Health Sector Progress Towards Universal Access. Progress Report 2011.

WICHUKCHINDA, N., et al, 2008. Effects of CCR2 and CCR5 Polymorphisms on HIV-1 Infection in Thai Females. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, vol. 47, no. 3, pp. 293.

WICHUKCHINDA, N., et al, 2006. Protective Effects of IL4-589T and RANTES-28G on HIV-1 Disease Progression in Infected Thai Females. *Aids*, vol. 20, no. 2, pp. 189.

WILLIAMS, A. and BLAKEMORE, W., 1990. Monocyte-mediated Entry of Pathogens into the Central Nervous System. *Neuropathology and Applied Neurobiology*, vol. 16, no. 5, pp. 377-392.

WILLIAMSON, C., et al, 2000. Allelic Frequencies of Host Genetic Variants Influencing Susceptibility to HIV-1 Infection and Disease in South African Populations. *Aids*, vol. 14, no. 4, pp. 449.

WINKLER, C., et al, 1998. Genetic Restriction of AIDS Pathogenesis by an SDF-1 Chemokine Gene Variant. *Science*, vol. 279, no. 5349, pp. 389-393.

WONG, H.L., CHATTOPADHYAY, N., WU, X.Y. and BENDAYAN, R., 2010. Nanotechnology Applications for Improved Delivery of Antiretroviral Drugs to the Brain. *Advanced Drug Delivery Reviews*, vol. 62, no. 4-5, pp. 503-517.

YU, Y., et al, 2004. Selective Assembly of HIV-1 Vif-Cul5-ElonginB-ElonginC E3 Ubiquitin Ligase Complex through a Novel SOCS Box and Upstream Cysteines. *Genes & Development*, vol. 18, no. 23, pp. 2867.

YUAN, C.C., et al, 2000. 5′ Nuclease Assays for the Loci CCR5- /Δ32, CCR2-V64I, and SDF1-G801A Related to Pathogenesis of AIDS. *Clinical Chemistry*, vol. 46, no. 1, pp. 24-30.

ZHANG, H., et al, 2003. The Cytidine Deaminase CEM15 Induces Hypermutation in Newly Synthesized HIV-1 DNA. *Nature*, vol. 424, no. 6944, pp. 94.

ZHAO, X.Y., et al, 2004. Effects of Single Nucleotide Polymorphisms in the RANTES Promoter Region in Healthy and HIV-infected Indigenous Chinese*. *European Journal of Immunogenetics*, vol. 31, no. 4, pp. 179-183.

ZHOU, L., et al, 2009. Evidence for Predilection of Macrophage Infiltration Patterns in the Deeper Midline and Mesial Temporal Structures of the Brain Uniquely in Patients with HIV-Associated Dementia. *BMC Infectious Diseases*, vol. 9, no. 1, pp. 192.

ZLOTNIK, A. and YOSHIE, O., 2000. Chemokines: A New Classification Review System and their Role in Immunity. *Immunity*, vol. 12, pp. 121-127.

Websites

http://hapmap.ncbi.nlm.nih.gov/ and NCBI dbSNP

http://www.ncbi.nlm.nih.gov/projects/SNP/databases

http://analysis2.bio-x.cn/SHEsisMain.htm

http://www.cbrc.jp/research/db/TFSEARCH.html

http://frodo.wi.mit.edu/

http://eu.idtdna.com/analyzer/Applications/OligoAnalyzer/

http://tools.neb.com/NEBcutter2/

http://www.ncbi.nlm.nih.gov/tools/primer-blast/

http://countryoffice.unfpa.org/zimbabwe/2010

http://data.unaids.org/publications/irc-pub06/zimbabwe_epi_report_nov05_en.pdf

http://www.worldmapper.org

APPENDICES

Appendix A: Ethics approval, Human Research Ethics Committtee, UCT

UNIVERSITY OF CAPE TOWN

Faculty of Health Sciences
Human Research Ethics Committee
Room E52-24 Groote Schuur Hospital Old Main Building
Observatory 7925
Ms S Arlefdien - Tel: [021]4066492 • Fax: [021]4066411
email: sumayah.arlefdien@uct.ac.za

13 December 2011

HREC REF: 572/2011

Dr C Dandara,
Clinical Lab Sciences
Medical Biochemistry

Dear Dr Dandara,

PROJECT TITLE: THE ROLE OF HOST GENETIC FACTORS IN HIV AND NEUROCOGNITIVE DEVELOPMENT AMONG CHILDREN BORN TO HIV INFECTED MOTHERS

Thank you for submitting your new study to the Faculty of Health Sciences Human Research Ethics Committee

It is a pleasure to inform you that the Ethics Committee has formally approved the above-mentioned study.

