

https://dspace.org/

Variation in CCR5, CXCR6 genotypes and differential control of HIV-1 in a cohort in Durban, KwaZulu Natal

Pakati, Nokwanda Faithfull

2021
https://hdl.handle.net/10539/33663

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

Variation in *CCR5*, *CXCR6* genotypes and differential control of HIV-1 in a cohort in Durban, KwaZulu Natal

Nokwanda Faithfull Pakati

A dissertation submitted to the Faculty of Health Sciences, University of Witswatersrand, Johannesburg, in fulfilment of the requirements for a degree of Master of Science (Med) Virology.

Supervisors:

Professor Caroline T. Tiemessen

Prof Koleka Mlisana

Dr Maria Paximadis

Johannesburg, 2021

Declaration

I, Nokwanda Faithfull Pakati (Student number: 2119205) am a student registered for a degree of MSc (Med) Virology in the academic year 2021.

I hereby declare the following:

- I am aware that plagiarism (the use of someone else‘s work without their permission and/or without acknowledging the original source) is wrong.
- I confirm that the work submitted for assessment for the above degree is my own unaided work except where I have explicitly indicated otherwise.
- I have followed the required conventions in referencing the thoughts and ideas of others.
- I understand that the University of the Witwatersrand may take disciplinary action against me if there is a conviction that this is not my own unaided work or that I have failed to acknowledge the source of the ideas or words in my writing.
- I have submitted my final research document to ‘Turnitin’ (i.e. an approved plagiarism detection software) and have attached report as an appendix.

Signature: _____ Date: _____

Dedication

This work is dedicated to my family for encouragement and always believing in me, your love and support is my strength. To my siblings Londeka and Samkele Pakati, success is not final, failure is not fatal: It is the courage to continue that counts.

Abstract

HIV-1 co-receptors such as CCR5 and CXCR6 have been shown to influence HIV-1 acquisition and disease progression to AIDS. Two *CCR5* (rs746492, rs553615728) and two *CXCR6* (rs2234355, rs2234358) single nucleotide polymorphisms (SNPs) that could potentially serve as protective (rs553615728 and rs2234355) and deleterious (rs746492 and rs2234358) markers in the context of HIV-1 control were reported in Johannesburg (JHB) (Picton *et al*., 2017; Koor *et al*., 2019) To investigate the roles of these variants further, this study was undertaken in Durban (DBN), with the aim of comparing the results of the two cohort studies, given that the epidemic dynamics may differ between the two geographic regions.

The study also describes the screening, recruitment as well as the characteristics and challenges of this process. A total of 287/590 met the eligibility criteria. The majority (73.7%) of the screened participants in this cohort were female with the median age of 29 years. The predominant mode of HIV-1 transmission was heterosexual (98.1%). Intermittent condom usage with regular partners and consistent condom usage with casual partners was reported.

The final cohort recruited was assigned to phenotypic groups based on their $CD4^+$ T-cell counts and viral load (VL) measurements -Progressors (N=200; $CD4^+$ T-cell <200 cells/µl and VL>10,000 RNA copies/ml) and Controllers (N=87; $CD4^+$ T-cell >500 cells/µl and VL <2000 RNA copies/ml). These categories were defined using data from a single time point only at an unknown stage after HIV-1 infection. The Controllers were further sub-grouped into elite controllers (ECs, N=31; VL <50 RNA copies/ml) and viraemic controllers (VCs, N=56; VL>50<2000 RNA copies/ml). The median age was comparable between the Controller and Progressor groups but the number of females was significantly more frequent in the Controllers (both ECs and VCs). For this reason gender was used to stratify some of the analyses below. Data for the same SNPs in a control HIV-1 negative cohort of 87 black South Africans and from HIV-1 infected cohorts from JHB (N=134 and N=113), were obtained from two separate studies for comparative purposes.

Comparison of allelic and genotypic frequencies of the *CCR5* rs746492, *CXCR6* rs2234355 and *CXCR6* rs2234358 SNPs between the HCs and other reference populations revealed some interesting results. Not unexpectedly, all three SNPs showed significantly different allelic representation compared to the EUR population, with *CXCR6* rs2224355 showing the largest difference (P<0.0001) due to the rarity of this SNP in the EUR population group. Compared to

the African populations, the HC group was more similar to the East African population and significantly different to the West African population.

Comparisons of both allelic and genotypic frequencies for all four variants in the DBN cohort revealed no significant differences for both the *CCR5* SNPs and the *CXCR6* rs223455 SNP. However, *CXCR6* rs2234358 minor allele (T) and minor allele homozygosity (TT) showed strong trends of underrepresentation in VCs compared to HCs (P=0.05 for both). When we compared representation of these variants across JHB and DBN cohorts, there were significant differences in the *CCR5* rs746492 minor (G) allele frequency in the Controller groups (P=0.02). The *CXCR6* rs2234355 genotypic frequency was significantly different between the JHB and DBN cohorts in the Progressor (P=0.03) and VC (P=0.04) groups. The allelic and genotypic frequencies for the *CCR5* and *CXCR6* SNPs in the DBN cohort were more aligned with HCs.

Assessing the effect of these variants on VL in the Progressor group showed no significant associations with the *CCR5* genotypes. Interestingly, the *CXCR6* rs2234358 SNP was significantly associated with higher VL (P=0.0003) with no significant associations of these variants on $CD4^+$ T-cell count.

The *CXCR6* rs2234355 SNP genotypes were significantly differentially distributed between genders in the Controllers, Progressors, and VCs. Given the difference in gender distribution, we assessed the effect on female VCs and Progressors. There was a strong trend of overrepresentation in VCs compared to Progressors in females (P=0.055) as well as the *CXCR6* rs2234355 SNP in dominant mode (GA+AA) (P=0.078).

The combined effect of the two *CXCR6* variants was assessed. The *CXCR6* rs2234355 protective genotype (GA+AA) in the absence of *CXCR6* rs2234358 deleterious genotype (TT) showed a significant higher representation in VCs versus Progressors (P=0.034) and this association was strengthened when females alone were compared between these two study groups (P=0.018). Frequency of 3-SNP genotype combination across *CCR5* and *CXCR6* were not significantly different between VCs and HCs.

In conclusion, the most significant findings in this study centred on CXCR6 and not CCR5, with a stronger effect seen in the female gender. The differences seen between the DBN and JHB cohorts suggest that either ethnic differences, or the epidemic dynamics, or a combination of both may be contributing factors, and that careful consideration needs to be applied in future

analyses going forward with regard to merging data across cohorts. Lastly and importantly, this study has served to yet again highlight an important role for CXCR6 in viraemic and not elite natural HIV-1 control.

Acknowledgments

My sincere gratitude goes to:

My supervisors: Prof Caroline Tiemessen, thank you for granting me the opportunity to undertake this course. Thank you for your support and guidance throughout this project.

Dr Maria Paximadis, I sincerely appreciate all the scientific advice, guidance, patience, commitment and time you have invested in me. Thank you for the training on the study methodologies, the knowledge I gained there was invaluable.

Prof Koleka Mlisana for your support, great ideas, motivation and opening your laboratory for the duration of my MSc. Dr Ravesh Singh and everyone in UKZN Medical Microbiology lab for excellent ideas, mentorship, guidance and training experience. I couldn't have chosen a better place to undertake this research.

Prof Photini Kiepiela and Shabashini Reddy, no words could ever describe my gratitude for such great mentors. Thank you for your commitment, time, ideas, guidance, encouragement and support. Thank you for the opportunity to be trained and serve as a counsellor on the study, which has provided me with such important insights into social aspects of HIV infection. God bless you.

Dr Tarlyee Reddy and Nobuhle for statistical analysis and interpretation.

Prof. Michele Ramsay and the H3Africa Consortium, for permitting the use of gene sequence data for the South African control group (HC) in this study.

Family for love and support, especially my parents MaNyanisa and baba Pakati your words of encouragement, constructive debates and laughter you shared with me.

Friends: Nqubeko L, Luleko C, Londi M, Sthe M, Sane M, Thando M and Sizanani M for listening to my frustrations and sharing my excitements. Good friends are like a four-leaf clover: hard to find and lucky to have.

The funders who made this study possible; the South African Research Chairs Initiative of the Department of Science and Technology and National Research Foundation of South Africa, and the Strategic Health Innovation Partnerships (SHIP) Unit of the South African Medical Research Council (a grantee of the Bill and Melinda Gates Foundation).

Our study participants, your commitment has made our efforts in understanding HIV-1 disease progression worthwhile.

Last but not least, I'd like to express my gratitude to God, the Almighty, for providing me with the opportunity to undertake such a monumental task and for providing me with the strength to complete it.

Table of contents

Declaration……..……i

Dedication……..ii

Abstract…….....iii

Acknowledgements…….....vi

Table of contents……viii

List of figures…….xiii

List of tables….....xviii

List of abbreviations……..…….....xix

CHAPTER 1: Introduction……1

1.1. The HIV-1 epidemic.…..……..…….2

1.2. HIV entry and integration into the target cell…….3

1.3. Immune response to HIV-1 infection……4

1.3.1. The innate immune response…….…..…...5

1.3.2. Adaptive immune response…….....…..5

1.4. The course of HIV-1 infection…….…..……...…6

1.5. Natural control of HIV-1 infection…….……..…….....9

1.6. Simian immunodeficiency viruses and HIV-1 disease progression……..……..10

1.7. HIV-1 treatment and cure……...…….11

1.8. The influence of host genetic variation on HIV-1 infection……..…….13

1.8.1. Chemokine receptors…….……..….14

1.8.1.1. CCR5 chemokine receptor…….……...15

1.8.1.1.1. The *CCR5* gene…….……...……16

1.8.1.1.2. Variation in the *CCR5* gene…….....17

A. *CCR5*Δ*32* polymorphism……..……...17

B. CCR5 haplotypes……...……...18
C. CCR5 promoter polymorphisms……..…..……20
D. Mutations affecting the coding sequence of *CCR5*……..…..……20
E. Other mutations found in within the *CCR5* gene……...…..…..…21
1.8.1.2. CCR2 polymorphism (V64I)…….…..…....23
1.8.1.3. CXCR6 chemokine receptor…….…….…...24
1.8.1.3.1. CXCR6 in HIV-1 infection…….…….……..…25
1.8.2. Chemokines…….……..…26
1.8.2.1. CC chemokines……..27
1.8.2.2. Copy number variation in CC chemokine ligands……..…….…27
1.8.2.3. CXC chemokines and ligands……..…....28
1.8.3. Variation in human leukocyte antigen (HLA) molecules…….…..…..29
1.8.4. Variation in killer immunoglobulin-like receptor (KIR) molecules in HIV-1 control……..……...…….……32
1.9. Recruitment of cohorts of HIV-1 infected individuals with long term follow-up…..…33
1.9.1. Loss to follow-up…….…….....34
1.9.2. Efforts to promote successful follow-up…….…….……35
1.10. Study rationale……..35
1.10.1. Aim of the study…….……..….36
1.10.2. Study specific objectives…….…….……36
CHAPTER 2: Materials and Methods…….…….…...37
2.1. Study design and participants…….……..…..38
2.1.1. Screening and enrolment procedures……..……..….39
2.1.2. Follow-up procedures…….……..…40
2.2. Laboratory methods……..……...42

2.2.1. $CD4^+$ T-cell count enumeration……42

2.2.2. HIV-1 RNA measurements……42

2.2.3. Specimen preparation and storage……42

2.2.4. Peripheral Blood Mononuclear Cells (PBMCs) separation and storage……42

2.2.5. Extraction of genomic DNA……42

2.2.6. Single nucleotide polymorphism (SNP) genotyping……43

2.2.6.1. Real-time C_T Shift Assay……43

2.2.6.2. Predesigned Taqman genotyping assay……48

2.3. Ethical clearance……49

2.4. Data collection and documentation……49

2.5. Data analysis……50

CHAPTER 3: Results……51

3.1. Clinical demographics……52

3.1.1. Screening and enrolment of eligible participants……52

3.1.1.1. HIV-1 Controllers……52

3.1.1.2. Progressors……53

3.1.2. Demographic, sexual and behavioral characteristics of all participants (N=590) screened……53

3.1.3. $CD4^+$ T-cell counts and viral load measurements for enrolled participants……55

3.2. *CCR5* (rs746492 and rs553615728*)* and *CXCR6* (rs2234355 and rs2234358) SNPs in the Black South African population and other reference populations……56

3.3. The role of *CCR5* (rs746492 and rs553615728*)* and *CXCR6* (rs2234355 and rs2234358) variants in HIV-1 control……62

3.3.1. Validation of the Taqman and C_T Shift assay for *CCR5* (rs746492) SNP……62

3.3.2. Comparisons of allelic and genotypic frequencies of *CCR5* and *CXCR6* SNPs between study groups……..……...……..64

3.3.2.1. Comparisons of allelic frequencies of *CCR5* and *CXCR6* SNPs between study groups……..……..66

3.3.2.2. Comparisons of genotypic frequencies of *CCR5* and *CXCR6* SNPs between study groups…….……..68

3.4. Investigation of *CCR5* and *CXCR6* variants and markers of disease progression: viral load and $CD4^+$ T cell count……...…..73

3.4.1. Effect of *CCR5* and *CXCR6* genotypes on viral load in the Progressor and VC groups……...73

3.4.2. Effect of *CCR5* and *CXCR6* SNP genotypes on $CD4^+$ T-cell counts……76

3.5. Analysis of the *CCR5* and *CXCR6* allelic and genotypic frequencies on HIV-1 control taking gender differences into account…….…..….79

3.5.1. Analysis of the *CXCR6* rs2234355 variant in viraemic HIV-1 control in the female gender groups…….……..87

3.6. The analysis of the combined effect of the *CCR5* and *CXCR6* SNPs…….……..88

3.6.1. Analysis of the combined effect of the two *CXCR6* variants (rs2234355 and rs2234358 SNPs)……..88

3.6.2. Analysis of the combined effect of protective genotypes of *CCR5* (rs746492 and *CXCR6* (rs2234355 and rs2234358) …….…….……..90

3.7. Comparison of the frequency of *CCR5* (rs746492 and rs553615728) and *CXCR6* (rs2234355 and rs2234358) SNPs between HCs, Durban and Johannesburg cohorts..91

3.7.1. Allelic comparisons of *CCR5* and *CXCR6* SNPs between HCs, Durban and Johannesburg cohorts……..….92

3.7.2. Genotypic comparisons of *CCR5* and *CXCR6* SNPs between HCs, Durban and Johannesburg cohorts……..…….94

CHAPTER 4: Discussion and conclusion….....99

REFERENCES……...……117

APPENDIX A: Ethical clearance…….……..……14 2

APPENDIX 1A: Human Research on Human Subjects Committee of the University of the Witwatersrand, Johannesburg……..…….142

APPENDIX 2A: Ethics clearance for the cohort in KZN was obtained from Medical Research Council Ethics Committee….....144

APPENDIX 3A: EThekwini Research Ethics Committee……..……145

APPENDIX 4A: KZN Department of Health approval letter…….……...146

APPENDIX B: Turnitin report….....147

List of figures

Figure 1.1: Schematic representing different conformational changes in gp120 and gp41 upon binding to the host cell surface receptors……..…4

Figure 1.2: Schematic showing the typical course of HIV-1 infection……7

Figure 1.3: Schematic representation of different rates of disease progression to AIDS in various individuals…….…….……9

Figure 1.4: Different antiretroviral regimens and their mode of actions…….……12

Figure 1.5: Schematic representation of the CCR5 showing amino acid sequence and seven transmembrane structure……...……15

Figure 1.6: Schematic representation of (A) *CCR5* gene and (B) CCR5 mRNA transcripts….17

Figure 1.7: Schematic representation of the *CCR5* haplotypes displaying 9 evolutionarily distinct haplotypes involving CCR5 promoter region SNPs, *CCR2*-V64I and CCR5Δ32 mutations……..19

Figure 1.8: Schematic showing chemokine subclasses as per number and spatial organization of conserved cysteine residues in their N-terminus…….……27

Figure 1.9: Schematic showing the HLA complex genes and their protein product classes (I, II, and III)…….…….…..30

Figure 1.10: Schematic representation of chromosome 19 (19q13.4) region where the *KIR* genes are located…….…..32

Figure 2.1: Flow diagram showing HIV-1 study design for the HIV-1 Functional cure study…….41

Figure 2.2: Example C_T-shift patterns shown for: (A) *CCR5* rs746492 SNP; (B) *CCR5* rs553615728 SNP; (C) *CXCR6* rs2234355 SNPs and (D) *CXCR6* rs2234358 SNP…….….47

Figure 2.3: An allelic discrimination plot depicting results for the predesigned Thermo Fisher

Taqman SNP genotyping assay…….…….……..49

Figure 3.1: Schematic flow diagram for screening and enrolment of participants for the HIV-1 Functional cure DBN cohort study…….……...52

Figure 3.2: Map of the African continent depicting the African countries represented in the 1000 Gemones Project……..……57

Figure 3.3: Map of the European continent depicting the European countries represented in the 1000 Genomes Project…….…….…58

Figure 3.4: An allele discrimination plot indicating 12 samples used to validate the C_T shift assay for the *CCR5* rs746492 SNP…….……...63

Figure 3.5: C_T shift results for the 12 samples validating the Taqman assay for the *CCR5* rs746492 SNP: A) homozygous wildtype (WT), B) heterozygous (Het.) and C) homozygous mutants (M)……..64

Figure 3.6: Bar graph showing allele frequency for, (A) *CCR5* rs746492 and (B) *CCR5* rs553615728 in the HCs, Controllers, Progressors, ECs and the VCs…….....67

Figure 3.7: Bar graph showing allele frequency for, (A) *CXCR6* rs2234355 and (B) *CXCR6* rs2234358 in the HCs, Controllers, Progressors, ECs and the VCs……..…….68

Figure 3.8: Pie charts showing genotypic representation (frequencies; %) for (A) *CCR5* rs746492 and (B) *CCR5* rs553615728 SNPs for the various groups……...69

Figure 3.9: Pie charts showing genotypic representation (frequencies; %) for (A) *CXCR6* rs2234355 and (B) *CXCR6* rs2234358 SNPs for the various groups…….…...70

Figure 3.10: Comparison of viral load in the Progressor and VC groups for: (A) *CCR5* rs746492 SNP genotypes and (B) *CCR5* rs553615728 SNP genotypes…….…….73

Figure 3.11: Comparison of viral load in the Progressor and VC groups for: (A) *CXCR6* rs2234355 SNP genotypes and (B) *CXCR6* rs2234355 SNP in the dominant mode (GA+AA)…….74

Figure 3.12: Comparison of viral load in the Progressor and VC groups for: (A) *CXCR6* rs2234358 SNP genotypes and (B) *CXCR6* rs2234358 SNP by grouping individuals with the major allele (i.e. GG+GT)……..……..…….75

Figure 3.13: Comparison of $CD4^+$ T-cell count for: between respective genotypes for: (A) *CCR5* rs746492 and (B) *CCR5* rs553615728 SNPs in the Controller, Progressor, EC and VC groups……...77

Figure 3.14: Comparison of $CD4^+$ T-cell count for: between respective genotypes for: (A) *CXCR6* rs2234355 and (B) *CXCR6* rs2234358 SNPs in the Controller, Progressor, EC and VC groups……..…….78

Figure 3.15: Pie charts showing representation of the *CCR5* rs746492 alleles in (A) Controller,(B) Progressor, (C) EC and (D) VC groups according to gender…….…….79

Figure 3.16: Pie charts showing representation of the *CCR5* rs553615728 alleles in (A) Controller, (B) Progressor, (C) EC and (D) VC groups according to gender…….80

Figure 3.17: Pie charts showing representation of the *CXCR6* rs2234355 alleles in (A) Controller, (B) Progressor, (C) EC and (D) VC groups according to gender…….81

Figure 3.18: Pie charts showing representation of the *CXCR6* rs2234358 alleles in (A) Controller, (B) Progressor, (C) EC and (D) VC groups according to gender…….82

Figure 3.19: Pie charts showing representation of the *CCR5* rs746492 genotypes in (A) Controller, (B) Progressor, (C) EC and (D) VC groups according to gender…….83

Figure 3.20: Pie charts showing representation of the *CCR5* rs553615728 genotypes (A) Controller, (B) Progressor, (C) EC and (D) VC groups according to gender……..84

Figure 3.21: Pie charts showing representation of the *CXCR6* rs2234355 genotypes (A) Controller, (B) Progressor, (C) EC and (D) VC groups according to gender……..85

Figure 3.22: Pie charts showing representation of the *CXCR6* rs2234358 genotypes (A) Controller, (B) Progressor, (C) EC and (D) VC groups according to gender……..86

Figure 3.23: Bar graphs showing *CXCR6* rs2234355 SNP: (A) distribution and comparisons of alleles in the VC and Progressor female groups and (B) comparison of the variant in the dominant mode (GA+AA) between VC and Progressor female groups……..…….88

Figure 3.24: Bar graphs showing comparison of combinations of (A) the *CXCR6* rs2234355 (355 GA+AA) and rs2234358 (-358TT) genotypes inHC, Progressor, EC and VC groups (combined females and males) and (B) the *CXCR6* rs2234355 (355GA) and rs2234358 (-358TT) in Progressor, EC and VC groups (females only)…….……89

Figure 3.25: Bar graphs showing comparison of combinations of the *CXCR6* rs2234355 (355 GA+AA), rs2234358 (-358TT) and *CCR5* rs746492 (+2919TT) genotypes in total HC, Progressor, EC and VC groups (combined females and males) as well as in Progressors, ECs and VCs females only)……..……91

Figure 3.26: Bar graphs showing the comparison of allele distributions between Durban and Johannesburg cohorts for the: (A) *CCR5* rs746492 SNP and (B) *CCR5* rs553615728 SNP in HC, Controller, Progressor, EC and VC groups……..…….93

Figure 3.27: Bar graphs showing the comparison of allele distributions between Durban and Johannesburg cohorts for the: (A) *CXCR6* rs2234355 SNP and (B) *CXCR6* rs2234358 SNP in HC, Controller, Progressor, EC and VC groups……..…….94

Figure 3.28: Comparison of genotypic distributions between Durban and Johannesburg cohorts for the: (A) *CCR5* rs746492 SNP and (B) *CCR5* rs553615728 SNP in HC, Controller, Progressor, EC and VC groups…….....97

Figure 3.29: Comparison of genotypic distributions between Durban and Johannesburg cohorts for the: (A) *CXCR6* rs2234355 SNP and (B) *CXCR6* rs2234358 SNP in HC, Controller, Progressor, EC and VC groups…….……...……..98

List of tables

Table 2.1: C_T shift assay real-time PCR reaction component and setup……..……45

Table 2.2: PCR primers and thermal cycle conditions used for the C_T shift assays for genotyping the *CXCR6* and *CCR5* SNPs……...46

Table 3.1: Characteristics of the recruited cohort (N=590) prior to screening out non-eligible individuals…….…….……..55

Table 3.2: $CD4^+$ T-cell counts, viral load measurements, gender and age of the enrolled participants (N=287)……....56

Table 3.3: Comparison of the representation of *CCR5* (rs746492) and *CXCR6* (rs2234355 and rs2234358) SNPs between a Black South African Healthy population to African and European populations……..…….61

Table 3.4: Number and frequency (%) of alleles and genotypes for the *CCR5* and *CXCR6* variants in the HIV-1 infected Controller, Controller sub-groups, Progressors as well as the HIV-1 negative healthy controls (HCs)…….……65

Table 3.5: Comparisons of the *CCR5* and *CXCR6* SNP allelic and genotypic frequencies in Controllers, controller sub-groups and Progressors to healthy controls (HCs)…….…...71

Table 3.6: Comparisons of the *CCR5* and *CXCR6* SNP allelic and genotypic frequencies in Controller, VC and EC groups to the Progressor group…….……..72

List of abbreviations

7TM	**Seven transmembrane**
ADCC	Antibody-dependent cellular cytotoxicity
AFR	African population
AGM	African green monkey
AIDS	Acquired immunodeficiency syndrome
APC	Antigen-presenting cells
APJ	Apelin receptor
ART	Anti-retroviral therapy
ARV	Antiretroviral treatment
ATF 1	Activating transcription factor 1
B-cells	B-lymphocytes
cART	Combination antiretroviral therapy
CCL	Chemokine ligand
CCL3	CC chemokine ligand 3
CCL3L	CC chemokine ligand 3-like
CCL3L1	CC chemokine ligand 3-like 1
CCL4	CC chemokine ligand 4
CCL4L	CC chemokine ligand 4-like
CCL5	CC chemokine ligand 5
CCR2	CC chemokine receptor 2
CCR5	CC Chemokine receptor 5
CCR5Δ32	CCR5 delta 32 (32 base pair deletion)
CD4	Cluster of differentiation 4
CD8	Cluster of differentiation 8
cDNA	Complementary deoxyribonucleic acid
CEU	Utah residents with European ancestry
CI	Confidence interval
CNV	Copy number variation
CpG	Cytidine phosphate guanidine
CREB1	cAMP responsive element binding protein 1
C_T	Cycle threshold
CTL	Cytotoxic T-lymphocytes
CXCL	CXC chemokine ligand
CXCR4	CXC chemokine receptor 4
CXCR6	CXC chemokine receptor 6
DBN	Durban
DC	Dendritic cell
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid

e.g.	For example
EC	Elite controller
EDTA	Ethylene-diamine-tetra acetic acid
Env	Envelope protein
ESN	Esan in Nigeria
et al.	And others
EUR	European population
FBS	Fetal bovine serum
FIN	Finnis in Finland
GBR	British in England and Scotland
gDNA	Genomic DNA
GIT	Gastrointestinal tract
Glu-Leu-Arg	Glutamic acid-leucine-arginine
gp120	Glycoprotein 120
gp41	Glycoprotein 41
GPCR	G protein coupled receptor
GWAS	Genome-wide association studies
GWD	Gambia in Western Division
H_2O	Water
HCC	Haem-filtrate CC chemokine
HCs	Healthy controls
HCV	Hepatitis C virus
HH(A)	Human haplotype A
HIV-1	Human immunodeficiency virus type 1
HIV-2	Human immunodeficiency virus type 2
HLA	Human leukocyte antigen
HTLV	Human T- lymphotropic retroviruses
i.e.	That is
IBS	Iberian populations in Spain
IC	Informed consent
IFN-α	Interferon-α
IFN-β	Interferon-β
IN	Integrase
IQR	Interquartile range
JHB	Johannesburg
KIR	Killer cell immunoglobulin-like receptor
KZN	KwaZulu Natal
LD	Linkage disequilibrium
LDL	Lower than detectable limit
LILRs	Leukocyte immunoglobulin-like receptors
LN	Lymph node
LNA	Lock nucleic acid

lncRNA	Long non-coding Ribonucleic acid
LRC	Leukocyte receptor complex
LTNP	Long term non-progressor
LWK	Luhya in Webuya, Kenya
MGB	Minor grove binder
MHC	Major histocompatibility complex
MIP-1α	Macrophage inflammatory protein-1α
MIP-1β	Macrophage inflammatory protein-1β
Ml	Millilitre
MSL	Mende in Sierra Leone
Mt	Mutant
MTCT	Mother-to-child-transmission
N	Number of study participants
NC	Non controllers
NHLS	National Health Laboratory Services
NHP	Non-human primate
NK	Natural killer
NNRTIs	Non-nucleoside reverse transcriptase inhibitors
NRTIs	Nucleoside reverse transcriptase inhibitors
NTC	Non template control
OR	Odds ratio
ORF	Open reading frame
PBMC	Peripheral blood mononuclear cell
PCP	*Pneumocystis jiroveci* pneumonia
PCR	Polymerase chain reaction
PD-1	Programmed death-1
pDCs	Plasmacytoid dendritic cells
PHC	Primary healthcare clinic
PIs	Protease inhibitors
PLG	Pan Leucogating
PMTCT	Prevention of mother-to-child transmission
ppt.	Participants
PTID	Participant's identity
qPCR	Quantitative real-time polymerase chain reaction
RANTES	Regulated upon Activation, Normal T-cell Expressed, and Secreted
RCMs	Red-capped mangabeys
RM	Rhesus macaques
RNA	Ribonucleic acid
RT	Reverse transcriptase
SA	South Africa
SAB	South African Black

SIV	Simian immunodeficiency virus
SM	Sooty mangabeys
SNP	Single nucleotide polymorphism
SR-PSOX	Scavenger receptor for and low density lipoprotein
SST	Serum separating tubes
TB	Tuberculosis
T-cell	T-lymphocytes
Tcm	Central memory T-cell
TCR	T-cell receptors
Tfh	T-follicular helper
Th17	T-helper 17
TLR7	Toll-like receptor 7
TNF-α	Tumour necrosis factor α
Tscm	Stem-cell memory T-cell
TSI	Toscani in Italy
UTR	Untranslated region
UTT	Universal test and treat
VC	Viraemic controller
VCT	Voluntary counselling and testing
VL	Viral load
VLS	Viral load set point
vs.	Versus
WGS	Whole Genome Sequencing
WHO	World Health Organization
WT	Wild type
YRI	Yoruba in Ibadan, Nigeria

CHAPTER 1: Introduction

1.1. The HIV-1 epidemic

The acquired immunodeficiency syndrome (AIDS) was first described in 1981 and was initially thought to be caused by a human T-lymphotropic retrovirus (HTLV) (Barre-Sinoussi *et al.*, 1983). Human immunodeficiency virus (HIV), a retrovirus found in 1986 which at that time was mainly associated with homosexual men is the causative of AIDS (Coffin *et al.*, 1986). HIV is classified into two different types, human immunodeficiency virus type 1 and 2 (HIV-1 and HIV-2). HIV-2 disease progression is slower than in HIV-1 and is less infectious. HIV-2 has limited effects on survival in most infected adults (Marlink *et al*., 1994), although it seemingly shows the same clinical symptoms to HIV-1 (Norrgren *et al*., 1995)**.** Globally, HIV-1 predominates and it is accountable for most infections worldwide, whereas HIV-2 is very rare outside of West Africa (Campbell-Yesufu and Gandhi, 2011).

Epidemiological and genetic evidence suggests that introduction of HIV into humans happened by zoonosis (human infection originating from animals) and by means of numerous independent transmission events from the sooty mangabey (SM, *Cercocebus atys*) and chimpanzee (*Pan troglodytes*) (Santiago *et al.*, 2005). The most common mode of transmission of the virus is through sexual transmission, with heterosexual transmission being the predominant mode in most developing countries. Other modes of transmission are through the sharing of contaminated needles, infected blood or blood products, mother to child and donated organs (Adler, 2001). Although there was a great deal of public concern when HIV/AIDS was discovered, the virus is not spread by casual or social contact.

Globally, in 2019, an estimated 37.9 million individuals were infected with HIV and about 1.7 million amongst them were children <15 years of age (UNAIDS, 2019). There was a global decline in deaths and in AIDS-related illness mainly due to the progress made by the availability of antiretroviral therapy (ART) in sub-Saharan Africa, mostly in eastern and southern Africa, with 53% of the world's HIV infected individuals (UNAIDS, 2018a). The gender gap is still prominent in sub-Saharan Africa, where women account for 56% of people living with HIV (UNAIDS, 2018a). South Africa (SA) has the highest HIV infection rate globally, with approximately 7.7 million individuals living with HIV in 2018 (UNAIDS, 2018b) and KwaZulu Natal (KZN) has the highest HIV-1 prevalence of the nine provinces in SA, estimated at 27% in 2017 (HSRC, 2018).

Without treatment, HIV-1 infection leads to functional impairment of lymphocyte populations (Brenchley *et al*., 2004) and impairment in CD4 positive T-lymphocytes ($CD4^+$ T-cells), CD8

positive T-lymphocytes ($CD8^+$ T-cells), Natural killer cells (NK cells) and B lymphocytes (B-cells) (Sauce *et al*., 2011), bringing about disease progression to AIDS-defining illnesses and loss of cellular immunity. The use of ART has resulted in a significant reduction in the morbidity and mortality associated with HIV-1 infection for persons with access to treatment (Lohse *et al.,* 2007). Most individuals on ART have achieved substantial reconstitution of their $CD4^+$ T-cell numbers and almost complete viral suppression. Despite this achievement, ART is not curative and it has numerous limitations, which include short and long term toxicities leading to discontinuation of therapy and rapid rebound of viral replication (Chawla *et al*., 2018; Chun *et al*., 1997). Therefore, further advances in treatment and new therapeutic interventions with the aim of attaining HIV remission, if not a "cure", are still urgently required.

1.2. HIV entry and integration into the target cell

The first step of the HIV life cycle is viral entry into the target cell. This is a multistep process which is dependent on the sequential interactions with two host receptors namely, the $CD4^+$ T-cell receptor and a co-receptor (a number of co-receptors have been identified). The CD4-domain of the viral glycoprotein 120 (gp120) binds to the CD4 receptor on the host cell, bringing about a conformational change in gp120 enabling it to bind to either the seven transmembrane CC chemokine receptor type-5 (CCR5) or the seven transmembrane CXC chemokine receptor type 4 (CXCR4 or fusin) on the cell surface depending on viral tropism (Dean *et al*., 1996; Feng *et al*., 1996). Binding of gp120 to the CD4 receptor and the co-receptor triggers another conformational change in glycoprotein 41 (gp41) (Luciw *et al*., 1996; Archin *et al*., 2014) enabling gp41 to fuse with the cellular membrane leading to the uncoating of the viral capsid and releasing its components into the cytoplasm (Balakrishna and Kondapi, 2016) as shown on Figure 1.1.

Activation of the viral reverse transcriptase (RT) enzyme occurs in the cytoplasm where it transcribes the single-strand HIV ribonucleic acid (RNA) genome into complementary DNA (cDNA), whereby the RNA strand is simultaneously degraded as the deoxyribonucleic acid (DNA) is being synthesised (Hu and Hughes, 2012). Thereafter the single-stranded cDNA is converted by the DNA polymerase into a double-stranded DNA which migrates into the cell nucleus (Sloan and Wainberg, 2011). The viral integrase (IN) enzyme then inserts the proviral genome at random into the human host cell genome. HIV infection in the cell and development of a persistent infection are finalized when the proviral DNA is integrated (Hindmarsh and

Leis, 1999). The main co-receptors utilised by HIV-1 are CCR5 and CXCR4, but other co-receptors have also been identified which can support HIV-1, HIV-2 and SIV infection as potential co-receptors such as -CCR1, CCR2B, CCR3, CCR4, CCR6, CCR8, CXCR1, CXCR3, CXCR6, CX3CR1, ChemR23, apelin receptor (APJ), and CXCR7 (Pollakis and Paxton, 2012; Arenzana-Seisdedos and Parmentier, 2006).

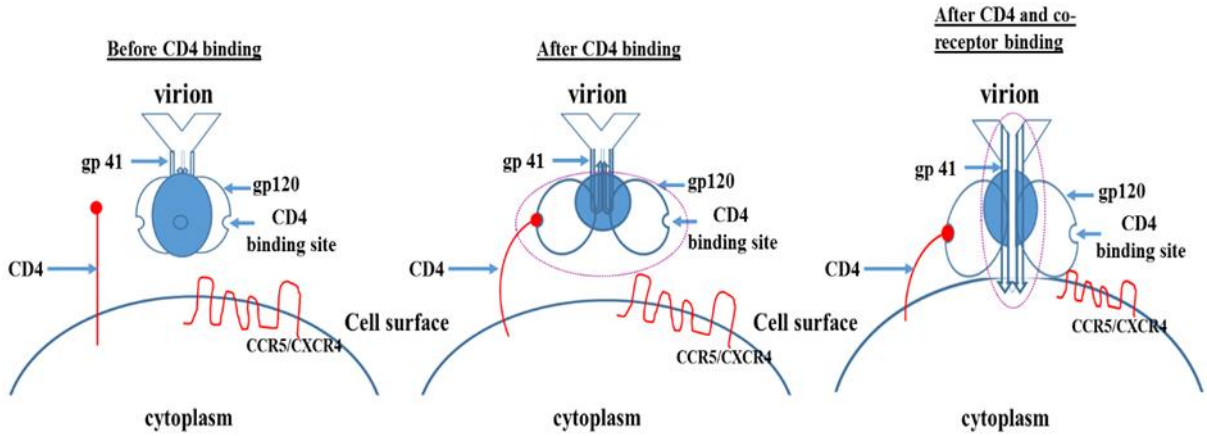


Figure 1.1. Schematic representing different conformational changes in gp120 and gp41 upon binding to the host cell surface receptors. Reproduced from Balakrishna and Kondapi (2016).

1.3. Immune response to HIV-1 infection

HIV-1 infection is characterized by the progressive decline of $CD4^+$ T-cell counts and surge of viral replication leading to high viral load (VL), with eventual stabilization to a viral load set point (VLS) (Pilcher *et al*., 2004), indicative of the immune system's partially successful attempt to fight off the virus in the absence of ART. The immune system employs multiple mechanisms in order to protect the body from diseases. The mechanisms fundamental to the establishment of VLS a couple of months after acute infection are likely to involve the interplay between host immune responses and viral replication (Mellors *et al.,* 2007). The immune system can broadly be divided into two components, namely the innate immune system and the adaptive immune system (Janeway, 1989).

1.3.1. The innate immune response

The early innate immune response, prior to the activation of the adaptive response, determines the course of HIV-1 disease progression. Innate immune cells (i.e. NK cells and dendritic cells (DCs), are part of the first line of defence when HIV-1 enters the body. NK cells, the counterpart of the T-lymphocytes (T-cells) of the adaptive immunity, produce pro-inflammatory cytokines and cytotoxic molecules which possess anti-HIV properties (Biron *et al.*, 1999) and also undertake the surveillance and killing of foreign or infected cells. NK cells have a diverse set of receptors on their surface and recognise human leukocyte antigen (HLA) molecules. A major group of these receptors are the killer immunoglobulin-like receptors (KIRs) (Long *et al*., 1996). DCs play an important role in linking innate and adaptive immunity by processing foreign antigens and presenting it to components of the adaptive immune system (Palucka and Banchereau, 1999). Plasmacytoid DCs (pDCs), a subset of DCs, secrete significant amounts of pro-inflammatory cytokines IL-6, tumour necrosis factor (TNFα) and macrophage inflammatory protein (MIP-1β) as well as inflammatory cytokines such as interferon alpha (IFN- α), and interferon beta (IFN-β) (Gilliet and Lande, 2008), which regulate T, B, and NK cell responses, together with IFN-α/β in response to viral infections. The secretion of these cytokines set up an antiviral environment thus inhibiting viral replication, activating and recruiting other immune cells and inducing adaptive immune responses (Carrington and Alter, 2012).

1.3.2. Adaptive immune response

The adaptive immune response which involves the T and B-cells is the most critical component in the control of HIV-1 infection (Bonilla and Ottegen, 2010). Entry of HIV into the target T-cells induces the cellular immune response and synthesis of viral proteins. An immune response is initiated when the major histocompatibility complex (MHC) class I on the cell surface presents the intracellularly degraded HIV peptide fragments which are recognised by T-cell receptors (TCR) on cluster of differentiation 8 cells ($CD8^+$ T-cells) (Hansson *et al*., 2002). $CD8^+$ T-cells play an important role in viral eradication by producing cytokines that inhibit viral replication (Cocchi *et al*., 1995), as well as secretion of perforin and granzymes A and B which are involved in the cytolysis of virus-infected cells resulting in the decline of viraemia after primary infection (Wherry *et al*., 2003). B-cell mediated antibody responses to HIV-1 infection are first present around 8 days post viral load detection in the form of immune complexes (Tomaras *et al*., 2008). Antibody-dependent cell-mediated cytotoxicity (ADCC) (a cytolytic mechanism based on antibody coating of an infected cell), is another B-cell associated

mechanism of clearing infectious agents which promotes lysis of target cells by effector cells (mainly NK cells) (Chung *et al.*, 2009). Longterm survivors of HIV-1 infection who maintain high of $CD4^+$ T-cell counts and low VLs have been shown to have high levels of neutralizing antibodies (Cao *et al*., 1995).

1.4. The course of HIV-1 infection

HIV-1 progressively weakens the body's ability to combat viruses and some cancers over time. Individuals infected with HIV-1 are more likely to develop life-threatening diseases and opportunistic infections, and if left untreated, they will eventually die. Clinically, there are three major phases of adult HIV-1 infection: acute, chronic, and AIDS (Figure 1.2). The acute or primary infection phase which is characterized by flu-like symptoms with fever, fatigue, muscle aches, nausea and diarrhoea, may persist for 2 to 4 weeks (Coffin and Swanstrom 2013). Acute infection is marked by high viraemia of up to 10^7 copies or more of viral RNA copies/ml of blood, peaking around 6 weeks after infection and severe $CD4^+$ T cell loss. HIV-1-specific antibodies and cytotoxic T-lymphocytes (CTLs, in the form of $CD8^+$ T-cells) that target HIV-1 antigens expressed on infected cells make an appearance after 3-6 weeks of infection (Coffin and Swanstrom, 2013). During this acute phase of peak viraemia, high titres of HIV-1 are detected in the blood using a p24 antigen test; however, HIV antibody tests such as enzyme immunoassay are often negative in the first three weeks (Fiebig *et al*., 2003). Individuals in this stage of infection are the most infectious because of the high VL in blood and genital secretions - over half of HIV transmissions may take place during this period (Shaw and Hunter, 2012).

At the end of the acute phase there is a marked drop in VL with the establishment of a VLS, likely as the result of partial immune control as well as exhaustion of activated $CD4^+$ T-cells. The rate of disease progression has been associated with VLS where a lower VLS was associated with a slower rate of disease progression and vice versa (Mellors *et al*., 1995). The asymptomatic chronic phase, or the latency period, ranges from 1-10 years where there is a gradual decrease in $CD4^+$ T-cell counts and an incremental increase in viraemia (Bashirova *et al*., 2011).

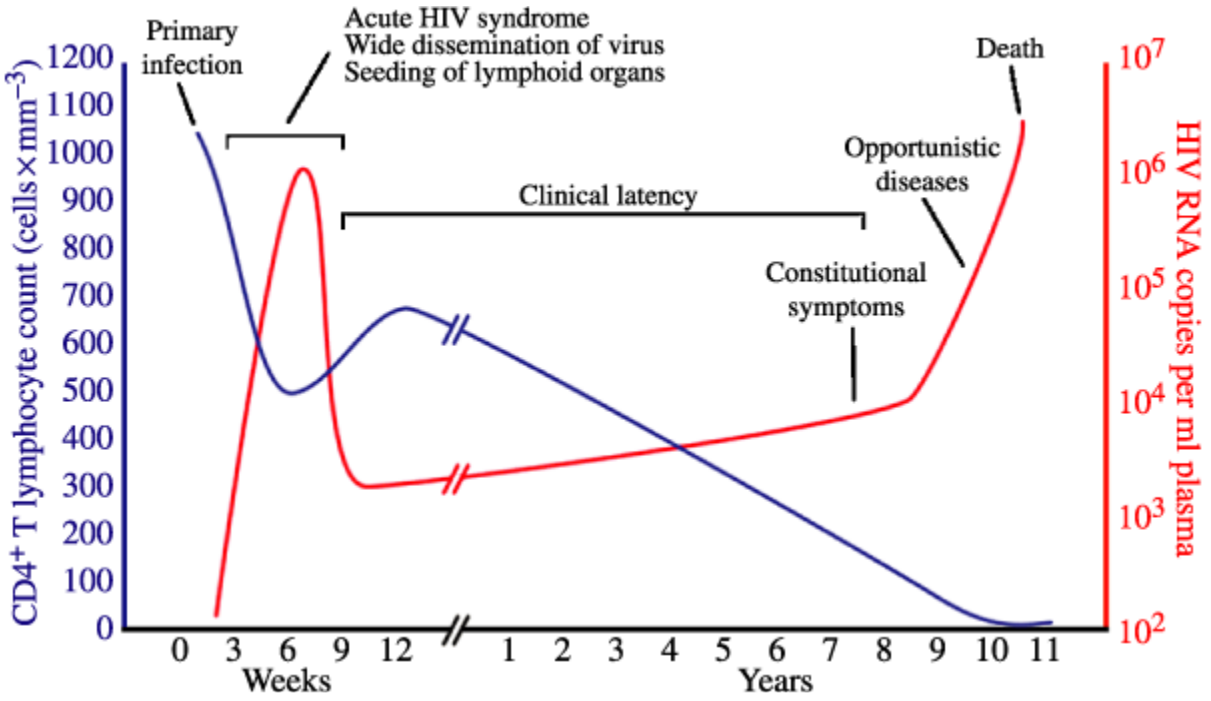


Figure 1.2. Schematic showing the typical course of HIV-1 infection. Reproduced from Kallings (2008).

The final stage, AIDS stage, is characterised by progressive loss of the $CD4^+$ T-cells (< 200 cells/ml of blood), severe impairment of immune function and high levels of viraemia - thereby consequently resulting in opportunistic infections (such as Candidiasis, cytomegalovirus and herpes simplex virus), neurological complications (AIDS dementia complex) and neoplasms (Kaposi‟s sarcoma), that would rarely occur in persons with intact immune function (Hutchinson, 2001). It has been reported that the period from initial infection to clinical AIDS in adults is approximately 8 to 11 years (Porter *et al.*, 1999), although about 10% of persons will rapidly progress to AIDS in 2 to 3 years following HIV infection (Klatt *et al.*, 2013). The numbers of $CD4^+$ T-cells continue to decrease, with anaemia and marked lymphopenia frequently detected during the AIDS phase.

While HIV-1 infection of $CD4^+$ T-cells is likely the primary cause of $CD4^+$ T-cell death during acute phase of HIV-1 infection, the relatively low frequency of HIV or simian immunodeficiency virus (SIV) infected $CD4^+$ T-cells (0.01-1%) during chronic infection appears insufficient to account for the observed ongoing $CD4^+$ T-cell depletion (Chun *et al.*, 1997; Mattapallil *et al.*, 2005). During the acute and chronic phases of infection, the main

cause of $CD4^+$ T-cell death appears to be cytolysis, apoptosis (Mattapallil *et al*., 2005) and pyroptosis (Doitsch *et al*., 2014) induced by HIV-1 infection, while the HIV-specific immune response also contributes by killinginfected $CD4^+$ T-cells (Walker *et al*., 1987).

Following the establishment of infection, HIV spreads locally at first, then quickly across the body to other lymphoid tissues. HIV-1 infects and substantially depletes $CD4^+$ T-cells in a variety of tissues during acute infection, especially in the gastrointestinal tract (GIT) and lymph nodes (LNs) (Schacker *et al*., 2001; Brenchley *et al*., 2004; Mattapallil *et al*., 2005). This massive infection also induces the homing of $CD4^+$ T-cells to lymphoid tissue (Chen *et al*., 2002), explaining the depletion of $CD4^+$ T-cells from peripheral blood. In addition, HIV-1 infection gradually reduces the immune system's capacity to compensate for the constant death of $CD4^+$ T-cells. One study of acute SIV infection around 80% of memory $CD4^+$ T-cells were lost initially, with up to 60% of memory $CD4^+$ T-cells infected at the peak of infection (Mattapallil *et al*., 2005).As a consequence, over time the continual death of $CD4^+$ T-cells overcomes the homeostatic capabilities of the immune system resulting in the gradual depletion of $CD4^+$ T-cells.

The depletion of $CD4^+$ T-cells has been shown to be due to ongoing and persistent T-cell activation even in virally suppressed HIV-1 infected individuals on ARVs (Bofill *et al*., 1995; Liu *et al*., 1996; Giorgi *et al*., 1999). Due to continual antigenic stimulation, resting T-cells are activated (Matsuyama *et al*., 1991; Swingler *et al*., 1999), undergo proliferation (Hellerstein *et al*., 1999; Deeks *et al*., 2002) and apoptosis (Gougeon *et al*., 1996; Meyaard *et al*., 1992; Ameisen *et al*., 1991), resulting in naïve and memory $CD4^+$ T-cells being continually lost (Grossman *et al*., 2002). The mechanism of this chronic systemic immune activation is driven by circulating microbial products derived from the GIT (Brenchley *et al*., 2006).

Immune exhaustion is characterised by the expression of programmed death 1 (PD-1) inhibitory receptor on effector $CD8^+$ T-cells in HIV chronic infection, rendering $CD8^+$ T-cells unresponsive leading to disease progression (Day *et al*., 2006). There is a close association between immune activation and exhaustion as expression of PD-1 is also found on activated $CD4^+$ and $CD8^+$ T-cells, B-cells and myeloid cells (Agata *et al*., 1996). Immune activation and exhaustion have also been related to the loss of regulatory T-cells (Sachdeva *et al*., 2010).

1.5. Natural control of HIV-1 infection

As already mentioned, a number of factors including viral, environmental and the host, influence HIV-1 disease outcome and thus infected individuals tend to exhibit differential disease progression. In the absence of ART, the majority of HIV-1 infected individuals progress to AIDS within 10 years of infection (Lyles *et al*., 2000). However, there exists a small group of HIV-infected individuals who can spontaneously control their VLs and/or preserve their $CD4^+$ T-cells for extended periods in the absence of ART - broadly termed HIV controllers (shown in Figure 1.3). This group may represent a natural model for a 'functional cure' whereby infected individuals control viral replication and remain asymptomatic for a long period of time in the absence of ART. Different models of natural control of HIV-1 have been previously described (Al-Jabri, 2007, Mens *et al*., 2010). Definitions employed for individuals at both ends of the disease progression spectrum tend to vary in different studies.

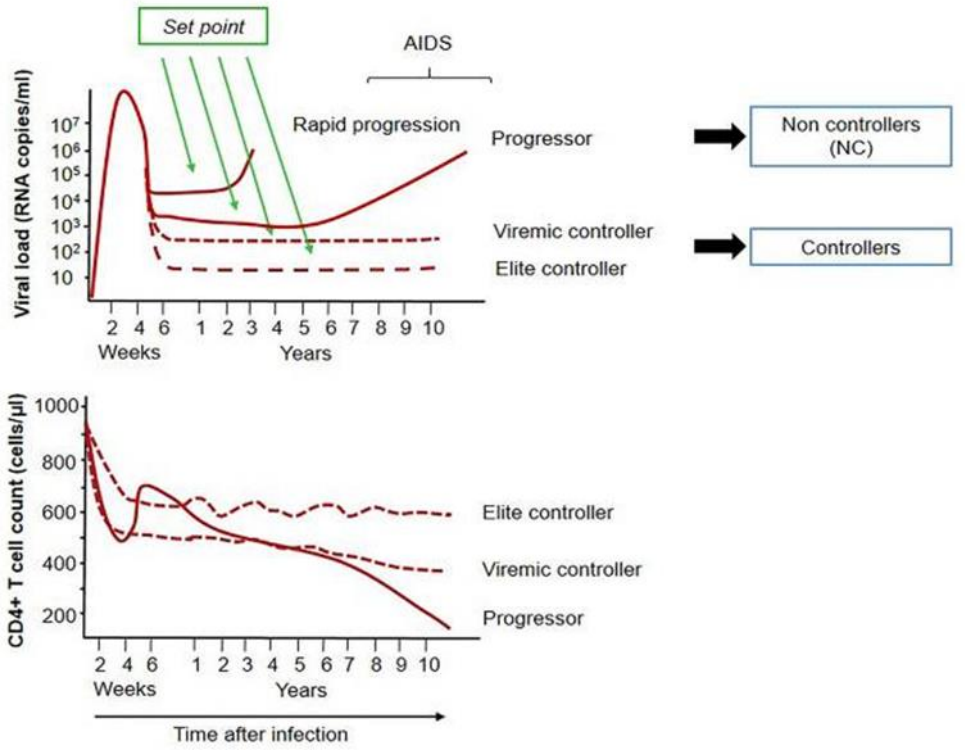


Figure 1.3. Schematic representation of different rates of disease progression to AIDS in various individuals. Reproduced from Gonzalo-Gil *et al.* (2017).

Elite controllers (ECs) are individuals who effectively suppress the virus to <50 RNA copies/ml and maintain high $CD4^+$ T-cell counts, generally >500 cells/µl (Okulicz and Lambotte, 2011; Deeks and Walker, 2007), in the absence of ART, and are thought to occur in approximately 1

in 300 people (Walker, 2007). While ECs have undetectable viraemia by standard clinical assays, almost all of them have detectable levels of plasma viraemia when tested using more sensitive RNA detection assays in plasma, and viral blips of >50 RNA copies/ml are often seen (Pereyra *et al*., 2009; Hatano *et al*., 2010 and Pereyra *et al*., 2008). The study of ECs could provide clues into the immunological pathways important for controlling HIV-1. Other distinct groups such as long term non-progressors (LTNPs), are individuals who can maintain a stable and normal $CD4^+$ T-cell count (generally >500 cells/µl) for ≥7 years, in the presence of detectable viraemia (Gurdasani *et al*., 2014), and viraemic controllers (VCs) tend to be defined by low VL set-points (generally VL>50<2000 RNA copies/ml) (Okulicz *et al*., 2009). HIV-specific $CD4^+$ T-cell responses have been found to display higher functionally avidity in ECs compared to progressors (Vingert *et al*., 2010).

The secretion of cytokine IL-21 by HIV-specific $CD4^+$ T-cells is associated with HIV-1 control and enhances the capacity of HIV-specific CTLs to suppress viral replication (Porichis and Kaufmann, 2012). Broader poly-functional $CD8^+$ T-cell responses, which are characterized by a high frequency of cells capable of secreting multiple cytokines simultaneously, have been associated with lower viral loads and slow disease progression (Borrow *et al*., 1994). Furthermore, ECs tend to have human leukocyte antigen (HLA) alleles that are significantly associated with viral control (Fellay *et al*., 2007; Pereyra *et al*., 2008; Pereyra *et al*., 2010; Gonzalo-Gil *et al*., 2017). Defective HIV strains do not explain the clinical status of most ECs or HIV controllers -these individuals control HIV-1 infection through immune or other unknown mechanisms. On the opposite end of the spectrum are individuals who progress to AIDS within 3 years of viral infection and are termed rapid progressors (de Medeiros *et al*., 2016). Most individuals however fall within the category of normal progressors. Therefore, studying individuals with variable HIV-control, from ECs to rapid progressors may facilitate in the understanding of factors associated with disease control and lack of control and may thus inform future vaccine, pharmaceutical and cure strategies.

1.6. Simian immunodeficiency viruses and HIV-1 disease progression

Simian immunodeficiency viruses (SIVs) are a group of viruses with a wide genetic range that infect a majority of African nonhuman primates (NHPs) - about 40 species of NHPs have been reported (Ansari and Silvertri, 2014). African NHPs such as African green monkeys (AGMs) and sooty mangabeys (SM) are natural host of SIV and despite high viral loads, infection in these primates is usually non-pathogenic (Raehtz *et al*., 2016). The lack of pathogenicity in

AGMs and SMs is thought to be a result of a million years of coevolution between the virus and the host (Keele *et al*., 2009).