Approval is granted until 15 January 2013

Please submit an annual progress report (FHS016) if the research continues beyond the expiry date. Please submit a brief summary of findings if you complete the study within the approval period so that we can close our file (FHS010).

Please note that the ongoing ethical conduct of the study remains the responsibility of the principal investigator.

Please quote the HREC. REF in all your correspondence.

Yours sincerely

PROFESSOR MARC BLOCKMAN
pp **CHAIRPERSON, FHS HUMAN RESEARCH ETHICS**
Federal Wide Assurance Number: FWA00001637.
Institutional Review Board (IRB) number: IRB00001938

This serves to confirm that the University of Cape Town Human Research Ethics Committee complies to the Ethics Standards for Clinical Research with a new drug in patients, based on the Medical Research Council (MRC-SA), Food and Drug Administration (FDA-USA), International Convention on Harmonisation Good Clinical Practice (ICH GCP) and Declaration of Helsinki guidelines.

The Human Research Ethics Committee granting this approval is in compliance with the ICH Harmonised Tripartite Guidelines E6: Note for Guidance on Good Clinical Practice (CPMP/ICH/135/95) and FDA Code Federal Regulation Part 50, 56 and 312.

Appendix B: Ethics Approval from Medical Research Council of Zimbabwe

Telephone: 791792/791193
Telefax: (263) - 4 - 790715
E-mail: mrcz@mrczimshared.co.zw
Website: http://www.mrcz.org.zw

Medical Research Council of Zimbabwe
Josiah Tongogara / Mazoe Street
P. O. Box CY 573
Causeway
Harare

MRCZ APPROVAL LETTER

Ref: **MRCZ/B/218** **10 June 2011**

Kudakwashe Mhandire
Department of Medical laboratory Sciences
College of Health Sciences
University of Zimbabwe
P.O.Box A 178 Avondale
Harare
Zimbabwe

RE: The Role of Host Genetic Factors in HIV and Neurocognitive Development Among Children Born to HIV Infected Mothers.

Thank you for the above titled proposal that you submitted to the Medical Research Council of Zimbabwe (MRCZ) for review. Please be advised that the Medical Research Council of Zimbabwe has **reviewed** and **approved** your application to conduct the above titled study. This is based on the following documents that were submitted to the MRCZ for review:

a) Study protocol.

- **APPROVAL NUMBER : MRCZ/B/218**

This number should be used on all correspondence, consent forms and documents as appropriate.

- **APPROVAL EFECTIVE DATE : 10 June 2011**
- **EXPIRATION DATE : 9 June 2012**
- **TYPE OF MEETING : Expedited**

After this date, this project may only continue upon renewal. For purposes of renewal, a progress report on a standard form obtainable from the MRCZ Offices should be submitted one month before the expiration date for continuing review.

- **SERIOUS ADVERSE EVENT REPORTING**: All serious problems having to do with subject safety must be reported to the Institutional Ethical Review Committee (IERC) as well as the MRCZ within 3 working days using standard forms obtainable from the MRCZ Offices.
- **MODIFICATIONS**: Prior MRCZ and IERC approval using standard forms obtainable from the MRCZ Offices is required before implementing any changes in the Protocol (including changes in the consent documents).
- **TERMINATION OF STUDY**: On termination of a study, a report has to be submitted to the MRCZ using standard forms obtainable from the MRCZ Offices.
- **QUESTIONS**: Please contact the MRCZ on Telephone No. (04) 791792, 791193 or by e-mail on mrcz@mrczimshared.co.zw.
- **Other**
- Please be reminded to send in copies of your research results for our records as well as for Health Research Database.
- You're also encouraged to submit electronic copies of your publications in peer-reviewed journals that may emanate from this study.

Yours Faithfully

.....

MRCZ SECRETARIAT
FOR CHAIRPERSON
MEDICAL RESEARCH COUNCIL OF ZIMBABWE

MEDICAL RESEARCH COUNCIL OF ZIMBABWE
10 JUN 2011
APPROVED
P. O. BOX CY 573 CAUSEWAY, HARARE

PROMOTING THE ETHICAL CONDUCT OF HEALTH RESEARCH
Registered with the USA Office for Human Research Protections (OHRP) as an International IRB (Number IRB00002409 IORG0001913)

Appendix C: Correlation between Haemoglobin levels and human genetic variation regardless of HIV status