High viral loads and rapid turnover of infected cells accompanied by the lack of disease progression in select HIV-infected humans and SIV-infected rhesus macaques (RMs) as well as SIV-infected SMs and AGMs, suggests that this does not result from viral control or cytopathogenicity (Rey-Cuille *et al*., 1998; Pandrea *et al*., 2008), but can be attributed to limited infection and/or loss of specific $CD4^+$ T-cell types, such as T-helper 17 (Th17) cells and T follicular helper (Tfh) cells in lymphoid tissue and central memory (Tcm) and stem-cell memory (Tscm) subsets (Paiardini *et al*., 2011; Cartwright *et al*., 2014; Brenchley *et al*., 2008). SIV-infected natural hosts lack key features of pathogenic infection, namely lymph node inflammation, fibrosis, progressive $CD4^+$ T-cell loss and immune activation (Lackner *et al*., 2012). The exact mechanisms by which natural hosts are able to avoid progression to AIDS in the presence of high viraemia is unknown and still need to be studied extensively.

1.7. HIV-1 treatment and cure

There is a marked decline in the morbidity and mortality associated with HIV infection for individuals with access to therapy as a result of the use of combination antiretroviral therapy (cART) (Lohse *et al*., 2007). Numerous ART-treated HIV-infected people gain almost complete viral suppression and significant reconstitution of $CD4^+$ T-cell counts. The introduction of cART with different modes of action (Figure 1.4) significantly decreased the mortality associated with HIV-1 infection and improved the life expectancy of patients who commence on treatment early, with life expectancy now approaching that of the general population (Laila *et al*., 2019). Previously, clinicians had to weigh the benefits of treatment outcome against its risks such as drug interaction, resistance, toxicity, inconvenience of lifelong treatment and costs when choosing to initiate patients on ART. In 2016, universal test and treat (UTT) was embraced in South Africa and other southern African countries (Meintjes *et al.*, 2015), where ART initiation was available for all individuals who tested HIV-positive, irrespective of their $CD4^+$ T-cell counts. However, despite this achievement, ART is not curative.

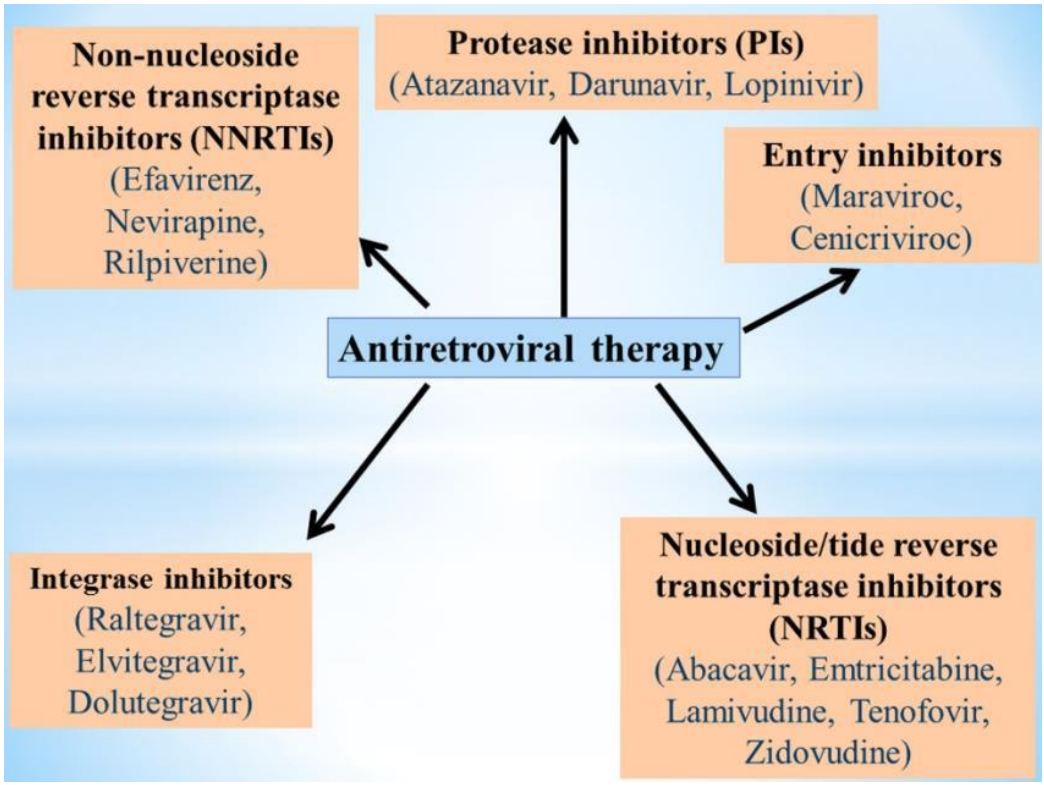


Figure 1.4. Different antiretroviral regimens and their mode of actions. Reproduced from Laila *et al*. (2019).

Interruption of therapy results in rapid viral rebound in most of HIV-1 infected persons due to replication-competent virus commonly refered to as the latent viral reservoir that resides in the long-lived infected cell populations (Chun *et al.*, 1997). Similarly, long-term treated patients on an effective regimen still experience risk of non-AIDS-related complications, such as neurocognitive, kidney, liver and heart diseases and persistent immune dysfunction (Volberding and Deeks, 2010). Although ART is life-saving it does not fully restore normal immune function and is not cost-effective for both HIV-infected individuals and the government. Thus, given all these considerations of life-long ART, an effective vaccine and/or cure for HIV-1 should remain urgent necessities.

To achieve either, it is important to characterize the drivers of immune activation and HIV-1 pathogenesis. As the HIV-1 cure research field has grown over the last few years, it has become necessary to revisit and define various concepts of the term ‘cure’ (Davenport *et al.*, 2019). Within this context, the concept of a ‘functional cure’ has emerged and is currently referred to

as ‘relapse-free remission’: the constant viral suppression without the need for ART. However, ‘functional cure’ does not result in complete eradication of HIV-1 in the body (Deeks, 2012).

Thus far the cure of the “Berlin patient”, together with the post-treatment control in a relatively large proportion adult individuals in the Visconti cohort and the SPARTAC trial (Sáez-Cirión *et al*., 2013; SPARTAC, 2013), the Mississippi child case of remission for a period of 27 months off ART prior to viral rebound (Persaud *et al*., 2013; Luzuriaga *et al*., 2015), and durable post-treatment control in the early-treated French (Frange *et al*., 2016) and South African child cases (Violari *et al*., 2019), has led to enthusiasm for clinical research that aims to achieve HIV-1 remission without the continued need for ART. However, total eradication of HIV-1 remains a long term and extremely challenging goal. Stem cell transplantation with CCR5-negative cells has demonstrated that a cure for HIV infection is possible. The ‘Berlin patient’ (Hütter *et al.,* 2009) and ‘London patient’ both attained HIV-1 remission (Gupta *et al*., 2019) post bone marrow transplantation from a CCR5 negative donor. The ‘Berlin patient’ had no detectable VL for more than 12 years and is considered to have experienced a successful sterilising cure (Allers *et al*., 2011), although sadly has succumbed to his cancer and passed away. A ‘sterilising cure’ is achieved when the replication-competent HIV provirus is completely eliminated from the body. The presence of a latent reservoir is the main barrier to attaining sterilizing cure (Finzi *et al*., 1997).

1.8. The influence of host genetic variation on HIV-1 infection

Host genetic variability plays an important role in determining susceptibility or resistance to pathogenic viral infections, yet the exact mechanisms underlying different levels of resistance and viral control in HIV-1 infected individuals are relatively unknown. According to a review by Chatterjee (2010), the host genes alleged to influence the rate of progression from HIV infection to AIDS are (i) genes that encode cell-surface receptors or affect their competence or serve as ligands, directly participating in HIV-1 entry into the cell and therefore influences the infection and the rate of disease progression; (ii) genes within HLA that regulate host immune response to infection and thus are important in inducing effective immune responses against any invading pathogens; and (iii) other cytokine or immune response genes that have been linked to a positive or negative impact on the rate of disease (Han *et al*., 1996).

1.8.1. Chemokine receptors

Chemokine receptors are a family of proteins found on the surface of certain cells. Their name is derived from their ability to induce directed chemotaxis in nearby responsive cells; they are chemotactic cytokines therefore they interact with cytokines known as chemokines (Premack and Schall, 1996). There have been about 18 proteins classified as chemokine receptors in humans based on their specific chemokine preference namely: CCR1 through CCR11, CXCR1 through CXCR7, XCR1 as well as CX_3CR1 (Murphy *et al*., 2000; Pollakis and Paxton, 2012). Each chemokine receptor has a rhodopsin-like seven- transmembrane (7TM) structure [termed G protein-coupled receptors (GPCRs)], that regulate leukocyte trafficking. The chemokine ligands that these receptors bind are small proteins (8– 12 kDa), mostly 70–80 amino acids long and the disulphide bonds between main cysteine residues give it a tertiary structure (Allen *et al*., 2007; Bachelerie *et al*., 2015). The 7TM structure consists of an extracellular N-terminus, an intracellular cytoplasmic C-terminal and alternating extracellular and intracellular loops in a structure illustrated on Figure 1.5. (Samson*et al*., 1996a).

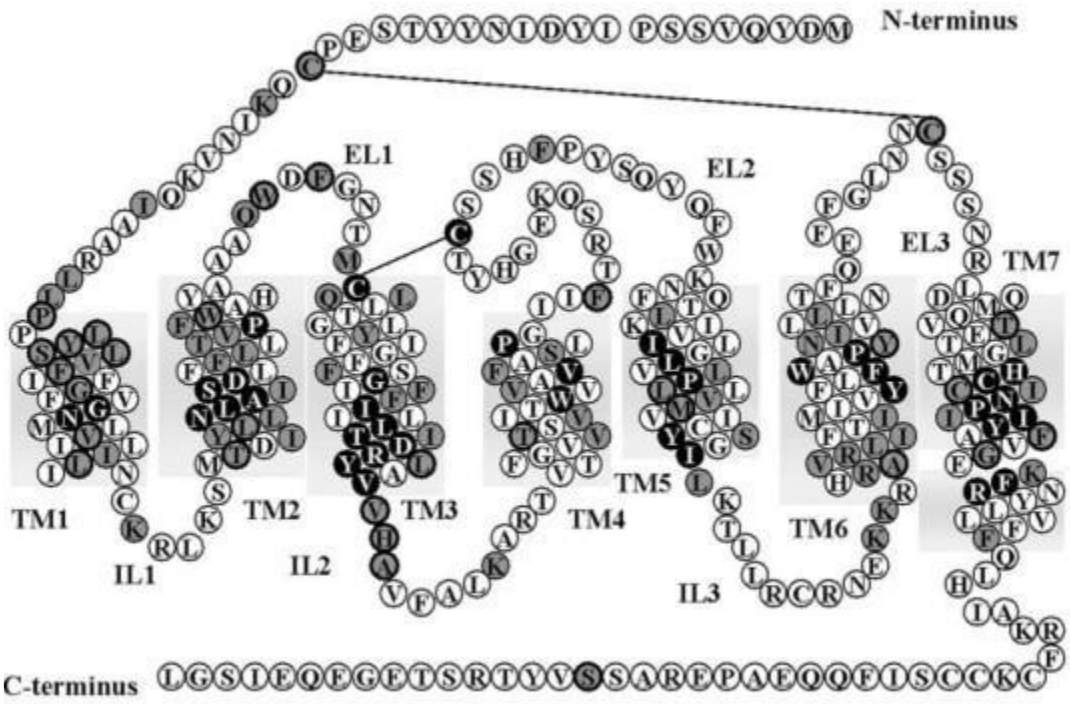


Figure 1.5. Schematic representation of the CCR5 showing amino acid sequence and seven transmembrane structure. Residues in the 7-TM region (TM1-TM7) shown in grey rectangles, spanning α-helices. Extracellular and intracellular loop regions shown as EL and IL, respectively. Disulfide bridges are shown as lines connecting cysteines. Conservation of strong groups in CCR1-CCR5 shown in grey circles while the identical residues in CCR1-CCR5 are shown in grey circles with bold outline and the highly conserved residues in the rhodopsin family of GPCR shown in black circles. Reproduced from Paterlini (2002).

1.8.1.1. CCR5 chemokine receptor

CCR5 is the co-receptor utilized by the R5 strain of HIV-1 to initiate infection (Alkhatib *et al*., 1996). It is a member of the 7TM receptor superfamily of GPCRs and is expressed on various immune cells (including macrophages, monocytes, T-cells and DCs), vascular smooth muscle and fibroblasts, endothelium, epithelium, and neurons, microglia and astrocytes in the central nervous system (Rottman *et al*., 1997). CCR5 plays an important role in the regulation of T-cells as well as the migration, proliferation and immune function of monocytes, especially in control of infiltration at the inflammation site (Eltayeb *et al*., 2007). The CCR5 co-receptor is used in the early phase of HIV-1 infection while the alternative co-receptor, CXCR4, is mainly found on naïve T-cells and is employed by HIV during the later disease stages in the typical

course of HIV/AIDS. CCR5 functions as the specific receptor for the chemokine ligands 3 (CCL3), 4 (CCL4), and 5 (CCL5), whose primary role is in the migration of immune cells to sites of inflammation (Khorramdelazad *et al*., 2013). The role of CCR5 in HIV-1 infection is not restricted to the level of virus entrance, but also to disease progression.

1.8.1.1.1. The *CCR5* gene

The human *CCR5* gene is localized on chromosome 3p21.3 encoding 352 amino acid protein (Liu *et al*., 1996; Samson *et al*., 1996b) and consists of two introns and three exons. Exons 1 and 2 cover most of the 5'-untranslated region (UTR) and exons 2A and 2B are not interrupted by an intron (Mummidi *et al*., 1997). Exons 2A and 2B are derived from exon 2 by alternative splicing and share the same 5' sequence but differ from each other by the use of two distinct donor splice sites 171 bp apart in the gene (D'Adamio *et al*., 1989). Exon 3 contains 11 base pairs of the 5'-UTR and the intron less open reading frame (ORF) (Moriuchi *et al*., 1997; Samson *et al*., 1996a). There are two promoter regions, an upstream promoter (P_U also called P2) which lies upstream of exon 1 and a downstream promoter (P_D also called P1) which lies in the region between intron 1 and intron 2 (Figure 1.6A) (Moriuchi *et al*., 1997; Mummidi *et al*., 1997).

Alternative splicing, as well as multiple transcription start sites, results in the production of several CCR5 mRNA transcripts differing in their 5'-UTR (Mummidi *et al*., 1997). The generation of multiple mRNA transcripts from the two CCR5 promoter sites is thought to regulate *CCR5* gene expression. CCR5 mRNA transcripts generated from the upstream promoter (P2) undergo alternative splicing, giving rise to two transcripts, CCR5A and CCR5B. CCR5A mRNA transcripts contain the untranslated exon 2A while CCR5B transcripts lack exon 2A (Figure 1.6B) (Mummidi *et al*., 1997). Transcripts generated from the downstream promoter (P1) are referred to as "truncated" isoforms (Mummidi *et al*., 2007).

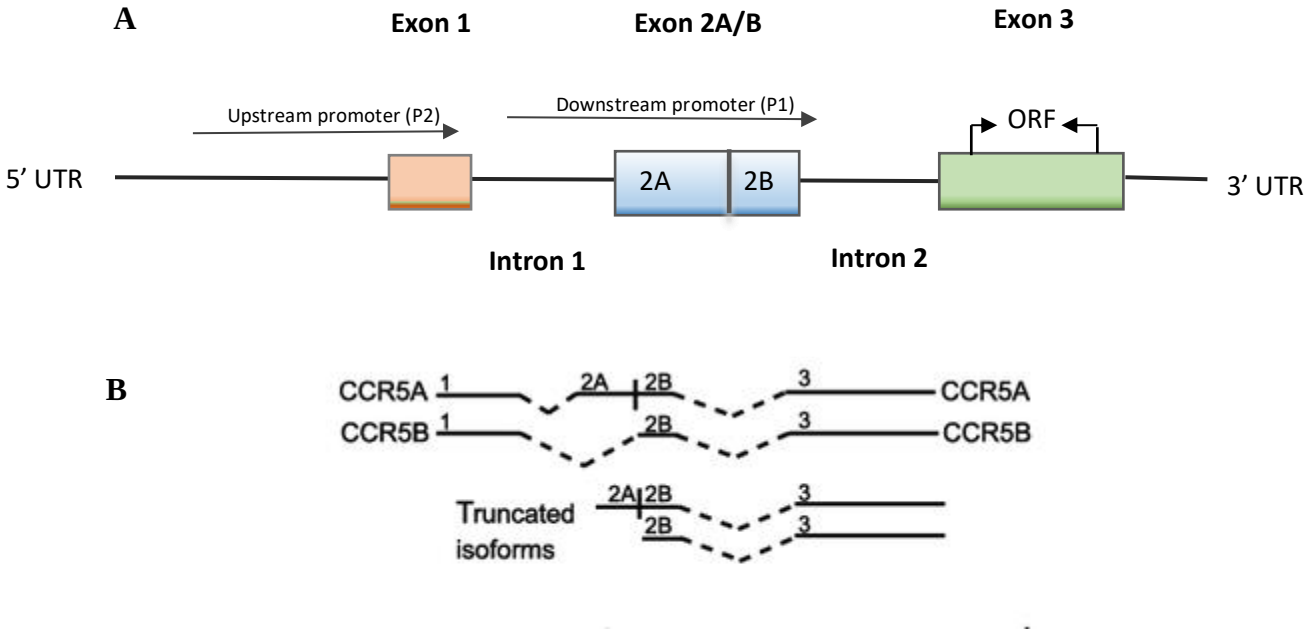


Figure 1.6. Schematic representation of: (A) the *CCR5* gene, *CCR5* consists of three exons (boxes). Exons 1, 2A, 2B, and exon 3 which encodes the ORF and the 3'-UTR. (B) CCR5 mRNA transcripts designated as CCR5A and CCR5B arising from P2 alternative splicing. P1 drives the expression of the transcripts that originate in exons 2A or 2B and this results in the formation of several "truncated" transcripts (Mummidi *et al*., 1997). Modified from Mummidi *et al*. (2007).

1.8.1.1.2. Variation in the *CCR5* gene

A. *CCR5Δ32* polymorphism

Polymorphisms found in the *CCR5* gene have been associated with resistance and susceptibility and to HIV-1 infection and disease progression. Genetic analysis of the ORF of the gene by Lui *et al*. (1996), revealed that *CCR5Δ32* (rs333) is characterized by a 32-base pair deletion in the region encoding the second extracellular loop of *CCR5* (exon 3/ORF region) (Liu *et al.*, 1996) which results in a truncation of the expressed protein (215 amino acid protein vs. 352 amino acid residues in the wild type), preventing the expression of CCR5 on the cell surface (Liu *et al.*, 1996). Individuals homozygous for the *CCR5Δ32* allele are protected against acquisition of R5 tropic HIV-1 (McDermott *et al*., 1998), while individuals heterozygous for the *CCR5Δ32* polymorphism have been associated with a significantly slower disease progression (Mulhenin *et al*., 2003; O'Brien *et al.*, 1997). Population studies report remarkable

ethnic and population differences in the representation of SNPs and other variants within the *CCR5* gene. The *CCR5Δ32* allele has been shown to be present in 10% of Europeans and it is rarely found or absent in Asians or Africans (Hladik *et al*., 2005). In addition to the association with HIV-1 infection, the *CCR5Δ32* allele is associated with various chronic inflammatory diseases (Barcellos *et al.*, 2000; Zernecke *et al.*, 2008; Rautenbach and Williams, 2020).

B. *CCR5* haplotypes

Polymorphisms in CCR5 and CCR2 genes are in strong LD (Carrington *et al*., 1999), performing combined analysis on these genes has assisted in creating extended haplotypes. Haplotypes have been described differently by various groups, based on SNPs/indels within the *CCR5* regulatory region as well as based on linkage with *CCR2*, they have been associated with variation in rates of HIV-1 disease progression (Gonzalez *et al*., 1999; Martin *et al*., 1998). The HHA-HHG*2 *CCR5* haplotype designation which includes the *CCR5Δ32* and *CCR2-V64I* polymorphisms in additionto select promoter variants (Gonzalez *et al.* 1999; Mummidi *et al.* 1998) has been the favourednomenclature in subsequent studies. The *CCR2-V64I* mutation resides in the CC-chemokine receptor 2 (*CCR2)* gene and results in a valine to isoleucine substitution (further discussed in 1.8.1.2).

Gonzalez *et al.* (1999) and Mummidi *et al*. (1998) identified nine evolutionarily distinct haplotype groups, *CCR5* human haplotype (HH) designated as HHA, HHB, HHC, HHD, HHE, HHF (F*1 and F*2) and HHG (G*1 and G*2) shown in Figure 1.7. These haplotypes were defined based on seven SNPs at positions, namely -2733, -2554, -2459, -2135, -2132, - 2086 and -1835 within the *CCR5* 5' regulatory region in combination with the *CCR5Δ32* and *CCR2-V64I* polymorphisms as mentioned above. HHG*1 and HHG*2 are differentiated by the presence of the *CCR5Δ32* mutation in the HHG*2 haplotype, while HHF*1 and HHF*2 are differentiated by the presence of *CCR2-V64I* mutation in HHF*1 (Gonzalez *et al*., 1999; Mummidi *et al*., 2000). The HHA haplotype was identified as the ancestral *CCR5* haplotype as it was shared with other NHPs such as chimpanzees (Mummidi *et al*., 2000). Haplotypes have been associated with varying rates of HIV-1 acquisition and disease progression and, furthermore, they are differently distributed with different disease modifying-effects between race groups (Gonzalez *et al*., 1999; Gonzalez *et al*., 2001).

The HHC haplotype was associated with disease acceleration in African Americans and it was shown to be associated with delayed HIV-1 progression to AIDS in Caucasian and Hispanic

populations (Li *et al*., 2005; Nguyen *et al*., 2004; Gonzalez *et al*., 1999; Martin *et al*., 1998). HHE homozygosity was associated with disease acceleration and rapid progression to death in Caucasians and Thais (although the HHE heterozygosity did not show a similar effect), however HHE homozygosity was not associated with disease modifying-effects in African Americans (Kaur *et al*., 2007). Pairing of HHE with HHC was associated with faster rate of progression to AIDS in African Americans (Gonzalez *et al*., 1999). The HHG*2 and HHF*2 haplotypes containing the *CCR5Δ32* and *CCR2-V64I* mutations, respectively, have been shown to be associated with slower disease progression in Caucasians as well as in African Americans (Gonzalez *et al*., 1999; Kaur and Mehra, 2009 The HHA/HHG*2, HHC/HHG*2 and HHF*2/HHG*2 genotypes in individuals heterozygous for *CCR5Δ32* have been associated with slow HIV-1 disease progression (Catano *et al*., 2011). In contrast the HHE/HHG*2 genotype is associated with accelerated disease progression (Catano *et al*., 2011).

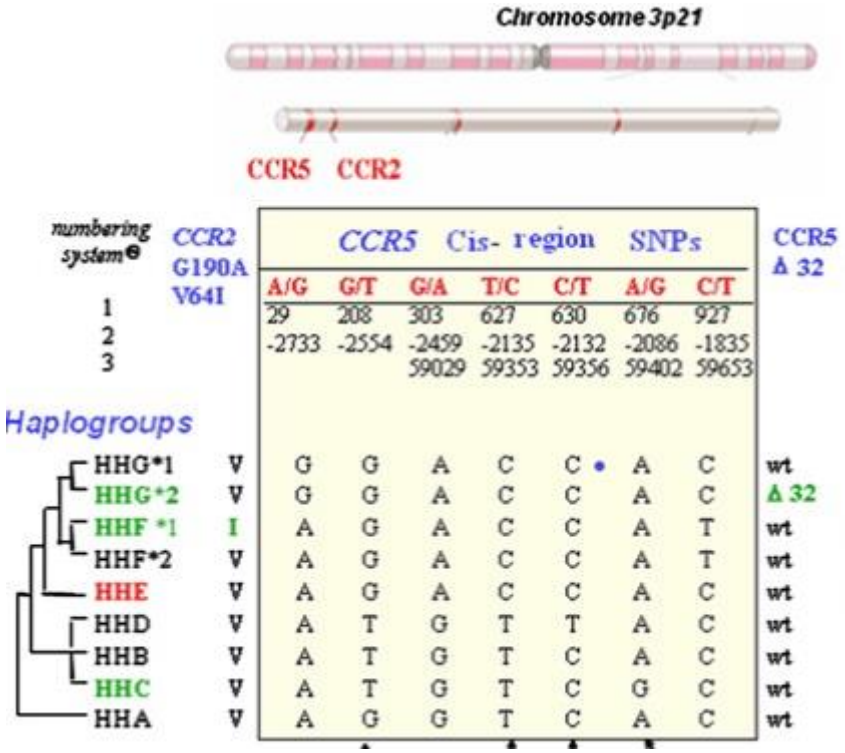


Figure 1.7. Schematic representation of the *CCR5* haplotypes displaying 9 evolutionarily distinct haplotypes involving *CCR5* promoter region SNPs, *CCR2-V64I* and *CCR5Δ32* mutations. Modified from Kaur and Mehra (2009).

Ten haplotypes in the CCR5 promoter region have been identified (*CCR5P1 to CCR5P10*) (Martin *et al.,* 1998). These authors demonstrated that homozygosity for the *CCR5P1* allele was associated with an accelerated progression to AIDS within the first 5 years of infection (Martin *et al.*, 1998).

C. *CCR5* promoter polymorphisms

The *CCR5*-2459G/A (rs1799987) polymorphism, also known as 59029G/A or 303G/A, a mutation found in the promoter region of the *CCR5* gene and part of the haplotypes described above, has been reported to be associated with variations in the clinical course of HIV-1 infection (Kaur *et al.*, 2007; An *et al*., 2000; McDermott *et al*., 1998; Martin *et al*., 1998). The *CCR5*-2459G allele has been associated with a negative effect on the expression of CCR5 in both T-helper cells and $CD14^{+}$ monocytes in individuals showing resistance to HIV-1 infection despite multiple and repeated exposures to the virus (Hladik *et al*., 2005; Salkowitz *et al*., 2003). The *CCR5*-2459G allele exhibits lower promoter activity than *CCR5*-2459A; individuals homozygous for -2459A allele progress to AIDS and death more rapidly compared to individuals homozygous for the -2459G allele (Knudsen *et al*., 2001; McDermott *et al*., 1998; Gong *et al*., 2017).

In the *CCR5*-2135C/T SNP (also a promoter variant part of described haplotypes), the T allele was also associated with a protective effect against vertical transmission of HIV-1 (Pedersen *et al*., 2007). The *CCR5*−2459G and *CCR5*−2135T alleles have been found to be in very strong LD with each other (Gonzalez *et al*., 1999; Clegg *et al*., 2000). Individuals without these genotypes demonstrate slower progression towards AIDS (Clegg *et al*., 2000). Promoter variants can affect transcription factor binding. This was shown by Bream *et al*. (1999) who reported that some *CCR5* variants bind select transcription factors more efficiently than others (Bream *et al*., 1999).

D. Mutations affecting the coding sequence of *CCR5*

In addition to the *CCR5*Δ*32* mutation, other mutations in the *CCR5* coding region have been described. The FS299 variant is found at a 3 to 4% allelic frequency in Asians and this mutation introduces frame-shifts which causes a premature termination of the CCR5 protein in the 7TM domain and complete absence of the C-terminal (Blanpain *et al*., 2000). This is similar to the *CCR5Δ32* variant, where there is failure of CCR5 transportation to the cell surface but still has the ability to bind HIV however with reduced efficiency (Blanpain *et al*., 2000), hence FS299

is associated with slow disease progression. Other mutations such as C20S, A29S and C269F disturb the formation of disulfide bridges, therefore modifyingsurface expression of the receptor and its ability to bind ligands (Carrington *et al*., 1999; Blanpain *et al*., 2000), while other mutations result in very low or undetectable levels of surfacereceptor (i.e. C269F, C101X and G106R).

The C101X (also known as T303A or *CCR5*m303), results from a (T-to-A) substitution at position 303 in the ORF of the *CCR5* gene and introduces a premature stop codon (Quillent *et al*., 1998). This variant leads to the lack of expression of a functional CCR5 protein and abrogation of its HIV-1 co-receptor function. C101X was identified in less than 1 % Caucasians (Quillent *et al*., 1998) and was not detected in the South African Black population (Williamson *et al*., 2000). Individuals with *CCR5*m303/Δ32 genotype are resistant to infection and it is assumed that *CCR5*m303 has similar effects to *CCR5Δ32* in delaying disease progression (Quillent *et al*., 1998).

Unlike *CCR5Δ32,* most of these *CCR5* variants are rare, with frequencies lower than 1–2%, and only existing in specific populations. Hence, their role in HIV-1 disease progression has not been properly established (Arenzana-Seisdedos and Parmentier, 2006; Blanpain *et al*., 2000; Capoulade-Métay *et al*., 2004). The rare C20S, C101X, and T303A alleles are present in Europeans also carrying the more common *CCR5Δ32* allele (Carrington *et al*., 1997; Quillent *et al*., 1998), which suggest that the effects of mutant alleles in these rare mutations may confer resistance to HIV-1 infection *in vitro* when they are in combination with *CCR5Δ32* (Quillent *et al*., 1998).

E. Other mutations found within the *CCR5* gene

Numerous polymorphisms have been reported in the *CCR5* gene (Martin *et al*., 1998; Dean *et al*., 1996; Mummidi *et al*., 1998; Smith *et al*., 1997). One of these is an intergenic SNP (rs1015164G/A) downstream of the *CCL2* gene which affects an antisense long noncoding RNA (lncRNA), termed *CCR5AS* that overlaps with the *CCR5*gene. The rs1015164G/A SNP has been shown to have an effect independent of the *CCR5*Δ*32* mutation (McLaren *et al*., 2015) and to be important in the outcome of HIV-1 infection in individuals of European descent. Recently, Kulkarni *et al*. (2019) demonstrated the deleterious effect of the rs1015164A allele in African Americans, Hispanics and Japanese HIV infected individuals who lack *CCR5*Δ*32*. These authors demonstrated that *CCR5AS* regulates CCR5 expression by interfering with RALY

an RNA-binding protein that binds to the *CCR5* 3'-UTR, hence enhancing CCR5 mRNA stability resulting in higher viral loads and increased susceptibility to HIV-1 infection. This discovery is important because noncoding genetic elements have an effect on the susceptibility and progression of infectious diseases (Ellwanger *et al*., 2018). A 24 base pair deletion, *CCR5*Δ*24*, was detected in Rwanda (Masquelier et al., 2007).

Recently, this *CCR5*Δ*24* genetic variant, similar to *CCR5*Δ*32*, was associated with protection against HIV-1 infection in Rwandans (Arendt et al., 2019), with 99.27% of individuals harbouring two wild-type CCR5 alleles but only 0.73% being heterozygous for *CCR5*Δ*24* while in Kenya the prevalence of *CCR5*Δ*24* was 0.55% in infants (Arendt et al., 2019). The *CCR5*Δ*24* polymorphism affects the CCR5 expression on the cell surface and it may have a similar effect in Africans as that of *CCR5*Δ*32* in Europeans.

CCR5 expression levels in the host are determined by *CCR5* gene and promoter polymorphisms (Samson *et al*., 1996a) as well as epigenetic factors (Gornalusse *et al*., 2015). CCR5 expression levels on cell surfaces impact several different stages of HIV-1 pathogenesis and disease outcomes with low CCR5 surface expression associated with a protective phenotype (Catano *et al*., 2011). The CCR5 expression level variability also can influence the process of envelope glycoprotein (Env) evolution thereby oppressing virus fitness and progression to AIDS (Jakobsen *et al*., 2013). Polymorphisms in the cis-regulatory regions of *CCR5* may explain the differential expression of CCR5 (van Damme, 2002). Despite the fact that the 5′-UTR has been widely studied, remarkably, the genetic variation in the CCR5 3′-UTR (a region increasingly presented to be significant for the regulation of mRNA) is poorly understood (Koor *et al*., 2019).

Single nucleotide polymorphisms in the *CCR5* gene have recently been shown to influence HIV-1 control in a South African cohort study of black South African HIV-1 infected controllers (N=71) and progressors (N=74) in Johannesburg, which showed that two SNPs in the *CCR5* gene, one in the 5'-UTR (-4223C/T, rs553615728) (identified by Gornalusse *et al*., 2015) and the other in the 3'-UTR (+2919T/G, rs746492) that can serve as protective and deleterious markers in the context of HIV-1 control, respectively (Koor *et al.,* 2019). These two SNPs occurred at relatively high allelic frequencies in most study groups, *CCR5* rs746492 minor allele (G): Progressors (45.3%); Controllers (30.8%); ECs (30.4%) and VCs (31.8%) and *CCR5* rs553615728 minor allele (T): Progressors (3.4%); Controllers (4.2%); VCs (6.8%),

with *CCR5* rs553615728-T allele being absent in ECs. The *CCR5* rs746492 and *CCR5* rs553615728 SNPs are the focus of the current study.

The *CCR5* rs553615728 SNP has been associated with the disruption of the cytidine phosphate guanidine (CpG) dinucleotide within the *cis*-region of *CCR5,* known as the CpG-41 (the binding site for DNA methylation). This mutation is exclusively found in individuals from southern Africa and suggested to be implicated in protection from HIV-1 acquisition and in HIV-1 control in black South Africans (Gornalusse *et al*., 2015, Koor *et al*., 2019).

The HHF haplotypes appear to be linked to the *CCR5* rs746492 SNP in black South Africans compared to Caucasians (Picton *et al*., 2010). Koor *et al*. (2019) report the *CCR5* rs746492 SNP in strong LD with the *CCR5* promoter HHE, HHF*1, HHF*2 and HHG*1 haplotypes as well as the two SNPs discussed earlier (-2135C/T and -2459G/A SNPs – termed 5′-UTR-2SNP-hap). These findings raised the question as to which polymorphism is responsible for the deleterious effect on HIV-1 control in black South Africans. Hence, information on *CCR5* genetic variation in HIV-1 infected individuals will provide a better understanding on how *CCR5* host gene expression influences the response to HIV-1 exposure in protection and/or disease acceleration.

1.8.1.2. CCR2 polymorphism (V64I)

The CC-chemokine receptor 2 (CCR2) is one of the few chemokine receptors that are considered as the minor co-receptors of HIV-1 cell entry (Smith *et al*., 1997; Kostrikis *et al*., 1998). The *CCR2* gene is located on the same chromosome as the *CCR5* gene (Samson *et al*., 1996a; Lee, 1998) and it has two isoforms as a result of alternative splicing: CCR2A and CCR2B. A mutation (G to A) resulting in an amino acid change, valine (GTC) to isoleucine (ATC) at codon 64 (*CCR2-V64I*), was found in the first transmembrane domain of CCR2A and CCR2B (Nakayama *et al*., 2004).

The *CCR2-V64I* polymorphism is present in all populations studied thus far, with an allelic frequency between 10 and 25% depending on the ethnic population and is associated with delayed progression to AIDS by approximately 2-4 years when compared to individuals possessing the homozygous wild type (Smith *et al*., 1997; Kaur and Mehra, 2009; Doms and Pieper, 1997; Mulherin *et al*., 2003). *CCR2-V64I* homozygosity has also been associated with reduced risk of acquiring HIV-1 in discordant couples in Thailand (Louisirirotchanakul *et al*., 2002) and was found to delay HIV-1 disease progression, after the child has been infected in the study conducted by Mangano *et al*. (2000). However, the association of chemokine and

chemokine receptor genetic polymorphisms like *CCR2-V64I* with HIV-1 transmission and/or rate of disease progression remains highly controversial as some studies have failed to show this association (Hladik *et al*., 2005; Schinkel *et al*., 1999). In yet another study the *CCR2-V64I* was found to have no effect on mother to child transmission in mothers in Western Kenya (Brouwer *et al*., 2005). Ioannidis *et al*. (2001) reported that HIV-1 infected adults with homozygous wild-type genotypes of both *CCR5* and *CCR2* progressed more rapidly to AIDS. Furthermore, the protective effects of *CCR2-V64I* mutation have been shown to be population specific, although here too results seem to be contradictory (Mummidi *et al*., 1998; Smith *et al*., 1997). Smith *et al*. (1997) reported the association of *CCR2 64I* mutation with delayed disease progression to AIDS in Caucasian, but not African HIV-1 infected individuals, while Mummidi *et al*. (1998) reported the opposite.

1.8.1.3. CXCR6 chemokine receptor

Cellular targeting by HIV or SIV is largely dependent on CD4 receptor and chemokine co-receptor expression levels. HIV-1 gains entry into cells by binding to CD4 and a co-receptor, typically CCR5 or CXCR4 as previously discussed in 1.2. In contrast to HIV-1, a number of SIV strains use CCR5 as a major co-receptor and rarely utilize CXCR4 for entry into host cells, however they have been shown, *in vitro*, to be capable of utilizing a number of human alternative co-receptors, including CCR2B, G protein-coupled receptor 1 (GPR1), GPR15 (BOB), and CC chemokine receptor type 6 (CXCR6) (Bonzo/STRL33/CD186/TYMSTR) (Deng *et al*., 1997; Alkhatib *et al*., 1997; Edinger *et al*., 1998).

CXCR6 is differently expressed by subsets of effector $CD4^+$ memory T-cells, plasma cells, NK cells and NK T-cells (Kim *et al*., 2001; Nakayama *et al*., 2003). A minority of HIV-1 variants have been shown to be capable of infecting cells expressing other co-receptors such as CXCR6 and CCR2B (Blaak *et al*., 2005). Primary isolates of HIV-2, that have been isolated from individuals with undetectable plasma viraemia, have been shown to infect cells using the CXCR6 co-receptor in addition to CCR5 and CXCR4 (Blaak *et al*., 2005). Primary isolates of HIV-1, which are able to use CXCR6 in virus entry have been reported by Zhang *et al*. (2001). As mentioned above, CXCR6 can also function as a co-receptor for SIVs (Liao *et al*., 1997; Alkhatib *et al*., 1997; Deng *et al*., 1997). CXC chemokine ligand 16 (CXCL16), also referred to as scavenger receptor for phosphatidylserine and low density lipoprotein (SR-PSOX), is the only known ligand of CXCR6 (Hofnagel *et al*., 2011) and will be discussed in more detail in section 1.8.2.

While many strains of SIV are thought to use CCR5 as the major entry co-receptor *in vivo*, there have been a fair amount of inconsistencies reported in this area. Firstly, Chen *et al*. (1998) reported that SIVrcm, isolated from red-capped mangabeys (RCMs), which have a high allelic frequency (86.6%) of a CCR5 inactivating deletion (RCM-*CCR5*Δ*24*), utilizes CCR2B rather than CCR5 for entry. Secondly, CCR5-deficient sooty mangabeys are susceptible to SIVsmm as they may use alternative co-receptors for entry if they carry mutant CCR5 alleles (SM-*CCR5*Δ*2*4 and SM-*CCR5*Δ*24*) that inhibit CCR5 surface expression (Chen *et al*., 1998; Riddick *et al*., 2010). Riddick *et al*. (2010) reported that about 50% SIVsmm infected animals carried homozygous *CCR5* mutant alleles and showed robust viral replication which was only slightly lower than replication in animals with WT *CCR5* alleles. CXCR6 was found to be the main co-receptor *in vitro* and *ex vivo* in SIVagmSab which is the SIV that infects sabaeus AGMs (Wetzel *et al*., 2017). SIVagmVer which infects vervet AGMs, also can utilize non-CCR5 entry pathways for primary cell infection *ex vivo* and uses CXCR6 efficiently *in vitro* (Riddick *et al*., 2015). These findings suggested that CCR5-independent pathways are a common feature for virus/host interactions in a number of natural hosts especially where the expression of CCR5 is low or genetically absent.

1.8.1.3.1. CXCR6 in HIV-1 infection

Polymorphisms in chemokine-encoding genes such as CXCR6 have been implicated in differential outcomes of HIV-1 infection and disease progression (Duggal *et al*., 2003; Liao *et al*., 1997, Limou *et al*., 2010). In a subset of HIV-infected African-American intravenous drug users, a polymorphism (*CXCR6-E3K,* rs2234355) that results in a non-conservative alteration in the CXCR6 N-terminus was associated with higher rates of survival after infection with *Pneumocystis carinii* (Duggal *et al*., 2003). Another polymorphism located within the CXCR6 gene (3'-UTR), the rs2234358 SNP, was linked to long-term non-progression of HIV-1 but not elite control in HIV-1 infected Caucasians (Limou *et al*., 2010). This association was confirmed in three other European-descent cohorts, making it the first non-HLA repeated association discovered using a GWAS method (Limou *et al*., 2010). Similarly, in a study conducted on a South African black population, the *CXCR6* rs2234358-T allele (minor allele) was significantly lower in VCs compared to both progressors as well as healthy uninfected controls, while minor allele homozygosity (TT) was also significantly lower in VCs compared to both progressors and healthy uninfected controls (Picton *et al*., 2017).

The most significant association in this South African study, however, was the combination of heterozygosity for the *CXCR6* rs2234355 SNP in the absence of the 3'-UTR rs2234358 SNP minor allele homozygosity, with VCs having 80% representation compared to 17% representation in progressors ($P<1x10^{-7}$) (Picton *et al*., 2017). Hence a better understanding of the role of CXCR6 co-receptor in viraemic HIV-1 control could inform vaccine development and novel therapeutic modalities.

1.8.2. Chemokines

Chemokines are small low-molecular-weight proteins that serve as ligands, activating and signaling through specific receptors to regulate a variety of cellular functions such as growth, immune response, leukocyte trafficking, and angiogenesis (Viola and Luster, 2008). They are also chemotactic cytokines as they induce directed chemotaxis of certain cells that are in close proximity. There are numerous sub-families of chemokines (i.e. α, β, γ and δ refers to CXC, CC, C and CX_3C respectively, where X is any amino acid) (Murphy *et al*., 2000) and well-defined by the positions of consecutively conserved NH_2-terminal cysteine residues (Figure 1.8) (Zlotnik and Yoshie, 2000).

Chemokines and their receptors play an important role in a number of physiological and pathological processes such as inflammation, HIV-1 infection as well as cancer (Szpakowska *et al*., 2012). Chemokines tend to be upregulated and the expression of their receptors altered in both acute and chronic HIV-1 infection (Wang *et al*., 2017). According to Angela-Covino *et al*. (2016), chemokines can either hinder or improve HIV replication, impacting both entry and post-entry occasions of the viral life cycle. Chemokines contribute to the variability in HIV pathogenesis, depending on the balance of their negative versus positive outcomes on HIV replication and spreading.

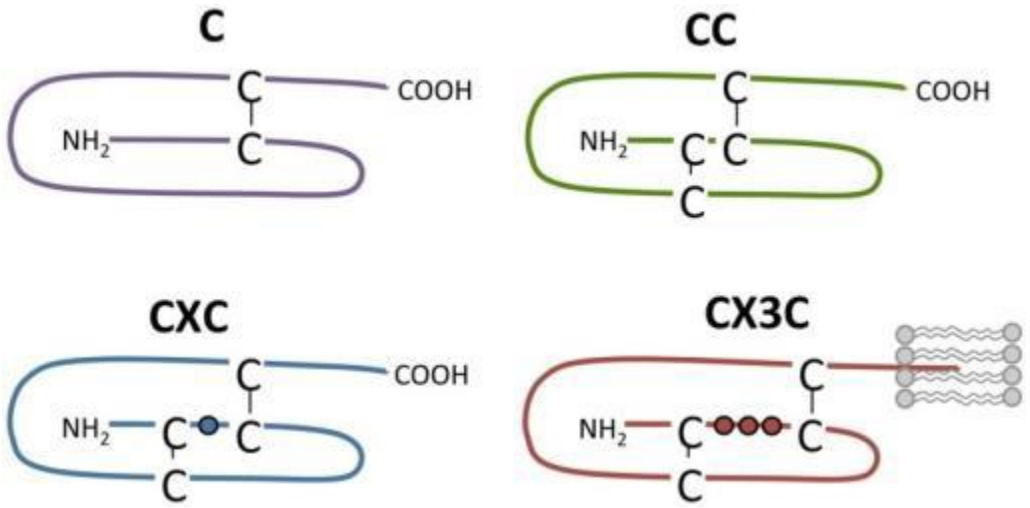


Figure 1.8. Schematic showing chemokine subclasses as per number and spatial organization of conserved cysteine residues in their N-terminus, black lines denote disulphide bridges. Reproduced from de Munnik *et al*. (2015).

1.8.2.1. CC chemokines

The CC chemokines (or β-chemokines) have two adjacent cysteines located close to the amino terminus (Broder and Collman, 1997; Horuk, 1999). The subfamily of the CC chemokines consists of 28 CC chemokine ligands (*CCL*) and they range from *CCL1-CCL28*. The genes encoding this cluster of CC chemokines are mainly located on chromosome 17 in the 17q11.2-q12 region. Findings over the past two decades have identified a close relationship between chemokines and HIV infection. This primary association was established by Cocchi *et al*. (1995) who found that a subset of these factors, namely the CC chemokines macrophage inflammatory protein-1α (MIP-1α)/*CCL3*, macrophage inflammatory protein-1β (MIP-1β)/*CCL4* and regulated upon activation, normal T-cell expressed and secreted (RANTES)/*CCL5*, the natural ligands of CCR5, could function as inhibitors of HIV *in vitro*.

1.8.2.2. Copy number variation in CC chemokine ligands

The human genome includes several insertions and deletions (indels) as well as duplications. These duplications can range from only a couple of base pairs long to whole chromosomal segments including whole gene clusters, e.g. the number of copies of a gene is replicated a variable number of times and the number of replications varies between individuals (Sharp *et al*., 2005). This polymorphism is termed copy number variation (CNV). The cluster of pro-inflammatory CC chemokines contains 16 genes at 17q11.2–q12. Four of these genes include

the two firmly related, paralogous sets *CCL3* -CC chemokine ligand 3-like (*CCL3L*) and *CCL4* -CC chemokine ligand 4-like (*CCL4L*) (Modi, 2004). At both the genomic and amino acid levels, members of each pair share 95% sequence identity. A particular characteristic of *CCL3L* and *CCL4L* is that they are present in variable copy numbers in the human genome. Townson *et al*. (2002) reported on the extent of *CCL3L/CCL4L* CNV in the Caucasian population. Although this association has been inconsistent, CNVs of the CC chemokine ligand 3-like1 gene (*CCL3L1*) have been implicated in HIV-1 susceptibility (Shrestha *et al*., 2010).

Copy numbers of these two chemokine genes range from 0-14 (Gonzalez *et al*., 2005) for *CCL3L* and 0-10 for *CCL4L* (Colobran *et al*., 2008). *CCL3L/CCL4L* copy number ranges also vary between different population groups. Individuals from African descent were found to have higher copy numbers with a median of 4-5 *CCL3L* copies (Gonzalez *et al*., 2005; Meddows-Taylor *et al*., 2006) compared to Caucasian individuals with a median of 2 copies (Colobran *et al*., 2010). Non-human primates, chimpanzees, AGMs and RMs have been found to have even higher copy numbers compared to humans with a median of 9 *CCL3L* copies (Gonzalez *et al*., 2005; Gornalusse *et al.,* 2009; Lim *et al.,* 2010). Elevated *CCL3L* and *CCL4L1* copy numbers were associated with resistance to HIV-1 infection and slower rate of disease progression while only individuals with *CCL3L/CCL4L* copy numbers lower than the median of a population-specific average are at the high risk of acquiring HIV (Gonzalez *et al*., 2005).
Studies have investigated the importance of these chemokines in the context of HIV-1 infection and have found that the production of both *CCL3* and *CCL4* by $CD8^+$ T-cells was significantly higher in asymptomatic HIV-1 infected individuals compared to uninfected individuals and infected individuals who have reached the AIDS phase of the infection, signifying that chemokine ligands play an important role in HIV-1 disease outcome (Cocchi *et al*., 2000). Associations between *CCL3L* and *CCL4L* CNV on HIV-1 disease progression are controversial with respect to the effect of *CCL3L/CCL4L* gene copy numbers on HIV-1 infection and disease progression (Shao *et al*., 2007; Bhattacharya *et al*., 2009; Urban *et al*., 2009). According to Shrestha *et al*. (2010) this controversy might be partly attributed to difference in methodologies used to quantify gene copy numbers.

1.8.2.3. CXC chemokines ligands

The CXC chemokines (α-chemokines) are a 16-member subfamily encoded by genes located mainly on the 4q12-21 chromosome region. A few individuals in this family have the tripeptide, glutamic acid-leucine-arginine (Glu-Leu-Arg or ELR sequence) situated close to the

N-terminus (Laing and Secombes, 2004). ELR-positive CXC chemokines (ELR^+), such as CXCL1, CXCL2, CXCL3, CXCL5, CXCL6, CXCL7, and CXCL8, are potent chemoattractants for neutrophils and promoters of angiogenesis, whereas ELR-negative CXC chemokines (ELR^-), such as CXCL4, CXCL9, and CXCL10, are strong inhibitors of angiogenesis (Keeley *et al*., 2011). The fundamental role of the ELR^+ CXC chemokine sub-group is to promote the adherence of neutrophils to endothelial cells and surfaces towards inflammatory sites (Laing and Secombes, 2004).

HIV-1 can cause increased cytokine and chemokine secretion through the innate immune response during the early stages of infection, promoting immune activation and resulting in a "cytokine storm". Immune activation was shown to arise between 1 and 3 days of hyper acute HIV-1 infection, while the “cytokine storm” is observed before peak viraemia (Stacey *et al*., 2009; Ndhlovu *et al*., 2015). Some chemokines are believed to block HIV-1 entry into T-cells and macrophages, as well as impact HIV-1 disease through their capacity to advance inflammation (Landrø *et al*., 2009; Wang *et al*., 2017).

CXCL16, is a member of the CXC chemokine family that is largely expressed on endothelial cells and it exists in both membrane-anchored and soluble forms mediating firm cell adhesion and chemotaxis (Matloubian *et al*., 2000; Wilbanks *et al*., 2001; Hofnagel *et al*., 2011). CXCL16 promotes the movement of activated T-cells into inflamed tissue and has been linked to the development of inflammatory diseases (Wilbanks *et al*., 2001; Wuttge *et al*., 2004). *In vitro* experiments in the study conducted by Landrø *et al*. (2009) suggest that soluble CXCL16 is not only a marker of inflammation, but also a pathogenic mediator in HIV-1 infected patients, enhancing inflammation and HIV-1 replication in those with elevated VL, irrespective of continuing ART. The authors also reported that increased CXCL16 serum levels are significantly correlated with disease progression in HIV-infected patients, and CXCL16 may reflect an interplay between inflammation and HIV-1 replication in untreated patients, leading to disease progression in HIV-1 infected individuals (Landrø *et al*., 2009).

1.8.3. Variation in human leukocyte antigen (HLA) molecules

The MHC is responsible for graft rejection as well as eliciting the strong responses to allogeneic tissues (Ingulli, 2010). The MHC genes, which are located in the short arm of chromosome 6, encode the human leukocyte antigens (HLAs) expressed on the cell surface. The MHC gene complex is the most polymorphic region of the human genome (Figure 1.9A)

(Ayala-García *et al*., 2012). There are two noteworthy classes of MHC molecules with different structure as well as function within the immune system, namely: (i) HLA class I molecules, expressed on all nucleated cells, comprised of a membrane-spanning α chain encoded by the MHC genes on chromosome 6, and a β chain encoded by the β2-microglobulin gene located on chromosome 15 (Figure 1.9C); and (ii) HLA class II molecules which are comprised of two membrane-spanning chains, α and β, of similar size and both encoded by the MHC genes (Figure 1.9B). HLA class II molecules are expressed on B-cells, activated T-cells, and antigen-presenting cells (APC) (Ingulli, 2010). HLA alleles have been associated with more than 100 diseases including many autoimmune and infectious diseases, such as malaria and influenza (Matzaraki *et al*., 2017), as well as an association with different outcomes of HIV-1 infection.

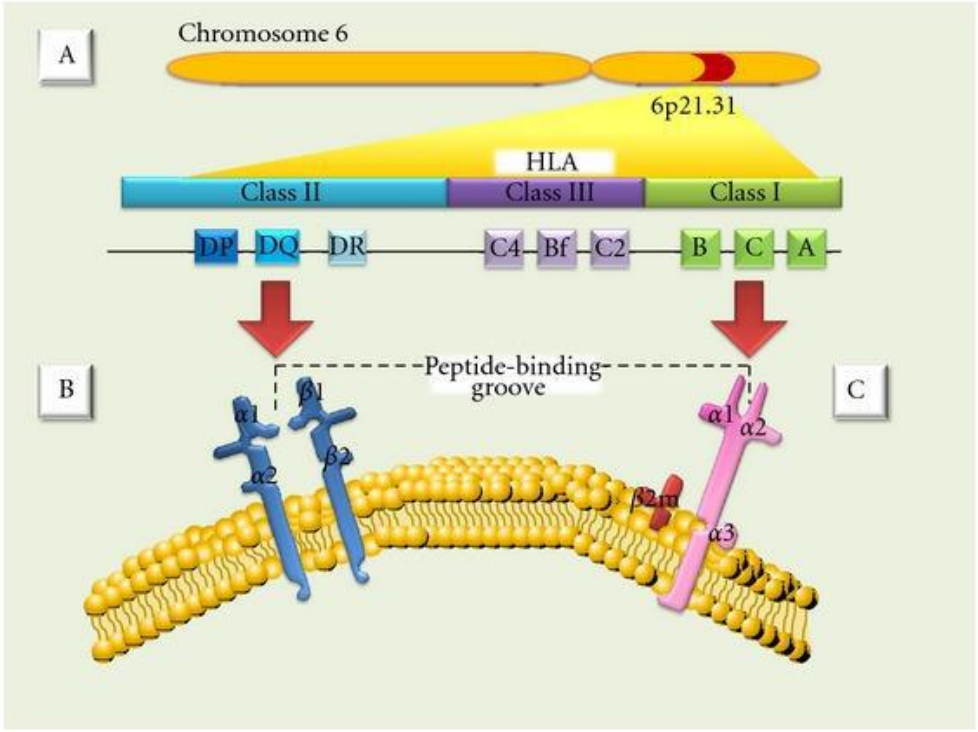


Figure 1.9. Schematic showing the HLA complex genes and their protein product classes (I, II, and III). Reproduced from Ayala-García *et al*. (2012).