Genotypes	**Low Hb** (< 11mg/dL)	**Normal Hb** (11-16mg/dL)	**O**dds **R**atio	P-**value**
CCR2 190G>A; rs17141036				
190G/G	8 (0.80)	50 (0.68)	1.92 (0.34-0.20)	0.42
190G/A	2 (0.20)	21 (0.28)	0.63 (0.06-3.56)	0.58
190A/A	0 (0.00)	3 (0.04)	-	0.52
CX3CR1 745G>A; rs3732379				
745G/G	7 (0.70)	61 (0.84)	0.46 (0.09-3.17)	0.30
745G/A	2 (0.20)	11 (0.15)	1.41 (0.13-8.49)	0.69
745A/A	1 (0.10)	1 (0.01)	8 (0.09-636)	0.09
RANTES In**1.1**T>C;rs2280789				
In1.1T/T	8 (0.89)	42 (0.57)	1	
In1.1T/C	1 (0.11)	30 (0.40)	0.18 (0.00-1.51)	0.08
In1.1C/C	0 (0.00)	2 (0.03	-	0.17
RANTES -403G>A; 2107538				
-403G/G	1 (0.10)	20 (0.27)	0.3 (0.01-2.44)	0.24
-403G/A	8 (0.80)	42 (0.57)	3.05 (0.55-31.02)	0.16
-403A/A	1 (0.10)	12 (0.16)	0.57 (0.01-4.91)	0.61
A3G 197T>C; rs3736685				
197T/T	4 (0.40)	28 (0.38)	1.10(0.21-5.09)	0.89
197T/C	4 (0.40)	34 (0.46)	0.78 (0.15-3.64)	0.72
197C/C	2 (0.20)	12 (0.16)	1.29 (0.12-7.68)	0.76
A3G **-571**G>C;rs5757463				
-571C/C	5 (0.50	57 (0.85)	0.17 (0.03-0.94)	**0.01**
-571C/G	5 (0.50)	10 (0.15)	5.7 (1.06-29.17)	**0.01**
-571G/G	0 (0.00)	0 (0.00)		
A3G **-90**C>G; rs5750743				
-90C/C	1 (0.10)	9 (0.10)	0.81 (0.02-7.29)	0.85
-90C/G	3 (0.30)	34 (0.47)	0.52 (0.08-2.5)	0.36
-90G/G	6 (0.60)	32 (0.44)	2.02 (0.43-10.46)	0.30
A3G 557A>G; rs8177832				
557A/A	3 (0.30)	27 (0.37)	0.73 (0.11-3.55)	0.67
557A/G	5 (0.50)	34 (0.47)	0.72 (0.17-2.78)	0.59
557G/G	2 (0.20)	12 (0.16)	1.27 (0.12-7.56)	0.78
A3G 199C>G;rs2294367				
199C/C	8 (0.89)	62 (0.91)	0.77 (0.08-40)	0.82
199C/G	1 (0.11)	5 (0.07)	1.58 (0.03-16.92)	0.69
199G/G	0 (0.00)	1 (0.02)	-	0.71

Appendix D: Comparison of demographic features between HIV exposed uninfected and HIV unexposed uninfected groups

Characteristics	HIV EU	HIV UEUI	P-value
Mean age in years ±Stdev (range)	8.11±0.47(7.5-9.08)	8.74±0.45(7.5-9.08)	1.00
Mean height in cm ±Stdev (range)	121.43±5.44(111-132)	119.32±6.42(108-133.5	0.168
Mean weight in kg ±Stdev (range)	22.53±2.80(16-27)	22.16±3.35(15-30.5)	0.646
Mean head circum in cm ±Stdev (range)	51.39±1.28(48-54)	51±1.19(49-54)	0.949
Body mass index	n=31	n=31	
Normal (5-85%)	28 (0.90)	27 (0.87)	
Underweight (<5%)	3 (0.10)	4 (0.13)	0.688
Sex	n=36	n=36	
Female	17 (0.47)	19 (0.53)	
Male	19 (0.53)	17 (0.47)	0.637
Haemoglobin	n=34	n=31	
Normal (11-16g/dL)	30 (0.88)	28 (0.90)	
Low (<11g/dL)	4 (0.12)	3 (0.10)	1.00
Morbidities			
Diarrhoea	n=36	n=36	
No	31 (0.80)	28 (0.78)	
Yes	5 (0.14)	8 (0.22)	0.358
Ear Discharge			
No	35 (0.97)	32 (0.89)	
Yes	1 (0.03)	4 (0.11)	0.357
Vomiting			
No	32 (0.89)	29 (0.81)	
Yes	4 (0.11)	7 (0.19)	0.514
Cough			
No	21 (0.58)	23 (0.64)	
Yes	15 (0.42)	13 (0.36)	0.629
Fever			
No	22 (0.61)	21 (0.58)	
Yes	14 (0.39)	15 (0.42)	0.810

Appendix E: Conference acceptance letters and abstracts