Strong association between select HLA class I alleles and HIV-1 disease outcome have consistently been reported in the literature and in addition has been verified in GWAS (Pereyra *et al*., 2009; Fellay *et al*., 2007). The protective effects of certain HLA class I alleles are

associated with enhanced $CD8^+$ T-cell mediated immunity (Kawashima *et al*., 2009). HLA class I presentation of HIV-1 epitopes derived from conserved regions of the HIV-1 genome to $CD8^+$ T-cells can pressurize the virus to mutate resulting in immune viral escape which can be associated with a cost to viral fitness (Altfeld and Allen, 2006).

A heterozygous HLA class I genotype has been associated with lowering HIV-1 VL by allowing the presentation of a larger range of peptides and reducing the risk of viral escape variants resulting in slower disease progression (Piacentini *et al*., 2009). In addition, various HLA alleles have been identified to be either beneficial or deleterious in the context of HIV-1 infection. According to Kiepiela *et al.* (2004), *HLA-B* 58:02* is deleterious whereas *HLA-B* 58:01* is associated with protection against HIV infection in a black South African cohort. The *HLA-B*57* and *HLA-B*27* alleles are mostly enriched in HIV ECs (Shasha and Walker, 2013) and have the strongest influence on VLS and disease progression (Kaslow *et al*., 1996; Goulder *et al*., 1996; Kiepiela *et al*., 2004; Thomas *et al*., 2009).

The most prevalent *HLA-B*57* allele in the sub-Saharan African population is *HLA-B*57:03* and is expressed in approximately 4–5% people (Kløverpis *et al.*, 2016). There is one amino acid difference between *HLA-B*57:03* and *HLA-B*57:02*, where the latter is the less common of the *HLA-B*57* alleles (Kløverpis *et al.*, 2016). The *HLA-B*57* alleles are also closely related to *HLA-B*58:01* sharing many similarities and present many of the same epitopes and protect against HIV-1 progression (Kløverpris *et al*., 2016). Studies performed in countries heavily impacted by the HIV-1 outbreak, such as Botswana and South Africa, indicate that increased use of ART, as well as viral adaptation to protective HLA alleles (i.e. *HLA-B*57* and *HLA-B*58:01*), both influence the reduced virulence of transmitted viruses (Payne *et al*., 2014).

In Caucasian HIV-1 infected individuals, the HLA class I alleles, *HLA-B*35* and *HLA-Cw*04* (Carrington *et al*., 1999), as well as *HLA-Cw*07* (Pereyra *et al.*, 2010) have been associated with rapid progression to AIDS. In Sub-Saharan Africa the deleterious alleles *HLA-B*18* and *HLA-B*5802* were associated with high viral loads in HIV-1 infected Zulu/Xhosa individuals (Kiepiela *et al*., 2004). The strong protective effect of select host HLA class I alleles on disease progression has also been associated with receptor families such killer immunoglobulin-like receptors (KIRs) recognizing specific HLA class I. (Borges and Cosman, 2000).

1.8.4. Variation in killer immunoglobulin-like receptor (KIR) molecules and HIV-1 control

The Killer immunoglobulin-like receptors (KIRs) genes are found within the leukocyte receptor complex (LRC) on chromosome 19 (19q13.4) shown in Figure 1.10, (Carrington and Norman, 2003). KIRs are found on NK cells and effector memory T cells (van Bergen and Koning, 2010), whereas leukocyte immunoglobulin-like receptors (LILRs), are expressed on APCs such as DCs, monocytes, macrophages, and B-cells, as well as on $CD4^+$ and $CD8^+$ T-cells (Borges and Cosman, 2000). The main regulators of NK cell recognition of virus-infected cells are thought to be the KIRs, since these are the natural receptors for HLA class I (Rajalingam, 2012).

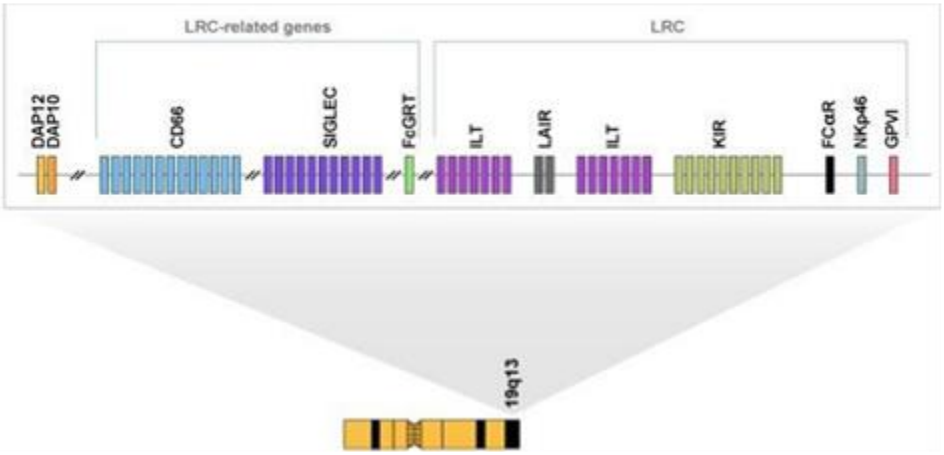


Figure 1.10. Schematic representation of chromosome 19 (19q13.4) region where the *KIR* genes are located. Reproduced from Carrington and Norman (2003).

Polymorphism and structural differences within the 14 *KIR* genes, and differences in expression confer additional differences in the ability of NK cells to identify and respond to virus-infected cells.(Alter *et al.*, 2011; Caligiuri, 2008; Jost and Altfeld, 2012; Lanier, 2008). KIRs display incomplete specificity in their recognition of HLA ligands: *KIR3DL1* and *KIR3DS1* recognise HLA-A and HLA-B molecules with the *Bw4* epitope, *KIR2DL2* recognizes HLA-C of both C1 and C2 groups, *KIR2DL3* recognises HLA-C of the C1 group exclusively and *KIR2DL1* recognises HLA-C molecules of the C2 group exclusively (Moesta and Parham, 2012). The effect of KIR-HLA combinations on clinical outcome of HIV-infected adults has

been reported. HIV-1-infected individuals expressing *KIR3DS1* and *HLA-Bw4* alleles with an isoleucine amino acid in position 80 (i.e. *HLA Bw4-80I*) progress slower to AIDS have lower VL and are protected against opportunistic infections compared to individuals without *HLA-Bw4-80I* (Martin *et al*., 2007; Qi *et al*., 2006; Martin *et al*., 2002).

Different cohort studies conducted in Thailand (Mori *et al*., 2015), Cote d'Ivoire (Jennes *et al*., 2006), and South Africa (Mori *et al*., 2019), showed that HIV-infected individuals with *HLA-C*16:01 KIR2DL3* have worse clinical outcome, rapidly progressing to AIDS compared to those without this genotype. Yet in a prevention of mother-to-child transmission (PMTCT) South African cohort (Paximadis *et al*., 2011) and in intravenous drug users in Vietnam (Ravet *et al*., 2007), the *HLA-C1-KIR2DL3* combination was protective against HIV-1 transmission. The possible explanation put forward for this discrepancy comes from a study examining the impact of KIRs in hepatitis C virus (HCV) infection (Khakoo *et al*., 2004). Those with *HLA-C1-KIR2DL3* combination who received a low dose HCV infection through needle stick transmission were able to resolve the infection compared to those receiving a high dose through blood transfusion. Although a plausible explanation of the protective role of *HLA-C1-KIR2DL3* in the PMTCT (Paximadis *et al*., 2011) and intravenous drug user (Ravet *et al*., 2007) studies could be due to the inoculum dose route of transmission, it is more likely due to the presence of the genotype combination *HLA-C*16:01 KIR2DL3*.

In another PMTCT study exploring HIV-transmission regarding the mother/infant pair, KIR concordance or discordance with respect to a functional *KIR2DS4*, a higher risk of intrapartum transmission was seen if *KIR2DS4* was present in the mother but absent in the infant while a higher intra-uterine transmission occurred if the soluble *KIR2DS4* was present in the infant but absent in the mother (Hong *et al*., 2013). Two studies in highly exposed-uninfected individuals (Vince *et al*., 2014; Naranbhai *et al*., 2017) assessing the risk of HIV acquisition reported no HLA-KIR pair association, although Naranbhai *et al*. (2017) found the *KIR2DL2/KIR2DS2* haplotype associated with HIV control. The identification of HLA-KIR combinations which protect or make an individual susceptible to HIV-infection would be important to delineate in HIV-1 control.

1.9. Recruitment of cohorts of HIV-1 infected individuals with long term follow up

Throughout the HIV epidemic, cohorts of seroconverters have been followed longitudinally to describe virological, clinical and immunological course of the untreated and treated HIV

infection. Other factors associated with HIV-1 infection and transmission such as behavioural and psychosocial factors have also been studied (Gore-Felton and Koopman, 2008). Long-term follow-up of HIV-1 infected individuals in natural history studies sheds light on issues such as long-term non-progression and host genetic factors of resistance, effect of ART on disease progression, transmission during and after seroconversion, and the effects of a positive test on sexual risk behaviour (Etter *et al*., 2013).

Snowball sampling, the use of the index participants or invitation letters and recruitment from existing studies or databases are some of the traditional recruitment methods that have been employed previously to invite patients to take part in research (Heerman *et al*., 2017; Rait *et al*., 2015), whereas the most recent common tracking methods include home visit, telephone/cell phone follow-up or contacting pre-identified contact persons (Hobden *et al*., 2011). Additionally, social media is also increasingly being used for recruitment and follow-up, however their success as well as feasibility in low-resource settings have not been reported (Gelinas *et al*., 2017. Numerous studies, both in paediatric and adult HIV-1 infection have been described with long term follow up. Studies in paediatric (Fox and Rosen, 2015; EPPICC study group, 2018) and adult (Rosen *et al*., 2007; Fox and Rosen, 2010) HIV-1 infection have been described with long term follow up.

1.9.1. Loss to follow-up

A review of treatment-naïve HIV-infected adults at ART initiation in a South African single site retrospective cohort study, defined a loss to follow-up as three months late for a scheduled appointment without a later visit (Butler *et al*., 2018). Loss and withdrawal from study follow-up is most common in longitudinal studies (Heerman *et al*., 2017). Therefore, for studies in medical research to be effective and successful, it is necessary to improve the capacity to maintain longitudinal cohorts taking cognisance of important cultural, logistical as well as social obstacles. Poverty associated factors (i.e. employment, informal housing, mobility, changing contact details as well as competing priorities such as child care), results in inefficient participant tracking and study participation (Onoya *et al*., 2011; Barakat-Haddad *et al*., 2009). HIV-1 associated stigma and risks associated to accidental HIV-status disclosure are the most common hurdles in HIV-related research (Onoya *et al*., 2011).

1.9.2. Efforts to promote successful follow-up

Maintaining regular communication has been identified as an important study retention strategy (Lee *et al*., 2000). Various communication approaches have been put forward that affect study completion rates (Beering *et al*., 2016). The use of all available contact mechanisms and channels, especially cell phones, have been shown to be important modalities for retaining follow-up, as fixed line telecommunications are not common in the highly mobile cohorts where most participants are in search for employment. These services have been used to both remind participants about study visits and to follow-up with missed visits (van Loggerenberg *et al*., 2008). Some of the strategies that usually make it easier to locate participants in the long term, thereby lowering the intensity and costs of contact attempts, are the provision and participant preference of addresses for relatives and rural homes and ensuring confidentiality is kept during all communication processes (Lee *et al*., 2000). Furthermore, relationships formed with the study staff members', knowledge that if there are any emergencies participants will be referred to a wellness clinic for medical attention, regular counselling sessions and the financial compensation for study visits are additional factors that have definitely contributed to the high retention rates (van Loggerenberg *et al*., 2008).

1.10. Study rationale

Studying models that represent HIV-1 control in comparison to HIV-1 progression may provide continued insights into mechanisms of host control, and may reveal specific targets or combination of factors that may be considered in developing new therapeutic strategies to attain a "functional cure". It has been demonstrated that HIV-1 has become less virulent over time, resulting from the selective viral adaptation to associated protective HLA alleles with slow disease progression as compared to previous studies (Payne *et al.*, 2014). The same could be true for other protective markers such as genetic variants associated with the co-receptors CCR5 and CXCR6.

As previously mentioned, in a cohort of HIV-1 infected Controllers and Progressors in Johannesburg, South Africa, four SNPs, two in the *CCR5* gene and two in the *CXCR6* gene were identified, that can serve as markers (both protective and deleterious depending on the SNP) in the context of HIV-1 control. However, we cannot assume that similar results will be found in the Durban cohort, South Africa, where the epidemic dynamics are different. It is therefore necessary to test patients from different sites separately to determine whether the

same associations are seen. If they are, then the sample numbers can be increased by combining data across sites in future studies, but if not, then the host and viral factors found in the Durban cohort, the epicentre of the HIV epidemic in SA could be investigated further.

1.10.1. Aim of the study

The overall aims of the study are:

1. To describe the experiences, opportunities and challenges encountered in the recruitment process of the Durban cohort.
2. To determine and compare the representation of protective/deleterious alleles and genotypes of chemokine receptors *CCR5* and *CXCR6*, previously identified in HIV-1 infected individuals showing variable levels of control in a Johannesburg cohort, to similarly HIV-1 infected individuals recruited in Durban.

1.10.2. Study specific objectives

1. To aid with recruitment and to describe the experiences, opportunities and challenges of recruiting HIV-1 infected Controllers and Progressors in Durban from primary healthcare clinics (PHCs).
2. To compare the representation of the *CCR5* (rs746492 and rs553615728) and *CXCR6* (2234355 and rs2234358) variants in a subset of healthy black South African population (N=87) (Ramsay *et al.*, 2016) to other reference population data obtained from the 1000 Genomes Project (www.1000genomes.org).
3. To genotype and subsequently determine if the protective and deleterious variants located in the *CCR5* (rs746492 and rs553615728) and *CXCR6* (2234355 and rs2234358) genes described in a Johannesburg cohort show similar findings with HIV-1 control in the Durban cohort.
4. To analyse the combined effect of protective/deleterious genotypes of *CCR5*/*CXCR6* and describe patterns of possession in HIV-1 infected Controllers and Progressors.

CHAPTER 2: Materials and Methods

2.1. Study design and participants

The Durban component of "HIV-1 Functional cure study" was conducted between October 2015 and August 2018 at two eThekwini PHCs in Durban, KZN, namely Lancers Road and Chesterville clinics. HIV-1 infected individuals that were ≥ 18 years of age were screened from both PHCs and preliminary recruitment was based on their $CD4^+$ T-cell count. These individuals had previously tested positive using HIV-1/2 rapid tests at the clinic's voluntary counselling and testing (VCT) site. Participants were recruited irrespective of time since seroconversion or diagnosis and assigned into two broad groups, namely (i) Controllers: if $CD4^+$ T-cell counts were >500 cells/µl with no additional age limit and (ii) Progressors: if $CD4^+$ T-cell count <200 cells/µl and ≤35 years of age. Final enrolment onto the study however depended on HIV-1 RNA VL counts with Controllers having VLs <2000 RNA copies/ml and Progressors having VLs >10 000 RNA copies/ml. However these categories were defined using $CD4^+$ T cell count and VL data from a single time point only with an unknown timing after HIV-1 infection. The study exclusion criteria included: current use of ART, evidence of an AIDS-defining illness except for tuberculosis (TB) and current pregnancy. The only known ART-exposed participants were the females who had been on the PMTCT programme prior to enrolment and disclosed the administration of single-dose nevirapine during labour.

In addition to the HIV-1 infected individuals recruited in the Durban cohort there was an additional 3573 HIV-1 uninfected and infected black South African individuals termed South African Blacks (SABs), whose genotypic data was available for the following three SNPs, (*CCR5* rs746492 and *CXCR6* rs2234355 and rs2234358 SNPs). This data was sourced from the AWI-Gen H3 Africa study (Ramsay *et al.*, 2016). Whole genome sequencing (WGS) was undertaken in a sub-group (N=100) of the 3573 individuals, which comprised HIV uninfected (N=87) and HIV infected (N=13) individuals. The HIV uninfected individuals (N=87), where in addition to the SNPs mentioned above, data for the *CCR5* rs553615728C/T SNP was available are termed healthy controls (HCs) and this group was used for comparative purposes in the current study. The South African individuals who were part of the AWI-Gen H3-Africa study were 40-60 years of age, represented Zulu, Xhosa, Venda, Tsonga, Pedi, Sotho and Tswana ethnolinguistic groups and were recruited across three South African Health and Demographic Surveillance System sites namely: Agincourt, Dikgale and Soweto in Mpumalanga, Limpopo and Gauteng, respectively. Genomic DNA extraction and genotyping was performed as detailed by Ali *et al*. (2018).

Furthermore, this study included ART-naïve HIV-1 infected black South African individuals from Johannesburg (JHB) for comparative purposes and data was obtained from two separate studies. The *CCR5* allelic and genotypic data was obtained from Koor *et al*. (2019), N=134 [Progressors (N=74) who were ART-naïve individuals requiring ART initiation upon enrolment and selected based on progression rate; Controllers (N=60) were characterised as: VCs (N=37) defined by $CD4^+$ T-cell counts >500 cells/µl with VL< 2000 RNA copies/ml and ECs (N=23) with $CD4^+$ T-cell counts >500 cells/µl and VL <50 RNA copies/ml. CXCR6 data was obtained from Picton *et al*. (2017), N=113 individuals who were assigned into different groups based on $CD4^+$ T-cell counts and VL data obtained over several time points: [Progressors (N=72), showed decline of $CD4^+$ T-cell counts from above 500 cells/µl to <350 cells/µl, initiated on ART and had VL >10 000 RNA copies/ml, VCs (N=30) defined by $CD4^+$ T-cell counts >500 cells/µl with VL <2000 RNA copies/ml and ECs (N=11) with $CD4^+$ T-cell counts >500 cells/µl and had at least one VL test yielding <50 RNA copies/ml. The JHB cohort had a comparable number of females between Controllers and Progressors.

2.1.1. Screening and enrolment procedures

Once the purpose of the study was explained in the participant's language of choice (either English or IsiZulu) and IsiZulu was mainly used, voluntary informed consent (IC) forms were signed by study participants. Documentation consenting to long term storage of samples was also obtained voluntarily from all study participants. Structured study questionnaires were also administered to ascertain the mode of HIV acquisition, estimate time of diagnosis, ART history as well as detailed information on sexual behaviour for all participants on their first study visit. Blood samples were taken for VL determination and participants asked to return to the PHC after a week for their VL results. Those who did not meet the study criteria were "exited" from any further follow up procedures within the study. Those who met the VL study criteria had blood samples collected and blood was processed: this included isolation and cryopreservation of peripheral blood mononuclear cells (PBMCs) and storage of serum, plasma and whole blood (with and without RNA-later). Enrolled participants also voluntarily signed an IC form for genetic testing. Study related questionnaires regarding the educational background, medical history, sociodemographic characteristics and locator information were also completed. Signed copies of all IC forms were given to all study participants. The study design is outlined in Figure 2.1.

This recruitment strategy was modified in September 2016 when the National Department of Health in South Africa, adopted the World Health Organisation (WHO) recommendations of universal treatment to all adults living with HIV, regardless of $CD4^+$ T-cell count (WHO, 2016). Upon identification in the "blood room", all documentation and study procedures for the first study visit and enrolment visit were conducted on the same day for the Progressor group due to participants in this group being initiated on antiretroviral (ARVs) immediately. We had estimated from previous VL estimations that most of these participants would meet the inclusion criteria. Controllers were counselled to commence ART after enrolment in keeping with the South African National Department of Health ART guidelines.

2.1.2. Follow-up procedures

Six monthly follow-up visits were scheduled for the Controller group at which whole blood was collected for $CD4^+$ T-cell count, VL determination and for the processing and storage of PBMCs, serum, plasma and whole blood for DNA and RNA isolation. Participant contact details were updated (if necessary) at each visit and for those who reported ART initiation, other study-related documents were completed and the date of ART initiation recorded. All participants, upon recruitment and follow-up visits, were financially compensated for their time and effort in the study, and/or for costs associated with travels during the study consultation at each study-related visit. Analysis of longitudinal follow-up data collected was still in progress at the time of submission of this dissertation and therefore was not included for the purposes of this study. One of the analyses included the testing of the presence of ARV drug levels in plasma samples for participants in the EC group to determine and confirm individuals as "true" Controllers.

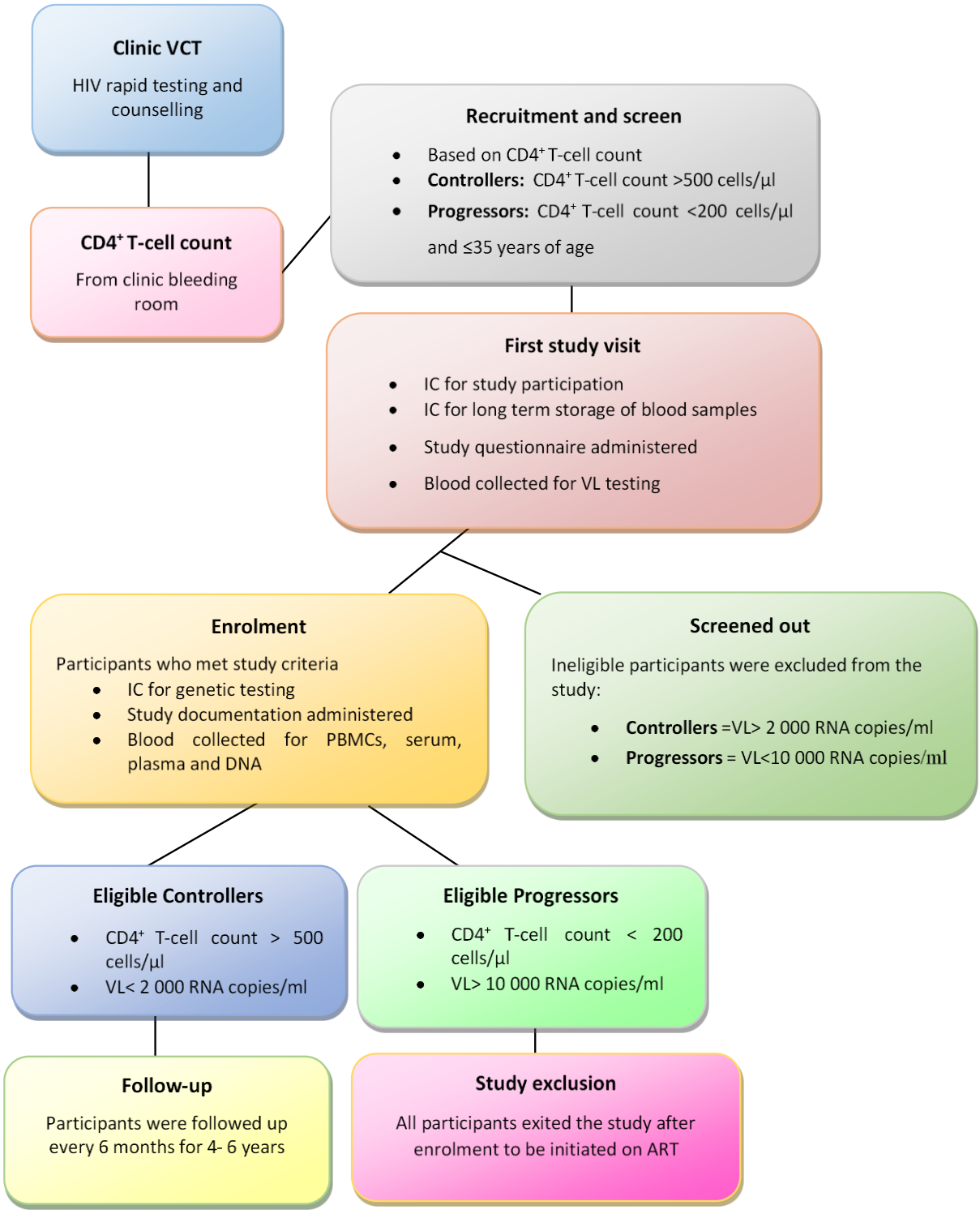


Figure 2.1. Flow diagram showing HIV-1 study design for the HIV-1 Functional cure study.

2.2. Laboratory methods

2.2.1. $CD4^+$ T-cell count enumeration

$CD4^+$ T-cell counts were performed by National Health Laboratory Services (NHLS) (Addington laboratory, Durban, KZN, South Africa) for all participants as part of HIV routine care and management. The NHLS used the Pan Leucogating (PLG) method for $CD4^+$ T-cell enumeration (Beckman Coulter) (Glencross *et al*., 2008).

2.2.2. HIV-1 RNA measurements

The HIV-1 viral loads were determined using the Roche Cobas® 6800/8800 System (Roche Diagnostic Systems, Mannheim, Germany) at the NHLS Department of Virology (Durban, KZN, South Africa). For statistical purposes, all VL results that were either lower than detectable limit (LDL) or <50 RNA copies/ml were imputed as half the limit of detection (20 RNA copies/ml). Therefore, a value of 10 RNA copies/ml, was imputed for all undetectable VL measurements.

2.2.3. Specimen preparation and storage

Blood samples were collected in ethylene-diamine-tetra-acetic acid (EDTA) and serum separating tubes (SSTs). An aliquot of EDTA whole blood was stored at -80°C and the remaining blood was processed for PBMC separation (see below) and cryopreservation. The plasma recovered from all EDTA blood collections was stored in aliquots at -80°C. Serum recovered post centrifugation of the SSTs was also stored at -80°C. Manual and an electronic sample inventory files were created for all sample aliquots stored from all study participants.

2.2.4. Peripheral Blood Mononuclear Cells (PBMCs) separation and storage

PBMCs were isolated from blood samples anticoagulated with EDTA by density gradient centrifugation (Ficoll-Paque) using standard methods (Ulmer *et al*., 1984) and stored in 90% fetal bovine serum (FBS), 10% dimethyl sulfoxide (DMSO) at -80ºC for short term and then in liquid nitrogen for long term storage.

2.2.5. Extraction of genomic DNA

Genomic DNA (gDNA) was isolated according to the manufacturer's instructions from stored whole blood sample using the QIAamp DNA Blood Mini kit (Qiagen, Dusseldorf, Germany). The DNA was quantified using Nano Drop 2000 (Thermo Fisher Scientific, Massachusetts, USA) and stored at -20°C until further use.

2.2.6. Single nucleotide polymorphism (SNP) genotyping

Two different assays were used to genotype the *CCR5* and *CXCR6* SNPs of interest. The one assay, namely the C_T shift assay, is an in-house designed assay that is SYBR green based, and the second assay, is a predesigned Taqman probe-based assay that is commercially available. Details are provided below.

2.2.6.1. Real-time CT Shift Assay

The real-time cycle threshold (C_T) shift assay use shifts in cycle threshold to delineate between different genotypes due to the mispriming 3' end nonspecific primer nucleotides. This assay uses allele-specific polymerase chain reaction (PCR) with two primers, one specific for the mutant allele (Mt), the other specific to the wild type (WT), as well as a common reverse or forward primer, whose orientation depends on the orientation of the allele-specific primers. Each of the specific WT and Mt primers has a locked nucleic acid (LNA) modified oligonucleotide at the 3' end nucleotide of the primer specific for either the WT or Mt allele, respectively. Using highly thermo-stable LNA oligonucleotides in a primer is advantageous for several reasons: one can use a shorter primer, obtain high sensitivity and affinity binding compared to non-LNA primers, the primer melting temperature is increased and there is less mispriming resulting in larger C_T shifts.

An in-house real-time SYBR green C_T shift assay was set up in 384 well-thermal block plates and amplified on an Applied Biosystems™ QuantStudio™ 5 Real Time System and analysed using the QuantStudio Design and Analysis software v.1.5.1, to genotype the *CCR5* rs553615728 (C/T) SNP at position -4223 (5'-UTR), and two *CXCR6* SNPs, namely rs2234358 (G/T) and rs2234355 (G/A) at positions +1071 (3'-UTR) and +7 (ORF), respectively (Jaumdally *et al.*, 2017, Picton *et al.,* 2017). In this genotyping assay, two qPCR reactions were performed per sample: one reaction using 10µM of wild type (WT) specific primer and a common pair primer (10µM) and the other reaction using10µM of mutant (Mt) specific primer and the same common pair primer (10µM). Each reaction was conducted in a 10µl final reaction volume containing, in addition to the primers, 5µl 1× FastStart Universal SYBR Green Master with ROX (Roche Diagnostics, Mannheim, Germany) and 2µl (~10-90ng) template DNA. Water (H_2O) was used as a non-template control (NTC). Details of assay set-up and primer sequences are shown in Tables 2.1 and 2.2, respectively.

An in-house C_T shift assay was designed and used for genotyping the *CCR5* (rs746492) SNP in 16 DNA samples/individuals where results were indeterminate using the predesigned Taqman genotyping assay employed to genotype this SNP (described under 2.2.6.2). This assay was validated using samples that were successfully genotyped using the predesigned Taqman assay. In this C_T shift assay, the reaction was comprised of 5µl 1× PowerUp SYBR green (Applied Biosystems; Thermo Scientific), 10µM of each forward and reverse primer which were non-LNA primers and 2µl (~10-90ng) template DNA made up to a final volume of 10µl with water. As above, H_2O was used as a NTC (Tables 2.1 and 2.2).

If there was a C_T shift difference of >8 cycles in the WT reaction or Mt reaction due to mispriming of either the WT and Mt primer respectively, then individuals were classified as homozygous for the WT or Mt allele, respectively. If the C_T shift difference was <2 cycles within each of the two PCR reactions, due to the optimal binding of both primers, this indicated that the individual was heterozygous having both the WT and Mt allele. Examples of these respective genotypes from select samples are shown in Figure 2.2. In these assays it is imperative that the misprimed reaction produces a curve so that one can be confident that in that well there was not a failed PCR reaction resulting in absence of curve –hence a mismatched base at the 3'end of the primer resulting in absence of amplification is not desired even though one would theoretically expect no amplification. Thus the mispriming of primers with unmatched 3'end nucleotides allows us to confidently genotype using this type of assay, and designing and optimising these assays involves making sure the C_T shifts are in the desired range with the use of LNA technology as described above.

Table 2.1. C_T shift assay real-time PCR reaction components and set-up

Reaction to detect wildtype (WT) allele			Reaction to detect mutant (Mt) allele		
PCR reaction constituent	**Volume (10µl)**	**Final [ ][a]**	**PCR reaction constituent**	**Volume (10µl)**	**Final [ ][a]**
WT-specific primer (3'end)	1µl	10µM	**Mt-specific primer (3'end)**	1µl	10µM
Common primer	1µl	10µM	**Common primer**	1µl	10µM
10X SYBR green[b]	5µl	5X	**10X SYBR green[b]**	5µl	5X
H_2O	1µl	-	**H_2O**	1µl	-
DNA template	2µl	~10-90ng	**DNA template**	2µl	~10-90ng

[ ][a]: denotes concentration; [b]: The SYBR green used was either FastStart Universal SYBR Green (Roche Diagnostics, Mannheim, Germany) or PowerUp SYBR green (Applied Biosystems; Thermo Scientific), depending on the SNPs being genotyped (see text 2.2.6.1).

Table 2.2. PCR primers and thermal cycle conditions used for the C_T shift assays for genotyping the *CXCR6* and *CCR5* SNPs

C_T Shift assay primers				PCR cycling conditions
Variant	Common Primer	WT Primer	Mt Primer	
CCR5 **rs746492**	5'- AAG TAC AAT TTA CCA GCC TCC GT -3' (F)	5'- CAA GTA TGT GCA CAA TCA TAT GAG AA -3' (R)	5'- CAA GTA TGT GCA CAA TCA TAT GAG AC -3' (R)	Initial denaturation 10 min at 95ºC followed by **55 cycles:** 15 sec denaturation at 95ºC, 59ºC annealing for 40 sec, 70ºC extension for 1 min
CCR5 **rs553615728**	5'- GTG GAG TAA CGC ACA CTG CAA -3' (F)	5'- CCA TTT CCT CAT CTG TTA AAT GAC [G] -3' (R)	5'- CCA TTT CCT CAT CTG TTA AAT GAC [A] -3' (R)	Initial denaturation 10 min at 95ºC followed by **40 cycles:** 15 sec denaturation at 95ºC, 59ºC annealing for 1 min, 72ºC extension for 1 min
CXCR6 **rs2234355**	5'- GAT ATG ACC AGC ACC AGA GAG -3' (R)	5'- CAG AAC AGA CAC CAT GGC A[G] - 3' (F)	5'- CAG AAC AGA CAC CAT GGC A[A] -3'(F)	Initial denaturation 10 min at 95ºC followed by **40 cycles:** 15 sec denaturation at 95ºC, 60ºC annealing for 40 sec, 72ºC extension for 1 min
CXCR6 **rs2234358**	5'- ACC AGC CAG AGT GTC TGA TAA C -3' (R)	5'- GCT CTG GAA TTT GCA AG[G] -3' (F)	5'- GCT CTG GAA TTT GCA AG[T] -3' (F)	

Square brackets ([ ]) denoted lock nucleic acid (LNA) modified 3'end nucleotide and ( ) denoted primer orientation (i.e. F-forward, R- reverse)

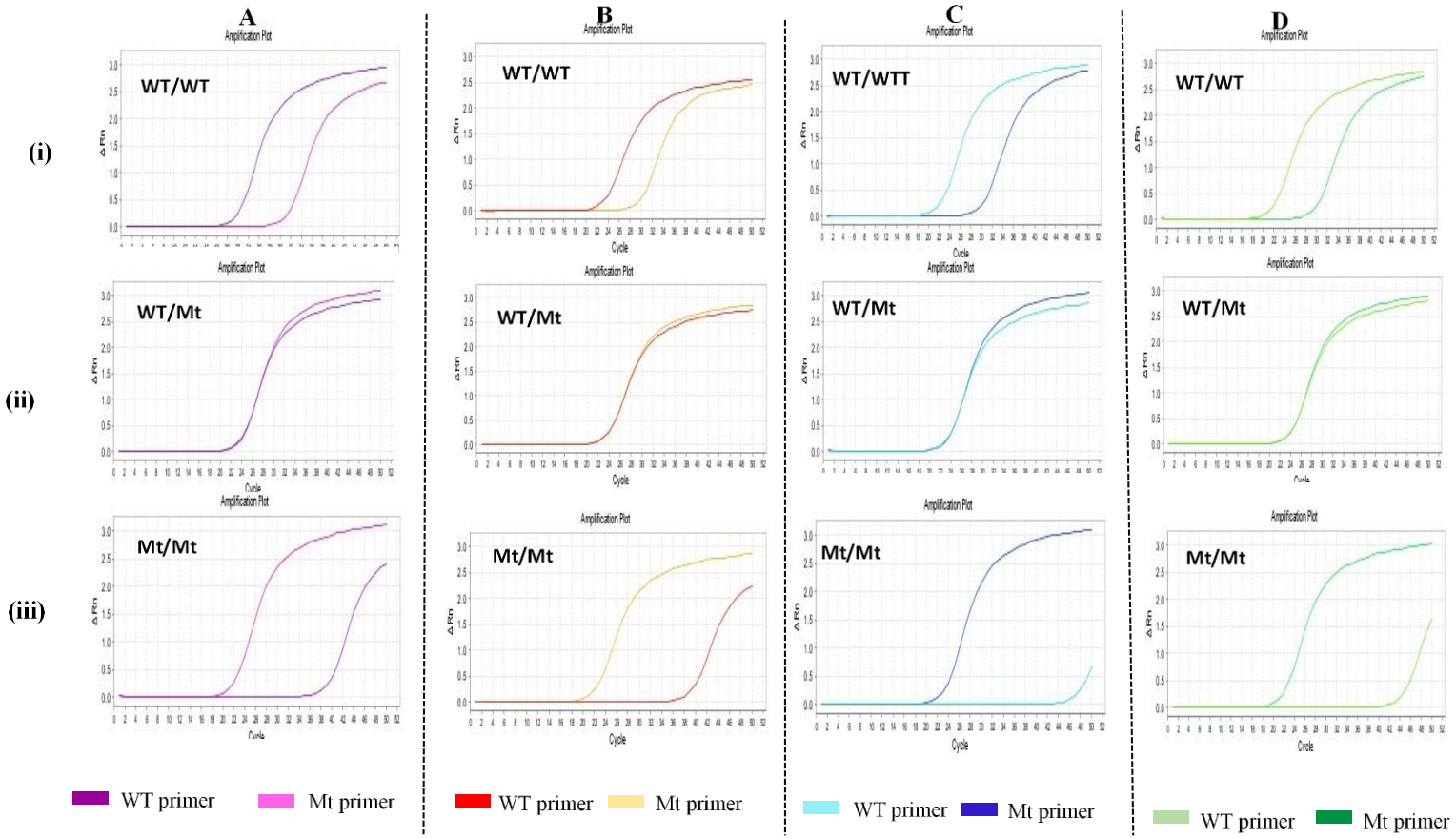


Figure 2.2. Example C_T-shift patterns shown for: (A) *CCR5* rs746492 SNP; (B) *CCR5* rs553615728 SNP; (C) *CXCR6* rs2234355 SNPs and (D) *CXCR6* rs2234358 SNP and where (i) depicts homozygous wildtype (WT/WT), (ii) depicts heterozygous (WT/Mt) and (iii) depicts homozygous mutant (Mt/Mt).

2.2.6.2. Predesigned Taqman genotyping assay

A predesigned commercial Taqman® assay (Thermo Fisher Scientific, USA: 4351379) was performed according to the manufacturer's instructions for the genotyping of *CCR5* rs746492T/G SNP at position +2919 (3'-UTR). Briefly, for each sample a mix was made containing 1µl of 10x *CCR5* +2919 Taqman Assay (containing probes and primers), and 5 µl of 2X Roche LightCycler 480 Probes Master Mix (Roche Diagnostics, Mannheim, Germany), 2µl nuclease-free water (Ambion) and 2µl (~10-90 ng) template DNA. The VIC™ labelled probe detects the G allele and the other FAM™ labelled probe detects the T allele. The sequence is as follows: [VIC/FAM]: AGA GAG AGA GGT TTT TTT CTG TTC T[G/T]TCT CAT ATG ATT GTG CAC ATA CTT G.

Reactions (10µl final volume) were set up in 96-well plates and amplified on an Applied Biosystems 7500 Real Time System with the 7500 software which determines the genotypes based on the assay details entered. The final amplification settings were as follows: an initial denaturation at 95°C for 10 minutes followed by 40 cycles of denaturation of 95°C for 15 seconds and annealing and extension at 60°C for 50 seconds. No template controls (sterile water) were included. Figure 2.3 shows schematic results of select individuals genotyped using this assay as generated by the 7500 software.

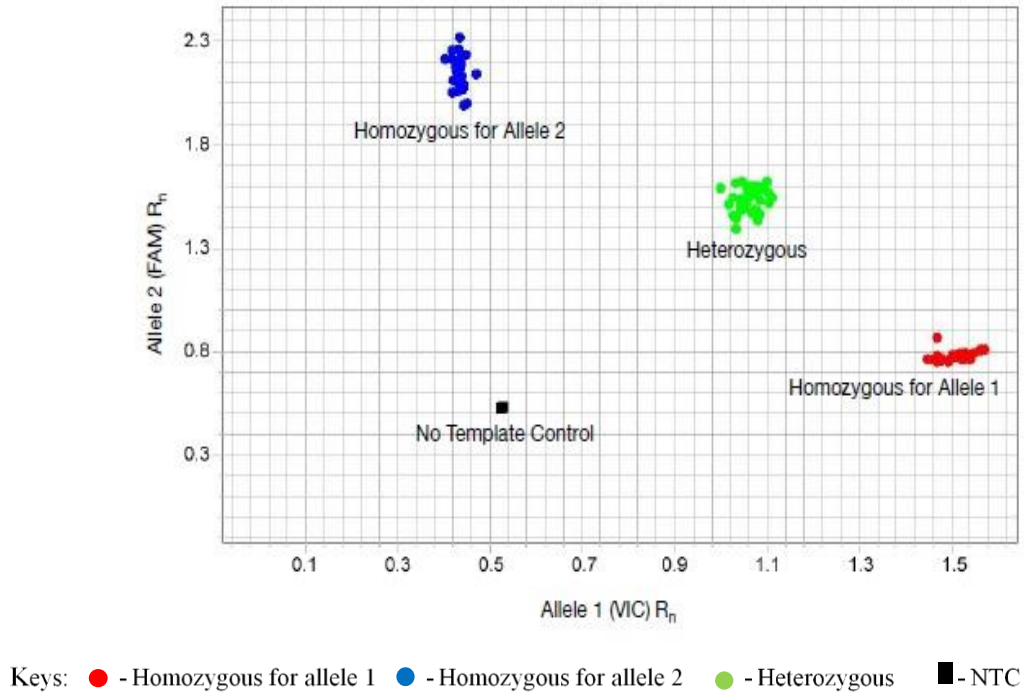


Figure 2.3. An allelic discrimination plot depicting results for the predesigned Thermo Fisher Taqman SNP genotyping assay. Reproduced from (http://www.fda.gov/media/119796/download).

2.3. Ethical clearance

Ethics approval for the "HIV-1 Functional Cure study" was obtained from the Human Research on Human Subjects Committee of the University of the Witwatersrand, Johannesburg, certificates M14026 and M190995 (Prof Caroline T. Tiemessen) (see Appendix 1A). Ethics clearance for the cohort in KZN was obtained from Medical Research Ethics Committee (EC017-11/14) (see Appendix 2A) and the EThekwini Research Ethics Committee (No.M.1/1/2 22 September 2015) (see Appendix 3A). For the 'Department of Health approval letter' by the KZN Department of Health (69/16) for research to be undertaken at Chesterville clinic, (see Appendix 4A).

2.4. Data collection and documentation

All data from the completed questionnaires as well as $CD4^+$ T-cell and VL estimations were entered into a specific database which was secured using a password-protected access system.

All data were de-identified to maintain participant confidentiality using a participant study number.

2.5. Data analysis

To compare SNP allelic and genotypic frequencies between groups, the Chi-square test or Fishers exact test, where appropriate, were used. The difference in allele frequencies between the groups was presented using odds ratios with 95% confidence intervals. The Shapiro Wilk test was used to test for normality of $CD4^+$ T-cell count and viral load. The Wilcoxon rank sum test was used to compare $CD4^+$ T-cell counts and viral load (VL) between two groups of interest. Kruskal-Wallis H test was used to analyse the relationship between *CCR5* and *CXCR6* SNP genotypes and $CD4^+$ T-cell counts while Mann-Whitney U test was used for VL. Analyses were performed using Graph Pad Prism version 5.0 for Windows (Graph pad Software, San Diego, California, USA (*www.graphpad.com*) and STATA 15.0. Two-sided tests were used and statistical significance considered set at $P<0.05$ and the $0.05<P<0.10$ were considered as trends of significance. No corrections for multiple testing were used for all p-values.

CHAPTER 3: Results

3.1. Clinical demographics

3.1.1. Screening and enrolment of eligible participants

A total of 590 HIV-1 infected participants were identified and screened via the clinic bleeding rooms based on their $CD4^+$ T-cell count documented in the clinic register. They were recruited between 2015 and 2018 into two groups namely Controllers (N=354) with $CD4^+$ T-cell absolute counts ≥500 cells/µl and Progressors (N=236) with $CD4^+$ T-cell absolute counts ≤200 cells/µl. Eligible participants (N=287) were enrolled based on VL measurements (Figure 3.1).

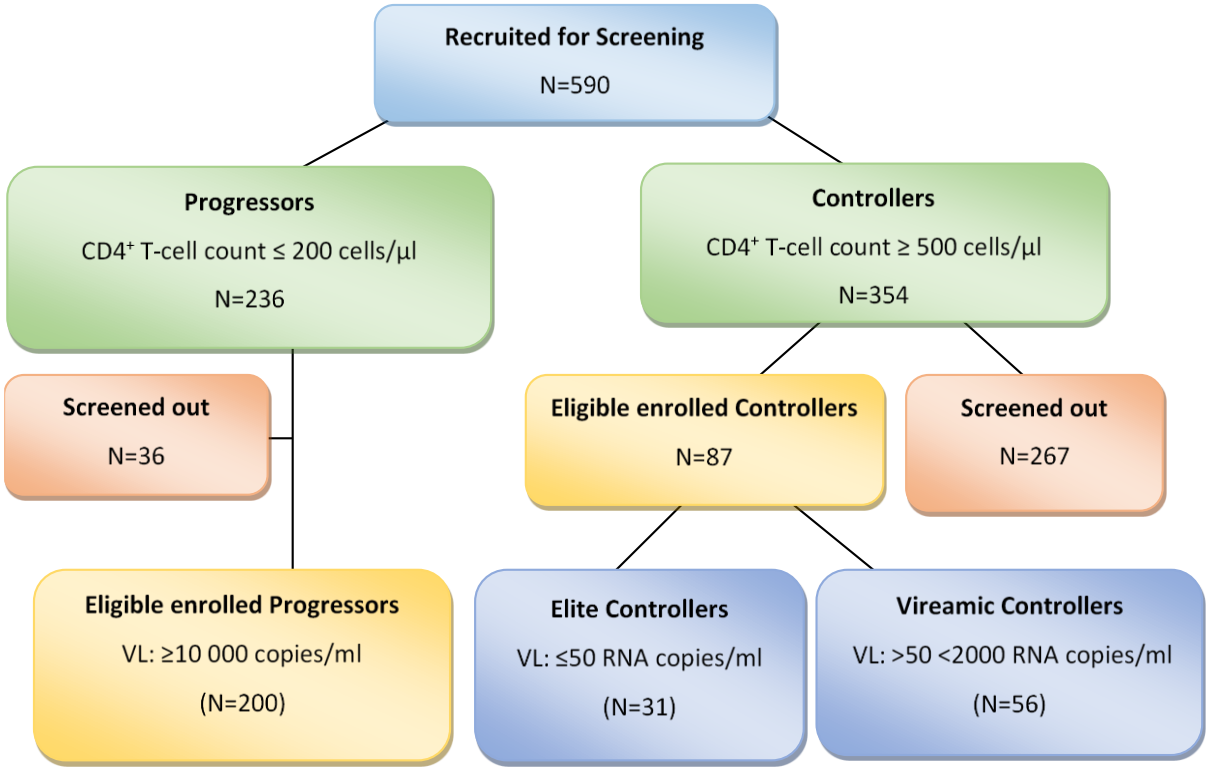


Figure 3.1. Schematic flow diagram for screening and enrolment of participants for the HIV-1 Functional cure DBN cohort study.

3.1.1.1. HIV-1 Controllers

Eighty- seven participants were eligible and enrolled in the Controller group of the study. Based on the VL results, the Controllers were further sub-grouped into 2 categories: Elite Controllers

(EC) (N=31) with viral load ≤50 RNA copies/ml and Vireamic Controllers (VC) (N=56) with 50 >VL<2000 RNA copies/ml (Figure 3.1).

Two hundred and sixty seven participants were screened out of the study because they did not meet the study criteria for enrolment for the following reasons: 245/267 (91.8%) had high VL measurements (>2000 RNA copies/ml); 13/267 (4.9%) had started ARV therapy after being screened but before being enrolled; 3/267 (1.1%) had relocated to another province; 1/267 (0.4%) did not wish to participate after being screened and 5/267 (1.9%) did not return to the clinic for enrolment although their VL measurements did meet the study criteria for enrolment.

3.1.1.2. Progressors

Two hundred Progressors were enrolled whilst 36 were screened out for the following reasons: 4/36 (11.1%) were older than 35 years, which was the age criteria for enrolment (these participants initially reported an incorrect date of birth at the screening visit which was verified against their identity document upon their return visit for enrolment); 5/36 (13.9%) had already commenced ARVs prior to the study, which was not reported when they were screened; 7/36 (19.4%) did not return to the clinic for any further follow-up for their clinical diagnosis and were uncontactable, and 20/36 (55.6%) participants had low VLs (<10 000 RNA copies/ml) that did not meet the study criteria for this group.

3.1.2. Demographic, sexual and behavioural characteristics of all participants (N=590) screened

The majority 73.7% (435/590) of the participants in this cohort were female with the median age of 29 years, 80.7% (476/590) of whom self-reported that their sexual debut occurred when they were older than 15 years of age while 12.0% (71/590) self-reported that they were younger than 15 years of age and 6.8% (40/590) did not know the actual age of their sexual debut, with a small number 0.5% (3/590) not willing to disclose this information (Table 3.1). A large portion of participants 68.3% (403/590) acknowledged having only one sexual partner in the last 6 months (prior to being screened to participate in the study) while 23.2% (137/590) had 2 or more partners, 8.0% (47/590) had no partner and only 0.5% (3/590) were unwilling to disclose. Heterosexual HIV-1 transmission was self-reported in the majority 98.1% (579/590) with only homosexual and other forms of transmission occurring in a minority, 0.5% (3/590) and 1.4% (8/590), respectively (Table 3.1).

Condom usage with regular partners was occasional in 44.2% (261/590), while 19.5% (115/590) reported never using condoms during sexual acts, with 23.7% (140/590) reporting to always using condoms and a minority 0.2% (1/590) did not want to disclose this information. In 12.4% (73/590) who were asked about condom usage upon recruitment replied that they had no regular partner implying not using condoms at the time (Table 3.1). A small number of individuals, 5.4% (32/590), reported taking ARVs in the past for PMTCT and the use of traditional medicine was reported by 3.9% (23/590) of individuals. The self-reported timing of HIV infection with respect to HIV diagnosis was 1 month to 3 years in 71.9% (424/590); 3-10 years in 18.6% (110/590) and > 10 years in 8.6% (51/590), and data was not available for 0.8% (5/590) (Table 3.1).

Table 3.1. Characteristics of the recruited cohort (N=590) prior to screening out of non-eligible individuals

Characteristics (demographic, sexual and behavioral)		N (%)
Gender	Female; 29 (26-34) *[Median (IQR) age in years]*	435 (73.7)
Age at sexual debut	≤15 years >15 years Do not know Not willing to disclose	71 (12.0) 476 (80.7) 40 (6.8) 3 (0.5)
Number of sexual partners ***[in the last 6 months]***	0 1 2+ Not willing to disclose	47 (8.0) 403 (68.3) 137 (23.2) 3 (0.5)
Route of HIV-1 transmission as deemed by participant	Heterosexual Homosexual Other	579 (98.1) 3 (0.5) 8 (1.4)
Pattern of condom usage	Always Occasionally Never Do not have a regular partner Not willing to disclose	140 (23.7) 261 (44.2) 115 (19.5) 73 (12.4) 1 (0.2)
Utilization of ARVs prior to the study		32 (5.4)
Utilization of traditional medicine for HIV		23 (3.9)
Timing of HIV-1 infection	1 month- 3years 3-10 years >10 years Not available	24 (71.9) 110(18.6) 51(8.6) 5(0.8)

3.1.3. CD4+ T-cell counts and viral load measurements for enrolled participants

A similar median age (in years) was found between the EC, VC, Controller and Progressor groups (31 years, 31 years, 30 years and 30 years, respectively) (Table 3.2). The number of females in the Progressor group 62.5% (125/200) was significantly lower compared to the Controllers 85.1% (74/87); EC 80.7% (25/31) and VC groups 87.5% (49/56) (P0.0006, P=0.007, P<0.0001, respectively).

As expected given the selection criteria, significantly higher median $CD4^+$ T-cell absolute count was observed in the Controllers (842 cells/µl, P<0.0001), ECs (954 cells/µl, P<0.0001) and VCs (774 cells/µl, P<0.0001) compared to Progressors (116 cells/µl). Similarly, the median viral load was significantly higher in the Progressor group (249 000 RNA copies/ml) compared to the Controllers (212 RNA copies/ml, P<0.0001), ECs (10 RNA copies/ml, P<0.0001), and VCs (884 RNA copies/ml, P<0.0001).

Table 3.2. $CD4^+$ T-cell counts, viral load measurements, gender and age of the enrolled participants (N=287)

Study groups	N	Age (years) (Median - IQR)	Gender (% Female)	$CD4^+$ T-cell Count (cells/µl) (Median - IQR)	VL (RNA copies/ml) (Median - IQR)	Log VL (Median - IQR)
Controllers	87	30 (26.5-33)	85.1	842 (708-1003)	212 (10-1065)	2.33 (1-3.03)
Elite Controllers	31	31 (26-39)	80.7	954 (833-1106)	10[a] (10-47)[b]	1 (1-1)
Viraemic Controllers	56	31 (27-37)	87.5	774 (664-941)	884 (227-1475)	2.95 (2.36-3.17)
Progressors	200	30 (26-32)	62.5	116 (59-165)	249 000 (87 550-78 1500)	5.40 (4.95-5.89)

[a]: Undetectable VLs were imputed with 10 RNA copies/ml; [b]: the normal range was used as the 1st and 3rd quartiles were equal

3.2. *CCR5* (rs746492 and rs553615728) and *CXCR6* (rs2234355 and rs2234358) SNPs in the Black South African population and other reference populations

It is important to study the genetic diversity characterized by mutations and frequencies of disease-related variants in the people from the African continent. Several African populations have been studied due to the availability of data through genomic browsers in the public domain. South Africa has one of the highest HIV/AIDS prevalence rates in the world with increasing rates of lifestyle diseases and rapid urbanisation (Ramsay *et al*., 2011).

Populations in Luhya in Webuya, Kenya (LWK), East Africa, Yoruba in Ibadan, Nigeria (YRI) West Africa, Esan in Nigeria (ESN), Gambia in Western Division, The Gambia (GWD) and Mende in Sierra Leone (MSL) (Figure 3.2) have been well studied and documented in the HapMap project (which documents a haplotype map of the human genome) (*http://hapmap.ncbi.nlm.nih.gov*) as well as the 1000 Genomes Project (a comprehensive data source of human genetic variation) (*www.1000genomes.org*). There are, however, limited population genetics data for the Southern African region (i.e. Swaziland, Lesotho, Botswana and Namibia) (May *et al*., 2013). Population data for three of the variants under study (*CCR5* rs746492, *CXCR6* rs2234355 and rs2234358) were available for a group of South African black individuals (SABs; N=3573) (Ramsay *et al*., 2016). In a subset of the same individuals (100/3573), WGS was undertaken of whom 87/100 were HIV-negative (termed healthy controls; HCs) and where population data was also available for the additional *CCR5* rs55361578 SNP and this latter group was therefore used for comparative purposes.

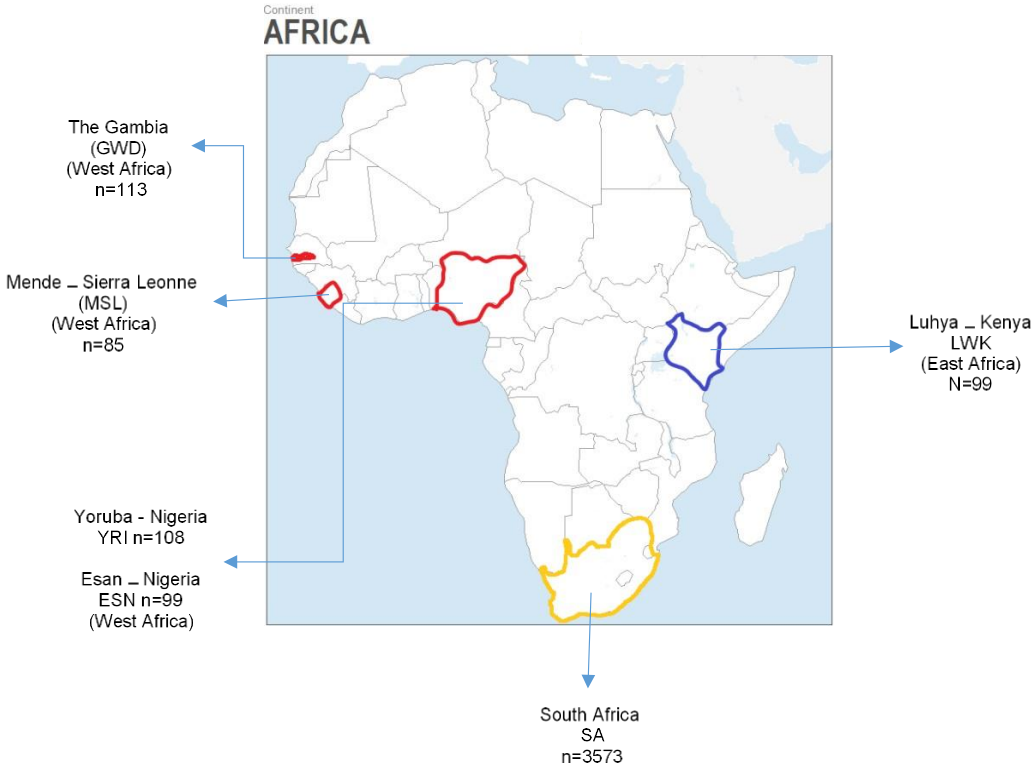


Figure 3.2. Map of the African continent depicting the African countries represented in the 1000 Genomes Project (*www.1000genomes.org*). Non-African based populations with African ancestry are not shown.

Data for only three variants (*CCR5* rs746492, *CXCR6* rs2234355 and rs2234358) was available for the total European population (*www.1000genomes.org*). This European population included Utah residents with European ancestry (CEU), Finnis in Finland (FIN), British in England and Scotland (GBR), Iberian populations in Spain (IBS) and Toscani in Italy (TSI) (Figure 3.3).

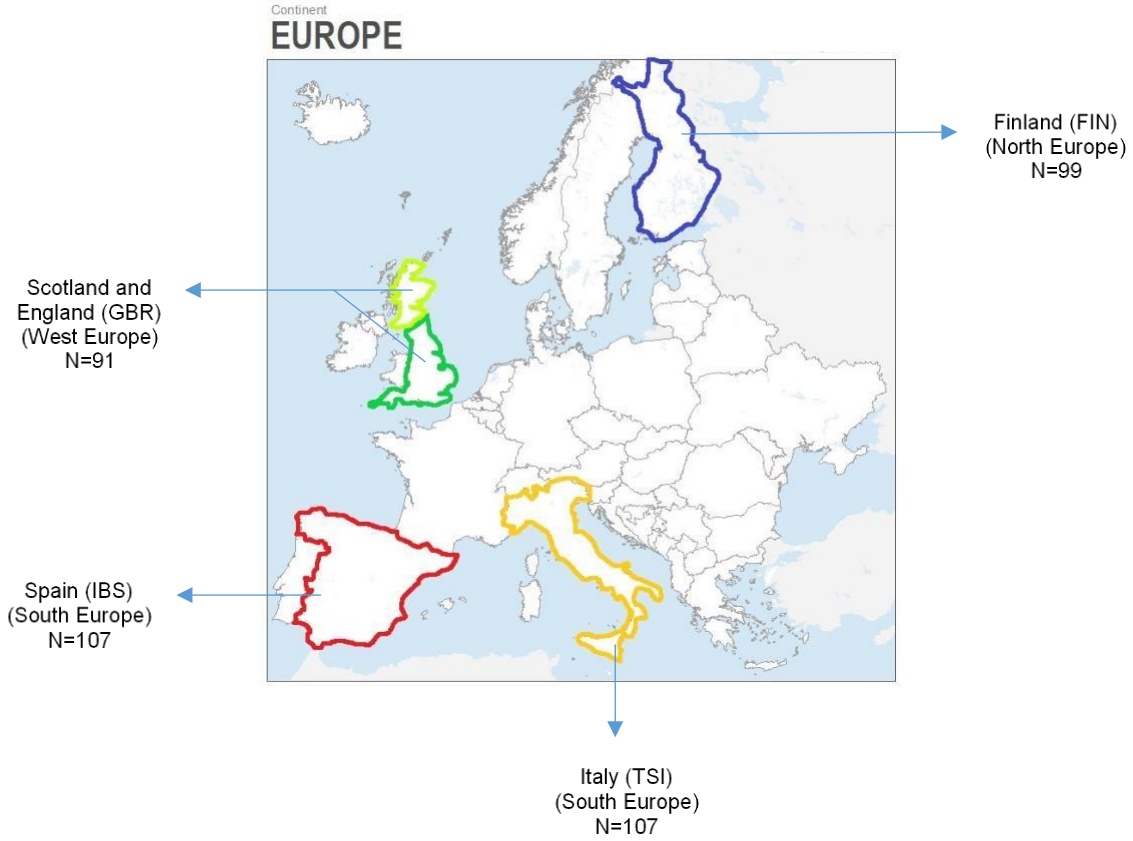


Figure 3.3. Map of the European continent depicting the European countries represented in the 1000 Genomes Project (*www.1000genomes.org*).

We therefore, compared the HCs (N=87) to the YRI (N=108), and LWK (N=99), as representative populations from west and east Africa, respectively, as well as the total African population (AFR) (N=661) which included: ESN (N=99), GWD (N=113), MSL (N=85) and populations with African ancestry [i.e. African Caribbean in Barbados (N=96) and African ancestry in Southwest USA (N=61)], in addition to YRI and LWK. HCs were also compared to the EUR populations (N=503) [i.e. CEU (N=99), FIN (N=99), GBR (N=91), IBS (N=107) and TSI (N=107)]. The allelic and genotypic frequencies for three of four variants (*CCR5*

rs746492, *CXCR6* rs2234355 and rs2234358) under study in HCs, LWK, YRI, AFR and EUR populations are listed in Table 3.3.

In the comparison of the allelic frequencies for the *CCR5* rs746492 SNP, the G allele was significantly higher in HCs compared to YRI (HCs: 42.0% vs. YRI: 30.6%, P=0.03) but significantly lower than the EUR (HCs: 42.0% vs. 53.5%, P=0.005) (Table 3.3). Similar frequencies were found when HCs were compared to LWK (HCs: 42.0% vs. LWK: 47.5%, P=0.30) and the total African population (HCs: 42.0% vs. AFR: 38.0%, P=0.32), Table 3.3.
When comparing the *CCR5* rs746492 genotypic frequencies for the various populations (Table 3.3), it was interesting to note that the YRI population showed a significant difference in genotype distribution compared to the SA Black HC population (P=0.028) and the LWK trended towards a significant difference compared to the HC population group (P=0.077). The total African population however did not differ significantly from the HC population (P=0.23) while the European population did show a significant difference in genotype distribution compared to the HC population (P=0.014). In the YRI vs. HC comparison, the two genotypes contributing to the significant difference were major allele homozygosity (TT; HC: 31% vs. YRI: 50%) and heterozygosity (TG; HC: 54% vs. YRI: 38.9%), while with the LWK population, minor allele homozygosity was largely responsible for the distribution difference (GG; HC: 14.9% vs. LWK: 27.3%) which was more similar to the EUR representation (GG; EU: 26.8%).

When comparing both the allelic and genotypic frequency distribution for the *CXCR6* rs2234355 SNP, HCs showed significant differences to both the EUR and YRI populations and were more similar to LWK and total African population groups. The minor allele (A) was significantly overrepresented in the YRI population compared to HCs (YRI: 60.2% vs. HCs: 46.6%, P=0.008) and significantly underrepresented in the EUR population compare to HCs (EUR: 0.5% vs. HCs: 46.6%, P<0.0001). Given the low representation of this SNP in the EUR population and the absence of minor allele homozygosity (AA), it is not surprising that the genotypic distributions differed very significantly between the HC and the EUR populations (P<0.0001). Major allele homozygosity (GG) was much higher in the HC population relative to the YRI population (GG; HC: 33.3% vs. YRI: 14.8%) which resulted in a significantly different genotypic distribution between the two population groups (P=0.009). The LWK and total African populations were more similar to the HC population and did not differ with respect to the distribution of genotypes (Table 3.3).

Both allelic and genotypic representation of the *CXCR6* rs2234358 SNP in the LWK population was similar to HCs. The *CXCR6* rs2234358-T allele was however significantly overrepresented in the YRI and total African populations compared to HCs (YRI: 75.5% vs. HCs: 56.9%, P<0.0001; AFR: 68.5% vs. HCs: 56.9%, P=0.003) while significantly underrepresented in EUR (EUR: 46.9% vs. HCs: 56.9%, P=0.02). Comparison of the genotype distributions between the YRI and total African populations to the HC populations both showed significant differences (YRI vs. HC: P=0.0003; AFR vs. HC: P=0.009) and the EUR population showed a strong trend of difference in genotypic distribution compared to the HC (P=0.054) (Table 3.3). For both the YRI and the total African populations there was similar heterozygosity representation to the HCs but both minor (TT) and major allele (GG) homozygosity were higher and lower compared to the HC population, respectively (Table 3.3), resulting in the significant differences in genotype distributions.

Table 3.3. Comparison of representation of *CCR5* (rs746492) and *CXCR6* (rs2234355 and rs2234358) SNPs between a Black South African Healthy population to African and European populations

SNP	Alleles and Genotypes	Black SA HC[a] (N=87)	Luhya (N=99)[b]	Black SA HC vs. LWK	Yoruba (N=108)[b]	Black SA HC vs. YRI	Total Africa (N=661)[b]	Black SA HC vs. AFR	Total European (N=503)[b]	Black SA HC vs. EUR
		N (%)	N (%)	P[c]	N (%)	P[c]	N (%)	P[c]	N (%)	P[c]
CCR5 rs746492	**Allele**									
	T	101 (58.0)	104 (52.5)		150 (69.4)		820(62.0)		468(46.5)	
	G	73 (42.0)	94 (47.5)	0.30	66 (30.6)	0.03	502(38.0)	0.32	538(53.5)	0.005
	Genotype									
	TT	27 (31.0)	32 (32.3)		54 (50.0)		262(39.6)		100(19.9)	
	TG	47 (54.0)	40 (40.4)	0.077	42 (38.9)	0.028	296(44.4)	0.23	268(53.3)	0.014
	GG	13 (14.9)	27 (27.3)		12 (11.1)		103(15.6)		135(26.8)	
CXCR6 rs2234355	**Allele**									
	G	93 (53.4)	115 (58.1)		86 (39.8)		673(50.9)		1001(99.5)	
	A	81 (46.6)	83 (41.9)	0.40	130(60.2)	0.008	649(49.1)	0.57	5(0.5)	<0.0001
	Genotype									
	GG	29 (33.3)	35 (35.3)		16(14.8)		177(26.8)		498(99.0)	
	GA	35 (40.2)	45 (45.5)	0.49	54(50.0)	0.009	319(48.2)	0.32	5(1.0)	<0.0001
	AA	23 (26.4)	19 (19.2)		38(35.2)		165(25.0)		0(0)	
CXCR6 rs2234358	**Allele**									
	G	75 (43.1)	72 (36.4)		53(24.5)		416(31.5)		534(53.1)	
	T	99 (56.9)	126 (63.6)	0.20	163(75.5)	<0.0001	906(68.5)	0.003	472(46.9)	0.02
	Genotype									
	GG	17 (19.5)	14 (14.2)		4(3.7)		69(10.4)		145(28.8)	
	GT	41 (47.1)	44 (44.4)	0.43	45(41.7)	0.0003	278(42.1)	0.009	244(48.5)	0.054
	TT	29 (33.3)	41 (41.4)		59(54.6)		314(47.5)		114(22.7)	

[a]: Black South Africa Healthy Controls; [b]: Data from the 1000 genomes project; Luhya (LWK): from Kenya; Yoruba (YRI): from Nigeria, EUR: Total European; [c]: For the genotypic comparisons, 2x3 chi-squared tests were used to compare the genotypic distribution between populations, hence a single P value; Grey shaded P values indicate significant differences (P<0.05).

3.3. The role of *CCR5* (rs746492 and rs553615728) and *CXCR6* (rs2234355 and rs2234358) variants in HIV-1 control

An in-house real-time C_T shift assay was used to genotype the *CCR5* rs553615728 and *CXCR6* (rs2234355 and rs2234358) SNPs, and a predesigned Taqman assay used for the *CCR5* rs746492 to determine the prevalence and contribution of these variants on HIV-1 disease control in the 287 HIV-1 infected individuals enrolled in this cohort.

Data from a sub-group of HIV-negative individuals (N=87) termed HCs, (mentioned under section 2.1) in whom population data was available for all the SNPs under study was used for comparative purposes.

3.3.1. Validation of the Taqman and C_T Shift assay for *CCR5* (rs746492) SNP

Two different assays were employed for genotyping the *CCR5* rs746492 SNP in select individuals due to the fact that 16/287 samples that were initially tested using the predesigned Taqman assay resulted in indeterminate results. Thus, an in-house C_T Shift assay for the rs746492 SNP was subsequently designed and validated using 12 samples that were conclusively genotyped using the Taqman assay (Figure 3.4). The C_T shift assay results using these same 12 samples are shown in Figure 3.5. After the verification of the C_T shift assay it was successfully used to test the 16 samples that initially resulted in indeterminate results.

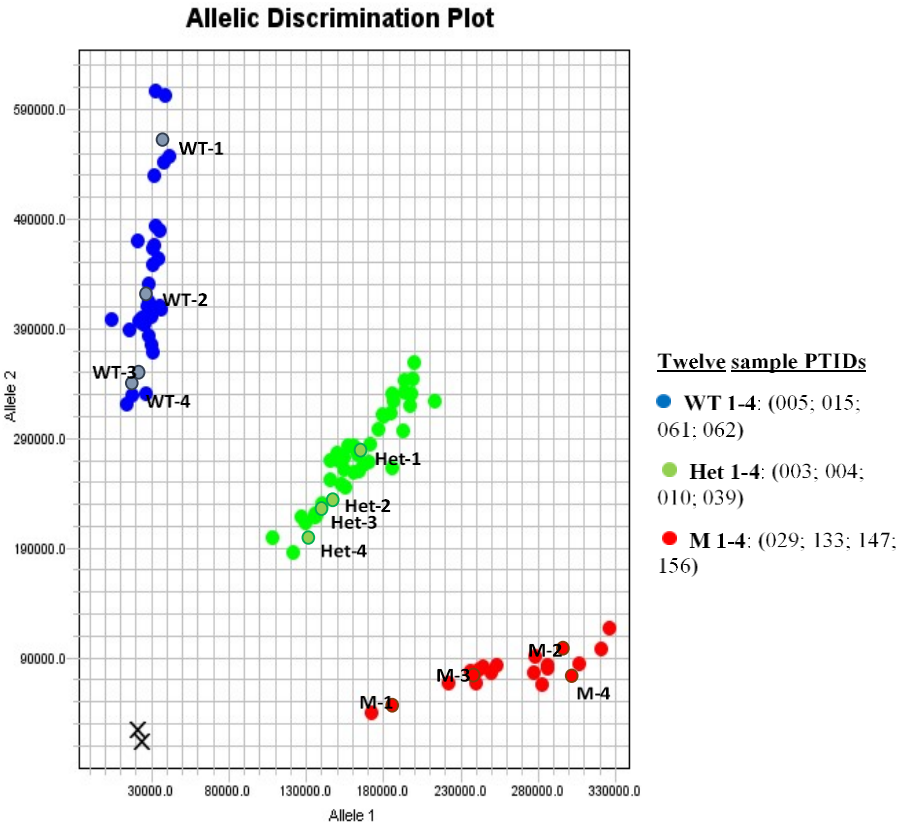


Figure 3.4. An allele discrimination plot indicating 12 samples used to validate the C_T shift assay for the *CCR5* rs746492 SNP. PTID: participant identity.

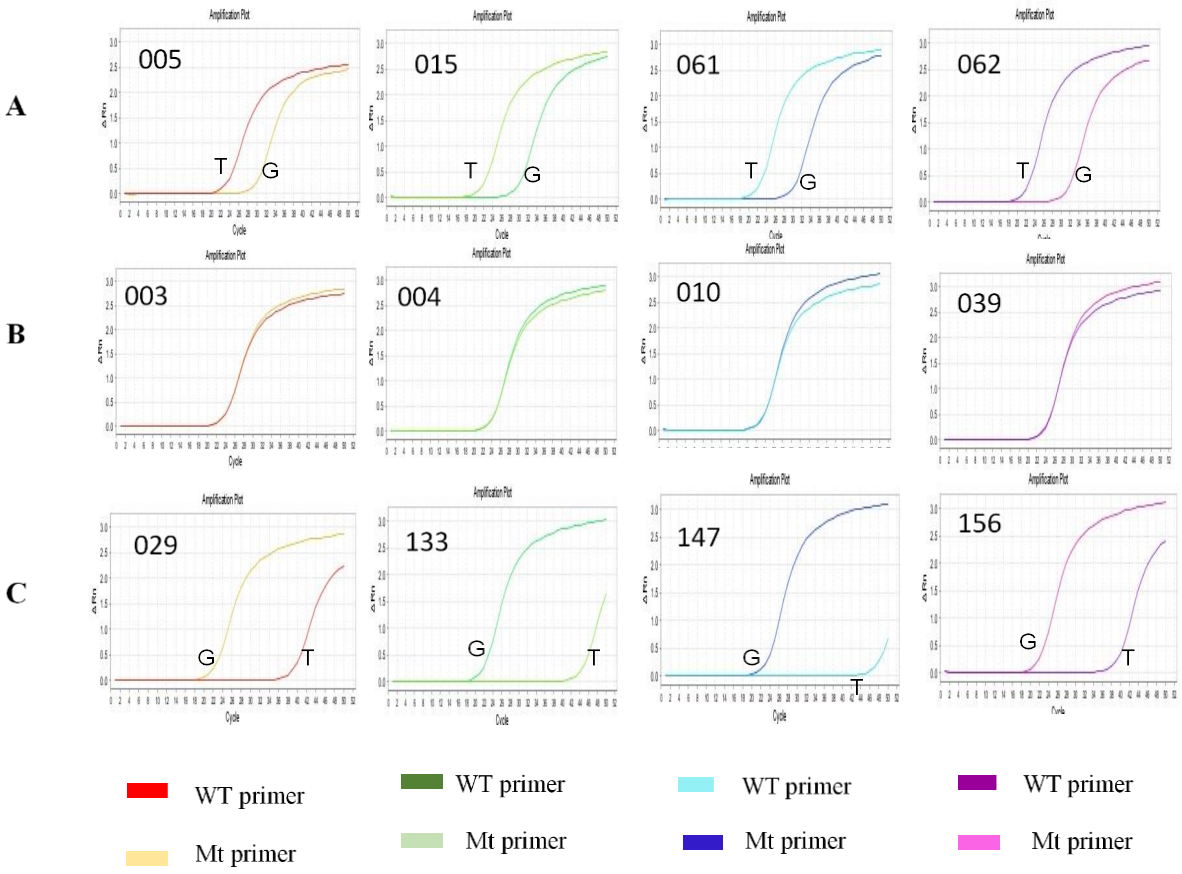


Figure 3.5. C_T shift results for the 12 samples validating the Taqman assay for the *CCR5* rs746492 SNP: (A) homozygous wildtype (WT), (B) heterozygous (Het) and (C) homozygous mutants (M).

3.3.2. Comparisons of allelic and genotypic frequencies of *CCR5* and *CXCR6* SNPs between study groups

The allelic and genotypic representation of *CCR5* (rs746492 and rs553615728) SNPs as well as *CXCR6* (rs2234355 and rs2234358) SNPs for the HIV-1 groups and healthy controls are shown in Table 3.4.

Table 3.4. Number and frequency (%) of alleles and genotypes for the *CCR5* and *CXCR6* variants in the HIV-1 infected Controller group, Controller sub-groups and Progressors as well as the HIV-1 negative healthy controls (HCs)

SNP	Alleles & Genotypes	Healthy Controls (N=87)	Controllers (ECs + VCs) (N=87)	Progressors (N=200)	Controller sub-groups	
					VCs (N= 56)	ECs (N=31)
CCR5	**Allele**					
rs746492	T	101 (58.0)	97 (55.8)	237 (59.3)	61 (54.5)	36 (58.1)
	G	73 (42.0)	77 (44.3)	163 (40.8)	51 (45.5)	26 (41.9)
	Genotype					
	TT	27 (31.0)	26 (29.9)	67 (33.5)	17 (30.4)	9 (29.0)
	TG	47 (54.0)	45 (51.7)	103 (51.5)	27 (48.2)	18 (58.1)
	GG	13(14.9)	16 (18.4)	30 (15.0)	12 (21.4)	4 (12.9)
CCR5	**Allele**					
rs553615728	C	171 (98.3)	170 (97.7)	390 (97.5)	109 (97.3)	61 (98.4)
	T	3 (1.7)	4 (2.3)	10 (2.5)	3 (2.7)	1 (1.6)
	Genotype					
	CC	84 (96.6)	83 (95.4)	190 (95.0)	53 (94.6)	30 (96.8)
	CT	3 (3.4)	4 (4.6)	10 (5.0)	3 (5.4)	1 (3.2)
	TT	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
CXCR6	**Allele**					
rs2234355	G	93 (53.4)	90 (51.7)	231 (57.8)	55 (49.1)	35 (56.5)
	A	81 (46.6)	84 (48.3)	169 (42.3)	57 (50.9)	27 (43.6)
	Genotype					
	GG	29 (33.3)	24 (27.6)	69 (34.5)	13 (23.2)	11 (35.5)
	GA	35 (40.2)	42 (48.3)	93 (46.5)	29 (51.8)	13 (41.9)
	AA	23 (26.4)	21 (24.1)	38 (19.0)	14 (25.0)	7 (22.6)
CXCR6	**Allele**					
rs2234358	G	75 (43.1)	90 (51.7)	194 (48.5)	62 (55.4)	28 (45.2)
	T	99 (56.9)	84 (48.3)	206 (51.5)	50 (44.6)	34 (54.8)
	Genotype					
	GG	17 (19.5)	22 (25.3)	48 (24.0)	16 (28.6)	6 (19.4)
	GT	41 (47.1)	46 (52.9)	98 (49.0)	30 (53.6)	16 (51.6)
	TT	29 (33.3)	19 (21.8)	54 (27.0)	10 (17.9)	9 (29.0)

3.3.2.1. Comparisons of allelic frequencies of CCR5 and CXCR6 SNPs between study groups

Comparisons of allelic frequencies was undertaken between the total Controller group (ECs plus VCs), the Controller sub-groups separately (ECs and VCs alone), HCs and Progressors. The latter two groups serve as the two extremes of HIV infection i.e. no HIV disease and AIDS phase of infection.

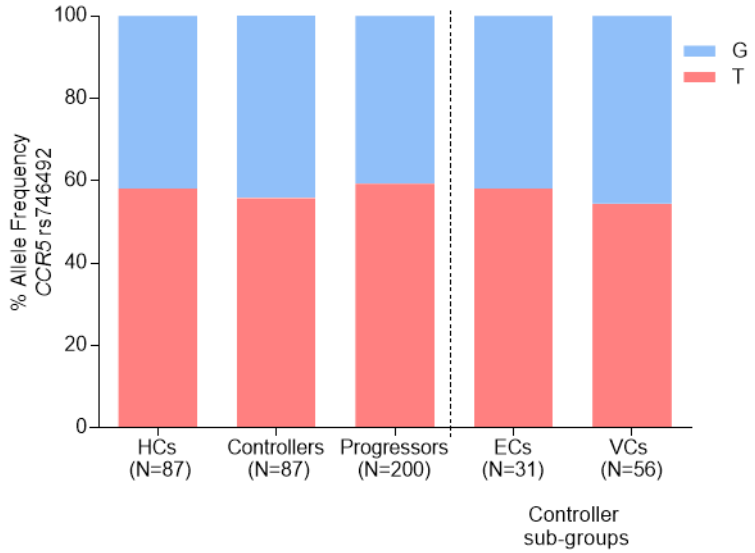
Figures 3.6 and 3.7 show the representation of the *CCR5* and *CXCR6* SNP alleles in all the study populations as well as the HCs, respectively. As can be seen in the bar graphs there is more variation between groups in the *CXCR6* variants, whereas *CCR5* representation seems more comparable between groups.

There was higher representation of the *CXCR6* rs2234355 A-allele in the VC group (50.9%) when compared to the Progressor group (42.3%) but this difference was not significant (P=0.11, OR=0.71, CI=0.46-1.07) (Table 3.4, 3.6).

Comparison of minor allele representation between the study groups and the HCs showed that only the *CXCR6* rs2234358 T-allele showed a strong trend of underrepresentation in the VCs compared to the HCs (VCs: 44.6% vs. HCs: 56.9%, P=0.05, OR=1.64, CI=1.01-2.64; Table 3.4; 3.5). The lower representation of this allele in the total Controller group (48.3%) compared to HCs (56.9%), although not significant (P=0.13, OR=1.41, CI=0.93-2.16) was due to the VCs since the ECs showed very similar representation to the HCs for this allele (Table 3.4; 3.5).

Comparison of minor allele representation between the Controller group and sub-groups to the Progressors for all the SNPs did not reveal any significant associations or trends (Table 3.6).

A



B

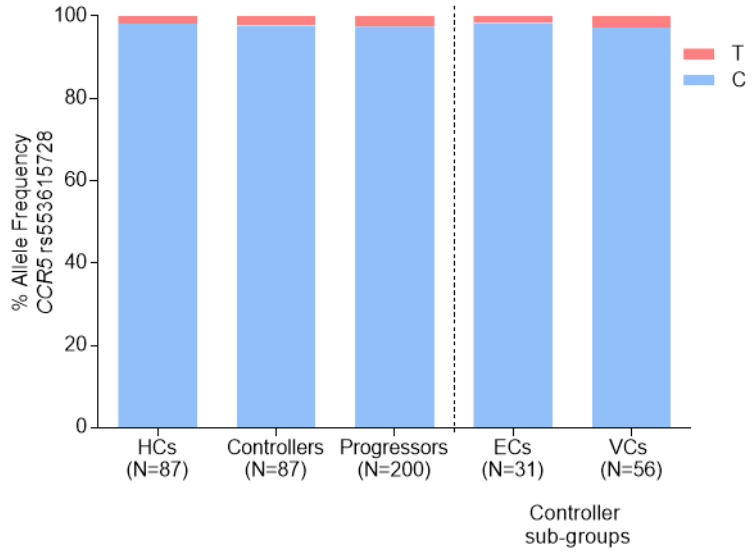


Figure 3.6. Bar graph showing allele frequency for, (A) *CCR5* rs746492 and (B) *CCR5* rs553615728 in the HCs (healthy controls), Controllers, Progressors, ECs (elite controllers) and the VCs (viraemic controllers).

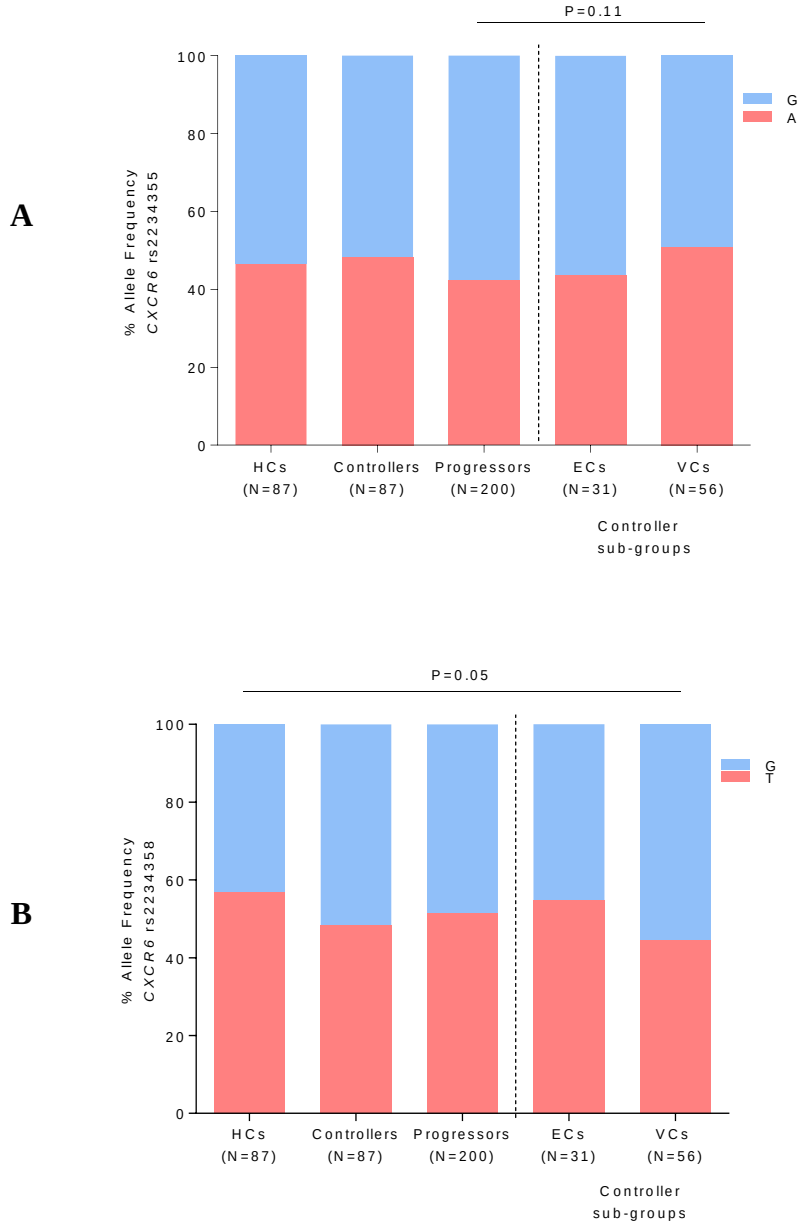


Figure 3.7. Bar graph showing allele frequency for, (A) *CXCR6* rs2234355 and (B) *CXCR6* rs2234358 in the HCs (healthy controls), Controllers, Progressors, ECs (elite controllers) and the VCs (viraemic controllers).

3.3.2.2. Comparisons of genotypic frequencies of *CCR5* and *CXCR6* SNPs between study groups

The *CCR5* rs746492 and rs553615728 genotypes showed no significant differences between study groups compared; due to the low representation of the rs553615728 SNP the *CCR5*

rs553615728 AA genotype was absent in all study groups. (Table 3.4; Figure 3.8; Table 3.5; Table 3.6).

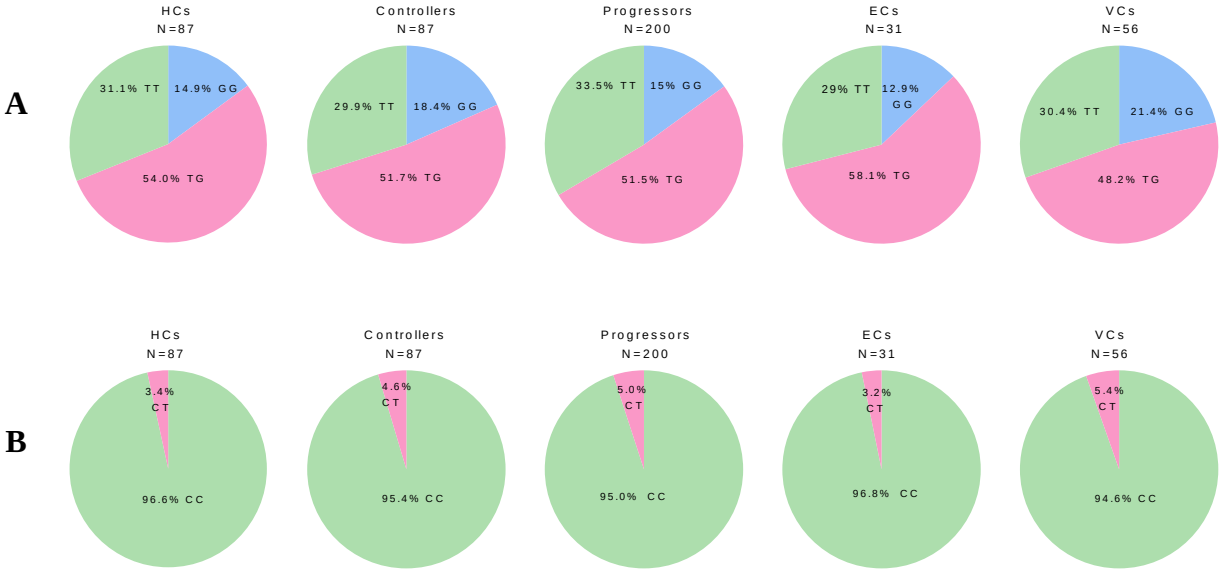


Figure 3.8. Pie charts showing genotypic representation (frequencies %) for (A) *CCR5* rs746492 and (B) *CCR5* rs553615728 SNPs for the various groups as shown. HC (healthy controls), Controllers, Progressors, EC (elite controllers) and VC (viraemic controllers).

Figure 3.9 shows the genotypic distribution of the two *CXCR6* SNPs (rs2234355 and rs2234358) in all the study populations as well as the reference HC group. As can be seen in Figure 3.9A, the *CXCR6* rs2234355 SNP showed similar distribution between groups, with no significant differences between study groups compared (Tables 3.5 and 3.6). Interestingly the VC group did have higher heterozygosity (GA) representation (51.8%) compared to both the HCs (40.2%) and the ECs (41.9%).

Similarly, there were no significant differences in over- or underrepresentation of the *CXCR6* rs2234358 genotypes between HCs, Controllers and Progressors, despite markedly less representation of minor allele homozygosity (TT) in VCs compared to HCs and less so compared to Progressors (VCs: 17.9% vs. HCs: 33.3%, P=0.05, OR=0.37, CI=0.14-0.99; vs. Progressors: 27.0%, P=0.19, OR=0.56, CI=0.23-1.34; Tables 3.4, 3.5 and 3.6).

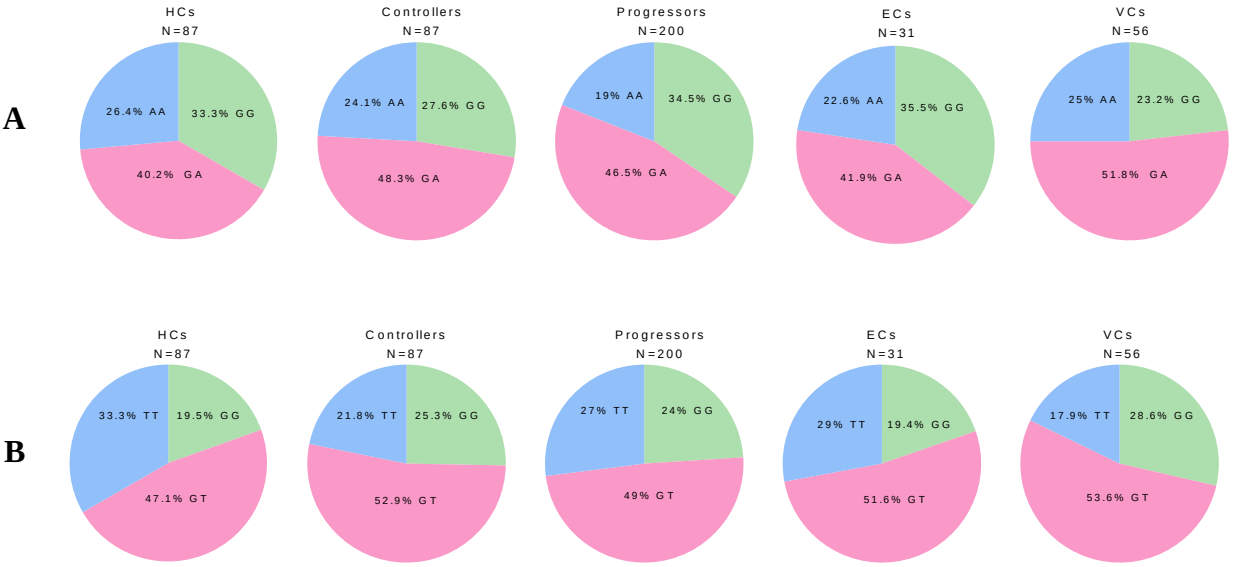


Figure 3.9. Pie charts showing genotypic representation (frequencies; %) for (A) *CXCR6* rs2234355 and (B) *CXCR6* rs2234358 SNPs for the various groups as shown. HC (healthy controls), Controllers, Progressors, VC (viraemic controllers) and EC (elite controllers).

Table 3.5. Comparisons of the *CCR5* and *CXCR6* SNP allelic and genotypic frequencies in Controllers, Controller sub-groups and Progressors to Healthy Controls (HCs)

Gene	SNP	Genotype	Controllers vs. HCs			Progressors vs. HCs			ECs vs. HCs			VCs vs. HCs		
			OR	CI	P	OR	CI	P	OR	CI	P	OR	CI	P
CCR5	rs746492	**Allele**												
		G	0.91	0.61-1.39	0.75	1.05	0.73-1.51	0.85	1.00	0.56-1.80	1.00	0.86	0.53-1.40	0.63
		Genotype												
		TG	0.99	0.51-1.95	0.99	1.13	0.64-1.99	0.67	1.15	0.45-2.91	0.77	0.91	0.42-1.97	0.82
		GG	1.28	0.52-3.17	0.60	1.08	0.49-2.37	0.86	0.92	0.24-3.56	0.91	1.47	0.54-3.95	0.45
		GT+GG	1.06	0.55-2.01	1.00	1.12	0.65-1.92	0.78	1.10	4.45-2.70	1.00	1.03	0.50-2.14	1.00
	rs53615728	**Allele**												
		T	1.47	0.32-6.69	0.69	1.34	0.36-4.99	0.71	2.10	0.21-20.70	0.65	1.25	0.25-6.36	1.00
		Genotype												
		CT	0.74	0.16-3.41	1.00	0.68	0.18-2.53	0.76	1.07	0.11-10.70	1.00	0.63	0.12-3.24	0.68
CXCR6	rs2234355	**Allele**												
		A	0.93	0.61-1.42	0.83	1.19	0.83-1.70	0.36	1.13	0.63-2.02	0.77	0.84	0.52-1.35	0.54
		Genotype												
		GA	1.45	0.72-2.93	0.30	0.90	0.50-1.60	0.71	0.98	0.38-2.51	0.97	1.85	0.82-4.19	0.14
		AA	1.10	0.49-2.46	0.81	1.44	0.73-2.83	0.29	0.80	0.27-2.40	0.69	1.36	0.53-3.45	0.52
		GA+AA	1.31	0.69-2.51	0.51	1.05	0.62-1.79	0.8	0.91	0.38-2.15	1.00	1.65	0.77-3.55	0.26
	rs2234358	**Allele**												
		T	1.41	0.93-2.16	0.13	1.24	0.87-1.78	0.24	1.09	0.61-1.95	0.88	1.64	1.01-2.64	0.05
		Genotype												
		GT	0.87	0.41-1.85	0.71	1.18	0.61-2.29	0.62	1.11	0.37-3.31	0.86	0.78	0.34-1.78	0.55
		TT	0.51	0.21-1.19	0.12	1.52	0.74-3.10	0.29	0.88	0.27-2.90	0.83	0.37	0.14-0.99	0.05
		GT+TT	0.72	0.35-1.47	0.47	1.30	0.70-2.42	0.45	1.01	0.36-2.85	1.00	0.61	0.28-1.33	0.23

*****WTgenotypes which were used as reference (not shown); Grey shaded value indicate relationship that show trends of under- or over representation (0.1>P≥0.05)

Table 3.6. Comparisons of the *CCR5* and *CXCR6* SNP allelic and genotypic frequencies in Controller, VC and EC groups to the Progressor group

Gene	SNP	Genotype	Controllers vs. Progressors			ECs vs. Progressors			VCs vs. Progressors		
			OR	CI	P	OR	CI	P	OR	CI	P
CCR5	rs746492	**Allele**									
		G	0.87	0.60-1.24	0.46	0.95	0.55-1.64	0.89	0.82	0.54-1.25	0.39
		Genotype									
		TG	1.37	0.64-2.93	0.41	0.99	0.28-3.48	0.99	1.58	0.67-3.71	0.30
		GG	1.13	0.64-1.10	0.69	1.30	0.55-3.07	0.55	1.03	0.52-2.04	0.93
		GT+GG	1.18	0.69-2.04	0.59	1.23	0.54-2.82	0.69	1.16	0.61-2.19	0.75
	rs53615728	**Allele**									
		T	1.09	0.34-3.52	1.00	1.56	0.20-12.44	1.00	0.93	0.25-3.44	1.00
		Genotype									
		CT	1.09	0.33-3.58	1.00	1.58	0.20-12.78	1.00	0.93	0.25-3.50	1.00
CXCR6	rs2234355	**Allele**									
		A	0.78	0.55-1.12	0.20	0.95	0.55-1.63	0.89	0.71	0.46-1.07	0.11
		Genotype									
		GA	1.30	0.72-2.34	0.39	0.88	0.37-2.70	0.77	1.66	0.81-3.42	0.17
		AA	1.59	0.78-3.22	0.20	1.16	0.41-3.23	0.78	1.96	0.83-4.59	0.12
		GA+AA	1.38	0.80-2.40	0.27	0.96	0.43-2.11	1.00	1.74	0.88-3.46	0.14
	rs2234358	**Allele**									
		T	1.14	0.80-1.62	0.53	0.87	0.51-1.50	0.68	1.32	0.86-2.01	0.24
		Genotype									
		GT	1.02	0.55-1.89	0.94	1.31	0.48-3.55	0.60	0.92	0.46-1.85	0.81
		TT	0.77	0.37-1.59	0.48	1.33	0.44-4.02	0.61	0.56	0.23-1.34	0.19
		GT+TT	0.93	0.52-1.67	0.88	1.31	0.51-3.40	0.65	0.79	0.41-1.53	0.60

3.4. Investigation of *CCR5* and *CXCR6* variants and markers of disease progression: viral load and CD4$^+$ T-cell count

3.4.1. Effect of *CCR5* and *CXCR6* genotypes on viral load in the Progressor and VC groups

We explored the association between *CCR5* and *CXCR6* genotypes with viral load, a marker of HIV-1 disease progression in the Progressor (N=200) and VC (N=56) groups.

No significant differences were seen in VL between individuals possessing either of the three *CCR5* rs746492 genotypes (Figure 3.10A). For the *CCR5* rs553615728 SNP, a much rarer variant, most individuals were homozygous for the major allele (CC) with less representation of the CT genotype and no representation of the TT genotype. There was no significant difference in VL between individuals with the CC vs. individuals with the CT genotype for this *CCR5* variant (Figure 3.10B).

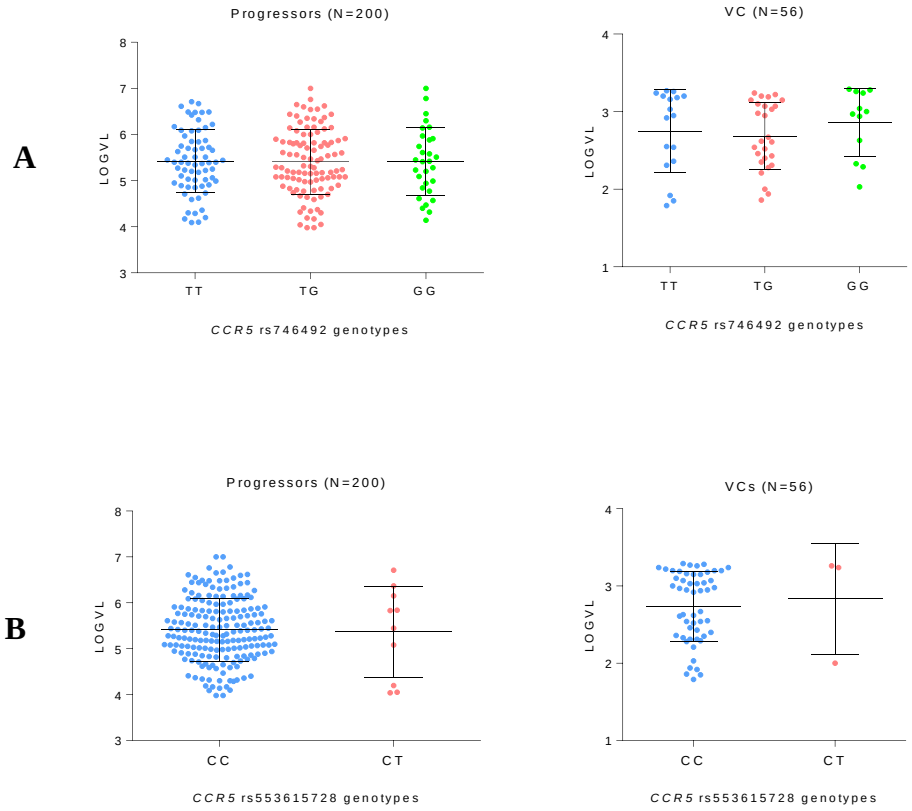


Figure 3.10. Comparison of viral load (log_{10} RNA copies/ml) in the Progressor and VC groups for: (A) *CCR5* rs746492 SNP genotypes and (B) *CCR5* rs553615728 SNP genotypes. Medians and interquartile ranges are shown, (Mann Whitney U-test).

Heterozygosity (GA) for the *CXCR6* rs2234355 SNP showed a weak trend (P=0.097) of higher representation in individuals with higher VL when compared to individuals homozygous for the major allele (GG) in the Progressor group but no significant difference in VL was observed in the VC group (P=0.75) (Figure 3.11A). Analysis of VL and this variant in the dominant mode (i.e. GA+AA) showed that VL was still marginally higher in the GA+AA group with a weaker association in both Progressors (P=0.17) and VCs (P=0.16) (Figure 3.11B).

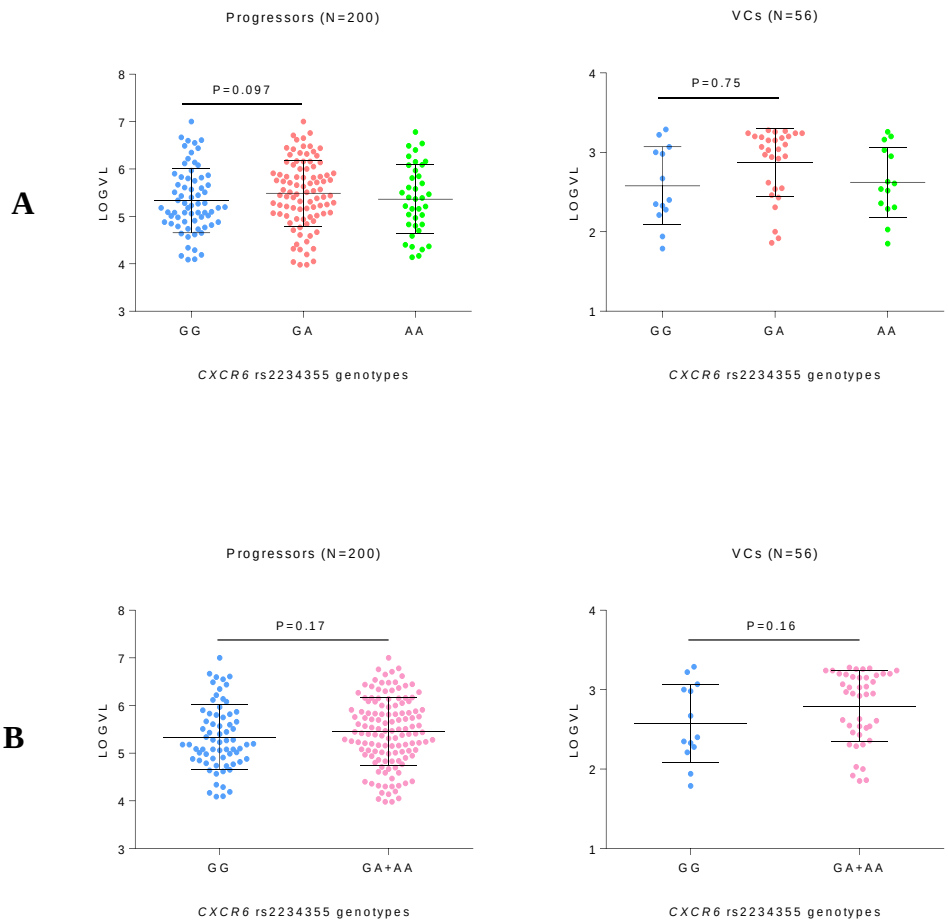


Figure 3.11. Comparison of viral load (log_{10} RNA copies/ml) in the Progressor and VC groups for: (A) *CXCR6* 2234355 SNP genotypes and (B) *CXCR6* rs2234355 SNP in the dominant mode (GA+AA). Medians and interquartile ranges are shown, (Mann Whitney U-test).

Interestingly when the effect of VL was evaluated for the *CXCR6* rs2234358 SNP, the possession of minor allele homozygosity (TT) was significantly associated with higher VL

compared to heterozygous (GT) individuals (P=0.0003) in Progressors but no significant differences was seen in VCs (P=.080) Figure 3.12A). This significant association was maintained when genotypes harbouring the major allele were grouped (i.e. GG+GT) and compared to individuals possessing the TT genotype in Progressors (P=0.0013) Figure 3.12B.

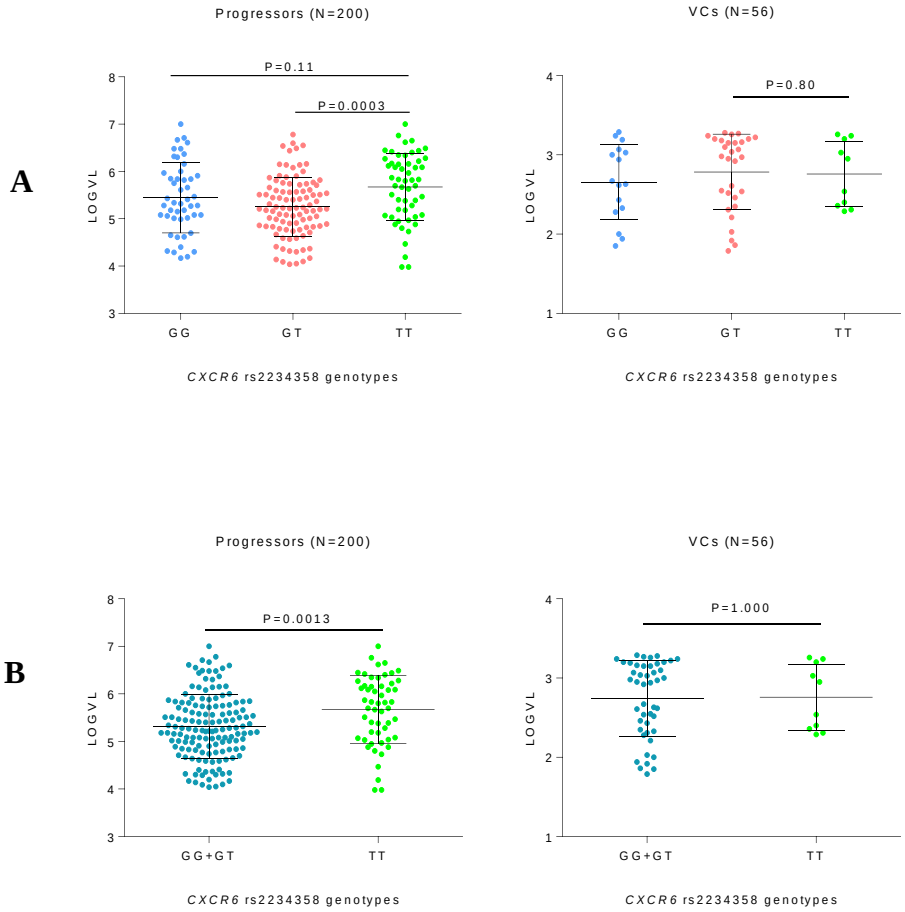


Figure 3.12. Comparison of viral load (log_{10} RNA copies/ml) in the Progressor and VC groups for: (A) *CXCR6* rs2234358 SNP genotypes and (B) *CXCR6* rs2234358 SNP by grouping individuals harbouring the major allele (i.e. GG+GT). Medians and interquartile ranges are shown, (Mann Whitney U-test).

3.4.2. Effect of *CCR5* and *CXCR6* SNP genotypes on $CD4^+$ T-cell counts

CCR5 and *CXCR6* SNP genotypes from Progressors (N=200), Controllers (N=87), VCs (N=56) and ECs (N=31) were stratified against their $CD4^+$ T-cell counts in order to determine if the respective variants had any impact on this marker of disease progression.

There was no significant difference between any of the genotypes with respect to the *CCR5* rs746492 SNP and $CD4^+$ T-cell counts in the Progressor, Controller, EC and VC groups (Figure 3.13A). Similarly, no significant differences in $CD4^+$ T-cell counts were observed between the GG and GA genotypes of the *CCR5* rs553615728 SNP in Progressors, Controllers and ECs. Although $CD4^+$ T-cell counts in VCs harbouring the GG genotype were significantly lower than individuals with the GA genotype (P=0.03), these results should be treated with caution due to small numbers (N=3) in those with the CT genotype (Figure 3.13B). There was no significant difference between any of the genotypes and $CD4^+$ T-cell counts with respect to the *CXCR6* rs2234355 or rs2234358 SNPs in all the groups that were analysed (Figures 3.14A and 3.14B).

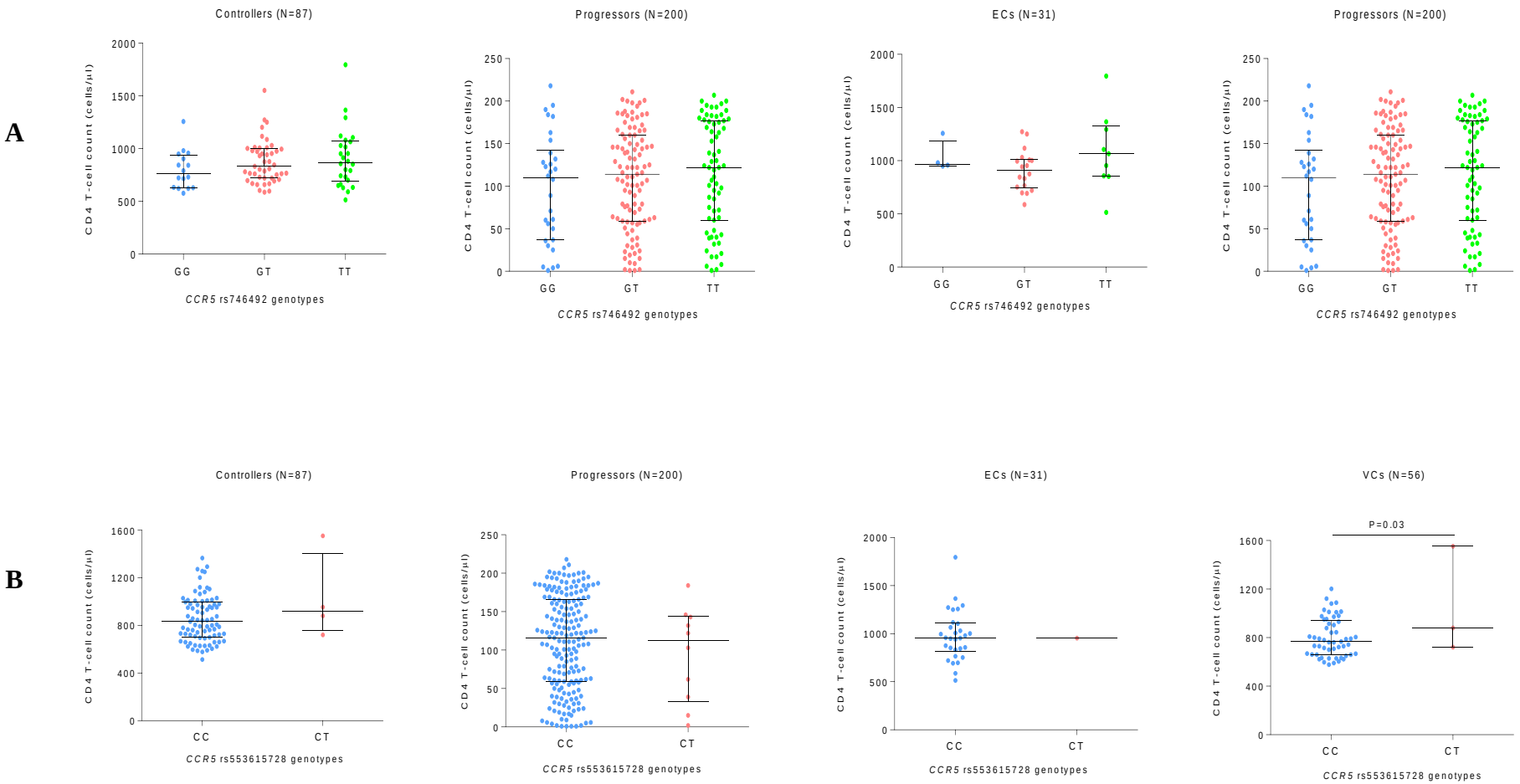


Figure 3.13. Comparison of $CD4^+$ T-cell count between respective genotypes for: (A) *CCR5* rs746492 and (B) *CCR5* rs553615728 SNPs in the Controller, Progressor, EC and VC groups. Medians and interquartile ranges are shown, Kruskal-Wallis H test**.**

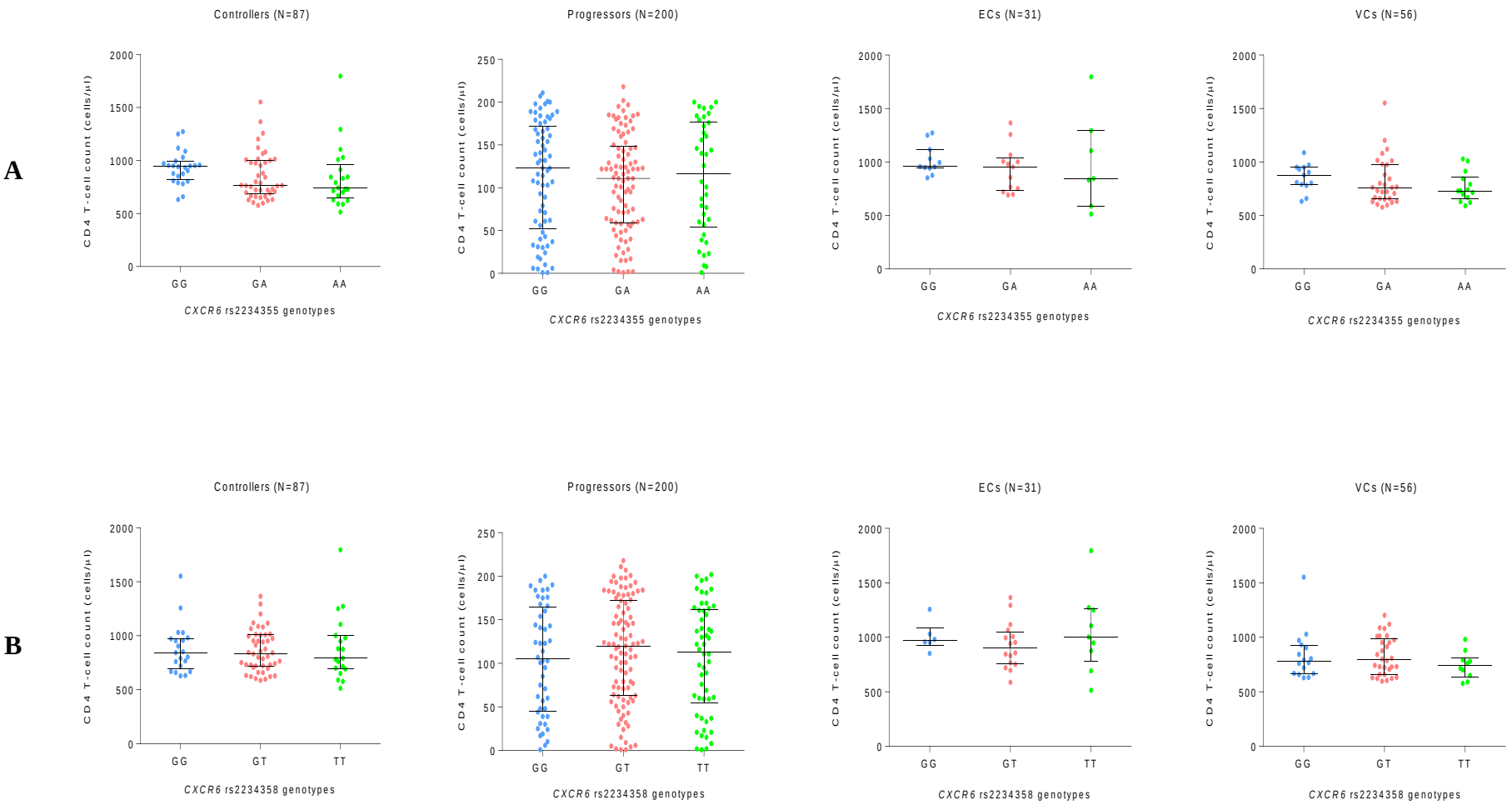


Figure 3.14. Comparison of $CD4^+$ T-cell count between respective genotypes for: (A) *CXCR6* rs2234355 and (B) *CXCR6* rs2234358 SNPs in the Controller, Progressor, EC and VC groups. Medians and interquartile ranges are shown, Kruskal-Wallis H test.

3.5. Analysis of the CCR5 and CXCR6 allelic and genotypic frequencies on HIV-1 control taking gender differences into account

As shown in Table 3.2, there was a significantly higher number of females recruited in the Controller group compared to the Progressors (85.1% vs. 62.5%; P<0.001). In order to determine whether the gender bias was affecting the results obtained from the genetic association analyses, we assessed the effects of the *CCR5* and *CXCR6* alleles and genotypes on HIV-1 control by taking gender into account, despite small sample sizes particularly of the EC and VC groups.

There was no significant difference in the *CCR5* rs746492 allelic distribution between males and females in the Controllers (P=1.00), Progressors (P=0.92), ECs (P=0.54) and VCs (P=0.78) (Figure 3.15).

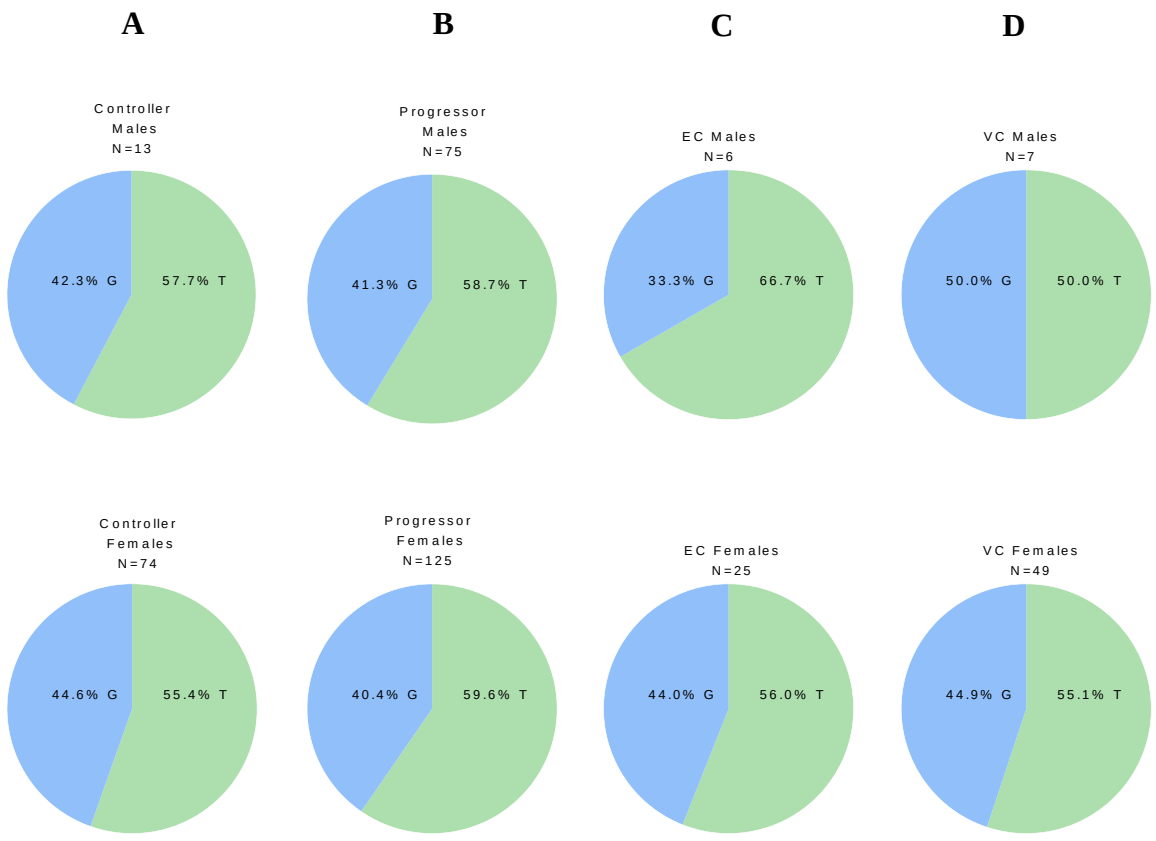


Figure 3.15. Pie charts showing representation of the *CCR5* rs746492 alleles in (A) Controller, (B) Progressor, (C) EC and (D) VC groups according to gender.

The *CCR5* rs553615728 alleles were not significantly differentially distributed between males and females in Controllers (P=0.11), Progressors (P=1.00) and ECs (P=1.00); Figure 3.16A, B and C). However, the *CCR5* rs553615728 minor allele (T) was significantly overrepresented in males (14.3%) compared to females (1.0%) in the VCs (P=0.04) (Figure 3.16D). Given the small number of VC males (N=7), this result needs to be treated with caution.

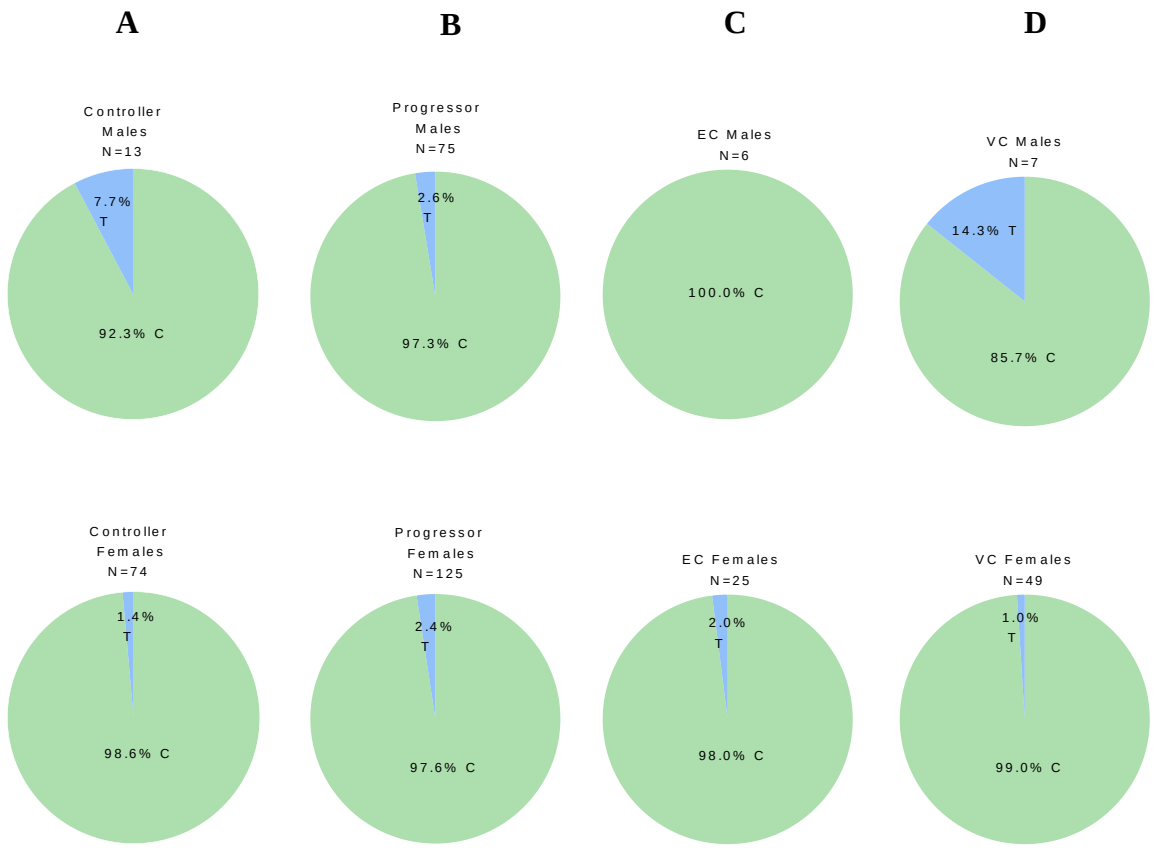


Figure 3.16. Pie charts showing representation of the *CCR5* rs553615728 alleles in (A) Controller, (B) Progressor, (C) EC and (D) VC groups according to gender.

The *CXCR6* rs2234355 minor allele (A) was comparable between males and females in the Controllers (50.0% vs. 48.0%, P=1.00) and Progressors (45.3% vs. 40.4%, P=0.35) (Figure 3.17A and B). However, it was more prevalent in males compared to females in the EC group (58.3% vs. 40.0%, P=0.34) while more prevalent in females compared to males in the VC group (52.0% vs. 42.9%, P=0.58) (Figure 3.17C and D).

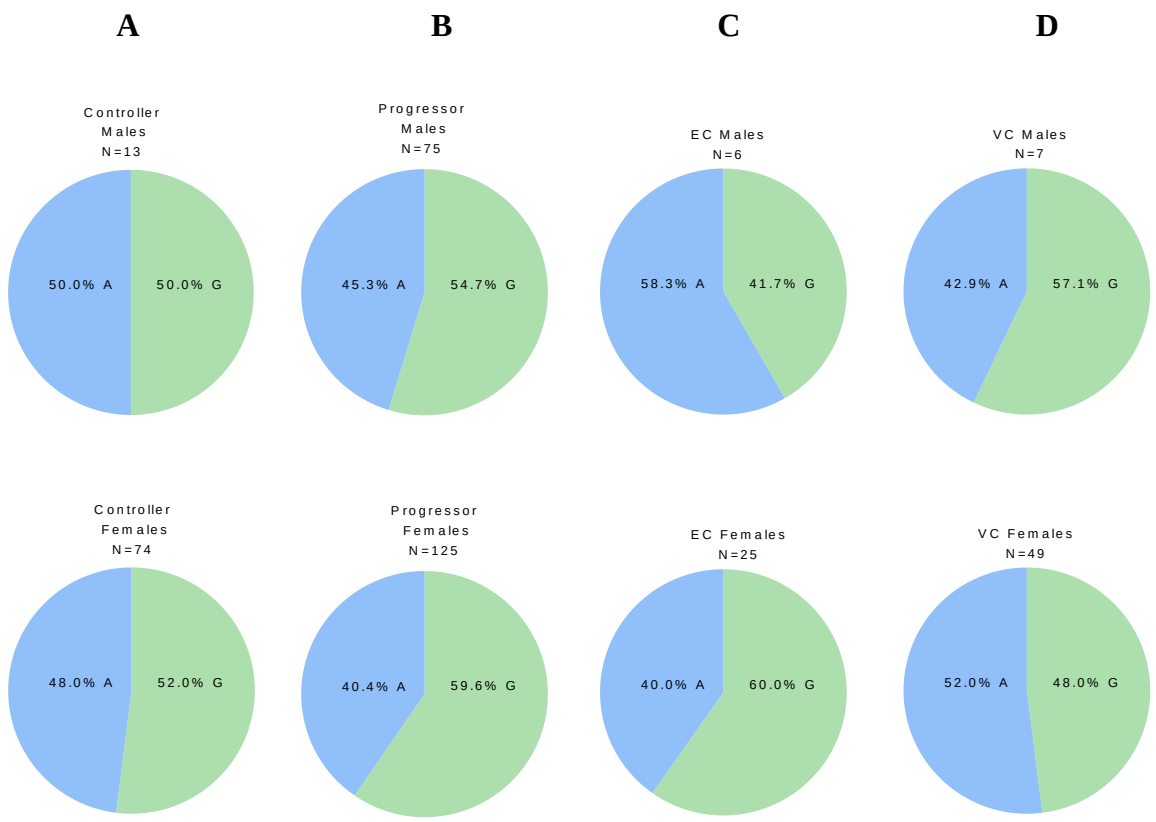


Figure 3.17. Pie charts showing representation of the *CXCR6* rs2234355 alleles in (A) Controller, (B) Progressor, (C) EC and (D) VC groups according to gender.

The *CXCR6* rs2234358 minor allele (T) was not significantly differentially distributed between males and females in the Progressor, EC and VC groups (P=0.18, P=0.20 and P=0.39, respectively) (Figure 3.18B, C and D). However, the *CXCR6* rs2234358 minor allele (T) showed strong trends of overrepresentation (P=0.09) in males compared to females (65.4% vs. 45.3%), in the Controllers (Figure 3.18A).

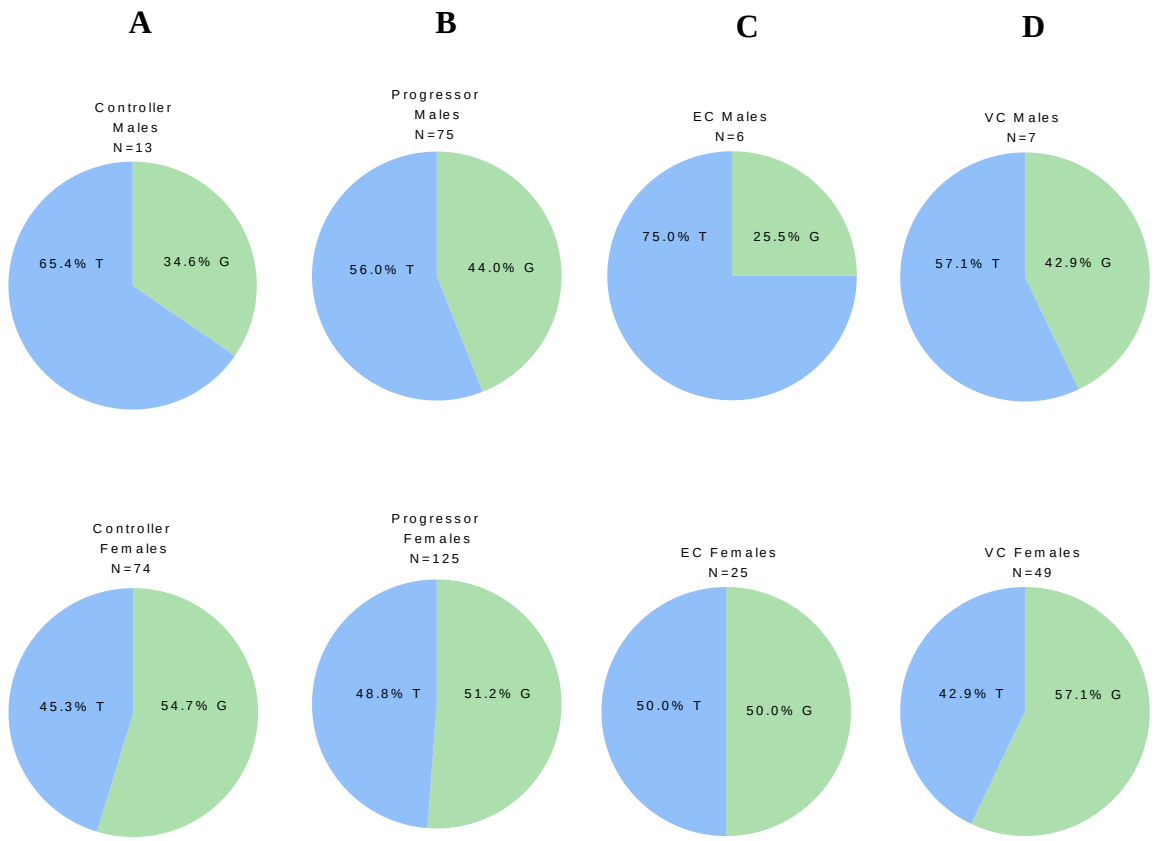


Figure 3.18. Pie charts showing representation of the *CXCR6* rs2234358 alleles in (A) Controller, (B) Progressor, (C) EC and (D) VC groups according to gender.

There was no significant difference for the *CCR5* rs746492 SNP genotypic distribution with respect to males vs. females in Progressor group (P=0.94; Figure 3.19B). Similarly, all three genotypes were comparable between males and females in the Controller (P=1.00) (Figure 3.19A) and VC groups (P=0.88) (Figure 3.19D). There was however no representation of minor allele homozygosity (GG) in males in the EC group, with heterozygosity (TG genotype) overrepresented in males compared to females (66.7% vs. 56.0%) and the TT genotype comparable between genders (33.3% vs. 28.0%). Given the small number of EC males (N=6) compared to EC females (N=25), these differences were not statistically significant in this group (P=0.69) (Figure 3.19C).

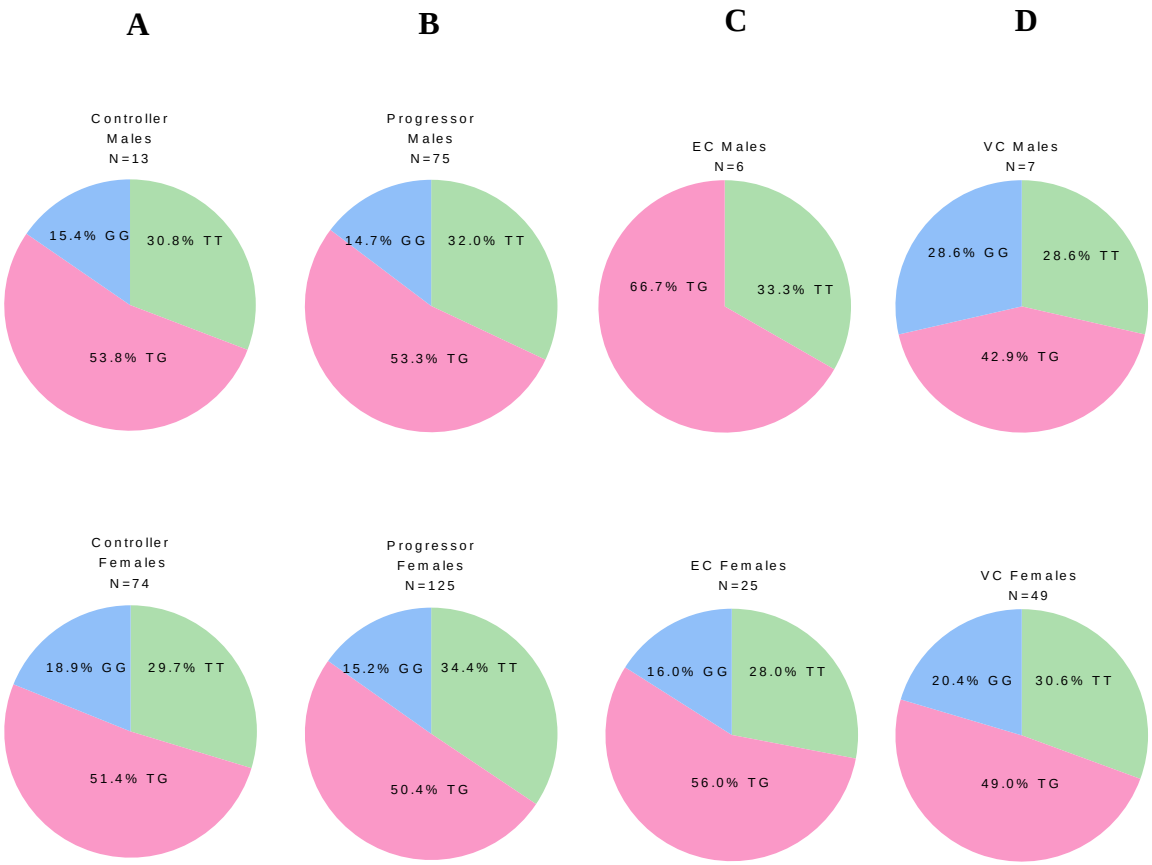


Figure 3.19. Pie charts showing representation of the *CCR5* rs746492 genotypes in (A) Controller, (B) Progressor, (C) EC and (D) VC groups according to gender.

The *CCR5* rs553615728 genotypes were not significantly differentially distributed between males and females in the Controllers (P=0.10), Progressors (P=1.00) and ECs (P=1.00) (Figure 3.20A, B and C, respectively). However, the *CCR5* rs553615728 genotypes were significantly differentially distributed between genders in the VCs (P=0.04) with heterozygosity (GA) being overrepresented in males (28.6%) compared to females (2.0%) and the GG genotype thus more prevalent in females (98.0%) compared to males (71.4%) (Figure 3.20D). Given the small number of VC males (N=7), this result needs to be treated with caution.



Figure 3.20. Pie charts showing representation of the *CCR5* rs553615728 genotypes in (A) Controller, (B) Progressor, (C) EC and (D) VC groups according to gender.

With regard to the *CXCR6* rs2234355 SNP, in Controllers, although genotypes differed between gender, the GG genotype was underrepresented in males compared to females (15.4% vs. 44.6%), similarly the AA genotype was underrepresented in males (15.4% vs. 25.7%) while the GA genotype was more prevalent in males compared to females (69.2% vs. 44.6%), this difference was not significant (P=0.34; Figure3.21A). In the Progressor group, the GG genotype was underrepresented in males compared to females (25.3% vs. 40.0%) whereas heterozygosity (GA) was more prevalent in males (58.6% vs. 39.2%), and AA genotype was more comparable between males and females (16.0% vs. 20.8%). Thus, overall the *CXCR6* rs2234355 genotypes were significantly differentially represented between genders (P=0.03; Figure 3.21B) in Progressors. In the VCs, although there was a distinct difference in genotype distribution between genders with absence of the AA genotype in males compared to females (0.0% vs. 28.6%), an overrepresentation of heterozygosity (85.7% vs. 46.9%) and underrepresentation of the GG genotype (14.3% vs. 24.5%), due to the small number of males

in the VC group this comparison did not reach significance (P=0.15; Figure 3.21D). In the EC group, the GG genotype was underrepresented in males compared to females (16.7% vs. 40.0%), GA was more prevalent in the males (50.0% vs. 40.0%) and the AA genotype, unlike in the VC group, was overrepresented in the males compare to females (33.3% vs. 20.0%). However, given the small sample numbers, these genotypes were not significantly different between the genders (P=0.61; Figure3.21C).

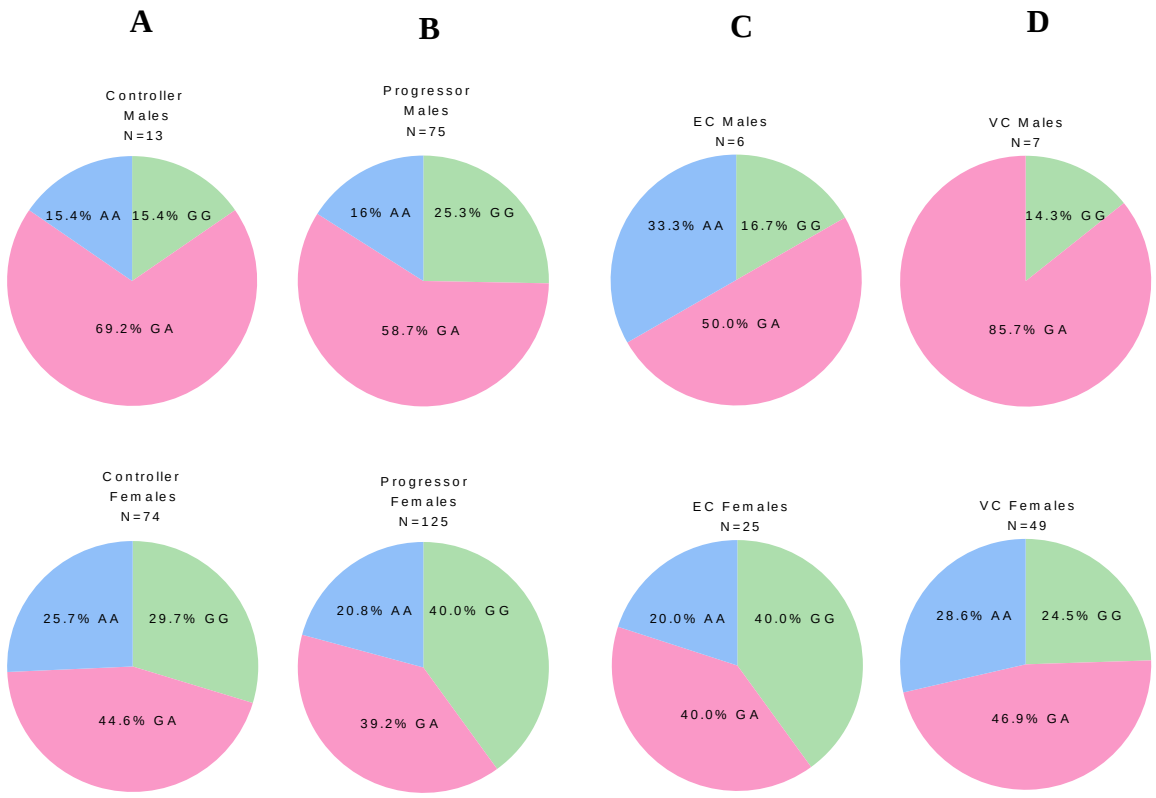


Figure 3.21. Pie charts showing representation of the *CXCR6* rs2234355 genotypes in (A) Controller, (B) Progressor, (C) EC and (D) VC groups according to gender.

When the *CXCR6* rs2234358 SNP genotype distribution was assessed between males and females in Controllers, GG was underrepresented in males compared to females (7.7% vs 28.4%), AA was overrepresented in males compared to females (38.5% vs. 18.9%) and GA genotype was comparable between males and females (53.8% vs. 52.7%), these differences were not significant (P=0.14; Figure 3.22A). In the Progressor group, GG was less prevalent in males (20.0% vs. 26.4%), TG was similarly represented between males and females (48.0%

vs. 49.6%) and the TT genotype was overrepresented in males (32.0% vs. 24.0%), however this variant's distribution was not significantly different between genders (P=0.40; Figure 3.22B). In the VC group, the GG genotype was less prevalent in males (14.3% vs. 30.6%), GT was similarly represented in males and females (57.1% vs. 53.1%) and the TT genotype was overrepresented in males compared to females (28.6% vs. 16.3%). Overall however, the representation of *CXCR6* rs334358 genotypes was not significant between genders in the VC group (P=0.57; Figure 3.22D). Similarly, there was no significant difference observed in the distribution of these genotypes in the EC group (P=0.34; Figure 3.22C), despite the GG genotype being absent in males vs. 24.0% representation in females and TT being overrepresented in males compared to females (50.0% vs. 24.0%), with similar TG representation (50.0% vs. 52.0%).

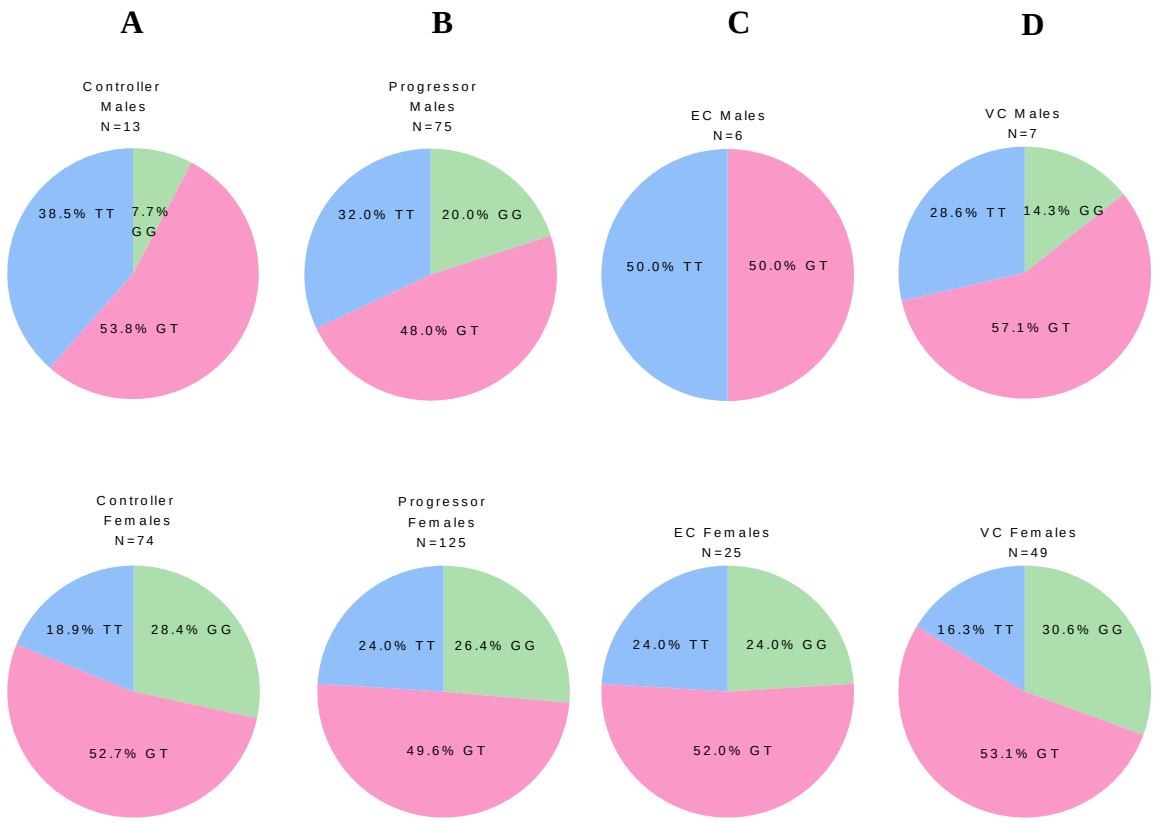


Figure 3.22. Pie charts showing representation of the *CXCR6* rs2234358 genotypes in (A) Controller, (B) Progressor, (C) EC and (D) VC groups according to gender.

3.5.1. Analysis of the *CXCR6* rs2234355 variant in viraemic HIV-1 control in the female gender groups

As shown in section 3.3.2.1 above (Table 3.4) for the *CXCR6* rs2234355 SNP, the frequency of the major allele (G) was higher in the Progressors (57.8%) compared to VCs (49.1%), whereas the minor allele (A) was underrepresented in Progressors (42.3%) compared to VCs (50.9%), however this was not a significant association (P=0.11, Table 3.6). Given the difference in gender distribution, overall the *CXCR6* rs2234355 genotypes were significantly differentially represented between genders (P=0.03) (section 3.5; Figure 3.21), we assessed the effect of this variant by exclusion of the smaller group (i.e. males) and analysed the effect of this variant in female VCs and Progressors. Figure 3.23 shows the comparison of allelic representation of this variant in female Progressors and VCs as well as comparison of the variant in the dominant mode (i.e. GG vs. GA+AA).

In females the *CXCR6* rs22343355 minor allele (A) showed a strong trend of overrepresentation in VC females compared to the Progressor females (VCs: 52.0% vs. Progressors: 40.4%, P=0.055, OR=0.62, CI=0.39-0.99) (Figure 3.23A) and comparison of this SNP in the dominant mode also showed a trend of higher representation in VC females compared to Progressor females (VCs: 75.5% vs. Progressors: 60.0%, P=0.078, OR=0.23, CI=0.49-1.02) (Figure 3.23B).

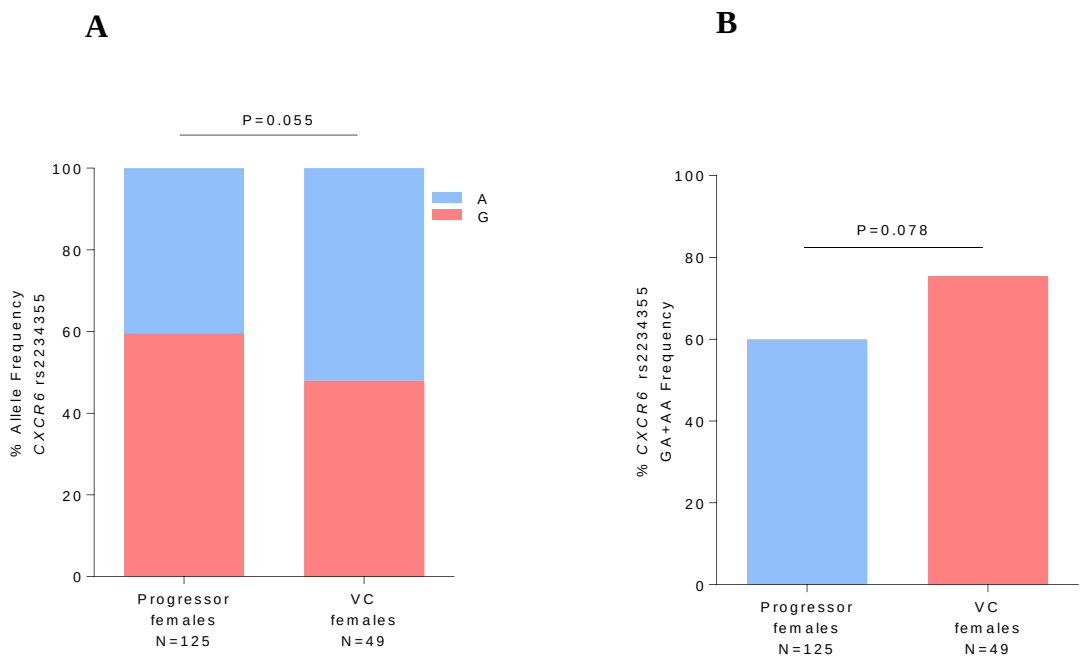


Figure 3.23. Bar graphs showing *CXCR6* rs2234355 SNP: (A) distribution and comparison of alleles in the VC and Progressor female groups and (B) comparison of the variant in the dominant mode (GA+AA) between VC and Progressor female groups.

3.6. The analysis of the combined effect of the *CCR5* and *CXCR6* SNPs

3.6.1. Analysis of the combined effect of the two *CXCR6* variants (rs2234355 and rs2234358 SNPs)

The *CXCR6* rs2234355 and rs2234358 SNPs have previously been associated with natural control in individuals with higher viraemia (VCs) compared to individuals progressing to HIV disease (Progressors) (Picton *et al*., 2017). Furthermore, we found an association of the rs2234358 TT genotype with the increased VL in the Progressors in this study (Figure 3.12). Since the individual *CXCR6* rs2234355 and rs2234358 SNPs alone did not reveal any significant associations when comparing ECs and VCs separately to Progressors in this cohort (Tables 3.5 and 3.6), as previously undertaken by Picton *et al*. (2017), we also analysed the combined effect of the two *CXCR6* SNPs to determine whether there was any association between the previously identified protective rs2234355 GA and deleterious rs2234358-TT genotypes of these two SNPs (Picton *et al*., 2017) in both the combined females and males as well as in the female group alone.

Figure 3.24A, shows that the combination of the *CXCR6* rs2234355 genotypes GA+AA (dominant mode) in the absence of the deleterious rs2234358-TT genotype (designated as: -358TT +355 GA+AA) was significantly overrepresented in the VC group compared to the Progressor and HC groups (VCs: 60.7% vs. Progressors: 44.0%, P=0.034; OR=0.51; CI=0.28-0.93; VCs: 60.7% vs. HCs: 36.8%, P=0.006; OR=0.38; CI=0.19-0.8) This significant difference between VCs and Progressors was more pronounced when we corrected for gender by comparing only the female groups. Female VCs were more significantly enriched for -358TT +355 GA+AA (61.2%) compared to Progressors (40.8%) (P=0.018; OR=0.44; CI= 0.22-0.86). Although female VCs had higher representation (40.8%) of the *CXCR6* rs2234355 genotype GA alone in the absence of the deleterious rs2234358-TT genotype (-358TT +355 GA) compared to Progressor females (27.2%), this association did not reach significance (P=0.101; OR=0.54; CI= 0.27-1.08) (Figure 3.24B).

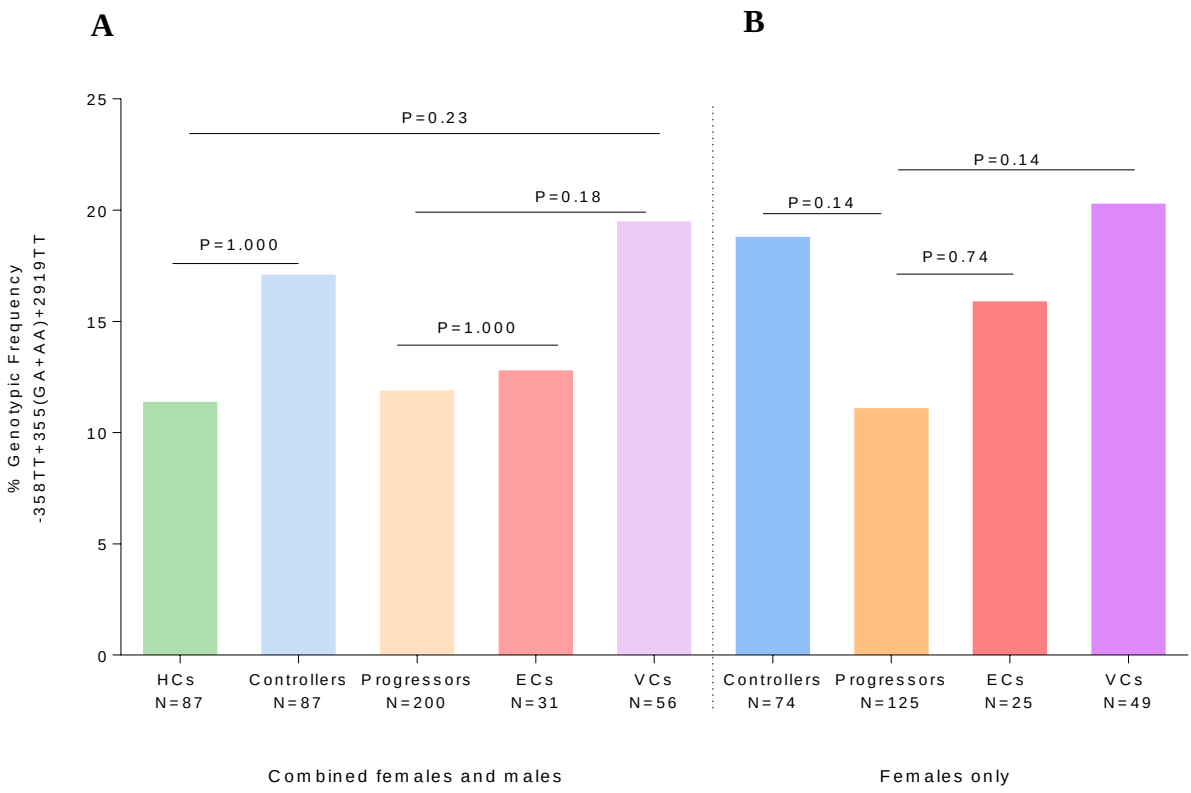


Figure 3.24. Bar graphs showing comparison of combinations of (A) the *CXCR6* rs2234355 (355 GA+AA) and rs2234358 (-358TT) genotypes in HC, Progressor, EC and VC groups (combined females and males) and (B) the *CXCR6* rs2234355 (355 GA) and rs2234358 (-358TT) in Progressor, EC and VC groups (females only).

3.6.2. Analysis of the combined effect of protective genotypes of *CCR5* (rs746492 and rs553615728) and *CXCR6* (rs2234355 and rs2234358)

Although no significant association was found in the individuals with *CCR5* (rs746492 and rs553615728) SNP genotypes with HIV-1 control in this study, it has been reported that *CCR5* rs746492 (TG+GG) (Koor *et al*., 2019) and *CCR5* rs553615728 (Gornalusse *et al*., 2015) can serve as deleterious and protective markers, respectively in HIV-1 control. We therefore analysed the combination of the protective genotypes of both CCR5 and CXCR6 co-receptors to determine their role in HIV-1 control in this study. Due to the low prevalence of the *CCR5* rs553615728 (CT and TT) genotypes in this cohort (Table 3.4), the *CCR5* rs553615728 SNP could not be included in this analysis.

The *CCR5* rs746492-TT designated as (+2919TT) genotype (protective) in combination with the *CXCR6* [-358TT+355(GA+AA)] (protective) was overrepresented in VCs compared to the Progressor and HC groups but this difference was not significant in combined females and males for each group (VCs: 19.6% vs. Progressors: 12.0%, P=0.18; OR=0.56;CI=0.25-1.22; VCs: 19.6% vs. HCs: 11.5%, P=0.23; OR=0.53; CI=0.21-1.35. There was no significant difference when EC, Controller, HC and Progressor groups were compared (ECs: 12.9% vs. Progressors: 12.0%, P=1.000; OR=0.92; CI=0.30-2.86); Controllers: 17.2% vs. Progressors: 12.0%, P=0.26; OR=0.65; CI=0.32-1.32; Progressors: 12.0% vs. HCs: 11.5%, P=1.000; OR=1.05; CI=0.48-2.30 as well as ECs vs HCs: ECs 12.9:% vs. HCs:11.5%, P=1.000; OR=0.88; CI=0.25-3.03). This difference between VCs and Progressors still remained when we corrected for gender by comparing the female groups only. Female VCs were still more enriched although not significant for -358TT+355(GA+AA) +2919TT compared to Progressors (VCs: 20.4% vs. Progressors: 11.2%, P=0.14; OR=0.49; CI=0.20- 1.20). Similarly, when the female groups were compared the -358TT+355(GA+AA) +2919TTcombination had no effect in ECs when compared to Progressors (ECs: 16.0% vs. Progressors:11.2%, P=0.74; OR=0.66; CI=0.20-2.21) and when Controllers were compared to Progressors (Controllers: 18.9% vs. Progressors: 11.2%, P=0.14; OR=0.54; CI=0.24-1.21) (Figure 3.25).

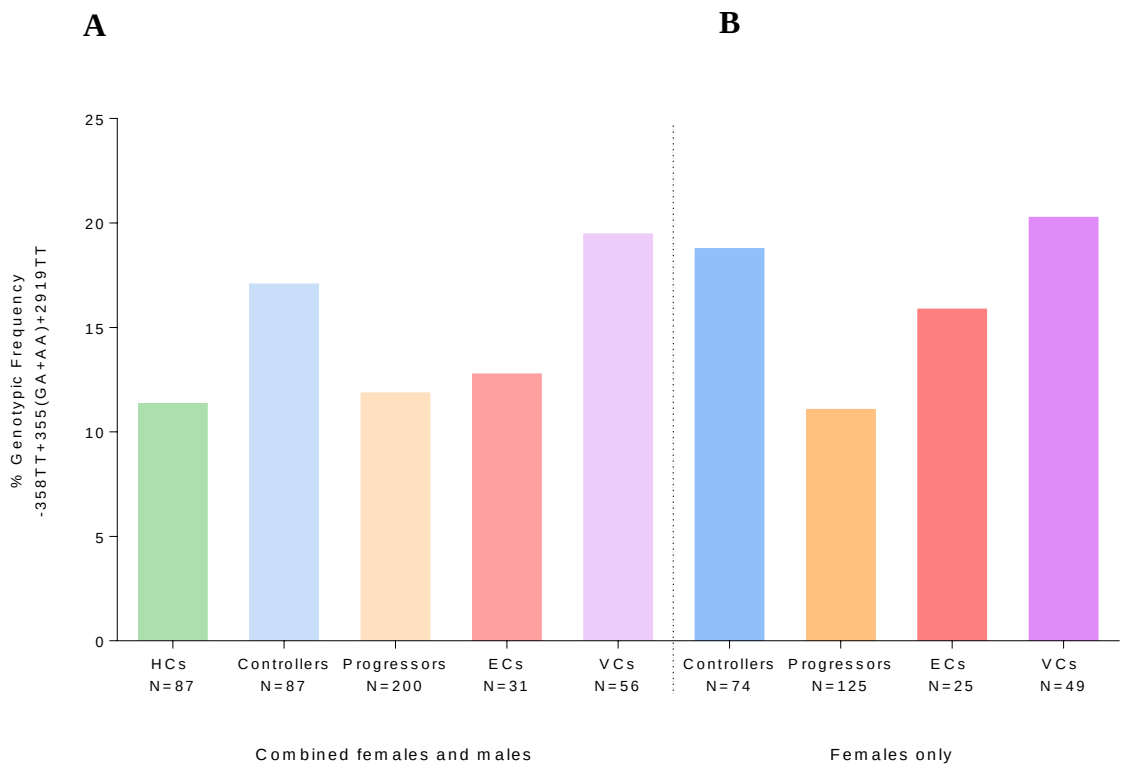


Figure 3.25. Bar graphs showing comparison of combinations of the *CXCR6* rs2234355 (355 GA+AA), rs2234358 (-358TT) and *CCR5* rs746492 (+2919TT) genotypes in HCs, Progressors, ECs and VCs (combined females and males) as well as in Progressors, ECs and VCs (females only).

3.7. Comparison of the frequency of *CCR5* (rs746492 and rs553615728) and *CXCR6* (rs2234355 and rs2234358) SNPs between HCs, Durban and Johannesburg cohorts

In order to ascertain whether the representations of the *CCR5* (rs746492 and rs553615728) and *CXCR6* (rs2234355 and rs2234358) SNPs in the DBN cohort were different to those in the JHB cohort, the allelic and genotypic frequencies for the *CCR5* and *CXCR6* SNPs in the DBN and JHB cohorts were compared. *CCR5* allelic and genotypic data for the JHB cohort was obtained from Koor *et al.* (2019) [Progressors (N=74), VCs (N=37) and ECs (N=23)] and *CXCR6* data for the JHB cohort was obtained from Picton *et al.* (2017) [Progressors (N=72), VCs (N=30) and ECs (N=11)]. The total Controller group (ECs and VCs) were also compared across cohorts. If significant differences were not found then the two cohorts could be combined, which would facilitate an increase in sample number of the Controller, Progressor, and VC and EC groups for future analyses. In addition we compared HCs (N=87) obtained from the AWI-

Gen H3 Africa study (Ramsay *et al*., 2016) to JHB and DBN Controllers to determine whether HCs are more similar to the JHB or DBN cohort.

3.7.1. Allelic comparisons of *CCR5* and *CXCR6* SNPs between HCs, Durban and Johannesburg cohorts

There was no significant difference in allelic representation of the *CCR5* rs746492 variant between the JHB and DBN cohorts in Progressors (G: 45.3% vs. 40.8%, P=0.38) and ECs (G: 30.4% vs. 41.9%, P=0.24), however, the minor allele (G) was significantly overrepresented in the DBN compared to JHB cohort in Controllers (G: 44.3% vs. 30.8%, P=0.02) with a strong trend of overrepresentation in the DBN cohort compared to JHB cohort in the VCs (G: 45.5% vs. 31.8%, P=0.06). Interestingly, the G-allele representation of the *CCR5* rs746492 SNP in HCs was more aligned to the DBN Controllers (P=0.75) and showed a strong trend of overrepresentation compared to the JHB Controllers (G: 30.8% vs. 42.0%, P=0.07). The DBN cohort VCs are more aligned with HCs in the representation of the *CCR5* rs746492 G-allele (Table 3.4, Figure 3.26A).

No significant difference was observed in the allelic representation of the *CCR5* rs553615728 variant between the JHB and DBN cohorts in Controllers (T: 4.2% vs 3.2%, P=0.49); Progressors (T: 3.4% vs 2.5%, P=0.56); VCs (T: 6.8% vs 3.0%, P=0.27) or ECs (T: 0% vs 2.0%, P=1.00) (Figure 3.26B). The allelic representation of the *CCR5* rs553615728 variant in Controllers from JHB and DBN cohorts was similar to HCs (T: 4.2% vs. 1.7%, P=0.28; 3.2% vs. 1.7%, P=1.00) (Figure 3.26B).

No significant difference was found in the allelic representation of the *CXCR6* rs2234355 variant between the JHB and DBN cohorts in Controllers (A: 51.2% vs. 48.3%, P=0.69); Progressors (A: 43.8% vs. 42.3%, P=0.77), VCs (A: 50.0% vs. 50.9%, P=1.00) or ECs (A: 54.5% vs. 43.5%, P=0.46) (Figure 3.27A). The allelic representation of the *CXCR6* rs2234355 variant in Controllers from JHB and DBN cohorts was similar to HCs (A: 51.2% vs. 46.6%, P=0.51; 48.3% vs. 46.6%, P=0.83) (Figure 3.27A).

Likewise, no significant difference was seen in allelic representation of the *CXCR6* rs2234358 variant between the JHB and DBN cohorts in Controllers (T: 45.1% vs. 48.3%, P=0.69); Progressors (T: 59.0% vs. 51.5%, P=0.14), VCs (T: 40.0% vs. 44.6%, P=0.63) or ECs (T: 59.1% vs. 54.8%, P=0.81). It is interesting to note that the T-allele representation of the *CXCR6* rs2234358 variant was comparable between DBN Controllers and HCs (T: 48.3% vs. 56.9%,

P=0.13) and showed a trend of underrepresentation in JHB Controllers compared to HCs (T: 45.1% vs. 56.9%, P=0.08) (Figure 3.27B).

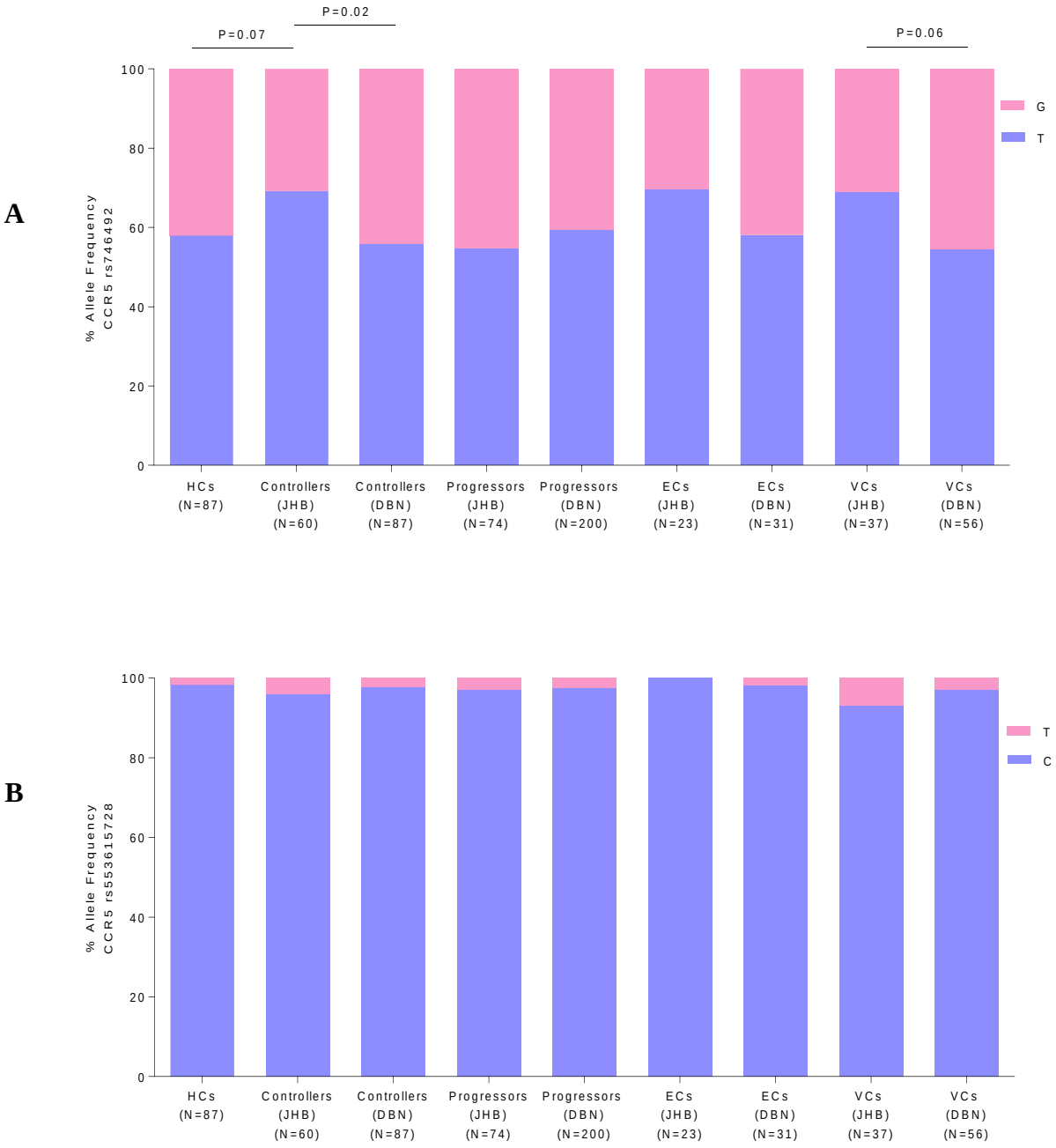


Figure 3.26. Bar graphs showing the comparison of allele distributions between Durban (DBN) and Johannesburg (JHB) cohorts for the: (A) *CCR5* rs746492 SNP and (B) *CCR5* rs553615728 SNP in HC, Controller, Progressor, EC and VC groups.

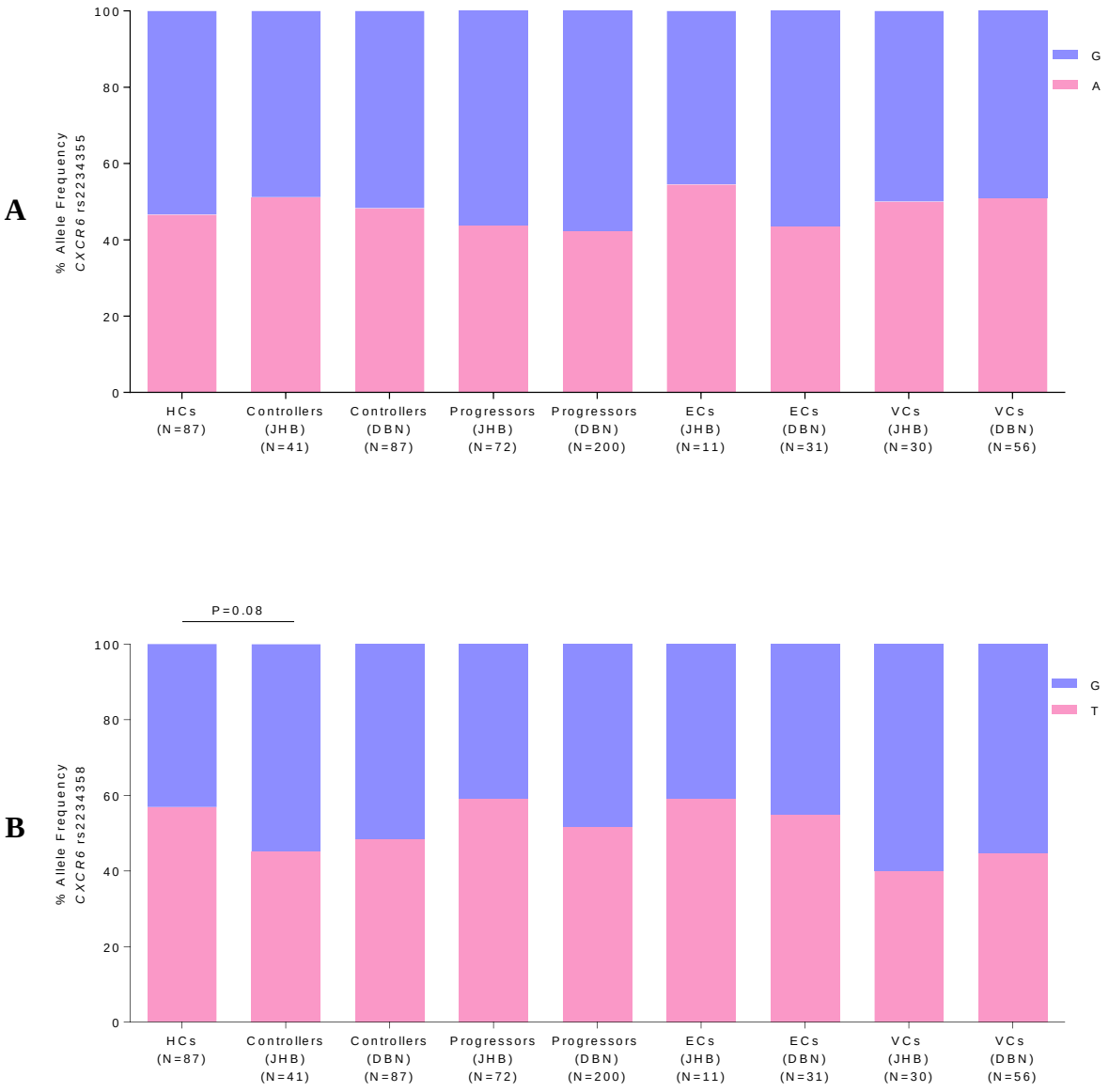


Figure 3.27. Bar graphs showing the comparison of allele distributions between Durban (DBN) and Johannesburg (JHB) cohorts for the: (A) *CXCR6* rs2234355 SNP and (B) *CXCR6* rs2234358 SNP in HC, Controller, Progressor, EC and VC groups.

3.7.2. Genotypic comparisons of *CCR5* and *CXCR6* SNPs between HCs, Durban and Johannesburg cohorts

Interestingly, when we compared the two cohorts in terms of genotypic representation, there appeared to be more distinct differences between the two cohorts.

CCR5 rs746492 major allele homozygosity (TT) was lower in JHB Progressors compared to DBN Progressors (22.2% vs. 33.5%) and higher in JHB Controllers, VCs and ECs compared

to DBN Controllers, VCs and ECs (Controllers: 50.0% vs. 29.8%; VCs: 51.4% vs. 30.4%; ECs: 47.8% vs. 29.0%). Heterozygosity (TG) was found at a higher frequency in Progressors in the JHB cohort compared to the DBN cohort (62.5% vs. 51.5 %,) and lower in Controllers, VCs and ECs in the JHB cohort compared to DBN cohort (Controllers: 38.3% vs.51.7%; VCs: 35.1% vs. 48.2%; ECs: 43.5% vs. 58.1%). Lastly, minor allele homozygosity (GG) was very similar between cohorts in Progressors (15.3% vs. 15.0%), however lower in JHB Controllers, VCs and ECs compared to DBN Controllers, VCs and ECs (Controllers: 11.7% vs. 18.4%; VCs: 13.5% vs. 21.4%; ECs: 8.7% vs. 12.9%) (Figure 3.28A). Overall however, comparing the *CCR5* rs746492 genotype distribution between the two cohorts did not show statistically significant distributions (Progressors: P=0.19; VCs: P=0.14 and ECs: P=0.40), but there was a trend of overrepresentation of the TG and GG genotypes in DBN cohort compared to JHB in Controllers (P=0.05). Overall, the genotypic representation of the *CCR5* rs746492 SNP in HCs (TT: 31.0%, TG: 54.0%, GG: 14.9%; Table 3.4) was more aligned to the DBN Controllers (P=0.86) and showed a strong trend of underrepresentation in JHB Controllers (P=0.07) compared to HCs.

For the *CCR5* rs553615728 SNP, major allele homozygosity (CC) was comparable between the JHB cohort and DBN cohort in Controllers (91.7% vs. 95.4%); Progressors (93.2% vs. 95.0%), in ECs (100% vs. 97%), however was lower in VCs in the JHB cohort (86.5% vs. 95.0%). Heterozygosity (CT) had slightly lower representation in Progressors from DBN (5.0%) compared to JHB (6.8%) cohort, while in Controllers and VCs the CT genotype had higher representation in the JHB cohort compared to DBN cohort (8.3% vs. 4.6%; 13.5% vs. 5.4% respectively). In the ECs, heterozygosity (CT) was absent in the JHB cohort (0%) and was less frequent in the DBN ECs (3.2%) compared to the DBN Progressor and VC groups. Minor allele homozygosity (TT) was absent in both cohorts, due to the rarity of this variant (Figure 3.28B). Overall the genotypic distribution of the *CCR5* rs553615728 SNP was not significantly different between cohorts in all study groups (Controllers: P=0.49; Progressors: P=0.56; VCs: P=0.26 and ECs: P=1.00), as well as when Controllers from JHB and DBN cohorts were compared to HCs (P=0.27; P=1.00, respectively).

For the *CXCR6* rs2234355 SNP, homozygosity for major allele (GG) was higher in the JHB cohort compared to the DBN cohort in the Progressors (42.0% vs. 34.5%), lower in the JHB cohort compared to the DBN cohort for Controllers and VCs (12.2% vs. 27.6%; 10.0% vs. 23.2%) as well as in ECs (18.0% vs. 35.9%) (Figure 3.29A). Heterozygosity (GA) had higher

representation in Progressors from DBN (46.5%) compared to JHB (29.0%), whereas in the Controllers and VC groups the GA genotype had higher representation in the JHB cohort compared to DBN cohort (73.2% vs. 48.3; 80.0% vs. 51.8% respectively). In the EC groups, heterozygosity was slightly more elevated in the JHB cohort compared to the DBN cohort (55.0% vs. 41.9%). Minor allele homozygosity (AA) was more highly represented in the JHB cohort compared to DBN cohort in Progressors (29.0% vs. 19.0%) and ECs (27.0% vs. 22.5%) although less so in the Controller and VC groups however, this genotype was substantially less prevalent in the JHB cohort compared to the DBN cohort (14.7% vs. 24.1%; 10.0% vs. 25.0% respectively) (Figure 3.29A). Interestingly, the overall distribution of the *CXCR6* rs2234355 genotypes were significantly different between JHB and DBN cohorts for the Controller, Progressor and VC groups, (P=0.03, P=0.03; P=0.04, respectively) but not for the EC groups (0.66). The genotypic representation of the *CXCR6* rs2234355 SNP in HCs (GG: 33.3%, GA: 40.2%, AA: 26.4%; Table 3.4) was more aligned to the DBN Controllers (P=0.56) and significantly differentially distributed when compared to JHB Controllers (P=0.002).

Lastly, regarding the *CXCR6* rs2234358 SNP, homozygosity for the major allele (GG) was underrepresented in the JHB cohort compared to the DBN cohort in the Progressor group (14.0% vs. 24.0%), was comparable between the cohorts in the Controllers and VCs (26.8% vs. 25.3%; 27.0% vs. 28.6% respectively), and had higher representation in the JHB cohort compared to the DBN cohort in the EC group (27.0% vs. 20.0%) (Figure 3.29B). Heterozygosity (GT) was relatively comparable between the JHB and DBN cohorts in Controllers and Progressors (56.1% vs. 52.9%, 54.0% vs. 49.0% respectively), higher in the JHB cohort compared to the DBN cohort in the VCs (66.0% vs. 53.6%) and considerably less prevalent in JHB cohort compared to the DBN cohort in the EC group (27.0% vs. 51.6%). Minor allele homozygosity of this variant (TT) was slightly more prevalent in the JHB cohort compared to the DBN cohort in Progressors (32.0% vs. 27.0%) but was slightly less prevalent in JHB compared to DBN in Controllers (17.1% vs. 21.8%), much less represented in the JHB cohort compared to the DBN cohort in VCs (7.0% vs. 17.9%) and more prevalent in the JHB cohort compared to the DBN cohort in the EC group (46.0% vs. 29.0%) (Figure 3.29B). Despite these differences, the distribution of these *CXCR6* rs2234358 genotypes between the two cohorts were not significantly different (Controllers: P=0.88; Progressors: P=0.19; VCs: P=0.33 and ECs: P=0.32). Similarly, HCs (GG: 19.5%, GT: 47.1%, TT: 33.3%; Table 3.4) were comparable to JHB and DBN Controllers (P=0.15; P=0.21, respectively).

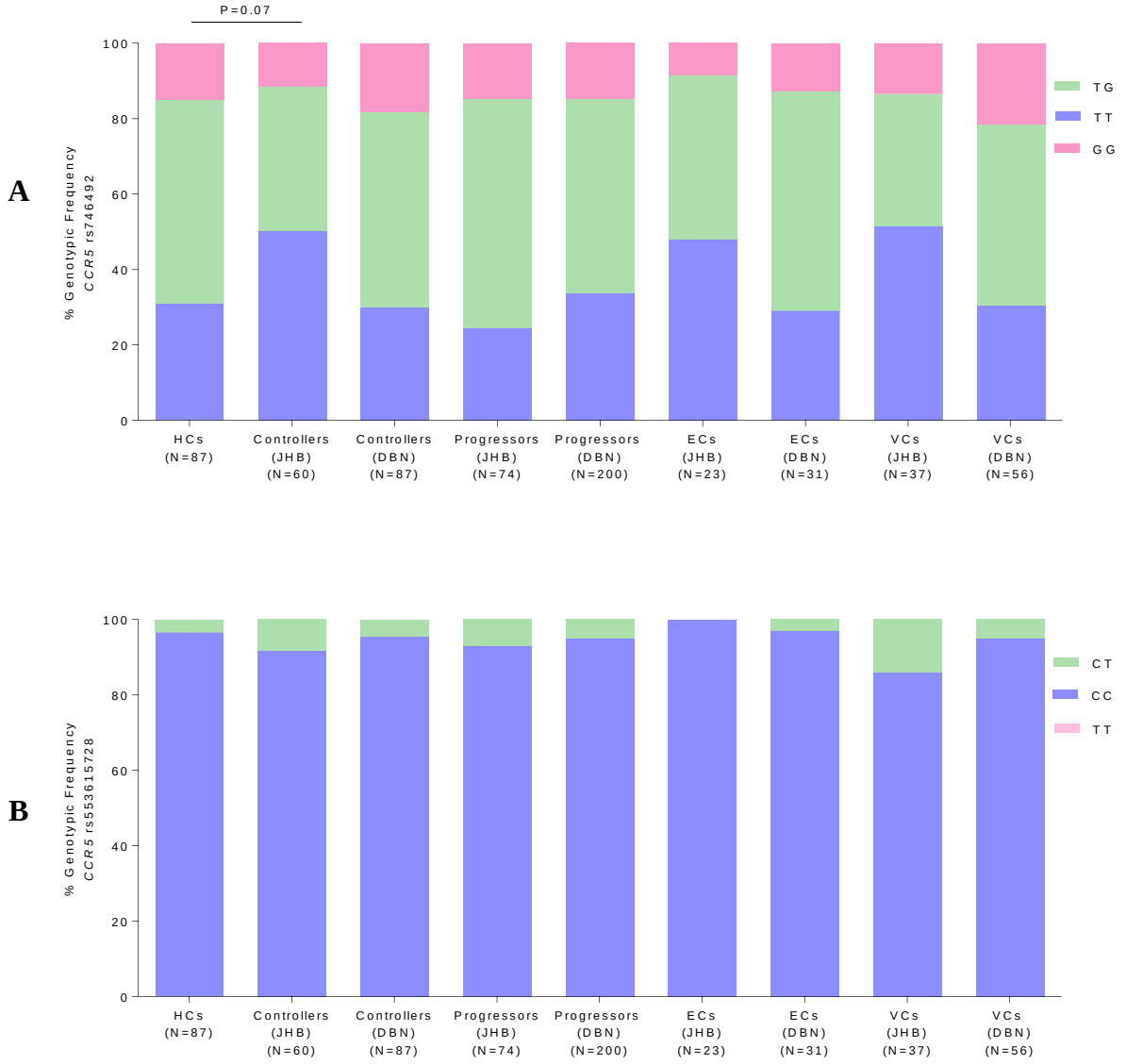
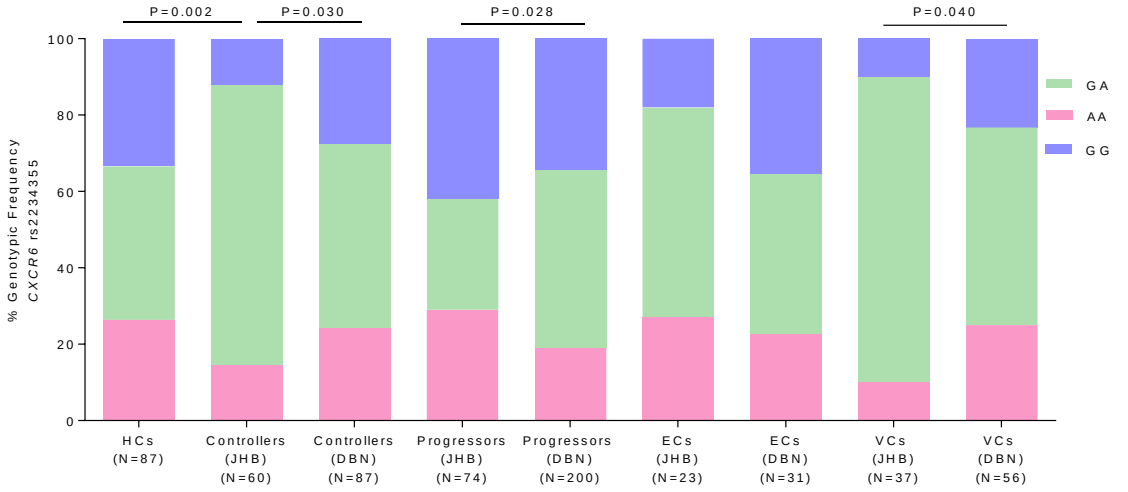


Figure 3.28. Comparison of genotypic distributions between HCs, Durban (DBN) and Johannesburg (JHB) cohorts for the: (A) *CCR5* rs746492 SNP and (B) *CCR5* rs553615728 SNP in HC, Controller, Progressor, EC and VC groups.

A



B

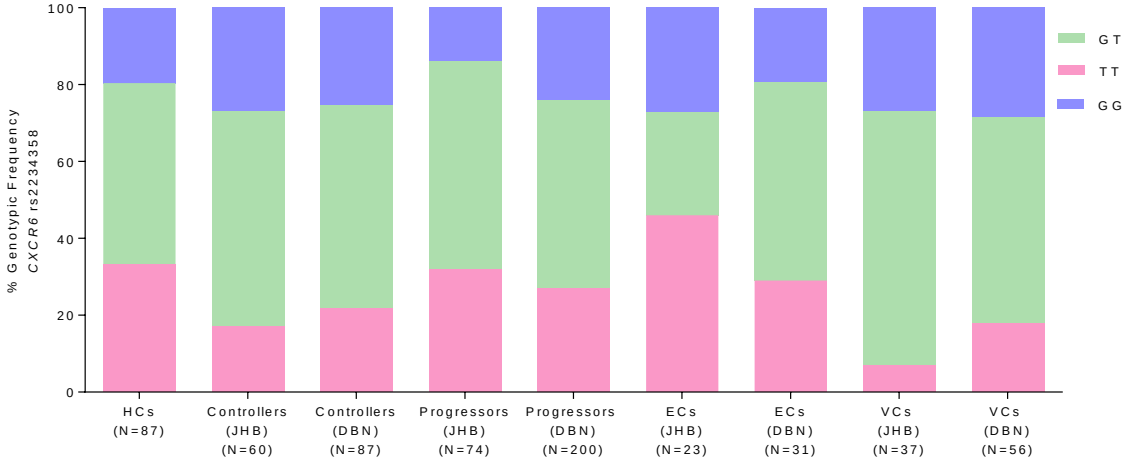


Figure 3.29. Comparison of genotypic distributions between HCs, Durban (DBN) and Johannesburg (JHB) cohorts for the: (A) *CXCR6* rs2234355 SNP and (B) *CXCR6* rs2234358 SNP in HC, Controller, Progressor, EC and VC groups.

CHAPTER 4: Discussion and conclusion

The overwhelming impact of HIV is most noticeable in Sub-Saharan Africa. In the absence of ART, most HIV-infected individuals have continuous viral replication and declining $CD4^+$ T-cell counts, progressing to AIDS, which inevitably results in death. Regardless of endeavours to develop a successful vaccine against HIV-infection, relatively toxic ARVs are still the best choice of treatment. There is however a subset of HIV-infected individuals who have the ability to naturally control their plasma viral replication without ARVs for extended periods of time - these individuals are broadly termed HIV Controllers. Elite controllers (ECs) comprise a small sub-group of HIV Controllers, with most HIV cohorts reporting a prevalence of <1% (Okulicz *et al*., 2009; Grabar *et al*., 2009). However the prevalence of ECs in this study was 5.2% which could possibly be attributed to the criteria used to define Controllers in this study based on a single time point. This high frequency of ECs in the 590 recruited participants could also be due to the already specific selection based on $CD4^+$ T cell count i.e. only individuals with $CD4^+$ T cell counts <200 and >500 cells/μl were included, therefore the estimates would be an over-estimate as the denominator was not based on participants representing all $CD4^+$ T cell counts within the period of recruitment. In addition, the numbers would be expected to be further reduced if Controllers categories were based on longitudinal data. The study of models that represent HIV-control can provide continued insights into the mechanisms of host control which may reveal specific targets or combination of factors that can be considered for the development of new therapeutic strategies in the endeavour to attain a "functional cure" for HIV infection.

In this study, 590 HIV-1 infected individuals were screened and recruited via the Lancers and Chesterville PHC bleeding rooms in Durban. These were convenient cohort recruitment locations as these individuals were already attending the PHC facilities and were scheduled to have their $CD4^+$ T-cell counts determined. We incorporated the study into the "clinic flow" so that participants eligible for our study were referred to us by the clinic staff. Of the 590 screened individuals, a total of 287 participants were enrolled into the current study and were categorised as HIV-1 infected Controllers or Progressors based on their $CD4^+$ T-cell counts and VL measurements, while the remaining 303 individuals were determined ineligible for study enrolment. There were several reasons for study ineligibility and these included the following: VL measurements and age exceeding the study criteria limit, commencement of ART before the enrolment visit, residence relocation and therefore not wanting to continue with participation in the study, and lastly, failure to return to the clinic for further clinical evaluation. The failure to return to the clinic could be related to either the participants not wanting to

commit to regular follow-up study-related clinic visits outside of their regular clinic visits, or secondly an inadequate understanding of the purpose of the study, or lastly that they felt pressured by health care workers into being initiated onto ARVs.

Several challenges were encountered during the recruitment process. At least 11% of the Progressors that were screened-out of the study were older than 35 years of age, the age that was set as the cut-off for eligibility of enrolment into the Progressor group. These participants had initially reported an incorrect date of birth at the screening visit which was then verified against their identity document upon their return visit for enrolment. Dishonesty or misleading medical information from the individuals attending PHC facilities is a frequent problem experienced within the clinics. Our results corroborate those of Sykes *et al*. (2019), who reported on "false elite controllers", where some individuals with the EC phenotype (HIV antibody-positive, RNA-negative) did not disclose ART use. Similarly, in this study there were individuals who had low $CD4^+$ T-cell counts and high VL displaying a Progressor phenotype, but upon being probed further at enrollment admitted to commencing ARVs prior to study recruitment but had not disclosed taking ARVs at the time of being screened.

Participants may sometimes be dishonest and falsify information for a number of reasons, these include wanting to keep a good relationship with the study research team and/or not wanting to lose the study reimbursement as well as the individual patient care from the study nurse or counsellor. Other reasons for false disclosures could also be related to experiencing peer pressure from friends within or outside the study. This can also be associated with non-commitment to the study as they only joined because they were giving in to the peer pressure. All of these issues tend to adversely affect the outcome of the study. One thing that we learned in this study is that study participants tend to give responses that are socially desirable especially when it pertains to sensitive or personal issues. Amongst the individuals that were identified as HIV Progressors, 16.7% did not return to the clinic for any further follow-up after being diagnosed with HIV-1. Although concerted effort was made to contact these individuals telephonically, this proved unsuccessful. We speculate that possible reasons for this were either individuals not understanding that they needed to be initiated on ARVs immediately, or that individuals were too ill to return, or the fear of being stigmatised about their status if they were seen at the ARV clinic upon collection of their ARVs.

Recruiting participants into a study can be an emotionally turbulent and maturing experience for the recruiter. Challenges encountered during the recruitment process were numerous and

included participants questioning how they would benefit from being recruited into the study with respect to the management of their health, to which the reply was that there may be no direct benefit to participants in such observational studies although participants and others may benefit in the future from the information gained from the study. When addressing such concerns and questions one needs to assure participants that their participation in the study is voluntary. Another example of a challenge included a male participant who was reluctant to answer questions on sexual behaviour due to the younger age of the recruiter, as he felt that it would be like disclosing such information to his daughter – in this case the recruiter respectfully asked him whether he would prefer to be handed over to an older colleague, who in this case successfully completed the recruitment process. Talking about HIV and AIDS poses difficulties since it involves discussion about sex which tends to be a very sensitive topic for many individuals, and this was more obvious where older male participants were concerned. Talking to participants about sex often resulted in some anxiety since it was often perceived as crossing status limits, determined by age, which in many African backgrounds it is a customary norm for the tone of conversation to be respectful when addressing older persons. Therefore, a recruiter needs to navigate that territory very cautiously and needs to avoid coming across as impolite or rude. We also noticed that some participants were more comfortable when the question regarding the numbers of partners they had in the last 3 months was posed by a younger non- professional (not a nurse or doctor) recruiter (which tended to be a higher number of partners), rather than when a mature nurse professional posed the question (where the answer tended to be one or two partners for the same individual). It has been reported that patients are rather reticent when it comes to their interaction with clinic nurses (Weitzman *et al*., 2012). Thus recruitment is a complex process which involves constant monitoring of the individual, the situation and the ethnic customs and boundaries.

There were also several opportunities that presented themselves during the recruitment process in this study, for example some participants in the Progressor group with low $CD4^+$ T-cell counts were too ill to be recruited and the study staff facilitated their immediate commencement of ARVs through the clinic sister in charge. Some individuals were concerned about the duration of the recruitment process as they also had to wait in the queue for group counselling prior to ARV initiation - due to the good working relationships between the study and clinic staff, we reassured participants that their place in the queue would not be affected as they would be taken to the clinic sister who was carrying out the patient ARV initiation after completion of the recruitment process without having to re-join the queue – sometimes a small gesture can

make the difference between a successful or failed recruitment. As another example, some individuals who after attending the group counselling session and had been initiated on ARVs sought more clarity about their disease - the study staff were thus able to empower them through education regarding their HIV disease. A few participants attended the clinic with their respective partners who questioned why one of them was being recruited and the other not – this allowed for the opportunity to explain the reason for conducting this study as well as to empower him/her with knowledge regarding their HIV disease. Lastly, participants appreciated the privacy the study provided during the recruitment process as they were alone with the recruiter and one other colleague.

The majority of participants enrolled in this cohort were female with the median age of 29 years. It has been reported that women are much more likely to seek health care because of the PMTCT programmes thereby enabling women to access HIV testing during routine antenatal visits (Shisana *et al*., 2015). Failure to visit the health facility for follow up care even after HIV-1 diagnosis amongst males has been strongly associated with stigma (Sprague and Simon, 2014).

Similarly, a study conducted in Johannesburg, South Africa reported that females were more knowledgeable about HIV than males, and healthcare facilities were frequented more by women (Johnston *et al*., 2010). According to Mburu *et al*. (2014) masculinity makes men want to be in control and powerful; the thought of being sick makes them feel invaluable so they prefer not to admit being ill.

Most individuals self-reported that their sexual debut occurred when they were older than 15 years of age while a minority (19.3%) self-reported being younger than 15 years, 6.8% of individuals did not know the actual age of their sexual debut and 0.5% of individuals were unwilling to disclose this information. Previously, most societies in sub-Saharan Africa were guided by norms and values that prohibited promiscuity and premarital sex acts (Bleek, 1981; Sarpong, 1977). In societies where premarital sex is prohibited or discouraged, individuals who engage in premarital sex stand the risk of being stigmatized should this become known to members of their society. Musheke *et al.* (2013) reported that the general public perceive HIV infected individuals as promiscuous and this is likely to make individuals feel insecure and afraid if they are seen attending an HIV testing site. The authors also mentioned that a positive HIV diagnosis has been associated with defining one's character in relation to moral issues. In

this study two thirds of individuals self-reported as having one partner in the last six months while the remainder had two or more or zero partners.

Heterosexual HIV transmission was self-reported by 98.1% of individuals in this cohort. In general, we found that participants in this study were knowledgeable about the different modes of HIV-1 transmission as well as the use of condoms to prevent HIV infection. However, the inconsistent condom usage reported by participants in this study points to the poor perception of actual infection risk in fact a fair number of individuals (19.5%) reported never using condoms during sexual acts. Most participants reported as tending to use condoms only with casual partners and not a regular partner, thereby displaying risky sexual behaviour with the regular partner. Similarly, Gibbs *et al*. (2014) showed that participants in new or casual relationships are more likely to use protection, and as the relationship becomes more stable, then condom usage declines. Although this behaviour is understandable and likely a natural progression in a stable relationship, it makes the obvious assumption that a regular partner is not having unprotected sex with other individuals (possibly HIV-infected) outside of the relationship, an assumption that can be very costly and should likely be discouraged in high risk societies. This is the most likely explanation in 23.7% of participants who reported always using condoms but yet still got infected.

Although individuals do seek alternative medical services, the utilization of traditional medicines by participants in this cohort was negligible.

The majority of individuals (71.9%), mostly females (69.0%), reported being diagnosed with HIV-1 infection within the last 3 years prior to consultation, while in approximately one quarter of individuals, diagnosis took place between 3-10 years prior to consultation. The more recent diagnoses by mostly women can likely be attributed to the health seeking behaviour of women through PMTCT programmes (Shisana *et al*., 2015).

Clinical manifestations of HIV-1 infection vary from one individual to another and those differences are independent of age, gender and mode of infection (Ahuja *et al*., 2008). Various studies have established that both viral and host genetic factors are essential determinants of the course of HIV-1 infection (Naranbhai and Carrington, 2017; McLaren *et al*., 2017; Liu *et al*., 1996). Without ART, the majority of HIV-1 infected individuals will develop uncontrolled viraemia and undergo progressive immune impairment thus progressing to AIDS within a few years of diagnosis. However, with that being said, there is a subset of individuals characterized

by a slow disease progression - individuals who can maintain a high $CD4^+$ T-cell count while supressing HIV replication (Fellay *et al.*, 2007). These individuals provide us with the opportunity to understand what pathways are important in natural control of HIV-1 in order for the design of better therapeutics, vaccine targets and ultimately a functional cure for this devastating disease.

It has been demonstrated that HIV-1 can become less virulent over time, resulting in protective HLA alleles no longer showing an association with slow disease progression as compared to previous, older studies (Payne *et al.*, 2014). This is largely due to the accumulation of escape mutants in critical viral epitopes important for immune control thereby lowering viral fitness at the population level (Kløverpis *et al.*, 2016). It could be postulated that the same may be occurring in response to other host factors, such as CCR5 and CXCR6, that have been found to be important in the HIV-1 immune response and that have been suggested to play a role in HIV-1 disease. For this reason, following findings made in a Johannesburg (JHB) based cohort (Picton *et al.*, 2010; Koor *et al.*, 2019), the current study was undertaken in a different region of South Africa (i.e. Kwazulu Natal; KZN), and results between cohorts compared and analysed to determine if findings from the JHB cohort studies were consistent with findings in KZN since the dynamics of the HIV epidemic may very well be different in the two regions.

Polymorphisms in the *CCR5* and *CXCR6* genes have been found to significantly influence the outcome of HIV infection (Bachelerie *et al.*, 2014). Therefore, information on *CCR5* and *CXCR6* genetic variation in HIV-1 infected individuals will provide a better understanding on how these chemokine receptors influence susceptibility to HIV-1 infection and/or disease acceleration. Different population groups show different associations to HIV-1 acquisition and disease progression, hence variation is also found among the various defined HIV-Controller groups. It is important to study genotypic variation in African populations to identify genes that play a major role in complex disease susceptibility, adaptation, and function in Africans and populations of recent African ancestry.

Africa is an important region to study human genetics because of the complex population history and exposure to infectious diseases which has resulted in various disease related genetic mutations in African populations (Campbell and Tishkoff, 2008; Campbell and Tishkoff, 2010). Availability of *CCR5* rs746492 and *CXCR6* (rs223435, rs2234358) SNP data from the 1000 genomes project (*www.1000genomes.org*), allowed us to compare representation of these variants in populations of European (Caucasian) ancestry and importantly to populations of

sub-Saharan African ancestry as well as to the black South African population (HCs) (Ramsay *et al.*, 2016). This allowed us to provide novel information as to how representation of these functionally important variants differed from other population groups.

The representation of the *CCR5* rs746492 and *CXCR6* rs2234355 and rs22334358 SNPs in the South African black population (HC) varied considerably from the European population (EUR), not unexpectedly since sub-Saharan African populations, given that they are older populations from an evolutionary perspective, tend to exhibit higher levels of genetic variation when compared to Caucasian populations, with different polymorphisms, differential variant distributions, and differences in patterns of linkage distribution (Picton *et al.*, 2010; Ramsay *et al.*, 2011; May *et al.*, 2013; Gentle *et al.*, 2013). What was somewhat surprising were the significant differences seen in all three SNPs, both in allelic genotypic representation, when we compared the South African black population to the West African Yoruba (YRI) population. The South African black population was more similar in representation of these three SNPs to the East African Luhya (LWK) population, especially with respect to the two *CXCR6* SNPs. This was similar to what was seen in the Johannesburg cohort with respect to these two *CXCR6* variants (Picton *et al.*, 2017). Picton *et al*. (2017) suggest that populations living in SIV endemic areas (West Africa) may have become resistant to immunodeficiency viruses due to undergoing genetic selection of certain alleles. While natural selection at these loci might be one explanation, the same pattern is often seen in other genes (Ramsay *et al.*, 2016) and might just be due to migration patterns during the Bantu Expansion. CXCR6 may well be one of these genes, as African green monkeys and sooty mangabeys display a non-pathogenic SIV infection despite significant viral replication, and use CXCR6 receptor efficiently (Riddick *et al.*, 2010; Wetzel *et al.*, 2017). However, what was worth noting was that *CXCR6* rs2234358 G-allele, which has been shown to be the beneficial allele in viraemic long term non-progression (Limou *et al.*, 2010; Picton *et al.*, 2017), had the highest allelic representation in the South African black population (HC: 43.1%) compared to the other African populations (LWK: 36.4%; YRI: 24.5% and total African: 31.5%). Although the HIV epidemic from a human genetics evolutionary perspective has not existed long enough to be considered as an extended period of time, it could be speculated that this difference in allele frequency may be attributed to the selection of this protective allele in the South African black population due to the country being the epicentre of the global HIV epidemic for most of the duration of the HIV epidemic.

The 3' and 5' untranslated regions (UTRs) as well as intronic regions are important in the regulation of the *CCR5* gene (Mummidi *et al*., 1997) since mutations in these regions may influence *CCR5* expression, an important feature that impacts HIV-1 disease outcome. Genetic variants in the *CCR5* gene have been shown to influence HIV-1 acquisition and progression to AIDS (Naranbhai and Carrington, 2017). A cohort study of HIV-1 infected Controllers and Progressors in JHB showed that two SNPs in the *CCR5* gene, namely one in the 3'-UTR region (+2919T/G, rs746492) (Koor *et al.,* 2019) and the other in the 5'-UTR (-4223C/T, rs553615728) (Gornalusse *et al*., 2015), could potentially serve as deleterious and protective markers in the context HIV-1 control, respectively (Koor *et al.,* 2019). In addition, Koor *et al*. (2019) reported on the effect of the allelic and genotypic representation of CCR5 5'-UTR-2SNP-hap (-2459G/A and-2135C/T) to be underrepresented in total Controllers and VCs compared to Progressors. Furthermore, these authors showed that the +2919T/G SNP is in LD with the 5'-UTR-2SNP-hap.

In this study, conducted on similar disease phenotypes (Controllers, ECs, VCs, and Progressors) in KZN, the *CCR5* rs746492 variant did not show a similar association. *CCR5* rs746492 minor allele (G) representation was comparable between Controllers, ECs, VCs and Progressors, and possession of the heterozygous genotype (TG) did not show an overrepresentation in Progressors compared to any of the Controller groups. These findings are different to those reported by Koor *et al.* (2019), who found that the minor allele was lower in ECs and VCs but was significantly underrepresented in total Controller group compared to Progressors, while heterozygosity (TG) was underrepresented in all Controller groups when compared to Progressors. The *CCR5* rs746492 SNP in the dominant mode (i.e. TG+GG) also did not reveal an association with HIV-1 disease progression in this study as it was found in the JHB cohort (Koor *et al.,* 2019). These results suggest that unlike the JHB cohort, the *CCR5* rs746492 SNP has no effect on HIV-1 control in the DBN cohort. This marked difference is an interesting finding and opens up questions regarding the role of CCR5 expression in HIV-1 control in the setting of the KZN epidemic. However, future work is needed to understand and confirm the effect of this polymorphism in the context of HIV-1 control in the greater black South African population.

The *CCR5* rs553615728 SNP, was only detected in individuals of sub-Saharan African descent in a study conducted in two South African populations – black and white South African populations (Picton *et al*., 2010). Although a fairly rare variant, the rs553615728 SNP has been

associated with the disruption of the CpG-41 site which causes changes in DNA methylation levels and alters the binding of the cAMP responsive element binding protein 1 (CREB1) transcription factor of CCR5 (Gornalusse *et al*., 2015). This polymorphism has been found to be more prevalent in individuals with the protective CCR5 HHA haplotype and there is a very strong LD between this SNP and HHA haplotype (Gornalusse *et al*., 2015; Koor *et al*., 2019). According to Gornalusse *et al*. (2015), the rs553615728 SNP is also associated (a strong trend) with protection from HIV-1 acquisition in southern Africans.

Two recent studies done in the South African black population have investigated the role of the *CCR5* rs553615728 SNP in HIV-1 disease (Koor *et al*., 2019; Jaumdally *et al*., 2017). Although the findings by Jaumdally *et al*. (2017) were not significant, these authors reported that exposed-uninfected individuals had a higher representation of the rs553615728 SNP (11%) compared to unexposed controls (5%). Koor *et al*. (2019) reported the *CCR5* rs553615728 SNP to be associated with HIV-1 control in the context of high viraemia (i.e. VCs) but not in elite control. In this study cohort we found no significant difference in representation of the *CCR5* rs553615728 SNP between the Controller, Controller sub-groups and Progressors. It should be noted however that in the Koor *et al*. (2019) cohort study, the effect of this SNP only showed significance once the groups were compared in the absence of the HHE haplotype (a known high transcriptional promoter haplotype). Since we had not sequenced the *CCR5* promoter regions in this study cohort, we could not carry out the same analysis and thus the results regarding this *CCR5* SNP and its effect on HIV-1 control in the KZN cohort are inconclusive.

The *CXCR6* rs2234355 SNP, also known as *CXCR6-E3K*, results in a non-synonymous amino acid transition within the N-terminus of CXCR6 (glutamic acid to lysine) and has been linked to increased survival from *Pneumocystis jiroveci* pneumonia (PCP) in ART-naive African-American HIV-infected patients (Duggal *et al*., 2003). Interestingly the initial VL suppression attributed to ART was associated with increased incidence of virological failure when the CXCR6-3K allele was present in Caucasians (Passam *et al*., 2007). This *CXCR6* rs2234355 SNP was found in higher percentages amongst Biaka Western Pygmies in comparison to Mbuti Eastern Pygmies in West Central Africa (Zhao *et al*., 2012) and has been attributed to being one of the protective alleles. The Biaka Western Pygmies used to live in the same region as the central African chimpanzee, which was the natural host for SIV strains from which HIV is thought to have originated (Zhao *et al*., 2012).

The *CXCR6* rs2234355 minor allele (A) was detected at a high frequency (>40%) in this cohort. This is in agreement with the findings by Picton *et al*. (2017) who also reported in excess of 40% frequency of the rs2234355-A allele in their black South African cohort. Although Picton *et al*. (2017) found that heterozygosity for this variant (GA) was highly significantly overrepresented in VCs compared to Progressors (80% vs. 29%, respectively) in their JHB cohort, both the allelic and genotypic comparisons conducted in this study showed no significant association of this polymorphism with HIV-1 control in the total Controller group or the VC and EC sub-groups.

The *CXCR6* rs2234358 polymorphism is located 42 base pairs downstream from the termination codon (3'-UTR) and may therefore influence mRNA stability, gene expression, mRNA splicing or mRNA regulation (Limou *et al*., 2010). However, the functional implications of sequence variants in the m-RNA 5′ leader (i.e., the region upstream of the ORF initiator codon) and 3′-UTRs (i.e., the region downstream of the ORF stop codon) are often uncharacterized as they are not always apparent (Robert and Pelletier, 2018).

The *CXCR6* rs2234358 polymorphism has been associated with HIV-1 disease control in the presence of viraemia (i.e. not in the context of elite control) in Caucasian individuals (Limou *et al*., 2010). This association was replicated in three additional independent cohorts of European descent reaching genome wide significance of $P_{combined}=9.7x10^{-10}$, and represented the first non-HLA replicated association discovered using a GWAS approach (Limou *et al*., 2010). Similarly, in a study conducted on a South African black population, the rs2234358 T-allele (minor allele) was significantly lower in viraemic controllers compared to both Progressors as well as healthy uninfected controls, and minor allele homozygosity (TT) was also significantly lower in viraemic controllers compared to both Progressors and healthy uninfected controls (Picton *et al*., 2017).

In this study the frequency of the minor allele (T) as well as minor allele homozygosity (TT) of the *CXCR6* rs2234358 SNP showed strong trends (P=0.05 for both) of lower representation in the VCs compared to the background black South African population (HCs). Progressors had higher representation of the *CXCR6* rs2234358 SNP T allele and TT genotype compared to VCs, however these were lower than the HC population and thus were not significant or trending towards significance. In the JHB study (Picton *et al*., 2017), the HC group also had higher representation of this SNP than the Progressors and the significance compared to VCs

was stronger relative to the HCs than to the Progressors. Thus, although the results in the current study do not show the same levels of significance, the relationships are similar and the shifts are in the same direction.

A number of host genetic factors have previously been associated with the classical markers of HIV-1 disease progression, i.e. viral load and $CD4^+$ T-cell count. Shrestha *et al*. (2010) reported that individuals with *HLA B*57* alleles had significantly lower mean VL compared to individuals without this allele. In the current study the association between *CCR5* and *CXCR6* genotypes and VL as a marker of HIV-1 disease progression was evaluated in the Progressor and VC groups. The relationship between genotypes and VL was not assessed in the Controllers and ECs as these individuals do not clinically progress and they have low or undetectable viral replication. The two *CCR5* and *CXCR6* SNP genotypes showed no association with VL in VCs. However, in Progressors, individuals possessing the *CXCR6* rs2234358 TT genotype had significantly higher VL when compared to individuals possessing *CXCR6* rs2234358 GT genotype. In the comparison of individuals with the TT genotype to all other individuals (i.e. TT vs. GG+GT), the individuals with *CXCR6* rs2234358 TT genotype still had significantly higher VL measurements. These results further support the deleterious effect of the *CXCR6* rs2234358 TT genotype – in the Limou *et al*. (2010) study, the possession of the TT genotype was associated with the most rapid progression to AIDS-related death in their cohorts using Kaplan-Meier survival curve analyses.

The most significant association reported by Picton *et al*. (2017) involved the combination of heterozygosity for the *CXCR6-E3K* SNP (rs2234355) in the absence of rs2234358 SNP minor allele homozygosity (TT), with VCs having 80% representation compared to 17% representation in Progressors ($P<1x10^{-7}$ after correction for multiple comparisons). This result also suggests that the significant underrepresentation of this genotype in VCs compared to Progressors in this study, and in the study by Picton *et al*., (2017), is likely attributable to direct control of viral load. It is also important to note that in the European-based studies, the T allele was also underrepresented in LTNPs (not ECs) compared to controls (Limou *et al*., 2010).

A recent transcriptomic study has shown that *CXCR6* gene expression was significantly downregulated in ECs in comparison to Progressors (Zhang *et al*., 2018), what was interesting is that the rs2234358 SNP was reportedly absent in their EC cohort. Given that European populations have a relatively high representation of this rs2234358 SNP (47% allelic frequency, 1000 Genomes Project), the absence of this SNP could be suggesting an association

with the lower CXCR6 expression seen. This may point to a mechanism in which decreased T-cell susceptibility to HIV-1 entry causes increased HIV-1 replication control in HIV-1 infected ECs. What is also worthy of mention is that in the Zhang *et al*. (2018) study, their definition of elite control did not seem as rigorous as one would expect and may have in fact been a mixture of ECs and VCs, however this would need to be confirmed and is speculative at this stage; nonetheless this study does differ in that it suggests a role for CXCR6 in elite HIV-1 control.

The *CCR5* (rs746492 and rs553615728) and *CXCR6* (rs2234355 and rs2234358) genotypes were stratified against $CD4^+$ T-cell counts in the Progressors, VCs and ECs as a measure of HIV control, in order to evaluate whether there was any marked difference in $CD4^+$ T-cell counts in the presence or absence of the respective SNPs. No significant associations were found between $CD4^+$ T-cell counts and *CCR5* (rs746492 and rs553615728) and *CXCR6* (rs2234355 and rs2234358) genotypes in Progressors, VCs and ECs, and although VCs had significantly higher $CD4^+$ T-cell count with the *CCR5* rs553615728 CT genotype compared to those with the GG genotype, due to the small sample number (N=3 with CT genotype), this result is likely not a true association. $CD4^+$ T-cell counts have been reported not to be a reliable marker for HIV-1 control even though there is an inverse relationship between $CD4^+$ T-cell count and VL (Shoko *et al*., 2019). As example, Hunt *et al*. (2008) reported a minority of individuals from an EC cohort (N=30), infected for about 16 years with VL<75 RNA copies/ml, had $CD4^+$ T-cell counts of ≤350 cells/µl (10%) while others (7%) met the clinical definition of AIDS phase at the time.

This study cohort had considerable differences between gender distributions amongst the study sub-groups, with the Progressor group having significantly less females compared to the Controller group confirming similar findings in another study (Madec *et al*., 2005), who demonstrated that females were overrepresented in the Controller group, had low baseline VL and high $CD4^+$ T-cell counts. Certain studies have also reported better control of HIV amongst females. Gandhi *et al*. (2002) reported that the VLS in females is 0.33-0.78 log copies RNA/ml lower than in males in untreated HIV-1 infection, while Yang *et al*., (2017) reported that women are 5 times more likely to become ECs compared to men. This was attributed to the fact that pDCs in females produce significantly more IFN-α in response to stimulation by toll-like receptor 7 (TLR7) ligands such as HIV-1 and other single-stranded RNA viruses. Souyris *et al*. (2018), recently showed that TLR7 is on the X-chromosome and is expressed at higher levels in females than males. However, Moore and Cheever, 1999, found no significant

differences in VL between females and males while in the JHB cohort studies gender was similarly distributed between the Progressor, Controller, VC and EC groups (Picton *et al.*, 2017; Koor *et al*., 2019).

We therefore assessed the allelic and genotypic frequencies for the *CCR5* and *CXCR6* SNPs on HIV control with respect to gender. There was no difference between males and females in *CCR5* and *CXCR6* allelic frequencies in any of the study groups. The *CCR5* rs746492 SNP did not reveal any significant differences in representation with respect to gender in any of the groups in this study. However, there was a significantly higher representation of the *CCR5* rs553615728 CT genotype in male VCs compared to female VCs, with the GG genotype more prevalent in female than male VCs. In the Progressor group the *CXCR6* rs2234355 GG genotype was significantly underrepresented in males compared to females and heterozygosity (GA) was overrepresented in males compared to females. These findings differ from the JHB cohort study (Picton *et al.,* 2017), where no differences were found with respect to gender between Progressors and Controllers.

Due to the gender differences discussed above, we subsequently analysed the effect of *CXCR6* rs2234355 variant in female VCs and Progressors. There was a trend of *CXCR6* rs2234355 minor allele (A) as well as the *CXCR6* rs2234355 in the dominant mode (GA+AA) overrepresentation in VC females compared to Progressors. Although we did not observe significant differences in the *CXCR6* rs2234355 and rs2234358 SNP representations when the genotypes were compared independently in Controllers, ECs, VCs and Progressors in this study, there was an association of the *CXCR6* s2234358 TT genotype with increased VL in Progressors. Analysis of the combined effect of the two *CXCR6* SNPs in males and females (combined group) as well as females alone was conducted to determine whether there was an association between the previously described combination of protective *CXCR6* rs2234355 GA genotype plus absence of the deleterious rs223438 TT genotype (Picton *et al.,* 2017). In this study we found that the effect of the two *CXCR6* SNPs in a slightly different combination (i.e. individuals lacking the *CXCR6* rs2234355 TT genotype in the presence of *CXCR6* rs2234355 in the dominant mode), designated as [-358TT+355(GA+AA)], showed significant higher representation in VCs compared to Progressors and compared to HCs in the total group comparison, with this association being enhanced when corrected for gender by comparing females only. These findings are thus consistent with those reported by Picton *et al.* (2017) and Limou *et al*. (2010), since *CXCR6* is impacting viraemic control but not elite control. Although

there was no significant overrepresentation in the combined effect of the *CCR5* and *CXCR6* protective genotypes, there was a higher representation in VCs compared to Progressors, HCs and ECs in the total and female only group comparisons. This is likely attributable to the protective effect of the *CCR5* genotype diluting the effect of the *CXCR6* protective genotype seen. These findings thus serve to corroborate that some protective factors may have an effect in Controllers with higher viraemia who have the ability to control their viral loads to some degree without a decline in their $CD4^+$ T-cell count, and show no association with ECs who display largely undetectable HIV-1 viraemia. This also highlights the importance of the study of natural control of HIV-1 by looking at the different modes of control and not simply grouping all modes under one group as it is likely that different or overlapping mechanisms will play a role in these modes of control. The use of gender stratification analysis to address the problem of gender as a confounder was hampered by the small sample sizes particularly in the EC and VC groups. Multivariate analysis to control for confounders might be useful to further interrogate the effect of *CCR5* and *CXCR6* variants on HIV control with respect to gender.

Gauteng is a highly ethnically heterogeneous province compared to KZN, which is relatively homogenous regarding the ethnicity with respect to the black population. The HIV-1 prevalence between these South African populations differs with KZN bearing the burden of HIV-1 (Shisana *et al.*, 2015). Factors responsible for these variations are poorly understood. Each location has its distinct characteristics as well as different patterns of individual risk factors. In the current study, we investigated whether the representation of the *CCR5* (rs746492 and rs553615728) and *CXCR6* (rs2234355 and rs2234358) SNPs in the DBN cohort differed from those in the JHB cohort. This comparison was carried out in order to determine if the two cohorts could be combined which would facilitate an increase in sample number of the Controllers, Progressor, EC and VC groups for future analyses. There was a significant overrepresentation of the *CCR5* rs746492 G allele in the DBN compared to the JHB cohort in Controllers and a strong trend of overrepresentation in VCs, but no other significant differences were observed in the allelic representation of *CCR5* rs553615728 SNP between the two cohorts. Likewise, no significant differences were found in *CXCR6* rs2234355 and rs2234358 SNP allelic comparisons between the two cohorts in all study groups. Comparison of genotypic distributions between the JHB and DBN cohorts showed no significant differences in *CCR5* rs746492 and rs553615728 SNPs.

The distribution of the *CXCR6* rs2234355 genotypes however were significantly different in the Progressor and VC groups between the two cohorts. With respect to the rs2234358 SNP, overall the genotypic distributions did not differ significantly between the two cohorts in any of the study populations. It is possible that the different HIV epidemic dynamics, in the background of different ethnic host differences, may be responsible for the differences seen between the JHB and KZN cohorts, particularly with respect to *CXCR6* SNP representation. It should be noted that the HC data that was used in both the JHB and the KZN comparisons was a mixture of black South African ethnicities. As the allelic and genotypic frequencies for the *CCR5* and *CXCR6* SNPs in the DBN cohort were more aligned with HCs, it would be informative to know what the background representation of these SNPs is in a predominantly healthy Zulu population group relative to the Progressor and Controller groups from KZN.

Susceptibility to HIV-1 infection and rate of progression to AIDS is influenced by a number of genetic variations. Understanding the contribution of the SNPs within the *CCR5* and *CXCR6* genes is crucial to elucidate the mechanisms of control of the virus. Limitations of this study included the following: First, we assigned participants into HIV phenotypic groups based on one time-point based on VL and $CD4^+$ T-cell count. The Controllers are therefore "potential" elite or viraemic controllers, and results might differ if these are defined by stricter criteria based on longitudinal clinical data. These longitudinal analyses are in progress as part of the larger ongoing study. The Progressors were also defined based on a single time-point, but restricted to individuals with progressive infection ($CD4^+$ T-cell counts <200 cells/µl, high VLs >10 000 RNA copies/ml) and who were ≤35 years of age. The age threshold for inclusion was to ensure a greater likelihood of excluding longterm controllers who have subsequently lost control. Second, with the utilization of self-reported questionnaires it became evident that sometimes participants do not disclose truthful information, therefore results obtained from questionnaires should be treated with some caution. In particular, as evident in the South African Blood Bank study, which identified "false elites" who did not disclose ART usage, it is important to test plasma samples for ARVs (Sykes *et al.*, 2019) to ensure the correct categorization of elite controllers. Third, there are a number of promoter haplotypes in *CCR5* that associate with control, but the *CCR5* and *CXCR6* SNPs genotyped and assessed in this study are not in linkage disequilibrium and thus no haplotype analysis was performed. The haplotype structures are unlikely to differ between these two cohorts, but haplotype frequencies may very well differ between the two cohorts, and may be differentially impacting on HIV control. Ongoing

studies of these cohorts will explore this possibility, as well as genetic variation of the chemokine ligands of CCR5 and CXCR6. There is a need to thoroughly investigate and understand the interactions of these important chemokine-chemokine receptor axes in the context of HIV-1 control in populations most affected by the HIV-1 epidemic. Lastly, small sample sizes especially in EC and VC groups, lack of knowledge of other genotypes in the *CCR5* or *CXCR6* or other genes such as HLA and lack of knowledge of viral strains (as some controllers could theoretically have less virulent strains) were additional weaknesses of this study.

In conclusion, these findings suggest that *CXCR6* rs2234358 TT is strongly associated with HIV-1 disease progression as this genotype was found to be significantly associated with higher VL in individuals that have progressed to disease. The *CCR5* SNPs did not demonstrate a marked effect in the DBN cohort as was reported in the JHB cohort (Koor *et al.*, 2019). This finding of the *CCR5* variants not impacting natural control in the DBN cohort raises several questions: (i) has the evolution of the HIV epidemic in KZN resulted in *CCR5* expression levels not being an important factor? (ii) has the virus in KZN evolved to be less virulent, thereby resulting in the less stringent identification of individuals who naturally control HIV, i.e. a Controller identified in KZN may in fact not be identified as a Controller in JHB, and hence we do not see the same effect of these variants as they have been 'diluted out' so to speak and therefore much larger numbers of participants would be required to see the same effect? As already stated there is evidence that as the HIV-1 epidemic evolves as the virus accumulates escape mutants in critical epitopes associated with immune control thereby rendering it less virulent, resulting in protective HLA alleles no longer showing an association with slow disease progression as compared to previous older studies (Payne *et al.*, 2014; Kløverpis e*t al.*, 2016). Thus one of the main objectives of the study, which was to determine whether these SNPs werehaving similar effects across cohorts so that cohorts could be combined and analysed as a largergroup has revealed a more complex situation than was anticipated. The contributing factors for the observed differences might be attributed to ethnic differences, or epidemic dynamics, or a combination of both. Therefore, a careful consideration needs to be applied in future analyses going forward with regards to merging data across cohorts. Lastly, this study has confirmed that it is important to know the genetic backgrounds of the populations being studied, even within the same geographic location and that the relationship between the host and virus is a complex one that is not necessarily the same in a constantly evolving epidemic.

REFERENCES:

1000 Genomes Project Consortium, 2015. A global reference for human genetic. Available at: (www.1000genomes.org).

Adler, M.W., 2001. Development of the epidemic. *Bmj*, *322*(7296), pp.1226-1229.

Agata, Y., Kawasaki, A., Nishimura, H., Ishida, Y., Tsubat, T., Yagita, H. and Honjo, T., 1996. Expression of the PD-1 antigen on the surface of stimulated mouse T and B lymphocytes. *International immunology*, *8*(5), pp.765-772.

Ahuja, S.K., Kulkarni, H., Catano, G., Agan, B.K., Camargo, J.F., He, W., O'connell, R.J., Marconi, V.C., Delmar, J., Eron, J. and Clark, R.A., 2008. CCL3L1-CCR5 genotype influences durability of immune recovery during antiretroviral therapy of HIV-1 infected individuals. *Nature medicine*, *14*(4), pp.413-420.

Ali, S.A., Soo, C., Agongo, G., Alberts, M., Amenga-Etego, L., Boua, R.P., Choudhury, A., Crowther, N.J., Depuur, C., Gómez-Olivé, F.X. and Guiraud, I., 2018. Genomic and environmental risk factors for cardiometabolic diseases in Africa: methods used for Phase 1 of the AWI-Gen population cross-sectional study. *Global health action*, *11*(sup2), p.1507133.

Al-Jabri, A.A., 2007. Mechanisms of host resistance against HIV infection and progression to AIDS. *Sultan Qaboos University Medical Journal*, *7*(2), p.82.

Alkhatib, G., Combadiere, C., Broder, C.C., Feng, Y., Kennedy, P.E., Murphy, P.M. and Berger, E.A., 1996. CC CKR5: A RANTES, MIP-1α, MIP-1◌̀ Receptor as a Fusion Cofactor for Macrophage-Tropic HIV-1. *Science*, *272*(5270), pp.1955-1958.

Alkhatib, G., Liao, F., Berger, E.A., Farber, J.M. and Peden, K.W., 1997. A new SIV coreceptor, STRL33. *Nature*, *388*(6639), p.238.

Allen, S.J., Crown, S.E. and Handel, T.M., 2007. Chemokine: receptor structure, interactions, and antagonism. *Annu. Rev. Immunol.*, *25*, pp.787-820.

Allers, K., Hütter, G., Hofmann, J., Loddenkemper, C., Rieger, K., Thiel, E. and Schneider, T., 2011. Evidence for the cure of HIV infection by CCR5Δ32/Δ32 stem cell transplantation. *Blood, The Journal of the American Society of Hematology*, *117*(10), pp.2791-2799.

Alter, G., Heckerman, D., Schneidewind, A., Fadda, L., Kadie, C.M., Carlson, J.M., Oniangue Ndza, C., Martin, M., Li, B., Khakoo, S.I. and Carrington, M., 2011. HIV-1 adaptation to NK-cell-mediated immune pressure. *Nature*, *476*(7358), pp.96-100.

Altfeld, M. and Allen, T.M., 2006. Hitting HIV where it hurts: an alternative approach to HIV vaccine design. *Trends in immunology*, *27*(11), pp.504-510.

Ameisen, J.C. and Capron, A., 1991. Cell dysfunction and depletion in AIDS: the programmed cell death hypothesis. *Immunology today*, *12*(4), pp.102-105.

An, P., Martin, M.P., Nelson, G.W., Carrington, M., Smith, M.W., Gong, K., Vlahov, D., O'Brien, S.J. and Winkler, C.A., 2000. Influence of CCR5 promoter haplotypes on AIDS progression in African–Americans. *Aids*, *14*(14), pp.2117-2122.

Angela-Covino, D., Sabbatucci, M. and Fantuzzi, L., 2016. The CCL2/CCR2 axis in the pathogenesis of HIV-1 infection: a new cellular target for therapy?. *Current drug targets*, *17*(1), pp.76-110.

Ansari, A.A. and Silvestri, G. eds., 2014. *Natural hosts of SIV: implication in AIDS*. Newnes.

Archin, N.M., Sung, J.M., Garrido, C., Soriano-Sarabia, N. and Margolis, D.M., 2014. Eradicating HIV-1 infection: seeking to clear a persistent pathogen. *Nature Reviews Microbiology*, *12*(11), pp.750-764.

Arendt, V., Amand, M., Iserentant, G., Lemaire, M., Masquelier, C., Ndayisaba, G.F., Verhofstede, C., Karita, E., Allen, S., Chevigné, A. and Schmit, J.C., 2019. Predominance of the heterozygous CCR 5 delta-24 deletion in African individuals resistant to HIV infection might be related to a defect in CCR 5 addressing at the cell surface. *Journal of the International AIDS Society*, *22*(9), p.e25384.

Arenzana-Seisdedos, F. and Parmentier, M., 2006, December. Genetics of resistance to HIV infection: Role of co-receptors and co-receptor ligands. In *Seminars in immunology* (Vol.18, No. 6, pp. 387-403). Academic Press.

Ayala García, M.A., González Yebra, B., López Flores, A.L. and Guaní Guerra, E., 2012. The major histocompatibility complex in transplantation. *Journal of transplantation*, *2012*.

Bachelerie, F., Ben-Baruch, A., Burkhardt, A.M., Combadiere, C., Farber, J.M., Graham, G.J., Horuk, R., Sparre-Ulrich, A.H., Locati, M., Luster, A.D. and Mantovani, A., 2014. International Union of Basic and Clinical Pharmacology. LXXXIX. Update on the extended family of chemokine receptors and introducing a new nomenclature for atypicalchemokine receptors. *Pharmacological reviews*, *66*(1), pp.1-79.

Bachelerie, F., Graham, G.J., Locati, M., Mantovani, A., Murphy, P.M., Nibbs, R., Rot, A., Sozzani, S. and Thelen, M., 2015. An atypical addition to the chemokine receptor nomenclature: IUPHAR Review 15. *British journal of pharmacology*, *172*(16), pp.3945-3949.

Balakrishna, L.S. and Kondapi, A.K., 2016. Role of Host Proteins in HIV-1 Early Replication. *Advances in Molecular Retrovirology*, p.21.

Barakat-Haddad, C.P., Elliott, S.J., Eyles, J. and Pengelly, D., 2009. Predictors of locating children participants in epidemiological studies 20 years after last contact: internet resources and longitudinal research. *European journal of epidemiology*, *24*(8), pp.397 405.

Barcellos, L.F., Schito, A.M., Rimmler, J.B., Vittinghoff, E., Shih, A., Lincoln, R., Callier, S., Elkins, M.K., Goodkin, D.E., Haines, J.L. and Pericak-Vance, M.A., 2000. CC chemokine receptor 5 polymorphism and age of onset in familial multiple sclerosis. *Immunogenetics*, *51*(4-5), pp.281-288.

Barré-Sinoussi, F., Chermann, J.C., Rey, F., Nugeyre, M.T., Chamaret, S., Gruest, J., Dauguet, C., Axler-Blin, C., Vézinet-Brun, F., Rouzioux, C. and Rozenbaum, W., 1983. Isolation of a T-lymphotropic retrovirus from a patient at risk for acquired immune deficiency syndrome (AIDS). *Science*, *220*(4599), pp.868-871.

Bashirova, A.A., Thomas, R. and Carrington, M., 2011. HLA/KIR restraint of HIV: surviving the fittest. *Annual review of immunology*, *29*, pp.295-317.

Beering, A.R., Sun, H., Cummings, S.E., Crodian, J.S., Dowden, S.D., Zhang, L., Haas, D.M., Casey, T.M. and Ahmed, A., 2016. Factors affecting retention and compliance in a longitudinal study of connected, low income, urban, primiparous mothers. *Clin Res Trial*, *2*(4), pp.192-198.

Bhattacharya, T., Stanton, J., Kim, E. Y., Kunstman, K. J., Phair, J. P., Jacobson, L. P. & Wolinsky, S. M. 2009. CCL3L1 and HIV/AIDS susceptibility. *Nature Medicine,* 15, 1112-1115.

Biron, C.A., Nguyen, K.B., Pien, G.C., Cousens, L.P. and Salazar-Mather, T.P., 1999. Natural killer cells in antiviral defense: function and regulation by innate cytokines. Annual review of immunology, 17(1), pp.189-220.

Blaak, H., Boers, P.H.M., Gruters, R.A., Schuitemaker, H., Van Der Ende, M.E. and Osterhaus, A.D.M.E., 2005. CCR5, GPR15, and CXCR6 are major coreceptors of human immunodeficiency virus type 2 variants isolated from individuals with and without plasma viremia. *Journal of virology*, *79*(3), pp.1686-1700.

Blanpain, C., Lee, B., Tackoen, M., Puffer, B., Boom, A., Libert, F., Sharron, M., Wittamer, V., Vassart, G., Doms, R.W. and Parmentier, M., 2000. Multiple nonfunctional alleles of CCR5 are frequent in various human populations. *Blood, The Journal of the American Society of Hematology*, *96*(5), pp.1638-1645.

Bleek, W., 1981. Avoiding shame: the ethical context of abortion in Ghana. *Anthropological quarterly*, pp.203-209.

Bofill, M, Gombert, W, Borthwick, N J, Akbar, AN, McLaughlin, JE, Lee, CA, Johnson, MA, Pinching, A and Janossy, G (1995) Presence of CD3+CD8*Bcl-2 "w lymphocytes undergoing apoptosis and activated macrophages in lymph nodes of HIV-I+ patients. Am J Pathol 146, 1542.

Bonilla, F.A. and Oettgen, H.C., 2010. Adaptive immunity. *Journal of Allergy and Clinical Immunology*, *125*(2), pp.S33-S40.

Borges, L. and Cosman, D., 2000. LIRs/ILTs/MIRs, inhibitory and stimulatory Ig-superfamily receptors expressed in myeloid and lymphoid cells. *Cytokine & growth factor reviews*, *11*(3), pp.209-217.

Borrow, P., Lewicki, H., Hahn, B.H., Shaw, G.M. and Oldstone, M.B., 1994. Virus-specific CD8+ cytotoxic T-lymphocyte activity associated with control of viremia in primary human immunodeficiency virus type 1 infection. *Journal of virology*, *68*(9), pp.6103 6110.

Bream, J.H., 1999. Promoter Alleles and Specific DNA Binding Factors CCR5. *Science*, *284*, pp.223a-223.

Brenchley, J.M., Paiardini, M., Knox, K.S., Asher, A.I., Cervasi, B., Asher, T.E., Scheinberg, P., Price, D.A., Hage, C.A., Kholi, L.M. and Khoruts, A., 2008. Differential Th17 CD4 T-cell depletion in pathogenic and nonpathogenic lentiviral infections. *Blood, The Journal of the American Society of Hematology*, *112*(7), pp.2826-2835.

Brenchley, J.M., Price, D.A., Schacker, T.W., Asher, T.E., Silvestri, G., Rao, S., Kazzaz, Z., Bornstein, E., Lambotte, O., Altmann, D. and Blazar, B.R., 2006. Microbial translocation is a cause of systemic immune activation in chronic HIV infection. *Nature medicine*, *12*(12), pp.1365-1371.

Brenchley, J.M., Schacker, T.W., Ruff, L.E., Price, D.A., Taylor, J.H., Beilman, G.J., Nguyen, P.L., Khoruts, A., Larson, M., Haase, A.T. and Douek, D.C., 2004. CD4+ T cell depletion during all stages of HIV disease occurs predominantly in the gastrointestinal tract. *Journal of Experimental Medicine*, *200*(6), pp.749-759.

Broder, C.C. and Collman, R.G., 1997. Chemokine receptors and HIV. *Journal of leukocyte biology*, *62*(1), pp.20-29.

Brouwer, K.C., Yang, C., Parekh, S., Mirel, L.B., Shi, Y.P., Otieno, J., Lal, A.A. and Lal, R.B., 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*, *21*(5), pp.358-362.

Butler, I., MacLeod, W., Majuba, P.P. and Tipping, B., 2018. Human immunodeficiency virus infection and older adults: A retrospective single-site cohort study from Johannesburg, South Africa. *Southern African Journal of HIV Medicine*, *19*(1).

Caligiuri, M.A., 2008. Human natural killer cells. *Blood, The Journal of the American Society of Hematology*, *112*(3), pp.461-469.

Campbell, M.C. and Tishkoff, S.A., 2008. African genetic diversity: implications for human demographic history, modern human origins, and complex disease mapping. *Annu. Rev. Genomics Hum. Genet. 9*, pp.403-433.

Campbell, M.C. and Tishkoff, S.A., 2010. The evolution of human genetic and phenotypic variation in Africa. *Current biology*, *20*(4), pp.R166-R173.

Campbell-Yesufu, O.T. and Gandhi, R.T., 2011. Update on human immunodeficiency virus (HIV)-2 infection. *Clinical infectious diseases*, *52*(6), pp.780-787.

Cao, Y., Qin, L., Zhang, L., Safrit, J. and Ho, D.D., 1995. Virologic and immunologic characterization of long-term survivors of human immunodeficiency virus type 1 infection. *New England Journal of Medicine*, *332*(4), pp.201-208.

Capoulade-Métay, C., Ma, L., Truong, L.X., Dudoit, Y., Versmisse, P., Nguyen, N.V., Nguyen, M., Scott-Algara, D., Barré-Sinoussi, F., Debré, P. and Bismuth, G., 2004. New CCR5 variants associated with reduced HIV coreceptor function in southeast Asia. *Aids*, *18*(17), pp.2243-2252.

Carrington, M. and Alter, G., 2012. Innate immune control of HIV. *Cold Spring Harbor perspectives in medicine*, *2*(7), p.a007070.

Carrington, M. and Norman, P. 2003. The KIR Gene Cluster. NCBI, Bethesda, MD. 48 pp.

Carrington, M., Dean, M., Martin, M.P. and O'Brien, S.J., 1999. Genetics of HIV-1 infection: chemokine receptor CCR5 polymorphism and its consequences. *Human molecular genetics*, *8*(10), pp.1939-1945.

Carrington, M., Kissner, T., Gerrard, B., Ivanov, S., O'Brien, S.J. and Dean, M., 1997. Novel alleles of the chemokine-receptor gene CCR5. *The American Journal of Human Genetics*, *61*(6), pp.1261-1267.

Cartwright, E.K., McGary, C.S., Cervasi, B., Micci, L., Lawson, B., Elliott, S.T., Collman, R.G., Bosinger, S.E., Paiardini, M., Vanderford, T.H. and Chahroudi, A., 2014. Divergent CD4+ T memory stem cell dynamics in pathogenic and non-pathogenic simian immunodeficiency virus infections. *The Journal of Immunology*, *192*(10), pp.4666- 4673.

Catano, G., Chykarenko, Z.A., Mangano, A., Anaya, J.M., He, W., Smith, A., Bologna, R., Sen, L., Clark, R.A., Lloyd, A. and Shostakovich-Koretskaya, L., 2011. Concordance of CCR5 genotypes that influence cell-mediated immunity and HIV-1 disease progression rates. *Journal of Infectious Diseases*, *203*(2), pp.263-272.

Chatterjee, K., 2010. Host genetic factors in susceptibility to HIV-1 infection and progression to AIDS. *Journal of genetics*, *89*(1), pp.109-116.

Chawla, A., Wang, C., Patton, C., Murray, M., Punekar, Y., de Ruiter, A. and Steinhart, C., 2018. A review of long-term toxicity of antiretroviral treatment regimens and implications for an aging population. *Infectious diseases and therapy*, *7*(2), pp.183 195.

Chen, J.J.Y., Huang, J.C., Shirtliff, M., Briscoe, E., Ali, S., Cesani, F., Paar, D. and Cloyd, M.W., 2002. CD4 lymphocytes in the blood of HIV+ individuals migrate rapidly to lymph nodes and bone marrow: support for homing theory of CD4 cell depletion. *Journal of leukocyte biology*, *72*(2), pp.271-278.

Chen, Z., Kwon, D., Jin, Z., Monard, S., Telfer, P., Jones, M.S., Lu, C.Y., Aguilar, R.F., Ho, D.D. and Marx, P.A., 1998. Natural infection of a homozygous Δ24 CCR5 red-capped mangabey with an R2b-tropic simian immunodeficiency virus. *The Journal of experimental medicine*, *188*(11), pp.2057-2065.

Chun, T.W., Carruth, L., Finzi, D., Shen, X., DiGiuseppe, J.A., Taylor, H., Hermankova, M., Chadwick, K., Margolick, J., Quinn, T.C. and Kuo, Y.H., 1997. Quantification of latent tissue reservoirs and total body viral load in HIV-1 infection. *Nature*, *387*(6629), p.183.

Chung, A.W., Rollman, E., Center, R.J., Kent, S.J. and Stratov, I., 2009. Rapid degranulation of NK cells following activation by HIV-specific antibodies. The Journal of Immunology, 182(2), pp.1202-1210.

Clegg, A.O., Ashton, L.J., Biti, R.A., Badhwar, P., Williamson, P., Kaldor, J.M. and Stewart, G.J., 2000. CCR5 promoter polymorphisms, CCR5 59029A and CCR5 59353C, are underrepresented in HIV-1-infected long-term non-progressors. *Aids*, *14*(2), pp.103 108.

Cocchi, F., DeVico, A.L., Garzino-Demo, A., Arya, S.K., Gallo, R.C. and Lusso, P., 1995. Identification of RANTES, MIP-1α, and MIP-1β as the major HIV-suppressive factors produced by CD8+Tcells. *Science*, *270*(5243), pp.1811-1815.

Cocchi, F., DeVico, A.L., Yarchoan, R., Redfield, R., Cleghorn, F., Blattner, W.A., Garzino Demo, A., Colombini-Hatch, S., Margolis, D. and Gallo, R.C., 2000. Higher macrophage inflammatory protein (MIP)-1α and MIP-1β levels from CD8+T-cells are associated with asymptomatic HIV-1 infection. *Proceedings of the National Academy of Sciences*, *97*(25), pp.13812-13817.

Coffin, J. and Swanstrom, R., 2013. HIV pathogenesis: dynamics and genetics of viral populations and infected cells. *Cold Spring Harbor perspectives in medicine*, *3*(1), p.a012526.

Coffin, J., Haase, A., Levy, J.A., Montagnier, L., Oroszlan, S., Teich, N., Temin, H., Toyoshima, K., Varmus, H. and Vogt, P., 1986. Human immunodeficiency viruses. *Science*, *232*(4751), p.697.

Colobran, R., Comas, D., Faner, R., Pedrosa, E., Anglada, R., Pujol-Borrell, R., Bertranpetit, J. & Juan, M. 2008. Population structure in copy number variation and SNPs in the CCL4L chemokine gene. *Genes & Immunity*, 9, 279-288.

Colobran, R., Pedrosa, E., Carretero-Iglesia, L. and Juan, M., 2010. Copy number variation in chemokine superfamily: the complex scene of CCL3L–CCL4L genes in health and disease. *Clinical & Experimental Immunology*, *162*(1), pp.41-52.

D'Adamio, L., Shipp, M.A., Masteller, E.L. and Reinherz, E.L., 1989. Organization of the gene encoding common acute lymphoblastic leukemia antigen (neutral endopeptidase

24.11): multiple miniexons and separate 5'untranslated regions. *Proceedings of the National Academy of Sciences*, *86*(18), pp.7103-7107.

Davenport, M.P., Khoury, D.S., Cromer, D., Lewin, S.R., Kelleher, A.D. and Kent, S.J., 2019. Functional cure of HIV: the scale of the challenge. *Nature Reviews Immunology*, *19*(1), pp.45-54.

Day, C.L., Kaufmann, D.E., Kiepiela, P., Brown, J.A., Moodley, E.S., Reddy, S., Mackey, E.W., Miller, J.D., Leslie, A.J., DePierres, C. and Mncube, Z., 2006. PD-1 expression on HIV specific T cells is associated with T-cell exhaustion and disease progression. *Nature*, *443*(7109), pp.350-354.

de Medeiros, R.M., Valverde-Villegas, J.M., Junqueira, D.M., Gräf, T., Lindenau, J.D., de Mello, M.G., Vianna, P., Almeida, S.E. and Chies, J.A.B., 2016. Rapid and slow Progressors show increased IL-6 and IL-10 levels in the pre-AIDS stage of HIV infection. *PloS one*, *11*(5), p.e0156163.

de Munnik, S.M., Smit, M.J., Leurs, R. and Vischer, H.F., 2015. Modulation of cellular signaling by herpesvirus-encoded G protein-coupled receptors. *Frontiers in pharmacology*, *6*, p.40.

Dean, M., Carrington, M., Winkler, C., Huttley, G.A., Smith, M.W., Allikmets, R., Goedert, J.J., Buchbinder, S.P., Vittinghoff, E., Gomperts, E. and Donfield, S., 1996. Genetic restriction of HIV-1 infection and progression to AIDS by a deletion allele of the CKR5 structural gene. *Science*, *273*(5283), pp.1856-1862.

Deeks, S.G. and Walker, B.D., 2007. Human immunodeficiency virus controllers: mechanisms of durable virus control in the absence of antiretroviral therapy. Immunity, 27(3), pp.406 416.

Deeks, S.G., 2012. Shock and kill. *Nature*, *487*(7408), pp.439-440.

Deeks, S.G., Hoh, R., Grant, R.M., Wrin, T., Barbour, J.D., Narvaez, A., Cesar, D., Abe, K., Hanley, M.B., Hellmann, N.S. and Petropoulos, C.J., 2002. CD4+ T Cell Kinetics and Activation in Human Immunodeficiency Virus—Infected Patients Who Remain Viremic Despite Long-Term Treatment with Protease Inhibitor—Based Therapy. *The Journal of infectious diseases*, *185*(3), pp.315-323.

Deng, H., Unutmaz, D., KewalRamani, V.N. and Littman, D.R., 1997. Expression cloning of new receptors used by simian and human immunodeficiency viruses. *Nature*, *388*(6639), p.296.

Doitsh, G., Galloway, N.L., Geng, X., Yang, Z., Monroe, K.M., Zepeda, O., Hunt, P.W., Hatano, H., Sowinski, S., Muñoz-Arias, I. and Greene, W.C., 2014. Cell death by pyroptosis drives CD4 T-cell depletion in HIV-1 infection. *Nature*, *505*(7484), pp.509-514.

Doms, R.W. and Peiper, S.C., 1997. Unwelcomed guests with master keys: how HIV uses chemokine receptors for cellular entry. *Virology*, *235*(2), pp.179-190.

Duggal, P., An, P., Beaty, T.H., Strathdee, S.A., Farzadegan, H., Markham, R.B., Johnson, L., O'Brien, S.J., Vlahov, D. and Winkler, C.A., 2003. Genetic influence of CXCR6 chemokine receptor alleles on PCP-mediated AIDS progression among African Americans. *Genes and immunity*, *4*(4), p.245.

Edinger, A.L., Hoffman, T.L., Sharron, M., Lee, B., O'Dowd, B. and Doms, R.W., 1998. Use of GPR1, GPR15, and STRL33 as coreceptors by diverse human immunodeficiency

virus type 1 and simian immunodeficiency virus envelope proteins. *Virology*, *249*(2), pp.367-378.

Ellwanger, J.H., Zambra, F.M.B., Guimarães, R.L. and Chies, J.A.B., 2018. MicroRNA-related polymorphisms in infectious diseases—tiny changes with a huge impact on viral infections and potential clinical applications. *Frontiers in immunology*, *9*, p.1316.

Eltayeb, S., Berg, A.L., Lassmann, H., Wallström, E., Nilsson, M., Olsson, T., Ericsson Dahlstrand, A. and Sunnemark, D., 2007. Temporal expression and cellular origin of CC chemokine receptors CCR1, CCR2 and CCR5 in the central nervous system: insight into mechanisms of MOG-induced EAE. *Journal of Neuroinflammation*, *4*(1), p.14.

Etter, P., Landovitz, R., Sibeko, S., Sobieszczyk, M.E., Riddler, S.A., Karg, C., Tsibris, A. and Schouten, J., 2013. Recommendations for the follow-up of study participants with breakthrough HIV infections during HIV/AIDS biomedical prevention studies. *AIDS (London, England)*, *27*(7), p.1119.

European Pregnancy and Paediatric HIV Cohort Collaboration (EPPICC) study group in EuroCoord, Judd, A., Chappell, E., Turkova, A., Le Coeur, S., Noguera-Julian, A., Goetghebuer, T., Doerholt, K., Galli, L., Pajkrt, D. and Marques, L., 2018. Long-term trends in mortality and AIDS-defining events after combination ART initiation among children and adolescents with perinatal HIV infection in 17 middle-and high-income.

Fellay, J., Shianna, K.V., Ge, D., Colombo, S., Ledergerber, B., Weale, M., Zhang, K., Gumbs, C., Castagna, A., Cossarizza, A. and Cozzi-Lepri, A., 2007. A whole-genome association study of major determinants for host control of HIV-1. *science*, *317*(5840), pp.944-947.

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*, *272*(5263), pp.872-877.

Fiebig, E.W., Wright, D.J., Rawal, B.D., Garrett, P.E., Schumacher, R.T., Peddada, L., Heldebrant, C., Smith, R., Conrad, A., Kleinman, S.H. and Busch, M.P., 2003. Dynamics of HIV viremia and antibody seroconversion in plasma donors: implications for diagnosis and staging of primary HIV infection. *Aids*, *17*(13), pp.1871-1879.

Finzi, D., Hermankova, M., Pierson, T., Carruth, L.M., Buck, C., Chaisson, R.E., Quinn, T.C., Chadwick, K., Margolick, J., Brookmeyer, R. and Gallant, J., 1997. Identification of a reservoir for HIV-1 in patients on highly active antiretroviral therapy. *Science*, *278*(5341), pp.1295-1300.

Fox, M.P. and Rosen, S., 2010. Patient retention in antiretroviral therapy programs up to three years on treatment in sub-Saharan Africa, 2007–2009: systematic review. *Tropical medicine & international health*, *15*, pp.1-15.

Fox, M.P. and Rosen, S., 2015. Systematic review of retention of pediatric patients on HIV treatment in low and middle-income countries 2008–2013. *Aids*, *29*(4), pp.493-502.

Frange, P., Faye, A., Avettand-Fenoël, V., Bellaton, E., Descamps, D., Angin, M., David, A., Caillat-Zucman, S., Peytavin, G., Dollfus, C. and Le Chenadec, J., 2016. HIV-1 virological remission lasting more than 12 years after interruption of early antiretroviral therapy in a perinatally infected teenager enrolled in the French ANRS EPF-CO10 paediatric cohort: a case report. *The lancet HIV*, *3*(1), pp.e49-e54.

Gandhi, M., Bacchetti, P., Miotti, P., Quinn, T.C., Veronese, F. and Greenblatt, R.M., 2002. Does patient sex affect human immunodeficiency virus levels?. *Clinical Infectious Diseases*, *35*(3), pp.313-322.

Gelinas, L., Pierce, R., Winkler, S., Cohen, I.G., Lynch, H.F. and Bierer, B.E., 2017. Using social media as a research recruitment tool: ethical issues and recommendations. *The American Journal of Bioethics*, *17*(3), pp.3-14.

Gentle, N.L., Paximadis, M., Puren, A. and Tiemessen, C.T., 2013. Genetic variability in markers of HLA-C expression in two diverse South African populations. *PloS one*, *8*(7), p.e67780.

Gibbs, L., Manning, W.D., Longmore, M.A. and Giordano, P.C., 2014. Qualities of romantic relationships and consistent condom use among dating young adults.

Gilliet, M. and Lande, R., 2008. Antimicrobial peptides and self-DNA in autoimmune skin inflammation. Current opinion in immunology, 20(4), pp.401-407.

Giorgi, J.V., Hultin, L.E., McKeating, J.A., Johnson, T.D., Owens, B., Jacobson, L.P., Shih, R., Lewis, J., Wiley, D.J., Phair, J.P. and Wolinsky, S.M., 1999. Shorter survival in advanced human immunodeficiency virus type 1 infection is more closely associated with T lymphocyte activation than with plasma virus burden or virus chemokine coreceptor usage. *The Journal of infectious diseases*, *179*(4), pp.859-870.

Glencross, D.K., Janossy, G., Coetzee, L.M., Lawrie, D., Aggett, H.M., Scott, L.E., Sanne, I., McIntyre, J.A. and Stevens, W., 2008. Large-scale affordable PanLeucogated CD4+ testing with proactive internal and external quality assessment: In support of the South African national comprehensive care, treatment and management programme for HIV and AIDS. *Cytometry Part B: Clinical Cytometry: The Journal of the International Society for Analytical Cytology*, *74*(S1), pp.S40-S51.

Gong, Z., Tang, J., Deng, C. and Hu, G., 2017. Effect of CCR5-59029A/G polymorphism on risk of HIV-1 infection: a meta-analysis based on 13 studies. International Journal of Clinical and Experimental Medicine, 10(11), pp.15051-15059.

Gonzalez, E., Bamshad, M., Sato, N., Mummidi, S., Dhanda, R., Catano, G., Cabrera, S., McBride, M., Cao, X.H., Merrill, G. and O'Connell, P., 1999. Race-specific HIV-1 disease-modifying effects associated with CCR5 haplotypes. *Proceedings of the National Academy of Sciences*, *96*(21), pp.12004-12009.

Gonzalez, E., Dhanda, R., Bamshad, M., Mummidi, S., Geevarghese, R., Catano, G., Anderson, S.A., Walter, E.A., Stephan, K.T., Hammer, M.F. and Mangano, A., 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*, *98*(9), pp.5199 5204.

Gonzalez, E., Kulkarni, H., Bolivar, H., Mangano, A., Sanchez, R., Catano, G., Nibbs, R.J., Freedman, B.I., Quinones, M.P., Bamshad, M.J. and Murthy, K.K., 2005. The influence of CCL3L1 gene-containing segmental duplications on HIV-1/AIDS susceptibility. *Science*, *307*(5714), pp.1434-1440.

Gonzalo-Gil, E., Ikediobi, U. and Sutton, R.E., 2017. Focus: Infectious Diseases: Mechanisms of Virologic Control and Clinical Characteristics of HIV+ Elite/Viremic Controllers. *The Yale journal of biology and medicine*, *90*(2), p.245.

Gore-Felton, C. and Koopman, C., 2008. Behavioral mediation of the relationship between psychosocial factors and HIV disease progression. *Psychosomatic Medicine*, *70*(5), pp.569-574.

Gornalusse, G., Mummidi, S., He, W., Silvestri, G., Bamshad, M. and Ahuja, S.K., 2009. CCL3L Copy number variation and the co-evolution of primate and viral genomes. *PLoS genetics*, *5*(1), p.e1000359.

Gornalusse, G.G., Mummidi, S., Gaitan, A.A., Jimenez, F., Ramsuran, V., Picton, A., Rogers, K., Manoharan, M.S., Avadhanam, N., Murthy, K.K. and Martinez, H., 2015. Epigenetic mechanisms, Tcell activation, and CCR5 genetics interact to regulate T-cell expression of CCR5, the major HIV-1 coreceptor. *Proceedings of the National Academy of Sciences*, *112*(34), pp.E4762-E4771.

Gougeon, M.L., Lecoeur, H., Dulioust, A., Enouf, M.G., Crouvoiser, M., Goujard, C., Debord, T.and Montagnier, L., 1996. Programmed cell death in peripheral lymphocytes from HIV infected persons: increased susceptibility to apoptosis of CD4 and CD8 T cells correlates withlymphocyte activation and with disease progression. *The Journal of Immunology*, *156*(9), pp.3509-3520.

Goulder, P.J., Bunce, M., Krausa, P., Mcintyre, K.A.R.L., Crowley, S., Morgan, B., Edwards, A., Giangrande, P., Phillips, R.E. And Mcmichael, A.J., 1996. Novel, cross-restricted, conserved, and immunodominant cytotoxic T lymphocyte epitopes in slow Progressors in HIV type 1 infection. *AIDS research and human retroviruses*, *12*(18), pp.1691-1698.

Grabar, S., Selinger-Leneman, H., Abgrall, S., Pialoux, G., Weiss, L. and Costagliola, D., 2009. Prevalence and comparative characteristics of long-term nonProgressors and HIV controller patients in the French Hospital Database on HIV. *Aids*, *23*(9), pp.1163-1169.

Grossman, Z., Meier-Schellersheim, M., Sousa, A.E., Victorino, R.M. and Paul, W.E., 2002. CD4+ T-cell depletion in HIV infection: are we closer to understanding the cause?. *Naturemedicine*, *8*(4), pp.319-323.

Gupta, R., Sultan, A.J., McCoy, L., Mok, H., Peppa, D., Salgado, M., Martinez-Picado, J., Nijhuis, M., Wensing, A., Lee, H. and Grant, P., 2019. HIV-1 remission following CCR5Δ32/Δ32 haematopoietic stem cell transplantation.*Nature*, p.1.

Gurdasani, D., Iles, L., Dillon, D.G., Young, E.H., Olson, A.D., Naranbhai, V., Fidler, S., Gkrania-Klotsas, E., Post, F.A., Kellam, P. and Porter, K., 2014. A systematic review of definitions of extreme phenotypes of HIV control and progression. *AIDS (London, England)*, *28*(2), p.149.

Han, X., Becker, K., Degen, H.J., Jablonowski, H. and Strohmeyer, G., 1996. Synergistic stimulatory effects of tumour necrosis factor α and interferon γ on replication of human immunodeficiency virus type 1 and on apoptosis of HIV-1-infected host cells. *European journal of clinical investigation*, *26*(4), pp.286-292.

Hansson, G.K., Libby, P., Schönbeck, U. and Yan, Z.Q., 2002. Innate and adaptive immunity in the pathogenesis of atherosclerosis. *Circulation research*, *91*(4), pp.281-291.

HapMap Project. Available at: (*http://hapmap.ncbi.nlm.nih.gov*).

Hatano, H., Delwart, E.L., Norris, P.J., Lee, T.H., Neilands, T.B., Kelley, C.F., Hunt, P.W.,Hoh R., Linnen, J.M., Martin, J.N. and Busch, M.P., 2010. Evidence of persistent low-level viremia in long-term HAART-suppressed, HIV-infected individuals. *AIDS (London, England)*, *24*(16), p.2535.

Heerman, W.J., Jackson, N., Roumie, C.L., Harris, P.A., Rosenbloom, S.T., Pulley, J., Wilkins, C.H., Williams, N.A., Crenshaw, D., Leak, C. and Scherdin, J., 2017. Recruitment methods for survey research: findings from the mid-south clinical data research network. *Contemporary Clinical Trials*, *62*, pp.50-55.

Hellerstein, M., Hanley, M.B., Cesar, D., Siler, S., Papageorgopoulos, C., Wieder, E., Schmidt, D., Hoh, R., Neese, R., Macallan, D. and Deeks, S., 1999. Directly measured kinetics of circulating T lymphocytes in normal and HIV-1-infected humans. *Nature medicine*, *5*(1), pp.83-89.

Hindmarsh, P. and Leis, J., 1999. Retroviral DNA integration. *Microbiology and Molecular Biology Reviews*, *63*(4), pp.836-843.

Hladik, F., Liu, H., Speelmon, E., Livingston-Rosanoff, D., Wilson, S., Sakchalathorn, P., Hwangbo, Y., Greene, B., Zhu, T. and McElrath, M.J., 2005. Combined effect of CCR5 Δ32 heterozygosity and the CCR5 promoter polymorphism− 2459 A/G on CCR5 expression and resistance to human immunodeficiency virus type 1 transmission. *Journal of virology*, *79*(18), pp.11677-11684.

Hobden, K., Curtis Forney, J., Wyszacki Durham, K. and Toro, P., 2011. Limiting attrition in longitudinal research on homeless adolescents: what works best?. *Journal of Community Psychology*, *39*(4), pp.443-451.

Hofnagel, O., Engel, T., Severs, N.J., Robenek, H. and Buers, I., 2011. SR-PSOX at sites predisposed to atherosclerotic lesion formation mediates monocyte-endothelial cell adhesion. *Atherosclerosis*, *217*(2), pp.371-378.

Hong, H.A., Paximadis, M., Gray, G.E., Kuhn, L. and Tiemessen, C.T., 2013. KIR2DS4 allelic variants: differential effects on in utero and intrapartum HIV-1 mother-to-child transmission. *Clinical Immunology*, *149*(3), pp.498-508.

Horuk, R., 1999. Chemokine receptors and HIV-1: the fusion of two major research fields. *Immunology today*, *20*(2), pp.89-94.

Http://www.fda.gov/media/119796/download. Accessed on the 28th January 2020.

Hu, W.S. and Hughes, S.H., 2012. HIV-1 reverse transcription. *Cold Spring Harbor perspective in medicine*, *2*(10), p.a006882.

Human Sciences Research Council (HSRC), 2018. The fifth South African national HIV prevalence, incidence, behaviour and communication survey, 2017: HIV impact assessment summary report.

Hunt, P.W., Brenchley, J., Sinclair, E., McCune, J.M., Roland, M., Shafer, K.P., Hsue, P., Emu, B., Krone, M., Lampiris, H. and Douek, D., 2008. Relationship between T cell activation and CD4+ T cell count in HIV-seropositive individuals with undetectable plasma HIV RNA levels in the absence of therapy. *The Journal of infectious diseases*, *197*(1), pp.126-133.

Hutchinson, J.F., 2001. The biology and evolution of HIV. *Annual review of anthropology*, pp.85-108.

Hütter, G., Nowak, D., Mossner, M., Ganepola, S., Müßig, A., Allers, K., Schneider, T., Hofmann, J., Kücherer, C., Blau, O. and Blau, I.W., 2009. Long-term control of HIV by CCR5 Delta32/Delta32 stemcell transplantation. *New England Journal of Medicine*, *360*(7), pp.692-698.

Ingulli, E., 2010. Mechanism of cellular rejection in transplantation. *Pediatric nephrology*, *25*(1), p.61.

Ioannidis, J.P., Rosenberg, P.S., Goedert, J.J., Ashton, L.J., Benfield, T.L., Buchbinder, S.P., Coutinho, R.A., Eugen-Olsen, J., Gallart, T., Katzenstein, T.L. and Kostrikis, L.G., 2001. Effects of CCR5-Δ 32, CCR2-64I, and SDF-1 3′ A alleles on HIV-1 disease progression: an international meta-analysis of individual-patient data. *Annals of internal medicine*, *135*(9), pp.782-795.

Jakobsen, M.R., Cashin, K., Roche, M., Sterjovs.ki, J., Ellett, A., Borm, K., Flynn, J., Erikstrup, C., Gouillou, M., Gray, L.R. and Saksena, N.K., 2013. Longitudinal analysis of CCR5 and CXCR4 usage in a cohort of antiretroviral therapy-naive subjects with progressive HIV-1 subtype C infection. *PLoS One*, *8*(6), p.e65950.

Janeway, C.A., 1989, January. Approaching the asymptote? Evolution and revolution in mmunology. In *Cold Spring Harbor symposia on quantitative biology* (Vol. 54, pp. 1 13). Cold Spring Harbor Laboratory Press.

Jaumdally, S.Z., Picton, A., Tiemessen, C.T., Paximadis, M., Jaspan, H.B., Gamieldien, H., Masson, L., Coetzee, D., Williamson, A.L., Little, F. and Gumbi, P.P., 2017. CCR 5 expression, haplotype and immune activation in protection from infection in HIV exposed uninfected individuals in HIV serodiscordant relationships. *Immunology*, *151*(4), pp.464-473.

Jennes, W., Verheyden, S., Demanet, C., Adjé-Touré, C.A., Vuylsteke, B., Nkengasong, J.N. and Kestens, L., 2006. Cutting edge: resistance to HIV-1 infection among African female sex workers is associated with inhibitory KIR in the absence of their HLA ligands. *The Journal of Immunology*, *177*(10), pp.6588-6592.

Johnston, L., O'Bra, H., Chopra, M., Mathews, C., Townsend, L., Sabin, K., Tomlinson, M. and Kendall, C., 2010. The associations of voluntary counseling and testing acceptance and the perceived likelihood of being HIV-infected among men with multiple sex partners in a South African township. *AIDS and Behavior*, *14*(4), pp.922-931.

Jost, S. and Altfeld, M., 2012. Evasion from NK cell-mediated immune responses by HIV 1. *Microbes and infection*, *14*(11), pp.904-915.

Kallings, L.O., 2008. The first postmodern pandemic: 25 years of HIV/AIDS. *Journal of internal medicine*, *263*(3), pp.218-243.

Kaslow, R.A., Carrington, M., Apple, R., Park, L., Muñoz, A., Saah, A.J., Goedert, J.J., Winkler, C., O'Brien, S.J., Rinaldo, C. and Detels, R., 1996. Influence of combinations of human major histocompatibility complex genes on the course of HIV–1 infection. *Nature medicine*, *2*(4), p.405.

Kaur, G. and Mehra, N., 2009. Genetic determinants of HIV-1 infection and progression to AIDS: susceptibility to HIV infection. *Tissue antigens*, *73*(4), pp.289-301.

Kaur, G., Singh, P., Rapthap, C.C., Kumar, N., Vajpayee, M., Sharma, S.K., Wanchu, A. and Mehra, N.K., 2007. Polymorphism in the CCR5 gene promoter and HIV-1 infection in North Indians. *Human immunology*, *68*(5), pp.454-461.

Kawashima, Y., Pfafferott, K., Frater, J., Matthews, P., Payne, R., Addo, M., Gatanaga, H., Fujiwara, M., Hachiya, A., Koizumi, H. and Kuse, N., 2009. Adaptation of HIV-1 to human leukocyte antigen class I. *Nature*, *458*(7238), p.641.

Keele, B.F., Jones, J.H., Terio, K.A., Estes, J.D., Rudicell, R.S., Wilson, M.L., Li, Y., Learn, G.H., Beasley, T.M., Schumacher-Stankey, J. and Wroblewski, E., 2009. Increased mortality and AIDS-like immunopathology in wild chimpanzees infected with SIVcpz. *Nature*, *460*(7254), pp.515-519.

Keeley, E.C., Mehrad, B. and Strieter, R.M., 2011. Chemokines as mediators of tumor angiogenesis and neovascularization. *Experimental cell research*, *317*(5), pp.685-690.

Khakoo, S.I., Thio, C.L., Martin, M.P., Brooks, C.R., Gao, X., Astemborski, J., Cheng, J., Goedert, J.J., Vlahov, D., Hilgartner, M. and Cox, S., 2004. HLA and NK cell inhibitory receptor genes in resolving hepatitis C virus infection. *Science*, *305*(5685), pp.872-874.

Khorramdelazad, H., Hakimizadeh, E., Hassanshahi, G., Rezayati, M., Sendi, H. and Arababadi, M.K., 2013. CCR5 Δ 32 mutation is not prevalent in Iranians with chronic HBV infection. *Journal of medical virology*, *85*(6), pp.964-968.

Kiepiela, P., Leslie, A.J., Honeyborne, I., Ramduth, D., Thobakgale, C., Chetty, S., Rathnavalu, P., Moore, C., Pfafferott, K.J., Hilton, L. and Zimbwa, P., 2004. Dominant influence of HLA-B in mediating the potential co-evolution of HIV and HLA. *Nature*, *432*(7018), p.769.

Kim, C.H., Kunkel, E.J., Boisvert, J., Johnston, B., Campbell, J.J., Genovese, M.C., Greenberg, H.B. and Butcher, E.C., 2001. Bonzo/CXCR6 expression defines type 1–polarized T cell subsets with extralymphoid tissue homing potential. *The Journal of clinical investigation*, *107*(5), pp.595-601.

Klatt, N. R., Chomont, N., Douek, D. C. & Deeks, S. G. 2013. Immune activation and HIV persistence: implications for curative approaches to HIV infection. *Immunol Rev,* 254**,** 326-42.

Kløverpris, H.N., Leslie, A. and Goulder, P., 2016. Role of HLA adaptation in HIV evolution. *Frontiers in immunology*, *6*, p.665.

Knudsen, T.B., Kristiansen, T.B., Katzenstein, T.L. and Eugen-Olsen, J., 2001. Adverse effect of the CCR5 promoter− 2459A allele on HIV-1 disease progression. *Journal of medical virology*, *65*(3), pp.441-444.

Koor, G.W., Paximadis, M., Picton, A.C., Karatas, F., Loubser, S., Waja, Z., Ebrahim, O., Martinsen, N. and Tiemessen, C.T., 2019, (manuscript submitted in October). CCR5 Expression-related Gene Variants and Natural HIV-1 Control in Black South Africans. In *AIDS Research And Human Retroviruses* (Vol. 34, Pp. 248-248).

Kostrikis, L.G., Huang, Y., Moore, J.P., Wolinsky, S.M., Zhang, L., Guo, Y., Deutsch, L., Phair, J., Neumann, A.U. and Ho, D.D., 1998. A chemokine receptor CCR2 allele delays HIV1 disease progression and is associated with a CCR5 promoter mutation. *Nature medicine*, *4*(3), pp.350-353.

Kulkarni, S., Lied, A., Kulkarni, V., Rucevic, M., Martin, M.P., Walker-Sperling, V., Anderson, S.K., Ewy, R., Singh, S., Nguyen, H. and McLaren, P.J., 2019. CCR5AS lncRNA variation differentially regulates CCR5, influencing HIV disease outcome. *Nature immunology*, *20*(7), pp.824-834.

Lackner, A.A., Lederman, M.M. and Rodriguez, B., 2012. HIV pathogenesis: the host. *Cold Spring Harbor perspectives in medicine*, *2*(9), p.a007005.

Laila, U., Akram, M., Shariati, M.A., Hashmi, A.M., Akhtar, N., Tahir, I.M., Ghauri, A.O., Munir, N., Riaz, M., Akhter, N. and Shaheen, G., 2019. Role of medicinal plants in

HIV/AIDS therapy. *Clinical and Experimental Pharmacology and Physiology*, *46*(12), pp.1063-1073.

Laing, K.J. and Secombes, C.J., 2004. Chemokines. *Developmental & Comparative Immunology*, *28*(5), pp.443-460.

Landrø, L., Damås, J.K., Halvorsen, B., Fevang, B., Ueland, T., Otterdal, K., Heggelund, L., Frøland, S.S. and Aukrust, P., 2009. CXCL16 in HIV infection–a link between inflammation and viral replication. *European journal of clinical investigation*, *39*(11), pp.1017-1024.

Lanier, L. L. 2008. Up on the tightrope: natural killer cell activation and inhibition. *Nat Immunol*, 9**,** 495-502.

Lee, B., Doranz, B.J., Rana, S., Yi, Y., Mellado, M., Frade, J.M., Martinez-A, C., O'Brien, S.J.,Dean, M., Collman, R.G. and Doms, R.W., 1998. Influence of the CCR2-V64I polymorphism on human immunodeficiency virus type 1 coreceptor activity and on chemokine receptor function of CCR2b, CCR3, CCR5, and CXCR4. *Journal of Virology*, *72*(9), pp.7450-7458.

Lee, C., Dobson, A., Brown, W., Adamson, L. and Goldsworthy, J., 2000. Tracking participants:lessons from the Women's Health Australia Project. *Australian and New Zealand journal of public health*, *24*(3), pp.334-336.

Li, M., Song, R., Masciotra, S., Soriano, V., Spira, T.J., Lal, R.B. and Yang, C., 2005. Association of CCR5 human haplogroup E with rapid HIV type 1 disease progression. *AIDS Research & Human Retroviruses*, *21*(2), pp.111-115.

Liao, F., Alkhatib, G., Peden, K.W., Sharma, G., Berger, E.A. and Farber, J.M., 1997. STRL33, A novel chemokine receptor–like protein, functions as a fusion cofactor for both macrophage-tropic and T-Cell line–tropic HIV-1. *Journal of Experimental Medicine*, *185*(11), pp.2015-2023.

Lim, S.Y., Chan, T., Gelman, R.S., Whitney, J.B., O'Brien, K.L., Barouch, D.H., Goldstein, D.B., Haynes, B.F. and Letvin, N.L., 2010. Contributions of Mamu-A* 01 status and TRIM5 allele expression, but not CCL3L copy number variation, to the control of SIVmac251 replication in Indian-origin rhesus monkeys. *PLoS genetics*, *6*(6), p.e1000997.

Limou, S., Coulonges, C., Herbeck, J.T., Van Manen, D., An, P., Le Clerc, S., Delaneau, O., Diop, G., Taing, L., Montes, M. and Van't Wout, A.B., 2010. Multiple-cohort genetic association study reveals CXCR6 as a new chemokine receptor involved in long-term no progression to AIDS. *The Journal of infectious diseases*, *202*(6), pp.908-915.

Liu, R., Paxton, W.A., Choe, S., Ceradini, D., Martin, S.R., Horuk, R., MacDonald, M.E., Stuhlmann, H., Koup, R.A. and Landau, N.R., 1996. Homozygous defect in HIV-1 coreceptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. *Cell*, *86*(3), pp.367-377.

Lohse, N., Hansen, A.B.E., Pedersen, G., Kronborg, G., Gerstoft, J., Sørensen, H.T., Vaeth, M. and Obel, N., 2007. Survival of persons with and without HIV infection in Denmark, 1995–2005. *Annals of internal medicine*, *146*(2), pp.87-95.

Long, E.O., Colonna, M. and Lanier, L.L., 1996. Inhibitory MHC class I receptors on NK and T cells: a standard nomenclature. Immunology today, 17(2).

Louisirirotchanakul, S., Liu, H., Roongpisuthipong, A., Nakayama, E.E., Takebe, Y., Shioda, T. and Wasi, C., 2002. Genetic analysis of HIV-1 discordant couples in Thailand: association of CCR2 64I homozygosity with HIV-1-negative status. *JAIDS Journal of Acquired Immune Deficiency Syndromes*, *29*(3), pp.314-315.

Luciw, P.A, 1996. Human immunodeficiency viruses and their replication. In: Fields BN, editor. *Philadelphia: Lippincott-Raven Virology. 3rd ed.*, pp. 1881–1952

Luzuriaga, K., Gay, H., Ziemniak, C., Sanborn, K.B., Somasundaran, M., Rainwater-Lovett, K., Mellors, J.W., Rosenbloom, D. and Persaud, D., 2015. Viremic relapse after HIV 1 remission in a perinatally infected child. *New England Journal of Medicine*, *372*(8), pp.786-788.

Lyles, R.H., Muñoz, A., Yamashita, T.E., Bazmi, H., Detels, R., Rinaldo, C.R., Margolick, J.B., Phair, J.P. and Mellors, J.W., 2000. Natural history of human immunodeficiency virus type 1 viremia after seroconversion and proximal to AIDS in a large cohort of homosexual men. *The Journal of infectious diseases*, *181*(3), pp.872-880.

Madec, Y., Boufassa, F., Rouzioux, C., Delfraissy, J.F., Meyer, L. and SEROCO Study Group, 2005. Undetectable viremia without antiretroviral therapy in patients with HIV seroconversion: an uncommon phenomenon?. *Clinical infectious diseases*, *40*(9), pp.1350-1354.

Mangano, A., Kopka, J., Batalla, M., Bologna, R. and Sen, L., 2000. Protective effect of CCR2 64I and not of CCR5-delta32 and SDF1-3'A in pediatric HIV-1 infection. *Journal of acquired immune deficiency syndromes (1999)*, *23*(1), pp.52-57.

Marlink, R., Kanki, P., Thior, I., Travers, K., Eisen, G., Siby, T., Traore, I., Hsieh, C.C., Dia, M.C. and Gueye, E.H., 1994. Reduced rate of disease development after HIV-2 infection as compared to HIV-1. *Science*, *265*(5178), pp.1587-1590.

Martin, M. P., Gao, X., Lee, J. H., Nelson, G. W., Detels, R., Goedert, J. J., Buchbinder, S., Hoots, K., Vlahov, D., Trowsdale, J., Wilson, M., O'brien, S. J. & Carrington, M. 2002. Epistatic interaction between KIR3DS1 and HLA-B delays the progression to AIDS. *Nat Genet*, 31**,** 429-34.

Martin, M. P., Qi, Y., Gao, X., Yamada, E., Martin, J. N., Pereyra, F., Colombo, S., Brown, E. E., Shupert, W. L., Phair, J., Goedert, J. J., Buchbinder, S., Kirk, G. D., Telenti, A.,Connors, M., O'Brien, S. J., Walker, B. D., Parham, P., Deeks, S. G., McVicar, D. W.& Carrington, M. 2007. Innate partnership of HLA-B and KIR3DL1 subtypes against HIV-1. *Nature Genetics,* 39, 733-740.

Martin, M.P., Dean, M., Smith, M.W., Winkler, C., Gerrard, B., Michael, N.L., Lee, B., Doms, R.W., Margolick, J., Buchbinder, S. and Goedert, J.J., 1998. Genetic acceleration of AIDS progression by a promoter variant of CCR5. *Science*, *282*(5395), pp.1907-1911.

Masquelier, C., Servais, J.Y., Rusanganwa, E., Roman, F., Havuga, E., Servais, J., Tuyizere, S., Omes, C., Karasi, J.C., Coruteille, O. and Karita, E., 2007. A novel 24-base pair deletion in the coding region of CCR5 in an African population. *Aids*, *21*(1), pp.111 113.

Matloubian, M., David, A., Engel, S., Ryan, J.E. and Cyster, J.G., 2000. A transmembrane CXC chemokine is a ligand for HIV-coreceptor Bonzo. *Nature immunology*, *1*(4), p.298

Matsuyama, T., Kobayashi, N. and Yamamoto, N., 1991. Is AIDS a TNF disease?. *Aids*, *5*(12), pp.1405-17.

Mattapallil, J.J., Douek, D.C., Hill, B., Nishimura, Y., Martin, M. and Roederer, M., 2005. Massive infection and loss of memory CD4+ T cells in multiple tissues during acute SIV infection. *Nature*, *434*(7037), pp.1093-1097.

Matzaraki, V., Kumar, V., Wijmenga, C. and Zhernakova, A., 2017. The MHC locus and genetic susceptibility to autoimmune and infectious diseases. *Genome biology*, *18*(1), p.76.

Maurano, M.T., Humbert, R., Rynes, E., Thurman, R.E., Haugen, E., Wang, H., Reynolds, A.P., Sandstrom, R., Qu, H., Brody, J. and Shafer, A., 2012. Systematic localization of common disease-associated variation in regulatory DNA. *Science*, *337*(6099), pp.1190-1195.

May, A., Hazelhurst, S., Li, Y., Norris, S.A., Govind, N., Tikly, M., Hon, C., Johnson, K.J., Hartmann, N., Staedtler, F. and Ramsay, M., 2013. Genetic diversity in black South Africans from Soweto. *BMC genomics*, *14*(1), p.644.

Mburu, G., Ram, M., Siu, G., Bitira, D., Skovdal, M. and Holland, P., 2014. Intersectionality of HIV stigma and masculinity in eastern Uganda: implications for involving men in HIV programmes. *BMC public health*, *14*(1), p.1061.

McDermott, D.H., Zimmerman, P.A., Guignard, F., Kleeberger, C.A., Leitman, S.F., Murphy, P.M. and Multicenter AIDS Cohort Study (MACS), 1998. CCR5 promoter polymorphism and HIV-1 disease progression. *The Lancet*, *352*(9131), pp.866-870.

McLaren, P.J., Coulonges, C., Bartha, I., Lenz, T.L., Deutsch, A.J., Bashirova, A., Buchbinder, S., Carrington, M.N., Cossarizza, A., Dalmau, J. and De Luca, A., 2015. Polymorphisms of large effect explain the majority of the host genetic contribution to variation of HIV-1 virus load. *Proceedings of the National Academy of Sciences*, *112*(47), pp.14658-14663.

McLaren, P.J., Pulit, S.L., Gurdasani, D., Bartha, I., Shea, P.R., Pomilla, C., Gupta, N., Gkrania Klotsas, E., Young, E.H., Bannert, N. and Del Amo, J., 2017. Evaluating the impact of functional genetic variation on HIV-1 control. *The Journal of infectious diseases*, *216*(9), pp.1063-1069.

Meddows-Taylor, S., Donninger, S.L., Paximadis, M., Schramm, D.B., Anthony, F.S., Gray, G.E., Kuhn, L. and Tiemessen, C.T., 2006. Reduced ability of newborns to produce CCL3 is associated with increased susceptibility to perinatal human immunodeficiency virus 1 transmission. *The Journal of general virology*, *87*(Pt 7), p.2055.

Meintjes, G., Black, J., Conradie, F., Dlamini, S., Maartens, G., Manzini, T.C., Mathe, M., Moorhouse, M., Moosa, Y., Nash, J. and Orrell, C., 2015. Southern African HIV Clinicians Society adult antiretroviral therapy guidelines: Update on when to initiate antiretroviral therapy. *Southern African journal of HIV medicine*, *16*(1), pp.1-4.

Mellors, J.W., Kingsley, L.A., Rinaldo, C.R., Todd, J.A., Hoo, B.S., Kokka, R.P. and Gupta, P., 1995. Quantitation of HIV-1 RNA in plasma predicts outcome after seroconversion. *Annals of internal medicine*, *122*(8), pp.573-579.

Mellors, J.W., Margolick, J.B., Phair, J.P., Rinaldo, C.R., Detels, R., Jacobson, L.P. and Muñoz, A., 2007. Prognostic value of HIV-1 RNA, CD4 cell count, and CD4 Cell count slope for progression to AIDS and death in untreated HIV-1 infection. Jama, 297(21), pp.2346-2350.

Mens, H., Kearney, M., Wiegand, A., Shao, W., Schønning, K., Gerstoft, J., Obel, N., Maldarelli, F., Mellors, J.W., Benfield, T. and Coffin, J.M., 2010. HIV-1 continues to

replicate and evolve in patients with natural control of HIV infection. *Journal of virology*, *84*(24), pp.12971-12981.

Meyaard, L., Otto, S.A., Jonker, R.R., Mijnster, M.J., Keet, R.P. and Miedema, F., 1992. Programmed death of T cells in HIV-1 infection. *Science*, *257*(5067), pp.217-219.

Modi, W.S., 2004. CCL3L1 and CCL4L1 chemokine genes are located in a segmental duplication at chromosome 17q12. *Genomics*, *83*(4), pp.735-738.

Moesta, A.K. and Parham, P., 2012. Diverse functionality among human NK cell receptors for the C1 epitope of HLA-C: KIR2DS2, KIR2DL2, and KIR2DL3. *Frontiers in immunology*, *3*, p.336.

Moore, R.D. and Cheever, L., 1999. Lack of sex difference in CD4 to HIV-1 RNA viral load ratio. *The Lancet*, *353*(9151), pp.463-464.

Mori, M., Leitman, E., Walker, B., Ndung'u, T., Carrington, M. and Goulder, P., 2019. Impact of HLA Allele-KIR Pairs on HIV Clinical Outcome in South Africa. *The Journal of infectious diseases*, *219*(9), pp.1456-1463.

Mori, M., Wichukchinda, N., Miyahara, R., Rojanawiwat, A., Pathipvanich, P., Tsuchiya, N., Miura, T., Yasunami, M., Ariyoshi, K. and Sawanpanyalert, P., 2015. The effect of KIR2D–HLA-C receptor–ligand interactions on clinical outcome in a HIV-1 CRF01_AE-infected Thai population. Aids, 29(13), pp.1607-1615.

Moriuchi, H., Moriuchi, M. and Fauci, A.S., 1997. Cloning and analysis of the promoter region of CCR5, a coreceptor for HIV-1 entry. *The Journal of Immunology*, *159*(11), pp.5441 5449.

Mulherin, S.A., O'Brien, T.R., Ioannidis, J.P., Goedert, J.J., Buchbinder, S.P., Coutinho, R.A., Jamieson, B.D., Meyer, L., Michael, N.L., Pantaleo, G. and Rizzardi, G.P., 2003. Effects of CCR5-Δ32 and CCR2-64I alleles on HIV-1 disease progression: the protection varies with duration of infection. *Aids*, *17*(3), pp.377-387.

Mummidi, S., Adams, L.M., VanCompernolle, S.E., Kalkonde, M., Camargo, J.F., Kulkarni, H., Bellinger, A.S., Bonello, G., Tagoh, H., Ahuja, S.S. and Unutmaz, D., 2007. Production of specific mRNA transcripts, usage of an alternate promoter, and octamer binding transcription factors influence the surface expression levels of the HIV coreceptor CCR5 on primary T cells. *The Journal of Immunology*, *178*(9), pp.5668 5681.

Mummidi, S., Ahuja, S.S., Gonalez, E., Anderson, S.A., Santiago, E.N., Stephan, K.T., Craig, F.E., O'Connell, P., Tryon, V., Clark, R.A. and Dolan, M.J., 1998. Genealogy of the CCR5 locus and chemokine system gene variants associated with altered rates of HIV 1 disease progression. *Nature medicine*, *4*(7), p.786.

Mummidi, S., Ahuja, S.S., McDaniel, B.L. and Ahuja, S.K., 1997. The Human CC Chemokine Receptor 5 (CCR5) gene multiple transcripts with 5′-end heterogeneity, dual promoter usage, and evidence for polymorphisms within the regulatory regions and noncoding exons. *Journal of Biological Chemistry*, *272*(49), pp.30662-30671.

Mummidi, S., Bamshad, M., Ahuja, S.S., Gonzalez, E., Feuillet, P.M., Begum, K., Galvis, M.C., Kostecki, V., Valente, A.J., Murthy, K.K. and Haro, L., 2000. Evolution of human and non-human primate CC chemokine receptor 5 gene and mRNA Potential roles for haplotype and mRNA diversity, differential haplotype-specific transcriptional activity, and altered transcription factor binding to polymorphic nucleotides in the

pathogenesis of HIV-1 and simian immunodeficiency virus. *Journal of Biological Chemistry*, *275*(25), pp.18946-18961.

Murphy, P.M., Baggiolini, M., Charo, I.F., Hébert, C.A., Horuk, R., Matsushima, K., Miller, L.H., Oppenheim, J.J. and Power, C.A., 2000. International union of pharmacology. XXII. Nomenclature for chemokine receptors. *Pharmacological reviews*, *52*(1), pp.145 176.

Musheke, M., Ntalasha, H., Gari, S., Mckenzie, O., Bond, V., Martin-Hilber, A. and Merten, S., 2013. A systematic review of qualitative findings on factors enabling and deterring uptake of HIV testing in Sub-Saharan Africa. *BMC public health*, *13*(1), p.220.

Nakayama, E.E., Tanaka, Y., Nagai, Y., Iwamoto, A. and Shioda, T., 2004. A CCR2-V64I polymorphism affects stability of CCR2A isoform. *Aids*, *18*(5), pp.729-738.

Nakayama, T., Hieshima, K., Izawa, D., Tatsumi, Y., Kanamaru, A. and Yoshie, O., 2003. Cutting edge: profile of chemokine receptor expression on human plasma cells accounts for their efficient recruitment to target tissues. *The Journal of Immunology*, *170*(3), pp.1136-1140.

Naranbhai, V. and Carrington, M., 2017. Host genetic variation and HIV disease: from mapping to mechanism. *Immunogenetics*, *69*(8-9), pp.489-498.

NCBI: www.ncbi.nlm.nih.gov. Accessed on the 19th September 2019.

Ndhlovu, Z.M., Kamya, P., Mewalal, N., Kløverpris, H.N., Nkosi, T., Pretorius, K., Laher, F., Ogunshola, F., Chopera, D., Shekhar, K. and Ghebremichael, M., 2015. Magnitude and kinetics of CD8+ T cell activation during hyperacute HIV infection impact viral set point. *Immunity*, *43*(3), pp.591-604.

Nguyen, L., Li, M., Chaowanachan, T., Hu, D.J., Vanichseni, S., Mock, P.A., van Griensven, F., Martin, M., Sangkum, U., Choopanya, K. and Tappero, J.W., 2004. CCR5 promoter human haplogroups associated with HIV-1 disease progression in Thai injection drug users. *Aids*, *18*(9), pp.1327-1333.

Norrgren, H., Andersson, S., Naucler, A., Dias, F., Johansson, I. and Biberfeld, G., 1995. HIV 1, HIV-2, HTLV-I/II and Treponema pallidum infections: incidence, prevalence, and HIV-2-associated mortality in an occupational cohort in Guinea-Bissau. *Journal of acquired immune deficiency syndromes and human retrovirology: official publication of the International Retrovirology Association*, *9*(4), pp.422-428.

O'Brien, T.R., Winkler, C., Dean, M., Nelson, J.A., Carrington, M., Michael, N.L. and White, G.C., 1997. HIV-1 infection in a man homozygous for CCR5△ 32. *The Lancet*, *349*(9060), p.1219.

Okulicz, J. F., Marconi, V. C., Landrum, M. L., Wegner, S., Weintrob, A., Ganesan, A., Hale, B., Crum-Cianflone, N., Delmar, J., Barthel, V., Quinnan, G., Agan, B. K., Dolan, M. J. & Infectious Disease Clinical Research Program, H. I. V. W. G. 2009. Clinical outcomes of elite controllers, viremic controllers, and long-term nonprogressors in the US Department of Defense HIV natural history study. *The Journal of Infectious Diseases* 200, 1714-1723.

Okulicz, J.F. and Lambotte, O., 2011. Epidemiology and clinical characteristics of elite controllers. *Current Opinion in HIV and AIDS*, *6*(3), pp.163-168.

Onoya, D., Reddy, P., Sifunda, S., Wingood, G.M., van den Borne, B. and Ruiter, R.A., 2011. Barriers to recruitment and retention of HIV-negative black South African women into

behavioral HIV prevention programs. *Journal of HIV/AIDS & Social Services*, *10*(3), pp.248-264.

Paiardini, M., Cervasi, B., Reyes-Aviles, E., Micci, L., Ortiz, A.M., Chahroudi, A., Vinton, C., Gordon, S.N., Bosinger, S.E., Francella, N. and Hallberg, P.L., 2011. Low levels of SIV infection in sooty mangabey central memory CD4+ T cells are associated with limited CCR5 expression. *Nature medicine*, *17*(7), p.830.

Palucka, K. and Banchereau, J., 1999. Dendritic cells: a link between innate and adaptive immunity. *Journal of clinical immunology*, *19*(1), pp.12-25.

Pandrea, I., Ribeiro, R.M., Gautam, R., Gaufin, T., Pattison, M., Barnes, M., Monjure, C., Stoulig, C., Dufour, J., Cyprian, W. and Silvestri, G., 2008. Simian immunodeficiency virus SIVagm dynamics in African green monkeys. *Journal of virology*, *82*(7), pp.3713-3724.

Passam, A.M., Sourvinos, G., Krambovitis, E., Miyakis, S., Stavrianeas, N., Zagoreos, I. and Spandidos, D.A., 2007. Polymorphisms of Cx3CR1 and CXCR6 receptors in relation to HAART therapy of HIV Type 1 patients. *AIDS research and human retroviruses*, *23*(8), pp.1026-1032.

Paterlini, M.G., 2002. Structure modeling of the chemokine receptor CCR5: implications for ligand binding and selectivity. *Biophysical journal*, *83*(6), pp.3012-3031.

Paximadis, M., Minevich, G., Winchester, R., Schramm, D.B., Gray, G.E., Sherman, G.G., Coovadia, A.H., Kuhn, L. and Tiemessen, C.T., 2011. KIR-HLA and maternal-infant HIV-1 transmission in sub-Saharan Africa. *PloS one*, *6*(2), p.e16541.

Payne, R., Muenchhoff, M., Mann, J., Roberts, H.E., Matthews, P., Adland, E., Hempenstall, A., Huang, K.H., Brockman, M., Brumme, Z. and Sinclair, M., 2014. Impact of HLA driven HIV adaptation on virulence in populations of high HIV seroprevalence. *Proceedings of the National Academy of Sciences*, *111*(50), pp.E5393-E5400.

Pedersen, B.R., Kamwendo, D., Blood, M., Mwapasa, V., Molyneux, M., North, K., Rogerson, S.J., Zimmerman, P. and Meshnick, S.R., 2007. CCR5 haplotypes and mother-to-child HIV transmission in Malawi. *PLoS One*, *2*(9), p.e838.

Pereyra F, Jia X, McLaren PJ, Telenti A, de Bakker PI, Walker BD, Ripke S, Brumme CJ, Pulit SL, Carrington M, 2010. The major genetic determinants of HIV-1 control affect HLA class I peptide presentation. *Science*, *330*(6010), pp.1551-1557.

Pereyra, F., Addo, M.M., Kaufmann, D.E., Liu, Y., Miura, T., Rathod, A., Baker, B., Trocha, A., Rosenberg, R., Mackey, E. and Ueda, P., 2008. Genetic and immunologic heterogeneity among persons who control HIV infection in the absence of therapy. *The Journal of infectious diseases*, *197*(4), pp.563-571.

Pereyra, F., Palmer, S., Miura, T., Block, B.L., Wiegand, A., Rothchild, A.C., Baker, B., Rosenberg, R., Cutrell, E., Seaman, M.S. and Coffin, J.M., 2009. Persistent low-level viremia in HIV-1 elite controllers and relationship to immunologic parameters. *The Journal of infectious diseases*, *200*(6), pp.984-990.

Persaud, D., Gay, H., Ziemniak, C., Chen, Y.H., Piatak Jr, M., Chun, T.W., Strain, M., Richman, D. and Luzuriaga, K., 2013. Absence of detectable HIV-1 viremia after treatment cessation in an infant. *New England Journal of Medicine*, *369*(19), pp.1828 1835.

Piacentini, L., Biasin, M., Fenizia, C. and Clerici, M., 2009. Genetic correlates of protection against HIV infection: the ally within. *Journal of internal medicine*, *265*(1), pp.110 124.

Picton, A.C., Paximadis, M. and Tiemessen, C.T., 2010. Genetic variation within the gene encoding the HIV-1 CCR5 coreceptor in two South African populations. *Infection, Genetics and Evolution*, *10*(4), pp.487-494.

Picton, A.C., Paximadis, M., Chaisson, R.E., Martinson, N.A. and Tiemessen, C.T., 2017. CXCR6 gene characterization in two ethnically distinct South African populations and association with viraemic disease control in HIV-1-infected black South African individuals. *Clinical Immunology*, *180*, pp.69-79.

Pilcher, C.D., Eron, J.J., Galvin, S., Gay, C. and Cohen, M.S., 2004. Acute HIV revisited: new opportunities for treatment and prevention. *The Journal of clinical investigation*, *113*(7), pp.937-945.

Pollakis, G. and Paxton, W.A., 2012. Use of (alternative) coreceptors for HIV entry. *Current opinion in HIV and AIDS*, *7*(5), pp.440-449.

Porichis, F. and Kaufmann, D.E., 2012. Role of PD-1 in HIV pathogenesis and as target for therapy. *Current HIV/AIDS Reports*, *9*(1), pp.81-90.

Porter, K., Johnson, A.M., Phillips, A.N. and Darbyshire, J.H., 1999. The practical significance of potential biases in estimates of the AIDS incubation period distribution _in the UK Register of HIV Seroconverters. *Aids*, *13*(14), pp.1943-1951.

Premack, B.A. and Schall, T.J., 1996. Chemokine receptors: gateways to inflammation and infection. *Nature medicine*, *2*(11), pp.1174-1178.

Qi, Y., Martin, M.P., Gao, X., Jacobson, L., Goedert, J.J., Buchbinder, S., Kirk, G.D., O'Brien, S.J., Trowsdale, J. and Carrington, M., 2006. KIR/HLA pleiotropism: protection against both HIV and opportunistic infections. *PLoS pathogens*, *2*(8), p.e79.

Quillent, C., Oberlin, E., Braun, J., Rousset, D., Gonzalez-Canali, G., Métais, P., Montagnier, L.,Virelizier, J.L., Arenzana-Seisdedos, F. and Beretta, A., 1998. HIV-1-resistance phenotype conferred by combination of two separate inherited mutations of CCR5 gene. *The Lancet*, *351*(9095), pp.14-18.

Raehtz, K., Pandrea, I. and Apetrei, C., 2016. The well-tempered SIV infection: Pathogenesis of SIV infection in natural hosts in the wild, with emphasis on virus transmission and early events post-infection that may contribute to protection from disease progression. *Infection, Genetics and Evolution*, *46*, pp.308-323.

Rait, M.A., Prochaska, J.J. and Rubinstein, M.L., 2015. Recruitment of adolescents for a smoking study: use of traditional strategies and social media. *Translational behavioral medicine*, *5*(3), pp.254-259.

Rajalingam, R., 2012. Overview of the killer cell immunoglobulin-like receptor system. *Immunogenetics* (pp. 391-414). Humana Press, Totowa, NJ.

Ramsay, M., Crowther, N., Tambo, E., Agongo, G., Baloyi, V., Dikotope, S., Gómez-Olivé, X., Jaff, N., Sorgho, H., Wagner, R. and Khayeka-Wandabwa, C., 2016. H3Africa AWI Gen Collaborative Centre: a resource to study the interplay between genomic and environmental risk factors for cardiometabolic diseases in four sub-Saharan African countries. *Global Health, Epidemiology and Genomics*, *1*.

Ramsay, M., Tiemessen, C.T., Choudhury, A. and Soodyall, H., 2011. Africa: the next frontier for human disease gene discovery?. *Human molecular genetics*, *20*(R2), pp.R214 R220.

Rautenbach, A. and Williams, A.A., 2020. Metabolomics as an Approach to Characterise the Contrasting Roles of CCR5 in the Presence and Absence of Disease. *International Journal of Molecular Sciences*, *21*(4), p.1472.

Ravet, S., Scott-Algara, D., Bonnet, E., Tran, H. K., Tran, T., Nguyen, N., Truong, L. X., Theodorou, I., Barre-Sinoussi, F., Pancino, G. & Paul, P. 2007. Distinctive NK-cell receptor repertoires sustain high-level constitutive NK-cell activation in HIV-exposed uninfected individuals. *Blood,* 109**,** 4296-305.

Rey-Cuillé, M.A., Berthier, J.L., Bomsel-Demontoy, M.C., Chaduc, Y., Montagnier, L., Hovanessian, A.G. and Chakrabarti, L.A., 1998. Simian immunodeficiency virus replicates to high levels in sooty mangabeys without inducing disease. *Journal of virology*, *72*(5), pp.3872-3886.

Riddick, N.E., Hermann, E.A., Loftin, L.M., Elliott, S.T., Wey, W.C., Cervasi, B., Taaffe, J., Engram, J.C., Li, B., Else, J.G. and Li, Y., 2010. A novel CCR5 mutation common in sooty mangabeys reveals SIVsmm infection of CCR5-null natural hosts and efficient alternative coreceptor use in vivo. *PLoS Pathog*, *6*(8), p.e1001064.

Riddick, N.E., Wu, F., Matsuda, K., Whitted, S., Ourmanov, I., Goldstein, S., Goeken, R.M., Plishka, R.J., Buckler-White, A., Brenchley, J.M. and Hirsch, V.M., 2015. Simian immunodeficiency virus SIVagm efficiently utilizes non-CCR5 entry pathways in African green monkey lymphocytes: potential role for GPR15 and CXCR6 as viral coreceptors. *Journal of virology*, *90*(5), pp.2316-2331.

Robert, F. and Pelletier, J., 2018. Exploring the impact of single-nucleotide polymorphisms on translation. *Frontiers in genetics*, *9*, p.507.

Rosen, S., Fox, M.P. and Gill, C.J., 2007. Patient retention in antiretroviral therapy programs in sub-Saharan Africa: a systematic review. *PLoS Med*, *4*(10), p.e298. Countries in Europe and Thailand: A cohort study. *PLoS medicine*, *15*(1), p.e1002491.

Rottman, J.B., Ganley, K.P., Williams, K., Wu, L., Mackay, C.R. and Ringler, D.J., 1997. Cellular localization of the chemokine receptor CCR5. Correlation to cellular targets of HIV-1 infection. *The American journal of pathology*, *151*(5), p.1341.

Sachdeva, M., Fischl, M.A., Pahwa, R., Sachdeva, N. and Pahwa, S., 2010. Immune exhaustion occurs concomitantly with immune activation and decrease in regulatory T cells in viremic chronically HIV-1 infected patients. *Journal of acquired immune deficiency syndromes (1999)*, *54*(5), p.447.

Sáez-Cirión, A., Bacchus, C., Hocqueloux, L., Avettand-Fenoel, V., Girault, I., Lecuroux, C., Potard, V., Versmisse, P., Melard, A., Prazuck, T. and Descours, B., 2013. Post treatment HIV-1 controllers with a long-term virological remission after the interruption early initiated antiretroviral therapy ANRS VISCONTI Study. *PLoS pathogens*, *9*(3).

Salkowitz, J.R., Bruse, S.E., Meyerson, H., Valdez, H., Mosier, D.E., Harding, C.V., Zimmerman, P.A. and Lederman, M.M., 2003. CCR5 promoter polymorphism determines macrophage CCR5 density and magnitude of HIV-1 propagation in vitro. *Clinical immunology*, *108*(3), pp.234-240.

Samson, M., Labbe, O., Mollereau, C., Vassart, G. and Parmentier, M., 1996a. Molecular cloningand functional expression of a new human CC-chemokine receptor gene. *Biochemistry*, *35*(11), pp.3362-3367.

Samson, M., Libert, F., Doranz, B.J., Rucker, J., Liesnard, C., Farber, C.M., Saragosti, S., Lapouméroulie, C., Cognaux, J., Forceille, C. and Muyldermans, G., 1996b. Resistance to HIV-1 infection in caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. *Nature*, *382*(6593), pp.722-725.

Santiago, M.L., Range, F., Keele, B.F., Li, Y., Bailes, E., Bibollet-Ruche, F., Fruteau, C., Noë, R., Peeters, M., Brookfield, J.F. and Shaw, G.M., 2005. Simian immunodeficiency virus infection in free-ranging sooty mangabeys (Cercocebus atys atys) from the Tai Forest, Cote d'Ivoire: implications for the origin of epidemic human immunodeficiency virus type 2. *Journal of virology*, *79*(19), pp.12515-12527.

Sarpong, P.K., 1977. Ashanti nubility rites. *Tema, Ghana: Ghana Publishing Corporation*.

Sauce, D., Larsen, M., Fastenackels, S., Pauchard, M., Ait-Mohand, H., Schneider, L., Guihot, A., Boufassa, F., Zaunders, J., Iguertsira, M. and Bailey, M., 2011. HIV disease progression despite suppression of viral replication is associated with exhaustion of lymphopoiesis. *Blood, The Journal of the American Society of Hematology*, *117*(19), pp.5142-5151.

Schacker, T., Little, S., Connick, E., Gebhard, K., Zhang, Z.Q., Krieger, J., Pryor, J., Havlir, D., Wong, J.K., Schooley, R.T. and Richman, D., 2001. Productive infection of T cells in lymphoid tissues during primary and early human immunodeficiency virus infection. *The Journal of infectious diseases*, *183*(4), pp.555-562.

Schinkel, J., M. W. Langendam, R. A. Coutinho, A. Krol, M. Brouwer et al., 1999 No evidence for an effect of the CCR5 delta32/+ and CCR2b 64I/+ mutations on human immunodeficiency virus (HIV)-1 disease progression among HIV-1-infected injecting drug users. J Infect Dis 179: 825-831.

Shao, W., Tang, J., Song, W., Wang, C., Li, Y., Wilson, C. M. & Kaslow, R. A. 2007. CCL3L1 and CCL4L1: variable gene copy number in adolescents with and without human immunodeficiency virus type 1 (HIV-1) infection. *Genes & Immunity*, 8, 224-231.

Sharp, A.J., Locke, D.P., McGrath, S.D., Cheng, Z., Bailey, J.A., Vallente, R.U., Pertz, L.M., Clark, R.A., Schwartz, S., Segraves, R. and Oseroff, V.V., 2005. Segmental duplications and copy-number variation in the human genome. *The American Journal of Human Genetics*, *77*(1), pp.78-88.

Shasha, D. and Walker, B.D., 2013. Lessons to be learned from natural control of HIV–future directions, therapeutic, and preventive implications. *Frontiers in immunology*, *4*, p.162.

Shaw, G.M. and Hunter, E., 2012. HIV transmission. *Cold Spring Harbor perspectives in medicine*, *2*(11), p.a006965.

Shisana, O., Rehle, T., Simbayi, L.C., Zuma, K., Jooste, S., Zungu, N., Labadarios, D. and Onoya, D., 2015. South African national HIV prevalence, incidence and behaviour survey, 2012.

Shoko, C. and Chikobvu, D., 2019. A superiority of viral load over CD4 cell count when predicting mortality in HIV patients on therapy. *BMC infectious diseases*, *19*(1), p.169.

Shrestha, S., Nyaku, M. and Edberg, J.C., 2010. Variations in CCL3L gene cluster sequence and non-specific gene copy numbers. *BMC research notes*, *3*(1), p.74.

Sloan, R.D. and Wainberg, M.A., 2011. The role of unintegrated DNA in HIV infection. *Retrovirology*, *8*(1), pp.1-15.

Smith, M.W., Dean, M., Carrington, M., Winkler, C., Huttley, G.A., Lomb, D.A., Goedert, J.J., O'Brien, T.R., Jacobson, L.P., Kaslow, R. and Buchbinder, S., 1997. Contrasting genetic influence of CCR2 and CCR5 variants on HIV-1 infection and disease progression. *Science*, *277*(5328), pp.959-965.

Souyris, M., Cenac, C., Azar, P., Daviaud, D., Canivet, A., Grunenwald, S., Pienkowski, C., Chaumeil, J., Mejía, J.E. and Guéry, J.C., 2018. TLR7 escapes X chromosome inactivation in immune cells. *Science immunology*, *3*(19).

SPARTAC Trial Investigators, 2013. Short-course antiretroviral therapy in primary HIV infection. *New England Journal of Medicine*, *368*(3), pp.207-217.

Sprague, C. and Simon, S.E., 2014. Understanding HIV care delays in the US South and the role of the social-level in HIV care engagement/retention: a qualitative study. *International journal for equity in health*, *13*(1), p.28.

Stacey, A.R., Norris, P.J., Qin, L., Haygreen, E.A., Taylor, E., Heitman, J., Lebedeva, M., DeCamp, A., Li, D., Grove, D. and Self, S.G., 2009. Induction of a striking systemic cytokine cascade prior to peak viremia in acute human immunodeficiency virus type 1 infection, in contrast to more modest and delayed responses in acute hepatitis B and C virus infections. *Journal of virology*, *83*(8), pp.3719-3733.

Swingler, S., Mann, A., Jacque, J.M., Brichacek, B., Sasseville, V.G., Williams, K., Lackner, A.A., Janoff, E.N., Wang, R., Fisher, D. and Stevenson, M., 1999. HIV-1 Nef mediates lymphocyte chemotaxis and activation by infected macrophages. *Nature medicine*, *5*(9), pp.997-1003.

Sykes, W., Van den Berg, K., Jacobs, G., Jauregui, A., Roubinian, N., Wiesner, L., Maartens, G., Swanevelder, R., Custer, B., Busch, M. and Jentsch, U., 2019. Discovery of False Elite Controllers: HIV Antibody-Positive RNA-Negative Blood Donors Found To Be on Antiretroviral Therapy. *The Journal of Infectious Diseases*, *220*(4), pp.643-647.

Szpakowska, M., Fievez, V., Arumugan, K., van Nuland, N., Schmit, J.C. and Chevigne, A., 2012. Function, diversity and therapeutic potential of the N-terminal domain of human chemokine receptors. *Biochemical pharmacology*, *84*(10), pp.1366-1380.

Thomas, R., Apps, R., Qi, Y., Gao, X., Male, V., O'huigin, C., O'Connor, G., Ge, D., Fellay, J., Martin, J.N. and Margolick, J., 2009. HLA-C cell surface expression and control of HIV/AIDS correlate with a variant upstream of HLA-C. *Nature genetics*, *41*(12), p.1290.

Tomaras, G.D., Yates, N.L., Liu, P., Qin, L., Fouda, G.G., Chavez, L.L., Decamp, A.C., Parks, R.J., Ashley, V.C., Lucas, J.T. and Cohen, M., 2008. Initial B-cell responses to transmitted human immunodeficiency virus type 1: virion-binding immunoglobulin M (IgM) and IgG antibodies followed by plasma anti-gp41 antibodies with ineffective control of initial viremia. *Journal of virology*, *82*(24), pp.12449-12463.

Townson, J.R., Barcellos, L.F. and Nibbs, R.J., 2002. Gene copy number regulates the production of the human chemokine CCL3-L1. *European journal of immunology*, *32*(10), pp.3016-3026.

Ulmer, A.J., Scholz, W., Ernst, M., Brandt, E. and Flad, H.D., 1984. Isolation and subfractionation of human peripheral blood mononuclear cells (PBMC) by density gradient centrifugation on Percoll. *Immunobiology*, *166*(3), pp.238-250.

UNAIDS data 2018a. *UNAIDS Web site* [Online]. Available: http://www.unaids.org

UNAIDS, 2018b. Fact sheet-latest global and regional statistics on the status of the AIDS epidemic. *Geneva: UNAIDS. UNAIDS Web site* [Online]. Available: http://www.unaids.org

UNAIDS, 2019 Global, H.I.V., 2019. AIDS statistics–2018 fact sheet. *UNAIDS Web site* [Online]. Available: https://www.unaids.org

Urban, T. J., Weintrob, A. C., Fellay, J., Colombo, S., Shianna, K. V., Gumbs, C., Rotger, M.,Pelak, K., Dang, K. K., Detels, R., Martinson, J. J., O'Brien, S. J., Letvin, N. L., McMichael, A. J., Haynes, B. F., Carrington, M., Telenti, A., Michael, N. L. & Goldstein, D. B. 2009. CCL3L1 and HIV/AIDS susceptibility. *Nature Medicine,* 15, 1110-1112.

van Bergen, J. and Koning, F., 2010. The tortoise and the hare: slowly evolving T-cell responses take hastily evolving KIR. *Immunology*, *131*(3), pp.301-309.

van Damme, M.P.W.A., J 2002 Macrophage inflammatory protein-1. *Cytokine & Growt Factor Reviews*, *13*, pp.455-481.

van Loggerenberg, F., Mlisana, K., Williamson, C., Auld, S.C., Morris, L., Gray, C.M., Karim, Q.A., Grobler, A., Barnabas, N., Iriogbe, I. and Karim, S.S.A., 2008. Establishing a cohort at high risk of HIV infection in South Africa: challenges and experiences of the CAPRISA 002 acute infection study. *PloS one*, *3*(4). Variation. *Nature*, *526*(7571), pp.68-74.

Vince, N., Bashirova, A.A., Lied, A., Gao, X., Dorrell, L., McLaren, P.J., Fellay, J. and Carrington, M., 2014. HLA class I and KIR genes do not protect against HIV type 1 infection in highly exposed uninfected individuals with hemophilia A. *The Journal of infectious diseases*, *210*(7), pp.1047-1051.

Vingert, B., Perez-Patrigeon, S., Jeannin, P., Lambotte, O., Boufassa, F., Lemaître, F., Kwok, W.W., Theodorou, I., Delfraissy, J.F., Thèze, J. and Chakrabarti, L.A., 2010. HIV controller CD4+ T cells respond to minimal amounts of Gag antigen due to high TCR avidity. *PLoS Pathog*, *6*(2), p.e1000780.

Viola, A. and Luster, A.D., 2008. Chemokines and their receptors: drug targets in immunity and inflammation. *Annu. Rev. Pharmacol. Toxicol. 48*, pp.171-197.

Violari, A., Cotton, M.F., Kuhn, L., Schramm, D.B., Paximadis, M., Loubser, S., Shalekoff, S., Dias, B.D.C., Otwombe, K., Liberty, A. and McIntyre, J., 2019. A child with perinatal HIV infection and long-term sustained virological control following antiretroviral treatment cessation. *Nature communications*, *10*(1), pp.1-11.

Volberding, P.A. and Deeks, S.G., 2010. Antiretroviral therapy and management of HIV infection. *The Lancet*, *376*(9734), pp.49-62.

Walker, B.D., 2007. Elite control of HIV Infection: implications for vaccines and treatment. *Topics in HIV medicine: a publication of the International AIDS Society, USA*, *15*(4), pp.134-136.

Walker, B.D., Chakrabarti, S., Moss, B., Paradis, T.J., Flynn, T., Durno, A.G., Blumberg, R.S., Kaplan, J.C., Hirsch, M.S. and Schooley, R.T., 1987. HIV-specific cytotoxic T lymphocytes in seropositive individuals. *Nature*, *328*(6128), pp.345-348.

Wang, Z., Shang, H. and Jiang, Y., 2017. Chemokines and Chemokine Receptors: Accomplices for Human immunodeficiency virus infection and Latency. *Frontiers in immunology*, *8*, p.1274.

Weitzman, E.R., Kelemen, S., Kaci, L. and Mandl, K.D., 2012. Willingness to share personal health record data for care improvement and public health: a survey of experienced personal health record users. *BMC medical informatics and decision making*, *12*(1), pp.110.

Wetzel, K.S., Yi, Y., Elliott, S.T., Romero, D., Jacquelin, B., Hahn, B.H., Muller-Trutwin, M., Apetrei, C., Pandrea, I. and Collman, R.G., 2017. CXCR6-mediated simian immunodeficiency virus SIVagmSab entry into Sabaeus African green monkey lymphocytes implicates widespread use of non-CCR5 pathways in natural host infections. *Journal of virology*, *91*(4), pp.e01626-16.

Wherry, E.J., Blattman, J.N., Murali-Krishna, K., Van Der Most, R. and Ahmed, R., 2003. Viral persistence alters CD8 T-cell immunodominance and tissue distribution and results in distinct stages of functional impairment. *Journal of virology*, *77*(8), pp.4911-4927.

Wilbanks, A., Zondlo, S.C., Murphy, K., Mak, S., Soler, D., Langdon, P., Andrew, D.P., Wu, L. and Briskin, M., 2001. Expression cloning of the STRL33/BONZO/TYMSTR ligand reveals elements of CC, CXC, and CX3C chemokines. *The Journal of Immunology*, *166*(8), pp.5145-5154.

Williamson, C., Loubser, S.A., Brice, B., Joubert, G., Smit, T., Thomas, R., Visagie, M., Cooper, M. and van der Ryst, E., 2000. Allelic frequencies of host genetic variants influencing susceptibility to HIV-1 infection and disease in South African populations. *Aids*, *14*(4), pp.449-451.

World Health Organization (WHO), 2016. *Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection: recommendations for a public health approach*.

Wuttge, D.M., Zhou, X., Sheikine, Y., Wågsäter, D., Stemme, V., Hedin, U., Stemme, S., Hansson, G.K. and Sirsjö, A., 2004. CXCL16/SR-PSOX is an interferon-γ–regulated chemokine and scavenger receptor expressed in atherosclerotic lesions. *Arteriosclerosis, thrombosis, and vascular biology*, *24*(4), pp.750-755.

Yang, O.O., Cumberland, W.G., Escobar, R., Liao, D. and Chew, K.W., 2017. Demographics and natural history of HIV-1-infected spontaneous controllers of viremia: extremes on a continuum. *AIDS (London, England)*, *31*(8), p.1091.

Zernecke, A., Shagdarsuren, E. and Weber, C., 2008. Chemokines in atherosclerosis: an update. *Arteriosclerosis, thrombosis, and vascular biology*, *28*(11), pp.1897-1908.

Zhang, W., Ambikan, A.T., Sperk, M., van Domselaar, R., Nowak, P., Noyan, K., Russom, A., Sönnerborg, A. and Neogi, U., 2018. Transcriptomics and targeted proteomics analysis to gain insights into the immune-control mechanisms of HIV-1 infected elite controllers. *EBioMedicine*, *27*, pp.40-50.

Zhang, Y.J., Zhang, L., Ketas, T., Korber, B.T. and Moore, J.P., 2001. HIV type 1 molecular clones able to use the Bonzo/STRL-33 coreceptor for virus entry. *AIDS research and human retroviruses*, *17*(3), pp.217-227.

Zhao, K., Ishida, Y., Oleksyk, T.K., Winkler, C.A. and Roca, A.L., 2012. Evidence for selection at HIV host susceptibility genes in a West Central African human population. *BMC evolutionary biology*, *12*(1), p.237.

Zlotnik, A. and Yoshie, O., 2000. Chemokines: a new classification system and their role in immunity. *Immunity*, *12*(2), pp.121-127.

APPENDIX A: Ethical clearance

APPENDIX 1A: Human Research on Human Subjects Committee of the University of the Witwatersrand, Johannesburg

R14/49 Prof CT Tiemessen et al

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)

CLEARANCE CERTIFICATE NO. M140926

NAME: (Principal Investigator)	Prof CT Tiemessen et al
DEPARTMENT:	School of Pathology Division of Virology National Institute for Communicable Diseases University of the Witwatersrand and MRC,Durban South African National Blood Services
PROJECT TITLE:	HIV-1 Positive South African Elite and Long Term Controllers: Viral and Host Targets for Functional Cure Strategies
DATE CONSIDERED:	03/10/2014
DECISION:	Approved unconditionally
CONDITIONS:	
SUPERVISOR:	
APPROVED BY:	Professor P Cleaton-Jones, Chairperson, HREC (Medical)
DATE OF APPROVAL:	17/10/2014

This clearance certificate is valid for 5 years from date of approval. Extension may be applied for.

DECLARATION OF INVESTIGATORS

To be completed in duplicate and **ONE COPY** returned to the Secretary in Room 10004, 10th floor, Senate House, University.

I/we fully understand the conditions under which I am/we are authorized to carry out the above-mentioned research and I/we undertake to ensure compliance with these conditions. Should any departure be contemplated, from the research protocol as approved, I/we undertake to resubmit the application to the Committee. **I agree to submit a yearly progress report.**

Principal Investigator Signature

Date 15 October 2014

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES

UNIVERSITY OF THE WITWATERSRAND JOHANNESBURG

R14/49 Professor CT Tiemessen, et al

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)
CLEARANCE CERTIFICATE NO. M190995

NAME: (Principal Investigator)	Professor CT Tiemessen, et al
DEPARTMENT:	National Institute for Communicable Diseases Centre for HIV and Sexually-Transmitted Infections Sandringham
PROJECT TITLE:	HIV-1 positive South African Elite and Long-term Controllers: viral and host targets for functional cure strategies
DATE CONSIDERED:	Ad hoc
DECISION:	Aproved unconditionally
CONDITIONS:	Renewal of M140926
SUPERVISOR:	Not applicable
APPROVED BY:	Dr CB Penny, Chairperson, HREC (Medical)
DATE OF APPROVAL:	2019/10/04

This clearance certificate is valid for 5 years from date of approval. Extension may be applied for.

DECLARATION OF INVESTIGATORS

To be completed in duplicate and **ONE COPY** returned to the Research Office Secretary on the 3rd Floor, Phillip Tobias Building, Parktown, University of the Witwatersrand, Johannesburg.

I/we fully understand the conditions under which I am/we are authorized to carry out the above-mentioned research and I/we undertake to ensure compliance with these conditions. Should any departure be contemplated from the research protocol as approved, I/we undertake to submit details to the Committee. I agree to submit a yearly progress report. When a funder requires annual re-certification, the application date will be one year after the date when the study was initially reviewed. In this case, the study was initially reviewed in **September** and will therefore reports and re-certification will be due early in the month of **September** each year. Unreported changes to the application may invalidate the clearance given by the HREC (Medical).

Principal Investigator Signature

5 October 2019
Date

PLEASE QUOTE THE CLEARANCE CERTIFICATE NUMBER IN ALL ENQUIRIES

APPENDIX 2A: Ethics clearance for the cohort in KZN was obtained from Medical Research Ethics Committee



ETHICS COMMITTEE

PO Box 19070, 7505 Tygerberg, South Africa
Francie van Zijl Drive, Parowvallei, 7500
Tel: +27 (0)21 938-0687; Fax: +27 (0) 866-854023
E-mail: adri.labuschagne@mrc.ac.za
http://www.mrc.ac.za/ethics/ethics.htm

19 May 2015

Prof P Kiepiela
HIV Prevention Research Unit
MRC Durban Westville

Dear Prof Kiepiela

Protocol ID: EC017-11/2014
Protocol title: HIV-1-positive South African elite and long-term controllers: viral and host targets for functional cure strategies
Meeting date: 28 April 2015

Thank you for your responses to the Committee, dated 2 February, 16 March, and 13 May 2015. The responses were satisfactory. I am pleased to inform you that ethics approval is now granted for the study.

Please note that the approval is valid for 1 year, i.e. from 28 April 2015 to 27 April 2016. Any changes to the research protocol must be submitted as an amendment. Any adverse events must be reported within 48 hours. Any protocol deviations have to be reported.

Wishing you well with your research.

Yours sincerely

PROF. D DU TOIT
CHAIRPERSON: MRC ETHICS COMMITTEE

MRC Ethics Committee: Prof D du Toit (chairperson), Prof D Kayongo, Dr NE Khomo, Ms N Morar, Prof N Morojele, Prof H Oosthuizen, Mr D Rebombo, Dr Y Sikweyiya, Prof A van Niekerk, Ms A Labuschagne

APPENDIX 3A: EThekwini Research Ethics Committee

ETHEKWINI MUNICIPALITY
Community & Emergency Services Cluster
Office of the Head of Health

9 Archie Gumede Place
Durban 4001
P O Box 2443
Durban 4000
Tel: (031) 311 3505
Fax: (031) 311 3710
Website: http://www.durban.org.za

21 April 2016

Dear Dr P Kiepiela

Subject: Approval of a research proposal.

The research proposal titled: **HIV -1 positive South African elite and long term controllers: viral and host targets for remission strategies** was reviewed by the eThekwini Municipal Health Department research Committee. The study is hereby **approved to be conducted at Lancers Road, Chesterville and CDC clinics.**

The following conditions need to be noted:

- Submission of the indemnity form obtainable from the EThekwini Municipality Health Unit before commencement of the study.
- Prior arrangements to be made with the facility and an assurance that all services will not be disrupted.
- No staff member should be used for collecting data for the researchers.
- Progress reports to be provided and the final report of the study to the eThekwini Municipality Health Unit or emailed to:**ntombifuthi.mangeni@durban.gov.za and Cc: grace.mufamadi@durban.gov.za**
- Obtain permission from the eThekwini municipality health department for press releases and release of results to communities/stakeholders.
- The department has to receive recognition for the assistance given.
- Any amended to the study to be communicated with the eThekwini Municipality Health Unit and the relevant amendment form obtainable from the unit to be submitted.
- Withdrawal of permission to conduct research will be left to the discretion of the eThekwini Municipality Health Unit.

Kind Regards

pp **Dr NI Gxagxisa**
Head: Health Unit

APPENDIX 4A: KZN Department of Health approval letter



health
Department:
Health
PROVINCE OF KWAZULU-NATAL

DIRECTORATE:
Health Research & Knowledge Management

Physical Address: 330 Langalibalele Street, Pietermaritzburg
Postal Address: Private Bag X9051
Tel: 033 395 2805/ 3186/ 3123 Fax: 033 394 3782
Email:
www.kznhealth.gov.za

Reference: 85/16

Date: 18 March 2016

Dear Dr P Kiepiela

Approval of research

1 The research proposal titled 'HIV-1 positive South African elite and long term controllers: viral and host targets for remission strategies' was reviewed by the Kwazulu-Natal Department of Health.

The proposal is hereby **approved** for research to be undertaken at Chesterville clinic.

2. You are requested to take note of the following:
   a. Make the necessary arrangement with the identified facility before commencing with your research project.
   b. Provide an interim progress report and final report (electronic and hard copies) when your research is complete.

3. Your final report must be posted to **HEALTH RESEARCH AND KNOWLEDGE MANAGEMENT, 10-102, PRIVATE BAG X9051, PIETERMARITZBURG, 3200** and e-mail an electronic copy to hrkm@kznhealth.gov.za

For any additional information please contact Mr X. Xaba on 033-395 2805.

Yours Sincerely

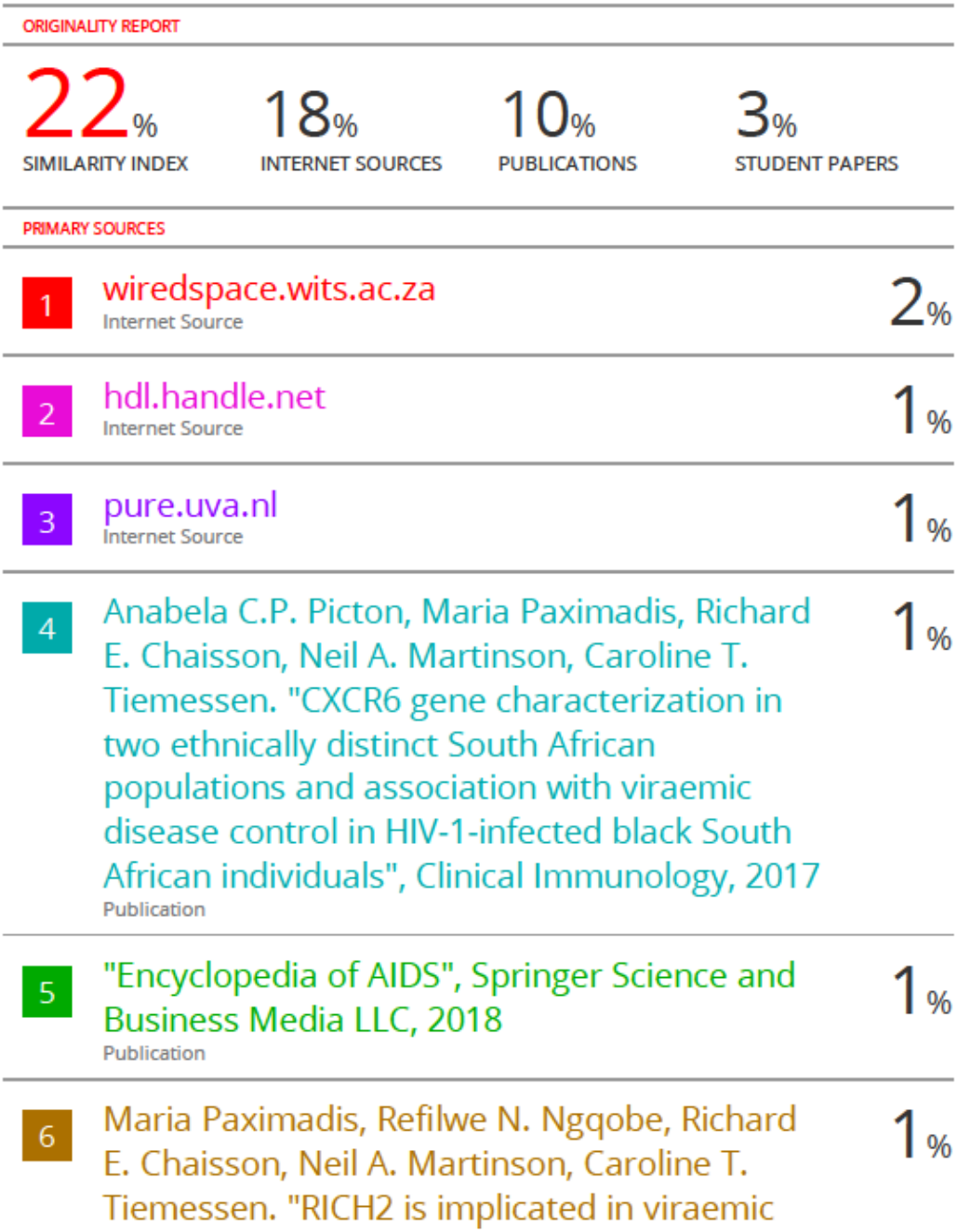
Dr E Lutge

Chairperson, Health Research Committee

Date: 18/03/16.

Fighting Disease, Fighting Poverty, Giving Hope

APPENDIX B: Turnitin report



ORIGINALITY REPORT

22%	18%	10%	3%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	wiredspace.wits.ac.za Internet Source	2%
2	hdl.handle.net Internet Source	1%
3	pure.uva.nl Internet Source	1%
4	Anabela C.P. Picton, Maria Paximadis, Richard E. Chaisson, Neil A. Martinson, Caroline T. Tiemessen. "CXCR6 gene characterization in two ethnically distinct South African populations and association with viraemic disease control in HIV-1-infected black South African individuals", Clinical Immunology, 2017 Publication	1%
5	"Encyclopedia of AIDS", Springer Science and Business Media LLC, 2018 Publication	1%
6	Maria Paximadis, Refilwe N. Ngqobe, Richard E. Chaisson, Neil A. Martinson, Caroline T. Tiemessen. "RICH2 is implicated in viraemic	1%